NORTHWESTERN UNIVERSITY

Physics-Informed and Data-Driven Methods for Computational Imaging and Sensing

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Computer Science

By

Manuel Ballester Matito

EVANSTON, ILLINOIS

August 2025

Abstract

This thesis introduces advanced computational methods that integrate optical physics with machine learning to improve the design and analysis of modern imaging and sensing systems. These methods combine physics-based modeling, global optimization, and data-driven learning to enhance the precision, stability, and interpretability of such systems under complex real-world conditions. The thesis is structured around three core projects, outlined below.

First, we present a Time-of-Flight infrared camera based on multi-wavelength interferometry, enabling high-resolution 3D depth sensing with sub-millimeter accuracy using conventional CMOS sensors. We further introduce machine learning models for single-shot phase unwrapping under noise conditions.

Second, we advance the spectroscopic analysis of thin solid films by improving transmittance and reflectance processing methods through global optimization and neural architectures that explicitly encode the underlying physics. These models enable accurate extraction of the complex refractive index of the thin films.

Third, we model and reverse engineer the functioning of semiconductor gamma-ray detectors, with a focus on CdZnTe-based technologies. Through a combination of PDE solvers, digital twin simulations, and hybrid deep learning frameworks, we achieve precise material defect characterization and sub-millimeter 3D event reconstruction.

Acknowledgments

This thesis represents the culmination of a long and deeply meaningful academic and personal journey. This path would not have been possible without the full support, encouragement, and guidance of many individuals. I am profoundly grateful to each of them. First and foremost, I extend my deepest thanks to my wife, Wildaline Serin, whose belief in my academic career and support sustained me through every challenge. To my brother, Marcos Ballester, for countless high-level discussions, academic insight, and support attending my key scientific events. To my parents, Marcos Antonio Ballester and Maria del Carmen Matito, for feeding my curiosity and encouraging my academic path from the very beginning. To my close friend Ignacio Chanivet, for countless stimulating conversations on modern physics and philosophy, which continually pushed my thinking beyond conventional boundaries. To my friend and classical music companion Javier de la Torre, who reminded me of the profound connection between science and art. To Joe Choat, Walt Stagner, and Dan Connolly, for introducing me to the roots of American music and culture, and for showing me the technology behind electric guitars and amps.

To my early mentor, Prof. Emilio Marquez, who welcomed me into physics research as a second-year math undergraduate and continues to collaborate with me to this day. Thank you also to Prof. Peter Petrik and Prof. Tivadar Lohner, from the Hungarian Academy of Sciences, with whom I had the opportunity to work during a summer research internship. Their mentorship and example of rigorous, passionate scientific inquiry deeply influenced my approach to research. To Prof. Nicolas Joly, from whom I learned most of what I know about optical theory. His guidance during my time in Germany had a lasting impact on my understanding of the field. To Prof. Oliver Cossairt, who served as my advisor during my predoctoral work and during the first year of my PhD path. His dynamic and insightful discussions helped shape my initial direction and thinking at the start of

this Ph.D. journey. To my Ph.D. co-advisor, Prof. Florian Willomitzer, whose mentorship and enthusiasm have been instrumental in the development of several projects presented in this thesis. And to my main Ph.D. advisor, Prof. Aggelos Katsaggelos, who welcomed me into his lab in my second year and has since exemplified what it means to be a truly interdisciplinary scientist, moving across engineering, physics, mathematics, medicine, and computer science. His mentorship has profoundly shaped both my technical work and my broader understanding of what it means to be a STEM scholar.

To my additional committee members, Prof. Tumblin, whose Socratic questioning helped me think more deeply, and Prof. Alexander, whose talk on bioinspired optical systems forever changed how I view the boundaries of optics. To Siemens Healthineers, especially Dr. Hans Vija, Dr. Jarek Kaspar, and Dr. Francesc Massanes, for both financial support and scientific mentorship over the past three years. You have taught me that research in industry can be as fundamental, ethical, and inspiring as in academia. Finally, I thank all the students and researchers I have had the privilege to closely work with during this Ph.D. journey, including Florian Schiffers, Jipeng Sun, Santiago Lopez Tapia, Asami Odate, Patrick Cornwall, Henry Chopp, John Bass, Hamidreza Hasani, Aniket Dashpute, Srutarshi Banerjee, Lionel Fiske, and many others. Your curiosity and collaborative spirit have greatly enriched my during these last four years of hard work.

Preface

This dissertation is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Computer Science at Northwestern University. This dissertation is a compilation of published and submitted works, all first-authored by the candidate, with minor adaptations and additions to unify notation, context, and narrative flow. Full citation details of each publication are provided in the bibliography.

The central theme of this work is the use of optimization techniques and (physics-based) machine learning for diverse problems of computational optics, imaging and sensing. The thesis includes three main contributions: (1) the development of a single-shot multi-wavelength interferometric 3D imaging system; (2) the spectrophotometric analysis and characterization of thin solid films; and (3) reverse engineering the properties and defects of semiconductor gamma-ray detectors. The developed methods find broad applicability in fields such as optical metrology, semiconductor processing, and medical imaging. Beyond these core chapters, this work reflects broader collaborations during my PhD, including the modeling of stellar evolution with PINNs, simulation of polarized light microscopy for material identification, the development of 3D dynamic holographic displays with spatial light modulators, and recovery of degraded holograms through neural rendering, among others.

During my PhD path, I was also fortunate to participate in interdisciplinary efforts at institutions such as The Art Institute of Chicago, The Hong Kong University, University of Cadiz (UCA), Eastern Switzerland University of Applied Sciences (OST), Friedrich-Alexander University of Erlangen-Nuremberg (FAU), and multiple departments within Northwestern University (NU), including Computer Science, Electrical and Computer Engineering, and Chemistry. I also had the opportunity to contribute to the mentoring of

undergraduate and master's students. Special thanks to Marcos Garcia Gutierrez (UCA), Noah Luechinger (OST), Rjano Ryser (OST), and Bao-An Nguyen (FAU).

My academic journey began with a B.S. in Mathematics, with a specialization in Mathematical Engineering, from the University of Cadiz (UCA), Spain, in 2018. I then pursued an M.S. in Advanced Optical Technologies at Friedrich-Alexander University (FAU), Germany, which I completed in 2021 with a double major on Physics of Light and Computational Optics. As part of this doctoral program, I also earned an M.S. in Computer Science from Northwestern University in 2023. This interdisciplinary background, spanning mathematics, physics, and computer science, has shaped the perspective and methodology behind this thesis, which sits at the intersection of computational techniques and physical modeling. This work is dedicated to all those who have guided, challenged, and supported me throughout this journey.

List of Abbreviations

- AI: Artificial Intelligence
- AR: Augmented Reality
- ASM: Angular Spectrum Method
- BRDF: Bidirectional Reflectance Distribution Function
- CdZnTe (CZT): Cadmium Zinc Telluride
- CMOS: Complementary Metal–Oxide–Semiconductor
- CNN: Convolutional Neural Network
- DL: Deep Learning
- FWHM: Full Width at Half Maximum
- GND: Generalized Normal Distribution
- IR: Infrared
- LED: Light Emitting Diode

- LIDAR: Light Detection and Ranging
- LSTM: Long Short-Term Memory (network)
- ML: Machine Learning
- NeRF: Neural Radiance Field
- NIR: Near-Infrared
- NLO: Nonlinear Optics
- PBML: Physics-Based Machine Learning
- PINN: Physics-Informed Neural Network
- PINO: Physics-Informed Neural Operator
- PLM: Polarized Light Microscopy
- RNN: Recurrent Neural Network
- SLM: Spatial Light Modulator
- SW: Synthetic Wavelength
- SWI: Synthetic Wavelength Interferometry
- SWH: Synthetic Wavelength Holography
- ToF: Time-of-Flight
- UV: Ultraviolet

- VAE: Variational Autoencoder
- Vis: Visible Spectrum
- VR: Virtual Reality

Glossary

- **Bandgap**: The energy difference between the conduction band and valence band of a material. Photons with energy below the bandgap cannot excite electrons across the gap, determining the optical absorption edge.
- **Dielectric Function**: A complex-valued function describing how a material responds to an electromagnetic field, directly related to the complex refractive index in non-magnetic materials.
- **Differentiable Model**: A forward model or simulator that is fully differentiable, enabling the use of gradient-based optimization for solving inverse problems. Typically implemented in frameworks like PyTorch.
- **Digital Twin**: A computational model that replicates the behavior of a physical system using simulations, enabling prediction, analysis, optimization, or control.
- **Dispersion Model**: A mathematical model that expresses the wavelength dependence of a material's dielectric function using a set of real-valued parameters.
- **Ellipsometry**: An optical technique that measures the change in polarization of light upon reflection, used to characterize thin film thickness and complex refractive index.

- **Fabry–Perot Interference**: Interference patterns arising from multiple reflections within a thin film, producing oscillations in the transmission or reflection spectra.
- **Film Sample**: A structure consisting of a thin target material (film) deposited on top of a typically transparent substrate, used for optical measurements.
- **Forward Problem**: The simulation of a physical process given known input parameters, often used to develop digital twins of sensors or imaging systems.
- **Inverse Problem**: The process of inferring the underlying causes or physical parameters from observed measurements, often ill-posed and requiring optimization or learning methods to solve.
- **Inverse Synthesis**: A computational method to determine the optical properties of thin films by fitting theoretical transmission or reflection spectra to experimental data.
- **Monte Carlo Method**: A numerical simulation technique that uses random sampling to approximate physical processes, often used for fast but statistically accurate approximations in radiation transport models.
- **Optical Characterization**: The process of determining a material's wavelength dependent optical functions, particularly the complex refractive index composed of the real part n and the extinction coefficient κ.
- **Phase Unwrapping**: A computational technique used to reconstruct a continuous phase map from wrapped phase data, which are typically recorded modulo 2π.
- **Probabilistic Model**: A simulator that incorporates randomness or uncertainty, meaning repeated runs with the same inputs may produce slightly different outputs. Use-

ful for modeling stochastic physical systems such as radiation detectors.

- **Scintillator**: A material that emits light when exposed to high-energy radiation, used in gamma-ray detection. Often compared with direct semiconductor detectors like CdZnTe.
- **Spectrophotometry**: The study of how materials reflect and transmit light intensity as a function of wavelength.
- **Swanepoel / Envelope Method**: A computational technique for extracting thin film optical properties by analyzing the upper and lower envelopes of oscillatory transmission or reflection spectra.
- **Synthetic Wavelength**: The long envelope wavelength of the beat wave calculated computationally from the interference of two optical waves.
- **Thin Solid Film**: A material layer (typically in the nanometer to micrometer range) deposited on a thick, usually transparent substrate.
- **Transmission / Reflection Spectra**: Spectral curves showing the wavelength-dependent transmittance or reflectance (ranging from 0 to 1) of a film sample, often used to perform the material characterization.

Dedication

To my family, friends, mentors, and collaborators: For your wisdom and guidance.

Table of Contents

Abstract **3**

Acknowledgments **4**

Preface **6**

List of Abbreviations **8**

Glossary **11**

Dedication **13**

Table of Contents **15**

List of Figures **18**

List of Tables **25**

1 Introduction **27**

2 Single-shot interferometric camera **31**

2.1 Introduction 32

2.2 Principles and computational framework 36

2.3 Experiments and Data Processing 41
2.4 Summary, discussion, and outlook 51
2.5 Conclusions 54

3 Spectrophotometry of thin solid films 56

3.1 Novel Formulas for Transmittance and Reflectance 58
3.2 Optimization based Envelope Detection Algorithm 102
3.3 Machine learning based characterization 146

4 Gamma ray semiconductor detectors 155

4.1 Novel Physics-Based Model for Detector Simulation 156
4.2 Defect detection with automatic differentiation 189
4.3 Event reconstruction and detector characterization 201

References 208

A Supplementary Material for Single-shot Interferometric Camera 241

A.1 Retrieving the optical fields 241
A.2 Synthetic Wavelength Interferometry 251
A.3 Limitations of the double-shot method 255
A.4 Non-Line-of-Sight experiment 257
A.5 Analysis and further results 259

B Supplementary Material for Thin Solid Films Spectroscopy 268

B.1 Swanepoel transmission and reflection coefficients 268
B.2 Substrate absorption model for simulations 270
B.3 High frequency noise in spectrophotometric measures 270

C Supplementary Material for Gamma-ray radiation detector **272**

C.1 Solving diffusion-repulsion equation numerically 272

C.2 Generalization of the diffusion-repulsion algorithm 274

Vita **274**

Vita **275**

List of Figures

2.1 Schematic setup and computational procedure of our novel single-shot ToF camera system. (a) Object (the camera FoV is 66 mm $\times$ 66 mm). (b) The lens system images the object onto a CCD/CMOS chip. An object beam at two wavelengths illuminates the object and two reference beams at two different wavelengths directly illuminate the chip at two different (vertical and horizontal) angles. (c) The captured image I(x,y) along with a zoom-in window. The image consists of speckles that are overlayed with crossed fringes. (d) Fourier transform of the captured image. (e) By shifting, filtering, and inverse Fourier transform, the speckled complex fields (amplitude and phase) at both optical wavelengths (see section 2.2) can be recovered from only one camera image. (f) Using the SWI principle, we generate the synthetic phasemap $\phi(\Lambda)$ to calculate the 3D depth map of the object. . . 36

2.2 Picture of our lab setup. Two identical tunable lasers emit light in the NIR range at around 850 nm. We use optical fibers and fiber beam splitters (BS) to direct the beams flexibly. We employ a lens system plus aperture to image the object into the sensor array. 39

2.3 Double-shot acquisition mode. The measured object is a painted clay figure of approximately 10 cm height (see Fig 2.1a), and the field of view is 66×66 mm^2. (a) Image of the small object cropped to the actual FoV. (b) Wrapped phasemap $\phi(\Lambda = 10\text{mm})$. (c) Acquired synthetic phasemap $\phi(\Lambda = 45\text{mm})$. (d) Phasemap $\phi(\Lambda = 10\text{mm})$, unwrapped with our multi-frequency unwrapping algorithm. (e) and (f) 3D model of the object calculated from $\phi(\Lambda = 45\text{mm})$ (e) and from unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 10\text{mm})$. . . 42

2.4 Single-shot acquisition mode. For the sake of comparability, the same object as in Fig. 2.3 is measured. (a) Image of the small object cropped to the actual FoV. (b) Wrapped phasemap $\phi(\Lambda = 10\text{mm})$. (c) Acquired synthetic phasemap $\phi(\Lambda = 50\text{mm})$. (d) Phasemap $\phi(\Lambda = 10\text{mm})$, unwrapped with our multi-frequency unwrapping algorithm. (e) and (f) 3D model of the object calculated from $\phi(\Lambda = 50\text{mm})$ (e) and from unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 10\text{mm})$. 45

2.5 Single-shot measurements of a dynamic object. (a) Object: Metronome sprayed with a scattering powder. (b, d) Phasemaps $\phi(\Lambda = 30\text{mm})$ obtained from two different video frames. (c, e) Corresponding 3D models. 46
2.6 Single-shot phase unwrapping using Deep Learning (DL). (a) Acquired wrapped synthetic phasemap $\phi(\Lambda = 5\text{mm})$. (b) Unwrapped phasemap predicted by our DL algorithm, $\phi^{\text{unwrap}}(\Lambda = 5\text{mm})$. (c) Comparison: Phasemap obtained from another single-shot unwrapping procedure [72], [73] (not DL-based). Unwrapping failed as phase jumps are still present. (d) 3D model calculated from our DL-based unwrapped phasemap (b). 48
2.7 Full-field, single-shot Non-Light-of-Sight (NLoS) imaging. (a) Schematic of the experiment. (b) Synthetic phasemap $\phi(\Lambda = 1\text{mm})$ of the object's hologram at the diffuser location. (c) The amplitude of backpropagated object (the point light source). 52

3.1 (a) Geometrical model of the sample under study. A thin solid film is deposited on top of a glass substrate. The inset shows a cross-sectional view of the film within the illuminated area, where the thin-film surface is approximated as a plane with a wedge parameter Δd. (b) Schematic diagram of the experimental spectrophotometric setup for transmission and reflection measurements. The setup includes two broadband light sources covering the UV and VIS-NIR spectral ranges, a diffraction optical element, a tunable slit that produces a quasi-monochromatic beam, and a beam splitter. 59
3.2 Scheme of the stratified media. Multiple interferences between incident and reflected waves occur because of the different interfaces. 71
3.3 Transmittance (a) and reflectance (b) spectra of our simulated sample (for an amorphous silicon thin film), assuming a weakly absorbing glass substrate. The results are shown for a monochromatic light source with infinite coherence (Eqs. 3.24 and 3.25) and for another light source with limited coherence (Eqs. 3.35 and 3.36), where $2d_1 < \ell < 2d_2$. 74
3.4 Sample (a) transmittance and (b) reflectance, using the exact formulae for uniform films (Eqs. 3.29-3.30) and for wedge films (Eqs. 3.37-3.45), with $\Delta d = 30$ nm. When the wedge increases, the contrast in the interferences decreases, as does the amplitude of the oscillations. 80
3.5 Simulated transmittance showing the effect of the correction factors N^+ and N^- (see Eqs. 3.41) for thin films with two representative wedge parameters, (a) $\Delta d = 30$ nm and (b) $\Delta d = 60$ nm. 80
3.6 Profile of a tilted film with a linear surface described by $S(x') = mx'$. The optical path length in the $z-$direction varies depending on the position along the $x'-$axis. 83

3.7 Sample transmittance (left column) and reflectance (right column) using the Swanepoel approximations. The upper and lower envelopes, as well as the tangent points where the spectrum intersects the envelopes, are shown. (a) and (b) represent uniform films (see Eqs. 3.74 and 3.90), while (c) and (d) correspond to films with a very high wedge, $\Delta d = 100$ nm (see Eqs. 3.80 and 3.99). Under high wedge condition, we see how the envelopes *cross over*, reaching a particular wavelength in which the spectrum and the two envelopes coincide. 91
3.8 Diagram of the spectroscopic measurements for the substrate alone. The measures of T_{s} and R_{s} allow us to extract the substrate properties n_2 and k_2. 97
3.9 (a) Schematic of the sample geometry, comprising a thin layer deposited on a glass substrate. (b) Multiple internal reflections within the thin film and the substrate. The total transmittance is measured sequentially throughout different wavelengths. (c) Simulated transmission spectra for the full sample and for the substrate alone. The ground-truth upper and lower envelopes, along with their tangent points, are indicated. (d) Magnified view of the strongly absorbing spectral region. The upper and lower tangent points are sometimes not associated with a nearby local extremum (first two on the left), or occur at positions significantly far from the local extrema (last two on the right). 106
3.10 Random initial guesses for the film parameters (such as thickness and the coefficients of the optical dispersion model) are selected. A theoretical transmittance spectrum is calculated from these parameters via Eq. 3.108. The simulated spectrum is then compared to the experimental spectrum, and the residual error is quantified using the cost function (RMSE) defined in Eq. 3.112. A global optimization algorithm iteratively adjusts the parameters to minimize this residual error. This process repeats until the optimal parameters that yield the best fit between the simulated and experimental spectra are found. 117
3.11 Comparison of experimental transmission measurements with simulated transmission curves for (a) sample #1 and (b) sample #2 of amorphous silicon, obtained using the inverse synthesis method. The bottom panels illustrate the residual differences or error between the experimental and simulated data, defined as $T^{\mathrm{exp}} - T^{\mathrm{sim}}$. 119

3.12 Summary of the Swanepoel method: the steps are marked in green, starting at the top left. (1) After computing the transmission envelopes, initial estimates of the refractive index n_1 are determined at the tangent wavelengths λ^{tan}. (2) A system of two equations (based on the interference formula) is solved at each pair of adjacent tangents point to estimate the film thickness, resulting in a set of possible thickness $\mathbf{d_1}$. The expected thickness value is calculated as $d_1 = \text{mean}(\mathbf{d_1}) \pm \text{std}(\mathbf{d_1})$. (3) The interference formula is then applied again to compute the interference orders $\mathbf{m_1}$. (4) These calculated decimal values should theoretically be half-integers, so they are updated to the nearest corrected values $\mathbf{m_2}$. (5) With these corrected interference orders, a more precise thickness estimate $d_2 = \text{mean}(\mathbf{d_2}) \pm \text{std}(\mathbf{d_2})$ is obtained using the interference formula. (6) This refined thickness allows for updated refractive index values n_2 at the tangent points. (7) Finally, an accurate extinction coefficient k_2 is computed at the tangent points using the precise d_2, the updated refractive index n_2, and information from the envelopes. 123
3.13 Graphical illustration of the iterative optimization process used to determine the upper envelope. The algorithm identifies upper tangent points so that the interpolated curve lies entirely above the experimental transmission data, ensuring that each concave region contains a unique tangent point. During the initial iteration (red envelope), the randomly selected tangent points yield a high cost (see bottom panel). In subsequent iterations (green dashed envelope), the cost gradually decreases, allowing the solution to converge to the desired upper envelope. 129
3.14 Envelopes and tangent points determined using our novel optimization procedure. In particular, (a) shows the reconstructions for simulated sample and (b) a magnified view of its strong absorption region, while (c) and (d) show transmittance curves for the real samples #1 and #2, respectively. These results satisfy the two constraints imposed in our optimization problem: first, the upper (lower) envelope consistently lies above (below) the transmittance curve, touching it only at upper (lower) tangent points; second, these tangent points are localized within the light-blue concave (upper) or light-green convex (lower) regions. 134
3.15 Refractive index and extinction coefficient for *a*-Si samples #1 and #2, found with the inverse synthesis approach (solid lines) and with our proposed improved Swanepoel method (circles). 144
3.16 (a) Geometrical model of the sample and illustration of the transmission spectroscopy experiment; (b) Simulated transmission spectra with its ground-truth upper envelope and the envelope predicted by the CNN model; (c) CNN architecture. 149

3.17 (a) Sample diagram (b) Simulated transmission spectra for representative samples. (c) Table with parameter boundaries and errors. (d) Model architecture. Operations: 1D convolutions, dense layers, LSTM module. (e) Training and validation cost history. 152

4.1 (a) Basic planar detector configuration. The absorption of incidence photon creates electron-hole pairs, and the free charges move through the bulk of the detector drift by the external E-field. (b) Simulated signals at the anode. We have used the specifications from [284] for CZT detectors. The simulator includes the charge drift due to the external field and the charge fluctuations (recombination, trapping and detrapping). (c) Spread of electrons in time due to diffusion and repulsion. If we compensate for the transport due to external E-field, the spread appears spherically symmetric around a central point. (d) Projected charge density on the $x-$direction at two different times. 159
4.2 Numerical solutions from the IFDM for diffusion-only, repulsion-only, and diffusion-repulsion combined. Further description provided in Section 3.3. 169
4.3 Root mean square for diffusion-only, repulsion-only, and diffusion-repulsion combined. We can see that the quadratic sum of diffusion-only and repulsion-only differs from the actual spread of the diffusion-repulsion processes. Parameters: $\mu = 1000$ cm/Vs, $D = 25$ cm^2/s, $P = 2 \times 10^5$ electrons, $\epsilon_r = 11$ relative permittivity, initial standard deviation $\sigma_0 = 10$ µm. 172
4.4 (a) Charge distribution (non-normalized) accounting for the total amount of elementary charges, plot at different times. We see the ground-truth distributions and those generated with the Benoit-Hamel model and with our proposed Generalized Normal Distribution (GND) model. (b) Error metrics of the BH and our GND models. 175
4.5 Monte Carlo simulation with $S = 200$ representative samples, using the simple Gaussian model (BH) and the GND model (ours). 177
4.6 Gaussian Mixture Model (with two Gaussian functions) matching the PDF of the charges at two different times. 183
4.7 Illustrating the strong linear correlation between the time-dependent parameters (a) μ^{GND} and σ^{BH}, (b) α^{GND} and σ^{BH}, for a representative scenario with $P = 2\times10^4 \in [2\times10^4, 2\times10^5]$ electron-hole pairs and with an electron mobility of $\mu = 1113 \in [700, 1300]$ cm^2/Vs. (c) Time-average β^{GND} for different (P, μ) scenarios compared with our proposed model from Eq. 4.33. . 185
4.8 Pixelated detector geometry with one cathode on the top and nine anodes on the bottom. 191
4.9 Weighting potential for anode k at the edge of the pixel array grid. (a) 3D view of the weighting potential. (b) 2D slice of the weighting potential at $x = 0.70$ cm. 194

4.10 Diagram illustrating the forward and inverse problems. In the forward problem (a), the model serves as a digital twin of the photon-counting detector, providing an accurate simulation of the detector signals based on the crystal properties, such as charge mobility (μ_e), or equivalently, the computational parameter R_e (mobility ratio). Conversely, the inverse problem (b) aims to deduce the material properties from the available information on the charges and resulting signals. 195
4.11 Comparison of ground-truth and predicted parameters for (a,b) noise-free, and (c,d) noisy input data. (a,c) Mobility ratio, and (b,d) electron trapping parameter. 197
4.12 (a) Average error in predicted parameters with varying noise levels in the input data. (b) Ground-truth signal compared to signals from predicted parameters under high noise conditions (std = 2%). 199
4.13 (a) Pixelated CdZnTe unit representation. (b) Signal at different electrodes with noise. (c) DL Architecture. (d) Normalized log amplitude of the signals. (e) Training history cost. 205

A.1 Setup diagram for the novel camera based on the synthetic wavelength concept. (a) The 3D object is illuminated with an object beam that contains light waves at wavelengths λ_1 and λ_2. The field reflected in the object will pass through an imaging system and reach the camera sensor chip. In addition, two reference beams directly illuminate the sensor. (b) The frontal view of the optical system: Looking at the objective from the sensor location. The horizontal reference uses wavelength λ_1 and vertical reference λ_2. We control the angle at which the reference beams reach the sensor by modifying the distances d_{hor} and d_{vert}. 241
A.2 (a) Photograph of our particular setup. Two identical tunable lasers will emit light in the NIR range at around 850 nm. We use optical fibers and beam splitters (BS) to direct the beams flexibly. We employ a 4f-system to image the object into the sensor array. (b) Diagram of the illumination engine. 245
A.3 (a) Image captured with our camera sensor. The spatial FoV of the camera is around 66mm × 66mm and the pixel resolution of the captured images is 1448×1448 pixels. We have cropped a window in the image at the central location pixels so that one can clearly see (b) the speckles (red window) and (c) the crossed vertical and horizontal fringes (blue window). The distance between fringes is about 3 pixels, near the sampling limit. 249
A.4 (a) Fourier transform of the captured image $I(x, y)$. (b) Shifting the Fourier spectrum so that the carrier frequency f_1 corresponding to the horizontal reference beam is displaced to the center. (c) Applying a Gaussian filter to the shifted Fourier spectrum to uniformly remove any signals outside the region. 250

A.5 (a) Amplitude and (b) phase of the object beam reflected by the object with wavelength λ_1. Similarly, (c) and (d) represent the amplitude and phase of the object beam reflected by the object with wavelength λ_2. Then, (e) and (f) show the calculated synthetic wave, Λ. 251
A.6 Limitations of the double-shot method: The top row displays an analysis for a synthetic wavelength $\Lambda = 5$ mm with translations d of (a) 0, (b) 0.01, (c) 0.04, and (d) 0.16 mm. Similarly, the bottom row presents an analysis for $\Lambda = 30$ mm with translations of (e) 0, (f) 0.01, (g) 0.04, and (h) 0.16 mm. . . 255
A.7 Double-shot acquisition mode. The measured object is a clay pot approximately 20 cm in height. The camera FoV was cropped to about 40×40 mm^2: (a) Image of the small object with a window (red) showing the FoV. (b) Wrapped phasemap $\phi(\Lambda = 3\text{mm})$. (c) Acquired synthetic phasemap $\phi(\Lambda = 20\text{mm})$. (d) Phasemap $\phi^{\text{unwrap}}(\Lambda = 3\text{mm})$, unwrapped with our multi-frequency unwrapping algorithm. (e) 3D model of the object calculated from $\phi(\Lambda = 20\text{mm})$, and (f) from unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 3\text{mm})$. 259
A.8 Single-shot acquisition mode. The measured object is a plaster bust of approximately 6 cm in height. (a) Image of the small object with a window (red) showing the FoV. (b) Wrapped phasemap $\phi(\Lambda = 5\text{mm})$. (c) Acquired synthetic phasemap $\phi(\Lambda = 50\text{mm})$. (d) Phasemap $\phi^{\text{unwrap}}(\Lambda = 5\text{mm})$, unwrapped with our multi-frequency unwrapping algorithm. (e) 3D model of the object calculated from $\phi(\Lambda = 50\text{mm})$, and (f) from the unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 5\text{mm})$. 261
A.9 Architecture of the Deep Learning algorithm. 262
A.10 Simulated data for the training of the Deep Learning unwrapping algorithm. We employed *Mitsuba* [351] to simulate realistic 3D objects. (a) Input: Wrapped noisy phasemap. (b) Output: Unwrapped phasemap. 263
A.11 Single-shot deep learning-based phase unwrapping method. (a) Wrapped phasemap $\phi(\Lambda = 10\text{mm})$. (b) Unwrapped phasemap predicted by the neural network approach. (c) Image of the small object with a window (red) showing FoV. (d) 3D model of the object calculated from unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 10\text{mm})$. 265

List of Tables

2.1 Depth precision analysis comparing the double-shot with the single-shot method. 47

3.1 Set of transmittance and reflection expressions analyzed in the present work, each with their corresponding levels of accuracy and approximations. 68
3.2 Errors associated with the approximated formulae. When $\Delta d = 0$ the formulae for uniform films are used, tested against the exact formulae from Eqs. 3.29 and 3.30. For $\Delta d = 10^{-5}$ nm, the formulae for wedged films are applied, demonstrating consistency as $\Delta d \to 0$. They were tested against the new formulae from Eqs. 3.37 and 3.45. Note that the two left blocks assume $\kappa_2 = 0$, while the two right blocks use the Swanepoel approximations ($\kappa_2 = 0$ and $n_1^2 \ll k_1^2$). 86
3.3 Optimized parameters for amorphous silicon samples obtained from the inverse synthesis method. The initial bounds for each parameter are also provided. 120
3.4 Results of the Swanepoel method applied to the transmission spectra of **Sample #1**. Estimated thicknesses: $d_1 = 1099.32 \pm 35.77$ nm (relative error = 3.25%), $d_2 = 1110.18 \pm 7.22$ nm (relative error = 0.65%). The last two data points are crossed out due to unreliable calculations at shorter wavelengths. 125
3.5 Results of the Swanepoel method applied to the transmission spectra of **Sample #2**. Estimated thicknesses: $d_1 = 1167.28 \pm 62.92$ nm (relative error = 5.39%), $d_2 = 1163.88 \pm 10.52$ nm (relative error = 0.90%). The last two data points are crossed out due to unreliable calculations at shorter wavelengths. 126
3.6 Comparing our proposed method with respect to the standard McClain algorithm within the same spectral region (between the first and last local extrema). 138

4.1 Comparison of metrics: Fitting ground-truth distributions using the simple Gaussian model by BH and our proposed GND method. 180

4.2 Common material properties of CdZnTe crystals [284], along with the corresponding dimensionless computational parameters used in the simulations. The table also shows the NRMSE between the predicted parameters and the ground-truth ones when employing noise-free and noisy input data. 193

Chapter 1

Introduction

Traditionally, the development of imaging and sensing technologies has relied on the foundation from optical physics and mathematical modeling. Over the last few decades, the increasing availability of computational resources, numerical methods, and large datasets has also opened the door to data-driven and advanced optimization approaches. These techniques can adapt to realistic noise, can model complex semi-empirical physical knowledge, and scale to high dimensional data. Across diverse fields (such as biomedical imaging, materials characterization, and astrophysics), sensing applications are becoming increasingly complex, driving the demand for novel computational methods that integrate the rigor of physics-based modeling with the flexibility and scalability of machine learning.

Our general approach combines the simulation of forward models that describe the physical process with inverse methods that estimate quantities of interest through optimization or learning. Central to this framework is the development of customized differentiable simulators, which not only allow for the generation of synthetic data for supervised training, but also enable the direct estimation of physical parameters via automatic differentiation. These simulators can be further embedded within machine learn-

ing models to constrain the solution space and enforce physical consistency during training. In this way, the thesis bridges analytical modeling and modern learning-based inference, providing robust and interpretable solutions to challenging inverse problems. The thesis is organized around three primary application areas, which are outlined below. The work presented in this thesis is based on a series of publications authored by the candidate, forming a coherent narrative centered around physics-informed computational imaging and inverse problem solving.

The first part focuses on the development of a single-shot 3D imaging system based on Synthetic Wavelength Interferometry (SWI) [1]–[3]. In this approach, a scene is illuminated with two coherent laser beams of slightly different wavelengths in the near-infrared range, and the interference pattern is recorded using a conventional CMOS sensor. Classical interferometric algorithms allow for the extraction of optical fields at each wavelength. However, due to the sub-micron scale of the optical wavelengths, the recovered phase information is severely affected by speckle noise, particularly when imaging everyday diffuse objects with surface roughness exceeding the wavelength. To overcome this, we computationally synthesize the beat wave formed by the superposition of the two original optical waves. This beat produces a synthetic wavelength much longer than the individual optical wavelengths. Therefore, the synthetic wave exhibits reduced speckle noise and preserves depth information, allowing for high-precision 3D surface reconstruction. The achievable accuracy depends on the synthetic wavelength selected, which can be typically tuned from a few microns to several centimeters depending on the application. One of the key challenges in this technique is phase unwrapping: recovering the continuous depthmap values from the wrapped synthetic phase, particularly under noisy conditions. To address this, we employ a customized learning-based single-shot unwrapping algorithm.

The second core project turns to the optical characterization of thin solid films using transmission and reflection spectroscopy. Thin-film metrology plays a central role in materials science, yet traditional methods often rely on idealized assumptions or manual tuning, limiting their applicability in real-world experimental conditions and mass characterization for industrial applications. We begin by revisiting classical curve-fitting approach, where material parameters are iteratively adjusted until the theoretical spectral response aligns with measured data. We extend the analytical formulation [4] of the forward model to handle more complex film geometries (such as highly wedged surfaces) and non-ideal substrates that exhibit partial absorption. These modifications improve the realism and fidelity of the simulations, allowing the forward model to serve as a digital twin of the spectrophotometric setup. For the inverse problem, we enhance the classical Swanepoel (envelope) method [5] using a global optimization procedure to estimate the envelope boundaries with greater robustness. We then transition to data-driven techniques [6], [7]: first, by emulating the envelope extraction process using convolutional neural networks, and later by directly solving the inverse problem via supervised learning, mapping transmission spectra to film parameters.

The third core project centers on the modeling and characterization of cadmium zinc telluride (CdZnTe) gamma-ray detectors. These detectors are widely used in nuclear medical imaging but are sensitive to bulk material defects and inhomogeneities in the material properties. First, we construct a precise digital twin of the detector using a physics-based forward model [8], [9] that incorporates the model for 3D charge transport (drift dynamics, thermal diffusion, and Coulomb repulsion), charge fluctuations (recombination, trapping, and detrapping), and signal generation (based on the Shockley-Ramo theorem). This simulator reproduces realistic electrical signals generated by individual $\gamma-$photon interactions. The forward differentiable model is then used to solve the inverse

problem [10]: localizing the 3D position of the gamma-ray event and estimating the deposited energy. We employ both data-drive solvers and optimization based approaches, with automatic differentiation and global solvers such as the Slime Mold Algorithm.

In addition to these three core domains, my Ph.D. journey also encompasses several complementary projects exploring broader applications of computational imaging and inverse modeling. These include: the identification and classification of modern art fibers using polarized light microscopy [11]; the development of 3D dynamic holographic displays based on spatial light modulators with extended etendue [12]; ellipsometric studies of thin films using polarized spectroscopy [13], [14]; the preservation of deteriorated Dali holograms through 3D generative AI models from multi-angle 2D images; and the application of physics-informed neural networks to the study of stellar structure via differential equation solvers in astrophysical imaging.

Chapter 2

Single-shot interferometric camera

We present a novel single-shot interferometric ToF camera targeted for precise 3D measurements of dynamic objects. The camera concept is based on Synthetic Wavelength Interferometry, a technique that allows retrieval of depth maps of objects with optically rough surfaces at submillimeter depth precision. In contrast to conventional ToF cameras, our device uses only off-the-shelf CCD/CMOS detectors and works at their native chip resolution (as of today, theoretically up to 20 Mp and beyond). Moreover, we can obtain a full 3D model of the object in single-shot, meaning that no temporal sequence of exposures or temporal illumination modulation (such as amplitude or frequency modulation) is necessary, which makes our camera robust against object motion.

In this chapter, we introduce the novel camera concept and show first measurements that demonstrate the capabilities of our system. We present 3D measurements of small (cm-sized) objects with > 2 Mp point cloud resolution (the native pixel resolution of our used detector) and up to sub-mm depth precision. We also report a "single-shot 3D video" acquisition and a first single-shot "Non-Line-of-Sight" measurement. Our technique has great potential for high-precision applications with dynamic object movement, e.g., in AR/VR, industrial inspection, medical imaging, and imaging through scattering

media like fog or human tissue.

2.1 Introduction

3D imaging techniques are progressively replacing conventional photographs, which cannot capture the complexity of real-world objects. A 3D model provides information about an object's surface shape, slope, or texture - essential properties in object analysis. Once we obtain a realistic 3D model of an object, we can reproduce it under different illuminations and rigid transformations using computer graphics techniques. These features make 3D imaging very useful for many applications, including medical diagnostic imaging, quality inspection of manufactured products, and autonomous driving.

There are numerous 3D imaging techniques in the literature, which can be mainly classified into three groups. The first group corresponds to triangulation-based techniques, which are one of the most commonly used techniques in industrial inspection or 3D metrology on macroscopic rough surfaces. Examples include active and passive stereo, line- or fringe-projection techniques, or focus-searching methods [15]–[19]. However, the depth precision of related methods scales with the stand-off distance, and occluded parts of the object cannot be measured. The second group refers to all reflectance-based methods, which typically use specific lighting conditions and then capture the light reflected or scattered off the object's surface to estimate its surface slope. Some examples include photometric stereo and deflectometry [20]–[25]. These systems often become highly complicated to calibrate, and the integration from the surface normal to surface shape is problematic in some cases. The third category corresponds to all Time-of-Flight-based (ToF) approaches [26]–[28], that calculate the optical pathlength between camera and object or scene to obtain its shape. ToF cameras of all types (pulsed, amplitude mod-

ulated, etc) and optical interferometers are prominent examples for this group. While conventional ToF cameras directly measure the travel time of light, interferometric cameras measure the optical pathlength difference via interference patterns produced by an object and a reference beam. ToF-based approaches are beneficial for many applications because they are robust to occlusions and the data precision does not depend on the standoff distance of the camera.

As mentioned, conventional “ToF cameras" typically exploit amplitude-modulated light sources either as pulsed illumination (LIDAR) [29], [30] or in the form of “continuous waves with amplitude modulation" (CWAM) [31]–[33]. These cameras generally require specific sensor architectures with a high temporal resolution. The main restriction of these cameras is their (still) fairly limited pixel resolution and their low depth precision, which is roughly on the order of centimeters. This makes conventional ToF cameras useful for basic estimation or detection tasks of larger objects (for example, to know if there is an object in front of an autonomous driving vehicle, or to estimate the pose of a human). However, they cannot be used for applications that require high data precision, such as specific tasks in medical imaging, industrial inspection, or optical metrology. Moreover, some ToF camera schemes require a sequence of exposures to calculate a 3D scene or exploit raster-scanning of a laser dot (i.e., are not “single-shot"). That makes them susceptible to fast object motions.

On the other hand, the main drawback of standard (single-wavelength) interferometric ToF cameras is that they are limited to the measurement of objects with optically smooth (specular) surfaces [34]–[36]. When an optically rough surface is illuminated with coherent light, the backscattered light field forms a “speckle pattern". Since the phase of this speckle field is randomized, we cannot recover information about the optical path-

length difference and hence, the depth map of the object. In practice, standard interferometric techniques become helpful only in particular fields, e.g., to study the nanomechanical motion of some objects, or for high-precision surface testing of lenses or smooth technical parts.

Multifrequency interferometric ToF cameras [35], [37]–[42] have shown that the interferometric principles can also extend to macroscopic objects with rough surfaces. These methods use the information from two (or more) interferometric measurements at different optical wavelengths to disambiguate the *random-phase fluctuations* of the object beam. Nevertheless, recent implementations of these techniques for Computer Vision applications [26], [43]–[46] still require sophisticated sensors or are limited to static objects (that is, they are not motion-robust).

In this contribution, we present a novel multifrequency interferometric ToF camera concept and device that overcomes the described limitations of the state-of-the-art (SOTA) ToF cameras shown above. Our demonstrated camera prototype scans object surfaces in single-shot with up to submillimeter depth precision. Here, “single-shot" means that only one camera image is necessary to generate a full-field high-quality 3D model. This feature makes our ToF camera robust against object motion. Moreover, the camera only uses off-the-shelf CMOS or CCD detector technology, that is, it theoretically even works with standard smartphone cameras. No specialized detector architecture (like PMDs [31], SPADs [47], or lock-in detection [48]) is required. The lateral point cloud resolution of our camera is equal to the pixel resolution of the used CMOS/CCD chip - theoretically up to 20Mp and beyond, given the current state of the art in CMOS camera technology [49]. These capabilities give our camera significant advantages with respect to the current SOTA in ToF camera technology and allow for many high-precision 3D imaging applications that

previously were not possible with ToF cameras. We summarize the specific contributions of our work as follows:

- We devised and developed a novel ToF camera technology. This technology is based on a previously introduced method called “Synthetic Wavelength Interferometry" (SWI) [26], [37]–[42]. Compared to previous demonstrations of SWI for Computer Vision applications [43]–[46], our method allows to scan optically rough object surfaces in single-shot using only off-the-shelf CMOS/CCD camera technology.

- We demonstrate ToF-based 3D measurements of static objects with rough surfaces with up to 330 μm depth precision and 2 Megapixel lateral point cloud resolution (equivalent to the pixel resolution of the employed CMOS camera). The measurements shown in Figs. 2.3 and 2.4 exploit multifrequency phase unwrapping procedures [50] to increase the depth sensitivity of the sensing principle, which means that these measurements are not single-shot. Later, Figs. 2.5 and 2.6 show real single-shot measurements. Finally, we demonstrate the 3D video reconstruction of a moving object from sequential single-shot measurements. To the best of our knowledge, this is the first demonstration of an SWI-based video sequence for Computer Vision applications.

- To combine the high-depth precision of multifrequency measurements with the motion robustness of single-shot measurements, we have developed a deep-learning-based algorithm for single-shot phase unwrapping. Compared to other SOTA phase unwrapping procedures [51]–[54], our algorithm is specifically tailored to our acquired data structures and integrates seamlessly into our data processing pipeline. We demonstrate first single-shot measurements unwrapped with our algorithm.

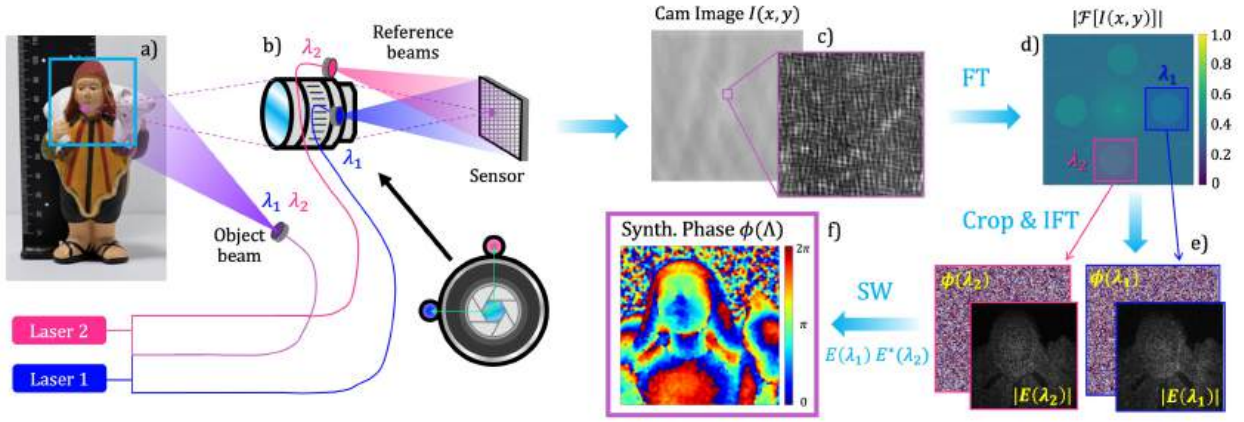


Figure 2.1: Schematic setup and computational procedure of our novel single-shot ToF camera system. (a) Object (the camera FoV is 66 mm $\times$ 66 mm). (b) The lens system images the object onto a CCD/CMOS chip. An object beam at two wavelengths illuminates the object and two reference beams at two different wavelengths directly illuminate the chip at two different (vertical and horizontal) angles. (c) The captured image I(x,y) along with a zoom-in window. The image consists of speckles that are overlayed with crossed fringes. (d) Fourier transform of the captured image. (e) By shifting, filtering, and inverse Fourier transform, the speckled complex fields (amplitude and phase) at both optical wavelengths (see section 2.2) can be recovered from only one camera image. (f) Using the SWI principle, we generate the synthetic phasemap $\phi(\Lambda)$ to calculate the 3D depth map of the object.

- We use our novel single-shot technique to acquire a “Non-Line-of-Sight" measurement of a point object hidden behind a scatterer. To the best of our knowledge, this experiment documents the first ever SWI-based single-shot measurement through scattering media.

2.2 Principles and computational framework

2.2.1 ToF camera concept

Our ToF camera concept is based on Synthetic Wavelength Interferometry (SWI), a method for multifrequency interferometry [37]–[42]. SWI exploits the spectral diversity of multiple measurements at multiple wavelengths to image objects with optically rough sur-

faces. The basic principle is summarized as follows: An optically rough surface is illuminated with coherent light (at wavelength λ_1), and the complex field $E(\lambda_1)$ scattered off the object's surface is measured using an optical interferometer. This field exhibits strong wavefront aberrations due to the mentioned speckle. Macroscopic optical pathlength information (i.e., the shape of the object) can not be recovered from $E(\lambda_1)$ (see Fig. 2.1e and Appendix A). Eventually, the static object is again illuminated with a slightly different wavelength λ_2, and a new field $E(\lambda_2)$ is obtained. Assuming that both illumination sources originate from the exact same location (e.g., the same fiber tip), the fields $E(\lambda_1)$ and $E(\lambda_2)$ are subject to the same microscopic and macroscopic pathlength variations. After calculating the difference between their phasemaps $\phi(\lambda_1) - \phi(\lambda_2)$, the phase aberrations imparted by the microscopic pathlength variations cancel each other out. The phasemap difference only contains the macroscopic pathlength variations on the order of a "Synthetic Wavelength" $\Lambda = \frac{\lambda_1 \lambda_2}{|\lambda_1 - \lambda_2|}$ (see Fig. 2.1f and Appendix A).

If λ_1 and λ_2 are spaced close enough (see analysis in [26], [55], [56]) the resulting "synthetic field" $E(\Lambda)$ does not exhibit speckle artifacts and can be processed like a "normal" ToF camera image or unspeckled optical interferogram. This means that the object's depth can be calculated by

$$z = \frac{1}{2} \frac{\phi(\Lambda) \cdot \Lambda}{2\pi} \quad . \tag{2.1}$$

One possible way to calculate the synthetic field $E(\Lambda)$ is *computational mixing* of $E(\lambda_1)$ and $E(\lambda_2)$,

$$\begin{aligned} E(\Lambda) = E(\lambda_1) \cdot E^*(\lambda_2) &= |E(\lambda_1)| \cdot |E(\lambda_2)| \cdot e^{i(\phi(\lambda_1) - \phi(\lambda_2))} \\ &= |E(\lambda_1)| \cdot |E(\lambda_2)| \cdot e^{i\phi(\Lambda)} \end{aligned} \tag{2.2}$$

where $E^*(\lambda_2)$ denotes the complex conjugate of $E(\lambda_2)$.

Recent work [43], [44], [46], [55] has introduced different ToF cameras for Computer Vision applications which are based on this synthetic wavelength principle. Despite the ability for “full-field" measurements and measurements showing high-depth precision, the introduced systems had one significant drawback: *at least two* sequentially captured camera images were needed to reconstruct the object’s depth maps. One would capture the field $E(\lambda_1)$ and then repeat the measurement using a different wavelength to obtain $E(\lambda_2)$. This approach works for static scenes but completely fails for dynamic scenes. If the object moves between the two sequential images, the optical pathlength variances between both images will differ. Therefore, the correlation between both speckle fields (needed to calculate the synthetic wavelength term) is lost. It turns out that this problem is much more severe than “conventional" motion artifacts (such as those from structured light imaging systems), as minimal object movements can induce significant changes in the observed speckle patterns. This presents a severe drawback because many real-world applications include scenes that naturally move at least on a microscopic scale (e.g., in medical imaging, autonomous driving, or AR/VR). We refer to section A.3 of Appendix A for related experiments and further description on this problem.

2.2.2 Proposed method

The significant limitation of current SWI principles has stimulated the idea for this contribution. Our novel idea draws inspiration from spatial heterodyning, off-axis holography, and single-sideband demodulation procedures [57]–[68]. Our new setup configuration (see Fig. 2.1a and 2.1b) only uses standard CCD/CMOS camera technology and can operate in single-shot. Our approach can be summarized as follows: We couple a portion of

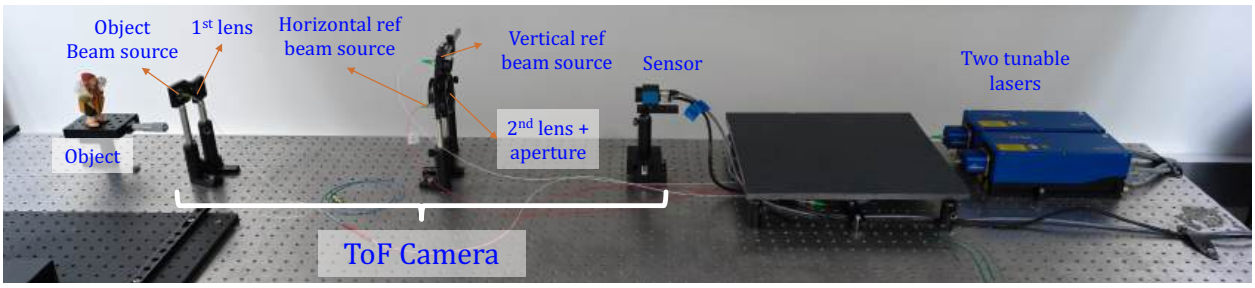


Figure 2.2: Picture of our lab setup. Two identical tunable lasers emit light in the NIR range at around 850 nm. We use optical fibers and fiber beam splitters (BS) to direct the beams flexibly. We employ a lens system plus aperture to image the object into the sensor array.

two laser beams (at two different wavelengths λ_1 and λ_2) together, which forms the object beam. The other portions of both beams remain separated and form two reference beams (see Fig. 2.1b). The reference beams are arranged so that they directly illuminate the camera chip at an angle. One reference beam (e.g., λ_1) encloses an angle with the horizontal x-axis, while the other reference beam (e.g., λ_2) encloses an angle with the vertical y-axis of the detector.

When the object beam is scattered off the object surface, it forms a speckle field at each wavelength: $E(\lambda_1)$ and $E(\lambda_2)$. These two speckle fields are incident on the detector and interfere with the reference beams. The speckle field at λ_1 produces a static interference pattern with the λ_1-reference beam (vertical fringes), while the speckle field at λ_2 produces a static interference pattern with the λ_2-reference beam (horizontal fringes). In other words, the camera image $I(x, y)$ (Fig. 2.1c) shows speckles that are overlayed by crossed fringes. For the sake of completeness, it should be mentioned that interference also happens between the fields at λ_1 and λ_2. These interferences result in temporally oscillating fringes which oscillate much faster than the integration time of our used camera and hence are not further discussed here (see Appendix A for details).

After acquiring the image $I(x, y)$ in single-shot, the complex speckle fields (phase

and amplitude) $E(\lambda_1)$ and $E(\lambda_2)$ required to form the synthetic field $E(\Lambda) = E(\lambda_1) \cdot E^*(\lambda_2)$ are retrieved via computational demodulation in the Fourier domain: Fig. 2.1d shows the 2D Fourier transform of the captured image, $\mathcal{F}[I(x, y)]$, where five spectral regions with high energy can be distinguished. The central region represents the DC component. The horizontal regions (left and right) are centered around the spatial carrier frequencies of the vertical fringes, which appear due to the interferences between the λ_1-reference beam and the object field $E(\lambda_1)$. Similarly, the vertical regions (up and down) are centered around the carrier frequencies for the horizontal fringes produced by the reference and object field at λ_2.

Eventually, the complex field $E(\lambda_1)$ is retrieved by performing the following operations:

- Find/evaluate the carrier frequency f_1 for the respective spectrum in the Fourier domain (blue box in Fig. 2.1d).
- Shift the Fourier spectrum to set the evaluated carrier frequency f_1 as the new center frequency (see Appendix A for further details).
- Filter the spectrum (either with a Hanning window or Gaussian kernel) so that only the frequency band around the new center frequency remains. The resulting filtered spectrum is denoted by $\mathcal{F}_{\text{hor}}[I(x, y)]$.
- An Inverse Fourier Transform (IFT) of $\mathcal{F}_{\text{hor}}[I(x, y)]$ eventually delivers the phase $\phi(\lambda_1)$ and amplitude $|E(\lambda_1)|$ of the complex field $E(\lambda_1)$ (see Fig. 2.1e and references [57], [59], [60]):

$$\mathcal{F}^{-1}\{\mathcal{F}_{\text{hor}}[I(x, y)]\} \propto |E(x, y; \lambda_1)| \exp\left(i\phi(x, y; \lambda_1)\right) \tag{2.3}$$

The complex field $E(\lambda_2)$ is retrieved from the *same image* in an analog fashion, using the vertically arranged regions in the Fourier spectrum. Eventually, the synthetic field $E(\Lambda)$ is formed via Eq. 2.2, the synthetic phase $\phi(\Lambda)$ is extracted, and the depth map of the object is calculated via Eq. 2.1. More details and mathematical calculations can be found in the Appendix A.

2.3 Experiments and Data Processing

2.3.1 Experimental setup

We first discuss the specific configuration of the setup used for our experiments, shown in Fig. 2.2. The optical system consists of a lens system that focuses the imaged object onto the camera chip. To incorporate the reference arms, it is required to leave a few centimeters between the sensor chip and its closest lens. This increases the back focal length of the system, affecting the Field of View (FoV) of the camera and making this first optical prototype system less compact. Further discussions on how to make the imaging system more compact are provided in section 2.4.

It should also be mentioned that multiple possible setup configurations exist. As mentioned before, we use a reference source aligned on the horizontal axis and another on the vertical axis (see Fig. 2.2). However, other configurations of spatially separated reference beams are possible as well. For instance, we could place two reference beams along one single axis enclosing different angles. Setups with >2 lasers (>2 wavelengths) are theoretically possible as well (e.g., for unwrapping, see sec. 2.4). The only essential condition is that the Fourier transform of the captured image (see Fig. 2.1d) must have separated spectral regions without overlap (aliasing) in distinct separated locations.

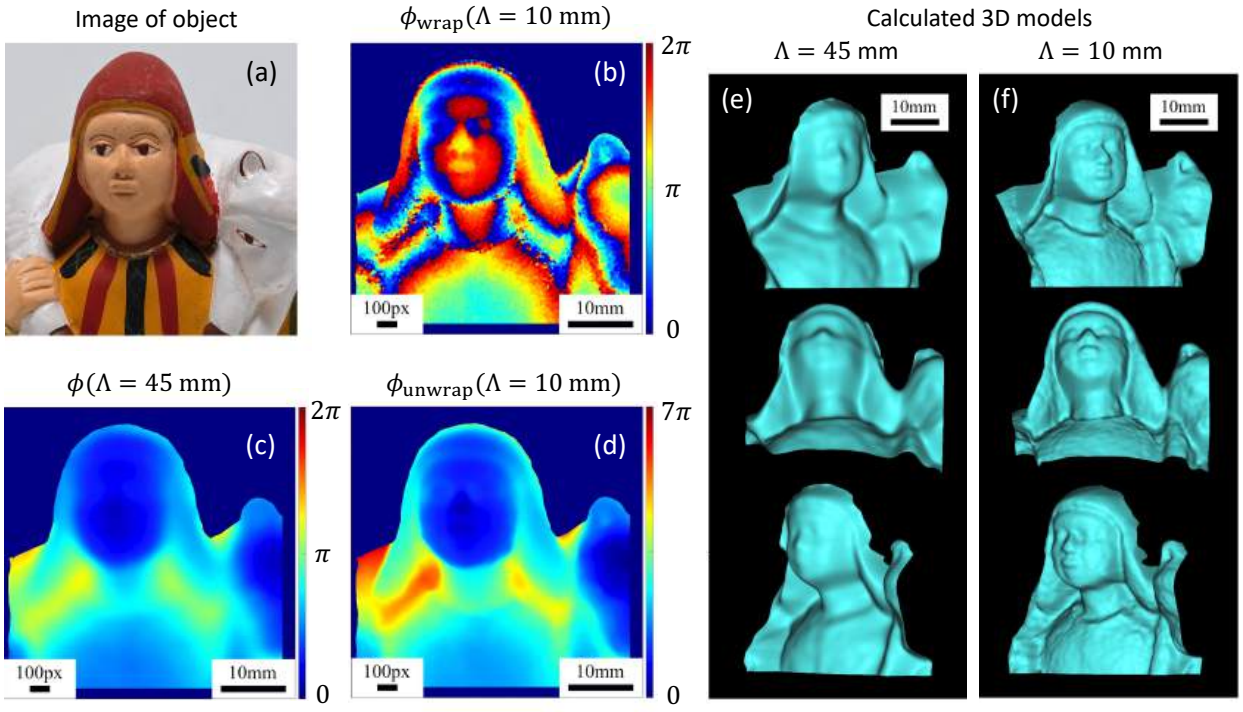


Figure 2.3: Double-shot acquisition mode. The measured object is a painted clay figure of approximately 10 cm height (see Fig 2.1a), and the field of view is 66×66 mm^2. (a) Image of the small object cropped to the actual FoV. (b) Wrapped phasemap $\phi(\Lambda = 10\text{mm})$. (c) Acquired synthetic phasemap $\phi(\Lambda = 45\text{mm})$. (d) Phasemap $\phi(\Lambda = 10\text{mm})$, unwrapped with our multi-frequency unwrapping algorithm. (e) and (f) 3D model of the object calculated from $\phi(\Lambda = 45\text{mm})$ (e) and from unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 10\text{mm})$.

To facilitate the reference and object beams at λ_1 and λ_2, two highly coherent light beams produced by two tunable lasers (*Toptica DFB pro* devices) are flexibly directed by optical fibers. We separate the light emitted from each laser into the object and reference arms using a fiber splitter, where 90% of the power goes to the object and 10% to the reference. In addition, we combine the two object arms so that the resulting object source contains both wavelengths, λ_1 and λ_2 (see Appendix A for further details).

Although the object beam has much more intensity than the two references, we must consider that the light scattered off the object's surface that passes through the optical system is considerably less intense than the initial object beam. ND filters were used

on the reference beams to match the intensity that reaches the sensor from the references and the object beam. When they match, the interference fringes have maximum contrast: destructive interferences with almost zero intensity and constructive fringes with four times the intensity of each beam [69].

Additionally, we place an aperture right behind the objective lens (see Fig. 2.1b) to control the size of the subjective speckles at the sensor location. This is essential, as our method requires several interference fringes within one speckle, so the speckle size needs to be controllable. Respective trade-offs related to the aperture size and the lateral resolution of our camera system will be discussed in sec. 2.4.

In the following, we use our described setup prototype to capture 3D images and 3D videos of different objects, using different single-shot and double-shot acquisition modes.

2.3.2 Double-shot acquisition

The “double-shot" acquisition mode still captures the fields $E(\lambda_1)$ and $E(\lambda_2)$ at both optical wavelengths sequentially. This means that *only one* reference arm (Fig. 2.1b) and *only one* direction in the Fourier spectrum (Fig. 2.1d) is exploited, and *two* camera images are required to calculate the synthetic field $E(\Lambda)$ (hence the word “double-shot"). However, it should be emphasized that each optical field $E(\lambda_1)$ or $E(\lambda_2)$ itself is acquired in *single-shot*, which is in stark contrast to conventional interferometry that normally relies on phase shifting and multiple exposures to capture one optical field.

Fig. 2.3 shows exemplarily double-shot acquisition results. The synthetic phasemap shown in Fig. 2.3a was acquired at a synthetic wavelength $\Lambda = 45$ mm. The respective calculated 3D model is shown in Fig. 2.3e from three different perspectives. Measurements at smaller synthetic wavelengths yield higher depth precision [26], [43], [44]. However,

for measurements at synthetic wavelengths Λ smaller than 2x the object depth, the reconstructed synthetic phasemaps observe phase wrapping. To solve this problem, we applied a multi-frequency unwrapping algorithm to recover unwrapped phasemaps at a small synthetic wavelength using *only a set of wrapped phasemaps* at different synthetic wavelengths. This means that a "guidance phasemap" at a large synthetic wavelength that does not observe wrapping is not required. Our applied approach works as follows: From two synthetic fields $E(\Lambda_1)$ and $E(\Lambda_2)$ with wrapped phasemaps, we calculate the beat-note field $E(\Lambda_B) = E(\Lambda_1) \cdot E^*(\Lambda_2)$, which could also be seen as the “synthetic-synthetic field". If $E(\Lambda_1)$ and $E(\Lambda_2)$ are chosen correctly, the respective phasemap $\phi(\Lambda_B)$ is not wrapped and can be now used as a guidance to unwrap fields at smaller synthetic wavelengths via the standard phase unwrapping procedure [70], [71]. Of course, this requires acquiring additional optical fields $E(\lambda_i)$ to form multiple synthetic fields at different synthetic wavelengths.

Fig. 2.3b displays a wrapped synthetic phasemap of the object acquired at a synthetic wavelength of $\Lambda = 10$ mm. Fig. 2.3d shows the same phasemap, unwrapped with the procedure described above. The respective rendered 3D model is shown in Fig. 2.3f. As expected, the 3D data display much higher depth resolution and details of the object surface. Fine details like the eyes or nose of the figure can be resolved with impressive quality (see scalebar for size comparison). Further results for the double-shot experiment are shown in the Appendix A (we encourage the reader to look at the measurement of the clay pot). It should also be mentioned that different smoothing operations are applied after each unwrapping step to facilitate “smoother" guidance phasemaps. The exact smoothing parameters are specified in Appendix A.

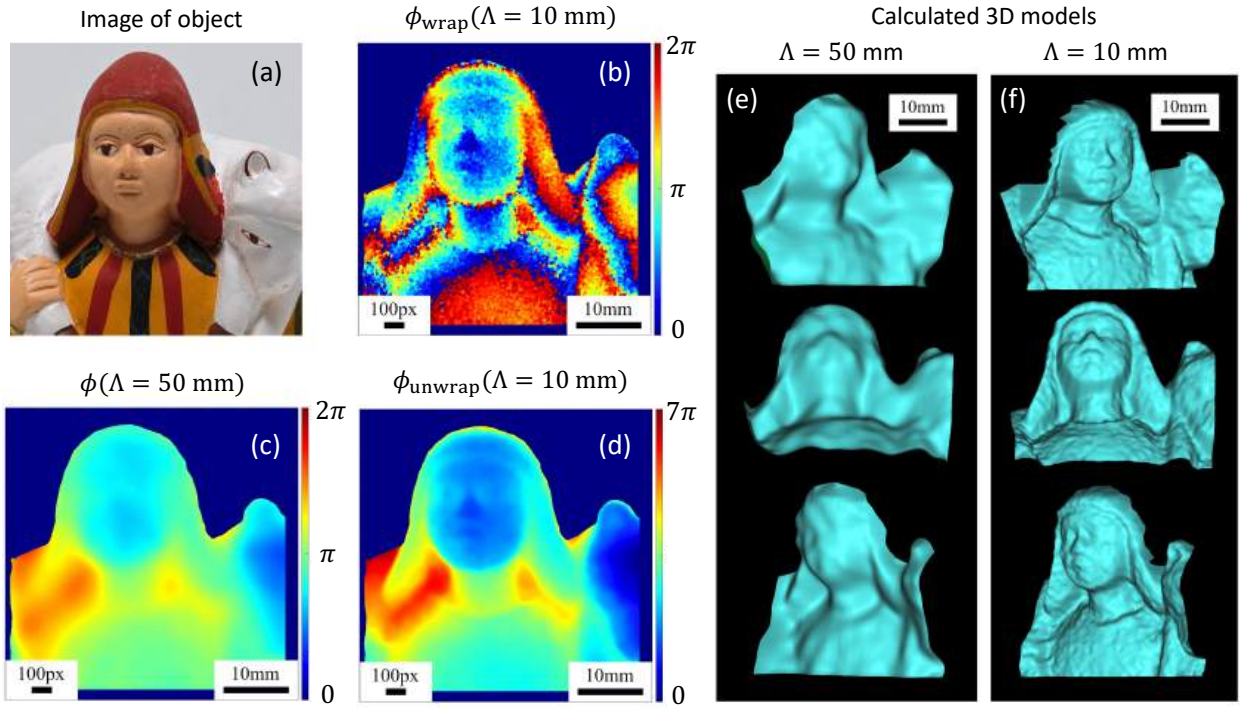


Figure 2.4: Single-shot acquisition mode. For the sake of comparability, the same object as in Fig. 2.3 is measured. (a) Image of the small object cropped to the actual FoV. (b) Wrapped phasemap $\phi(\Lambda = 10\text{mm})$. (c) Acquired synthetic phasemap $\phi(\Lambda = 50\text{mm})$. (d) Phasemap $\phi(\Lambda = 10\text{mm})$, unwrapped with our multi-frequency unwrapping algorithm. (e) and (f) 3D model of the object calculated from $\phi(\Lambda = 50\text{mm})$ (e) and from unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 10\text{mm})$.

2.3.3 Single-shot acquisition

The “single-shot" acquisition mode captures the synthetic field $E(\Lambda)$ with *only one* camera image, using the full Fourier demodulation procedure (with two reference beams) described in section 2.2. The results in Fig. 2.4 are intentionally presented for the same object and in an identical subfigure arrangement to Fig. 2.3, to allow for comparing both acquisition modes.

Figure 2.4a shows a single-shot phasemap captured at a synthetic wavelength of $\Lambda = 50$ mm - large enough that the phasemap is “unique", i.e., not subject to phase wrapping. The respective 3D model is shown in Fig. 2.4e from three different perspectives.

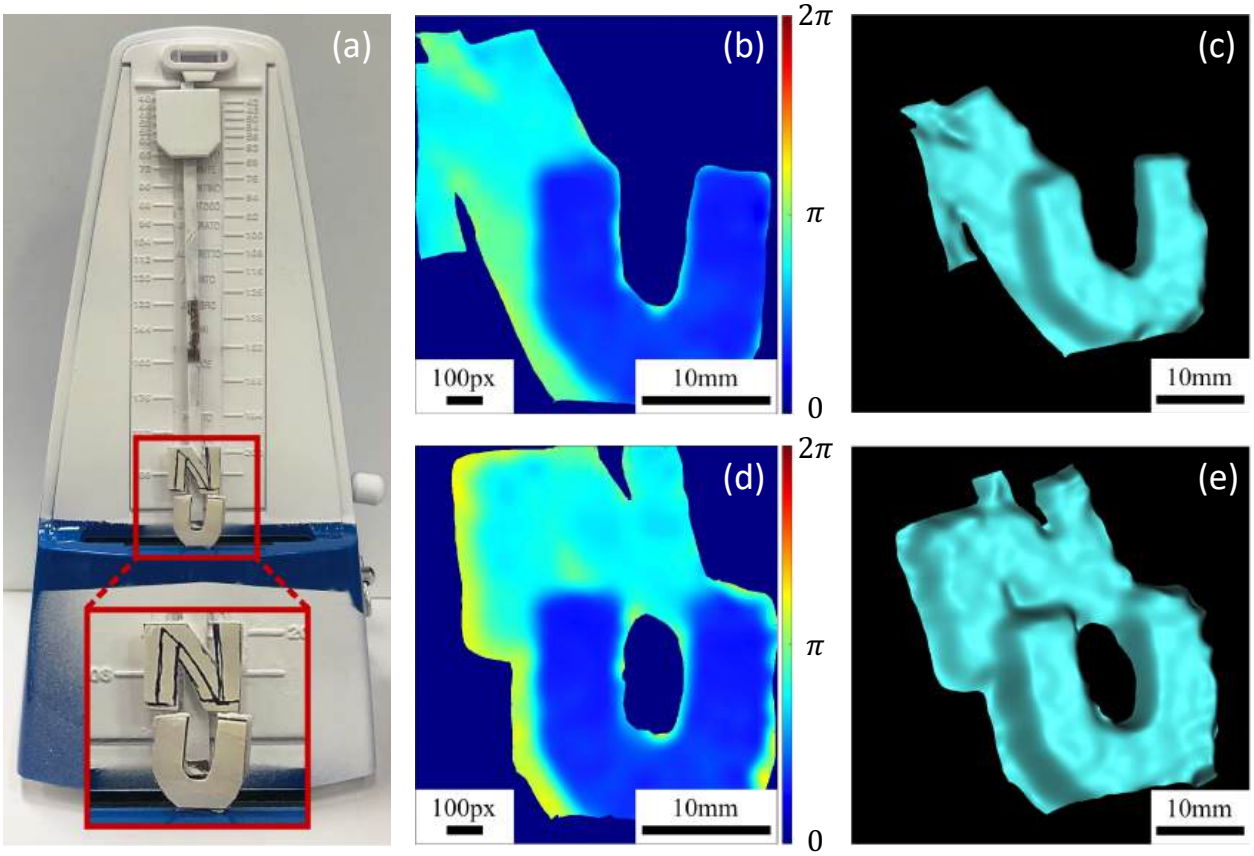


Figure 2.5: Single-shot measurements of a dynamic object. (a) Object: Metronome sprayed with a scattering powder. (b, d) Phasemaps $\phi(\Lambda = 30\text{mm})$ obtained from two different video frames. (c, e) Corresponding 3D models.

Figure 2.4b shows a wrapped single-shot phasemap captured at $\Lambda = 10$ mm. The same phasemap processed with the unwrapping procedure described above is displayed in Fig. 2.4d, and the respective 3D model is shown in Fig. 2.4f. It can be seen that the quality and detail richness of the measurement does not significantly decrease for the “single-shot" compared to the previously shown “double-shot" method.

For a quantitative depth precision evaluation of our camera and a comparison between single-shot and double-shot acquisition, we evaluated single-shot and double-shot measurements of a planar surface (cardboard) at different synthetic wavelengths. A small

Λ [mm]	40	10	5	3	1
δz_{double} [mm]	5.30	1.95	1.14	0.82	0.43
δz_{single} [mm]	5.56	1.78	1.64	0.79	0.33

Table 2.1: Depth precision analysis comparing the double-shot with the single-shot method.

region of the cardboard (roughly 23mm $\times$ 23mm) has been picked for evaluation. After subtracting a best-fit plane, the standard deviation of the (unfiltered) point cloud was calculated. This value represents the depth precision δz of the respective measurement. The results are displayed in Tab. 2.1: It can be seen that the single-shot measurements achieve roughly the same depth precision as their double-shot counterparts at the same synthetic wavelength. For $\Lambda \leq 3$ mm, our camera achieves sub-mm precision with a best-reported precision of only $\delta z_{single} = 330 \mu m$ for $\Lambda = 1$ mm. This value outperforms the precision of conventional ToF cameras roughly by a factor of **100x**.

To further demonstrate the single-shot capability of our novel Synthetic Wavelength Interferometry (SWI) ToF camera, we recorded a 3D video of a moving object. Each video frame consists of a single-shot measurement of the object at a synthetic wavelength of $\Lambda = 30$ mm. The object consists of a metronome whose pendulum moves left-to-right and then back right-to-left with multiple oscillations. To make the movement more obvious, two letters are attached to the metronome - one letter at the pendulum and one letter at the metronome body. During oscillation, the back letter moves laterally with respect to the front letter.

Figure 2.5 b and d show two synthetic phasemaps from two different time instances of the video. The respective 3D reconstructions are shown in Fig. 2.5 c and e. The full video of the captured phasemap and 3D reconstruction can be seen in link [74]. We emphasize again that such a 3D video sequence would not have been possible with earlier

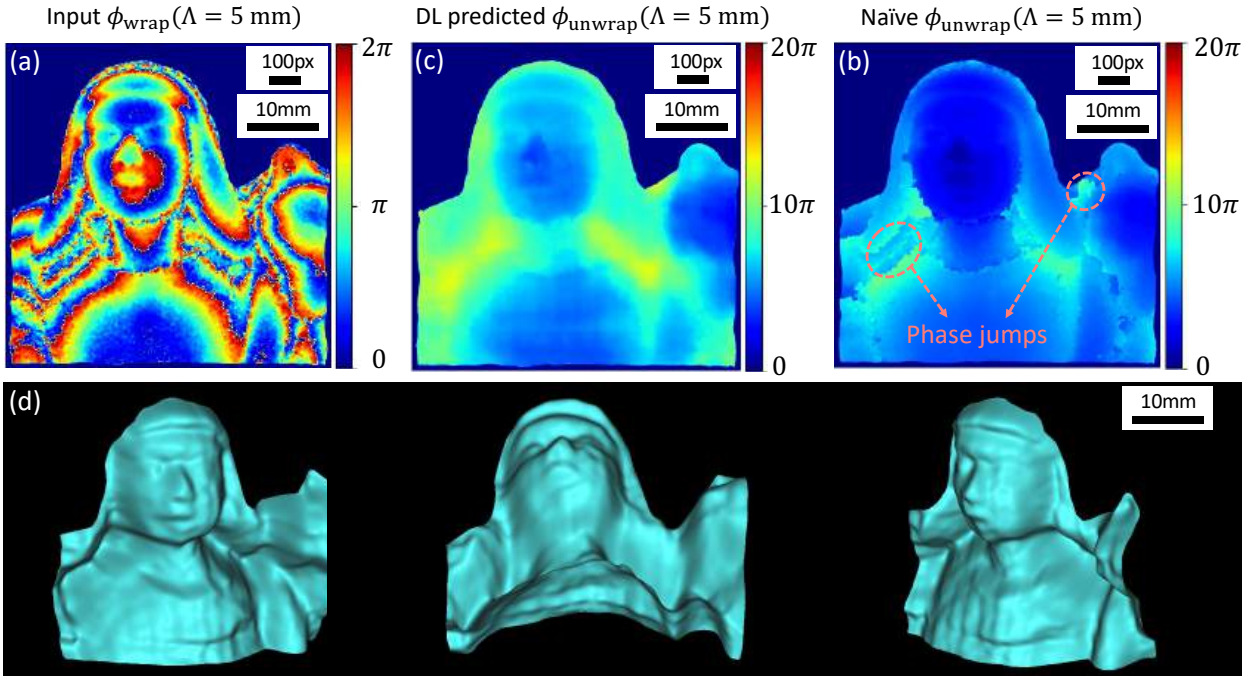


Figure 2.6: Single-shot phase unwrapping using Deep Learning (DL). (a) Acquired wrapped synthetic phasemap $\phi(\Lambda = 5\text{mm})$. (b) Unwrapped phasemap predicted by our DL algorithm, $\phi^{\text{unwrap}}(\Lambda = 5\text{mm})$. (c) Comparison: Phasemap obtained from another single-shot unwrapping procedure [72], [73] (not DL-based). Unwrapping failed as phase jumps are still present. (d) 3D model calculated from our DL-based unwrapped phasemap (b).

implementations of Synthetic Wavelength Interferometry for Computer Vision applications shown, e.g., in [43], [44], [46], [55].

In light of the encouraging single-shot results that have been demonstrated so far, one could plausibly argue that the "unwrapped single-shot" results shown in Fig. 2.4 are *not* genuinely single-shot, as we needed multiple sequentially captured single-shot phasemaps to apply the multifrequency unwrapping algorithm. And indeed, if we want to exploit the main feature of our novel camera (namely precise single-shot 3D acquisition of dynamic objects), multifrequency phase unwrapping from sequential images would not be possible[1]. For this reason, we developed a Deep Learning (DL) based algorithm to

[1]It should be emphasized, however, that the discussed motion artifact problem is now reduced to "con-

unwrap our captured (wrapped) single-shot phasemaps at smaller synthetic wavelengths.

Compared to other SOTA DL-based phase unwrapping procedures [51]–[54], our algorithm is specifically tailored to our acquired data structures and integrates seamlessly into our data processing pipeline. We defined a Convolutional Neural Network (CNN) architecture combined with a Long-Short-Term-Memory (LSTM) module [75]–[79]. The network is trained using both experimentally acquired and simulated wrapped phasemaps at different synthetic wavelengths as input and unwrapped phasemaps as output. The experimentally acquired unwrapped phasemaps for the training process were obtained using the multifrequency unwrapping algorithm described above. The supervised learning machine algorithm has a loss function that measures the mean squared error (MSE) between the real output image and the output that the CNN architecture generates from the input data. During the training process, the parameters of the CNN are optimized so that the model produces accurate unwrapped output phasemaps out of the wrapped input phasemaps. Once the training is complete, we apply the model to a newly measured wrapped phasemap to predict the unwrapped version. Neither the measured phasemap nor any other measurement of the evaluated object surface have been seen before by the algorithm.

Although not the main focus of this paper, we here report the first single-shot measurements unwrapped with our DL-based algorithm in Fig. 2.6. It can be seen that the algorithm correctly unwraps and reconstructs the overall shape of the object (see Fig. 2.6b). Fine details like eyes and nose are also reconstructed, although the reconstruction quality is slightly lower as for the measurements unwrapped with the multifrequency approach

ventional motion artifacts" as, e.g., observed in multi-shot structured light principles. This is because the multifrequency unwrapping with two single-shot synthetic phasemaps does not require the speckle patterns to be correlated between the different time instances (see the previous discussion).

(see Figs. 2.3 and 2.4). For comparison, we also attempted to unwrap the phasemap with an often-used single-shot unwrapping procedure (not DL-based) described in [72], [73]. The result is shown in Fig. 2.6c. It can be seen that this other method was not able to unwrap our synthetic phasemap correctly, and phase jumps are still present in the result.

2.3.4 Single-shot NLoS imaging through scattering scenes

So far, this paper was focused on "Line-of-Sight" (LoS) measurements, where the object is directly visible and can be directly imaged onto the sensor. However, the same technique can be used to image objects hidden from direct view, e.g., behind a scattering medium or around a corner. These so-called “Non-Line-of-Sight” (NLoS) measurements are potentially useful for many future applications in autonomous driving, medical diagnosis, or quality inspection of manufacturing products. For instance, we could think of an imaging system that could see through fog, that could see through human tissue to observe the inner organs, or that could look “around corners” to see some industrial machinery in operations.

Multi-shot NLoS measurements using Synthetic wavelength imaging principles have been demonstrated in [55]. However, it turns out that the single-shot ability of our camera is particularly crucial for NLoS measurements - even if the hidden object does not move! The reason is that most scattering media, such as fog, smoke, turbid water, or even human tissue, are in constant microscopic motion. For the double-shot method introduced in [55], this would lead to the already discussed decorrelation of speckle patterns which would result in a complete loss of information (please see related experiments and details in the Appendix A). In Fig. 2.7, we demonstrate a first single-shot NLoS measurement through a scattering scene. As shown in Fig. 2.7a, the object behind the scatterer is a simple point light source (fiber tip), which emits light at both used optical wavelengths

λ_1 and λ_2. The scatterer is a 220-grit ground glass diffuser. The synthetic phasemap $\phi(\Lambda)$ (see Fig 2.7b) is captured at the diffuser surface, and the related field is computationally backpropagated at the synthetic wavelength to reconstruct the point source in the hidden volume behind the diffuser. We refer to [55] for details about the used reconstruction method. An image of the reconstructed point light source is shown in Fig. 2.7c. The diameter of the bright spot in the middle roughly matches the theoretical expectations about the reconstruction resolution of the point light source. Further explanations are provided in the Appendix A.

Although very basic, this experiment documents (to the best of our knowledge) the first ever SWI-based single-shot full-field measurement through scattering media.

2.4 Summary, discussion, and outlook

We first summarize the advantages of our novel method with respect to other ToF-based imaging systems. While standard (single-wavelength) interferometric ToF cameras are limited to smooth object surfaces, our novel system can image objects with optically rough (e.g., fully lambertian) surfaces. Other existing multifrequency interferometric ToF cameras for Computer Vision applications were limited to static objects and/or used sophisticated sensors (e.g., PMDs [31], SPADs [47], or lock-in detectors [48]). In contrast, this work demonstrates a new setup design that enables high-precision measurements with conventional CCD/CMOS sensors. Our main contribution to the field is the introduction of our single-shot SWI method, which allows us to record videos of moving objects and reconstruct the dynamic 3D model at high precision.

In addition, we have demonstrated “unwrapped dual-shot measurements" and “unwrapped single-shot measurements" with a depth precision of up to $330\mu m$ - roughly 100x

better than the precision of conventional ToF cameras. Due to this immense performance difference, we focused our experiments around demonstrating the capabilities of *our* system and refrained from comparisons with conventional ToF camera systems (such as *Microsoft Kinect V2* [80]). We hope that it became obvious to the reader that such systems would not have been able to sufficiently resolve any of our small measured objects at all.

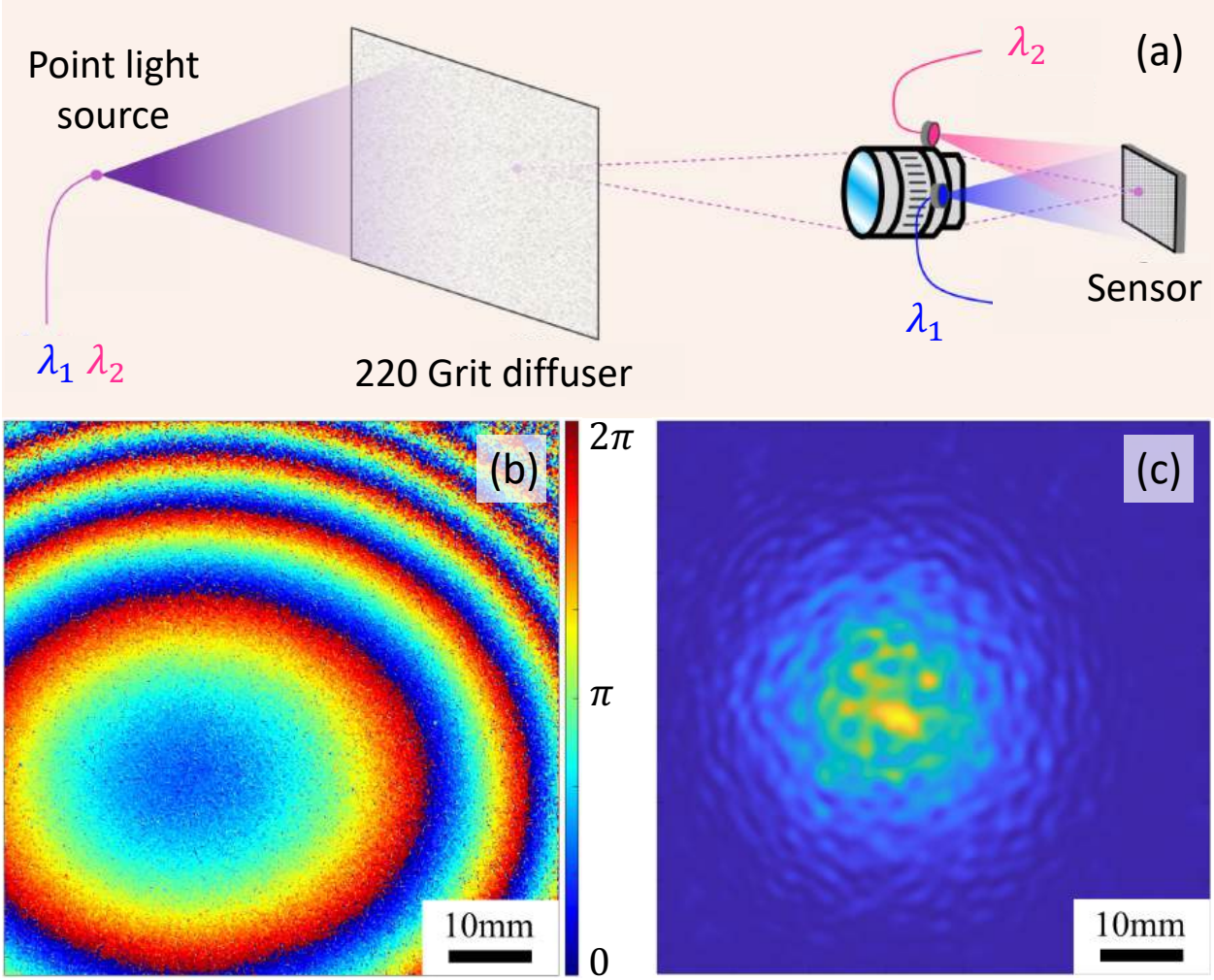


Figure 2.7: Full-field, single-shot Non-Light-of-Sight (NLoS) imaging. (a) Schematic of the experiment. (b) Synthetic phasemap $\phi(\Lambda = 1\text{mm})$ of the object's hologram at the diffuser location. (c) The amplitude of backpropagated object (the point light source).

We also want to mention that our single-shot method is not only limited to the simultaneous measurement of two optical wavelengths, i.e., one synthetic wavelength. Indeed,

we could theoretically operate three lasers simultaneously and illuminate the object with λ_1, λ_2, and λ_3. For instance, we could place each of the three reference beams with a 60-degree difference around the aperture. With this procedure, one could retrieve the object beams $E(\lambda_1)$, $E(\lambda_2)$, and $E(\lambda_3)$ from a single-shot image following our proposed method (see section 2.2). This new configuration enables the construction of three different synthetic fields: $E(\Lambda_{12})$, $E(\Lambda_{23})$, and $E(\Lambda_{13})$. Consequently, one can then use the "standard" multi-frequency unwrapping technique to retrieve a precise depth map from an unwrapped phasemap at a small synthetic wavelength *purely in single-shot*. Of course, the principle can be extended to even more optical wavelengths as long as the different regions in the Fourier domain stay separated. Further developments in this direction will be part of our future work.

Despite the promising first results shown in this paper, our imaging system is not without limitations. On the one hand, there is a fundamental limitation related to the aperture size of the optical system. A wider aperture size would allow higher spatial frequencies from the object beam to pass, resulting in higher lateral resolution during the imaging process but also increasing the size of the spectral regions in the Fourier domain. According to this tradeoff, we must carefully calibrate the aperture size so that the spectral regions in the Fourier domain do not overlap. Besides the tradeoff mentioned above, a small aperture also leads to decreased light throughput and hence a longer required exposure time. This issue becomes more severe when we record videos. We need a short exposure time to capture unblurred images of moving objects, which decreases the amount of light arriving at the sensor. The “optimal" aperture size is reached shortly before the spectral regions overlap in the Fourier domain.

Another limitation of our current system is its limited field of view (FoV). The back focal length of a camera is defined by the distance between the camera sensor and the closest lens from the optical system (see Fig. 2.2). As we need to place the reference beams pointing directly toward the camera sensor chip, we have to increase the back focal length of our camera. A short back focal length would imply a larger angle between the reference beams and the optical axis, which reduces the spacing between the fringes (see Fig. 2.1c). As we are already operating near the sampling limit (with about three pixels per fringe), the back focal length of our system cannot be shortened, and the total size of the setup becomes large. In the future, we will explore different solution approaches to this problem which can be found in optical metrology literature (e.g., in [61]).

2.5 Conclusions

We have developed a novel ToF single-shot imaging system based on Synthetic Wavelength Interferometry (SWI). Our contributions extend previous works on SWI-based ToF cameras, which were either limited to static objects or used sophisticated sensor architectures. We have shown the design of the new camera prototype and presented the computational procedure to obtain high-precision depth maps for several different introduced acquisition modes. We have tested our imaging system by measuring different static and dynamic objects and have acquired a “single-shot 3D video" of a moving object. Ultimately, we have exposed another potential new application of our technology by using our camera for a first-ever SWI-based single-shot NLoS measurement.

In the future, we hope that our introduced technique becomes part of a new wave of imaging devices that allow for completely new procedures and methods on how we perceive and interact with our environment. Beyond the current work, several promis-

ing directions have emerged in the broader context of synthetic wavelength techniques, including (a) light-in-flight imaging to temporally and spatially resolve light propagation through complex scenes [81], [82], (b) imaging through scattering media for nLoS applications [83], [84], and (c) high-resolution imaging via multi-mode fibers [85] (with potential applications in medical endoscopy).

Chapter 3

Spectrophotometry of thin solid films

This chapter presents a comprehensive framework for the optical modeling and characterization of thin solid films using transmission and reflection spectroscopy. The work is organized into three main sections: forward modeling, inverse retrieval based on optimization techniques, and deep learning based characterization. Together, these contributions form a robust pipeline for simulating and extracting material properties from spectrophotometric measurements.

The first section introduces a differentiable forward model that replicates the spectral transmittance and reflectance curves of thin film samples deposited on glass substrates. Given the geometric and optoelectronic properties of the film (including thickness, wedge angle, bandgap energy, Urbach energy, and dispersion parameters) the model simulates the experimental transmission and reflection spectra across a range of wavelengths. This forward simulator serves as a digital twin of the spectrophotometry setup, allowing for accurate reproduction of the light–matter interaction under realistic conditions. Our original contribution in this section lies in the derivation of new analytical expressions, based on the transfer matrix method, that extend traditional thin-film optics models to handle more complex and experimentally relevant scenarios. In particular, we

address two significant non-ideal conditions: (i) non-uniform film geometries with high wedge slopes, which can arise due to an intentional deposition technique or unintentional surface irregularities, and (ii) partially absorbing substrates, which introduce deviations from the assumption of transparent supports. These extensions improve the model's fidelity to real-world data and enable more accurate optical simulations.

The second section focuses on solving the inverse problem of optical characterization: determining the material properties of a thin film given its experimental transmission spectrum. We build upon the classical Swanepoel method, which uses envelope-based algebraic formulas to estimate refractive index and thickness from interference fringes. Recognizing that the accuracy of the method is highly sensitive to the correct identification of the transmission curve's upper and lower envelopes, we introduce a novel global optimization algorithm to robustly estimate these envelopes, thus significantly improving the reliability of the Swanepoel method in practical scenarios.

In the final section, we explore machine learning–based methods to further enhance the inverse modeling process. In subsection 3.1, we present a hybrid analytical and deep learning approach that improves the Swanepoel method with a convolutional neural network. Subsection 3.2 introduces a fully data-driven method in which a deep neural network directly infers the material parameters from transmission spectra. Finally, subsection 3.3 describes a physics-informed deep learning approach that incorporates the forward optical model as a regularizing constraint in the network training loss, ensuring physical consistency and improving generalization. Together, these techniques illustrate the transition from classical analytical methods to optimization based and data-driven methods for thin-film optical characterization.

3.1 Novel Formulas for Transmittance and Reflectance

Historically, spectroscopic techniques have been essential for studying the optical properties of thin solid films. However, existing formulae for both normal transmission and reflection spectroscopy often rely on simplified theoretical assumptions, which may not accurately align with real-world conditions. For instance, it is common to assume (1) that the thin solid layers are deposited on completely transparent thick substrates and (2) that the film surface forms a specular plane with a relatively small wedge angle. While recent studies have addressed these assumptions separately, this work presents an integrated framework that eliminates both assumptions simultaneously. In addition, the current work presents a deep review of various formulae from the literature, each with their corresponding levels of complexity. Our review analysis highlights a critical trade-off between computational complexity and expression accuracy, where the newly developed formulae offer enhanced accuracy at the expense of increased computational time. Our user-friendly code, which includes several classical transmittance and reflectance formulae from the literature and our newly proposed expressions, is publicly available in both Python and Matlab at this link.

3.1.1 Introduction

Thin solid films are essential in a wide range of modern industries, especially in the development of efficient transistors and photodiodes, as well as in the fabrication of protective metal and dielectric coatings. Thin film technologies directly impact the performance and efficiency of several everyday devices, including active matrix LCDs, photovoltaic systems, flash memory chips, photonic integrated circuits, and semiconductor batteries [13], [86]–[89]. It is important to remark that the specific conditions under which thin films are

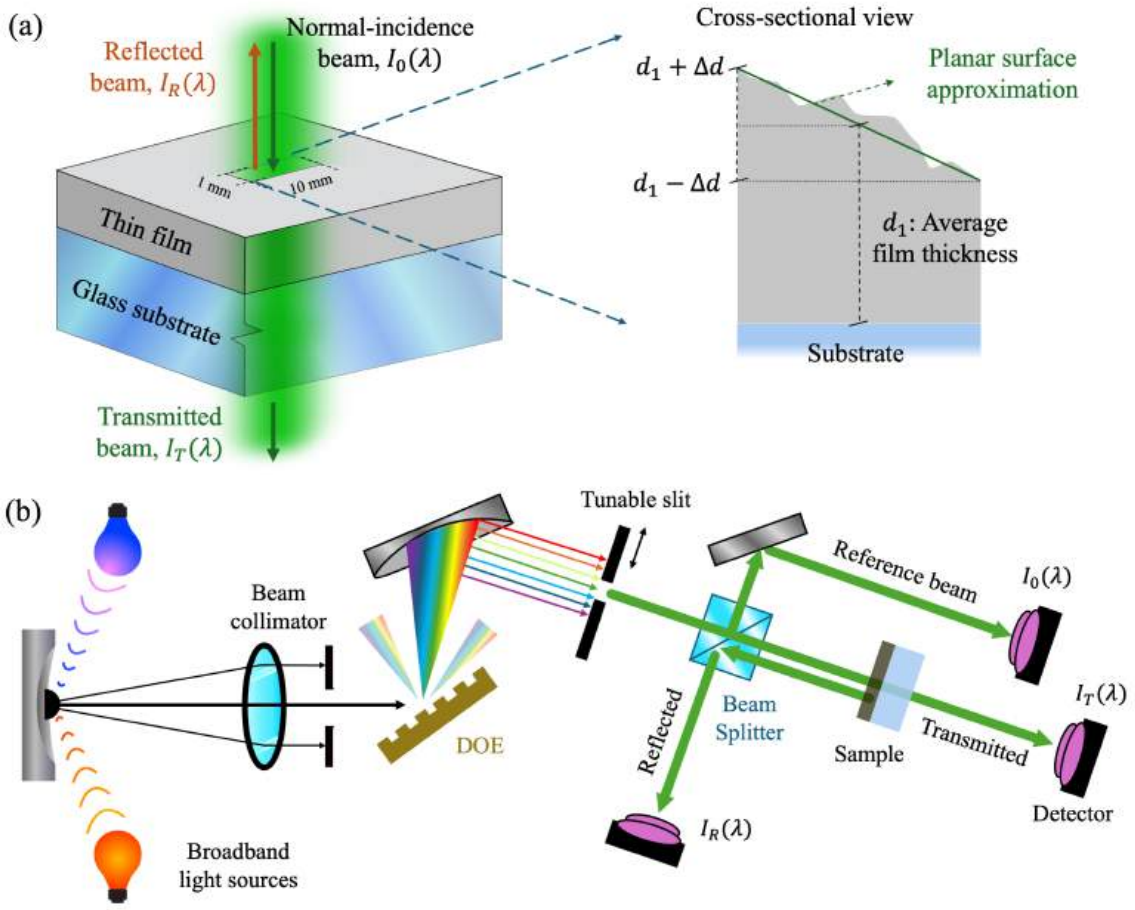


Figure 3.1: (a) Geometrical model of the sample under study. A thin solid film is deposited on top of a glass substrate. The inset shows a cross-sectional view of the film within the illuminated area, where the thin-film surface is approximated as a plane with a wedge parameter Δd. (b) Schematic diagram of the experimental spectrophotometric setup for transmission and reflection measurements. The setup includes two broadband light sources covering the UV and VIS-NIR spectral ranges, a diffraction optical element, a tunable slit that produces a quasi-monochromatic beam, and a beam splitter.

prepared have a significant impact on their final properties. Factors such as deposition techniques, preparation time, growth temperature, and working pressure all play critical roles in determining the characteristics of the resulting films [90]–[92]. Consequently, a precise analysis of their optoelectronic properties prior to mass production is crucial.

Thin films, with thicknesses often ranging from nanometers to several microme-

ters, are commonly deposited onto thick glass substrates (often on the millimeter scale), as seen in Fig. 3.1a. Our main objective is to determine the optical properties of the film, specifically the refractive index $n_1(\lambda)$ and the extinction coefficient $\kappa_1(\lambda)$, as functions of the wavelength λ, within a broad spectral range of interest, typically in the UV-Vis-NIR region. While there are multiple methods for optically characterizing these thin solid films, *spectroscopic ellipsometry* often emerges as the preferred choice due to its high accuracy. This technique measures how the polarization state of an incident beam changes upon reflection on the layer surface [93], [94]. Spectroscopic ellipsometry techniques provide comprehensive data collection through different reflection angles and polarization states, which allows for highly precise thin-film material characterization. It should be noted that this technique is particularly effective for analyzing complicated sample structures, including multilayered [95]–[97] or anisotropic materials [98], [99].

Despite the significant benefits of ellipsometry techniques, it is still very common to perform the optical characterization of simple thin films using measurements of normal transmittance and/or reflectance [100]. Several compelling factors contribute to the continued reliance on these latter measures. First, transmission and reflection spectrophotometers are indispensable instruments with applications in a wide range of fields, including biology and chemistry [101]. Second, the widespread availability of these instruments in research institutes and industries makes them a practical choice. Third, these instruments are generally more economical than spectroscopic ellipsometers, offering a cost-effective alternative for many laboratories [102]. In addition, many commercial instruments are designed to integrate both optical transmittance and reflectance measurements, facilitating efficient data acquisition [103]. Fourth, collecting a single set of light intensity data at normal incidence is generally easier, simpler, and faster than collecting a comprehensive ellipsometric dataset.

In the spectral ranges where the films exhibit medium-to-low absorption, most of the incident light is transmitted, enabling highly sensitive and accurate optical characterizations based on these transmission measurements. Conversely, for characterizations within the strong absorption regions of a film material, reflection measurements (such as those performed in reflectometry or ellipsometry) offer more advantages. It should be noted that ellipsometric measures are sensitive to surface roughness [104]–[107], as rough surfaces lead to a certain degree of depolarization upon reflection. While there are well-established methods to account for these effects, such as effective medium approximations (EMA) [108], [109], incorporating depolarization and surface roughness can add significant complexity to the ellipsometric analyses. In contrast, transmitted and reflected intensity measurements are generally less affected by surface roughness [110]. Moreover, one can still measure the total transmission and reflection in fully optically rough surfaces by using integrating sphere components [103] in the spectrophotometers, which allows to capture the scattered light due to diffuse reflectance and transmittance.

A representative schematic of a double-beam spectrophotometer can be found in Fig. 3.1b. A lamp emits broad-spectrum light. Typically, a deuterium arc lamp is used for the UV region, and a tungsten-halogen lamp for the Vis-NIR spectrum. This light is then dispersed by a diffraction optical element (DOE), as seen in the figure. Consequently, it passes through a slit, allowing for a controllable spectral bandwidth that usually ranges from 0.1 to 10 nm, which determines the coherent length of the beam. Various apertures are available to adjust the beam diameter, generally in the millimeter range. The normal-incidence reflected and transmitted light intensities, denoted as I_{T} and I_{R}, are then measured by a photodetector. Double-beam instrumentation splits the input beam into two parts: one to analyze the sample and the other to serve as a reference, measuring the baseline intensity I_0 with a separate sensor. Variations in the light source intensity are then

recorded simultaneously in both the reference and sample beams, allowing an accurate intensity ratio for transmission $T = I_{\mathrm{T}}/I_0$ and reflection $R = I_{\mathrm{R}}/I_0$. Then, N sequential transmittance and reflectance experimental measurements, denoted $\{T^{\exp}(\lambda_i), R^{\exp}(\lambda_i)\}$, are obtained at discrete specific wavelengths λ_i, $i \in \{1, 2, ..., N\}$, with a typical step size of a few nanometers. Our goal is then to develop a computational method that determines the two optical constants, $n_1(\lambda_i)$ and $\kappa_1(\lambda_i)$, from that experimental dataset.

Since the proliferation of advanced spectroscopic techniques, numerous analytical expressions have been formulated for the spectral transmission and reflection of thin-film samples, denoted $T^{\mathrm{theory}}(\lambda; n_1, \kappa_1)$ and $R^{\mathrm{theory}}(\lambda; n_1, \kappa_1)$. The formulae were then used to computationally fit the real-world optical measurements, finding the most reasonable values of (n_1, κ_1) that match the experiment. However, these theoretical expressions generally relied on rough assumptions or approximations. The approximations introduced were necessary in order to simplify the analysis and reduce computational complexity, particularly given the limited computing resources available at the time. Among the simplifications that still persist nowadays, two are particularly critical: (1) the glass substrate is presumed to be completely transparent throughout the spectral range of interest, and (2) the film surface is regarded as a specular plane with a minor wedge angle, usually only a few nanometers in height.

Current instruments are capable of detecting very subtle variations in light intensity measurements, revealing the existing weak substrate absorption, because of the presence of inherent glass absorption at specific spectral ranges or impurities in the substrate. Additionally, although it is true that most films are conveniently prepared to have quasi-uniform thicknesses, more 'exotic' thin-film samples may display complex surface geometries [111] or being deliberatively prepared with a certain wedge [112], [113]. Such

variations in thickness can lead to a pronounced wedge effect on the illuminated spot under analysis (see Fig. 3.1a). It should be highlighted that the modern literature has already addressed simplifications (1) and (2) separately, as explained in the next section in further detail. In contrast, the current work presents a unified and comprehensive theoretical model that addresses both problems *simultaneously*, bridging the gap in the current literature. Our code, written both in Python 3.0 [114] and Matlab 2021a [115], is available to the public at this link. Please note that this code also includes a collection of previously established formulae in this field, accounting for the results of T^{theory} and R^{theory} under different approximations.

3.1.2 Historical context and previous works

The study of the physics of thin films has profoundly influenced the fields of optics and photonics, particularly in the analysis of optical interference effects during the nineteenth century [36], [69]. Relevant advancements in this field include the development of the Fabry-Perot interferometer, the analysis of the etalon effect, and the formulation of Airy's equations. These contributions have facilitated advances in instruments for precise wavelength determination and enabled the development of accurate optical components. In addition, the examination of absorption bands in these thin layer materials significantly advanced atomic physics in the early twentieth century, improving our understanding of material properties at the quantum level. During the latter half of the twentieth century, there was a proliferation of accurate optoelectronic analyses of film semiconductor materials. The key advances that propelled these characterizations included *(i)* the introduction of *automated* spectrophotometers [116], [117], *(ii)* the integration of modern computer technologies and numerical methods, and *(iii)* the continuous development and refinement of film preparation techniques [118].

3.1.2.1 Computational methods to find optical properties

It is essential to review the historical context of well-established computational methods that perform optical characterization. We mentioned above a straightforward approach that consists of directly fitting the experimental transmittance (or reflectance) measurements with the derived theoretical formulae. By minimizing the least-squares error between the experimental data and our theoretical models, we find the optimal values for the optical functions, n_1 and κ_1. This approach is commonly known as "inverse synthesis" or "reverse engineering", and involves solving a complicated global optimization problem [119]–[125].

In contrast, Hall and Ferguson (1955) [126], Lyashenko and Miloslavskii (1964) [127], and Manifacier *et al.* (1975) [128] developed an alternative approach to find optical properties, initially known as the method of "successive iterations". This method does not operate for all the discrete measure wavelengths, but rather works for those particular wavelengths at which the minimum and maximum thin-film interferences occur. This algorithm requires the calculation, during the intermediate steps, of the lower and upper *envelopes* of the spectral transmission and/or reflection curves.

Ryno Swanepoel significantly refined this technique, introducing more precise formulae in two seminal works in 1983 [129] and 1984 [130], respectively for uniform and wedged films. Subsequently, the algorithm was universally renamed as the *Swanepoel method* to his honor. Originally confined to transmission measurements, this methodology was later extended to include reflectometry studies. This expansion began with the pioneering work of Kushev et al. in 1986 [131], and was further developed through a series of studies by Minkov et al. [132]–[136]. Since then, numerous significant works have further advanced this method for both transmission and reflection spectroscopy [137]–

[146].

A comparison between the *inverse synthesis* and *envelope method* [147], [148] indicates that, although inverse synthesis generally achieves greater precision, it comes with significantly higher computational costs. In addition to these two traditional methods, new research directions are emerging. For example, a novel deep learning (DL) technique has shown promise in accurately predicting optical properties of thin solid films from reflection [149] and transmission [6] spectra. Furthermore, a hybrid method has been proposed that combines the traditional Swanepoel method with DL enhancements [7]. In that work, a neural network automatically determines the envelopes of a transmission spectra, and then an automatic algorithm determines the optical constants using the information contained in the envelopes. Another approach recently proposed by our group is a numerical method that compares the experimental transmittance with the transmittance generated by a *numerical simulator* [150]–[152]. The simulator closely replicates wave propagation under the realistic conditions of a spectrophotometer, effectively acting as a digital twin of the actual setup.

3.1.2.2 Existing transmission and reflection formula

Revisiting now the early theoretical analysis of transmission and reflection formulae for thin solid films, it is essential to highlight the contributions of A.W. Crook in 1948 [153] and F. Abelès in the 1950s [154]–[156]. They independently laid the foundation for the transmission and reflection analysis in stratified media. The compilation work in the original Heavens book "Optical Properties of Thin Solid Films" from 1955 [157], [158] further consolidated the theoretical formulations employed in this field.

Despite these incipient accurate analyzes, it was common practice during that pe-

riod to simplify the theoretical expressions and assume that the film was deposited on an *infinitely thick* substrate [128], [159], [160]. Observe that the approximation overlooks the impact of back-reflection at the later substrate-air interface. The approximation significantly simplified the formulae to favor a straightforward optical characterization employing the envelope method [128]. In 1983 and 1984, the two aforementioned seminal works by Swanepoel [129], [130] formally challenged the notion of neglecting the actual thickness of the substrate. These studies demonstrated that such simplifications could result in errors in transmittance values of up to 3-4%. Subsequent studies highlighted the importance of accounting for substrate back-reflection in reflectance analyses [161].

In 1971, Potapov and Rakov proposed a pioneer algorithm to account for the effect of *slightly absorbing* thick substrates [162] on the transmittance and reflectance of samples with uniform films. Another relevant subsequent work on the topic was later reported by Vriens and Rippens in 1983 [163]. A later work by Swanepoel in 1989 [164] meticulously derived the comprehensive formulae for slightly absorbing substrates, explaining now in depth the effect of a quasi-coherent light source. It should be noted that these close-form friendly expressions and methodology proposed by Swanepoel have been adopted in the present work. In 1997, Kotlikov and Terechenko [165], [166] independently reached findings identical to those of Swanepoel, although in the context of antireflection coatings. Continuing the discussion on substrate absorption, it is important to highlight a study [167] by Nichelatti (2002), which introduces a very efficient numerical technique to directly characterize the substrate optical properties through spectroscopic measurements of that isolated substrate (explained in Section 9). In 2010, Barybin and Shapovalov [168] offered an alternative derivation of the transmission and reflection formulae for uniform films using the matrix formalism, specifically accounting for the effects of highly absorbing substrates.

The analysis of transmission and reflection formulae has continually evolved to accommodate increasingly complex sample characteristics. First, research has broadened from straightforward planar surfaces (tilted or not) to consider more complex, higher-order surface topographies [150], [169]–[172]. Second, considerable attention has been devoted to multilayer configurations [173], [174], where multiple thin films are stacked. Rather than closed-form analytical solutions, these scenarios often require numerical approaches based on the Transfer Matrix Method. These multi-film analyses enable the study of film coatings, surface oxidation, and other multilayer structures with applications in areas like color filters for display technologies [175], [176]. Third, while most spectrophotometric thin-film characterizations assume normal incidence, some studies have investigated quasi-normal [144], [150] and oblique angles [177], broadening the applicability of these techniques. Fourth, additional research has explored how light scattering from surface roughness, texture, and defects influences the measured transmission and reflection [178]–[180].

A new significant contribution came from Ruiz-Perez (RP) *et al.* [181] in 2020, adapting the transmission formula to account for films with a *pronounced wedge* deposited on transparent substrates. The RP formula accounts for 'exotic' films with a high-wedge condition (explained below in detail), which leads to significant changes in the shape of the spectrum, shrinking the characteristic oscillations of the transmission curves. This particular transmission expression has been successfully used in the recent literature and was named as the "universal transmission formula" [182], [183]. Built upon the RP approach, our current work presents both transmission and reflection formulae, which are now able to account not only for complex films with a high wedge angle but also for thick absorbing glass substrates. Table 3.1 summarizes the contributions of several relevant formulae from the literature, together with their associated approximations and their cor-

responding degree of precision.

Formula	Equation	Work	Coherence	Uniform	Approximations
T	Eq. 3.24	Established [184]	∞	Yes	Exact
R	Eq. 3.25	Established [184]	∞	Yes	Exact
T_ℓ	Eq. 3.29	Swanepoel 1989 [164]	$2d < L < 2d_\mathrm{s}$	Yes	Exact
R_ℓ	Eq. 3.30	Swanepoel 1989 [164]	$2d < L < 2d_\mathrm{s}$	Yes	Exact
$T_{\Delta d}^{\mathrm{new}}$	Eq. 3.37	Current	$2d < L < 2d_\mathrm{s}$	No	Exact
$R_{\Delta d}^{\mathrm{new}}$	Eq. 3.45	Current	$2d < L < 2d_\mathrm{s}$	No	Exact
$T_{\Delta d}^{\mathrm{RP20}}$	Eq. 3.55	Ruiz Perez 2020 [181]	$2d < L < 2d_\mathrm{s}$	No	$\kappa_2 = 0$
R^{Mink89}	Eq. 3.69	Minkov 1989 [132]	$2d < L < 2d_\mathrm{s}$	No	$\kappa_2 = 0$
T^{Swan83}	Eq. 3.74	Swanepoel 1983 [129]	$2d < L < 2d_\mathrm{s}$	Yes	$\kappa_2 = 0$ and $\kappa_1^2 \ll n_1^2$
$T_{\Delta d}^{\mathrm{Swan84}}$	Eq. 3.80	Swanepoel 1984 [130]	$2d < L < 2d_\mathrm{s}$	No	$\kappa_2 = 0$ and $\kappa_1^2 \ll n_1^2$
R^{RP01}	Eq. 3.90	Ruiz Perez 2001 [145]	$2d < L < 2d_\mathrm{s}$	Yes	$\kappa_2 = 0$ and $\kappa_1^2 \ll n_1^2$
$R_{\Delta d}^{\mathrm{RP01}}$	Eq. 3.99	Ruiz Perez 2001 [145]	$2d < L < 2d_\mathrm{s}$	No	$\kappa_2 = 0$ and $\kappa_1^2 \ll n_1^2$
$T_\mathrm{s}^{\mathrm{approx}}$	Eq. 3.99	Established [86]	$L < 2d_\mathrm{s}$	Yes	$\kappa_2 = 0$
$R_\mathrm{s}^{\mathrm{approx}}$	Eq. 3.99	Established [138]	$L < 2d_\mathrm{s}$	Yes	$\kappa_2 = 0$
T_s	Eqs. 3.103, 3.107	Nichelatti 2002 [167]	$L < 2d_\mathrm{s}$	Yes	Exact
R_s	Eqs. 3.104, 3.107	Nichelatti 2002 [167]	$L < 2d_\mathrm{s}$	Yes	Exact

Table 3.1: Set of transmittance and reflection expressions analyzed in the present work, each with their corresponding levels of accuracy and approximations.

3.1.3 Theoretical background

The present study assumes that the thin film is made of a homogeneous material with isotropic properties. We describe the complex dielectric function (the electric permittiv-

ity) across the stratified structure as

$$\boldsymbol{\epsilon}(z) = \begin{cases} \boldsymbol{\epsilon_0}, & z < 0 & (\text{Air}) \\ \boldsymbol{\epsilon_1}, & 0 < z < d_1 & (\text{Film}) \\ \boldsymbol{\epsilon_2}, & d_1 < z < d_1 + d_\text{s} & (\text{Substrate}) \\ \boldsymbol{\epsilon_0}, & d_1 + d_\text{s} < z & (\text{Air}) \end{cases} \tag{3.1}$$

The bold notation represents the complex nature of the functions. In this formula, d is the thickness of the film and d_s the thickness of the substrate. The initial sections 2-4 assume a film of uniform thickness $\Delta d = 0$. Later on, we explore films with a certain tilt, including those with the high wedge condition $\Delta d > \lambda/(4n)$, which requires a careful and distinct mathematical treatment, as discussed in detail in [145]. Although multiple surface geometries could be considered for the film, a plane surface with a potential wedge parameter Δd is reasonably assumed for the small illuminated area. This illuminated spot is often represented as a rectangular region with only a few millimeters of width and height . This planar assumption can be thought of as a Taylor first-order approximation of the actual much more intricate film surface (see Fig. 3.1a).

Maxwell equations express a relationship between the electric and magnetic fields [69] , denoted $\mathbf{E}$ and $\mathbf{H}$. Note that an external electric field $\mathbf{E}$ can influence an internal electric polarization $\mathbf{P}$ within a dielectric or semiconductor material, described as

$$\mathbf{P} = \epsilon_0 \boldsymbol{\chi}^{(1)} \mathbf{E} + \epsilon_0 \boldsymbol{\chi}^{(2)} \mathbf{E}^2 + \epsilon_0 \boldsymbol{\chi}^{(3)} \mathbf{E}^3 + ... \tag{3.2}$$

The electric flux density within the material is then given by $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$. For non-magnetic materials, the magnetic flux density is simply $\mathbf{B} = \mu_0 \mathbf{H}$, where μ_0 is the magnetic permeability in vacuum. In Eq. 3.2, the Taylor approximation around $\mathbf{E} = 0$ was

used to approximate an arbitrary polarization function, considering complex susceptibility coefficients $(\boldsymbol{\chi}^{(1)}, \boldsymbol{\chi}^{(2)}, \boldsymbol{\chi}^{(3)}, ...)$. In practice, most materials exhibit a linear response, although second- and third-order effects have also been studied in depth [185] since the invention of laser in 1960, especially in the context of high intensity light beams. Our research focuses exclusively on linear materials, where $\mathbf{D} = \epsilon_0(1 + \boldsymbol{\chi}^{(1)})\mathbf{E} = \epsilon_0\boldsymbol{\epsilon_r}\mathbf{E} = \boldsymbol{\epsilon}\mathbf{E}$. In these particular cases, the complex relative permittivity $\boldsymbol{\epsilon_r}$ fully describes the optoelectric properties of the material.

In linear non-magnetic isotropic film materials, the complex refractive index is defined as $\mathbf{n}(z) = \sqrt{\boldsymbol{\epsilon_r}} = n(z) + i\kappa(z)$. Here, n denotes the real part of the refractive index, proportional to the phase velocity of light through the medium. The extinction (or attenuation) coefficient κ is closely related to the absorption of the medium. The complex dielectric function reveals key information about the electronic transitions of the material. In turn, this allows us to deduce several critical aspects relevant to the film industry. For instance, it provides estimations of the band gap energy [182], the material conductivity [186], the dissipation factor [187], the compound stoichiometry [14], and the level of structural disorder [86]. Understanding structural disorder is particularly important for the analysis of the physical properties of amorphous materials and the identification of defects in crystalline structures.

3.1.4 Light propagation through a sample

Without loss of generality, let us consider a simple linearly-polarized light beam oriented in the x-direction, expressed as $\mathbf{E} = (E_x, 0, 0)$, perpendicular to the film's incidence plane (see Fig. 3.2). Considering that this planar light wave propagates in the $z-$direction, the propagation vector becomes $\mathbf{k} = (0, 0, 2\pi/\lambda) = n(0, 0, 2\pi/\lambda_0) = nk_0\hat{z}$, where $\lambda = \lambda_0/n$ is the wavelength in the material, λ_0 the wavelength in vacuum, and k_0 the wavenumber.

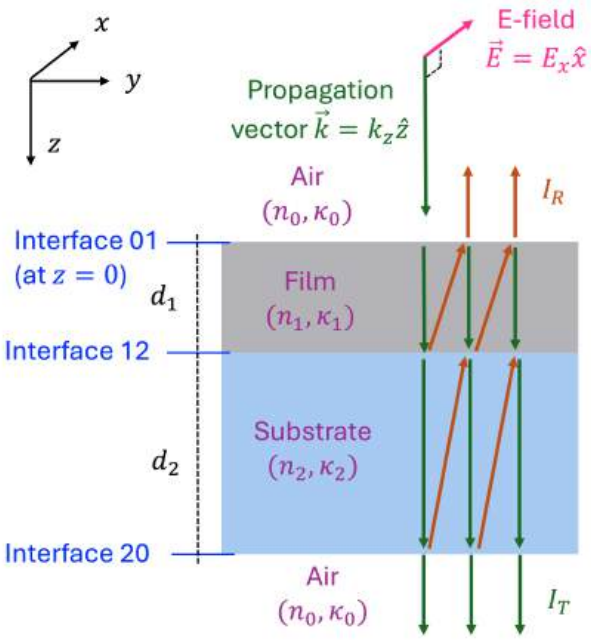


Figure 3.2: Scheme of the stratified media. Multiple interferences between incident and reflected waves occur because of the different interfaces.

Isotropic linear materials ensure the orthogonality between the propagation vector and the electric and magnetic fields [36], thus $\mathbf{H} = (0, H_y, 0)$.

If no external current is applied to the material $(J = 0)$ and no free charges are induced in the bulk of the material $(\rho = 0)$, the Maxwell equations simplify as follows [36]:

$$\partial_x E_x = 0 \tag{3.3}$$

$$\partial_y H_y = 0 \tag{3.4}$$

$$\partial_z E_x = \mu_0 \partial_t H_y \tag{3.5}$$

$$-\partial_z H_y = \tilde{\epsilon} \partial_t E_x \tag{3.6}$$

3.1.4.1 Propagation matrix

We will now focus only on propagation through layer 1 (the thin film). When the incident beam is a monochromatic continuous wave, the temporal dependency of the light wave becomes $e^{-i\omega t}$ [36], leading to an electric field of the form $E_x(x, y, z, t) = A(x, y, z)e^{-i\omega t}$. Here, $\omega = 2\pi c/\lambda_0$ represents the angular frequency and c the speed of light in vacuum. Incorporating this field expression into Eqs. 3.3-3.6 and assuming constant material properties $\boldsymbol{\epsilon_1}$ and μ_0 within the bulk of the layer, one can derive the so-called Helmholtz equation:

$$(\partial_{zz} + \boldsymbol{n}_1 k_0)A(x, y, z) = 0 \tag{3.7}$$

A well-known solution [36] for the amplitude $A(x, y, z)$ from Eq. 3.7 is a combination of two planar waves propagating in the $z-$ direction,

$$\begin{aligned} A(x, y, z) = A_{\mathrm{T}}(z) + A_{\mathrm{R}}(z) = \\ A_0 \mathbf{t}_{02} e^{ik_0 \mathbf{n_1} z} + A_0 \mathbf{r}_{12} e^{-ik_0 \mathbf{n_1} z} \end{aligned} \tag{3.8}$$

Eq. 3.8 shows two distinct components, corresponding to a transmission wave (moving forward) and a reflection wave (moving backward) [184]. Note that $\mathbf{t}_{01}$ and $\mathbf{r}_{12}$ represent the complex amplitude Fresnel coefficient for transmission and reflection, respectively. The transmission occurs between layer 0 and layer 1, while the reflection takes place from layer 2 toward layer 1. At normal incidence, these coefficients are defined as

$$\mathbf{t}_{01} = \frac{2\mathbf{n}_0}{\mathbf{n}_0 + \mathbf{n}_1} = \mathbf{r}_{01} + 1 \tag{3.9}$$

$$\mathbf{r}_{12} = \frac{\mathbf{n}_1 - \mathbf{n}_2}{\mathbf{n}_1 + \mathbf{n}_2} = \mathbf{t}_{12} - 1 \tag{3.10}$$

Considering the amplitudes $A_{\mathrm{T}}(z)$ and $A_{\mathrm{R}}(z)$ immediately after the air-film interface (see Fig. 3.2), our goal is to determine the transmitted and reflected wave amplitudes just before they encounter the next interface: the film-substrate boundary at $z+d_1$. By applying Eq. 3.8 at $z+d_1$, we derive the following expressions using matrix formalism [188]:

$$\begin{pmatrix} A_{\mathrm{T}}(z+d_1) \\ A_{\mathrm{R}}(z+d_1) \end{pmatrix} = P_1 \begin{pmatrix} A_{\mathrm{T}}(z) \\ A_{\mathrm{R}}(z) \end{pmatrix} \tag{3.11}$$

$$P_1 = \begin{pmatrix} e^{ik_0\mathbf{n}_1 d_1} & 0 \\ 0 & e^{-ik_0\mathbf{n}_1 d_1} \end{pmatrix} \tag{3.12}$$

The transfer matrix P_1 represents the *propagation matrix* for layer 1. Note that the complex refractive index (located in the exponents of Eq. 3.11) accounts for both phase delay and absorption through that homogeneous layer. Indeed, we can derive from Eq. 3.11 that a collimated light beam passing through the film during a single forward trip becomes

$$A_{\mathrm{T}}(z+d_1) = A_T(z)\exp\left(-\frac{\alpha_1}{2}d_1\right)\exp\left(-\frac{i\delta_1}{2}\right) \tag{3.13}$$

Where

$$\alpha_1 = \tfrac{4\pi}{\lambda_0}\kappa_1 \tag{3.14}$$

$$\phi_1 = \tfrac{2\pi}{\lambda_0}n_1 d \tag{3.15}$$

$$\delta_1 = 2\phi_1 \tag{3.16}$$

Here, ϕ_1 refers to the phase delay introduced during the propagation of the beam through the film. It is more convenient to work instead with the variable δ_1, which repre-

sents the phase delay during a whole round trip. Another relevant wavelength-dependent optical parameter is α_1, known as the absorption coefficient.

Considering the well-known relation [69] between light intensity and wave amplitude, $I \propto |A|^2$, one can see from Eq. 3.13 that the transmitted intensity decreases exponentially with depth as $I_{\mathrm{T}}(z+d_1) = I_{\mathrm{T}}(z)e^{-\alpha_1 d_1}$, following the Beer-Lambert law [69]. Note that $x_1 = I_{\mathrm{T}}(z+d_1)/I_{\mathrm{T}}(z) = e^{-\alpha_1 d}$ represents the transmittance corresponding to a single light trip. However, it should be noted that the multiple beam interferences within the two stratified media will lead to a different overall sample transmittance.

3.1.4.2 Dynamic matrix

Let us now consider the case where the beam reaches the interface between two layers, such as the thin film (layer 1) and the thick substrate (layer 2).

Because the incident beam is normal to the sample surface, the electric and magnetic fields have only tangential components, represented as (E_{1x}, H_{1y}) for the film layer

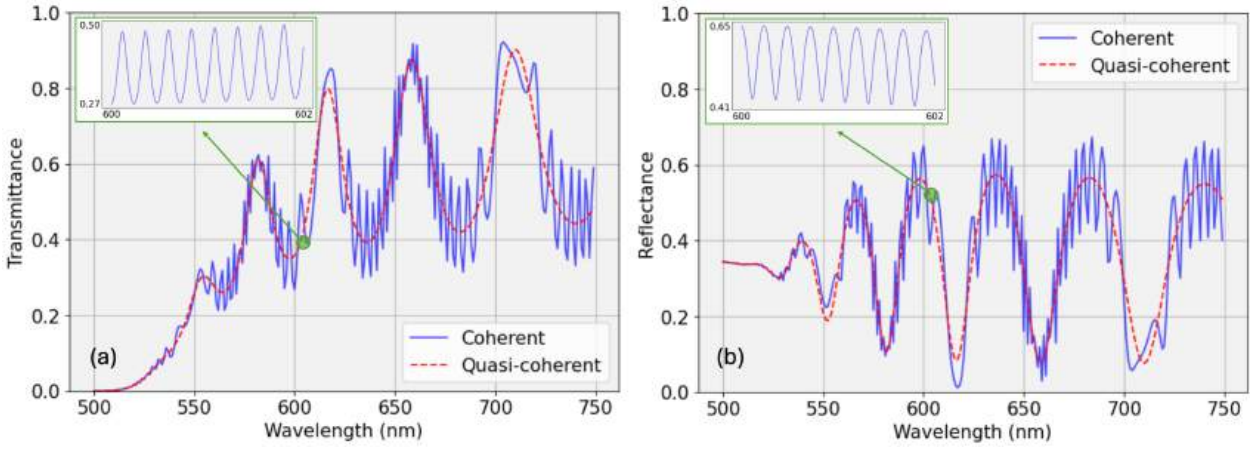


Figure 3.3: Transmittance (a) and reflectance (b) spectra of our simulated sample (for an amorphous silicon thin film), assuming a weakly absorbing glass substrate. The results are shown for a monochromatic light source with infinite coherence (Eqs. 3.24 and 3.25) and for another light source with limited coherence (Eqs. 3.35 and 3.36), where $2d_1 < \ell < 2d_2$.

and (E_{2x}, H_{2y}) for the substrate layer. The tangent fields must remain continuous across the interface, meaning that $E_{1x} = E_{2x}$ and $H_{1y} = H_{2y}$ at $z = d$. Imposing these boundary conditions in Eqs. 3.3-3.6 gives us the following relations [184]:

$$A_{\mathrm{T}}^{(1)} + A_{\mathrm{R}}^{(1)} = A_{\mathrm{T}}^{(2)} + A_{\mathrm{R}}^{(2)} \tag{3.17}$$

$$\mathbf{n_1}(A_{\mathrm{T}}^{(1)} - A_{\mathrm{R}}^{(1)}) = \mathbf{n_2}(A_{\mathrm{T}}^{(2)} - A_{\mathrm{R}}^{(2)}) \tag{3.18}$$

Here, $A_{\mathrm{T}}^{(i)}$ and $A_{\mathrm{R}}^{(i)}$ represent the amplitudes of the transmitted and reflected light waves within the layer $i \in \{1, 2\}$, right next to the film-substrate interface. Eqs. 3.17-3.18, when transformed to matrix form, yield

$$\begin{pmatrix} 1 & 1 \\ \boldsymbol{n_1} & -\boldsymbol{n_1} \end{pmatrix} \begin{pmatrix} A_{\mathrm{T}}^{(1)} \\ A_{\mathrm{R}}^{(1)} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \boldsymbol{n_2} & -\boldsymbol{n_2} \end{pmatrix} \begin{pmatrix} A_{\mathrm{T}}^{(2)} \\ A_{\mathrm{R}}^{(2)} \end{pmatrix} \tag{3.19}$$

After a few calculations, one can finally find

$$\begin{pmatrix} A_{\mathrm{T}}^{(2)} \\ A_{\mathrm{R}}^{(2)} \end{pmatrix} = D_{12} \begin{pmatrix} A_{\mathrm{T}}^{(1)} \\ A_{\mathrm{R}}^{(1)} \end{pmatrix} \tag{3.20}$$

$$D_{12} = \frac{1}{1 + \mathbf{r_{12}}} \begin{pmatrix} 1 & -\mathbf{r_{12}} \\ -\mathbf{r_{12}} & 1 \end{pmatrix} \tag{3.21}$$

The transfer matrix D_{12} is called the *dynamic matrix* and accounts for amplitude changes between the film-to-substrate (interface 12).

3.1.4.3 Resulting transmittance and reflectance

By sequentially computing the wave propagation through the layers and the interactions at the interfaces, the transmitted and reflected fields through the whole sample can be determined as:

$$\begin{pmatrix} A_{\mathrm{T}}^{(3)} \\ A_{\mathrm{R}}^{(3)} \end{pmatrix} = M \begin{pmatrix} A_{\mathrm{T}}^{(0)} \\ A_{\mathrm{R}}^{(0)} \end{pmatrix} \tag{3.22}$$

Where the transfer matrix is now given by

$$M = D_{23} \times P_2 \times D_{12} \times P_1 \times D_{01} \tag{3.23}$$

Note that there is no light absorption throughout layers 0 and 3 (as it corresponds to air). Therefore, the propagation in these layers only leads to a constant phase delay, which will not affect the overall transmitted or reflected light intensity. The total transmittance and reflection can be calculated from the elements of the transfer matrix M, as explained in [184]. We then obtain the following formulae:

$$T = \left|\frac{1}{M^{(1,1)}}\right|^2 = \frac{\mathbf{h}}{\boldsymbol{\tau}\bar{\boldsymbol{\tau}}} = \frac{h}{a+2b\cos\delta_2+2c\sin\delta_2} \tag{3.24}$$

$$R = \left|\frac{M^{(1,2)}}{M^{(1,1)}}\right|^2 = \frac{\boldsymbol{\eta}\bar{\boldsymbol{\eta}}}{\boldsymbol{\tau}\bar{\boldsymbol{\tau}}} = \frac{e+2f\cos\delta_2+2g\sin\delta_2}{a+2b\cos\delta_2+2c\sin\delta_2} \tag{3.25}$$

Where the complex coefficients are defined as follows:

$$\begin{aligned} \mathbf{h} &= x_1x_2(\mathbf{r}_{01}+1)(\bar{\mathbf{r}}_{01}+1)(\mathbf{r}_{12}+1)(\bar{\mathbf{r}}_{12}+1)(\mathbf{r}_{23}+1)(\bar{\mathbf{r}}_{23}+1) \\ \boldsymbol{\tau} &= \mathbf{r}_{01}(\mathbf{r}_{12}+\mathbf{r}_{23}x_2e^{i\delta_2})x_1e^{i\delta_1}+(\mathbf{r}_{12}\mathbf{r}_{23}x_2e^{i\delta_2}+1) \\ \boldsymbol{\eta} &= -\mathbf{r}_{01}(\mathbf{r}_{12}\mathbf{r}_{23}x_2e^{i\delta_2}+1)-(\mathbf{r}_{12}+\mathbf{r}_{23}x_2e^{i\delta_2})x_1e^{i\delta_1} \end{aligned} \tag{3.26}$$

Note that the overall sample absorption can then be computed as $\mathcal{A} = 1 - (T + R)$ [186]. The overbar on bold symbols, such as $\bar{\boldsymbol{\tau}}_2$, indicates the conjugate of complex numbers. Note that the sinusoidal functions from Eqs. 3.24-3.25 depend on δ_2 and account for the beam interferences within the substrate. Since the sample's transmittance and reflectance are *scalar* wavelength-dependent functions, Eqs. 3.24-3.25 can be equivalently written either in terms of complex coefficients $(\mathbf{h}, \boldsymbol{\tau}, \boldsymbol{\eta})$ or either in terms of real coefficients (h, a, b, c, e, f, g), which are all defined in Appendix B.1. Although Eqs. 3.24-3.25 becomes more extensive when working with the real coefficients, it is much more convenient for computational purposes. It should also be clarified that the coefficients (a, b, c), which appear in the denominator of Eqs. 3.24-3.25, contain sinusoidal functions that depend on the phase δ_1. These sinusoidal functions account for the Fabry-Perot light beam interferences within the thin film.

Figure 3.3 (see curves in blue) shows the resulting transmission and reflection within the spectral range of 500 to 750 nm. In this work, we consider a simulated thin-layer sample that has a slightly absorbing glass substrate (see Appendix B.2 for further details), with fix values $n_2 = 1.5$, $\kappa_2 = 10^{-6}$, and $d_2 = 0.5$ mm. The optical properties of an amorphous silicon thin film, with thickness $d_1 = 1$ μm, were simulated employing the empirical dispersion expressions $n_1 = 2.6 + 3 \cdot 10^5/\lambda^2$, $\log_{10}(\alpha_1) = -8 + 1.5 \cdot 10^6/\lambda^2$ [129]. Note that the extinction coefficient κ_1 can be directly determined from the absorption coefficient α_1 by using Eq. 3.14.

3.1.5 Exact formulae for uniform films

So far, all the calculations have assumed a purely monochromatic light source, which introduces some unwanted high-frequency noise in the transmission and reflection spectra (see Figs. 3.3). In the actual experiment, the light sources have a finite spectral band-

width, which inherently limits the coherence length of the beam. By carefully selecting the bandwidth, we can avoid coherent interference within the substrate, effectively eliminating the associated noise.

For a standard light beam with Gaussian spectral shape, the coherence length ℓ can be determined from the specific central wavelength λ_c and the bandwidth $\Delta\lambda$ (measured at full width half maximum). The established formula for this relationship is $\ell = 4\ln(2)/\pi\lambda_c^2/\Delta\lambda$ [189]. The pure monochromatic light source ($\Delta\lambda \approx 0$) considered in Eqs. 3.24-3.25 inherently led to an infinite coherence length. In this case, the interreflections both within the thin film and within the thick substrate resulted in the superposition of the electromagnetic waves, causing two distinct interference effects.

According to the etalon interference condition [69], constructive interference occurs at wavelengths where $m = 2nd/\lambda$ is an integer, leading to high values in the transmission spectra. Note that when m is a half-integer, there is destructive interference, and we observe valleys in the spectra. According to this interference equation, the relatively small thickness of the film $d \approx 1\,\mu$m leads to low-frequency oscillations. These long oscillations appear as a characteristic sinusoidal pattern in the transmission spectrum, as illustrated in Figs. 3.3a and 3.3b (see dashed red lines). In contrast, the considerable thickness of the substrate, $d_s \approx 1$ mm, gives rise to high-frequency oscillations that can be interpreted as noise.

The curves presented in Figs. 3.3a and 3.3b were plotted at each discrete wavelengths, 500, 501, 502, ..., and 750 nm. However, this spectral resolution is insufficient to clearly resolve the high-frequency oscillations originated by light wave interferences at the substrate, as illustrated in the insets of the figures (top left). Indeed, the insets show that there are eight oscillation peaks in a range of two nanometers. This limitation in sam-

pling leads to the emergence of *aliasing effects*. In spectral regions where the oscillation frequency is roughly a multiple of the *sampling rate,* the curve appears to have fewer oscillations (e.g. 700 to 725 nm). Conversely, regions with greater misalignment exhibit rapid noisy oscillations (e.g. 725 to 750 nm).

Consequently, it is standard practice to deliberately adjust the bandwidth such that the coherence length ℓ is less than $2\,d_{\mathrm{s}}$, twice the thickness of the substrate. This adjustment mitigates the unwanted etalon effect from the glass, and it effectively removes the noise in the spectra, as the beam will become incoherent after a single round trip. Recall that the coherence length defines the distance over which the electromagnetic waves maintain their sinusoidal nature [69]. When two coherent light beams with electric fields E_1 and E_2 overlap, they interfere linearly, resulting in a combined electric field, $E = E_1 + E_2$ [69]. The resulting light intensity is given by $I = |E_1 + E_2|^2 = I_1 + I_2 + 2E_1E_2\cos(\Delta\theta)$, where $\Delta\theta$ is the phase difference between the two light beams. The interference term $2E_1E_2\cos(\Delta\theta)$ oscillates rapidly with changes in wavelength. In contrast, for incoherent beams, the electric fields do not interfere, and the total intensity is simply $I = I_1 + I_2$, with no oscillatory interference term.

In many real experiments, the coherence length is set within $2\,d < \ell < 2\,d_{\mathrm{s}}$. The multiple thin-film interreflections then result in the linear superposition of the electric fields (see Figs. 3.2). However, interreflections within the substrate lead to an incoherent superposition, producing just a simple sum of their intensities. To incorporate this complicated finite coherence model into our expressions from Eqs. 3.24-3.25, the *spectral averaging* method [184] proposes to average out the transmittance and reflectance over the corresponding phase delay introduced by the substrate, $\delta_2 = 4\pi n_2 d_2/\lambda$. For instance, consider an incident beam with a reasonable spectral bandwidth of 1 nm, center around

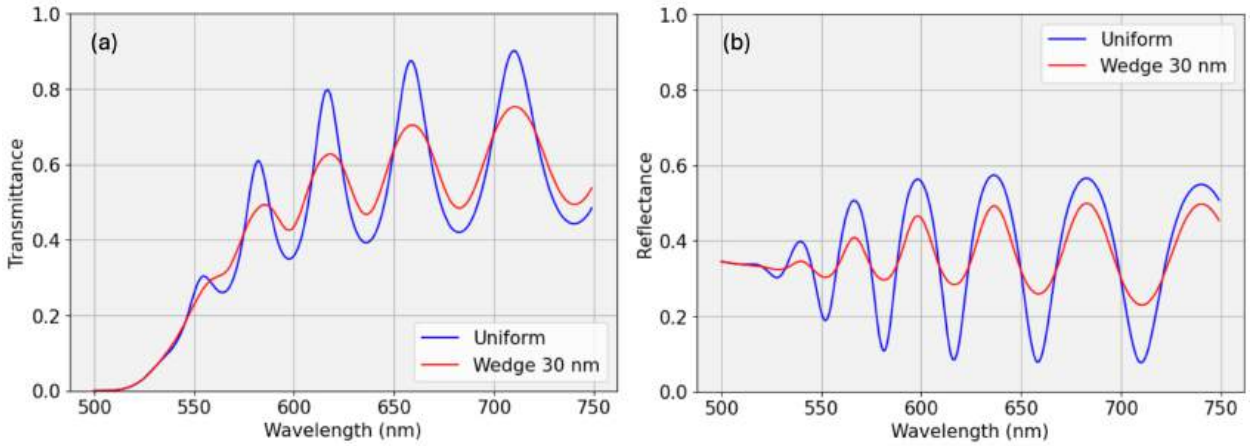


Figure 3.4: Sample (a) transmittance and (b) reflectance, using the exact formulae for uniform films (Eqs. 3.29-3.30) and for wedge films (Eqs. 3.37-3.45), with $\Delta d = 30$ nm. When the wedge increases, the contrast in the interferences decreases, as does the amplitude of the oscillations.

the wavelength $\lambda_0 = 600$ nm. The spectrum of the light beam spans from $\lambda_0^- = 599.5$ nm to $\lambda_0^+ = 600.5$ nm, which leads to a phase delay change of $\Delta\delta_2 = |\delta_2^- - \delta_2^+| = 8\pi$. The superposition of infinite waves with a phase delay range exceeding 2π will effectively average out (see Appendix B.3 for further details). Therefore, we can simply integrate the transmittance and reflectance from Eqs. 3.24-3.25 within the substrate phase delay range

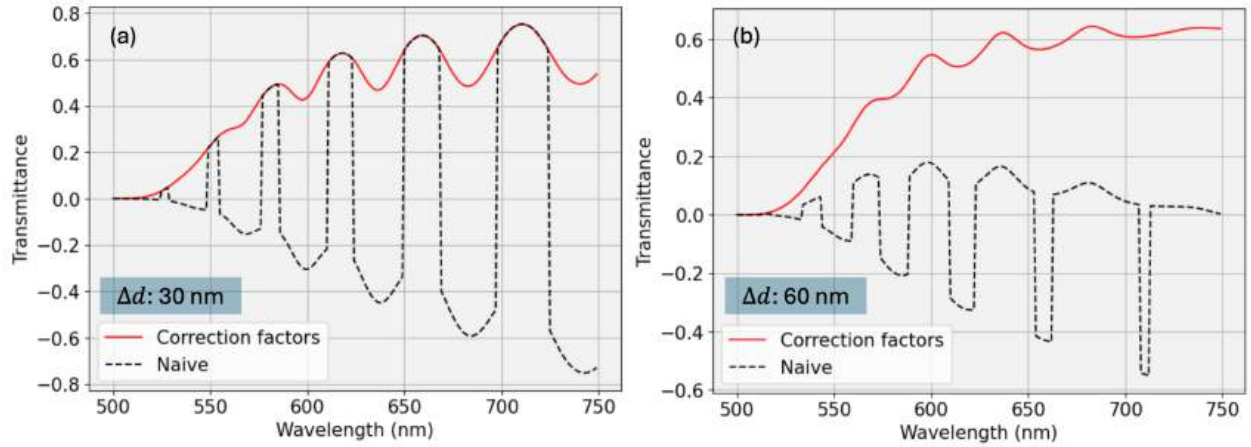


Figure 3.5: Simulated transmittance showing the effect of the correction factors N^+ and N^- (see Eqs. 3.41) for thin films with two representative wedge parameters, (a) $\Delta d = 30$ nm and (b) $\Delta d = 60$ nm.

$[0, 2\pi]$, yielding

$$T_\ell = \frac{1}{2\pi}\int_0^{2\pi} \frac{h\ \mathrm{d}\delta_2}{a+2b\cos\delta_2+2c\sin\delta_2} \tag{3.27}$$

$$R_\ell = \frac{1}{2\pi}\int_0^{2\pi} \frac{(e+2f\cos\delta_2+2g\sin\delta_2)\ \mathrm{d}\delta_2}{a+2b\cos\delta_2+2c\sin\delta_2} \tag{3.28}$$

Taking into account a change of variable, $\theta = e^{i\delta_2}$, we can conveniently calculate the phase-average reflectance and transmittance as

$$T_\ell = \frac{1}{2\pi}\oint_{|\theta|=1} \frac{h\ \mathrm{d}\theta}{F(\theta)} \tag{3.29}$$

$$R_\ell = \frac{1}{2\pi}\oint_{|\theta|=1} \frac{G(z)\ \mathrm{d}\theta}{F(\theta)} \tag{3.30}$$

$$F(\theta) = \theta^2(c+ib) + \theta(ia) + (-c+ib) \tag{3.31}$$

$$G(\theta) = \theta^2(g+if) + \theta(ie) + (-g+if) \tag{3.32}$$

The denominator can now be factor as $F(\theta) = (c+ib)(\theta-\theta_1)(\theta-\theta_2)$, where

$$\theta_{1,2} = \frac{-ai \pm \sqrt{-a^2+4(b^2+c^2)}}{2(c+ib)}, \tag{3.33}$$

$$|\theta_{1,2}| = \left|-\gamma \pm \sqrt{1-\gamma^2}\right|,\ \ \gamma = \frac{a}{2\sqrt{b^2+c^2}} \in [0,1] \tag{3.34}$$

Note that the discriminant in Eq. 3.33 is negative, as $a^2 \gg b^2, c^2$, making the numerator a pure imaginary number. This allows for a straightforward calculation of $|\theta_{1,2}|$ in Eq. 3.34. It then becomes evident that θ_1 (with a positive square root) lies within the unit disk of the complex plane, while θ_2 lies beyond the disk. Applying the Cauchy Residue Theorem [190] to Eq. 3.29, we finally obtain the expression for the transmittance,

$$T_\ell = h\,i\,\mathcal{R}\left(\frac{1}{F(z)}, z_1\right) = \frac{h\,i}{(c+ib)}\ \lim_{z\to z_2}\frac{(z-z_2)}{(z-z_1)(z-z_2)} = \frac{h}{\sqrt{a^2-4(b^2+c^2)}} = \frac{h}{u} = \frac{h}{a'+2b'\cos\delta_1+2c'\sin\delta_1} \tag{3.35}$$

The variables (u, a', b', c') are defined in Appendix B.1 . The coefficients (a', b', c') do not have any sinusoidal components. Therefore, the oscillations in transmittance T_ℓ come solely from the phase delay δ_1 (see the sine and cosine in the denominator of Eq. 3.35). These oscillations are now only introduced by the thin film and not the substrate.

Employing an absolutely similar methodology, the analysis of reflectance results in the following formula:

$$R_\ell = \frac{e}{u} - \frac{2(bf + cg)}{uw} \tag{3.36}$$

The expressions (w, e, b, f, c, f) are also defined in the Appendix B.1. Fig. 3.3 shows the transmission and reflection for a quasi-coherent beam. It can now be seen in that figure that the high-frequency noisy oscillations caused by light beam interference within the substrate have been correctly removed both in the transmission and reflection spectra.

3.1.6 Novel formulae for highly-wedged films

We now examine a film characterized by a wedge-shaped planar surface, denoted as $\Delta d > 0$. This wedge implies that the film thickness at the illuminated spot varies within the range $[d_1 - \Delta d, d_1 + \Delta d]$. In that case, the phase delay introduced by the film exhibits a variation from $\delta_1^- = 4\pi n_1(d_1 - \Delta d)/\lambda$ to $\delta_1^+ = 4\pi n_1(d_1 + \Delta d)/\lambda$, depending on the particular region within the illuminated spot. Considering a reasonable wedge $\Delta d \ll d_1$, the change in phase delay is typically $\Delta\delta_1 = |\delta_1^+ - \delta_1^-| \ll 2\pi$. As a result, while the interference effects from the thin film are not completely eliminated, they can be significantly reduced depending on the value of Δd. A very large wedge (such as $\Delta d \approx d$) will eventually cancel out all interference effects.

Additionally, the film's absorption will vary depending on the light path, whether it passes through $d_1 - \Delta d_1$, or $d_1 + \Delta d_1$, or somewhere in between (see Fig. 3.6). Therefore,

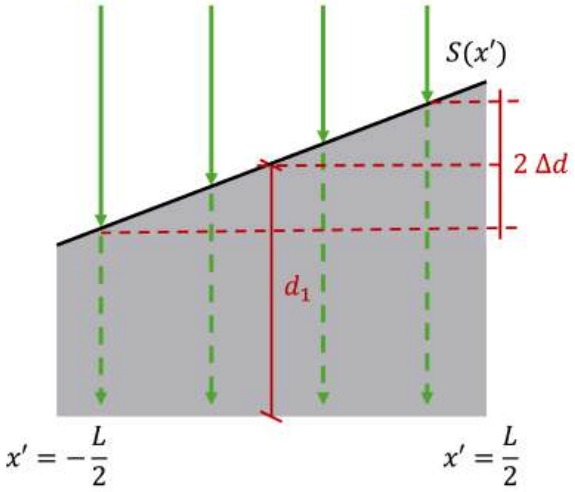


Figure 3.6: Profile of a tilted film with a linear surface described by $S(x') = mx'$. The optical path length in the $z-$direction varies depending on the position along the $x'-$axis.

the single-trip transmittance x_1 should now be integrated from $x_1^- = e^{-\alpha_1(d_1-\Delta d)}$ to $x_1^+ = e^{-\alpha_1(d_1+\Delta d)}$. With the assistance of the symbolic software Wolfram Alpha Mathematica [191], we were able to calculate the total transmittance of the sample,

$$T_{\Delta d}^{\text{new}} = \frac{1}{(\delta_1^+-\delta_1^-)(x_1^+-x_1^-)} \int_{x_1^-}^{x_1^+} \int_{\delta_1^-}^{\delta_1^+} \frac{h \, \mathrm{d}\delta_1 \, \mathrm{d}x_1}{a'+2b'\cos\delta_1+2c'\sin\delta_1}$$

$$\approx \frac{1}{\delta_1^+-\delta_1^-} \int_{\delta_1^-}^{\delta_1^+} \frac{h \, \mathrm{d}\delta_1}{a'+2b'\cos\delta_1+2c'\sin\delta_1} = \tag{3.37}$$

$$\frac{2\cdot h}{K\cdot(\delta_1^+-\delta_1^-)} \left[\arctan\left(\frac{I^+}{K}\right) + N^+\pi - \arctan\left(\frac{I^-}{K}\right) - N^-\pi\right]$$

$$K = \sqrt{a'^2 - 4\cdot(b'^2+c'^2)} \tag{3.38}$$

$$I^- = (a'-2b')\tan\left(\frac{\delta_1^-}{2}\right) + 2c' \tag{3.39}$$

$$I^+ = (a'-2b')\tan\left(\frac{\delta_1^+}{2}\right) + 2c' \tag{3.40}$$

$$N^+ = \text{round}\left(\frac{\delta_1^+}{2\pi}\right), \quad N^- = \text{round}\left(\frac{\delta_1^-}{2\pi}\right) \tag{3.41}$$

Note that in the intermediate steps, we applied the identity $\tanh^{-1}(i\,q) = i\tan^{-1}(q)$, valid for any value $q \in \mathbb{R}$.

The average light absorption x_1 in a tilted film is approximately equivalent to that observed in a uniformly thick film, which is the reason for the approximation in Eq. 3.37. Although this assumption has been commonly used in various contexts [181], it has not been previously supported by a rigorous mathematical proof. We will now explore the conditions under which this statement holds true. We introduce a local coordinate system (x', y', z) such that the linear wedge is expressed solely in terms of the coordinate x'. The film profile can then be described as $S(x') = mx'$, where $m = 2\Delta d/L$ represents the thin film slope. According to the Lambert-Beer law, the light intensity after the collimated beam passes through the film depth is given by $I(x') = I_0 e^{-\alpha_1(d_1+mx')}$. By integrating over the $x'-$range $[-L/2, L/2]$, we obtain

$$\langle I_{\text{out}}(x')\rangle = \tfrac{1}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} I_0 e^{-\alpha_1[d_1+mx']}\,dx' \tag{3.42}$$

$$= \tfrac{I_0}{L} e^{-\alpha_1 d_1}\left(\tfrac{-1}{\alpha_1 m} e^{-\alpha_1 m x'}\Big|_{-\frac{L}{2}}^{\frac{L}{2}}\right) \tag{3.43}$$

$$= I_0 e^{-\alpha_1 d_1}\tfrac{\sinh(\alpha_1 mL/2)}{\alpha_1 mL/2} \tag{3.44}$$

We must take into account that $q = \alpha_1 mL/2 = \alpha_1\Delta d$. One should notice that $\sinh(q)/q \approx 1$ for $q \approx 0$, which justifies the aforementioned assumption for relatively small wedges, where q remains close to zero. To quantify this, even in the case of a relatively large wedge $\Delta d = 60$ nm and high absorption $\alpha_1 = 10^4$ cm^{-1}, the wavelength-dependent term $\sinh(q)/q$ has an average value of 1.00060. Therefore, we can say that $I \approx I_0 e^{-\alpha_1 d_1}$ with an average error of around 0.06%, which is negligible in practice, comparable to the typical photometric noise of the spectrophotometer. In these common scenarios, the single-pass light transmittance x_1 in a tilted film can be considered approximately equivalent to that observed in a uniform film. It should be noted that the film presents high absorption at 500 nm, as seen in Fig. 3.3, where the transmittance eventually approaches

zero. At that wavelength, the term $\sinh(y)/y$ finds the highest value, 1.015, and the approximation becomes less accurate.

Figs. 3.4a and 3.4b present the transmittance and reflectance simulated curves, respectively, for films with different wedges, $\Delta d = 0, 30, 60$ nm. We can see how higher wedges lead to less interference contrast, reducing the amplitude of the oscillations in the spectra.

Eqs. 3.41 introduce the correction factors N^+ and N^-, which account for the different branches of the tangent function and facilitate the application of the formula to high wedges, where $\Delta d > \lambda/(4n)$. These factors were previously motivated in [181] in the context of sample transmittance for non-absorbing substrates. To clearly see the importance of the correction factors in the transmittance formula (see Eq. 3.41) when dealing with relatively high Δd parameters, one can also see in Figs. 3.5a and 3.5b the plot of the spectra with and without the correction factors for the two representative cases $\Delta d = 30, 60$ nm. As the wedge increases, the naive approach (without correction factors) disagrees more with respect to the expected curves.

To the best of our knowledge, there is no equivalent *closed-form* expression for the reflectance. Consequently, we propose using the trapezoidal rule [192] for efficient and accurate numerical estimation:

$$\begin{aligned} R_{\Delta d}^{\text{new}} = \frac{1}{\delta_1^+ - \delta_1^-} \int_{\delta_1^+}^{\delta_1^+} \left(\frac{e}{u} - \frac{2(bf + cg)}{uw} \right) d\delta_1 = \\ \frac{1}{2} \Bigg[R_L \left(\delta_1^- + 2\Delta d \frac{0}{N} \right) + 2R_L \left(\delta_1^- + 2\Delta d \frac{1}{N} \right) + \dots \\ \dots + 2R_L \left(\delta_1^- + 2\Delta d \frac{N-1}{N} \right) + R_L \left(\delta_1^- + 2\Delta d \frac{N}{N} \right) \Bigg] \end{aligned} \tag{3.45}$$

κ_2	Δd (nm)	Transmission (RP) Eqs. 3.55 RMSE (%)	Reflection (Minkov) Eqs. 3.69 RMSE (%)	Transmission (Swanepoel) Eqs. 3.74-3.80 RMSE (%)	Reflection (Swanepoel) Eqs. 3.90-3.99 RMSE (%)
0	0	0	10^{-4}	0.035	0.068
0	10^{-5}	0	-	0.035	0.068
0	30	0	-	0.012	0.067
0	60	0	-	0.007	0.067
10^{-6}	0	0.496	0.034	0.497	0.074
10^{-6}	10^{-5}	0.496	-	0.497	0.074
10^{-6}	30	0.487	-	0.487	0.073
10^{-6}	60	0.488	-	0.488	0.073

Table 3.2: Errors associated with the approximated formulae. When $\Delta d = 0$ the formulae for uniform films are used, tested against the exact formulae from Eqs. 3.29 and 3.30. For $\Delta d = 10^{-5}$ nm, the formulae for wedged films are applied, demonstrating consistency as $\Delta d \to 0$. They were tested against the new formulae from Eqs. 3.37 and 3.45. Note that the two left blocks assume $\kappa_2 = 0$, while the two right blocks use the Swanepoel approximations ($\kappa_2 = 0$ and $n_1^2 \ll k_1^2$).

Where $R_L(\delta_1)$ is the reflectance for a uniform film as a function of the phase δ_1, as given by Eq. 3.36. Our findings indicate that setting $N = 40$ points for the phase already yields reasonable results, since we obtain a total absorptance value $\mathcal{A}$ effectively zero when *no* film or substrate absorption is added ($\kappa_1 = \kappa_2 = 0$), $\mathcal{A} = |1 - T_{\Delta d} - R_{\Delta d}| < 10^{-2}$. In any case, note that we can define a sufficiently fine sampling that meets any desired level of accuracy.

3.1.7 Particular case: Non-absorbing substrate

3.1.7.1 Transmittance

The newly derived transmittance formula, Eq. 3.37, closely matches that found by Ruiz-Perez (RP) *et al.* in [181]. However, our expression now incorporates the possibility of substrate absorption. Indeed, the RP equation (without substrate absorption) can be alternatively derived now[1] by simply setting $\kappa_2 = 0$ in our new Eq. 3.37. Following the

[1]We found an errata in the RP formula presented in [181]. In particular, their Eq. 6 incorrectly omits a minus sign before the term "$C_{22}x\sin(\phi)$". This error can be confirmed by comparing it with Eq. A1 in [129]. Consequently, Eqs. 14 in [181], corresponding to the formulae for the coefficients "K_1" and "K_2", should

notation from RP, we then get the following simplified formulae

$$A = 16(n_1^2 + k_1^2)n_2, \tag{3.46}$$

$$B = ((n_1 + 1)^2 + k_1^2)((n_1 + 1)(n_1 + n_2^2) + k_1^2), \tag{3.47}$$

$$C_1 = 2\left((n_1^2 + k_1^2 - 1)(n_1^2 + k_1^2 - n_2^2) - 2k_1^2(n_2^2 + 1)\right), \tag{3.48}$$

$$C_2 = 2k_1\left(2(n_1^2 + k_1^2 - n_2^2) + (n_1^2 + k_1^2 - 1)(n_2^2 + 1)\right), \tag{3.49}$$

$$D = ((n_1 - 1)(n_1 - n_2^2) + k_1^2)((n_1 - 1)^2 + k_1^2), \tag{3.50}$$

$$T^{\mathrm{RP20}} = \frac{A\,x_1}{B - C_1\,x_1\,\cos(\delta_1) + C_2\,x_1\,\sin(\delta_1) + D\,x_1^2}, \tag{3.51}$$

$$F = (B + Dx_1^2 + C_1x_1)\tan\left(\frac{\delta_1^-}{2}\right) + C_2x_1, \tag{3.52}$$

$$G = (B + Dx_1^2 + C_1x_1)\tan\left(\frac{\delta_1^+}{2}\right) + C_2x_1, \tag{3.53}$$

$$H = B^2 - x_1^2(C_1^2 + C_2^2 - 2BD - D^2x_1^2), \tag{3.54}$$

$$T_{\Delta d}^{\mathrm{RP20}} = \frac{2Ax_1}{(\delta_1^+ - \delta_1^-)\sqrt{H}}\left[\arctan\left(\frac{G}{\sqrt{H}}\right) + N^+\pi - \arctan\left(\frac{F}{\sqrt{H}}\right) - N^-\pi\right] \tag{3.55}$$

Note that T^{RP20} from Eq. 3.51 corresponds to the transmittance of the sample with a uniform film thickness, without taking into account substrate absorption. Similarly, $T_{\Delta d}^{\mathrm{RP20}}$ from Eq. 3.55 represents the case in which the sample has a certain wedge $\Delta d > 0$. We have confirmed numerically that, when no substrate absorption is considered (i.e., $\kappa_2 = 0$), the root mean square error (RMSE) is zero for any wedge parameter between *(i)* our new exact formula from Eq. 3.37, and *(ii)* the approximate expression from Eq. 3.55. When $\kappa_2 = 10^{-6}$ and $\Delta d = 30$ nm, the RMSE of the approximated formulae by RP goes up to 0.465%. This finding clearly shows the importance of incorporating substrate absorption into the models. More results for different wedges are summarized in Table 3.2. One can see that, as the wedge increases, the error decreases. This happens because

also include a minus sign before "$C_{22}x$".

the smaller the oscillation amplitude, the smoother the curve, getting lower peak values. Similarly, we have checked for the case of uniform films ($\Delta d = 0$) that the exact formula Eq. 3.29 and Eq. 3.51 produce exactly the same numerical results when no substrate absorption is considered.

3.1.7.2 Reflectance

Based on the foundational work of Grebenstikov *et al.* [193], Minkov *et al.* [132] derived in 1989 an accurate expression for reflectance (which excludes substrate absorption). While Minkov's equation is only applies to films with uniform thickness, an analytical expression for wedge-shaped films can be simply derived if we further assume $\kappa_1 \ll n_1$ (as will be explaine in detail in the next section).

One can observe that Minkov's formula, Eq. 3.69, is *not* completely equivalent to the exact expression derived from Eq. 3.36 when substrate absorption is disregarded ($\kappa_2 = 0$). However, the numerical discrepancy between them is minimal. For instance, we found a value of discrepancy error as low as 10^{-4}% when using our simulated sample. The discrepancy arises because, in addition to the assumption that $\kappa_2 = 0$, second-order approximations involving n_1 and κ_1 were also utilized in Minkov's expression.

However, when considering the weakly absorbing substrate from our simulated sample, the RMSE of Minkov's formula rises up to 0.034%. It should be highlighted that the reflection formula, Eq. 3.69, is considerably less affected by the substrate absorption than the transmission formula from Eq. 3.55 (see Table 3.2): As mentioned earlier, the overall reflectance and transmittance can be expressed as the addition of infinite light wave components, which appear due to the inter-reflections at each interface (see Fig. 3.2). When analyzing the overall reflectance, the two *main* components come from the first

air-film interface and the second film-glass interface. The third component from the last substrate-glass interface (thus affected by the substrate absorption) is certainly weaker. In contrast, all transmission components account for the beam passing through the whole substrate volume. Therefore, transmittance measurements are notably affected by the existing substrate absorption.

The main advantage of Minkov's reflectance formula, and the reason of its popularity [194], [195], is that it provides a simpler and more accessible expression than simply setting $\kappa_2 = 0$ in our new complicated formula, Eq. 3.36. Additionally, Minkov's formula facilitates the analysis of reflection for the envelope method, as explored in the next section.

Minkov's formula is shown below:

$$A' = ((n_1 - 1)^2 + k_1^2)\left((n_1 + n_2)^2 + k_1^2\right) \tag{3.56}$$

$$B_1' = 2\left((n_1^2 + k_1^2 - 1)(n_1^2 + k_1^2 - n_2^2) + 4k_1^2 n_2\right) \tag{3.57}$$

$$B_2' = 4k_1\left(n_2(n_1^2 + k_1^2 - 1) - (n_1^2 + k_1^2 - n_2^2)\right) \tag{3.58}$$

$$C' = ((n_1 + 1)^2 + k_1^2)\left((n_1 - n_2)^2 + k_1^2\right) \tag{3.59}$$

$$A'' = ((n_1 + 1)^2 + k_1^2)\left((n_1 + n_2)^2 + k_1^2\right) \tag{3.60}$$

$$B_1'' = 2\left((n_1^2 + k_1^2 - 1)(n_1^2 + k_1^2 - n_2^2) - 4k_1^2 n_2\right) \tag{3.61}$$

$$B_2'' = 4k_1\left(n_2(n_1^2 + k_1^2 - 1) + (n_1^2 + k_1^2 - n_2^2)\right) \tag{3.62}$$

$$C'' = ((n_1 - 1)^2 + k_1^2)\left((n_1 - n_2)^2 + k_1^2\right) \tag{3.63}$$

$$G' = 64 n_2 (n_2 - 1)^2 (n_1^2 + k_1^2)^2 \tag{3.64}$$

$$D'' = ((n_1 + 1)^2 + k_1^2)\left((n_1 + 1)(n_1 + n_2^2) + k_1^2\right) \tag{3.65}$$

$$E_1'' = 2\left((n_1^2 + k_1^2 - 1)(n_1^2 + k_1^2 - n_2^2) - 2k_1^2(n_2^2 + 1)\right) \tag{3.66}$$

$$E_2'' = 2k_1\left((n_1^2 + k_1^2 - n_2^2) + (n_2^2 + 1)(n_1^2 + k_1^2 - 1)\right) \tag{3.67}$$

$$F'' = ((n_1 - 1)^2 + k_1^2)\left((n_1 - 1)(n_1 - n_2^2) + k_1^2\right) \tag{3.68}$$

$$R^{\text{Mink89}} = \frac{A' - (B_1' \cos(\delta_1) - B_2' \sin(\delta_1))x_1 + C' x_1^2}{A'' - (B_1'' \cos(\delta_1) - B_2'' \sin(\delta_1))x_1 + C'' x_1^2} + \frac{G' x_1^2}{A'' - (B_1'' \cos(\delta_1) - B_2'' \sin(\delta_1))x_1 + C'' x_1^2} \times \frac{1}{D'' - (E_1'' \cos(\delta_1) - E_2'' \sin(\delta_1))x_1 + F'' x_1^2} \tag{3.69}$$

3.1.8 The Swanepoel Approximations

To ensure the completeness of this work, and to facilitate comparison with our results, we will now examine other previously established formulae for T and R that have been widely employed in the literature. As mentioned before, Swanepoel refined an algebraic

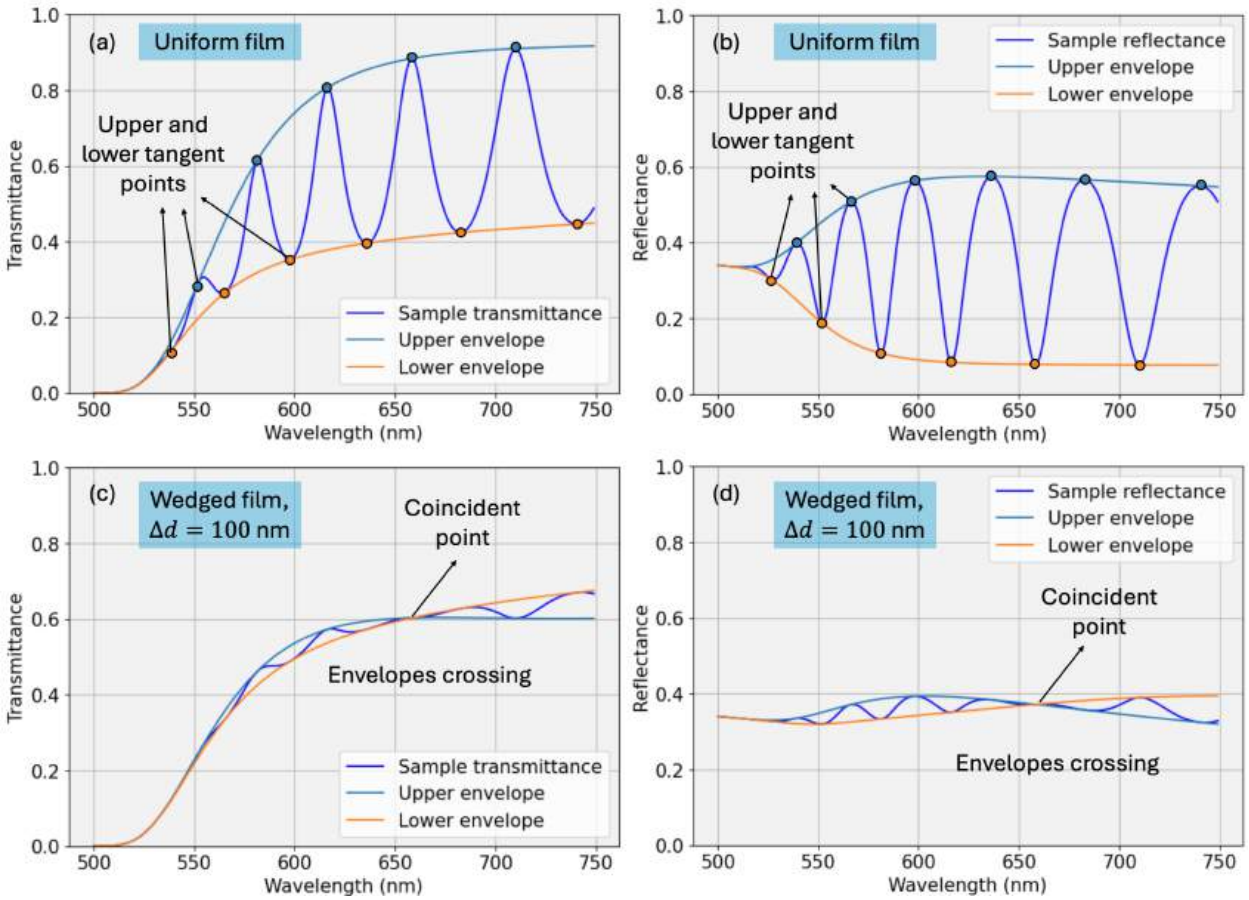


Figure 3.7: Sample transmittance (left column) and reflectance (right column) using the Swanepoel approximations. The upper and lower envelopes, as well as the tangent points where the spectrum intersects the envelopes, are shown. (a) and (b) represent uniform films (see Eqs. 3.74 and 3.90), while (c) and (d) correspond to films with a very high wedge, $\Delta d = 100$ nm (see Eqs. 3.80 and 3.99). Under high wedge condition, we see how the envelopes *cross over*, reaching a particular wavelength in which the spectrum and the two envelopes coincide.

procedure, commonly known as the envelope method, to directly determine (n_1, κ_1, d). However, this method finds the optical properties only at specific wavelengths λ_i corresponding to the maxima or minima Fabry-Perot interferences. In particular, these critical points are situated near the peaks and valleys of the transmission spectra (see Figs. 5a and 5b). In practice, they are identified by the intersection of the experimental spectra with their respective envelopes (reason why these intersection points are often described as the “tangent points”).

This envelope method operates under two assumptions: *(i)* the substrate does not absorb ($\kappa_2 = 0$), and (2) the extinction coefficient of the film is much weaker than the refractive index $\kappa_1^2 \ll n_1^2$. These two conditions are known as the Swanepoel approximations. Note that we analyzed condition *(i)* only in Section 7. Condition *(ii)* is typically valid for spectral regions with medium-to-weak absorption, where the transmittance alone indeed contains sufficient information for characterization. This second assumption, however, does not hold in areas of strong absorption, particularly near the optical band gap, where the transmission T clearly decreases, eventually approaching and reaching zero.

3.1.8.1 Transmission

By setting $\kappa_2 = 0$ in Eq. 3.35, the transmission through a sample with a uniform-thickness film simplifies [129] to the following expression:

$$A_0 = 16 n_1^2 n_2 \tag{3.70}$$

$$B_0 = (n_1 + 1)^2 (n_1 + 1)(n_1 + n_2^2) \tag{3.71}$$

$$C_0 = 2(n_1^2 - 1)(n_1^2 - n_2^2) \tag{3.72}$$

$$D_0 = (n_1 - 1)^3 (n_1 - n_2^2) \tag{3.73}$$

$$T^{\text{Swan83}} = \frac{A_0 x_1}{B_0 - C_0 x_1 \cos(\delta_1) + D_0 x_1^2} \tag{3.74}$$

$$T_{\text{M}}^{\text{Swan83}} = \frac{A_0 x_1}{B_0 - C_0 x_1 + D_0 x_1^2} \tag{3.75}$$

$$T_{\text{m}}^{\text{Swan83}} = \frac{A_0 x_1}{B_0 + C_0 x_1 + D_0 x_1^2} \tag{3.76}$$

It must be pointed out that the upper and lower envelopes of the spectrum from Eqs. 3.75 and 3.76 are found by fixing the sinusoidal component of the transmittance in Eq. 3.74 to the maximum or minimum value, that is, $\cos(\delta_1) = \pm 1$.

The set of expressions from Eq. 3.74 were first derived in the Swanepoel seminal

paper from 1983 [129]. When no substrate absorption is considered ($\kappa_2 = 0$), the RMSE between our exact Eq. 3.37 and the approximated Eq. 3.74 (which assumes $\kappa_1^2 \ll n_1^2$) is 0.039% for our simulated thin-film sample. When substrate absorption is considered, the error increases to 0.472%.

For non-uniform thin films with a certain wedge parameter $\Delta d > 0$, the expressions for the transmission and envelopes are rewritten as follows:

$$F_0 = \tfrac{A_0 x_1}{B_0 + D_0 x_1^2}, \quad G_0 = \tfrac{C_0 x_1}{B_0 + D_0 x_1^2} \tag{3.77}$$

$$I_0^+ = \tfrac{1+G_0}{\sqrt{1-G_0^2}} \tan\left(\tfrac{\delta_1^+}{2}\right) \tag{3.78}$$

$$I_0^- = \tfrac{1+G_0}{\sqrt{1-G_0^2}} \tan\left(\tfrac{\delta_1^-}{2}\right) \tag{3.79}$$

$$T_{\Delta d}^{\text{Swan84}} = \tfrac{\lambda}{4\pi n_1 \Delta d} \tfrac{F_0}{\sqrt{1-G_0^2}} \Big[\arctan(I_0^+) + N^+\pi - \arctan(I_0^-) - N^-\pi \Big] \tag{3.80}$$

$$I_{\text{M}} = \tfrac{1+G_0}{\sqrt{1-G_0^2}} \tan\left(\tfrac{2\pi n_1 \Delta d}{\lambda}\right) \tag{3.81}$$

$$I_{\text{m}} = \tfrac{1-G_0}{\sqrt{1-G_0^2}} \tan\left(\tfrac{2\pi n_1 \Delta d}{\lambda}\right) \tag{3.82}$$

$$N_\delta = \text{round}\left(\tfrac{\delta_1^+ - \delta_1^+}{2\cdot 2\pi}\right) \tag{3.83}$$

$$T_{\text{M}}^{\text{Swan84}} = \tfrac{\lambda}{2\pi n_1 \Delta d} \tfrac{F_0}{\sqrt{1-G_0^2}} (\arctan(I_{\text{M}}) + N_\delta\, \pi) \tag{3.84}$$

$$T_{\text{m}}^{\text{Swan84}} = \tfrac{\lambda}{2\pi n_1 \Delta d} \tfrac{F_0}{\sqrt{1-G_0^2}} (\arctan(I_{\text{m}}) + N_\delta\, \pi) \tag{3.85}$$

These formulae for wedged-shaped films were initially found in the subsequent seminal paper by Swanepoel from 1984 [130]. Please consider that we have also added the correction numbers N^+ and N^- to account for highly tilted films, $\Delta d > \lambda/(4n)$, and we have also included a correction factor N_δ for the envelopes. An interesting detail is that when $\Delta d > \lambda/(4n)$ holds, the upper and lower envelopes cross: the lower envelope will stay

above the spectra and the upper envelope below (see Fig. 5c). Over a wide spectral range, one can see how the two envelopes alternate sequentially. Therefore, there are particular wavelengths at which the spectra $T_{\Delta d}^{\mathrm{Swan84}}$ and its two envelopes precisely coincide, as seen in Figs. 3.7c-d. The crossover points are located at $\lambda_{\mathrm{cross}} = 4n\Delta d/N$, for $N = 1, 2, 3, ...$, as discussed in [181]. These particular points contain essential information and allow us to extract accurate information for the optical properties.

The RMSE between the exact Eq. 3.37 and the approximated Eq. 3.80 is 0.465% when substrate absorption is taken into considering and $\Delta d = 30$ nm. The error from this formula is roughly equivalent to the error from the RP formula in Eq. 3.55. The RP formula (which removes the assumption $\kappa_1^2 \ll n_1^2$) offers improvements only up to the fourth decimal place.

Setting $\kappa_2 = 0$ and $\Delta d = 10^{-5}$ nm in Eq. 3.80 leads to a transmission error of 0.039%. As expected, the result obtained was identical to that previously obtained by using Eq. 3.74 for uniform films. This consistency occurs because Eq. 3.80 converges to Eq. 3.74 as $\Delta d \to 0$.

3.1.8.2 Reflection

Setting $\kappa_1^2 \ll n_1^2$ in Minkov's reflection formula (see Eq. 3.69), we derive the following simplified expression for uniform films:

$$a_0 = n_1 - 1, \qquad b_0 = n_1 + 1 \tag{3.86}$$

$$c_0 = n_1 - n_2, \qquad d_0 = n_1 + n_2 \tag{3.87}$$

$$e_0 = n_1 - n_2^2, \qquad f_0 = n_1 + n_2^2 \tag{3.88}$$

$$g_0 = 64\, n_2 \left(n_2 - 1\right)^2 n_1^4 \tag{3.89}$$

$$R^{\mathrm{RP01}} = \frac{(a_0 d_0)^2 + (b_0 c_0 x_1)^2 - 2 a_0 b_0 c_0 d_0 x_1 \cos(\delta_1)}{(b_0 d_0)^2 + (a_0 c_0 x_1)^2 - 2 a_0 b_0 c_0 d_0 x_1 \cos(\delta_1)} + \frac{g_0 x_1^2}{(b_0 d_0)^2 + (a_0 c_0 x_1)^2 - 2 a_0 b_0 c_0 d_0 x_1 \cos(\delta_1)} \times \frac{1}{b_0^3 f_0 + a_0^3 e_0 x_1^2 - 2 a_0 b_0 c_0 d_0 x_1 \cos(\delta_1)} \tag{3.90}$$

$$R_{\mathrm{M/m}}^{\mathrm{RP01}} = \frac{(a_0 d_0 \pm b_0 c_0 x_1)^2}{(b_0 d_0 \pm a_0 c_0 x_1)^2} + \frac{g_0 x_1^2}{(b_0 d_0 \pm a_0 c_0 x_1)^2} \times \frac{1}{\left(b_0^3 f_0 + a_0^3 e_0 x_1^2 \pm 2 a_0 b_0 c_0 d_0 x_1\right)} \tag{3.91}$$

We have verified that these results precisely match the exact formulae from Eq. 3.36 when $\kappa_2 = 0$ and $\kappa_1^2 \ll n_1^2$. Note that the way to effectively set this last approximation is to write $\kappa_1^2 = 0$ when that term is compared to n_1^2 within summations. The upper and lower envelopes defined in Eq. 3.91 correspond to the positive $(+)$ and negative $(-)$ signs, respectively. These formulae were partially developed by Minkov *et al.* [132] in 1989. The RMSE of Eq. 3.90 is 0.42% for absorbing substrates and 0.02% for transparent substrates.

The work [145] by RP *et al.* from 2001 also incorporates equations that address the

scenario of weakly-wedge shaped films:

$$L_0 = b_0^2 d_0^2 + a_0^2 c_0^2 x_1^2, \quad L_1 = b_0^3 f_0 + a_0^3 e_0 x_1^2 \tag{3.92}$$

$$L_2 = 2a_0 b_0 c_0 d_0 x_1, \quad L_3 = \sqrt{L_2 + L_1} \tag{3.93}$$

$$L_4 = \sqrt{L_1 - L_2}, \quad L_5 = a_0^2 d_0^2 + b_0^2 c_0^2 x_1^2 \tag{3.94}$$

$$L_6 = \sqrt{L_0 + L_2}, \quad L_7 = \sqrt{L_0 - L_2} \tag{3.95}$$

$$L_8 = \tan\left(\tfrac{\delta_1^-}{2}\right), \; L_9 = \tan\left(\tfrac{\delta_1^+}{2}\right), \; L_{10} = \tan\left(\tfrac{\delta_1^+ - \delta_1^-}{4}\right) \tag{3.96}$$

$$T_1 = \arctan\left(\tfrac{L_3 L_9}{L_4}\right) + N^+\pi - \arctan\left(\tfrac{L_3 L_8}{L_4}\right) - N^-\pi \tag{3.97}$$

$$T_2 = \arctan\left(\tfrac{L_6 L_9}{L_7}\right) + N^+\pi - \arctan\left(\tfrac{L_6 L_8}{L_7}\right) - N^-\pi \tag{3.98}$$

$$R_{\Delta d}^{\mathrm{RP01}} = 1 - \tfrac{2}{(\delta_1^+ - \delta_1^-)\cdot(L_1 - L_0)} \left[\tfrac{g_0 x_1^2}{L_3 L_4} T_1 + \tfrac{(L_0 - L_5)(L_1 - L_0) - g_0 x_1^2}{L_6 L_7} T_2\right] \tag{3.99}$$

$$T_{1\mathrm{M}} = \arctan\left(\tfrac{L_4 L_{10}}{L_3}\right) + N_\delta \pi, \; T_{2\mathrm{M}} = \arctan\left(\tfrac{L_7 L_{10}}{L_6}\right) + N_\delta \pi$$

$$T_{1\mathrm{m}} = \arctan\left(\tfrac{L_3 L_{10}}{L_4}\right) + N_\delta \pi, \; T_{2\mathrm{m}} = \arctan\left(\tfrac{L_6 L_{10}}{L_7}\right) + N_\delta \pi$$

$$R_{\Delta d,\,\mathrm{M/m}}^{\mathrm{RP01}} = 1 - \tfrac{4}{(\delta_1^+ - \delta_1^-)(L_1 - L_0)} \left(\tfrac{g_0 x_1^2}{L_3 L_4} T_{1\mathrm{M/1m}} + \tfrac{(L_0 - L_5)(L_1 - L_0) - g_0 x_1^2}{L_6 L_7} T_{2\mathrm{M/2m}}\right) \tag{3.100}$$

It should be noted that we have now added the corresponding correction factors to the original reflection formula and its envelopes, accounting also for the highly-wedged shaped films. Figs. 3.7d display the reflectance for a highly-tilted film. In this spectral range, one can also see one shift of the lower and upper envelopes for the reflectance spectrum.

3.1.9 Substrate transmission and reflection

The thin film is typically deposited on a commercial glass substrate, which is relatively thick (with a known $d_2 \sim 1$ mm) and has plane-parallel surfaces, ensuring uniformity. Be-

fore performing the optical characterization of the film under study, it is a common practice to first characterize the optical properties of the substrate (n_2, k_2) using transmittance and/or reflectance measurements of the substrate alone, denoted $\{T_{\mathrm{s}}^{\mathrm{exp}}(\lambda_i), R_{\mathrm{s}}^{\mathrm{exp}}(\lambda_i)\}$. This preliminary step allows us to counteract the influence of the substrate on the overall transmittance and reflectance of the sample, allowing a precise determination of the desired properties of the thin film, namely (n_1, k_1, d_1).

Under the approximation $k_2 = 0$, the thickness of the substrate d_{s} becomes irrelevant for the calculation of the transmission and reflection intensities, as the glass substrate will not absorb light. However, the real refractive index n_2 remains important because it determines the amount of light reflected on the film-to-substrate (interface 02) and substrate-to-air (interface 20), as seen in Fig. 3.2, and described in Section 4.A.

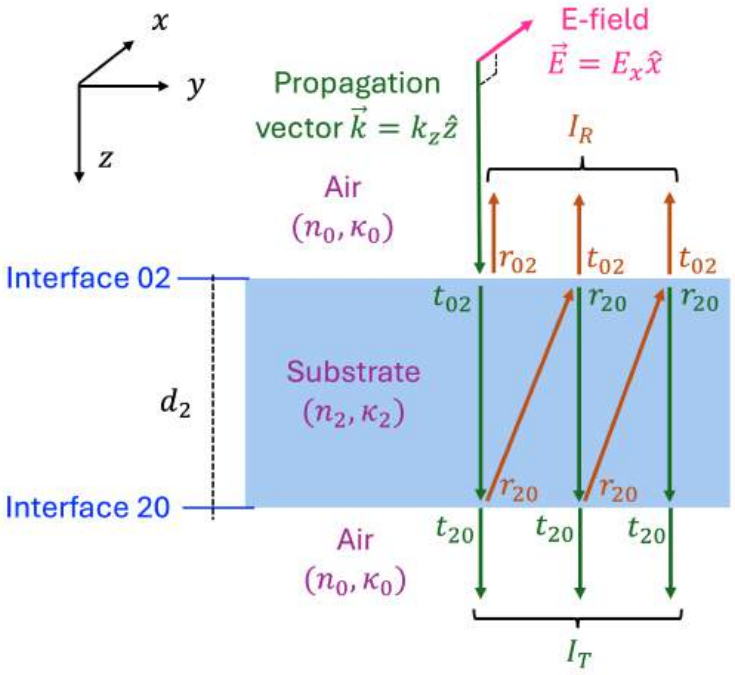


Figure 3.8: Diagram of the spectroscopic measurements for the substrate alone. The measures of T_{s} and R_{s} allow us to extract the substrate properties n_2 and k_2.

Figure 3.8 shows a representative diagram of the spectroscopic measurements for the substrate alone. In the simplistic scenario in which $\kappa_2 = 0$, n_2 can be uniquely determined [184] from either transmission-only or reflection-only measurements by using the following relations:

$$T_{\mathrm{s}}^{\mathrm{approx}} = \frac{2n_2}{n_2^2+1},\ n_2 = \frac{1}{T_{\mathrm{s}}^{\mathrm{approx}}} + \left(\frac{1}{(T_{\mathrm{s}}^{\mathrm{approx}})^2} - 1\right)^{1/2} \tag{3.101}$$

$$R_{\mathrm{s}}^{\mathrm{approx}} = \frac{(n_2-1)^2}{n_2^2+1},\ n_2 = \frac{1+\sqrt{R_{\mathrm{s}}^{\mathrm{approx}}(2-R_{\mathrm{s}}^{\mathrm{approx}})}}{1-R_{\mathrm{s}}^{\mathrm{approx}}} \tag{3.102}$$

When the glass substrate *does* absorb light, both transmission and reflection measurements of the substrate are necessary to uniquely determine $n_2(\lambda)$, $k_2(\lambda)$, and d_{s}. The formulae can be found using the same Abele transfer matrix formalism as in Section 4, and then using the *spectral averaging* method to account for the limited coherence length, $L < 2d_{\mathrm{s}}$. However, for the particular case of a simple slab surrounding by air, it becomes more convenient to use the traditional technique of infinite incoherent summation [167]:

$$T_{\mathrm{s}} = T_{02}^2 e^{-\alpha_2 d_2} \times \sum_{m=0}^{\infty} \left[R_{02} e^{-\alpha_2 d_2}\right]^{2m} = \tag{3.103}$$

$$\frac{T_{02}^2 e^{-\alpha_2 d_2}}{1-R_{02}^2 e^{-2\alpha_2 d_2}}$$

$$R_{\mathrm{s}} = R_{02} + T_{02}^2 R_{02} e^{-2\alpha_2 d_2} \times \sum_{m=0}^{\infty} \left[R_{02} e^{-\alpha_2 d_2}\right]^{2m} = \tag{3.104}$$

$$R_{02} + \frac{R_{02}\, T_{02}^2 e^{-2\alpha_2 d_2}}{1-R_{02}^2 e^{-2\alpha_2 d_2}}$$

Here, $R_{02} = |\mathbf{r}_{02}|^2 = |\mathbf{r}_{20}|^2$ and $T_{02} = |\mathbf{t}_{02}|^2 = |\mathbf{t}_{20}|^2$ represent the square norm of the Fresnel's coefficients. From Eqs. 3.10-3.9, we can see that, at normal incidence, the normed square of these coefficients is the same for the air-to-substrate and substrate-to-air interfaces. Following the diagram from Fig. 3.8, we can see that the first addend "R_{02}" from

Eq. 3.104 corresponds to the first bounce back of the light beam at the air-to-substrate interface. The second addend (for $m = 0$) occurs when: *(i)* the light passes through the air-to-substrate interface (T_{02}), *(ii)* it then gets reflected at the substrate-to-air interface (R_{02}), and *(iii)* it is finally transmitted through the substrate-to-air ($T_{20} = T_{02}$) interface bouncing back to the detector above. During the round-trip of the beam within the slab, the Lambert-Beer law describes the amount of light absorption; this absorption is the responsible for the exponential term for $m = 0$ in the Eq. 3.104. Further interreflections (for $m \geq 1$) just include other addends that can be reasoned in a similar fashion.

These *direct* formulae that define T_{s} and R_{s} as functions of the substrate optical properties have been known for more than a century. However, determining the substrate properties (n_2, k_2, d_2) from T_{s} and R_{s} is a more intricate *inverse* problem. The work by Nichelatti [167] in 2002 provided the first-ever *analytical* expressions for that scenario. The following relations were found:

$$R_{02}(T_{\mathrm{s}}, R_{\mathrm{s}}) = \frac{2+T_{\mathrm{s}}^2-(1-R_{\mathrm{s}})^2}{2(2-R_{\mathrm{s}})} - \frac{\sqrt{\left(2+T_{\mathrm{s}}^2-(1-R_{\mathrm{s}})^2\right)^2-4R_{\mathrm{s}}(2-R_{\mathrm{s}})}}{2(2-R_{\mathrm{s}})} \tag{3.105}$$

$$k_2 = \frac{\lambda}{4\pi h}\ln\left(\frac{R_{02}T_{\mathrm{s}}}{R_{\mathrm{s}}-R_{02}}\right) \tag{3.106}$$

$$n_2 = \frac{1+R_{02}}{1-R_{02}} \pm \sqrt{\frac{4R_{02}}{(1-R_{02})^2} - k_2^2} \tag{3.107}$$

Both the positive and negative signs in Eq. 3.107 define correct mathematical solutions to Eq. 3.104. However, only one solution makes physical sense. While it is possible to discriminate the correct solution on a case-by-case basis by using commonly reported values, in our particular spectroscopic analysis, only the *positive* solution is physically meaningful. Note that the film can be analyzed in spectral regions of high absorption (when $\kappa_1^2 \not< n_1^2$), as it can be very thin and still permit some transmission. However, using

fully absorbing substrates with a high κ_2 in the spectral region of interest is *not* desirable, as their significant thickness would then prevent any transmission entirely. Therefore, $\kappa_2^2 \ll n_2^2$ is always the practical scenario for the case of the substrate. That implies that n_2 from Eq. 3.107 is greater than one for the positive square-root (the correct physical solution) and less than one for the negative square-root (which must be discarded).

We can now compare the approximate formulae (with $\kappa_2 = 0$) from Eqs. 3.101 and 3.102 with the exact expressions from Eqs. 3.103 and 3.104. It is important to note that for an absorbing substrate with $\kappa_2 = 10^{-6}$, the RMSE of the approximated formulae is 0.3% for transmittance and 0.2% for reflectance. Deriving n_2 from the simpler Eqs. 3.101 and 3.102 leads to an RMSE of 0.010% and 0.005%, respectively, for our simulated data.

3.1.10 Concluding remarks

In the present article, we derived the formulae for transmittance and reflectance of a normal-incident quasi-coherent light beam in a sample with a thin film, possibly with a high wedge profile, deposited on a thick absorbing substrate. Our model assumes a homogeneous, isotropic, and non-magnetic film material with linear response.

In addition, we review other relevant formulae commonly used in the literature, examining their approximations and the level of accuracy. Rougher approximations lead to formulae that are easier to implement and compute. Then, a trade-off can be found between the accuracy of more complex expressions and the efficiency of approximate ones. Depending on the specific sample and the desired accuracy level, one can choose the most suitable model. For instance, for many conveniently-prepared films with uniform-thickness analyzed in the medium-to-low-absorption region, the simple Swanepoel formulae provide good accuracy with minimal computation time. However, our newly-developed

formula proves to be more accurate in other more complicated scenarios: For instance, it is particularly effective when dealing with an 'exotic' film surface, modeled with a large wedge parameter, $\Delta d > \lambda/(4n)$, over a small illumination spot. Additionally, it is beneficial for the analysis across wide spectral ranges that encompass regions with strong film absorption ($\kappa_1^2 \not< n_1^2$) and substrate absorption ($\kappa_2 > 0$).

A review of the formulae analyzed in our work can be found in Table 3.1. These expressions have been written with a consistent notation, tested numerically and analytically, and compared with each other. These formulae have been coded both in Python and in Matlab, and they are publicly available link. Future work will focus on further comparison of the developed formulas with respect to the traditional ones with experimental measurements, providing an empirical proof of their practical utility. Additional future work is also needed to include absorbing substrates and tilted films in multilayer configurations, as it will enhance the applicability of the formulas to more complex thin-film technologies, such as those employed in organic light-emitting diodes (OLED) and organic photovoltaic cells (OPV). Investigating oblique beam incidence and non-specular surfaces, where there is light scattering due to the film surface roughness, will broaden the potential applications to encompass a wider range of film structure analyses.

3.2 Optimization based Envelope Detection Algorithm

The Swanepoel method is a widely used optical technique for characterizing thin films through normal-incidence transmission measurements. A critical step in this approach involves extracting the upper and lower envelopes of the measured oscillatory spectrum. By analyzing the transmission spectrum and its corresponding envelopes, the Swanepoel procedure enables precise determination of the film refractive index and extinction coefficient. However, even minor inaccuracies in envelope construction can propagate significant errors into the final characterization results. To address this challenge, we present a novel physics-informed optimization algorithm for envelope detection. Our mathematical model reformulates the envelope detection problem as a global optimization task that enforces the physical properties of the Swanepoel envelopes. Extensive validation on fifty randomly generated transmission spectra demonstrates unprecedented accuracy: the method achieves root-mean-square errors (RMSE) below 0.10% for the upper envelope and 0.06% for the lower envelope (more than doubling the accuracy of current state-of-the-art approaches). Furthermore, the proposed model has been assessed with two experimental transmission spectra, demonstrating its robustness and accuracy with real noisy data. We have developed an open-source Python software package, which is available online. This software includes not only our innovative envelope construction algorithm but also additional envelope drawing algorithms for comparison and an efficient implementation of the Swanepoel method, enabling complete optical characterization of thin films.

3.2.1 Introduction

Thin solid films are indispensable components in a vast array of optical systems and optoelectronic devices. *Semiconductor* thin films are foundational in the fabrication of transistors, diodes, and photodiodes [196]–[199]. For instance, silicon thin films are extensively employed for cost-effective optical sensors [200]–[202] and solar cells [89], [182], [203] due to their efficient light-absorbing capabilities and scalable manufacturing processes. *Conductive* thin films are highly valued for their exceptional electrical conductivity and reflectivity, enabling their use in electrode fabrication, optical mirrors, protective coatings, and forming electrical connections within integrated circuits [204]–[207]. In particular, transparent conductive films play a crucial role in the development of touch panels and liquid crystal displays. *Dielectric* thin films, engineered with specific refractive indices and thicknesses, create thin-film capacitors, act as insulating layers, and form anti-reflection coatings that enhance light transmission [208], [209]. Dielectric coatings are integral to optical instruments such as lenses, bandpass filters, dichroic mirrors, and beamsplitters, providing precise control over light propagation and wavelength selection [210].

The manufacturing processes of the thin film material critically influence their optical properties [211], [212], which in turn affect the performance of the target devices. Therefore, precise determination and control of these properties are essential for optimizing the mentioned optoelectronic devices and optical systems. By performing optical characterization of thin films, we can obtain their refractive index and extinction coefficient, denoted as $n(\lambda)$ and $k(\lambda)$, respectively. These wavelength-dependent optical properties combine to form the complex refractive index ($\mathbf{n} = n + ik$) or the complex dielectric function ($\boldsymbol{\epsilon} = \mathbf{n}^2$). The dielectric function offers insights into the optoelectronic proper-

ties of the film material and allows for precise quantification of the electronic transitions. Therefore, one can extract information about critical parameters such as bandgap energy [213]–[215], molecular disorder [216], and charge mobilities [217].

A myriad of techniques exist to determine the optical properties of thin solid films. Spectroscopic analyses are common [13], [218], [219] for characterizing materials across a broad spectral range (UV-VIS-NIR), employing transmittance, reflectance, or ellipsometric measurements. Spectroscopic ellipsometry, in particular, is a commonly employed technique that measures the changes in the polarization state of light upon reflection, at varying angles of incidence. This technique can produce a comprehensive phase-sensitive dataset, allowing for the precise characterization of complex film structures, including surface roughness, thickness variations, multilayer systems, and anisotropic materials. In contrast, conventional spectrophotometric techniques, which rely on direct measurements of the transmitted or reflected light intensity, offer a simpler approach for obtaining an overall average estimation of a material's optical properties. While spectrophotometry is less sensitive to fine structural variations, it remains a widely used method due to its straightforward implementation and broad applicability. The present study focuses on characterizations derived solely from normal-incidence transmission spectra, which are optimal for assessing material properties at photon energies below the optical bandgap, where moderate to low absorption enables highly sensitive transmission data. Typically, the thin-film material is deposited on a thick glass substrate with negligible absorption ($k_{\mathrm{s}} \approx 0$) and a near-constant refractive index (n_{s}) in the spectral region of interest (ROI). Before the thin-film preparation, the transmittance of the substrate alone, denoted $T_{\mathrm{s}}(\lambda)$, is sequentially collected over different wavelengths using a spectrophotometer, allowing us to easily account for the effect of the substrate when analyzing the whole sample. As illustrated in Fig. 3.9a, the transmittance of the complete sample (thin film over the sub-

strate), $T(\lambda)$, is subsequently measured.

The Swanepoel method and other characterization approaches: In 1983, Ryno Swanepoel [129] introduced an numerical procedure, widely known nowadays as Swanepoel method or Envelope method, to determine the thin-film thickness d and its optical functions (n and k) at specific wavelengths. These particular wavelength of interest are called *tangent points* (see Fig. 3.9c), as they appear where the transmission spectrum intersects (touch but not cross) its upper or lower envelopes. Although earlier envelope-based approaches had been developed before Swanepoel[126]–[128], his contribution significantly enhanced the precision and applicability of the technique, leading to a widespread algorithm extensively used today. Renowned for its simplicity, reliability, and computational efficiency [147], the Swanepoel method operates in the order of a few seconds on standard computers, making it ideal for rapid mass characterization of samples. Additionally, it is considered a "model-free" method, meaning that it does not rely on pre-established optical dispersion models for $n(\lambda)$ and $k(\lambda)$.

This method analyzes the interference patterns created by a quasi-monochromatic beam as it passes through a thin film, where the patterns result from the superposition of multiple internal reflections within the film (see Fig. 3.9b). For the desired etalon (Fabry-Perot) effect to manifest prominently within the thin film, the coherence length of the beam must exceed the film thickness (typically on the micrometer scale), denoted d. This ensures that multiple internal reflections within the film lead to significant constructive and destructive interference, depending on the wavelength. Consequently, the transmission spectrum $T(\lambda)$ exhibits characteristic sinusoidal oscillations with distinct peaks and valleys, as illustrated in Fig. 3.9c. The Swanepoel method exploits these spectral oscillations to accurately determine the film properties. Note that in order to prevent the sub-

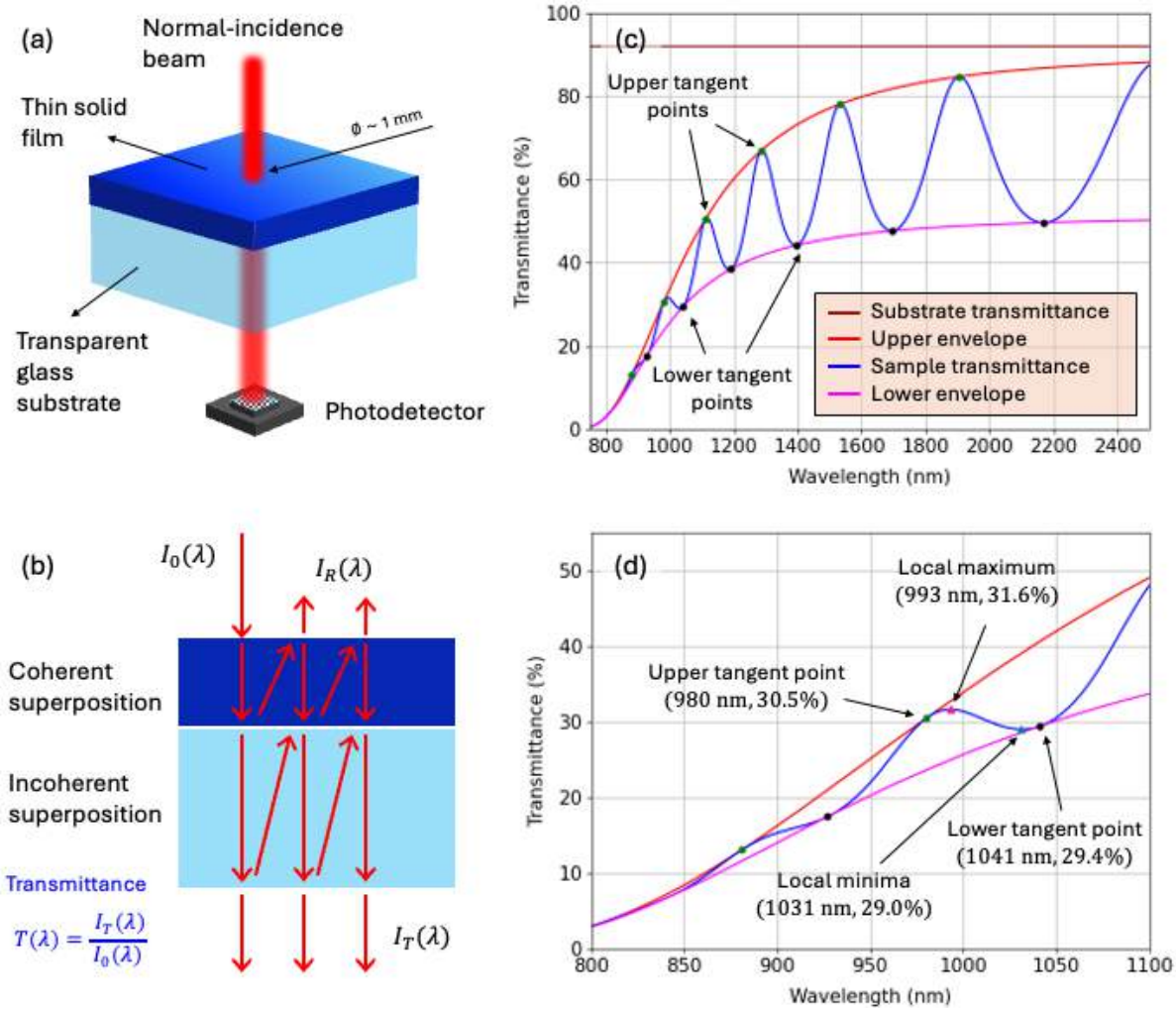


Figure 3.9: (a) Schematic of the sample geometry, comprising a thin layer deposited on a glass substrate. (b) Multiple internal reflections within the thin film and the substrate. The total transmittance is measured sequentially throughout different wavelengths. (c) Simulated transmission spectra for the full sample and for the substrate alone. The ground-truth upper and lower envelopes, along with their tangent points, are indicated. (d) Magnified view of the strongly absorbing spectral region. The upper and lower tangent points are sometimes not associated with a nearby local extremum (first two on the left), or occur at positions significantly far from the local extrema (last two on the right).

strate from generating its own etalon effects, the coherence length must be set shorter than the substrate thickness (typically on the millimeter scale), denoted d_s. As the coherence length increases with wavelength, it may eventually exceed the substrate thickness at longer wavelengths. This can lead to very high-frequency oscillations [4], causing

aliasing when the oscillation frequency surpasses the spectral sampling rate used in the experiment. Such aliasing manifests as characteristic noise at the spectrum's end if the experiment is not carefully calibrated. This condition ensures that any multiple reflections within the substrate do not contribute to coherent interference patterns, resulting instead in an incoherent superposition of the beams. The exact value of d_{s} is therefore not relevant for the characterization of the sample.

Since its inception, the Swanepoel method has served as a key computational method for the analysis of thin-film optical analysis, inspiring numerous enhancements that have significantly broadened its applicability. For instance, shortly after introducing his original algorithm, Swanepoel himself extended the methodology in 1984 [130], enabling the analysis of films with tilted planar-specular surfaces. The algorithm can then apply to samples with non-uniform thickness, generally characterized by a wedge Δd. Recognizing the limitations of spectrophotometric measurements for accurate film analysis, Swanepoel proposed an improved method in 1985 [220]. This approach involved acquiring transmission spectra at two different angles of incidence: one at normal incidence and the other at an oblique angle. By analyzing the shift in the spectral domain between these two measurements, the method enabled the direct determination of both the film thickness d and refractive index n without relying on the absolute transmittance values. In 1986, Kushev *et al.* generalized the envelope method to the field of reflectometry [131], enabling the determination of thin-film optical properties from normal-incidence reflectance measurements. In 1989, another work by Swanepoel now accounted for the substrate absorption [164]. In 1995, Nowak [221] investigated the effects of realistic band-limited incident beams with either linear or Gaussian spectral profiles, further refining the accuracy of the method. Later, Wu *et al.* (2008) [222] extended the method to analyze inhomogeneous films with a linearly graded refractive index from the substrate interface to the film

surface. Jin *et al.* contributed two very significant advancements in 2017. First, they analyzed the spectral regions in which the film presents strong absorption [143], where the original Swanepoel method produced substantially less accurate solutions. Then, they expanding the algorithm to determine the refractive index n and extinction coefficient k at arbitrary wavelengths, not just at the mentioned tangent points [223].

In addition to the Swanepoel method, the *inverse synthesis* approach serves as a powerful alternative for the optical characterization of thin films [119], [121], [224], [225]. This method involves fitting experimentally measured transmission spectra with a simulated transmission curve, which is derived from pre-defined dispersion models for the refractive index $n(\lambda)$ and the extinction coefficient $k(\lambda)$. One of the key advantages of the inverse synthesis approach is its ability to provide precise optical characterization over a continuous broad spectral range [148]. Furthermore, this method allows for the modeling of more complex physical scenarios, including surface roughness, refractive index gradients, and non-uniform film thickness with specific geometries. The conventional Swanepoel method approximates these realistic films as a simple, idealized homogeneous layers. Neglecting these film imperfections leads to correlated effects on the spectral envelopes, making it impossible to distinguish such features based on spectro photometric measurements alone. Consequently, the Swanepoel method is primarily useful as a first-order approximation for optical characterization [148], [226]. Envelope-based characterization can serve as a preliminary estimate for subsequent refinement using inverse synthesis.

Despite its advantages, it should be noted that the inverse synthesis approach can be computationally intensive. Additionally, solving the inverse problem does not always guarantee physically meaningful solutions, as the optimization algorithm may converge

to incorrect local minima, leading to non-physical parameter estimations (this problem is further discussed in the next section). Given these challenges, it is often advisable to use the simpler Swanepoel method as a complementary tool for validation and cross-checking of results, a practice commonly reported in the literature [183], [226]. Implementing the inverse synthesis approach is non-trivial and demands careful parameter selection. This reverse engineering approach relies on an appropriate dispersion formula that accurately describes the refractive index and absorption behavior of the film material. In some cases, determining a suitable model requires a detailed understanding of the material's electronic transitions to correctly incorporate contributions from different optical oscillators. A variety of dispersion models are available, ranging from the simple empirical Cauchy model (suitable for transparent materials such as glass), to more sophisticated physically-motivated models, like the Tauc-Lorentz (TL) or Cody-Lorentz (CL) models [227], which account for absorption. The development of more advanced dispersion models remains an active area of research, with recent efforts focusing on models such as the Analytical Tauc/Cody-Lorentz Urbach (ATLU or ACLU) model [228]–[230] or the Universal Dispersion Model (UDM) [231]–[234], which provides broad applicability across different material classes and spectral ranges.

Besides the envelope method and the inverse synthesis, further computational approaches are emerging. For instance, there is a Deep Learning architecture that directly predicts a film properties from transmission data [6]. Another recent technique employs a digital twin [150] to numerically simulate wave propagation through the spectrophotometer.

Envelope detection algorithms: The Swanepoel method assumes precise knowledge of both the upper and lower envelopes of the transmission spectrum (refer to Fig. 3.9b).

In practice, accurately identifying these envelopes is a significant challenge on its own. The accuracy of the Swanepoel method is highly dependent on the precise determination of these envelopes, as any discrepancies in their construction propagate through the entire algebraic procedure, ultimately affecting the reliability of the film characterization. In 1991, Marjorie McClain *et al.* [235] addressed this problem by introducing a robust iterative algorithm for envelope construction. This seminal work has since established the McClain algorithm as the standard approach, extensively utilized in the literature [236]–[238]. The McClain algorithm is both efficient and reasonably accurate in practical applications. It determines the envelopes using B-spline interpolation of the mentioned tangent points, which are seen in Fig. 3.9c. As anticipated, the accuracy of this interpolation-based method improves with an increasing number of interpolation nodes, requiring at least four spectral oscillations with its four associated tangent points to produce reliable results. Beyond the original McClain algorithm, several other interpolation-based methodologies have been proposed to enhance envelope detection [239]–[242].

Alternative approaches have been developed to address spectra with limited number of oscillations, sometimes displaying only two or three tangent points. A common strategy in such scenarios involves fitting the envelopes by connecting local extrema (maxima for the upper envelope and minima for the lower envelope) using predefined exponential [243], [244] or parabolic functions [142]. However, it should be mentioned that, although the tangent points typically lie near the extrema of the transmission spectrum, they do *not* coincide precisely. Indeed, observe that the film absorption induces a decay in the transmittance curve as wavelengths decrease and approximates the optical band gap, causing the tangent points to shift slightly away from the extrema, as seen in Fig. 3.9d. In the regions of strong film absorption (where the transmission spectrum remains low) the tangent points may not correspond to any local minimum or maximum at all,

as illustrated by the first lower and upper tangent points of the simulated spectrum, in Fig. 3.9d. Consequently, the mentioned parabolic and exponential fits often lead to noticeable errors in the constructed envelopes. A recently proposed deep learning–based method [7] developed by our team shows promise in accurately determining envelopes for transmission spectra, even in these challenging scenarios with as few as two or three tangent points. However, this approach has so far been validated only with simulated data and not with real, noisy experimental measurements.

Enhancing the accuracy of envelope computation remains a significant and ongoing challenge in the field. In the present work, we propose a novel interpolation-based algorithm for the precise determination of Swanepoel envelopes. Our approach integrates the physical properties of the Swanepoel envelopes into a global optimization framework, which is then efficiently solved using an advanced optimizer. We validated our method using fifty computer-generated transmission spectra, achieving a root-mean-square error (RMSE) of less than 0.10% for the upper envelope and below 0.06% for the lower envelope. These results demonstrate a significant enhancement in accuracy compared to existing methods, reducing errors by more than half with respect to the standard McClain algorithm, as detailed in the subsequent comparison section. While our proposed method is more computationally expensive than the existing algorithms, it still performs in a reasonable timeframe and runs in less than a minute using standard computer hardware.

Furthermore, we have integrated our innovative envelope construction algorithm with an automated adaptation of the original Swanepoel method within a user-friendly Python package named *swanpy*. This open-source software is publicly accessible at this link. Note that *swanpy* does not only includes our novel envelope construction algorithm but also incorporates additional standard envelope detection techniques for comparative

analysis. Our automated and efficient implementation of the Swanepoel method facilitates the comprehensive optical characterization of thin films. It is worth noting that the present work demonstrates the effectiveness of *swanpy* in practical applications by fully characterizing two real amorphous silicon (*a*-Si) samples, which were previously analyzed by our team [183] and used here as a benchmark.

3.2.2 Thin film transmission equation and inverse-synthesis challenge

As discussed earlier, the inverse synthesis method requires a simulator to reproduce the transmission curve, which is then fitted to experimental data. To develop such a simulator, we begin with the theoretical expression for normal-incidence transmittance, derived from the Fresnel equations [36]. Following extensive derivations in [4], the final formula relates the wavelength λ and the film's optical properties (n, k, d) to the transmittance. Assuming a uniform, isotropic, and homogeneous film deposited on a transparent substrate with a known refractive index n_s, the resulting expression is:

$$T^{\text{sim}}(\lambda; n, k, d) = \frac{A'x}{B' + x(C_1' \sin\varphi - C_2' \cos\varphi) + D'x^2}, \tag{3.108}$$

where

$$\alpha = 4\pi k/\lambda, \tag{3.109a}$$

$$x = e^{-\alpha d}, \tag{3.109b}$$

$$\varphi = 4\pi n d/\lambda, \tag{3.109c}$$

$$A' = 16(n^2 + k^2)n_s, \tag{3.109d}$$

$$B' = \big((n+1)^2 + k^2\big)\big((n+1)(n+n_s^2) + k^2\big), \tag{3.109e}$$

$$C_1' = 2k\big(2(n^2 + k^2 - n_s^2) + (n^2 + k^2 - 1)(n_s^2 + 1)\big), \tag{3.109f}$$

$$C_2' = 2(n^2 + k^2 - 1)(n^2 + k^2 - n_s^2) - 4k^2(n_s^2 + 1), \tag{3.109g}$$

$$D' = \big((n-1)(n-n_s^2) + k^2\big)\big((n-1)^2 + k^2\big). \tag{3.109h}$$

Observe that the refractive index of the substrate, n_s, is known because it can be computed beforehand from the transmission of the substrate alone, T_s, using the equation [4]:

$$n_s = 1/T_s + \sqrt{1/T_s^2 - 1}. \tag{3.110}$$

It should also be emphasized that α (1/cm) in Eq. 3.109a corresponds to the absorption coefficient. Additionally, the term x in Eq. 3.109b denotes the a fraction of the transmitted intensity: this parameters is calculated from the Lambert-Beer law [36] and ranges from 0 to 1, expressing the amount of light that passes through the film in one single pass (not accounting for multiple round trips of the beam within the film). Lastly, the two trigonometric functions in the denominator of Eq. 3.108 depend on the phase φ and account for the Fabry-Perot thin-film interferences, corresponding to the low-frequency oscillations noticeable in the transmission spectrum.

One could naively develop a least-squares regression program to fit the experimental transmission spectrum $T(\lambda)$ to the theoretical model $T^{\text{sim}}(\lambda; n(\lambda), k(\lambda), d)$, thereby determining the film's optical and geometrical properties $n(\lambda)$, $k(\lambda)$, and d. However, n and k are intricate wavelength-dependent functions and not simple constants. Multiple combinations of n and k can produce the same theoretical transmission spectrum, making this fitting problem *ill-posed* in nature. As Kushev and Zheleva [131] realized in 1986, "the system is satisfied by many solutions and it is difficult to identify the real physical one".

The *inverse synthesis* method takes this fitting idea a step further and eases the problem by assuming some pre-defined dispersion models for the optical functions $n(\lambda)$ and $k(\lambda)$. Consequently, these complicated optical functions are expressed in terms of just a few real coefficients. For instance, one can employ the simple Wemple-DiDomenico (WDD) dispersion equation [245] for n and the Urbach rule [246] for the absorption coefficient α (the extinction coefficient k can then be derived using Eq. 3.109a). We then have the following formulas:

$$n(E; E_0, E_{\text{d}}) = \sqrt{\frac{E_0 E_{\text{d}}}{E_0^2 - E^2} + 1}, \tag{3.111a}$$

$$\alpha(E; \alpha_0, E_{\text{u}}) = \alpha_0 \exp\left(\frac{E}{E_{\text{u}}}\right). \tag{3.111b}$$

The parameter E_0 (eV) represents the oscillator energy in a single-oscillator electron model. In practice, this number quantifies an average energy gap of the different electronic transitions and provides a good approximation of the optical bandgap energy, $E_{\text{g}} \approx E_0/2$ [247], [248]. E_{d} (eV) is the dispersion energy, which quantifies the strength of interband optical transitions and provides an extra degree of freedom to account for the refractive index value as the photon energy approaches zero $E \to 0$. The parame-

ter α_0 (1/cm) defines the absorption coefficient in the limit of zero photon energy. The Urbach energy, denoted as E_{u} (eV), characterizes the exponential decay of the absorption coefficient with decreasing photon energy below the bandgap, indicating the degree of structural disorder in amorphous materials responsible for sub-bandgap absorption. Please note that Eqs. 3.111a-3.111b can be similarly expressed as functions of wavelength using the Planck-Einstein relation $E = hc/\lambda$, where h is Planck's constant and c is the speed of light in vacuum.

The combined WDD-Urbach model offers a foundational estimate of the electronic transitions in thin-film materials. This model is primarily applicable to amorphous materials with two main energy levels, encompassing a single dominant electronic transition between the valence and conduction bands, which defines the bandgap energy. However, it disregards weaker transitions that may play a significant role in crystalline materials. Furthermore, this dispersion model is limited to specific spectral regions, particularly for photon energies that are well below the optical bandgap but remain sufficiently separated from phonon resonances in the near-infrared range. In the simulated example illustrated in Fig. 3.9, we used the following parameter values: $E_0 = 4.28$ eV, $E_{\mathrm{d}} = 30.45$ eV, $\alpha_0 = 269.46$ mm^{-1}, $E_{\mathrm{u}} = 229.18$ meV, and $d = 1323$ nm.

Employing more precise models that are consistent with the Kramers-Kronig relations between n and k increases computational complexity considerably. This complexity is particularly evident when utilizing various modifications and adaptations of the Tauc-Lorentz-Urbach (TLU) [183], [227], [228], [249] and Cody-Lorentz-Urbach (CLU) [227], [229], [250] models, which are valid across the entire spectral range. These advanced models, enhanced with numerous refinements, necessitate extensive computations and introduce additional subtle and intricate variables into the simulator.

Regardless of the optical dispersion model chosen, the inverse synthesis approach enables the simplification of the transmittance formula $T(\lambda; n(\lambda), k(\lambda), d)$ from Eq. 3.108 in terms of only a few input coefficients. In our case, the WDD-Urbach model allows us to rewrite the simulated transmission curve as a function of $(E_0, E_\mathrm{d}, \alpha_0, E_\mathrm{u}, d) \in \mathbb{R}^5$. With these scalar values, we can establish the following least-squares optimization problem (outlined in Fig. 3.10):

$$\begin{aligned}
&\text{minimize} \quad \mathcal{C}(E_0, E_\mathrm{d}, \alpha_0, E_\mathrm{u}, d) = \mathrm{RMSE}(T, T^\mathrm{sim}) = \sqrt{\frac{\sum_{i=0}^{N} \left[T(\lambda_i) - T^\mathrm{sim}(\lambda_i; E_0, E_\mathrm{d}, \alpha_0, E_\mathrm{u}, d)\right]^2}{N}}, \\
&\text{subject to} \quad \begin{cases} E_0 \in [a_1, b_1], \quad E_\mathrm{d} \in [a_2, b_2], \\ \alpha_0 \in [a_3, b_3], \quad E_\mathrm{u} \in [a_4, b_4], \\ d \in [a_5, b_5]. \end{cases}
\end{aligned} \tag{3.112}$$

The cost function $\mathcal{C}$ quantifies the discrepancy between the predicted and experimental spectra across the discrete wavelengths within the spectral region of interest (ROI). Achieving cost values below 1-1.5% typically indicates a good fit between the simulated and measured data [183], [251]. However, the acceptable threshold for the cost function depends on the spectral range length and the specific characteristics or noise level of the measurements. Note that in Eq. 3.112, the box constraints $[a_i, b_i]$ define the acceptable ranges for the parameters, which vary depending on the specific thin-film material under investigation.

Although this inverse synthesis approach is a popular choice for optical characterization, it is not without limitations. For instance, there is the crucial task of setting feasible constraint ranges for a_i and b_i, as the computational effort increases notably when

these ranges are broadened. Furthermore, the inverse problem is inherently non-convex in nature, marked by numerous local minima and maxima, and necessitating the use of sophisticated global optimization algorithms. These global optimizers are often affected by the curse of dimensionality, resulting in substantially increased computational costs as the number of variables grows. Additionally, in practical scenarios with noisy data and approximated simulation models, these algorithms can still produce solutions that are mathematically acceptable (satisfying constraints and exhibiting low cost values) but physically inaccurate, corresponding to local minima rather than the global optimum. To mitigate the risk of converging to suboptimal local minima, it is recommended to execute these optimization algorithms multiple times with different initial points or random seeds and to compare the resulting solutions. This approach enhances the likelihood of identifying the most accurate and reliable characterization of the system.

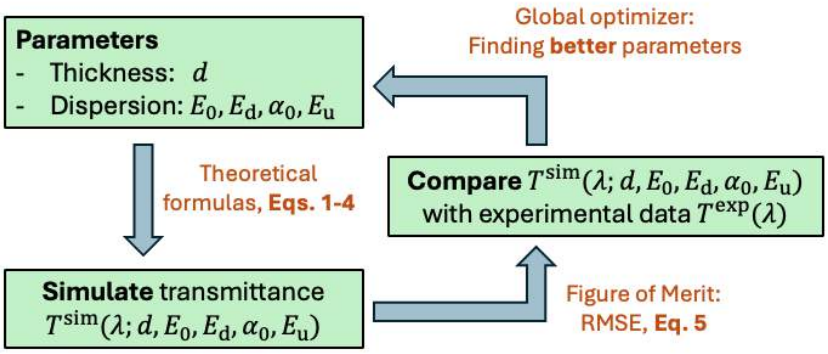


Figure 3.10: Random initial guesses for the film parameters (such as thickness and the coefficients of the optical dispersion model) are selected. A theoretical transmittance spectrum is calculated from these parameters via Eq. 3.108. The simulated spectrum is then compared to the experimental spectrum, and the residual error is quantified using the cost function (RMSE) defined in Eq. 3.112. A global optimization algorithm iteratively adjusts the parameters to minimize this residual error. This process repeats until the optimal parameters that yield the best fit between the simulated and experimental spectra are found.

The runtime of the inverse synthesis method is influenced by several factors, including the particular dispersion model employed and its number of variables, the choice of

the global optimizer and its hyperparameters (such as the number of iterations or initial seeds), the range of the box constraints, and the number of times the optimization is executed to corroborate results. Consequently, solving the inverse problem described by Eq. 3.112 can become a computationally intensive task, sometimes requiring hours of manual fine-tuning and computation.

For a later comparison of our improved Swanepoel method, we use now the inverse synthesis method to characterize two real spectra. Specifically, two *a*-Si samples were prepared using radio-frequency magnetron sputtering (RFMS) on top of a transparent Corning Eagle XG glass substrate. The target p-type silicon had a 5N purity and a density of 2.32 g/cm^3. During thin-film deposition, the RFMS chamber was maintained at room temperature (25 °C) with an RF power of 525 W. The deposition rates were 1.269 nm/s for sample #1 and 1.339 nm/s for sample #2. The chamber pressures were 4.4 Pa and 3.2 Pa, and the argon gas flow rates were 70 sccm and 50 sccm for samples #1 and #2, respectively.

Transmission measurements of both the substrate alone and the full sample were collected using an Agilent Cary 5000 double-beam spectrophotometer, covering the UV-VIS-NIR range from 300 to 2500 nm. The initial spectral range (up to 700 nm) was cropped out to avoid the interval where the substrate exhibits considerable absorption and where the thin film shows very strong absorption (leading to a near-zero transmission), which provides no useful information for characterization. The fitting problem was then solved using the Simulated Dual Annealing algorithm, an efficient global optimizer available in the Python *scipy* package [252]. Figure 3.11 illustrates the agreement between the experimental and simulated spectra, with root mean square errors (RMSE) of 0.85% for sample #1 and 0.98% for sample #2. The bottom panel of the figure displays the residual disagreement or fitting error, defined as $T^{\text{exp}} - T^{\text{sim}}$. Notably, in the NIR region beyond 2100 nm,

where the substrate exhibits slight absorption, a stronger disagreement is observed in Fig. 3.11a.

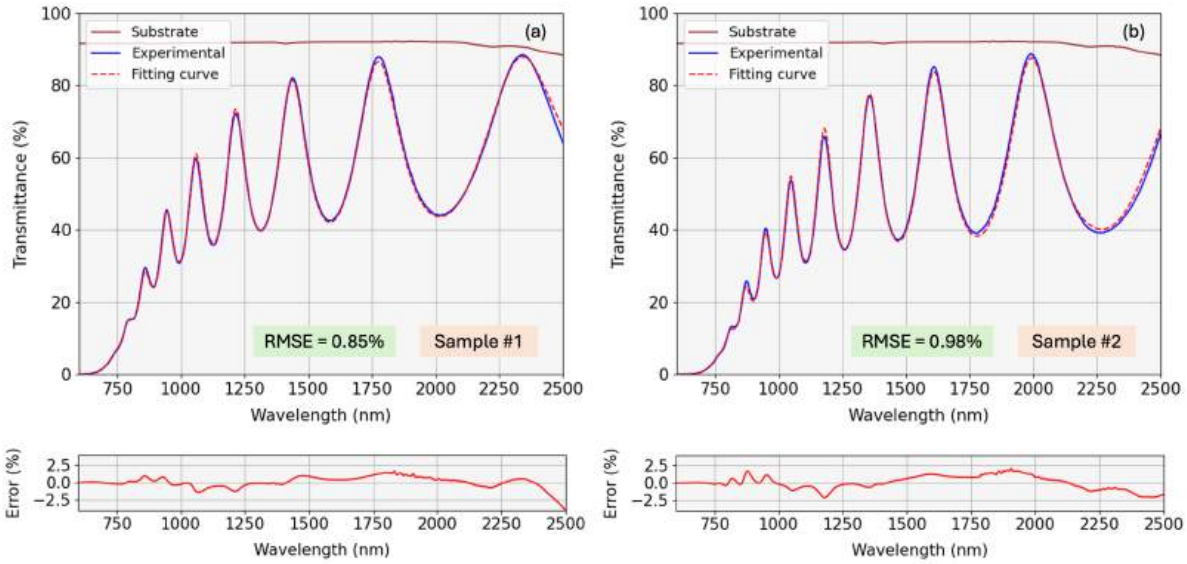


Figure 3.11: Comparison of experimental transmission measurements with simulated transmission curves for (a) sample #1 and (b) sample #2 of amorphous silicon, obtained using the inverse synthesis method. The bottom panels illustrate the residual differences or error between the experimental and simulated data, defined as $T^{\text{exp}} - T^{\text{sim}}$.

The optimal parameters obtained from the inverse synthesis method for samples #1 and #2 are summarized in Table 3.3. Initially, the points for the optimization were randomly chosen within the specified bounds for each parameter. In this particular case, the optimization process required only a few minutes of computation time and manual adjustments of the bounds, thanks to the relatively low dimensionality of the simplistic models for n and k employed here for illustrative purposes.

3.2.3 Automated Swanepoel method

As previously noted, the Swanepoel method provides a more straightforward alternative to the inverse synthesis approach. It relies on the principles of Fabry-Perot interference

Table 3.3: Optimized parameters for amorphous silicon samples obtained from the inverse synthesis method. The initial bounds for each parameter are also provided.

Parameter	**Unit**	**Lower Bound**	**Upper Bound**	**Sample #1**	**Sample #2**
Thickness d	nm	500	1500	1117	1160
Oscillator energy E_0	eV	3.2	5.0	3.22	3.21
Dispersion energy E_d	eV	25	35	27.96	33.28
Absorption coefficient α_0	cm^{-1}	0.1	0.5	0.225	0.186
Urbach energy E_u	eV	0.1	0.3	0.247	0.234

and employs the simple algebraic interference formula to perform the optical characterization. Although the method's approximations can reduce the precision of the estimated film properties, it offers several key advantages: it requires far less computational effort, imposes no dispersion assumptions, and avoids multiple non-physical solutions. In other words, instead of encountering the ambiguity often faced with inverse synthesis methods, the Swanepoel approach yields a unique, physically meaningful set of values for n, k, and d.

Our automated program is based on the original Swanepoel algorithm from 1983 [129], [253], designed specifically for plane-parallel thin films of uniform thickness, composed of isotropic and homogeneous materials. As previously discussed, numerous adaptations of this algorithm have been proposed to handle more complex scenarios [130], [131], [143], [221]–[223], and these enhancements could potentially be integrated into our software framework in the future.

In spectral regions where the sample presents medium-to-weak optical absorption, the condition $n^2 \gg k^2$ holds. These spectral regions, frequently examined in transmission spectroscopy studies, correspond to photon energies below the optical bandgap. Un-

der this circumstance, Eq. 3.108 can be simplified as follows:

$$T^{\text{sim}}(\lambda; n, k, d) = \frac{Ax}{B - Cx\cos\varphi + Dx^2}, \tag{3.113}$$

Observe that the fraction of transmitted intensity x and the phase ϕ were previously defined in Eq. 3.109b and 3.109c, respectively. Additionally, we use the following new expressions:

$$\begin{aligned} A &= 16n^2 n_s, \\ B &= (n+1)^3(n+n_s^2), \\ C &= 2(n^2-1)(n^2-n_s^2), \\ D &= (n-1)^3(n-n_s^2). \end{aligned} \tag{3.114}$$

In contrast with the more accurate transmittance expression from Eq. 3.108, the formula for T^{sim} given in Eq. 3.113 now only comprises a single sinusoidal function. Therefore, we can find the two envelopes of the oscillating transmission curve by assigning $\cos(\varphi) = 1$ for the upper envelope T_M^{sim}, and $\cos(\varphi) = -1$ for the lower envelope T_m^{sim}. These envelopes are then expressed as follows:

$$T_M^{\text{sim}}(\lambda; n, k, d) = \frac{Ax}{B - Cx + Dx^2}, \tag{3.115a}$$

$$T_m^{\text{sim}}(\lambda; n, k, d) = \frac{Ax}{B + Cx + Dx^2}. \tag{3.115b}$$

The intersection points between the simulated transmission spectrum T^{sim} and its upper envelope T_M^{sim} are identified as the upper tangent points, which are represented by λ^{M}, as shown in Fig. 3.9b. Similarly, λ^{m} denotes the lower tangent points, which are as-

sociated with the intersections of T^{sim} with its lower envelope T_m^{sim}. Due to the sinusoidal characteristic of Fabry-Perot oscillations, these upper and lower tangent points occur in an alternating pattern. The sorted sequence of these tangent points is expressed here as λ^{tan}.

While determining the simulated envelopes of a simulated transmission curve is straightforward, following their definition from Eq. 3.115 (as done for the initial Fig. 3.9), identifying the Swanepoel envelopes of an *experimental* curve poses a significant challenge. The Swanepoel method assumes that the true envelopes, T_M and T_m , of a measured transmission spectrum T have already been (somehow) accurately determined. Based on this assumption, the optical properties n and k are extracted using a series of physically meaningful equations, evaluated exclusively at the tangent points. Below, we detail our automated step-by-step procedure for implementing the Swanepoel method, which is summarized in Fig. 3.12.

3.2.3.1 Preliminary estimation of n_1 and d_1

In this preliminary step, we carry out a rough estimation of the refractive index (n_1) and film thickness (d_1). One can obtain a formula for the refractive index using Eq. 3.113 and Eqs. 3.115a-3.115b, leading to

$$\begin{aligned} n_1 &= \sqrt{N + \sqrt{N^2 - n_s^2}}, \text{ where} \\ N &= 2n_s \frac{T_M - T_m}{T_M T_m} + \frac{n_s^2 + 1}{2} \end{aligned} \tag{3.116}$$

Eq. 3.116 is then evaluated at each tangent point, yielding the series of values $n_1(\lambda^{\text{tan}})$. Tables 3.4-3.5 lists these computed values $n_1(\lambda^{\text{tan}})$ corresponding to the experimental spectra illustrated in Fig. 3.11b. Subsequently, we utilize the fundamental Fabry-Perot inter-

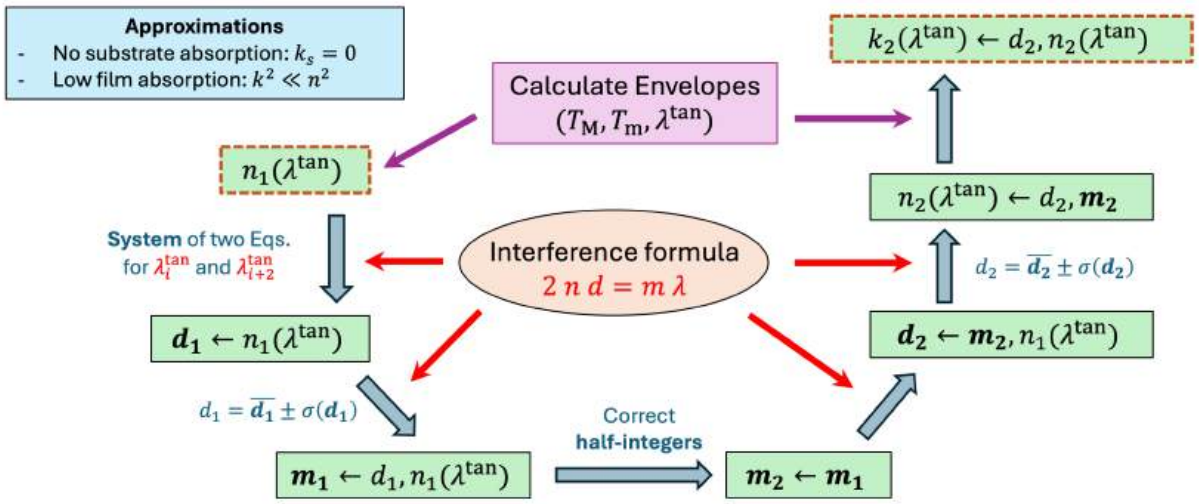


Figure 3.12: Summary of the Swanepoel method: the steps are marked in green, starting at the top left. (1) After computing the transmission envelopes, initial estimates of the refractive index n_1 are determined at the tangent wavelengths λ^{tan}. (2) A system of two equations (based on the interference formula) is solved at each pair of adjacent tangents point to estimate the film thickness, resulting in a set of possible thickness $\mathbf{d_1}$. The expected thickness value is calculated as $d_1 = \text{mean}(\mathbf{d_1}) \pm \text{std}(\mathbf{d_1})$. (3) The interference formula is then applied again to compute the interference orders $\mathbf{m_1}$. (4) These calculated decimal values should theoretically be half-integers, so they are updated to the nearest corrected values $\mathbf{m_2}$. (5) With these corrected interference orders, a more precise thickness estimate $d_2 = \text{mean}(\mathbf{d_2}) \pm \text{std}(\mathbf{d_2})$ is obtained using the interference formula. (6) This refined thickness allows for updated refractive index values n_2 at the tangent points. (7) Finally, an accurate extinction coefficient k_2 is computed at the tangent points using the precise d_2, the updated refractive index n_2, and information from the envelopes.

ference [69] equation:

$$2n(\lambda)d = m\lambda. \tag{3.117}$$

This well-established relation holds only at the *tangent* points ($\lambda = \lambda^{\text{tan}}$), which explicitly denote the transition between constructive and destructive interferences, or vice-versa. The variable m symbolizes the interference order, a number in the set $\mathbb{N}/2 = \{1/2, 1, 3/2, 2, 5/2, ...\}$ and therefore referred to as a "half-integer." Note that the successive tangent points, arranged from longer to shorter wavelengths, correspond to consecutive interference orders. For instance, Tables 3.4-3.5 lists the values of m that range from 3.0

to 10.0 for sample #1, with increments of 0.5.

Two adjacent upper tangent points (or two adjacent lower points), denoted as λ_i^{tan} and $\lambda_{i+2}^{\text{tan}}$, have orders of interference m and $m+1$, respectively. Therefore, we can define the following system of equations:

$$\begin{cases} 2n_1(\lambda_i^{\text{tan}})d_1 = m\lambda_i^{\text{tan}} \\ 2n_1(\lambda_{i+2}^{\text{tan}})d_1 = (m+1)\lambda_{i+2}^{\text{tan}} \end{cases} \tag{3.118}$$

Knowing now λ_i^{tan} and $n_1(\lambda_i^{\text{tan}})$ for each tangent point i, we can solve Eq. 3.118 for the unknown variable d_1. The equation can be solved to find the d_1 values for the different tangent points, except for the two greater ones (as the system needs the points i and $i+2$). The set of values $\mathbf{d_1}$ represents the thickness estimations using data from the multiple tangent points (see Tables 3.4-3.5). Ideally, each of these d_1 values should be identical. In practice, the thickness is approximated as $d_1 = \text{mean}(\mathbf{d_1}) \pm \text{std}(\mathbf{d_1})$.

Sometimes, it is advisable to exclude the elements of $\mathbf{d_1}$ corresponding to the shorter wavelengths if they clearly differ from the the mean value, as they are often not very consistent with the other values (see the marked numbers on the Tables 3.4-3.5). This discrepancy arises when there exists a strong absorption at shorter wavelengths, as the Swanepoel approximation ($n \gg k$) becomes less accurate.

Table 3.4: Results of the Swanepoel method applied to the transmission spectra of **Sample #1.** Estimated thicknesses: $d_1 = 1099.32 \pm 35.77$ nm (relative error = 3.25%), $d_2 = 1110.18 \pm 7.22$ nm (relative error = 0.65%). The last two data points are crossed out due to unreliable calculations at shorter wavelengths.

Tan (nm)	T_m (%)	T_M (%)	n_1	d_1 (nm)	Half int inexact	Half int exact	d_2 (nm)	n_2	k_2 ($\times 10^3$)	α_2 ($\times 10^3$ cm^{-1})	x_2
2340.0	44.35	88.46	3.2049	–	3.0113	3.0	1095.19	3.1616	2.9	0.1550	0.9829
2020.0	43.94	88.28	3.1535	–	3.4324	3.5	1120.97	3.1842	4.2	0.2614	0.9714
1776.0	43.35	87.90	3.1818	1185.00	3.9390	4.0	1116.34	3.1995	4.0	0.2836	0.9690
1588.0	42.56	85.61	3.1940	1110.70	4.4222	4.5	1118.68	3.2184	5.6	0.4458	0.9517
1436.0	41.28	82.01	3.2289	1094.23	4.9437	5.0	1111.85	3.2337	7.9	0.6948	0.9258
1316.0	39.72	77.51	3.2443	1101.41	5.4203	5.5	1115.50	3.2598	11.0	1.0545	0.8895
1216.0	37.90	72.17	3.2718	1130.93	5.9157	6.0	1114.98	3.2859	14.5	1.4990	0.8467
1128.0	35.72	65.99	3.3037	1078.69	6.4394	6.5	1109.67	3.3022	18.6	2.0713	0.7946
1056.0	33.39	59.50	3.3333	1073.15	6.9401	7.0	1108.81	3.3292	23.0	2.7408	0.7377
996.0	30.84	52.76	3.3677	1105.09	7.4342	7.5	1109.06	3.3643	28.0	3.5269	0.6760
944.0	27.89	45.57	3.4117	1092.79	7.9461	8.0	1106.78	3.4012	33.9	4.5095	0.6061
896.0	24.33	37.19	3.4421	1086.14	8.4463	8.5	1106.31	3.4301	42.3	5.9310	0.5177
856.0	20.29	29.20	3.5074	1034.42	9.0088	9.0	1098.25	3.4697	52.3	7.6742	0.4266
820.0	15.99	21.16	3.5336	~~1068.99~~	9.4747	9.5	–	–	–	–	–
788.0	11.62	14.25	3.5865	~~1101.34~~	10.0070	10.0	–	–	–	–	–

Table 3.5: Results of the Swanepoel method applied to the transmission spectra of **Sample #2.** Estimated thicknesses: $d_1 = 1167.28 \pm 62.92$ nm (relative error = 5.39%), $d_2 = 1163.88 \pm 10.52$ nm (relative error = 0.90%). The last two data points are crossed out due to unreliable calculations at shorter wavelengths.

Tan (nm)	T_m (%)	T_M (%)	n_1	d_1 (nm)	Half int inexact	Half int exact	d_2 (nm)	n_2	k_2 ($\times 10^3$)	α_2 ($\times 10^3$ cm^{-1})	x_2
2260.0	39.13	89.72	3.4974	–	3.6128	3.5	1130.84	3.3981	1.3	0.0706	0.9918
1992.0	39.12	88.71	3.4246	–	4.0136	4.0	1163.33	3.4230	3.2	0.2050	0.9764
1776.0	39.09	87.20	3.4086	1345.01	4.4806	4.5	1172.33	3.4333	4.3	0.3011	0.9656
1608.0	38.38	85.11	3.4256	1216.15	4.9734	5.0	1173.53	3.4540	5.6	0.4359	0.9505
1472.0	37.27	81.60	3.4566	1165.49	5.4822	5.5	1171.08	3.4780	7.8	0.6642	0.9256
1356.0	35.93	76.99	3.4765	1153.44	5.9854	6.0	1170.13	3.4952	10.8	0.9974	0.8904
1260.0	34.40	71.62	3.5049	1153.61	6.4940	6.5	1168.36	3.5184	14.1	1.4067	0.8490
1180.0	32.75	65.89	3.5296	1170.04	6.9830	7.0	1170.12	3.5485	17.7	1.8860	0.8029
1108.0	30.85	59.89	3.5632	1151.50	7.5077	7.5	1166.08	3.5699	21.6	2.4504	0.7519
1048.0	28.89	53.84	3.5932	1142.88	8.0044	8.0	1166.64	3.6017	25.7	3.0864	0.6982
992.0	26.54	46.93	3.6257	1138.98	8.5326	8.5	1162.82	3.6224	31.1	3.9394	0.6322
948.0	24.05	40.39	3.6679	1135.25	9.0326	9.0	1163.07	3.6653	36.8	4.8776	0.5668
904.0	20.81	32.28	3.6903	1170.02	9.5302	9.5	1163.58	3.6894	45.8	6.3617	0.4769
868.0	17.24	24.95	3.7659	~~1065.01~~	10.1286	10.0	–	–	–	–	–
836.0	13.55	18.09	3.7984	~~1083.89~~	10.6072	10.5	–	–	–	–	–

3.2.3.2 Refining the values of n_2 and d_2

The previous estimates are highly sensitive to noise from the experimental transmission measurements and errors in the two corresponding envelopes. Therefore, it is important to go through a second refinement step to find more suitable values of n_2 and d_2. First, note that we can now calculate the "inexact" half-integers (denoted by $\mathbf{m_1}$) by inserting the predicted rough values $n_1(\lambda^{\text{tan}})$ and d_1 in Eq. 3.117, yielding $\mathbf{m_1} = 2n_1(\lambda^{\text{tan}})d_1/\lambda^{\text{tan}}$. The set of approximated values $\mathbf{m_1}$ are shown in the 6th column of Tables 3.4-3.5.

We know that $\mathbf{m_1}$ should theoretically be half-integers, so we can improve the found

values. Indeed, one will now search for the sequence of consecutive half-integers $\mathbf{m_2}$ that best match the approximated data $\mathbf{m_1}$. This process can be automated by formulating a fitting problem and employing the direct search method, an effective global optimization technique well-suited for this straightforward, low-dimensional scenario. Specifically, we identify the optimal sequence of consecutive half-integers that minimizes the absolute error distance relative to our set of approximate interference orders m_1. The refined values $\mathbf{m_2}$ are presented in the seventh column of Tables 3.4-3.5.

Setting now $n_1(\lambda^{\text{tan}})$ and the *exact* values of $\mathbf{m_2}$ into Eq. 3.117, we can determine a new array $\mathbf{d_2}$ with more accurate film-thickness estimates: $\mathbf{d_2} = \mathbf{m_2}\lambda^{\text{tan}}/2n_1(\lambda^{\text{tan}})$. The corrected film thickness is then computed as $d_2 = \text{mean}(\mathbf{d_2}) \pm \text{std}(\mathbf{d_2})$. These values are shown in the 8th column of Tables 3.4-3.5. In the two example shown in the table, one can see noticeable improvements from d_1 to d_2, reducing the uncertainty of the thickness value by an order of magnitude.

Finally, one can employ the interference formula once again to compute the corrected refractive index $n_2(\lambda^{\text{tan}})$. Note that the known values that we introduce in the interference formula are the exact set of values $\mathbf{m_2}$ and the highly accurate thickness d_2. In particular, we now compute $n_2(\lambda^{\text{tan}}) = \mathbf{m_2}\lambda^{\text{tan}}/2d_2$.

3.2.3.3 Finding k_2

With the refined values of the refractive index $n_2(\lambda^{\text{tan}})$ and thickness d_2, we now aim to find precise values of the extinction coefficient $k_2(\lambda^{\text{tan}})$ at each tangent point. Observe that since the extinction coefficient k, the absorption coefficient α, and the variable x are directly related through Eqs. 3.109a–3.109b, it is sufficient to estimate any one of these quantities. We will then analyze x.

The variables (A, B, C, D) from Eq. 3.114 depend only on n and d, and can now be accurately calculated using the refined values $n_2(\lambda^{\text{tan}})$ and d_2. Considering that the envelopes T_M and T_m were (somehow) previously calculated, we can now solve either Eqs. 3.115a or 3.115b for $x_2(\lambda^{\text{tan}})$ at each wavelength λ^{tan}. It is important to note that these equations are quadratic, resulting in two possible ambiguous results for x_2. However, it must be remarked again that the parameter x_2 represents a fraction of transmitted intensity: the beam intensity transmitted in a single pass through the thin film, neglecting the multiple inter-reflections in the film and the substrate. Of course, such quantity must lie in the range from 0 and 1, and we can then select the unique physically meaningful solution within this range, out of the two possible mathematical results.

3.2.4 Novel procedure to find the Swanepoel envelopes

The previous section explained how to obtain the optical functions using our automated implementation of the Swanepoel method. Please note that the method did *not* describe how to find the envelopes from an experimental transmission spectrum or their corresponding tangent points, but rather assumed that these envelopes were known *a priori*. We now propose our novel approach to construct the envelopes and solve that initial and particularly challenging problem.

3.2.4.1 Data treatment and convexity regions

Initially, we need to select only the adequate spectral region in which the substrate exhibits negligible absorption ($k_s \approx 0$), and where the thin-layer sample has *non-zero* transmission (excluding areas of clear strong absorption), therefore confining our analysis to the spectral region valid for the Swanepoel method. We then employ a convenient Gaussian filter to reduce noise in the transmission spectrum of the sample. Finally, we in-

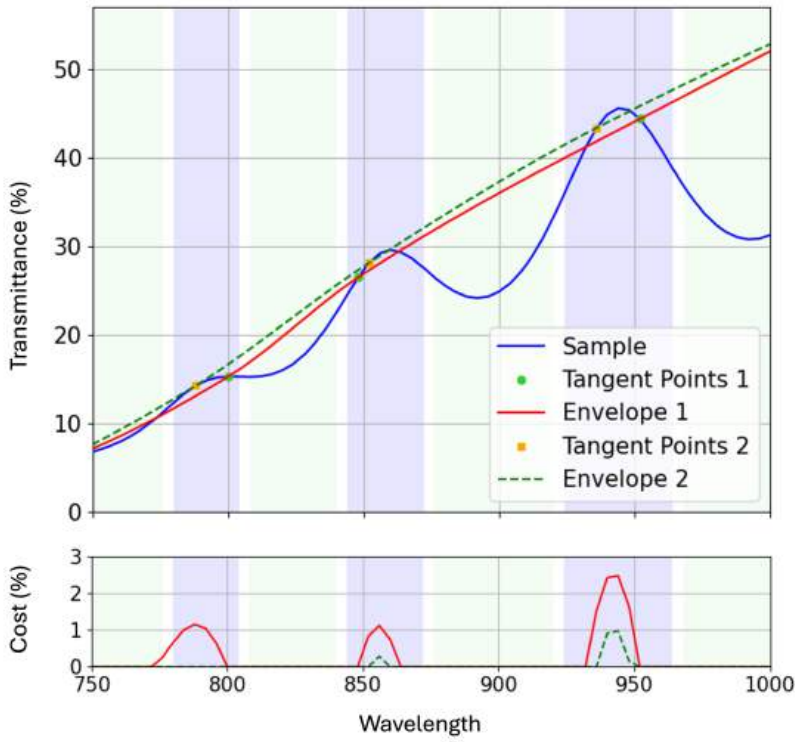


Figure 3.13: Graphical illustration of the iterative optimization process used to determine the upper envelope. The algorithm identifies upper tangent points so that the interpolated curve lies entirely above the experimental transmission data, ensuring that each concave region contains a unique tangent point. During the initial iteration (red envelope), the randomly selected tangent points yield a high cost (see bottom panel). In subsequent iterations (green dashed envelope), the cost gradually decreases, allowing the solution to converge to the desired upper envelope.

terpolate the transmission curve over the full wavelength range with *cubic splines*, which ensure the second-order differentiability of the spectrum. This first step is crucial to find the tangent points with high precision, as the experimental transmittance measurements might have a relatively low spectral resolution (e.g., with measured wavelengths every 10 nm).

Following this initial data treatment, we now identify downward concave intervals within the transmission spectrum by employing the second-order numerical derivatives. The remaining spectral intervals not identified as concave are consequently categorized

as upward convex intervals. Figure 3.13 show the convex regions in light green and the concave ones light blue. We will designate the concave intervals as $[u_t^{\min}, u_t^{\max}]$, for $t \in \{1, ..., U\}$. Similarly, the convex intervals are denoted as $[v_t^{\min}, v_t^{\max}]$, for $t \in \{1, ..., V\}$. The resulting intervals must alternate, ensuring the relation $U = V \pm 1$.

Adopting the reasoning presented in [235], we will only incorporate the intervals at the end (either concave or convex) if they possess a local extremum, given that such intervals could potentially encompass a tangent point. In other words, if last interval is severely cropped, not including its minimum or maximum, it should be discarded. For instance, a short interval at the end of both graphs from Fig. 3.13 appears blank, as it does not contain an extrema. Additionally, the regions in the left-hand side in which $T(\lambda)$ is below a certain numerical threshold (such as 0.5%) are crop out, as no useful information about the tangent point is found in the regions of complete absorption.

In practice, noise in the transmission spectra and small absorption of the substrate typically lead to incorrect interval classification of the convexity. To compensate for this problem, we introduce a "window" variable that establishes the minimum size of the intervals (the highest frequency allowed). Intervals shorter than this window size are considered spurious and are merged with adjacent intervals to form larger, more significant regions. This process effectively filters out inconsistent high-frequency noisy oscillations in the transmission spectrum and its derivatives, ensuring that only meaningful convexity intervals are retained. The selection of this window variable must be customized depending on the unique characteristics of the experimental data and its noise.

3.2.4.2 Finding the tangent points with an optimization algorithm

After these steps, we formulate a global optimization problem that allows us to simultaneously determine both the tangent points and the envelopes. We incorporate into our model two essential properties of the Swanepoel envelopes, detailed below:

- **Envelopes bound the transmission curve**: The upper envelope consistently lies above the transmission curve, $T^{\text{sim}}(\lambda) \leq T_M^{\text{sim}}(\lambda)$, and the lower envelope lies below it, $T_m^{\text{sim}}(\lambda) \leq T^{\text{sim}}(\lambda)$. Consequently, the transmission curve and its envelopes touch at the tangent points but do not cross each other. This behavior follows from the definition of envelopes (see Eq. 3.114) and can be readily proved under the reasonable assumption that $n \geq n_s$, which ensures the variable C in Eq. 3.114 is non-negative.

- **Convexity at tangent points**: Each upper tangent point is located in a concave region of the transmission curve, while each lower tangent point is found within a convex region. To justify this intuitive property, we used symbolic computation software (*Wolfram Alpha Mathematica*) to determine the second derivative of the transmission $T^{\text{sim}}(\lambda)$ at the tangent points. In spectral regions where the film exhibits medium-to-weak absorption, the optical functions $n(\lambda)$ and $k(\lambda)$ vary only slightly. It is reasonable to consider the variables (A, B, C, D, x) from Eq. 3.113 as approximately constant around a specific tangent point. Consequently, variations in the transmission spectrum near that point primarily stem from the term $\cos(\varphi)$ in Eq. 3.113, due to the rapid change of the phase $\varphi = 4\pi nd/\lambda$ with wavelength. This observation allows us to analytically estimate the derivatives of $T^{\text{sim}}(\lambda_{\text{tan}})$. We found that the second derivative is positive at the lower tangent points and negative at the upper ones.

Considering these two reasonable properties for the transmission spectra of thin uniform films, we can now define the following minimization problems:

$$\begin{cases} \text{minimize} \quad \mathcal{C}_M(\lambda_1^M, \lambda_2^M, ..., \lambda_U^M) = \displaystyle\sum_{i=1}^{k} \text{ReLU}\Big[T(\lambda_i) - T_M(\lambda_i; \lambda_1^M, \lambda_2^M, ..., \lambda_U^M)\Big], \\ \text{subject to} \quad \lambda_t^M \in [u_t^{\text{min}}, u_t^{\text{max}}], \text{ for } t \in \{1, 2, ..., U\}. \end{cases} \tag{3.119a}$$

$$\begin{cases} \text{minimize} \quad \mathcal{C}_m(\lambda_1^m, \lambda_2^m, ..., \lambda_V^m) = \displaystyle\sum_{i=1}^{k} \text{ReLU}\Big[T_m(\lambda_i; \lambda_1^m, \lambda_2^m, ..., \lambda_V^m) - T(\lambda_i)\Big], \\ \text{subject to} \quad \lambda_t^m \in [v_t^{\text{min}}, v_t^{\text{max}}], \text{ for } t \in \{1, 2, ..., V\}. \end{cases} \tag{3.119b}$$

Observe that non-uniform films are often modeled using a Taylor first-order approximation with a planar surface exhibiting a certain tilt. A more general expression for the transmission formula in these scenarios is derived in [4]. It should be noted that for films with a very high wedge, $\Delta d > \lambda/(4n)$, the two properties described above do no longer hold. In such cases, the upper and lower envelopes eventually cross over each other, resulting in the upper envelope lying below the spectrum with upper tangent points in convex regions (and vice versa for the lower envelope). Our current study does not apply to this more complex scenario.

The rectified linear unit (ReLU) function, frequently employed in machine learning contexts, is defined as $\text{ReLU}(x) = x$ if $x > 0$, and $\text{ReLU}(x) = 0$ if $x \leq 0$. This function ensures that the cost $\mathcal{C}_M(\lambda_1^M, \lambda_2^M, ..., \lambda_n^M)$ penalizes only the instances when the upper envelope is beneath the transmission spectrum (breaking the first property), while $\mathcal{C}_m(\lambda_1^m, \lambda_2^m, ..., \lambda_u^m)$ penalizes only the cases where the lower envelope is above the spec-

trum. Note that the constraint of the optimization problem guarantees the presence of one (and only one) upper tangent point for each concave region, and one (and only one) lower tangent point for each convex region, as established in the second property.

The predicted top and bottom envelopes, denoted by $T_M(\cdot)$ and $T_m(\cdot)$, are obtained using cubic Hermite interpolation [254] on the tangent points. For instance, the upper envelope corresponds to the interpolation of the points $\{\lambda_t^M, T(\lambda_t^M)\}_t$, for $t \in \{1, 2, ..., U\}$. Once these points are interpolated, we can evaluate the upper $T_M(\lambda)$ envelope function at any wavelength λ within the continuous spectral region, and vice-versa for the lower envelope.

The Hermite interpolation provides a piece-wise polynomial approximation of the envelopes, while maintaining monotonicity within each piece. By utilizing cubic Hermite interpolation, we not only ensure that the interpolated envelopes pass through the given tangent points but also that they conform to the specific slopes at those points. This method incorporates both the function values and their first derivatives, resulting in a piecewise cubic polynomial that is continuous in both value and slope. The inclusion of derivative information allows the interpolation to capture the local behavior of the data more accurately, preserving essential features such as peaks and inflection points. Moreover, cubic Hermite interpolation minimizes oscillations between data points, a common issue with higher-degree polynomial interpolations that do not consider slope information. This reduction in unwanted oscillatory behavior leads to a more stable and reliable approximation of the envelopes.

To solve the optimization problems presented in Eqs. 3.119a-3.119b, we utilize the Simulated (Dual) Annealing algorithm [252]. It should be noted that conventional local optimizers (such as Gradient-Descent, Levenberg-Marquardt, or Nelder-Mead), fail

to converge to the correct solution due to the presence of several local minima. Very significantly, our algorithm converges in 30-60 seconds on a standard computer, given the simplicity of the cost functions. Figs. 3.14a-b show a the results of our envelope-detection algorithm for a simulated sample and a zoom-in of the strong-absorption region. Figs. 3.14c-d show the results of our approach for the real samples #1 and #2, respectively.

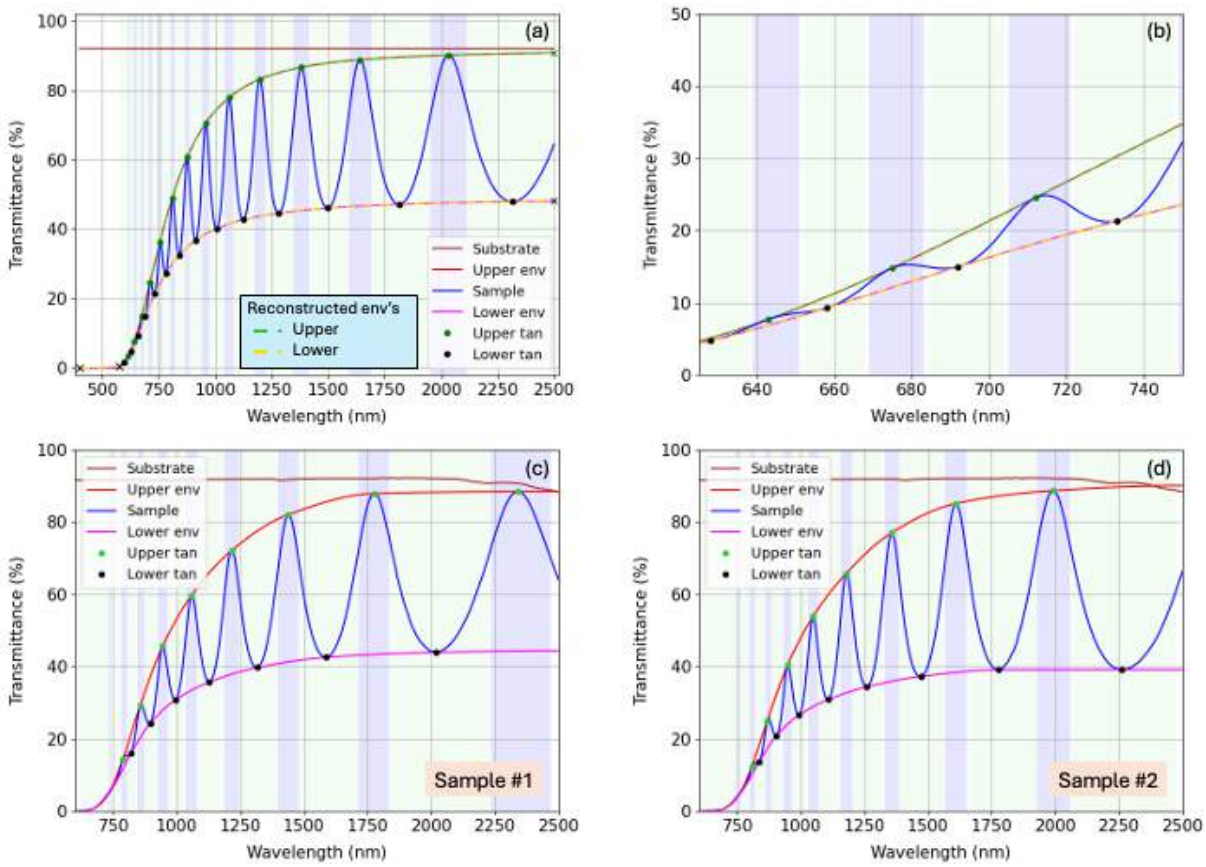


Figure 3.14: Envelopes and tangent points determined using our novel optimization procedure. In particular, (a) shows the reconstructions for simulated sample and (b) a magnified view of its strong absorption region, while (c) and (d) show transmittance curves for the real samples #1 and #2, respectively. These results satisfy the two constraints imposed in our optimization problem: first, the upper (lower) envelope consistently lies above (below) the transmittance curve, touching it only at upper (lower) tangent points; second, these tangent points are localized within the light-blue concave (upper) or light-green convex (lower) regions.

3.2.4.3 Boundary points

Interpolation near the edges of the spectral region of interest (ROI) can be challenging, particularly at the first and last tangent points. To improve accuracy, we propose extending the envelope calculation to include boundary points at both ends of the ROI rather than stopping at the first and last tangent points. This ensures that the first and last tangent points become interior nodes rather than endpoints, improving the stability of the interpolation. The selection of boundary points is informed by the film's behavior at both extremes of the spectrum.

First boundary point(s): At shorter wavelengths, where absorption is strong, the amplitude of the oscillations decrease, and the upper and lower envelopes converge toward the measured transmittance $T(\lambda_{\text{init}}) \approx 0$, indicating that the film is essentially opaque (see Fig. 3.14a below 600 nm). Thus, an additional boundary node $(\lambda_{\text{init}}, T(\lambda_{\text{init}}))$ is placed at the initial wavelength for both the upper and lower envelopes. In cases where the region of near-complete absorption is not fully cropped, one can place this boundary node at the earliest wavelength where $T(\lambda_{\text{init}}) \approx 0$ still holds true. One can see these first two added boundary points in Fig. 3.14a.

Last boundary point: At longer wavelengths, where absorption weakens and the envelopes flatten, we approximate the last boundary point $(\lambda_{\text{last}}, y_{\text{last}})$ using a simple parabolic model based on the *two final tangent points,* denoted (λ_1, y_1) and (λ_2, y_2). In this ending region, the tangent points indeed lie very close to the local extrema, making this approximation convenient in practice. Specifically, we assume the mathematical model:

$$y_{\text{last}} = T_0 + \frac{T_1}{\lambda - \lambda_0}, \tag{3.120}$$

Where the amplitude $T_1 \in [0, 1]$ and wavelength λ_0 are determined analytically from the two last maxima (λ_1, y_1) and (λ_2, y_2) for the upper envelope, or minima for the lower envelope. The parameter T_0 is fix to the mean substrate absorption T_s in the case of the upper envelope: in the weak absorption region, the upper envelope naturally converges to the almost flat curve of the transmittance. In the case of the lower envelope, we calculate the distance between the last maximum and minimum values, which approximates $T_\mathrm{M}^{-1} - T_\mathrm{m}^{-1}$ at the end of the spectrum. We then know that, in the region of weak absorption, the lower envelope must eventually convergent for longer wavelengths to the value $T_0 = T_\mathrm{s} - (T_\mathrm{M}^{-1} - T_\mathrm{m}^{-1})$.

By integrating this physically informed boundary treatment, we ensure a smoother and more reliable envelope estimation across the entire ROI. Note that alternative boundary conditions may be employed for other type of spectra, such as those not reaching near-zero transmittance at the short-wavelength limit, not approaching a medium-to-weak region at the long-wavelength end, or reaching the completely non-absorption region at the long-wavelength end. Our implementation allows users to select different boundary models, such as a flat envelope at the long-wavelength side when complete transparency is reached, exponential and linear fits from the two first or last tangent points, and even manual selection, for more pathological cases.

3.2.5 Other envelope construction techniques

We now present an overview of three envelope construction approaches, including (i) the widely employed McClain iterative algorithm [235], (ii) the practical parabolic or exponential fitting methods, and the recently devised (iii) deep learning technique [7], all of which will serve as benchmarks for evaluating our approach in the subsequent section. The implementation of other methodologies proposed in the literature [239]–[242], for

which no openly available code exists, is beyond the scope of this study.

3.2.5.1 McClain algorithm

The most widely employed and established algorithm for envelope construction was developed by McClain et al. in 1991 [235]. This technique consists of two distinct phases: first estimating the tangent points and then refining them.

In the *first* phase, one identifies the local minima and maxima of the transmission curve, tentatively assigning them as the lower and upper tangent points. Since not all tangent points coincide with peaks and valleys, an alternative approach involves selecting the midpoints of the concave and convex regions. These estimated tangent points are then used to construct the first approximated lower and upper envelopes via B-spline interpolation. The initial estimation of the tangent points during the first phase is the most critical part to find accurate envelopes.

Observe that the algorithm requires the first and last tangent points, located at the boundaries of the interval, to coincide with local extrema (i.e., maxima for the upper envelope and minima for the lower envelope). While this approach helps avoid including invalid or spurious points, as noted in the seminal reference [235], it also leads to drawbacks. For one, it may produce larger errors, as evidenced by the substantial shift from the first extremum to the corresponding tangent point (see Fig. 3.9d). Moreover, it excludes many potentially useful tangent points in the shorter-wavelength region where small oscillations often occur without forming clear local maxima or minima. In broadband transmission measurements, the spectrum typically transitions from strong absorption near the optical bandgap to moderate or weak absorption further away. Early in the spectrum, small oscillations frequently arise without distinct local extrema. By requiring

that the first tangent point be a local extremum, the algorithm effectively discards these non-extremal oscillations, undermining the intended benefit of using the midpoints of concave/convex intervals instead of the peaks/valleys as initial tangent point candidates.

In the *second* phase, the algorithm iteratively refines the tangent points to improve the accuracy of the envelopes. Focusing on the upper envelope, the algorithm identifies "incorrect" regions in which the estimated upper envelope $T_M(\lambda)$ falls below the actual transmission spectrum $T(\lambda)$. The upper tangent points are then relocated to the midpoints of these incorrect regions, as done in the bisection method. A new upper envelope is recalculated based on the updated tangent points. This iterative process is repeated until convergence (or up to a maximum of 20 times), progressively reducing the size of the incorrect regions at each step. Note that the same refinement procedure is applied to the lower tangent points and envelope. In this case, the incorrect regions are those where the estimated lower envelope $T_m(\lambda)$ lies above the transmission spectrum $T(\lambda)$.

Method	Upper env.	Lower env.	Upper tan	Lower tan
McClain	0.202 ± 0.129	0.090 ± 0.040	2.59 ± 1.75	2.48 ± 1.45
Ours	0.080 ± 0.078	0.043 ± 0.034	0.27 ± 0.33	0.55 ± 0.72

Table 3.6: Comparing our proposed method with respect to the standard McClain algorithm within the same spectral region (between the first and last local extrema).

3.2.5.2 Parabolic or exponential fitting

Another prevalent technique involves the direct fitting of local maxima (or minima) using either parabolic or exponential curves to construct the upper (or lower) envelopes. The

envelopes then follow one of these mathematical expressions:

$$f^{\text{para}}(\lambda; T_0, T_1, T_2, \Lambda_1, \Lambda_2) = T_0 + \frac{10^2 \cdot T_1}{\lambda - \Lambda_1} + \frac{10^5 \cdot T_2}{(\lambda - \Lambda_2)^2} \tag{3.121a}$$

$$f^{\text{exp}}(\lambda; T_0, T_1, T_2, \Lambda_1, \Lambda_2) = T_0 + T_1 \cdot e^{-\lambda/\Lambda_1} + T_2 \cdot e^{-\lambda/\Lambda_2} \tag{3.121b}$$

The optimal parameters $(T_0, T_1, T_2, \Lambda_1, \Lambda_2)$ can be obtained by solving a nonlinear least-square fitting problem. Although generally less precise than McClain algorithm, it is crucial to point out that this fitting approach is notably beneficial for transmission spectra that exhibit a restricted number of oscillations and tangent points, a condition often encountered with ultra-thin films. The piece-wise interpolations used in both the McClain algorithm and our newly proposed methods typically require at least four tangent points for a reliable estimation of each envelope.

The fitting method has also proved valid [243] for crystalline film materials, where the materials present a distinct absorption band and are almost transparent beyond the absorption region. These materials have an almost nonexistent Urbach tail, given their low level of molecular disorder. In this cases, the subsequent very flat envelope curves can be conveniently modeled by Eqs. 3.121a-3.121b, independently of the amount of present oscillations in the transmission spectrum.

Our method introduces the artificial boundary points at the ends of the spectral range using a similar approach, but applying first-order parabolic or exponential functions rather than a second-order ones. This adjustment ensures that the fitting procedure is well-defined and depends only on the final two nodes, enabling a direct analytical solution.

3.2.5.3 Deep learning approach

While both McClain and the curve-fitting algorithms offer reasonable outcomes in practice and converge to an envelope of an oscillating function, they do not necessarily converge to the exact expression of the Swanepoel envelopes, as defined in Eqs. 3.115. To overcome this issue, our group developed a deep learning (DL) approach [7] that calculates the envelopes in alignment with the Swanepoel theoretical expressions. In particular, we created a dataset of 100,000 input-output pairs $\{T(\lambda), T_M(\Lambda)\}$ for training using the exact Eqs. 3.113-3.115a. A supervised Convolutional Neural Networks (CNNs) model was trained with this dataset to predict the upper and lower envelopes from a transmission spectrum in real-time. While this DL model does not have limitations on the number of tangent points, it is susceptible to noise and artifacts from real-world transmission measurements. This is caused because the distribution of the simulated data differs significantly, which is a known issue for DL approaches used in inverse problems like image restoration or super-resolution [255]–[258].

The noise in real measures strongly depends on the instrumentation and the conditions of the particular experiment. For instance, one potential sources of deviance between the simulated and real spectra is the photodetector noise, which can be responsible up to the 0.2% transmittance error. Second, the effect of the realistic limited coherent length of the beam in the thin film, which affects the peaks and valleys. Third, the increase of coherent length with the wavelength, which may eventually lead to the high frequency noise due to the undesired thick-substrate interferences. Fourth, the change in the emitting source for UV versus VIS-NIR, which may produces a significant jump in the transmittance. Fifth, the substrate absorption of at particular wavelengths, and the absorption of air components if the experiment is not carefully performed in a vacuum.

Observe that modeling these noise sources in general is challenging and not yet fully addressed in the current literature. Incorporating a comprehensive theoretical noise model could enhance the current DL approach by providing deeper insights and improving its overall performance with real data.

3.2.6 Validation and comparison

We began by validating our envelope construction procedure using 50 computer-generated transmission spectra, each accompanied by a known theoretical envelope. In particular, we generated 50 random set of parameters $(d, E_0, E_\mathrm{d}, \alpha_0, E_\mathrm{u})$ within the bounds specified in Table 3.3. These data were then employed in Eq. 3.113 to produce the transmittance, and Eqs. 3.115a-3.115b for their corresponding envelopes. The substrate transmittance is assumed constant, with the typical value $T_\mathrm{s} = 92\%$.

Accurate envelope determination relies on the inclusion of boundary points. For the first boundary point, corresponding to a regime of nearly complete absorption ($T \approx 0$), the estimation is straightforward in our simulated samples, as the chosen spectral range encompasses the strongly absorbing region near the band gap. In contrast, determining the last boundary point is a more challenging task. Comparison of our default parabolic-fitting model with the known ground-truth last boundary point resulted in a transmittance error of $0.143 \pm 0.214\%$ for the upper boundary point and $0.142 \pm 0.125\%$ for the lower boundary point.

Applying our envelope detection procedure to the simulated spectra yielded an RMSE of $0.098 \pm 0.084\%$ for the upper envelope and $0.058 \pm 0.045\%$ for the lower envelope. The slightly higher RMSE for the upper envelope is attributed to its intrinsically larger values. Notably, *all* tangent points were correctly identified, with an average absolute error

of 1.06 ± 0.95 nm for the upper tangent points and 2.99 ± 2.62 nm for the lower tangent points. Table 3.6 compares our method performance to that of the standard McClain algorithm. Because McClain method identifies only tangent points associated with local extrema, the spectral range in which it computes the envelopes is considerably shorter than the full region of interest. To ensure a fair comparison, we restrict both methods to the same shorter spectral region. Our method achieves less than half the RMSE for the envelopes and an absolute error on tangent point location that is about an order of magnitude smaller.

For a further comparison with curve-fitting approaches, we restrict analysis to spectra with fewer than five tangent points associated with extrema and generate 50 computer-simulated spectra by varying the sample thickness from 200 to 500 nm only, considering ultra-thin film samples. The reported RMSE for the envelopes is computed between the first and last identified tangent points. Under these conditions, the exponential fit yields an error of 2.11% for the upper envelope and 0.31% for the lower envelope, whereas the parabolic fit produces errors of 1.19% (upper) and 0.42% (lower). Because the tangent points in these fits are directly (and somewhat imprecisely) tied to local extrema, their errors increase considerably: 6.10 nm for the exponential fit and 7.66 nm for the parabolic fit. As noted previously, the DL algorithm achieves very low errors (0.14% for the upper and 0.05% for the lower envelopes) over the entire spectral range, but it remains incompatible with realistic noisy data and does not directly provide the locations of the tangent points.

It is important to recognize that the exact optical properties and ground-truth envelopes are obviously unavailable for the case of real spectra, unlike the case with simulated data. To address this limitation, we leverage the results of inverse synthesis as a

benchmark for assessing the optical characterization analysis performed by *swanpy*. This comparison should thus be viewed as a *consistency check* rather than a strict accuracy evaluation.

1. First, Fig.3.15 shows good agreement between the optical functions n and k extracted via inverse synthesis (continuous curves) and those evaluated at specific tangent points by our improved Swanepoel method.

2. Second, the Swanepoel method demonstrates reliable accuracy for determining film thickness. As shown in Tables 3.4-3.5, sample #1 exhibits a relative thickness error of 0.65%, while sample #2 has an error of 0.90%. Notably, the thicknesses from inverse synthesis ($d = 1117$ nm for sample #1 and $d = 1160$ nm for sample #2) both lie within the corresponding Swanepoel confidence intervals ($d = 1110 \pm 7$ nm for sample #1 and $d = 1163 \pm 11$ nm for sample #2).

3. Third, an indirect but informative way to validate envelope detection in real measurements is to analyze the interference orders from Tables 3.4-3.5. We find that the computed orders m lie very close to half-integer values (columns 6 and 7), indicating an accurate determination of tangent points and envelopes. When misidentified tangent points do occur, their associated interference orders tend to align with other half-integer values rather than the expected one.

This consistency across multiple metrics supports the robustness of our approach for real experimental data.

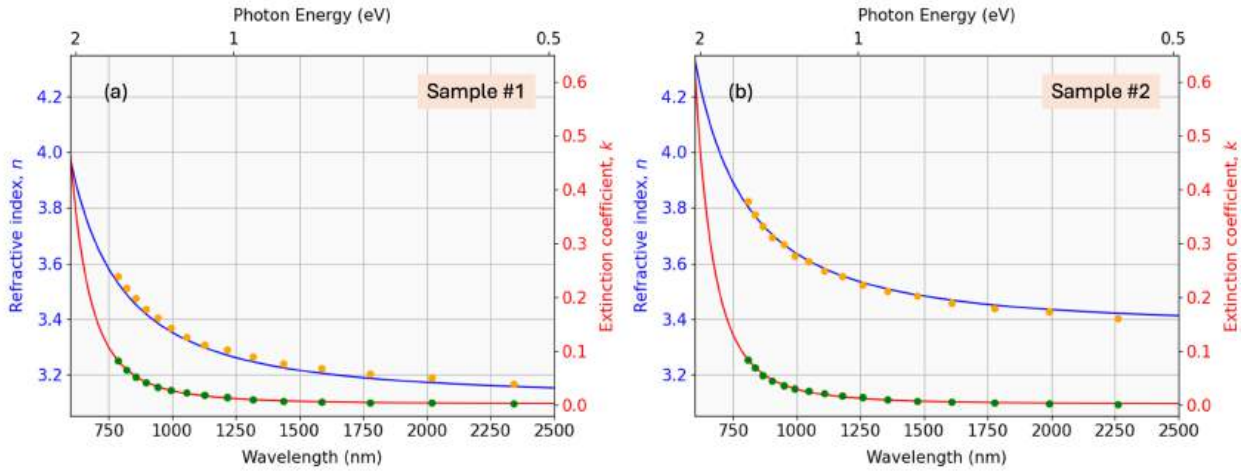


Figure 3.15: Refractive index and extinction coefficient for *a*-Si samples #1 and #2, found with the inverse synthesis approach (solid lines) and with our proposed improved Swanepoel method (circles).

3.2.7 Conclusions

In this work, we present a novel algorithm for Swanepoel envelopes construction, an essential aspect in the optical characterization of thin films. Our approach integrates the physical properties of the Swanepoel envelopes into a global optimization problem. After evaluation with fifty computer-generated spectra, our algorithm demonstrated an RMSE below 0.01% for the upper envelope and 0.06% for the lower envelope, significantly surpassing the performance of existing state-of-the-art methods. We also successfully applied our approach to characterize two real samples of amorphous silicon.

Additionally, we integrated our envelope construction algorithm with our automated implementation of the Swanepoel method in an open-source Python package, *swanpy*. Given its noise robustness, high accuracy, and speed, our program emerges as a valuable tool for the optical characterization of thin solid films. It should be noted that the envelope construction presented in this work can be coupled with more advanced envelope methods. Therefore, future research could explore the applicability of our approach in

more complex and realistic films, such as those with non-uniform thickness, multi-layer samples, or absorbent substrates.

3.3 Machine learning based characterization

3.3.1 Hybrid classical method with deep learning assistance

Dielectric thin films play a crucial role in various applications in modern electronics, including solar cells and OLED thin-film transistors. The dielectric properties of these films depend heavily on the specific material used and the conditions in which they were prepared [100]. Therefore, accurate optical characterization of these films is essential to improving their efficiency. In this work, we propose a novel Deep Learning (DL) approach that enhances the Swanepoel Method (SM), a computational algorithm used to determine thin film properties from their transmission spectra. Specifically, we demonstrate the effectiveness of Convolutional Neural Networks (CNN) in accurately and efficiently calculating the Swanepoel envelopes.

3.3.1.1 Methods

Various experimental techniques can be used to characterize dielectric thin films optically. Among them, transmission spectroscopy stands out as a relatively inexpensive and non-invasive accurate technique [171]. It allows the determination of the thickness d, the refractive index $n(\lambda)$, and the extinction coefficient $k(\lambda)$ of the film through its depth. The film is typically deposited on a thick transparent substrate. This sample is then placed in a spectrophotometer that measures the transmittance, i.e., the ratio of transmitted versus incident light intensity, $T(\lambda) = I_{\mathrm{f}}(\lambda)/I_0$ (see Fig. 3.16a) . Knowing the transmission spectrum of the sample and the substrate alone, one can use two broad categories of methods to carry out the optical characterization: the Inverse Synthesis (IS) approach [183], and the previously mentioned Swanepoel Method (SM) [129]. IS defines a complicated non-convex optimization problem to fit the experimental spectra with an approximate

theoretical curve. On the other hand, SM is based on the interferences generated from the inter-reflections of the coherent beam within the thin film [86]. These interferences strongly depend on the light wavelength and can be observed as oscillations in the transmission spectrum (see Fig. 3.16b). Although IS is considered more accurate in practical applications [148], SM is orders of magnitude faster and does not rely on any theoretical approximations for the dispersion models, being widely used for efficient thin films characterizations.

Swanepoel [129] devised a simple algebraic procedure to find the optical properties of the films using the interference fringe equation [86]. However, this algorithm is only applicable to certain wavelengths called "tangent points", which correspond to the intersections of the transmission spectrum with its upper or lower envelopes (see Fig. 3.16b). The accuracy of this method strongly depends on the precise determination of the envelopes, which has been a challenging problem for decades. Our work presents the first-ever Deep Learning technique to find these envelopes. First, we developed a program that produces computer-generated transmission spectra along with their corresponding envelopes using the formulae from [259]. We generated a large dataset with 1000 pairs $\{T(\lambda), T_{\text{upper}}(\lambda)\}$ to enable supervised learning. We then developed the symmetrical CNN architecture shown in Fig. 3.16c, which inputs the transmittance spectrum and returns its upper envelope. These networks are specifically designed to process structured data, such as spectra or temporal signals, where the sequence of data points is important and each point is related to the previous one [260]. We followed an identical approach for the lower envelopes considering the dataset $\{T(\lambda), T_{\text{lower}}(\lambda)\}$.

3.3.1.2 Results, discussions, and conclusions

Due to the simple architecture, training was completed in less than 20 minutes on a standard computer. We used the MSE loss function and ADAM optimizer with a learning rate of 5×10^{-5} and trained for 2000 epochs in batches of 100. After training the network, we evaluated its performance using a test dataset that had not been seen by the program previously. Our method reached high accuracy, with an average root-mean-square-error (RMSE) between the ground-truth envelopes and the ones generated using the NN model equal to 0.25%. Spectral fits with a RMSE below 1% are generally considered acceptable in the literature [86], [183], [259]. Fig. 3.16b shows an example of one of the predicted envelopes, presenting a RMSE of 0.40 %.

As a qualitative analysis with respect to other conventional techniques, it is worth noting that some of the most common approaches to find the envelopes are based on the idea developed by McClain *et al.* [235] and related approaches [261]. These algorithms first select the local maxima of the spectrum and then use an interpolation curve to find the upper envelope. The curve is then iteratively corrected in several steps until it passes through the desired tangent points and converges to the upper envelope. A similar approach is followed for the lower envelope (from the local minima). These previous direct approaches have two serious problems: (1) The interpolation is significantly less effective when the spectrum has only a few tangent points, and (2) some of the tangent points (typically high photon energies) are not associated or close to the local peaks. Our method does not suffer from these limitations.

Our results demonstrate the potential of a new standardized method for calculating Swanepoel envelopes with high accuracy, leading to improved optical characterization of thin films. However, our project is still in development and has not yet been trained with

noisy or real experimental spectra at this stage. Further work is also needed to conduct a systematic study and quantitatively compare our approach with previous methodologies.

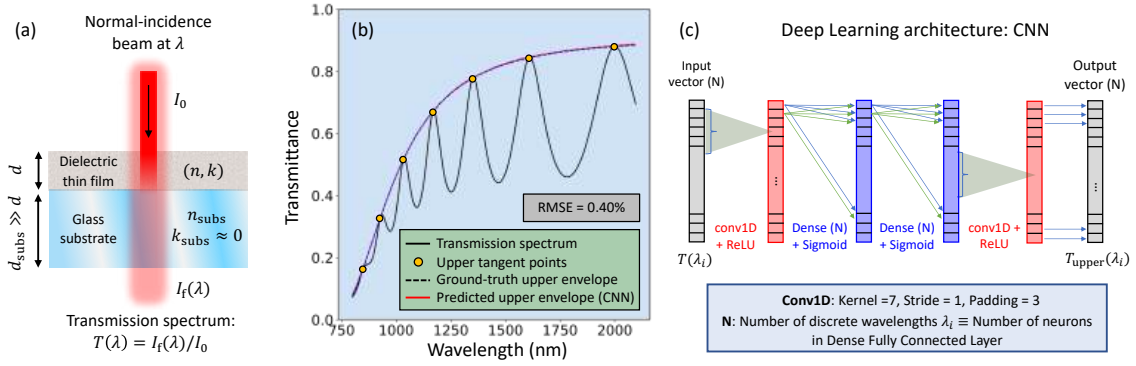


Figure 3.16: (a) Geometrical model of the sample and illustration of the transmission spectroscopy experiment; (b) Simulated transmission spectra with its ground-truth upper envelope and the envelope predicted by the CNN model; (c) CNN architecture.

3.3.2 Direct estimation of optical properties with neural networks

Semiconductor and dielectric thin film materials are essential in many technological industries, including the fabrication of photovoltaic cells, solid-state batteries, and thin-film transistors for LCDs [100]. The optical characterization of such films becomes important for the performance assessment of devices. For example, solar cells will need photodiodes with very high absorption over the large solar spectral range. To evaluate the precise optical properties, one searches for the optical properties $n(\lambda)$ and $k(\lambda)$, which denote the real refractive index and the extinction (or attenuation) coefficient as wavelength-dependent functions. These two functions can be combined to provide the complex relative permittivity of the film, $\epsilon = n + ik$, which contains detailed information on the molecular structure of the material. Information on the reflectance, transmittance, and absorbance of the material can be obtained directly. In addition, electronic transitions, band gap energy, and level of molecular disorder can also be extracted [262]. These are essential features in the semiconductor industry. There are different experimental and

theoretical methods for carrying out the optical characterization of thin-film materials. The material is typically prepared over a transparent glass substrate (see Fig. 3.17a) with a particular deposition technique (e.g. thermal evaporation, sputtering, chemical vapor deposition, and more). Note that the process is performed under precise measurement conditions. For instance, the pressure and temperature in the preparation chamber, the deposition rate, or the crystal growth rate could be measured. Slightly different conditions may lead to different optical properties: Each condition produces a unique material that can be analyzed [91]. Although there are a myriad of experimental techniques for performing optical characterization, this particular study focuses on transmission spectroscopy.

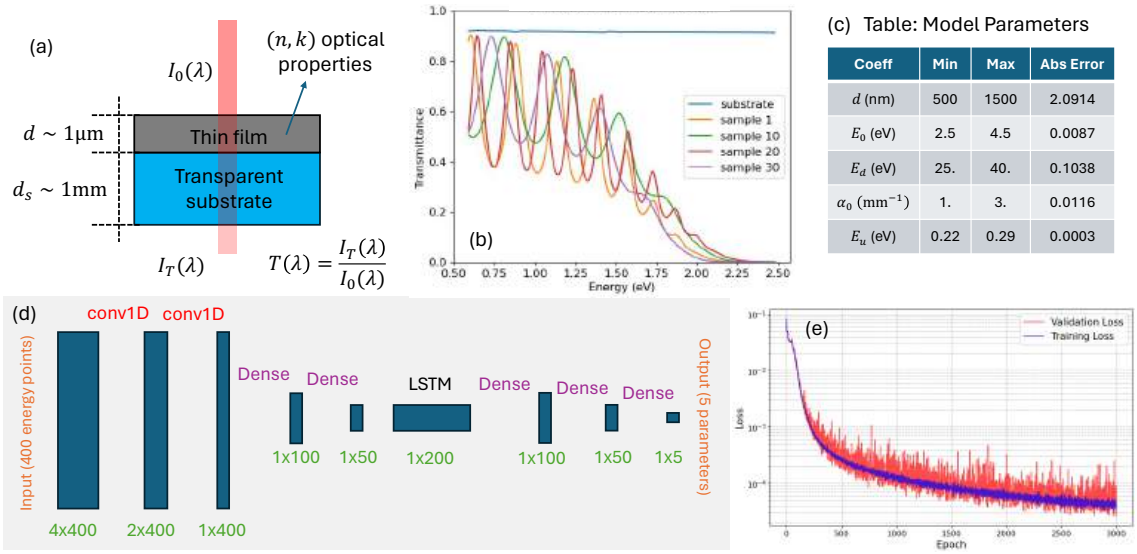
First, the transmittance measurements $\{T(\lambda_i)\}_i$ are captured at discrete wavelengths $\lambda_i \in [\lambda_{\text{min}}, \lambda_{\text{max}}]$, with a step size typically of a few nanometers. Note that transmittance can be expressed equivalently as a function of photon energies E_i. Fig. 3.17b shows the simulated transmittance spectra for some representative samples and for the substrate alone. It is common to take these data within the UV-VIS-NIR range for a comprehensive optical characterization. Then, there are two main computational algorithms to find $n(\lambda_i)$ and $k(\lambda_i)$ at the measured wavelengths λ_i. One of them is the so-called *envelope Method*, a technique strongly pushed by R. Swanepoel [129]. This methodology is highly efficient, allowing characterizations to be performed in just a few milliseconds. However, it imposes some strong approximations (e.g. completely transparent substrates and $k \ll n$) and requires the additional computation of the lower and upper envelopes of the transmission spectrum, which often lead to errors that propagate throughout the algorithm. Recently, our group developed a Deep Learning (DL) method to accurately obtain envelopes, reducing this error [7]. The other main method consists of *fitting* the transmittance data with a theoretical formula. Although this method is highly accurate [148], it can take up to 20-30

min to characterize a single film on a standard computer. In this work, we present the first ever full characterization using a pure DL-based algorithm, with the advantage of highly accurate and instant characterizations once the model is trained.

We have developed a DL architecture (see Fig. 3.17d) that integrates a Convolutional Neural Network (CNN) with a Long-Short-Term Memory (LSTM) module. This integration aims to exploit CNN capabilities for the extraction of local features in the spectral domain, effectively capturing the characteristics of sinusoidal signals associated with transmittance. Simultaneously, the LSTM module is used to extract overarching signal features, using its capacity to retain information throughout the length of the signal. Previous studies have demonstrated the efficacy of this combined methodology in various applications [75]. To generate the datasets, we assumed the Wemple-Didomenico dispersion model [263], expressed as $n(E) = \sqrt{1 + (E_0 E_d)/(E_0^2 - E^2)}$, along with the Urbach equation: $\alpha(E) = \alpha_0 \exp(E/E_u)$, with $k(\lambda) = \alpha(\lambda)\lambda/(4\pi)$ [262]. It is also desired that the thickness of the film d be included in the characterization. We established a reasonable range for these coefficients, as shown in Fig. 3.17c. We generate 100,000 random vectors $\{\Theta_i\}_i$, each containing the five coefficients, within these limits. Then, they were normalized within the interval [0, 1] to serve as the output of our network. For each set of output data, we generate the corresponding input. The transmittance vector $\{T(\lambda_i; \Theta_i)\}_i$ is calculated using the theoretical transmission equations [129]. Note that we additionally consider four channels as input: Substrate transmittance, sample transmittance, and normalized log-magnitude and phase of the sample transmittance. The data set was divided into 60-20-20% between training, validation, and a test dataset (which remained completely unseen by the network). The model training was carried out using the ADAM optimizer for 3000 epochs, with an initial learning rate of 5×10^{-4} and standard exponential decay rates of 0.9 for the first moment and 0.999 for the second moment. Training was

completed with a mean squared error (MSE) of 4.6×10^{-5} for the training set and 4.5×10^{-5} for the validation set, showing that there was no overfitting (Fig. 3.17e shows the complete cost history). Once the network was trained, the predictions with the test dataset revealed an average MSE of 4.2×10^{-5}, indicating a high level of network proficiency. The absolute error for each coefficient is detailed in Fig. 3.17c.

To the best of our knowledge, this is a pioneer demonstration of complete characterization using only transmission data. It is important to note that this research is ongoing and, so far, only simulated data have been analyzed. Future efforts will aim to validate these findings with experimental measurements.



Coeff	Min	Max	Abs Error
d (nm)	500	1500	2.0914
E_0 (eV)	2.5	4.5	0.0087
E_d (eV)	25.	40.	0.1038
α_0 (mm^{-1})	1.	3.	0.0116
E_u (eV)	0.22	0.29	0.0003

Figure 3.17: (a) Sample diagram (b) Simulated transmission spectra for representative samples. (c) Table with parameter boundaries and errors. (d) Model architecture. Operations: 1D convolutions, dense layers, LSTM module. (e) Training and validation cost history.

3.3.3 Physics-Informed Neural Cycles (PINCs)

To enhance the interpretability and physical consistency of the machine learning models used for thin-film characterization via transmission spectroscopy, we incorporated a physics-informed regularization term into the loss function. Specifically, we introduced

a physics-based cost of the form:

$$C_{\text{PB}}(W) = \|\theta - g_{\text{net}}(F(\theta); W)\|_2^2 , \tag{3.122}$$

where $\theta \in \mathbb{R}^n$ denotes the vector of ground-truth material parameters, $F(\theta)$ is a differentiable forward simulator (programmed in Pytorch) that generates the corresponding transmission spectra, based on our analytical transmission expression previously derived. The neural network, denoted as g_{net}, is trained to solve the inverse problem, recovering θ from a given spectrum. This cycle-consistency constraint enforces that the learned inverse model is physically coherent with the forward process, imposing that $g_{\text{net}}(F(\theta); W) \approx \theta$ once the weights W are correctly learned. In other words, this cost penalizes when the composition of the inverse (neural network) and forward (classical) model is not the identity function.

Unlike traditional Physics-Informed Neural Networks (PINNs), which typically incorporate differential equation residuals into the loss function, our approach does not rely on governing PDEs. Instead, we utilize a physically accurate algebraic forward model F, composed of differentiable closed-form expressions. We refer to this approach as **Physics-Informed Neural Cycles (PINCs)**, since it leverages a cyclic structure in the loss to enforce physical fidelity without explicit PDE constraints.

We validated the proposed method using synthetic data generated from the forward model, under both noise-free and noisy conditions (white Gaussian noise with a standard deviation of 0.5%). Several network architectures were tested, including standard multilayer perceptrons (MLPs), convolutional neural networks (CNNs), and deep variational neural networks (DVNNs) [264]. The best performance was obtained using a DVNN model, achieving highly accurate parameter recovery with R^2 scores as follows:

- Thickness: 0.999915
- Wedge angle: 0.998820
- Peak energy E_0: 0.998056
- Dispersion energy E_d: 0.994622
- Absorption coefficient α_0: 0.998093

For the reconstructed transmission spectra, we observed mean relative errors of 0.007% for noise-free inputs and 0.035% for noisy inputs. These results demonstrate that the PINC framework offers significant improvements over purely data-driven models, particularly in enforcing physically plausible solutions in this inverse spectroscopic problem.

Chapter 4

Gamma ray semiconductor detectors

This chapter presents a comprehensive framework for modeling and characterizing the behavior of semiconductor $\gamma-$ray detectors using both physics-based and data-driven techniques. The content is organized into three main sections, each addressing a critical aspect of detector simulation and inverse reconstruction.

The first section introduces a forward physics-based model that serves as a digital twin of the detector. Given a $\gamma-$ray interaction event, defined by the spatial coordinates (x, y, z) of the interaction and the deposited energy E_d, along with the physical properties and potential defects of the detector material, the model simulates the electrical signals produced by the sensor electrodes. This modeling framework captures the full physics of charge generation and transport within the semiconductor material following photon interaction. Unlike traditional models, which simply consider that the charge carriers move under the influence of an external electric field, our formulation also incorporates second-order effects that are often neglected but can significantly impact signal formation. In particular, we model both thermal diffusion and Coulomb repulsion between charge carriers, leading to spatial expansion of the charge cloud during drift. This expansion may increase the likelihood of charge sharing with neighboring pixels, ultimately

resulting in crosstalk and signal distortion.

Building on our forward models, the second section focuses on solving an inverse problem to characterize material inhomogeneities and defects. Using a differentiable implementation of the forward simulator, we apply automatic differentiation and gradient-based optimization to recover depth-dependent variations in detector properties and defects, such as carrier mobility and trapping-detrapping lifetimes. The detector is modeled as a multilayer structure, with each layer parameterized independently, allowing for the recovery of fine-grained internal variations that influence signal behavior. This learning-based approach enables the proper calibration of the forward model to better reflect the actual, and potentially nonuniform, behavior of the semiconductor crystal.

The third and final section addresses two complementary problems: (i) estimating global average material properties of the detector crystal and (ii) reconstructing the physical parameters of individual $\gamma-$ray interaction events. For global material characterization, we employ a global optimization techniques (such as the Slime Mold Algorithm) that uses the temporal signals to infer average detector properties. In parallel, we introduce a convolutional neural network (CNN) architecture trained to reconstruct the 3D interaction position and the deposited energy of $\gamma-$ray events. Once trained, the CNN enables rapid event inference on millisecond timescales, making it suitable for potential real-time applications.

4.1 Novel Physics-Based Model for Detector Simulation

Semiconductor detectors for high-energy sensing (X/γ-rays) play a critical role in fields such as astronomy, particle physics, spectroscopy, medical imaging, and homeland security. The increasing need for precise detector characterization highlights the importance

of developing advanced digital twins, which help optimize the design and performance of imaging systems. Current simulation frameworks primarily focus on modeling electron-hole pair dynamics within the semiconductor bulk after the photon absorption, leading to the current signals at the nearby electrodes. However, most simulations neglect charge diffusion and Coulomb repulsion, which spatially expand the charge cloud during propagation, due to the high complexity they add to the physical models. Although these effects are relatively weak, their inclusion is essential for achieving high-fidelity replication of real detector behavior. There are some existing methods that successfully incorporate these two phenomena with minimal computational cost, developed by Gatti in 1987, and by Benoit and Hamel in 2009. The present work evaluates these two approaches and proposes a novel Monte Carlo technique that offers higher accuracy in exchange for increased computational time. Our new method enables more realistic performance predictions while remaining within practical computational limits.

4.1.1 Introduction

Single-photon semiconductor detectors are widely used in fields such as astronomy, particle physics, spectroscopy, radiology, and homeland security. These detectors are highly effective in precisely identifying both the 3D location of photon-detector interactions and the deposited energy of incoming high-energy photons. These detectors use thick crystals to stop most of the incoming radiation, and they are often made of materials such as Ge, Si, CdTe or CdZnTe, as well as emerging materials like HgI_2, PbI_2 and TlBr [265]–[270].

In recent years, there has been significant progress in the development of accurate *digital twins* that closely replicate the functioning of such detectors [271]–[284]. These simulators enhance our ability to reconstruct the events (3D position and energy) more efficiently. For instance, some of these computational models incorporate spatially de-

pendent crystal properties and defects, providing refined event reconstruction for customized detector configurations and compensating for crystal impurities [10], [285]–[289].

A basic design for these detectors is the planar configuration, featuring a cathode on one side and an anode on the opposite side, as seen in Fig. 4.1a. Nevertheless, it is important to emphasize that our presented model is not influenced by a particular detector type and could be applied to more advanced configurations, such as the common pixelated geometry (that takes advantage of the *small pixel effect* on the anode side) [290] or Frisch grid structures [291].

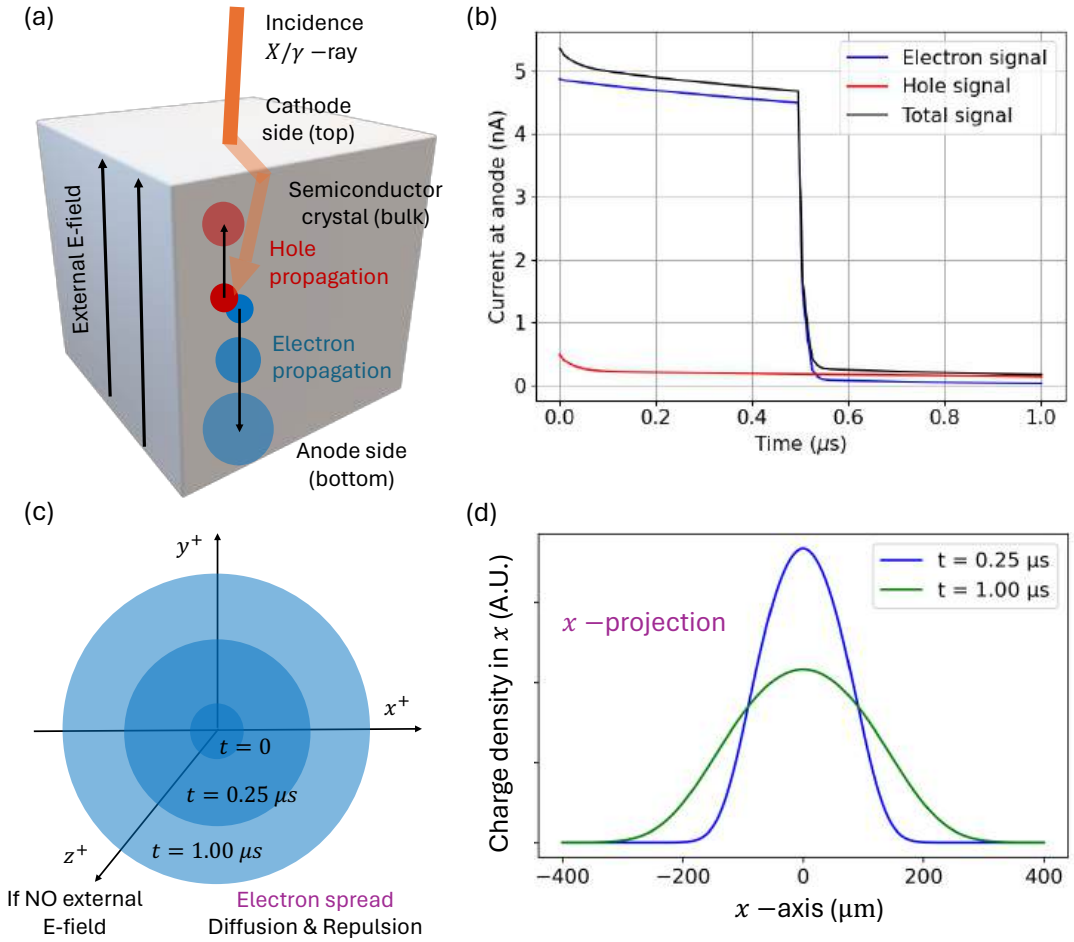


Figure 4.1: (a) Basic planar detector configuration. The absorption of incidence photon creates electron-hole pairs, and the free charges move through the bulk of the detector drift by the external E-field. (b) Simulated signals at the anode. We have used the specifications from [284] for CZT detectors. The simulator includes the charge drift due to the external field and the charge fluctuations (recombination, trapping and detrapping). (c) Spread of electrons in time due to diffusion and repulsion. If we compensate for the transport due to external E-field, the spread appears spherically symmetric around a central point. (d) Projected charge density on the $x-$direction at two different times.

Between the detector electrodes, a bias electric field E_b is established (see Fig. 4.1a). When a photon is absorbed, it undergoes an initial conversion process, such as the photoelectric or Compton effect, which then generates P electron-hole pairs, typically ranging from tens to hundreds of thousands, depending on the deposited energy and the ioniza-

tion energy of the crystal. These free electrons and holes are then driven toward their corresponding electrodes by the bias E-field: the holes migrate toward the cathode along the field lines, while the electrons are pushed toward the anode. The digital twins must accurately replicate how these electrons and holes move as they are drifted through the detector's bulk. Understanding the temporal and spatial propagation of the charges enables the accurate reconstruction of the current signals at the electrodes. In particular, the so-called Shockley-Ramo theorem explains how to emulate the detector signals knowing the movement of the free electron and holes [292]–[296].

Fig. 4.1b displays the simulated current signals at the anode location produced with a modern simulator [284]. For this visual example, we consider a representative planar CZT detector with the reasonable dimensions of $1 \times 1 \times 1$ cm^3. Note that photon interaction event here occurred right at the center of the cube, at the coordinates $(x_0, y_0, z_0) = (0.5, 0.5, 0.5)$ cm of the detector, and the presumed deposited energy was 140.5 keV, corresponding to the gamma photons typically used in medical SPECT applications and emitted by a source of Technetium-99m. To simplify the model, we assume a strong bias E-field, approximated as a uniform constant field of $E_\mathrm{b} = 1000$ V/cm throughout the crystal, created by the applied linear voltage induced at the electrodes (see Fig. 4.1a) [297]–[299]. While this planar detector configuration and constant-field assumption serve as a practical scenario for demonstration, they are not limitations of our method for analyzing diffusion and repulsion processes, which can be extended to more complex detector geometries [300] and variable E-fields [301]–[303].

In this example, the free electrons generated within the CZT crystal reach the anode more quickly than the holes reach the cathode. This is due to the significantly higher mobility of electrons, which is approximately 1000 cm^2/Vs, compared to the mobility of

holes, which ranges between 10 and 100 cm^2/Vs [304]. Note that the amount of charges fluctuates with time and decreases as a result of the trap-induced recombination [305], [306], leading to a slight decay in the signals over time (see Fig. 4.1b). One can additionally see a drop in the electron signal right after the electron collection at 0.5 μs[1]. Observe in Fig. 4.1b that the electrons contribute more than the holes to the total anode signal, which happens because of their higher drift velocity and their increasing proximity to the anode over time. It must be emphasized that another similar signal, although with a negative sign, could be read at the cathode location (not shown in the figure).

However, many of the simulators from the literature [271], [272], [278], [283], [284], [289], [307], fully overlook the effects of charge diffusion and Coulomb repulsion. They rather consider the cloud as a point-charge that moves following the *one-dimensional* paths of the bias E-field. Other simulators omit only the self-repulsion effect [277], [280], as this effect is particularly demanding from the computationally point of view [308]. While we should acknowledge that diffusion and repulsion are second-order effects that often have a relatively low impact on the current signal, they are responsible for spreading the charge cloud in *three-dimensional* space as it traverses the crystal, as shown in Fig. 4.1c.

Many simulators exclude these two effects not because they are insignificant, but due to the increased model complexity and computational time associated with their inclusion. In many applications, neglecting them still yields acceptable approximations. However, as new-generation detectors demand higher energy and spatial resolution, these subtle effects are becoming increasingly important in refined simulations [282]. Quantifying the specific contribution of diffusion and repulsion is challenging, as their impact

[1]The electron drift velocity can be estimated as the product of the charge mobility with the bias E-field, $v = (10^3 cm^2/Vs)(10^3 V/cm) = 10^6$ cm/s in the simulated scenario plot in Fig. 4.1b. This implies that electrons travel from the midpoint of the detector (at $z = 0.5$ cm) to the anode (at $z = 0$ cm) in 0.5 μs, as seen in Figs. 4.1a-b.

strongly depends on numerous factors, including the detector geometry, the semiconductor material, the energy of incident photons, the location of the interaction, and the operation conditions (such as temperature). In any case, to highlight the importance that these two effects have in a real scenarios, consider an event occurring near or over the interpixel gap between two anodes. In this situation, a basic model that treats the charge injection as a delta-distribution could lead to complete current leakage or their collection at only one anode pixel. In contrast, accounting for the expansion of the charge cloud (which can grow to hundreds of microns) could result in the realistic phenomenon of shared collection of the electron cloud at the adjacent anodes, significantly changing the signals shape.

Our present study analyzes these second-order effects in depth. First, we will evaluate two prevalent models previously employed to integrate these effects effectively: *(i)* the analytical study developed by Gatti *et al.* [309] in 1987, and *(ii)* the stochastic method devised by Benoit and Hamel [310] in 2009. Gatti's model treated diffusion and Coulomb repulsion as separate phenomena, each modeled by a distinct partial differential equation. This work has influenced many subsequent relevant analyses [311]–[315].

Built upon Gatti's results, the more recently developed Benoit-Hamel (BH) method consolidates both the diffusion and repulsion effects together into a single mathematical model. This approach assumes that the charge distribution maintains a simple three-dimensional Gaussian shape over time, with a mean located at the center of the cloud and an evolving standard deviation that slowly increases. Our proposed approach further extends the BH model and proposes that the charges spread out following a Generalized Normal Distribution (GND). This distribution offers greater flexibility than the simple Gaussian model and allows us to accurately match the shape of the real intricate

charge spread.

4.1.2 Theoretical Background

For the moment, we will focus solely on the diffusion and repulsion of charges, neglecting *(i)* the charge drift due to the external E-field, and *(ii)* the charge fluctuation due to additional pair generations, recombination, or trapping-detrapping processes. Observe that it is possible to adopt a local coordinate system centered on the electron cloud, which moves at the electron drift velocity [311]. This choice of reference frame can simplify the analysis by considering the electron cloud as stationary relative to its surrounding environment (see Fig. 4.1c-d). The core of our analysis then becomes the well-known diffusion-repulsion equation, formulated as:

$$\partial_t n = D\nabla^2 n - \nabla \cdot (\mu n E) \tag{4.1}$$

Here, μ represents the mobility of the charge carriers, D denotes the Einstein diffusion coefficient, and E is the field generated by the cloud itself. Importantly, D is not an independent parameter, as it can be derived from the semiconductor temperature T and the carrier mobility. Following the Einstein–Smoluchowski relation [316], $D = \mu k_{\mathrm{B}} T/q$, where k_{B} is the Boltzmann constant and q is the elementary charge. Equation 4.1 thus describes the time- and space-dependent evolution of the electron concentration $n(x, y, z, t)$. A similar equation and approach are applicable to holes, considering their corresponding mobility and diffusion coefficient. According to Gauss's law, the repulsive Coulomb field

generated by the charges within the cloud[2] is expressed as

$$\nabla \cdot E = \frac{\rho}{\epsilon} \tag{4.2}$$

Here $\rho = q\,n$ is the electron density (C/cm^3) and ϵ is the permittivity of the crystal. We then aim to solve a system with two coupled partial differential equations (PDE), with Eqs. 4.1 and 4.2.

Both the diffusion process and Coulomb repulsion inherently lack directional preference. When the P generated electron-hole pairs are located deep within the central bulk of the detector, they are minimally affected by boundary effects and it is expected that they will maintain radial uniformity [309]. The characteristic broadening of the distribution is then expected as the concentration evolves in time (see Fig. 4.1c-d). We can now rewrite the equations and variables using spherical coordinates (r, θ, ϕ). If we multiply both sides of Eq. 4.1 by q and perform the 3D integral over a ball with radius R, denoted by the open set $\Omega = B(0, R) \subset \mathbb{R}^3$, we obtain the following result:

$$\int_\Omega q\,\partial_t n\,dV = \int_\Omega q\left(\frac{D}{r^2}\partial_r\big(r^2\partial_r n\big) - \mu\nabla\cdot(nE)\right)dV \tag{4.3}$$

The left-hand side of Eq. 4.3 can then be reformulated as $\partial_t Q$, using the total charge $Q(R, t)$ contained in a volume. The right-hand side of Eq. 4.3, however, presents a more complicated integrand composed of two parts, one responsible for diffusion and the other

[2]The self-repulsion field of the cloud should not be confused with the externally applied bias electric field E_b between the electrodes, as shown in Fig. 4.1a.

for repulsion. Integrating the diffusion part, one gets the following expression:

$$\int_{\Omega} q\left(\frac{D}{r^2}\partial_r\left(r^2\partial_r n\right)\right)dV = D\left(\partial_{RR}\,Q - \frac{2\partial_R\,Q}{R}\right) \tag{4.4}$$

The repulsion part on the right-hand side of Eq. 4.3 can be simplified using the Gauss divergence theorem [317],

$$\int_{\Omega} q\Big(\mu\nabla\cdot(nE)\Big)dV = \mu\frac{\partial_R\,Q}{4\pi R^2}\frac{Q}{\epsilon} \tag{4.5}$$

Note that the Coulomb field E was calculated, as a function of the total charge Q, by solving analytically the Gauss law [318] over the conveniently selected ball Ω. We have finally derived a simplified reformulation of the initial coupled partial differential equations from Eqs. 4.1 and 4.2 as

$$\partial_t Q = D\left(\partial_{RR}\,Q - \frac{2\partial_R\,Q}{R}\right) - \mu\frac{\partial_R\,Q}{4\pi R^2}\frac{Q}{\epsilon} \tag{4.6}$$

Eq. 4.6 does not have a closed-form analytical solution given a general initial condition $Q(R, t = 0)$. However, we can solve it using numerical tools. Despite its apparent simplicity, special care should be taken due to the nonlinear behavior. We have employed a customized implicit finite difference Method (IFDM) [319] and solved the resulting nonlinear discrete equations (at each time step) with the Newton-Raphson method. More details are provided in Appendix A. We analyze the propagation of charges in a time range of $(0, 1]$ µs, with $T = 200$ time steps, and a radial space interval of $(0, 400]$ µm, where $N = 400$ represents the spatial steps.

While the current numerical approach successfully provides the desired solution, it presents two significant challenges. First, it is important to note that standard computers

require approximately 3 to 7 seconds to compute this solution. Although this might not appear excessive at first glance, efficient simulators are typically designed to operate within a timeframe of tens of milliseconds, as they often must be evaluated thousands of times to address inverse problems. Second, these digital twins do not only account for diffusion and repulsion, but also for the charge drift due to the external field and the charge fluctuations, and incorporating the numerical solutions for diffusion-repulsion into a generally complicated simulator framework is not straightforward.

As a result, research efforts focus on developing approximate solutions that are computationally efficient and can be easily integrated into broader simulation frameworks. In this work, we will utilize the precise numerical results from the IFDM method solely for comparison purposes, allowing us to evaluate the accuracy of two commonly used efficient approximations (the Gatti and BH methods) alongside our novel approach.

4.1.3 Gatti Model: Decoupling Processes

4.1.3.1 Diffusion

Gatti *et al.* [309] analytically examined diffusion and repulsion as independent processes. In particular, note that we can extract the diffusion part from Eq. 4.6 as

$$\partial_t Q^{\text{diff}} = D\left(\partial_{RR} Q^{\text{diff}} - \frac{2\partial_R Q^{\text{diff}}}{R}\right) \tag{4.7}$$

The associated diffusion equation for charge concentrations then becomes

$$\partial_t n^{\text{diff}} = D\nabla^2 n^{\text{diff}} \tag{4.8}$$

For a initial delta distribution, $n(r,0) = \delta(r)$, Eq. 4.8 has a well-known solution:

$$n^{\text{diff}}(r,t) = \frac{N}{(4\pi Dt)^{3/2}} \exp\left(-\frac{r^2}{4Dt}\right) \tag{4.9}$$

In other words, the charge concentration propagates following a 3D Gaussian shape with variance that progressively increases in time, as $\sigma(t)^2 = 2Dt$. If we consider instead an initial Gaussian distribution at the initial time, $n(r,0) \sim N(0,\sigma_0)$, the same propagation shape is still valid. In that case, the concentration becomes

$$n^{\text{diff}}(r,t) = \frac{N}{[2\pi(\sigma_0^2 + 2Dt)]^{3/2}} \exp\left(-\frac{r^2}{2(\sigma_0^2 + 2Dt)}\right) \tag{4.10}$$

We can then calculate by integration [309] the solution of Eq. 4.7, the initial diffusion equation for total charges.

4.1.3.2 Repulsion

On the other hand, the repulsion part is extracted from Eq. 4.1 as

$$\partial_t Q^{\text{rep}} = -\mu \frac{\partial_R Q^{\text{rep}}}{4\pi R^2} \frac{Q^{\text{rep}}}{\epsilon} \tag{4.11}$$

Using the separation of variables method[3], Gatti *et al.* [309] found the following solution:

$$Q^{\text{rep}}(R,t) = \frac{R^3}{3\left(\frac{\mu}{4\pi\epsilon}\right)t}\big[H(R) - H(R - R_0(t))\big] + qN\, H(R - R_0(t)) \tag{4.12}$$

[3]An alternative approach to solving the repulsion-only equation can be found in terms of charge density [310], $\partial_t \rho^{\text{rep}} = -\nabla \cdot (\mu \rho^{\text{rep}} E)$, instead of using Eq. 4.11 for total charge. From the numerical result shown in Fig. 4.2h, it is evident that the charge density $\rho(r,t)$ must be uniform, which we adopt as an ansatz. Thus, $\partial_t \rho^{\text{rep}} = -\mu\big[E \cdot \nabla \rho^{\text{rep}} + \rho^{\text{rep}} \nabla \cdot E\big] = -\mu \rho^{\text{rep}}\, \nabla \cdot E = -(\rho^{\text{rep}})^2 \mu/\epsilon$. Here, we have used the Gauss's law from Eq. 4.2 and the identity $\nabla \rho^{\text{rep}} = 0$, which derived from the ansatz. This allows us to solve the ODE directly, yielding the same solution as in Eq. 4.15

Where the time-dependent function $R_0(t)$ is defined as

$$R_0(t) = \left[3\left(\frac{\mu}{4\pi\epsilon}\right) t\, qN\right]^{1/3} \tag{4.13}$$

Note that beyond the radius $R_0(t)$, the charge distribution collects all the charges (qN coulombs). Here, we have used the Heaviside step function,

$$H(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases} \tag{4.14}$$

Observe that, in the limit when $t \to 0$, we get $R_0(t) = 0$, and the total charge becomes $Q^{\text{rep}}(R, 0) = qN$ at $R = 0$ and zero elsewhere. This corresponds to the delta distribution for the initial condition of the charge density, $\rho(r, 0) = q\, N\, \delta(r)$. From Eq. 4.12, we can calculate the charge concentration for $t > 0$ using partial derivatives as

$$\rho^{\text{rep}}(r, t) = \begin{cases} \frac{\epsilon}{\mu t} & \text{if } r < R_0(t), \\ 0 & \text{if } r \geq R_0(t). \end{cases} \tag{4.15}$$

In conclusion, at any given time, the charge density remains constant within the radius R_0 and zero beyond it. Thus, in contrast to the 3D Gaussian spread resulting from diffusion, repulsion cause the charges to propagate as a 3D uniform distribution.

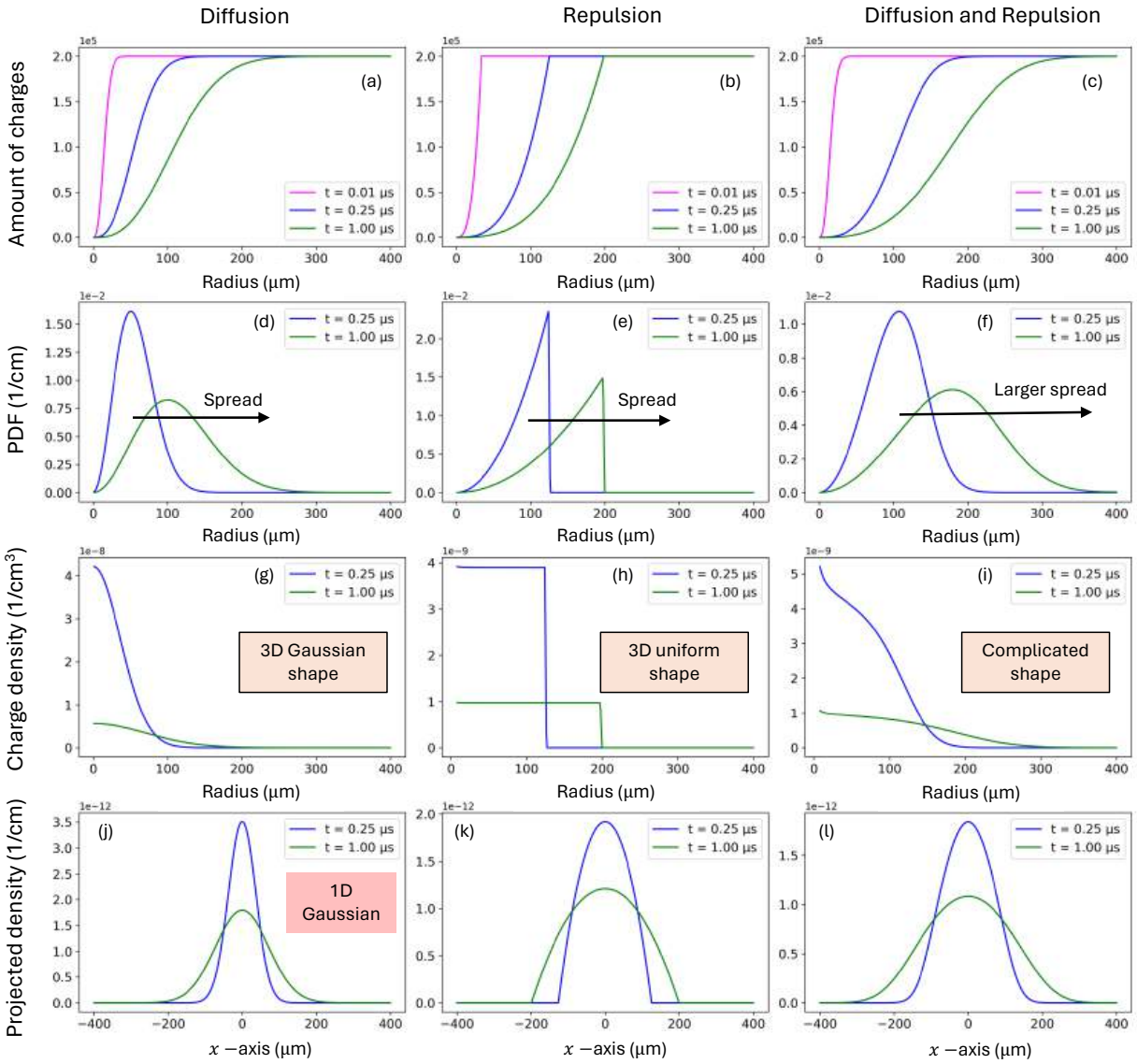


Figure 4.2: Numerical solutions from the IFDM for diffusion-only, repulsion-only, and diffusion-repulsion combined. Further description provided in Section 3.3.

4.1.3.3 Root Mean Square Metric

Figure 4.2 shows the ground-truth numerical results for diffusion-only, repulsion-only, and diffusion-repulsion, which were found using the aforementioned IFDM approach. We used the common values for electrons in CZT crystals [284] at room temperature, with

$\mu = 1000$ cm/Vs and $D = 25$ cm^2/s. There are four common alternative to present the results, which are shown in the different rows of Fig. 4.2:

- The number of elementary charges, $N(R)$, within a given radial distance R from the center of the charge cloud, as displayed in Figs. 4.2a-c. For R values larger than the charge cloud, $N(R)$ will approximate the total number of generated electron-hole pairs (in this representative case, $P = 2 \times 10^5$ charges). Alternatively, this variable can be expressed as the total Coulomb charge, $Q(R) = qN(R)$, or as the normalized charge distribution, $\hat{Q}(R) = Q(R)/qP$.

- The probability density function (PDF) of the charge distribution, $f(r,t) = \partial_R \hat{Q}(R,t)$, shown in Figs. 4.2d-f.

- The charge densities $\rho(r,t) = \partial_R Q(R)/4\pi R^2$, expressed in C/cm^3, as displayed in Figs. 4.2g-i.

- The charge densities projected over the $x-$coordinate (due to symmetry, this projection is identical for all coordinates) become

$$\rho_x(x,t) = \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz \rho(x,y,z,t) = \int_{|x|}^{\infty} dr \int_0^{2\pi} d\phi\, \rho(r,\theta,\phi,t) r \sin(\theta) = \int_{|x|}^{\infty} \rho(r,t) 2\pi r dr \tag{4.16}$$

  Where one needs to consider the spherical symmetry of the density and $\theta = \pi/2$.

We previously mentioned that the density $\rho^{\text{diff}}(r,t) = q\, n^{\text{diff}}(r,t)$ for the diffusion-only process is characterized by the shape of a 3D Gaussian, as derived by Gatti and presented in Eq. 4.9. A fundamental principle in statistics [320] states that the marginal distribution of a multivariate normal distribution is also normally distributed. Thus, $\rho_x^{\text{diff}}(x,t)$

is a 1D Gaussian (with the same mean and standard deviation), as observed in Fig. 4.2j. This property will play an important role in the subsequent Benoit-Hamel approach.

Note that so far our analysis has focused on the diffusion and repulsion of the charges distributed around a central point, the origin. To measure the spread of this distribution, we then consider the standard deviation of $\rho_x(x,t)$ around the zero mean, which provides the root mean squared:

$$\text{RMS}(t) = \int_{\infty}^{\infty} x^2 \rho_x(x,t) dx \tag{4.17}$$

Figure 4.3 indicates that the effect of repulsion on the cloud spreading is slightly stronger than that of diffusion. This behavior is influenced by the initial generation of electron-hole pairs, where diffusion dominates when the deposited energy is lower, while repulsion becomes the predominant factor as the deposited energy increases.

One can calculate the quadratic sum of the RMS for diffusion-only and repulsion-only to obtain a *rough* estimate of the RMS for the combined effects of diffusion and repulsion, as see Fig. 4.3. Note that by decoupling both processes, we can provide an approximate efficient solution that no longer requires to solve PDEs, as we could directly use the analytical solutions.

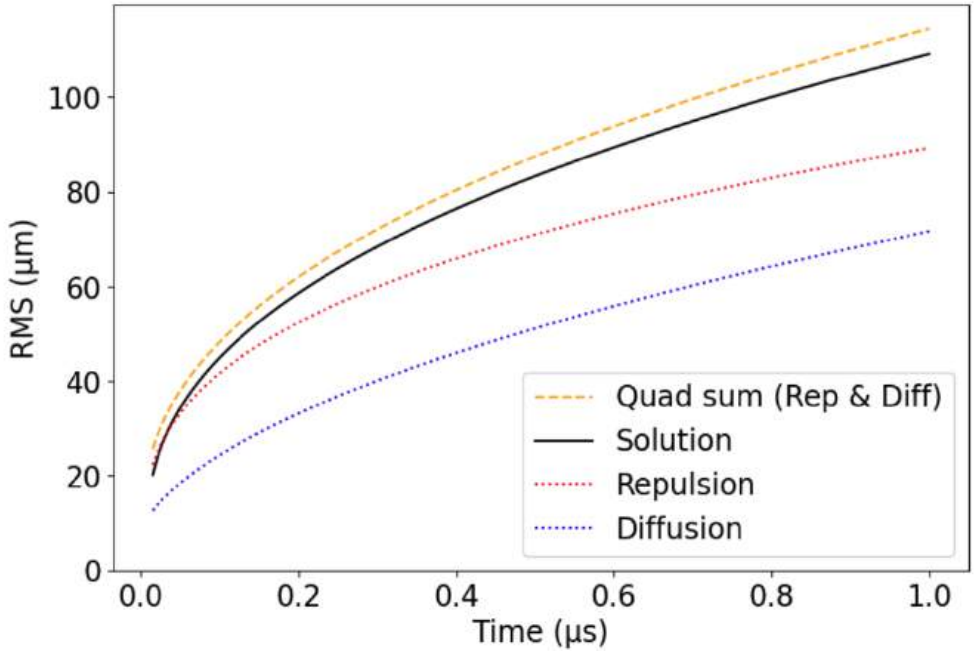


Figure 4.3: Root mean square for diffusion-only, repulsion-only, and diffusion-repulsion combined. We can see that the quadratic sum of diffusion-only and repulsion-only differs from the actual spread of the diffusion-repulsion processes. Parameters: $\mu = 1000$ cm/Vs, $D = 25$ cm^2/s, $P = 2 \times 10^5$ electrons, $\epsilon_r = 11$ relative permittivity, initial standard deviation $\sigma_0 = 10$ µm.

4.1.4 BH Model: Gaussian Distribution

4.1.4.1 Analytical Derivations

Inspired by Gatti's model, Benoit and Hamel (BH) [310] proposed an alternative approximate solution for the original system of Eqs. 4.1 and 4.2, but now effectively coupling both the diffusion and repulsion processes. They started with the initial assumption that the intricate projected charge density along the $x-$axis, resulting from the combined effects of diffusion and repulsion (refer to Fig. 4.2l), can be approximated by a simple one-dimensional Gaussian density. Consequently, the following formula applies:

$$\rho_x^{\mathrm{BH}}(r,t) = \frac{Nq}{[2\pi\sigma_x^{\mathrm{BH}}(t)^2]^{1/2}} \exp\left(-\frac{r^2}{2\sigma_x^{\mathrm{BH}}(t)^2}\right) \tag{4.18}$$

Where $\sigma_x^{\text{BH}} = \sigma_x^{\text{diff}} + \sigma_x^{\text{rep}}$ is an unknown function responsible for the spread in time. Note that this assumption is equivalent to saying that the three-dimensional charge density $\rho(x, y, z, t)$ can be approximated as a 3D multivariate normal density. We will now determine each of the two addends of the standard deviation:

- As mentioned above, charges diffusing in three dimensions adopt a Gaussian distribution with a variance of $\sigma^{\text{diff}}(t)^2 = 2Dt$. The marginal density ρ_x^{diff} along the x-axis remains Gaussian and has the same identical variance, so $\sigma_x^{\text{diff}}(t)^2 = 2Dt$. For reasons that will become apparent in Section 4.3, expressing this formula as a time-dependent differential equation is more advantageous. Thus, we could represent it as

$$\partial_t \sigma_x^{\text{diff}}(t)^2 = 2D \tag{4.19}$$

- On the other hand, the charges spread as a result of repulsion in a uniform three-dimensional shape. The charge density $\rho^{\text{rep}}(r)$ due to repulsion-only is described in Eq. 4.15, but note that the marginal density $\rho_x^{\text{rep}}(x)$ along the $x-$axis has a more intricate mathematical description: As seen in Fig. 4.2k, it is no longer uniform. Despite its complicated shape, the *variance* of this marginal density can still be calculated by definition. For instance, after a few calculations in the intermediate steps, one can derive the following:

$$\sigma_x^{\text{rep}}(t)^2 = \int_{\infty}^{-\infty} (x-0)^2 \rho_x^{\text{rep}}(x,t)dx = \frac{R_0^2(t)}{5} \tag{4.20}$$

  As mentioned in the point above, we are now looking for an expression of the tempo-

ral derivative of the variance. Therefore, we easily derive the following expression:

$$\partial_t \sigma_x^{\text{rep}}(t)^2 = \frac{2R_0'(t)R_0(t)}{5} = \frac{\mu N q}{10\sqrt{5}\pi\epsilon\,\sigma_x^{\text{rep}}(t)} \tag{4.21}$$

Assuming that these two processes can be coupled together using a simple Gaussian distribution from Eq. 4.18 with an evolving standard deviation $\sigma_x^{\text{BH}}(t)$, one gets

$$\partial_t \sigma_x^{\text{BH}}(t)^2 = 2\left(D + \frac{\mu N q}{20\sqrt{5}\pi\epsilon\,\sigma_x^{\text{BH}}(t)}\right) = 2\tilde{D}(t) \tag{4.22}$$

In conclusion, the BH approach suggests that repulsion can also be modeled as a second-order process similar to diffusion. The model now includes a time-dependent coefficient, denoted by $\tilde{D}(t)$, which accounts for *both* diffusion and repulsion[4].

4.1.4.2 Numerical Evaluation

The BH approach provides extremely accurate predictions for the *spread* of charges. The average absolute error for RMS is as low as 0.40 μm, while the RMS error found with the Gatti model (see Fig. 4.3) was 4.20 μm. Fig. 4.4a illustrates the comparison between the ground-truth distributions and those calculated using the BH approach. Although the BH model provides a good overall fit of the charge distributions, there is room for improvement: This method does not fully capture the exact *shape* of the charge cloud.

It is important to note that fitting the ground-truth numerical solutions with a simple Gaussian distribution is fully equivalent to using the Benoit-Hamel (BH) approach. In

[4]Note that another strength of the BH approach is that a similar mathematical procedure applies for elliptical (non-spherically symmetric) charge distributions. We can model these shapes using asymmetric 3D Gaussian distributions, where the diagonal covariance matrix may have $\sigma_x \neq \sigma_y \neq \sigma_z$. The asymmetry for initialization provides higher degrees of freedom and flexibility, which generally allows better fits to the experimental data.

essence, the BH method inherently provides the optimal fit to the data. This property is evident because the Gaussian distribution involves only two parameters: A known mean of zero and a standard deviation, which is precisely derived from the ODE in Eq. 22 to match the standard deviation of the data.

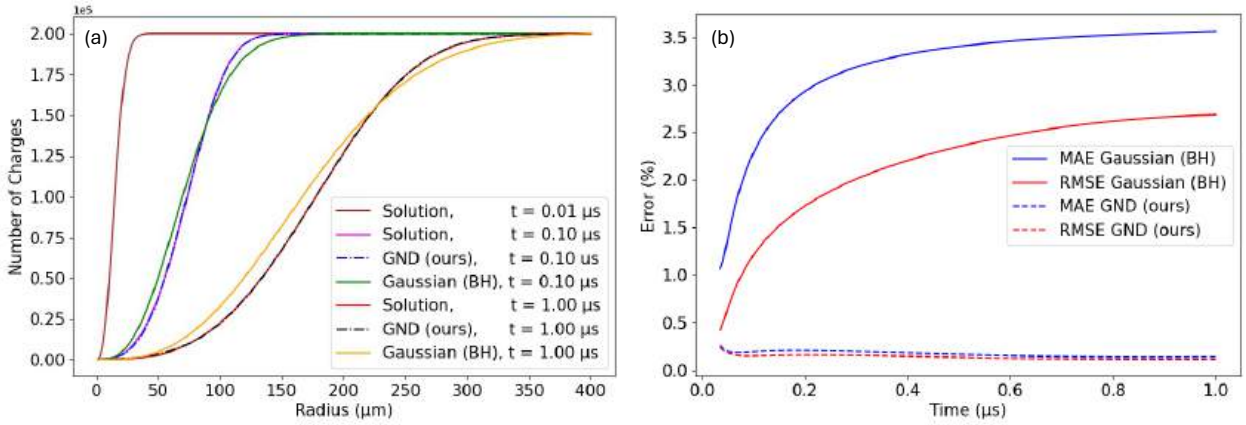


Figure 4.4: (a) Charge distribution (non-normalized) accounting for the total amount of elementary charges, plot at different times. We see the ground-truth distributions and those generated with the Benoit-Hamel model and with our proposed Generalized Normal Distribution (GND) model. (b) Error metrics of the BH and our GND models.

To show the discrepancy between the ground-truth and approximated normal distributions, we will measure the root mean squared error (RMSE) and the mean absolute error (MAE) at each time,

$$\text{RMSE}(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\hat{Q}^{\text{gt}}(R_i,t) - \hat{Q}^{\text{approx}}(R_i,t)\right]^2} \tag{4.23a}$$

$$\text{MAE}(t) = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{Q}^{\text{gt}}(R_i,t) - \hat{Q}^{\text{approx}}(R_i,t)\right| \tag{4.23b}$$

The time-averaged errors for the BH model were: An MSE of 3.21% and an RMSE of 2.19%. Fig. 4.4b shows the RMSE and MAE over time. Moreover, one can see how these errors tend to increase as the charges evolve over time and continue to spread.

4.1.4.3 Monte Carlo Algorithm

The advantage of using a simple Gaussian distribution approximation for both processes lies in the ability to simulate the diffusion-repulsion process as a *random walk*. Note that the marginal charge distribution along the $x-$axis, $\rho_{\text{x}}(x,t)$, follows $N\big[0,\sigma_x^{\text{BH}}(t)^2\big]$ at any time step $t \in \{1,...,T=200\}$. To sequentially simulate this marginal distribution in a stochastic manner, we will now consider $X_i(t)$ random variables, for $i \in \{1,2,...,S\}$ and $S = 200$ the number of samples, each distributed as

$$X_i(t) \sim X_i(t-1) + N\big[0,\Delta\sigma^{\text{BH}}(t)^2\big] \tag{4.24}$$

The initial delta distribution at $t = 0$ produced right after the photon-absorption is often represented numerically as a Gaussian distribution with a small standard deviation. In our simulations, we initialize $X_i(0) \sim N\big[0,\sigma^{\text{BH}}(0)^2\big]$, where the standard deviation $\sigma^{\text{BH}}(0)$ is 10 μm, as done in [310]. Therefore, the progression iteratively expands to:

$$\begin{aligned} X_i(t) \sim N\big[0,\sigma^{\text{BH}}(0)^2\big] + \\ \sum_{i=0}^{M} N\big[0,\sigma^{\text{BH}}(t-i)^2 - \sigma^{\text{BH}}(t-(i-1))^2\big] \end{aligned} \tag{4.25}$$

Note that the sum of independent Gaussian variables follows a Gaussian distribution [321], with the mean and variance being the sum of the individual means and variances. Applying this property to Eq. 4.25, it can be seen that the random variables $X_i(t)$ will indeed follow $N\big[0,\sigma^{\text{BH}}(t)^2\big]$, in agreement with the charge distribution described by the Benoit-Hamel method.

Observe that we initially had to solve the complicated non-linear PDE from Eq. 4.6, which required the computationally intensive IFDM. In contrast, we can now iteratively

propagate the S representative charges stochastically at each time step. Before each step, we need to find $\Delta\sigma^{\mathrm{BH}}(t)^2 = \sigma^{\mathrm{BH}}(t)^2 - \sigma^{\mathrm{BH}}(t-1)^2 = 2\tilde{D}(t)\Delta t$, as described in Eq. 4.22. It becomes now clear why we reformulated the standard deviation using a time-dependent ODE. Moreover, we can then efficiently solve that equation using a first-order explicit finite-difference scheme.

Of course, the same approach followed for the propagation of charge in the $x-$ direction can also be implemented for the y and z axes, due to the spherical symmetry. The complete BH method is outlined in Algorithm 1, which takes only about 9-16 µs on a standard computer (hundreds of times faster than with the numerical IFDM)[5]. The results are shown in Fig. 4.5.

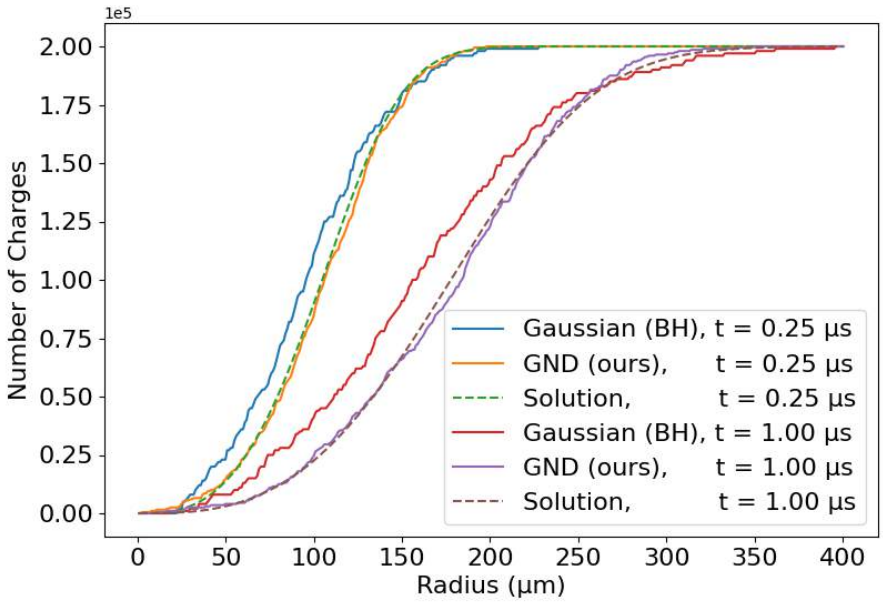


Figure 4.5: Monte Carlo simulation with $S = 200$ representative samples, using the simple Gaussian model (BH) and the GND model (ours).

[5]In the algorithm, the notation $\mathcal{N}_S(\cdot,\cdot)$ denotes the generation of S normally distributed samples with a certain mean and variance.

Algorithm 1 BH approach: Diffusion and repulsion

Require: Simulation parameters
Number of samples, $S = 200$
Time steps, $T = 200$
Time increment, $\Delta t = 5 \times 10^{-9}$ s
Require: Physical constants
Electron charge, $q = 1.60 \times 10^{-19}$ C
Permittivity, $\epsilon_0 = 8.85 \times 10^{-12}$ F/m
Relative permittivity, $\epsilon_r = 11$
Diffusion coefficient, $D = 25\,\mathrm{cm}^2/\mathrm{s}$
Electron mobility, $\mu = 1000\,\mathrm{cm}^2/\mathrm{Vs}$
Number of electrons, $P = 2 \times 10^5$
Initial standard deviation, $\sigma_0 = 10\,\mu\mathrm{m}$

$\epsilon \leftarrow \epsilon_0 \cdot \epsilon_r$
$\sigma^2_{\mathrm{BH}}[0] \leftarrow \sigma_0^2$
$x[0], y[0], z[0] \sim \mathcal{N}_{\mathrm{S}}(0,1) \cdot \sigma_0$

for $t \leftarrow 1$ to T **do**
    $\tilde{D} \leftarrow D + \frac{\mu P q}{20\sqrt{5}\,\pi\,\epsilon\sqrt{\sigma^2_{\mathrm{BH}}[t-1]}}$
    $\Delta\sigma^{\mathrm{BH}} \leftarrow \sqrt{2 \cdot \tilde{D} \cdot \Delta t}$
    $x[t] \leftarrow x[t-1] + \mathcal{N}_{\mathrm{S}}(0,1) \cdot \Delta\sigma^{\mathrm{BH}}$
    $y[t] \leftarrow y[t-1] + \mathcal{N}_{\mathrm{S}}(0,1) \cdot \Delta\sigma^{\mathrm{BH}}$
    $z[t] \leftarrow z[t-1] + \mathcal{N}_{\mathrm{S}}(0,1) \cdot \Delta\sigma^{\mathrm{BH}}$
    $\sigma^{\mathrm{BH}}[t]^2 \leftarrow \Delta\sigma^{\mathrm{BH}}[t-1]^2 + 2 \cdot \tilde{D} \cdot \Delta t$
end for

Additionally, it should be emphasized that the BH approach is compatible with existing detector simulators that account for additional charge displacements and fluctuations. At each time step, one can drift the S representative charges due to the bias E-field, then add a secondary weaker walk due to diffusion-repulsion, and finally recalculate the amount of charges after the trap-induced recombination process, as explained in more details in [310].

4.1.5 Our Proposed GND Model

4.1.5.1 Generalized Normal Distributions

Rather than approximating the normalized charge distribution $\hat{Q}(R,t)$ as a 3D Gaussian shape, we now propose that it more closely resembles a Generalized Normal Distribution (GND). The probability density function (PDF) of the GND is defined as follows:

$$f^{\text{GND}}(r,t) = \frac{\beta^{\text{GND}}}{2\alpha^{\text{GND}}\Gamma(1/\beta^{\text{GND}})} \exp\left(\frac{|x-\mu^{\text{GND}}|}{\alpha^{\text{GND}}}\right)^{\beta^{\text{GND}}} \tag{4.26}$$

Note that we must consider the parameters as time-dependent, being for instance $\mu^{\text{GND}}(t)$ the mean. Additionally, the mean is now not necessarily zero (see Figs. 4.2d-f), since Eq. 4.26 represents the radial density $f(r,t) = \partial_R \hat{Q}(R,t)$ and *not* the marginal density on the $x-$axis (Figs. 4.2j-l). The parameter $\beta^{\text{GND}}(t)$ modulates the peakedness of the distribution, categorizing a family of distributions. Specifically, $\beta^{\text{GND}} = 2$ corresponds to the standard normal distribution with mean $\mu^{\text{GND}}(t)$ and variance $\alpha^{\text{GND}}(t)^2/2$. Values of β^{GND} less than two result in distributions with sharper peaks and heavier tails, whereas values greater than two yield flatter densities with lighter tails.

Figure 4.4a presents the ground-truth numerical solutions alongside their fits using a Gaussian distribution (solid lines) and the GND (dashed lines, superimposed on the numerical solutions). As anticipated, the greater flexibility of the GND allows for a markedly improved fit to the numerical solutions. The RMS error decreases now even more than with the BH approach, reaching 0.18 μm with the GND. Similarly, Fig. 4.4b illustrates the error between the ground-truth numerical solutions and their fits, comparing the simple Gaussian distribution (solid blue and red lines) with the GND (dashed blue and red lines). The corresponding time-average MAE and RMSE are exceptionally low for the GND fits,

with values of 0.17% and 0.14%, respectively. These errors are an order of magnitude lower than those obtained from simple Gaussian fits. More importantly, the error appears to be stable over time (see Fig. 4.4b). We also evaluated additional metrics for comparison, such as the Kullback-Leibler (KL) divergence, the Pearson correlation coefficient, and the cosine similarity. The results, detailed in the Table 4.1, further highlight the advantages of adopting the Generalized Normal Distribution (GND) over the simple Gaussian distribution. Observe that for each time step, the parameters were found by solving a least squares error problem, which fits the ground-truth distribution to our proposed model using the Trust Region Reflective Optimizer defined in the Python package *SciPy* [322].

Table 4.1: Comparison of metrics: Fitting ground-truth distributions using the simple Gaussian model by BH and our proposed GND method.

Metrics	Gaussian (BH)	GND (Ours)
RMS absolute error (μm)	0.40	0.18
MAE (%)	3.21	0.17
RMSE (%)	2.19	0.14
Cosine similarity	$1 - 2.19 \cdot 10^{-3}$	$1 - 8.92 \cdot 10^{-6}$
Correlation coefficient	$1 - 9.20 \cdot 10^{-4}$	$1 - 2.14 \cdot 10^{-6}$
KL divergence	$1.94 \cdot 10^{-4}$	$1.32 \cdot 10^{-5}$

4.1.5.2 Gaussian Mixture Model

We aim to develop a Monte Carlo simulation based on Gaussian processes, as previously done in Section 4.3. This will enable the integration of our approximated results for repulsion-diffusion into a broader simulator that also accounts for the additional effects of charge drift and fluctuations. To build the stochastic method, we will first apply a Gaussian Mixture Model (GMM) to model the generalized normal distribution (GND) as a sum

of Gaussian distributions. Our findings indicate that the sum of only two Gaussian components are sufficient to accurately represent the GNDs that appear in our problem. The PDF of the GMM is then expressed as follows:

$$f^{\text{GMM}}(r) = w_1\phi(r;\mu_1,\sigma_1^2) + w_2\phi(r;\mu_2,\sigma_2^2) \tag{4.27}$$

Here, ϕ represents the PDF of a Gaussian distribution, and the parameters (μ_1, σ_1^2) and (μ_2, σ_2^2) specify the means and variances of the two Gaussian components, respectively. Note that the GND is symmetric around its central point μ_{GND}. We will maintain such a property by ensuring that both Gaussian functions have: (1) the same weight $w_1 = w_2 = 0.5$, (2) the same standard deviation $\sigma_1 = \sigma_2$, and (3) symmetric means with $|\mu^{\text{GND}} - \mu_1| = |\mu^{\text{GND}} - \mu_2|$. Let us now determine the exact value for these four unknown parameters:

- We can find a reasonable value for μ_1 (and therefore μ_2 also) by matching the first three moments (mean, variance, and skewness) of the GMM with the GND. In particular, if we define

  $$\mu_1(t) := \sqrt{\sigma^{\text{GND}}(t)^2 - \sigma_1(t)^2} + \mu^{\text{GND}} \tag{4.28}$$

  Then we can derive, from Eq. 4.28, the following

  $$\mu^{\text{GMM}} = w_1\mu_1 + w_2\mu_2 = \mu^{\text{GND}} \tag{4.29a}$$

  $$\begin{aligned}(\sigma^{\text{GMM}})^2 &= w_1\sigma_1^2 + w_2\sigma_2^2 + w_1(\mu_1 - \mu^{\text{GMM}})^2 \\ &\quad + w_2(\mu_2 - \mu^{\text{GMM}})^2 = (\sigma^{\text{GND}})^2\end{aligned} \tag{4.29b}$$

  Note that both results are valid for any σ_1 value (we will decide on a particular value

below). Because both the GMM and the GND are symmetric, there is also a match of the third moment (the skewness is zero).

- We can now find a convenient σ_1 value by matching the peak of the GMM density with the peak of the GND density at the central point, $r_0 = \mu^{\text{GND}}$, meaning

$$f^{\text{GND}}(r_0; \mu^{\text{GND}}, \sigma^{\text{GND}}) \equiv f^{\text{GMM}}(r_0; \mu_1, \sigma_1) \tag{4.30}$$

Substituting the value of μ_1 from Eq. 4.28 and considering the known peak value c_{peak} of the GND distribution, one finds:

$$f^{\text{GMM}}(\sigma_1) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(\frac{(\sigma^{\text{GND}})^2 - \sigma_1^2}{2\sigma_1^2}\right) = c_{\text{peak}} \tag{4.31}$$

This non-linear algebraic equation can be efficiently solved as a least squares minimization problem, where the objective function represents the residuals of the equation. Specifically, we employed the Python package *SciPy* and utilized again the efficient trust-region optimizer, incorporating box constraints to ensure that σ_1 remains within the physically meaningful range of positive values less than the known parameter σ^{GND}.

Fig. 4.6 shows the fits of the GND probability density functions at two different times using the proposed mixture model.

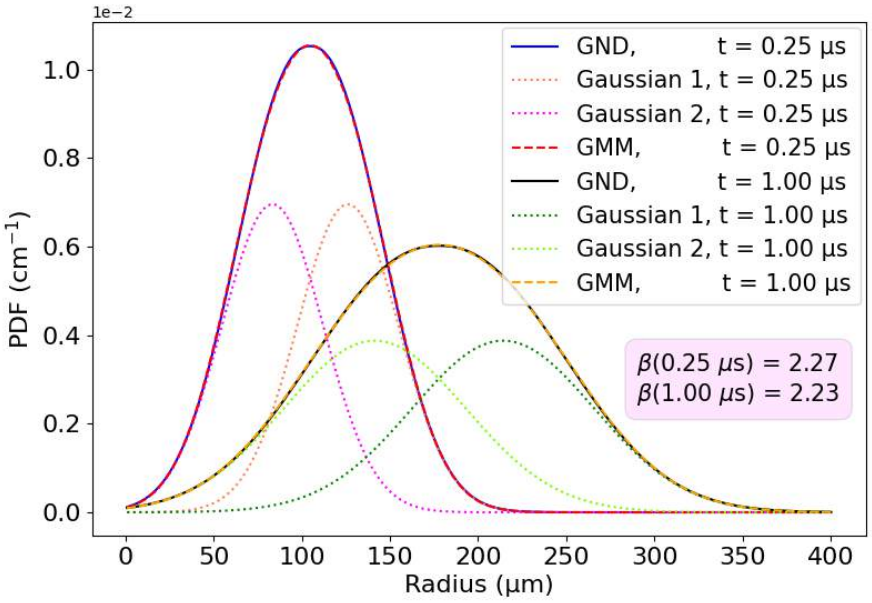


Figure 4.6: Gaussian Mixture Model (with two Gaussian functions) matching the PDF of the charges at two different times.

4.1.5.3 Monte-Carlo Algorithm

Using Eqs. 4.28 and 4.30, we can also compute the values for (μ_1, σ_1) and (μ_2, σ_2). We now have all the necessary information to develop a Monte Carlo simulator with the proposed GND approach. This simulator follows a methodology analogous to the one described in Algorithm 1, with the slight difference of (1) having two simultaneous Gaussian processes instead of one, and (2) doing the updates of the radial distance instead of the Cartesian coordinates. The new methodology is outlined in Algorithm **??**. The computational cost of this modified algorithm ranges between 16-19 μs on a standard computer, approximately double that of the original BH approach. The charge distributions found with this simulation are illustrated in Fig. 4.5.

Algorithm 2 Novel approach: Diffusion and repulsion

Require: Simulation parameters (as in Algorithm 1)
Require: Physical constants (as in Algorithm 1)
Require: Pre-computed parameters: $\mu_1, \mu_2, \sigma_1, \sigma_2$

$\epsilon \leftarrow \epsilon_0 \cdot \epsilon_r$
$r_1[0] \leftarrow \mu_1[0] + \mathcal{N}_S(0, \sigma_1[0])$
$r_2[0] \leftarrow \mu_2[0] + \mathcal{N}_S(0, \sigma_2[0])$
$\phi_1, \phi_2 \leftarrow \text{rand}_S[0, 2\pi]$
$\theta_1, \theta_2 \leftarrow \text{rand}_S[0, \pi]$
$x_1[0], x_2[0] \leftarrow r_1[0] \sin\phi_1 \cos\theta_1, r_2[0] \sin\phi_2 \cos\theta_2$
$y_1[0], y_2[0] \leftarrow r_1[0] \sin\phi_1 \sin\theta_1, r_2[0] \sin\phi_2 \sin\theta_2$
$z_1[0], z_2[0] \leftarrow r_1[0] \cos\phi_1, r_2[0] \cos\phi_2$
for $t = 1$ to T **do**
    $\Delta\mu_1[t] \leftarrow \mu_1[t] - \mu_1[t-1]$
    $\Delta\mu_2[t] \leftarrow \mu_2[t] - \mu_2[t-1]$
    $\Delta\sigma_1[t] \leftarrow \sigma_1[t] - \sigma_1[t-1]$
    $\Delta\sigma_2[t] \leftarrow \sigma_2[t] - \sigma_2[t-1]$
    $\Delta r_1[t] \leftarrow \Delta\mu_1[t-1] + \mathcal{N}_S(0,1) \cdot \Delta\sigma_1$
    $\Delta r_2[t] \leftarrow \Delta\mu_2[t-1] + \mathcal{N}_S(0,1) \cdot \Delta\sigma_2$
    $x_1[t] \leftarrow x_1[t-1](1 + \Delta r_1 / r_1[t-1])$
    $x_2[t] \leftarrow x_2[t-1](1 + \Delta r_2 / r_2[t-1])$
    $y_1[t] \leftarrow y_1[t-1](1 + \Delta r_1 / r_1[t-1])$
    $y_2[t] \leftarrow y_2[t-1](1 + \Delta r_2 / r_2[t-1])$
    $z_1[t] \leftarrow z_1[t-1](1 + \Delta r_1 / r_1[t-1])$
    $z_2[t] \leftarrow z_2[t-1](1 + \Delta r_2 / r_2[t-1])$
    $r_1[t] \leftarrow \sqrt{x_1[t]^2 + y_1[t]^2 + z_1[t]^2}$
    $r_2[t] \leftarrow \sqrt{x_2[t]^2 + y_2[t]^2 + z_2[t]^2}$
end for

4.1.5.4 Extending the Algorithm for Different Scenarios

One can observe that there is a high correlation between the parameter $\sigma^{\text{BH}}(t)$ found with the simple Gaussian fit and the parameters $\mu^{\text{GND}}(t)$ and $\alpha^{\text{GND}}(t)$ of the GND approach. This correlation is clearly illustrated in Figs. 4.7a-b. The consistency between the two models allows us to generalize our proposed GND for other scenarios (with different material properties and deposited energies) without having to pre-compute the GND parameters for each scenario beforehand, rather using the very convenient results from the Benoit-Hamel model.

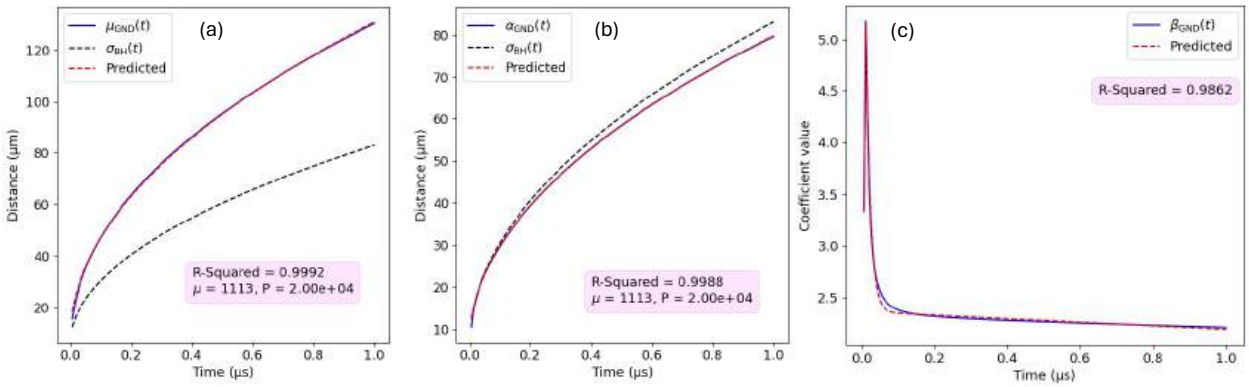


Figure 4.7: Illustrating the strong linear correlation between the time-dependent parameters (a) μ^{GND} and σ^{BH}, (b) α^{GND} and σ^{BH}, for a representative scenario with $P = 2 \times 10^4 \in [2 \times 10^4, 2 \times 10^5]$ electron-hole pairs and with an electron mobility of $\mu = 1113 \in [700, 1300]$ cm^2/Vs. (c) Time-average β^{GND} for different (P, μ) scenarios compared with our proposed model from Eq. 4.33.

For a specific material with permittivity ϵ, the dynamic of the charges due to diffusion and repulsion dynamics are solely[6] influenced by the charge mobility μ and the number of electron-hole pairs generated P. Taking for example a wide range of interest, such as $\mu \in [700, 1300]$ for electrons in CZT and $P \in [2 \times 10^4, 2 \times 10^5]$, we have established the following relationships:

$$\mu^{\text{GND}}(t; \mu, P) = [\mu a_1 + P b_1 + c_1] \cdot \sigma^{\text{BH}}(t) + [\mu d_1 + P e_1 + f_1] \tag{4.32a}$$

$$\alpha^{\text{GND}}(t; \mu, P) = [\mu a_2 + P b_2 + c_2] \cdot \sigma^{\text{BH}}(t) + [\mu d_2 + P e_2 + f_2] \tag{4.32b}$$

The particular values of the coefficients are provided in Appendix B for these particular range of interest. Similar analyses could be done by the reader in other different ranges of interest for μ and P to analyze different crystals.

[6]It is important to remark again that the diffusion coefficient D is not an independent variable - it is linearly related to μ by the Einstein-Smoluchowski relation [316]

It is important to note that shaping parameter β^{GND} shows a minimal dependency on the physical parameters (μ, P), and therefore a correlation with σ^{BH} was not found. This time-dependent coefficient influences the peakedness of the distribution, and a value greater than two allows for the steeper distributions observed in our problem (see Fig. 4.4a). Our findings indicate that the time evolution of $\beta^{\text{GND}}(t)$ aligns closely with a model that features an initial exponential decay followed by a gradual transition to a linear decrease. We have modeled this parameter as a piecewise function with $\beta^{\text{GND}}(t = 0) = a_3$, matching the initial distribution at the starting time, and then modeling the time evolution for $t > 0$ as

$$\beta^{\text{GND}}(t) = b_3 e^{-c_3 t} + (d_3 t + e_3) \tag{4.33}$$

The particular coefficients are also provided in the Appendix B. The R-squared of the β^{GND} parameter from Eq. 4.33, calculated with respect to different ground-truth $\beta^{\text{GND}}(t; \mu, P)$, is in average 0.986, showing an excellent fit, as one can observe in Fig. 4.7c.

Employing the time-dependent parameters from Eqs. 4.32-4.33 using the presented correlation analysis, we obtained the time-averaged MAE of 0.44% and RMSE of 0.35% for the charge cloud distribution. These results are slightly worse than those obtained from pre-computed parameters (shown in Table. 4.1), but still about an order of magnitude better than the results from the BH approach. In exchange for losing some accuracy, this correlation-based method offers greater flexibility: It extends our novel algorithm to different scenarios without the need to pre-compute any values at all, just using the simple solutions for $\sigma^{\text{BH}}(t)$ from the Benoit-Hamel algorithm.

4.1.6 Concluding Remarks

In this study, we examine the charge dynamics in semiconductors, focusing on diffusion and repulsion effects, with specific applications to high-energy semiconductor detectors.

First, we treated diffusion and repulsion as decoupled processes and used the theoretical findings from Gatti et al [309] to get analytical expressions for the charge distributions. While the charges propagate following a multivariate (3D) Gaussian distribution under the influence of diffusion, they propagate as a multivariate (3D) uniform distribution due to Coulomb repulsion. Considering the RMS as the quadratic sum of both separate processes leads to considerable inaccuracies. In particular, our simulations for electron dynamics in a CZT detector, the time-average RMS mismatch for the charge cloud distribution was found to be 4.20 μm. On the bright side, these approximated results can be computed very efficiently from the close-form expressions.

Second, we analyzed how the Benoit-Hamel (BH) approach [310] improves upon the previous model by coupling both processes, approximating the overall charge distribution as a simple multivariate 3D Gaussian function. This methodology enables an efficient implementation of the distributions using a Monte Carlo technique. It also allows for the easy integration of these two weak processes in more general simulation frameworks that account for other effects. Note that the standard deviation of the Gaussian distribution increases with time, reflecting the spatial expansion of the charges. In our simulations, the BH approach yielded a time-averaged RMS error of only 0.40 μm. Although this method provides high accuracy in predicting the spread of charges, the actual shape of the charge distribution does not closely resemble a Gaussian function. For instance, the MAE and RMSE of the charge distribution were 3.21% and 2.19%, respectively.

Third, we have developed a novel method that naturally extends the BH approach. Rather than taking a simple Gaussian as an *ansatz*, we consider a Generalized Normal Distribution (GND). This new assumption provides higher degrees of freedom that allow better fits of the actual ground-truth distributions. We approximated the GND as a Gaussian Mixture Model (GMM) with two Gaussian functions, enabling an efficient simulation of these processes using another Monte Carlo algorithm. Fitting the GND to the ground-truth numerical solutions lead to a RMS error of the cloud of 0.18%. The MAE and RMSE of the charge distribution (cloud shape) is also reduced to 0.17% and 0.14%. These results then demonstrate the strong capability of this more complex distribution to accurately model the system. The GND algorithm can be easily extended to various scenarios (with different material properties and deposited energies) by determining the GND parameters from the BH parameters via correlation analysis. Using this correlation approach, the error between the ground-truth numerical solution and the GND results remains low, with a MAE of 0.44% and a RMSE of 0.35% for the cloud shape. The computational cost of the GND approach is about double of the BH approach, though still reasonable and in the order of tens of milliseconds in standard computers.

These findings facilitate the refinement of semiconductor detector simulators and contribute to the analysis of next-generation detectors. Using these simulators, future research could focus on accurately solving inverse problems considering diffusion and repulsion, either by performing event reconstructions or by characterizing the material properties of the detector bulk.

4.2 Defect detection with automatic differentiation

CdZnTe-based detectors are highly valued because of their high spectral resolution, which is an essential feature for nuclear medical imaging. However, this resolution is compromised when there are substantial defects in the CdZnTe crystals. In this study, we present a learning-based approach to determine the spatially dependent bulk properties and defects in semiconductor detectors. This characterization allows us to mitigate and compensate for the undesired effects caused by crystal impurities. We tested our model with computer-generated noise-free input data, where it showed excellent accuracy, achieving an average RMSE of 0.43% between the predicted and the ground truth crystal properties. In addition, a sensitivity analysis was performed to determine the effect of noisy data on the accuracy of the model.

4.2.1 Introduction

Single-photon detectors are essential in nuclear medical imaging techniques, where high spectral resolution is critical for tissue differentiation. In particular, detector technologies based on cadmium zinc telluride (CdZnTe) are increasingly favored in modalities that operate with relatively low photon flux, such as *single photon emission computed tomography* (SPECT) [323]. CdZnTe crystals exhibit a high stopping power (allowing the absorption of most incoming photons) and a wide band gap of 1.6 eV (allowing detector operation at room temperature) [324]. These photon-counting detector (PCDs) offer a rapid temporal response (less than 10 ns) and can be configured with a pixelated geometry, enabling sub-millimeter spatial resolution. However, the energy resolution of these detectors is often compromised due to the high defect density typically found in the thick CdZnTe crystals [325]–[327]. These defects include point defects, dislocations, and grain boundaries [312],

which adversely influence the optoelectric properties of the crystal.

In this study, we employ a reverse engineering approach to determine the material properties and defects in the CdZnTe crystals used for radiation detectors. Building upon prior physics-based learning models developed by our team [285]–[289], our novel model accurately identifies fundamental spatially-dependent properties, such as charge mobility and the lifetimes for charge recombination, trapping, and detrapping. By precisely characterizing these features, the study seeks to mitigate the detrimental effects of crystal impurities. The model takes as input the charge-induced signals and the charge concentrations after a particular photon-detector event interaction. We initially validate our model using computer-generated noise-free data to establish a baseline for performance. Subsequently, we now assess for the first time the impact of noise in the input data on the accuracy of this characterization.

4.2.2 Detector Simulator

We have built a physics-driven model that simulates the functioning of the CdZnTe photon-counting detectors. Our simulations consider a detector with dimensions $1 \times 1 \times 1$ cm^3 and a standard pixelated configuration: a single cathode on top and nine anodes on the bottom (see Fig. 4.8). When a $\gamma-$ray penetrates the cathode and interacts with the crystal, several electron-hole (e-h) pairs are generated. Drifted by an external electric field E, these charge carriers move and produce charge-induced signals at the nearby electrodes [294], [328].

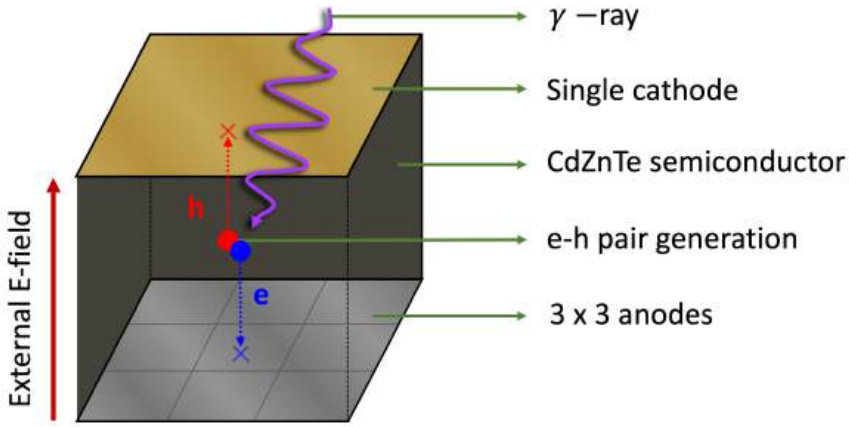


Figure 4.8: Pixelated detector geometry with one cathode on the top and nine anodes on the bottom.

Following a delta-like photon-detector event $\delta(x_0, y_0, z_0, t_0)$ occurring at time t_0 and specific location $(x_0, y_0, z_0) \in [0, 1]^3$, the electron concentration $n_{\mathrm{e}}(x, y, z, t)$ can be model with a system of partial differential equations (PDEs) [273], [329]:

$$\begin{cases} \partial_t n_{\mathrm{e}} - \nabla \cdot (\mu_{\mathrm{e}} E n_{\mathrm{e}}) = -\frac{1}{\tau_{\mathrm{eR}}} n_{\mathrm{e}} - \frac{1}{\tau_{\mathrm{Te}}} n_{\mathrm{e}} + \frac{1}{\tau_{\mathrm{eD}}} \tilde{n}_{\mathrm{e}} + \delta \\ \partial_t \tilde{n}_{\mathrm{e}} = \frac{1}{\tau_{\mathrm{eT}}} n_{\mathrm{e}} - \frac{1}{\tau_{\mathrm{eD}}} \tilde{n}_{\mathrm{e}} \end{cases} \tag{4.34}$$

Note that an equivalent system and notation can also be applied to model the dynamics of holes, where the concentration is denoted as n_{h}. Eq. 4.34 account for various dynamic processes: charge drift, charge generation-recombination, and trapping-detrapping effects. The variable $\tilde{n}_{\mathrm{e}}$ represents the concentration of electrons within the trapping energy levels, which includes shallow and deep traps. These traps are conceptualized as an infinite *wells* that attracts (and eventually expels) a percentage of the charges, as described by the Shockley-Read-Hall theory [330].

The mean lifetimes for electron charge trapping, detrapping, and recombination

are denoted by τ_{eT}, τ_{eD} and τ_{eR}, respectively. The electron mobility is given by μ_{e}, with its drift velocity being $v_{\mathrm{e}} = -\mu_{\mathrm{e}}E$. In the simulations, we considered a uniform vertical field with magnitude $E = 850$ V/cm. Our model simplifies the charge dynamics by assuming a purely vertical trajectory for the charges along the $z-$axis, omitting minor deviations that could arise from interpixel gaps [331] or polarization effects [332]. Additionally, we incorporate an empirical relationship between the mobilities of electrons and holes, estimated as $\mu_{\mathrm{h}} \approx \mu_{\mathrm{e}}/10$, based on the findings reported in [333]. Our model neglects the weak effects of diffusion and Coulomb repulsion, which is reasonable under a high electric field strength [281], [310].

Eq. 4.34 can be efficiently solved numerically using the explicit finite-difference method [284]. We defined a spatial step of $\Delta z = 0.01$ cm and a time step of $\Delta t = 10$ ns, capturing the high temporal response of the real CdZnTe-PCDs. Note that the crystal properties (charge mobility and lifetimes) can be equivalently reformulated using dimensionless computational parameters, as detailed in Table 4.2. These parameters naturally emerge during the discretization of the PDE system, as further explained in [288]. To accommodate the spatial variability of properties, we define a stratified media: There are $N = 100$ layers (stacked sequentially from the cathode to the anode), each with distinct parameter values.

Table 4.2: Common material properties of CdZnTe crystals [284], along with the corresponding dimensionless computational parameters used in the simulations. The table also shows the NRMSE between the predicted parameters and the ground-truth ones when employing noise-free and noisy input data.

Material properties	Symbol	Value	Parameter	Value	NRMSE (%)	
					Noise-free	Noisy
Charge mobility [cm^2/Vs]	μ_{e}	1120	$R_{\mathrm{e}} = (\mu_{\mathrm{e}}E)\Delta t/\Delta z$	0.95	$< 10^{-2}$	0.22
	$\mu_{\mathrm{h}} \approx \mu_{\mathrm{e}}/10$	112	$R_{\mathrm{h}} = (\mu_{\mathrm{h}}E)\Delta t/\Delta z$	0.095	$< 10^{-2}$	0.22
Recombination lifetime [μs]	τ_{eR}	10	$P_{\mathrm{eR}} = \Delta t/\tau_{\mathrm{eR}}$	0.001	2.81	23.81
	τ_{hR}	1	$P_{\mathrm{hR}} = \Delta t/\tau_{\mathrm{hR}}$	0.01	$< 10^{-2}$	2.86
Trapping lifetime [μs]	τ_{eT}	10	$P_{\mathrm{eT}} = \Delta t/\tau_{\mathrm{eT}}$	0.001	0.18	3.46
	τ_{hT}	0.067	$P_{\mathrm{hT}} = \Delta t/\tau_{\mathrm{hT}}$	0.15	$< 10^{-2}$	0.74
Detrapping lifetime [μs]	τ_{eD}	0.4	$P_{\mathrm{eD}} = \Delta t/\tau_{\mathrm{eD}}$	0.025	0.01	4.52
	τ_{hD}	0.067	$P_{\mathrm{hD}} = \Delta t/\tau_{\mathrm{hD}}$	0.15	$< 10^{-2}$	0.88

After numerically solving the system from Eq. 4.34 for both electron and hole concentrations, we obtain the charge densities as functions of time and space. The current signals for each electrode, indexed by $k \in \{1, 2, ..., 10\}$, are subsequently calculated using the Shockley-Ramo theorem [294]:

$$i^{(k)}(t) = qn_{\mathrm{e}}(t, x, y, z)E_{\mathrm{w}}^{(k)}(x, y, z), \tag{4.35}$$

where the weighting electric field $E_{\mathrm{w}}^{(k)}$ is derived from the corresponding weighting potential $\phi_{\mathrm{w}}^{(k)}$. The weighting potential is determined by solving the Poisson equation $\nabla^2\phi_{\mathrm{w}}^{(k)}(x, y, z) = 0$, subject to the boundary conditions $\phi_{\mathrm{w}}^{(k)} = 1$ at the location of the electrode k, and $\phi_{\mathrm{w}}^{(k)} = 0$ at the areas of all other electrodes. Figure 4.9 illustrates the weighting potential for an anode located at the edge of the bottom pixel array. It is worth noting that the

commonly measured charge-induced signals are simply calculated as $Q^{(k)}(t) = \int i^{(k)}(t)dt$.

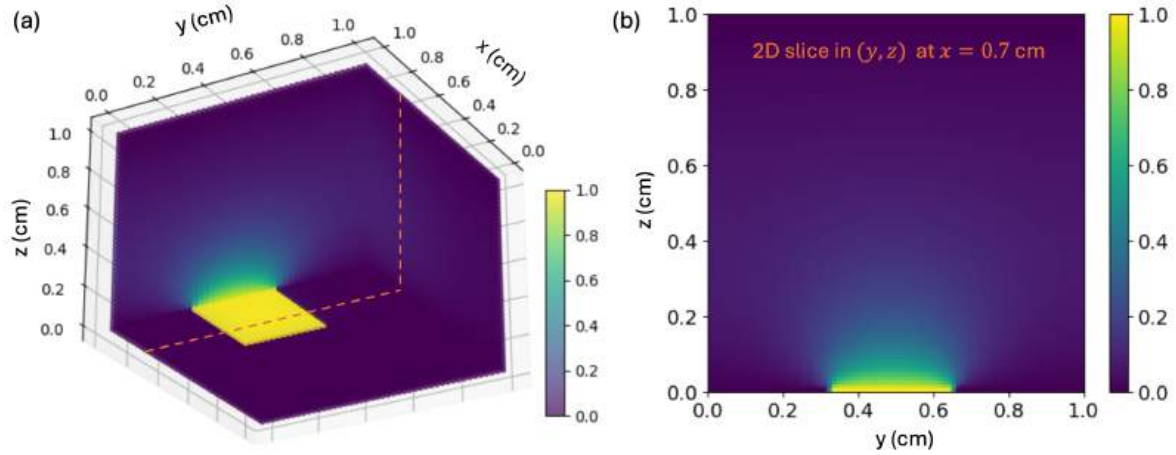


Figure 4.9: Weighting potential for anode k at the edge of the pixel array grid. (a) 3D view of the weighting potential. (b) 2D slice of the weighting potential at $x = 0.70$ cm.

4.2.3 Detector Characterization

Given the detector signals and charge concentrations, we can now construct an inverse model that provides us with the material properties and defects of the CdZnTe crystal at different locations, as outlined in Fig. 4.10. As proposed in [288], we aim to solve the following optimization problem:

$$\begin{aligned}\min C(\Theta) = &\|n_{\mathrm{e}}^{\mathrm{sim}}(t,z;\Theta) - n_{\mathrm{e}}^{\mathrm{given}}(t,z)\|_{\mathrm{F}}^2 \\ &+ \|n_{\mathrm{h}}^{\mathrm{sim}}(t,z;\Theta) - n_{\mathrm{h}}^{\mathrm{given}}(t,z)\|_{\mathrm{F}}^2 \\ &+ \|\tilde{n}_{\mathrm{e}}^{\mathrm{sim}}(t,z;\Theta) - \tilde{n}_{\mathrm{e}}^{\mathrm{given}}(t,z)\|_{\mathrm{F}}^2 \\ &+ \|\tilde{n}_{\mathrm{h}}^{\mathrm{sim}}(t,z;\Theta) - \tilde{n}_{\mathrm{h}}^{\mathrm{given}}(t,z)\|_{\mathrm{F}}^2 \\ &+ \lambda \sum_{i=1}^{10} \|Q_i^{\mathrm{sim}}(t;\Theta) - Q_i^{\mathrm{given}}(t)\|_{\mathrm{F}}^2\end{aligned} \tag{4.36}$$

The 2D array $n_{\mathrm{e}}^{\mathrm{sim}}(t, z; \Theta)$ depicts the simulated concentration of electrons at discrete times and positions, and $\| \cdot \|_{\mathrm{F}}$ the Frobenius matrix norm. The simulated concentration and signals depend on the computational parameters, which are encapsulated by $\Theta \in \mathbb{R}^{7\times100}$, seven parameters over the $N = 100$ voxels in depth. Equation 4.36 describes a problem in which we try to fit the simulator's output to the provided data. It is crucial to note that the data used in this study were generated in a computational manner, not from experiments. This means that we have prior knowledge of the ground-truth parameters, enabling a straightforward evaluation of our inverse solver algorithm. The errors between the calculated and given charge-induced signals $Q^{(k)}$ at each electrode $k \in \{1, 2, ..., 10\}$ are incorporated into the Eq. 4.36 as a regularization term with a relatively low coefficient ($\lambda = 0.1$), as proposed in [288].

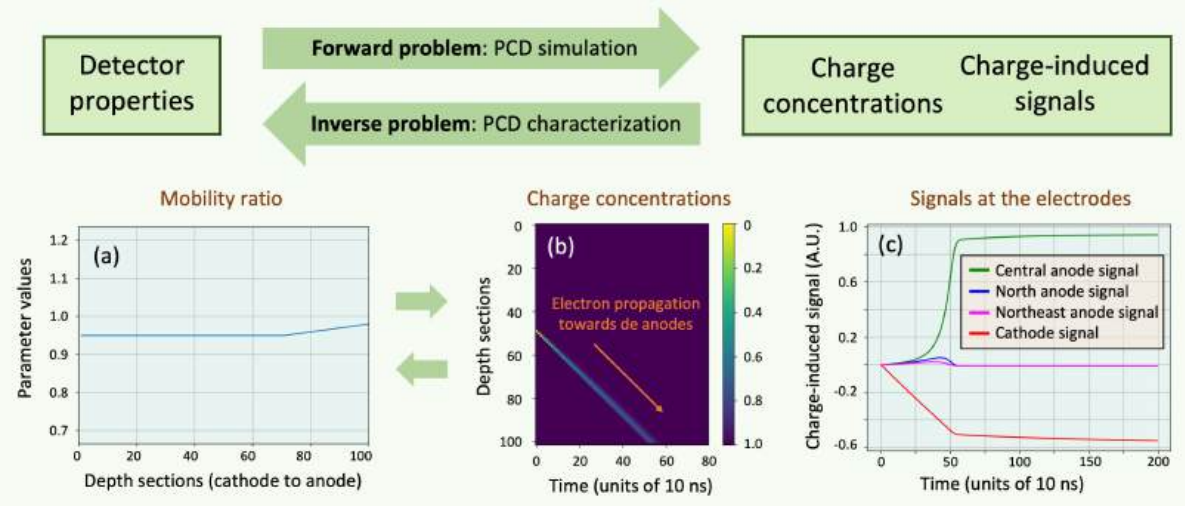


Figure 4.10: Diagram illustrating the forward and inverse problems. In the forward problem (a), the model serves as a digital twin of the photon-counting detector, providing an accurate simulation of the detector signals based on the crystal properties, such as charge mobility (μ_{e}), or equivalently, the computational parameter R_{e} (mobility ratio). Conversely, the inverse problem (b) aims to deduce the material properties from the available information on the charges and resulting signals.

The non-linear fitting problem from Eq. 4.36 displays a cost function $C(\Theta)$ with sev-

eral local minima, and therefore describes a complicated global optimization problem. Due to the curse of dimensionality [260], conventional global optimizers have a relatively high computational cost and are less suitable for our problem. To address this challenge, we employ the Adam optimizer [334], a momentum-based gradient descent method designed to handle non-convex optimization problems efficiently while avoiding stagnation at local minima. For that optimizer, we set the hyperparameters with a learning rate of $5 \cdot 10^{-4}$, a first moment $\beta_1 = 0.9$, and a second moment $\beta_2 = 0.999$.

Please observe that we used automatic differentiation to accelerate gradient computations with the *PyTorch* library. We also utilized an NVIDIA Tesla GPU to run our program, which is specifically designed to accelerate AI operations. We are able to perform 20,000 iterations, until convergence, in approximately 3 hours.

To evaluate the model outcomes, we adopt the Normalized Root Mean Square Error (NRMSE) of each parameter within a specified Region of Interest (ROI). For an event at the center of the detector (depth voxel 50), the ROI for the electron properties spans voxels 50-100, as most of the generated free electrons reach the anode location (voxel 100). In contrast, due to the relative lower drift velocity of holes, they predominantly undergo recombination or trapping when traveling through the first 10 voxels, being their ROI between voxels 40-50. Finally, it is important to note that we enforce box constraints using the projected gradient method [335]. The NRMSE could range $\pm 25\%$ relative to the values of the ground-truth parameters.

4.2.4 Results

We first want to extract detector characteristics and defects considering ideally noise-free computer-generated input data. Figure 4.11 a shows both the predicted and ground-

truth computational parameter R_e (see Table 4.2) in each depth voxel, indicative of charge mobility. Consequently, we are using a single mobility parameter for both electrons (with ROI highlighted in blue) and holes (ROI in red). Figure 4.11b offers a fitting result for another representative parameter (associated with the electron trapping lifetime). Please find in Table 4.2 all the NRMSE found for the multiple computational parameters. The average error was only 0.43% for all the parameters, indicating the correct convergence.

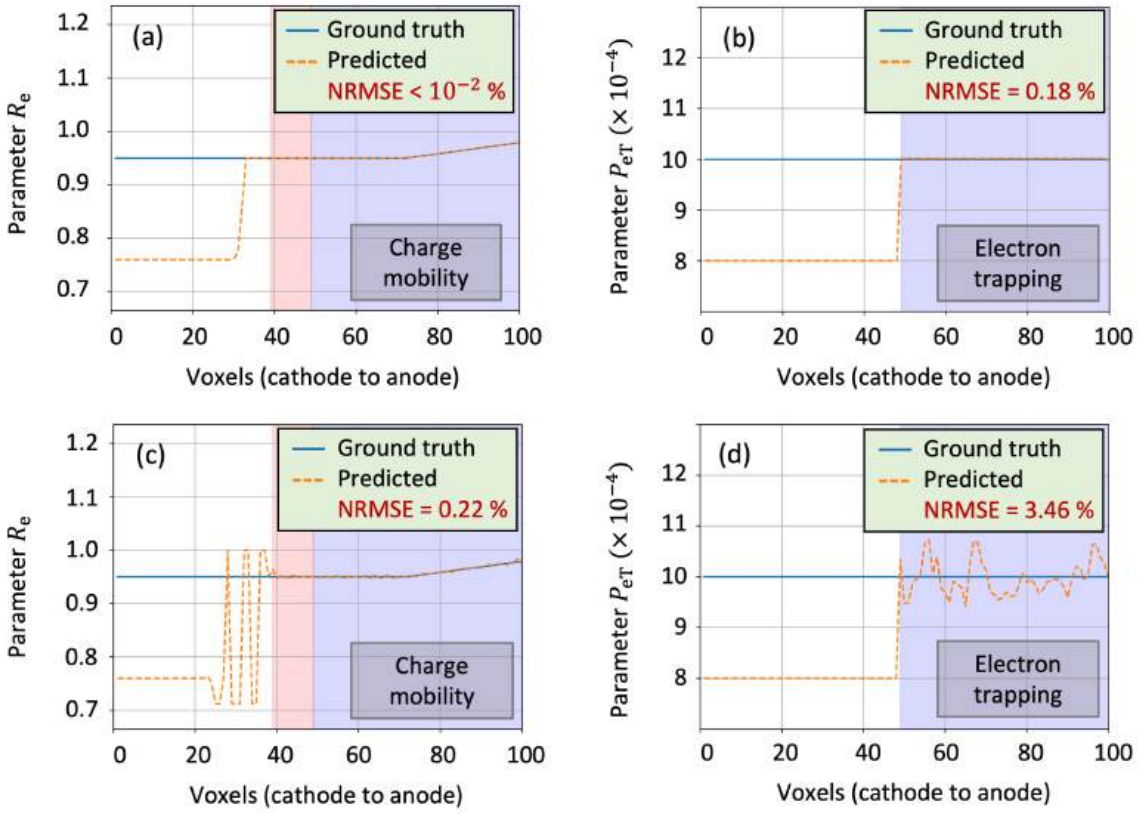


Figure 4.11: Comparison of ground-truth and predicted parameters for (a,b) noise-free, and (c,d) noisy input data. (a,c) Mobility ratio, and (b,d) electron trapping parameter.

We will now analyze the case in which an additive Gaussian noise is introduced to the computer-generated input datasets. The standard deviation of the Gaussian noise is 0.1% of the maximum signal value. It is crucial to note that the cost function C should not

reach zero in these noisy scenarios: A cost below the threshold given by the ground truth parameters, $C(\Theta^{\mathrm{gt}})$, would indicate an overfit, suggesting that the model has learned the noise. The regularizer in Eq. 4.36 prevents such overfitting, as it inhibits the algorithm from adjusting both the noise in the concentrations and the signals simultaneously. Figures 4.11c-4.11d show a comparison between the predicted parameters R_{e} and P_{eT} and their ground-truth values within the regions of interest for this case with noisy data. The average NRMSE of the parameters determined became 4.89%. Although R_{e} maintained a relatively low error of 0.22%, the error for P_{eT} increased to 3.46% (see Fig. 4.11). The last column of the Table 4.2 presents the results for all other parameters, in this case of noisy input data. In particular, we can see that the electron recombination parameter, P_{eR}, exhibits a significantly larger error. Its precise prediction during training becomes less critical because small variations in this parameter do not significantly impact the overall concentrations and signals [284].

Figure 4.12a indicates the correlation between the increase in the standard deviation in the input data and the increase in the average error in the predicted parameters. With a standard deviation ranging from 0% to 2%, the highest observed NRMSE was 16.19%. Interestingly, the random nature of the added noise and the ADAM optimizer itself led to the peak at 1.5% and not 2.0%, as one might initially anticipate. In contrast, the NRMSE for the mobility ratio increased almost linearly with the added noise level and found a maximum value of only 3.01%. In Fig. 4.12b, one sees the ideal ground-truth signals compared to the signals generated by the predicted parameters. Even with an input noise level that has a standard deviation of 2%, the predicted signals align closely with the ground truth, resulting in an error of 0.25%.

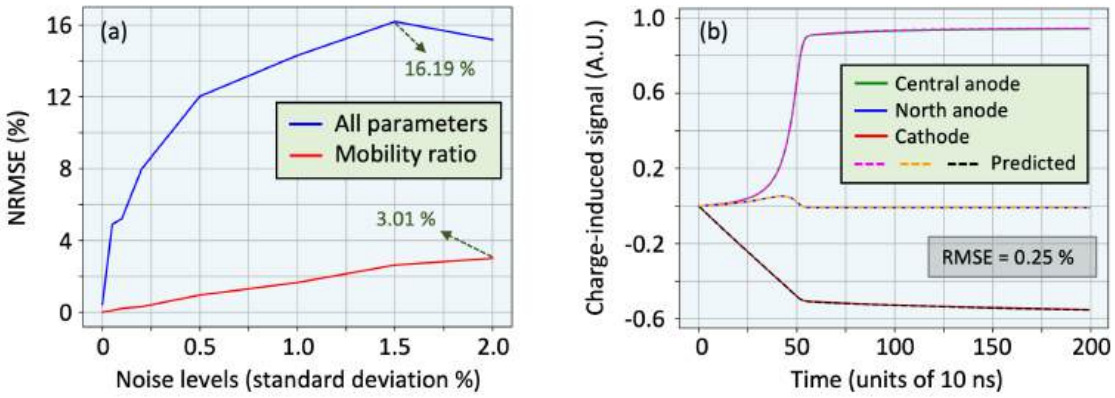


Figure 4.12: (a) Average error in predicted parameters with varying noise levels in the input data. (b) Ground-truth signal compared to signals from predicted parameters under high noise conditions (std = 2%).

4.2.5 Discussions

Although our inverse model demonstrated high accuracy with noise-free input data, the introduction of additive Gaussian noise to the charge concentrations and signals led to more pronounced errors. It is important to emphasize that the model still yields valuable insights even under high noise levels, as explained in the next two arguments.

First, the initial parameters deviate by 20% from the ground truth values, and the limit of the box constraint was up to $\pm 25\%$. In contrast, the average error of our model remained well below 16.2% for all scenarios (see Fig. 4.12a). These results demonstrate that even under pathologically noisy conditions, the algorithm still refines the initial estimates.

Second, throughout all noise conditions, the mobility ratio maintained a relatively low error. Slight variations in this parameter induce large changes in the signal production, and therefore in the cost function value [284]. Therefore, this high sensitivity en-

ables the correct characterization of the charge mobility.

However, our approach is not without limitations. On the one hand, we recognize that our approach involves the so-called "inverse crime", since we employ the same simulator for the generation of input data (with or without noise) and to solve the optimization problem. Future studies should validate the model using real-world experimental data. On the other hand, it must be noted that, while our detector simulator [284] is efficient, it operates under simplified assumptions. As mentioned, it considers a uniformly directed external E-field from the anodes to the cathode (resulting in one-dimensional charge trajectories, in the $z-$ direction) and also neglects the charge cloud expansion in the 3D space due to the second-order effects of charge diffusion and Coulomb repulsion.

4.2.6 Conclusions

This study introduces a reverse engineering model to infer spatially varying material properties and defects in CdZnTe detectors using charge concentrations and signals. Our developed software shows a relatively high accuracy (with an average error of 0.43%) when using noise-free input data. However, the error increased to 4.89% when processing data inputs affected by the additive Gaussian noise, with a standard deviation of 0.1% of the signal peak. Despite the performance degradation with higher noise levels, the derived parameters still provide a refined understanding compared to the initial assumptions. Furthermore, our model accurately predicted the mobility ratio even under conditions of very high noise levels. Finally, it is important to highlight that the parameters related to the lifetimes of recombination, trapping, and detrapping are highly susceptible to noise disturbances in the input data.

4.3 Event reconstruction and detector characterization

4.3.1 Slime Mold Algorithm for detector characterization

Semiconductor detectors for high-energy detection ($X-$rays and $\gamma-$rays) are generally categorized as either energy-integrating detectors (EIDs) or photon-counting detectors (PCDs). EIDs accumulate the total energy deposited by incoming photons to produce a single-channel output, whereas PCDs register individual photon interactions, enabling energy discrimination and multi-channel image formation. This energy-resolved capability is particularly beneficial when photon fluxes are low enough to allow distinct pulse identification [336].

These semiconductor detectors have a wide range of applications, including high-energy astrophysics, particle physics, homeland security, and nuclear medical imaging. The latter includes modalities such as radiography, computed tomography (CT), single-photon emission computed tomography (SPECT), and positron emission tomography (PET). While EIDs are commonly used in traditional CT, PCDs are essential in SPECT and PET, where the low photon count and broad energy spectrum of gamma emissions allow enhanced tissue contrast and tumor localization. Among PCD technologies, cadmium zinc telluride (CdZnTe) stands out for its spectroscopic performance [337]. CdZnTe detectors also offer a high stopping power and can also operate at room temperature, eliminating the need for cryogenic cooling systems required by high-resolution germanium-based detectors. This makes CdZnTe an attractive material for scalable and cost-effective detector systems.

However, a major challenge with CdZnTe detectors is the inherent non-uniformity of their crystalline structure, especially in thick detector configurations [10]. Spatial vari-

ations in charge transport properties, along with localized defects, introduce complex position-dependent responses that degrade performance. To address this, we develop a method to recover material parameters and defect characteristics from the observed detector signals, effectively mitigating these intrinsic imperfections.

The forward model simulates time-dependent electrode signals in response to a gamma-ray interaction, defined by the 3D injection position (x_0, y_0, z_0), deposited energy E_d, and a set of eight key material parameters, collectively denoted as θ. For the test scenario presented in this study, we simulate an interaction at the center of the detector unit, with an injection position at $(0, 0, 0.5)$ cm in a cubic detector of 1 cm^3, and a deposited energy of 140 keV. The material properties used are as follows: electron mobility $\mu_\mathrm{e} = 1000$ cm^2/Vs and hole mobility $\mu_\mathrm{h} = 100$ cm^2/Vs; recombination lifetimes $\tau_\mathrm{eR} = 10\ \mu$s and $\tau_\mathrm{hR} = 1\ \mu$s; trapping times $\tau_{eT} = 10\ \mu$s and $\tau_\mathrm{hT} = 0.07\ \mu$s; and detrapping times $\tau_\mathrm{eD} = 0.4\ \mu$s and $\tau_\mathrm{hD} = 0.08\ \mu$s. These values represent usual operating conditions for a CdZnTe detector and serve as the baseline for inversion and optimization [284].

The charge transport is modeled using a probabilistic particle-based scheme inspired by the approach of Benoit and Hamel (2009) [310]. A representative number of elements for electrons and holes are propagated under the influence of the internal electric field, while second-order effects, diffusion and Coulomb repulsion, are incorporated through a time-evolving Gaussian approximation. The diffusion coefficient is given by the Einstein–Smoluchowski relation, $D = \mu k_B T/q$, where k_B is Boltzmann constant and q the elementary charge. The initial charge cloud formed by a gamma-ray interaction is assumed to follow a 3D Gaussian distribution with standard deviation $\sigma_0 = A\,E\,(1 - B/(1 + C\,E_\mathrm{d}))$, where E_d is the deposited energy in keV, and $A = 0.95\,\mu$m/keV, $B = 0.98$, and $C = 0.003\,\mathrm{keV}^{-1}$. For an energy deposition of 140 keV, the initial spread is approxi-

mately $\sigma_0 \approx 41\,\mu$m. The evolution of this spread is then tracked over time by evaluating $\sigma(t)$ at time steps of $\Delta t = 10$ ns, accounting for the combined broadening due to diffusion and charge repulsion [8]. Carrier fluctuation loss mechanisms throughout the path (recombination, trapping, and detrapping) are modeled using exponential decay laws derived from the Shockley–Read–Hall theory. Finally, the time-dependent current induced at each detector electrode is computed using the Shockley–Ramo theorem [281].

The model outputs the time-dependent induced currents at ten electrodes: the central pixel anode, eight neighboring anodes, and the common cathode. Given a set of observed signals, the inverse problem consists of identifying the unknown parameters $(z_0, E_\mathrm{d}, \theta)$ that best reproduce these signals via the forward model. While (x_0, y_0) are assumed to be known from careful experimental calibration, (z_0, E_d) and the eight parameters in θ are estimated within a bounded uncertainty of $\pm 20\%$ from estimated values provided by the manufacturer. We formulate this inverse problem as a minimization of the root mean square error (RMSE) between the measured and simulated signals. To tackle this complicated non-linear, and non-convex optimization problem, we employ the Slime Mold Algorithm, a recently discovered metaheuristic optimizer inspired by the behavior of the *Physarum polycephalum* [338]. Compared to other traditional heuristic optimizers, SMA has demonstrated superior global search capabilities in complex error over a large wide range of problems [339].

Our numerical experiments, based on simulated data, show promising results: The depth of interaction (z_0) is recovered with an accuracy better than 2 μm, and the energy E_d is reconstructed with an error of just 0.71%. Among all parameters, electron mobility μ_e exhibits the strongest influence on the signal shape and is recovered with high precision (0.24% error). Other parameters, while less sensitive, are also successfully estimated

within acceptable ranges (average error of 7.2%), with a hole trapping time error of 4.7%, hole mobility of 6.3%, and electron trapping of 13.2%. These results demonstrate that a small subset of parameters dominate signal formation, allowing accurate calibration even when less critical parameters are recovered with lower precision.

In conclusion, we present a robust strategy to infer average material properties and defect parameters in CdZnTe detectors using a biologically-inspired global optimization framework. Our results, although based on simulated data, validate the feasibility of signal-based detector characterization and open the door for future work with experimental measurements. This framework has the potential to enhance detector calibration and compensate for intrinsic crystal defects, ultimately improving the performance of room-temperature semiconductor detectors.

4.3.2 Convolutional neural networks for event reconstruction

Room-temperature semiconductor detectors (RTSD) are widely used for X-ray and gamma-ray detection in medical imaging and nuclear instrumentation due to their ability to operate without cooling systems while maintaining their spectroscopic capabilities. Among these, CdTe and CdZnTe detectors are of particular relevance owing to their low intrinsic noise and high atomic numbers, which provide efficient photon stopping power and good energy resolution at room temperature [290]. However, thick CdZnTe crystals are often affected by material defects that introduce charge trapping centers, degrading energy and position resolution [289], [340]. A typical detector configuration involve a planar cathode on the top surface and a matrix of anode on the bottom (such as 11×11), forming a sub-pixelated configuration (Fig. 4.13a). For signal interpretation, analysis is usually restricted to a local 3×3 subset of anodes centered around the interaction site, as signals from more distant electrodes are negligible. In this study, we consider a $1 \times 1 \times 1$ cm^3

CdZnTe crystal to investigate charge transport and signal formation.

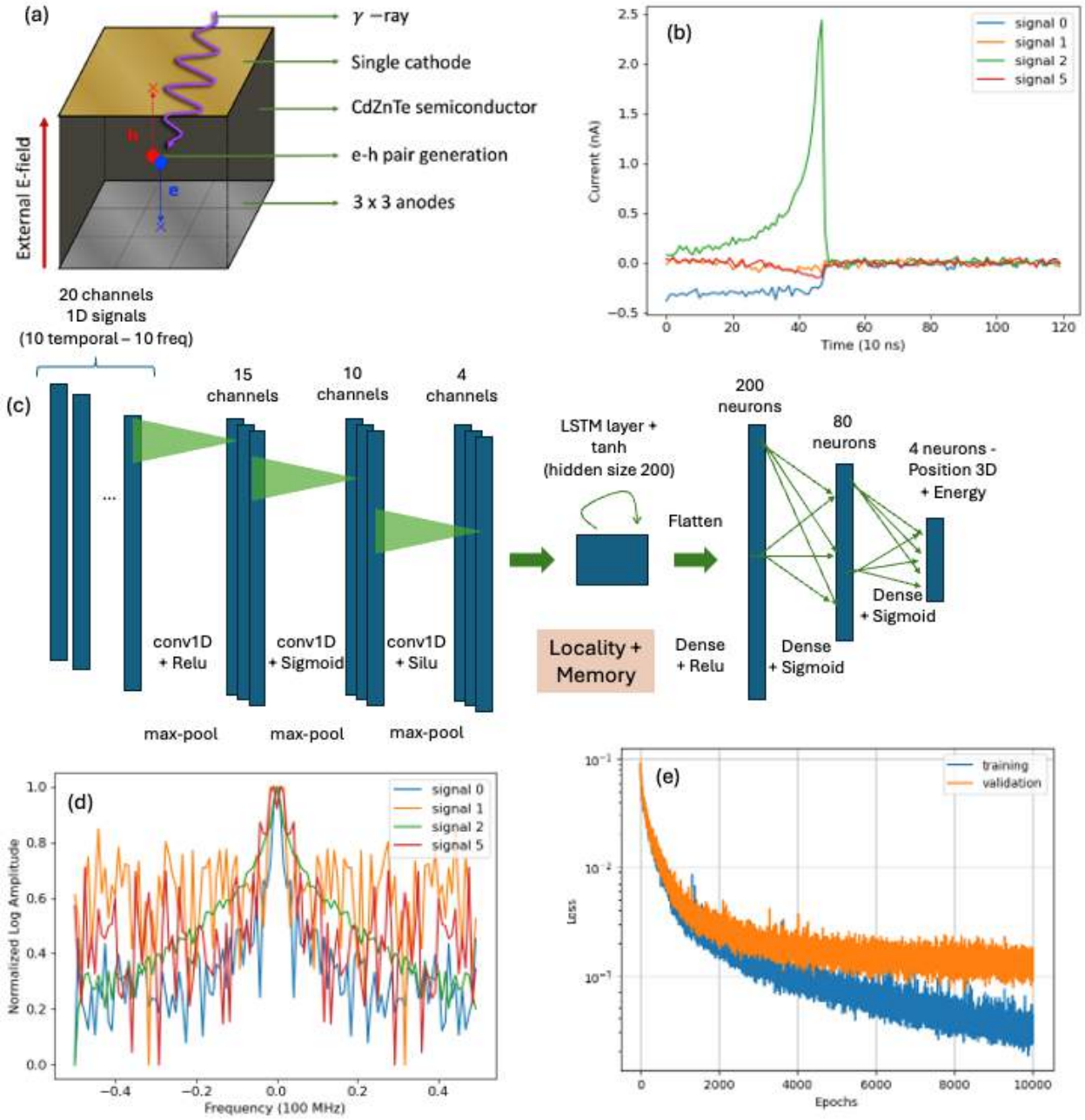


Figure 4.13: (a) Pixelated CdZnTe unit representation. (b) Signal at different electrodes with noise. (c) DL Architecture. (d) Normalized log amplitude of the signals. (e) Training history cost.

When a high-energy photon interacts within the crystal (usually through the pho-

toelectric absorption and Compton scattering), it generates a cloud with thousands of electron-hole pairs. These carriers drift under an applied electric field (-1000 V at the cathode with grounded anodes in our simulations), inducing current signals at the electrodes. According to the Shockley-Ramo theorem, the induced current depends on the carrier position and the specific weighting potential of each electrode [294]. This dimensionless potential quantifies the contribution of carrier motion at each point in the detector to the signal on each electrode. Fig. 4.13b shows the signals at different electrodes (signal 0 is for cathode and 1-9 for anodes from north-west to south-east). The sub-pixelated configuration enhances the small pixel effect [290], in which the weighting potential becomes sharply peaked near the collecting anode. As a result, the signal is dominated by the final portion of the electron trajectory, making it largely insensitive to slower hole related components and reducing the impact of hole trapping. This effect is particularly beneficial given the significantly higher electron mobility in CdZnTe ($\mu_e \approx 1000$ cm^2/Vs $\approx 10\mu_h$) and longer lifetime compared to holes.

Accurate reconstruction of the 3D interaction position and deposited energy (x, y, z, E_d) remains a intricate task nowadays, especially when accounting for material inhomogeneities, charge trapping-detrapping, and recombination. In previous work, our group developed a convolutional neural network (CNN) that infers these quantities from the temporal signals of the ten anodes surrounding the interaction [341]. Here, we extend this framework to incorporate more realistic charge transport physics and noisy signal conditions. In particular, we now introduce a probabilistic forward model in which the carrier cloud evolves as a 3D Gaussian distribution under the combined effects of electrostatic repulsion and diffusion, as detailed in [8]. This model accounts for dynamic cloud expansion in time and incorporates carrier loss mechanisms such as trapping, detrapping, and recombination, modeled through the exponential decay processes (parameterized by their

specific lifetimes). Additionally, to better characterize the noise structure, the input to the network includes not only the raw temporal signals but also the normalized logarithmic amplitude of their Fourier transforms. To capture both local and temporal signal features, we augment the CNN with a recurrent architecture based on Long Short-Term Memory (LSTM) units, enabling the network to exploit temporal dependencies without suffering from vanishing gradient issues. This hybrid CNN-RNN structure (Fig. 4.13c) combines the strengths of spatial and sequential learning and has shown strong performance in related image processing applications [342].

The network was trained on 5000 simulated events with random interaction locations and energies between 80 and 160 keV. Signal were considered with additive Gaussian white noise at levels up to 0.1% of the maximum amplitude (see Fig. 4.13b). The dataset was split 85/15/15% for training, validation, and testing, respectively. Training was performed on a T4 GPU using the ADAM optimizer with a learning rate of 10^{-3} and momentum parameters $\beta = (0.9, 0.999)$ in a PyTorch framework, during 10000 epochs (Fig. 4.13e). The model achieved a mean absolute position error of 225 μm and an energy resolution error of 2.42%.

In summary, we present a significant enhancement to event reconstruction in CdZnTe detectors by integrating the physically informed 3D charge cloud expansion, the frequency-domain signal representation, and the newly proposed hybrid CNN-RNN deep learning architecture. The resulting system achieves sub-millimeter spatial resolution and accurate energy estimation under realistic noise and defect conditions, advancing the capabilities of RTSDs for precise imaging and spectroscopy.

References

[1] M. Ballester, H. Wang, J. Li, O. Cossairt, and F. Willomitzer, "Single-shot synthetic wavelength imaging: Sub-mm precision tof sensing with conventional cmos sensors," *Optics and Lasers in Engineering*, vol. 178, p. 108 165, 2024.

[2] M. Ballester, H. Wang, O. Cossairt, and F. Willomitzer, "Dynamic high-precision 3d measurements with an interferometric single-shot tof camera," in *Computational Optical Sensing and Imaging*, Optica Publishing Group, 2023, CTh2B–2.

[3] M. Balleste, *Dynamic High-Precision ToF Imaging utilizing Single-Shot Synthetic Wavelength Interferometry on Rough Surfaces*. Deutsche Gesellschaft für angewandte Optik (DGaO), 2023.

[4] M. Ballester, E. Marquez, J. Bass, C. Wuersch, F. Willomitzer, and A. K. Katsaggelos, "Review and novel formulae for transmittance and reflectance of wedged thin films on absorbing substrates," *arXiv preprint arXiv:2409.02323*, 2024.

[5] M. Ballester, A. P. Marquez, S. Lopez-Tapia, *et al.*, "Enhancing the swanepoel method: Precise envelope detection of thin-film transmission spectra," *Optics Express*, vol. 33, no. 6, pp. 13 376–13 400, 2025.

[6] M. Ballester, C. Würsch, E. Marquez, F. Willomitzer, and A. Katsaggelos, "Optical characterization of thin films from transmission data using deep learning," in *Optical Sensors*, Optica Publishing Group, 2024, FD1–1.

[7] M. Ballester, E. Marquez, A. P. Marquez, *et al.*, "Deep learning method to find swanepoel envelopes in transmission spectroscopy," in *Computational Optical Sensing and Imaging*, Optica Publishing Group, 2023, CTh2A–5.

[8] M. Ballester, J. Kaspar, F. Massanés, A. H. Vija, and A. K. Katsaggelos, "Charge diffusion and repulsion in semiconductor detectors," *Sensors*, vol. 24, no. 22, p. 7123, 2024.

[9] M. Ballester, J. Kaspar, F. Massanés, A. H. Vija, and A. K. Katsaggelos, "Simulating diffusion and repulsion of charges in single photon semiconductor detectors," in *Laser Applications to Chemical, Security and Environmental Analysis*, Optica Publishing Group, 2024, FD1–6.

[10] M. Ballester, J. Kaspar, F. Massanes, S. Banerjee, A. H. Vija, and A. K. Katsaggelos, "Characterization of crystal properties and defects in cdznte radiation detectors," *arXiv preprint arXiv:2409.06738*, 2024.

[11] M. Ballester, Z. Ho, A. Odate, M. Walton, and A. Katsaggelos, "Simulation of polarized light microscopy for multiple analyzer angles," in *Preprints*, 2024.

[12] M. Ballester, F. Schiffers, Z. Wang, *et al.*, "Fast simulations in computer-generated holograms for binary data storage," in *Computational Optical Sensing and Imaging*, Optica Publishing Group, 2021, CTh4A–7.

[13] E. Márquez, E. Blanco, J. M. Mánuel, *et al.*, "Mid-infrared (mir) complex refractive index spectra of polycrystalline copper-nitride films by ir-vase ellipsometry and their fib-sem porosity," *Coatings*, vol. 14, no. 1, p. 5, 2023.

[14] E. Márquez, E. Blanco, M. García-Gurrea, *et al.*, "Optical properties of reactive rf magnetron sputtered polycrystalline cu_3n thin films determined by uv/visible/nir spectroscopic ellipsometry: An eco-friendly solar light absorber," *Coatings*, vol. 13, no. 7, p. 1148, 2023.

[15] V. Srinivasan, H.-C. Liu, and M. Halioua, "Automated phase-measuring profilometry of 3-d diffuse objects," *Applied optics*, vol. 23, no. 18, pp. 3105–3108, 1984.

[16] M. Schaffer, M. Grosse, and R. Kowarschik, "High-speed pattern projection for three-dimensional shape measurement using laser speckles," *Applied optics*, vol. 49, no. 18, pp. 3622–3629, 2010.

[17] M. Takeda and K. Mutoh, "Fourier transform profilometry for the automatic measurement of 3-d object shapes," *Applied optics*, vol. 22, no. 24, pp. 3977–3982, 1983.

[18] J. L. Schonberger and J.-M. Frahm, "Structure-from-motion revisited," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 4104–4113.

[19] F. Willomitzer and G. Hausler, "Single-shot 3d motion picture camera with a dense point cloud," *Optics express*, vol. 25, no. 19, pp. 23 451–23 464, 2017.

[20] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical engineering*, vol. 19, no. 1, pp. 139–144, 1980.

[21] B. K. Horn, "Height and gradient from shading," *International journal of computer vision*, vol. 5, no. 1, pp. 37–75, 1990.

[22] F. Willomitzer, C.-K. Yeh, V. Gupta, *et al.*, "Hand-guided qualitative deflectometry with a mobile device," *Optics express*, vol. 28, no. 7, pp. 9027–9038, 2020.

[23] C. Faber, E. Olesch, R. Krobot, and G. Hausler, "Deflectometry challenges interferometry: The competition gets tougher!" In *Interferometry XVI: techniques and analysis*, SPIE, vol. 8493, 2012, pp. 232–246.

[24] J. Wang, B. Xu, T. Wang, *et al.*, "Vr eye-tracking using deflectometry," in *Computational Optical Sensing and Imaging*, Optical Society of America, 2021, CF2E–3.

[25] L. Huang, M. Idir, C. Zuo, and A. Asundi, "Review of phase measuring deflectometry," *Optics and Lasers in Engineering*, vol. 107, pp. 247–257, 2018.

[26] F. Willomitzer, "Synthetic wavelength imaging–utilizing spectral correlations for high-precision time-of-flight sensing," *arXiv preprint arXiv:2209.04941*, 2022.

[27] G. Hausler and F. Willomitzer, "Reflections about the holographic and non-holographic acquisition of surface topography: Where are the limits?" *Light: Advanced Manufacturing*, vol. 3, no. 2, pp. 1–10, 2022.

[28] F. Heide, L. Xiao, W. Heidrich, and M. B. Hullin, "Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3222–3229.

[29] R. Collis, "Lidar," *Applied optics*, vol. 9, no. 8, pp. 1782–1788, 1970.

[30] C. Weitkamp, *Lidar: range-resolved optical remote sensing of the atmosphere*. Springer Science & Business, 2006, vol. 102.

[31] R. Schwarte, Z. Xu, H.-G. Heinol, *et al.*, "New electro-optical mixing and correlating sensor: Facilities and applications of the photonic mixer device (pmd)," in *Sensors, Sensor Systems, and Sensor Data Processing*, SPIE, vol. 3100, 1997, pp. 245–253.

[32] R. Lange and P. Seitz, "Solid-state time-of-flight range camera," *IEEE Journal of quantum electronics*, vol. 37, no. 3, pp. 390–397, 2001.

[33] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (tof) cameras: A survey," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 1917–1926, 2011.

[34] P. Hariharan, *Optical Interferometry, 2e*. Elsevier, 2003.

[35] J. C. Wyant, "Interferometric optical metrology," *OSA Century of Optics*, 2015.

[36] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Elsevier, 2013.

[37] A. F. Fercher, H. Z. Hu, and U. Vry, "Rough surface interferometry with a two-wavelength heterodyne speckle interferometer," *Applied optics*, vol. 24, no. 14, pp. 2181–2188, 1985.

[38] U. Vry and A. Fercher, "Higher-order statistical properties of speckle fields and their application to rough-surface interferometry," *JOSA A*, vol. 3, no. 7, pp. 988–1000, 1986.

[39] R. Dandliker, R. Thalmann, and D. Prongue, "Two-wavelength laser interferometry using superheterodyne detection," *Optics letters*, vol. 13, no. 5, pp. 339–341, 1988.

[40] R. Dandliker, Y. Salvade, and E. Zimmermann, "Distance measurement by multiple-wavelength interferometry," *Journal of Optics*, vol. 29, no. 3, p. 105, 1998.

[41] P. De Groot, "Interferometric laser profilometer for rough surfaces," *Optics letters*, vol. 16, no. 6, pp. 357–359, 1991.

[42] P. J. De Groot, "A review of selected topics in interferometric optical metrology," *Reports on Progress in Physics*, vol. 82, no. 5, p. 056 101, 2019.

[43] F. Li, F. Willomitzer, P. Rangarajan, M. Gupta, A. Velten, and O. Cossairt, "Sh-tof: Micro resolution time-of-flight imaging with superheterodyne interferometry," in *2018 IEEE International Conference on Computational Photography (ICCP)*, IEEE, 2018, pp. 1–10.

[44] F. Li, F. Willomitzer, M. M. Balaji, P. Rangarajan, and O. Cossairt, "Exploiting wavelength diversity for high resolution time-of-flight 3d imaging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 7, pp. 2193–2205, 2021.

[45] Y. Wu, F. Li, F. Willomitzer, A. Veeraraghavan, and O. Cossairt, "Wished: Wavefront imaging sensor with high resolution and depth ranging," in *2020 IEEE International Conference on Computational Photography (ICCP)*, IEEE, 2020, pp. 1–10.

[46] A. Kotwal, A. Levin, and I. Gkioulekas, "Swept-angle synthetic wavelength interferometry," *arXiv preprint arXiv:2205.10655*, 2022.

[47] M. Perenzoni, L. Pancheri, and D. Stoppa, "Compact spad-based pixel architectures for time-resolved image sensors," *Sensors*, vol. 16, no. 5, p. 745, 2016.

[48] T. Spirig, P. Seitz, O. Vietze, and F. Heitger, "The lock-in ccd-two-dimensional synchronous detection of light," *IEEE Journal of quantum electronics*, vol. 31, no. 9, pp. 1705–1708, 1995.

[49] H. Kim, J. Park, I. Joe, *et al.*, "5.6 a 1/2.65 in 44mpixel cmos image sensor with 0.7 μm pixels fabricated in advanced full-depth deep-trench isolation technology," in *2020 IEEE International Solid-State Circuits Conference-(ISSCC)*, IEEE, 2020, pp. 104–106.

[50] J. M. Huntley and H. Saldner, "Temporal phase-unwrapping algorithm for automated interferogram analysis," *Applied Optics*, vol. 32, no. 17, pp. 3047–3052, 1993.

[51] S. Su, F. Heide, G. Wetzstein, and W. Heidrich, "Deep end-to-end time-of-flight imaging," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6383–6392.

[52] S.-H. Baek, N. Walsh, I. Chugunov, Z. Shi, and F. Heide, "Centimeter-wave free-space neural time-of-flight imaging," *ACM Transactions on Graphics (TOG)*, 2022.

[53] D. Droeschel, D. Holz, and S. Behnke, "Multi-frequency phase unwrapping for time-of-flight cameras," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2010, pp. 1463–1469.

[54] V. Pascazio and G. Schirinzi, "Multifrequency insar height reconstruction through maximum likelihood estimation of local planes parameters," *IEEE Transactions on Image Processing*, vol. 11, no. 12, pp. 1478–1489, 2002.

[55] F. Willomitzer, P. V. Rangarajan, F. Li, M. M. Balaji, M. P. Christensen, and O. Cossairt, "Fast non-line-of-sight imaging with high-resolution and wide field of view using synthetic wavelength holography," *Nature communications*, vol. 12, no. 1, pp. 1–11, 2021.

[56] F. Willomitzer, P. V. Rangarajan, F. Li, M. M. Balaji, M. P. Christensen, and O. Cossairt, "Synthetic wavelength holography: An extension of gabor's holographic principle to imaging with scattered wavefronts," *arXiv preprint arXiv:1912.11438*, 2019.

[57] M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," *JosA*, vol. 72, no. 1, pp. 156–160, 1982.

[58] S. Ri, N. Agarwal, Q. Wang, and Q. Kemao, "Comparative study of sampling moire and windowed fourier transform techniques for demodulation of a single-fringe pattern," *Applied Optics*, vol. 57, no. 36, pp. 10 402–10 411, 2018.

[59] H. Tiziani, A. Rothe, and N. Maier, "Dual-wavelength heterodyne differential interferometer for high-precision measurements of reflective aspherical surfaces and step heights," *Applied optics*, vol. 35, no. 19, pp. 3525–3533, 1996.

[60] P. J. De Groot, "Extending the unambiguous range of two-color interferometers," *Applied optics*, vol. 33, no. 25, pp. 5948–5953, 1994.

[61] T. Tahara and Y. Arai, "Multiwavelength off-axis digital holography with an angle of more than 40 degrees and no beam combiner to generate interference light," *Applied optics*, vol. 56, no. 13, F200–F204, 2017.

[62] P. Bergstrom, D. Khodadad, E. Hallstig, and M. Sjödahl, "Dual-wavelength digital holography: Single-shot shape evaluation using speckle displacements and regularization," *Applied Optics*, vol. 53, no. 1, pp. 123–131, 2014.

[63] D. Khodadad, E. Hallstig, and M. Sjodahl, "Shape reconstruction using dual wavelength digital holography and speckle movements," in *Optical Measurement Systems for Industrial Inspection VIII*, SPIE, vol. 8788, 2013, pp. 106–116.

[64] D. Khodadad, P. Bergstrom, E. Hallstig, and M. Sjodahl, "Single shot dual-wavelength digital holography: Calibration based on speckle displacements," *International Journal of Optomechatronics*, vol. 8, no. 4, pp. 326–339, 2014.

[65] A. Groger, G. Pedrini, F. Fischer, D. Claus, I. Aleksenko, and S. Reichelt, "Two-wavelength digital holography through fog," *Journal of the European Optical Society-Rapid Publications*, vol. 19, no. 1, p. 25, 2023.

[66] W. Osten, A. Faridian, P. Gao, *et al.*, "Recent advances in digital holography," *Applied optics*, vol. 53, no. 27, G44–G63, 2014.

[67] G. Pedrini, W. Osten, and M. E. Gusev, "High-speed digital holographic interferometry for vibration measurement," *Applied optics*, vol. 45, no. 15, pp. 3456–3462, 2006.

[68] B. Javidi, A. Carnicer, A. Anand, *et al.*, "Roadmap on digital holography," *Optics express*, vol. 29, no. 22, pp. 35 078–35 118, 2021.

[69] E. Hecht, *Optics*. Pearson Education India, 2012.

[70] W. Xu, E. C. Chang, L. K. Kwoh, H. Lim, W. Cheng, and A. Heng, "Phase-unwrapping of sar interferogram with multi-frequency or multi-baseline," in *Proceedings of IGARSS'94-1994 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, vol. 2, 1994, pp. 730–732.

[71] M. Gupta, S. K. Nayar, M. B. Hullin, and J. Martin, "Phasor imaging: A generalization of correlation-based time-of-flight imaging," *ACM Transactions on Graphics (ToG)*, vol. 34, no. 5, pp. 1–18, 2015.

[72] M. A. Herraez, D. R. Burton, M. J. Lalor, and M. A. Gdeisat, "Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path," *Applied optics*, vol. 41, no. 35, pp. 7437–7444, 2002.

[73] *"conventional" (non nn-based) unwrapping algorithm predefined in the python library scikit-image*, `https://scikit-image.org/docs/stable/auto_examples/filters/plot_phase_unwrap.html`, Accessed: 2022-11-28.

[74] *3d video of the dynamic object (metronome)*, `https://drive.google.com/drive/folders/1hj8g9Od3LE1fBSYYNflO5gBS-jgm-8GS`, Accessed: 2022-12-01.

[75] M. V. Perera and A. De Silva, "A joint convolutional and spatial quad-directional lstm network for phase unwrapping," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2021, pp. 4055–4059.

[76] G. Spoorthi, R. K. S. S. Gorthi, and S. Gorthi, "Phasenet 2.0: Phase unwrapping of noisy data based on deep learning approach," *IEEE Transactions on Image Processing*, vol. 29, pp. 4862–4872, 2020.

[77] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.

[78] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[79] K. Wang, Q. Kemao, J. Di, and J. Zhao, "Deep learning spatial phase unwrapping: A comparative review," *Advanced Photonics Nexus*, vol. 1, no. 1, pp. 014 001–014 001, 2022.

[80] A. Corti, S. Giancola, G. Mainetti, and R. Sala, "A metrological characterization of the kinect v2 time-of-flight camera," *Robotics and Autonomous Systems*, vol. 75, pp. 584–594, 2016.

[81] P. Cornwall, M. Ballester, H. Wang, and F. Willomitzer, "Towards synthetic light-in-flight," in *Computational Optical Sensing and Imaging*, Optica Publishing Group, 2023, CTh2B–3.

[82] P. Cornwall, M. Ballester, S. Forschner, M. M. Balaji, A. Katsagelos, and F. Willomitzer, "Computational light-in-flight imaging using synthetic waves," in *Computational Optical Imaging and Artificial Intelligence in Biomedical Sciences II*, SPIE, vol. 13333, 2025, pp. 12–17.

[83] P. Cornwall, S. Forschner, K. Kassem, *et al.*, "Single-shot synthetic wavelength imaging through scattering media," in *Laser Science*, Optica Publishing Group, 2024, JTu7A–1.

[84] P. Cornwall, M. M. Balaji, P. Liu, *et al.*, "Single-shot imaging through scatter or turbulence using synthetic waves," in *Three-Dimensional Imaging, Visualization, and Display 2025*, SPIE, vol. 13465, 2025, 134650O.

[85] S. Forschner, P. Cornwall, M. Ballester, M. Madabhushi Balaji, J. Czarske, and F. Willomitzer, "Towards synthetic wavelength imaging through multi-mode fibers," in *Proc. of SPIE Vol*, vol. 13258, 2024, pp. 1 325 801–212.

[86] E. Márquez, J. J. Ruíz-Pérez, M. Ballester, *et al.*, "Optical characterization of h-free a-si layers grown by rf-magnetron sputtering by inverse synthesis using matlab: Tauc–lorentz–urbach parameterization," *Coatings*, vol. 11, no. 11, p. 1324, 2021.

[87] M. Birkett, "Optical properties of earth-abundant semiconductors for renewable energy," Ph.D. dissertation, The University of Liverpool (United Kingdom), 2016.

[88] M. I. Rodríguez-Tapiador, J. Merino, T. Jawhari, A. L. Muñoz-Rosas, J. Bertomeu, and S. Fernández, "Power effect on the properties of copper nitride films as solar

absorber deposited in pure nitrogen atmosphere," *Applied Research*, vol. 3, no. 1, e202200105, 2024.

[89] A. J. Lopez-Garcia, G. Alvarez-Suarez, E. Ros, *et al.*, "Enhanced selective contact behavior in a-si: H/oxide transparent photovoltaic devices via dipole layer integration," *Solar RRL*, vol. 8, no. 14, p. 2 400 276, 2024.

[90] E. Márquez, J. Ramirez-Malo, P. Villares, R. Jimenez-Garay, P. Ewen, and A. Owen, "Calculation of the thickness and optical constants of amorphous arsenic sulphide films from their transmission spectra," *Journal of Physics D: Applied Physics*, vol. 25, no. 3, p. 535, 1992.

[91] E. Márquez, E. Saugar, J. Díaz, *et al.*, "The influence of ar pressure on the structure and optical properties of non-hydrogenated a-si thin films grown by rf magnetron sputtering onto room-temperature glass substrates," *Journal of Non-Crystalline Solids*, vol. 517, pp. 32–43, 2019.

[92] M. I. Rodríguez-Tapiador, J. Merino, T. Jawhari, A. Muñoz-Rosas, J. Bertomeu, and S. Fernández, "Impact of the rf power on the copper nitride films deposited in a pure nitrogen environment for applications as eco-friendly solar absorber," *Materials*, vol. 16, no. 4, p. 1508, 2023.

[93] R. M. Azzam, N. M. Bashara, and D. T. Burns, "Ellipsometry and polarized light," *Analytica Chimica Acta*, vol. 199, pp. 283–284, 1987.

[94] H. G. Tompkins and J. N. Hilfiker, *Spectroscopic ellipsometry: practical application to thin film characterization*. Momentum Press, 2015.

[95] G. Bader, P. Ashrit, and V.-V. Truong, "Transmission and reflection ellipsometry of thin films and multilayer systems," *Applied optics*, vol. 37, no. 7, pp. 1146–1151, 1998.

[96] P. Petrik, T. Lohner, L. Égerházi, and Z. Geretovszky, "Optical models for the ellipsometric characterization of carbon nitride layers prepared by inverse pulsed laser deposition," *Applied surface science*, vol. 253, no. 1, pp. 173–176, 2006.

[97] S. Jena and N. Sahoo, "Evolutionary design, deposition and characterization techniques for interference optical thin-film multilayer coatings and devices," in *Recent Advances in Thin Films*, Springer, 2020, pp. 281–343.

[98] D. D. Engelsen, "Ellipsometry of anisotropic films," *JOSA*, vol. 61, no. 11, pp. 1460–1466, 1971.

[99] B. Fodor, T. Defforge, E. Agócs, M. Fried, G. Gautier, and P. Petrik, "Spectroscopic ellipsometry of columnar porous si thin films and si nanowires," *Applied Surface Science*, vol. 421, pp. 397–404, 2017.

[100] E. Marquez, M. Ballester, M. Garcia, *et al.*, "Complex dielectric function of h-free a-si films: Photovoltaic light absorber," *Materials Letters*, vol. 345, p. 134 485, 2023.

[101] M. Gore, *Spectrophotometry and spectrofluorimetry: a practical approach*. OUP Oxford, 2000, vol. 225.

[102] S. Bosch, "Spectrophotometry, ellipsometry, and computer simulation in thin film developments," in *ROMOPTO'97: Fifth Conference on Optics*, vol. 3405, SPIE, 1998, pp. 1120–1131.

[103] E. Marceau, C. Perier, F. L. Ulis, and T. Burt, *Measurement of diffuse reflection from catalyst powders using the praying mantis accessory and agilent cary 5000 uv-vis-nir spectrophotometer*.

[104] D. E. Aspnes, "Optical properties of thin films," *Thin solid films*, vol. 89, no. 3, pp. 249–262, 1982.

[105] D. Stroud, "The effective medium approximations: Some recent developments," *Superlattices and microstructures*, vol. 23, no. 3-4, pp. 567–573, 1998.

[106] S. Adachi and H. Mori, "Optical properties of fully amorphous silicon," *Physical Review B*, vol. 62, no. 15, p. 10 158, 2000.

[107] D. Lehmann, F. Seidel, and D. R. Zahn, "Thin films with high surface roughness: Thickness and dielectric function analysis using spectroscopic ellipsometry," *SpringerPlus*, vol. 3, pp. 1–8, 2014.

[108] L. Rolland, C. Vallée, M.-C. Peignon, and C. Cardinaud, "Roughness and chemistry of silicon and polysilicon surfaces etched in high-density plasma: Xps, afm and ellipsometry analysis," *Applied surface science*, vol. 164, no. 1-4, pp. 147–155, 2000.

[109] X. Zhou, S. Arakawa, S. Tunmee, *et al.*, "Structural analysis of amorphous carbon films by bema theory based on spectroscopic ellipsometry measurement," *Diamond and Related Materials*, vol. 79, pp. 46–59, 2017.

[110] J. A. Pradeep and P. Agarwal, "Determination of thickness, refractive index, and spectral scattering of an inhomogeneous thin film with rough interfaces," *Journal of Applied Physics*, vol. 108, no. 4, 2010.

[111] M. Ohlídal, I. Ohlídal, P. Klapetek, D. Nečas, and A. Majumdar, "Measurement of the thickness distribution and optical constants of non-uniform thin films," *Measurement Science and Technology*, vol. 22, no. 8, p. 085 104, 2011.

[112] M.-C. López-Santos, R. Alvarez, A. Palmero, *et al.*, "Micron-scale wedge thin films prepared by plasma enhanced chemical vapor deposition," *Plasma Processes and Polymers*, vol. 14, no. 12, p. 1 700 043, 2017.

[113] H. Shimizu, T. Ogura, T. Maeda, and S. Suzuki, "A wedge-shaped au thin film: Integrating multiple surface plasmon resonance sensors in a single chip and enhancing the figure of merit," *Frontiers in Nanotechnology*, vol. 3, p. 724 528, 2021.

[114] G. V. Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009, ISBN: 1441412697.

[115] *MATLAB version 9.10.0.1613233 (R2021a)*, The Mathworks, Inc., Natick, Massachusetts, 2021.

[116] R. D. Simoni, R. L. Hill, M. Vaughan, and H. Tabor, "A classic instrument: The beckman du spectrophotometer and its inventor, arnold o. beckman," *Journal of Biological Chemistry*, vol. 278, no. 49, pp. 79–81, 2003.

[117] K. Vedam, "Spectroscopic ellipsometry: A historical overview," *Thin Solid Films*, vol. 313, pp. 1–9, 1998.

[118] M. A. Butt, "Thin-film coating methods: A successful marriage of high-quality and cost-effectiveness—a brief exploration," *Coatings*, vol. 12, no. 8, p. 1115, 2022.

[119] J. Dobrowolski, F. Ho, and A. Waldorf, "Determination of optical constants of thin film coating materials based on inverse synthesis," *Applied Optics*, vol. 22, no. 20, pp. 3191–3200, 1983.

[120] T. Buffeteau and B. Desbat, "Thin-film optical constants determined from infrared reflectance and transmittance measurements," *Applied Spectroscopy*, vol. 43, no. 6, pp. 1027–1032, 1989.

[121] N. Kamble, R. Tokas, A. Biswas, S. Thakur, D. Bhattacharyya, and N. Sahoo, "Determination of the optical constants of hfo_2–sio_2 composite thin films through reverse fitting of transmission spectra," *Vacuum*, vol. 86, no. 4, pp. 422–428, 2011.

[122] L. Gao, F. Lemarchand, and M. Lequime, "Reverse engineering from spectrophotometric measurements: Performances and efficiency of different optimization algorithms," *Applied Physics A*, vol. 108, pp. 877–889, 2012.

[123] M. M. A.-G. Jafar, M. H. Saleh, M. J. A. Ahmad, B. N. Bulos, and T. M. Al-Daraghmeh, "Retrieval of optical constants of undoped amorphous selenium films from an analysis of their normal-incidence transmittance spectra using numeric puma method," *Journal of Materials Science: Materials in Electronics*, vol. 27, pp. 3281–3291, 2016.

[124] A. Tejada, L. Montañez, C. Torres, *et al.*, "Determination of the fundamental absorption and optical bandgap of dielectric thin films from single optical transmittance measurements," *Applied Optics*, vol. 58, no. 35, pp. 9585–9594, 2019.

[125] J. C. Alonso-Huitron and A. Garcia-Valenzuela, *A comprehensive model for the optical transmission for determining the optimal thickness and figure of merit of al-doped zno films as transparent conducting coatings*, Unpublished or in press, 2024.

[126] J. F. Hall and W. Ferguson, "Optical properties of cadmium sulfide and zinc sulfide from 0.6 micron to 14 microns," *JOSA*, vol. 45, no. 9, pp. 714–718, 1955.

[127] S. Lyashenko and V. Miloslavsky, "A simple method for semiconductor and dielectric layers thickness and optical constants determining," *Optika i Spektroskopiya*, vol. 16, p. 151, 1964.

[128] J. Manifacier, J. Gasiot, and J. Fillard, "A simple method for the determination of the optical constants n, k and the thickness of a weakly absorbing thin film," *Journal of Physics E: Scientific Instruments*, vol. 9, no. 11, p. 1002, 1976.

[129] R. Swanepoel, "Determination of the thickness and optical constants of amorphous silicon," *Journal of Physics E: Scientific Instruments*, vol. 16, no. 12, p. 1214, 1983.

[130] R. Swanepoel, "Determination of surface roughness and optical constants of inhomogeneous amorphous silicon films," *Journal of Physics E: Scientific Instruments*, vol. 17, no. 10, p. 896, 1984.

[131] D. Kushev, N. Zheleva, Y. Demakopoulou, and D. Siapkas, "A new method for the determination of the thickness, the optical constants and the relaxation time of weakly absorbing semiconducting thin films," *Infrared Physics*, vol. 26, no. 6, pp. 385–393, 1986.

[132] D. Minkov, "Calculation of the optical constants of a thin layer upon a transparent substrate from the reflection spectrum," *Journal of Physics D: Applied Physics*, vol. 22, no. 8, p. 1157, 1989.

[133] D. Minkov, "Method for determining the optical constants of a thin film on a transparent substrate," *Journal of Physics D: Applied Physics*, vol. 22, no. 1, p. 199, 1989.

[134] D. Minkov, "Computation of the optical constants of a thin dielectric layer on a transmitting substrate from the reflection spectrum at inclined incidence of light," *JOSA A*, vol. 8, no. 2, pp. 306–310, 1991.

[135] D. Minkov, "Errors made in the computation of the optical constants of a thin dielectric layer from the envelopes of the reflection spectrum at inclined incidence of the light," *Optik (Stuttgart)*, vol. 87, no. 4, pp. 137–140, 1991.

[136] D. Minkov and R. Swanepoel, "Comparative study of the use of the matrix approach and the flow-graph approach for optical analysis of isotropic stratified planar structures," in *Current Developments in Optical Design and Engineering V*, vol. 2540, SPIE, 1995, pp. 131–138.

[137] E. Marquez, J. Ramirez-Malo, P. Villares, R. Jimenez-Garay, and R. Swanepoel, "Optical characterization of wedge-shaped thin films of amorphous arsenic trisulphide based only on their shrunk transmission spectra," *Thin Solid Films*, vol. 254, no. 1-2, pp. 83–91, 1995.

[138] J. Gonzalez-Leal, E. Marquez, A. Bernal-Oliva, J. Ruiz-Perez, and R. Jimenez-Garay, "Derivation of the optical constants of thermally-evaporated uniform films of binary chalcogenide glasses using only their reflection spectra," *Thin Solid Films*, vol. 317, no. 1-2, pp. 223–227, 1998.

[139] J. Ruiz-Pérez, E. Márquez, J. González-Leal, D. Jiménez-Garay, and D. Minkov, "A new method for the optical characterization of inhomogeneous thin films based only on spectroscopic reflection measurements," in *Materials Science Forum*, vol. 287, Trans Tech Publ, 1998, pp. 363–366.

[140] B. Richards, A. Lambertz, and A. Sproul, "Determination of the optical properties of non-uniformly thick non-hydrogenated sputtered silicon thin films on glass," *Thin Solid Films*, vol. 460, no. 1-2, pp. 247–255, 2004.

[141] E. Shaaban, I. Yahia, and E. El-Metwally, "Validity of swanepoel's method for calculating the optical constants of thick films," *Acta Physica Polonica A*, vol. 121, no. 3, pp. 628–635, 2012.

[142] D. Dorranian, L. Dejam, and G. Mosayebian, "Optical characterization of cu_3n thin film with swanepoel method," *Journal of Theoretical and Applied Physics*, vol. 6, pp. 1–9, 2012.

[143] Y. Jin, B. Song, Z. Jia, *et al.*, "Improvement of swanepoel method for deriving the thickness and the optical properties of chalcogenide thin films," *Optics express*, vol. 25, no. 1, pp. 440–451, 2017.

[144] D. Minkov, G. Angelov, E. Marquez, *et al.*, "Increasing the accuracy of the characterization of a thin semiconductor or dielectric film on a substrate from only one quasi-normal incidence uv/vis/nir reflectance spectrum of the sample," *Nanomaterials*, vol. 13, no. 17, p. 2407, 2023.

[145] J. Ruíz-Pérez, J. González-Leal, D. Minkov, and E. Márquez, "Method for determining the optical constants of thin dielectric films with variable thickness using only their shrunk reflection spectra," *Journal of Physics D: Applied Physics*, vol. 34, no. 16, p. 2489, 2001.

[146] V. Filippov, "Method of the ratio of envelopes of the reflection spectrum for measuring optical constants and thickness of thin films," *Optics and Spectroscopy*, vol. 88, pp. 581–585, 2000.

[147] H. Liu, D. Hou, Z. Wang, Y. Ji, Y. Fan, and R. Fan, "Comparison of envelope method and full spectra fitting method for determination of optical constants of thin films," in *Seventh International Conference on Thin Film Physics and Applications*, vol. 7995, SPIE, 2011, pp. 381–384.

[148] M. Ballester, A. P. Márquez, S. Banerjee, *et al.*, "Comparison of optical characterization methods for transmission spectroscopy," in *Computational Optical Sensing and Imaging*, Optica Publishing Group, 2022, JW5D–4.

[149] Z. Wang, Y. C. Lin, K. Zhang, W. Wu, and S. Huang, "Measuring complex refractive index through deep-learning-enabled optical reflectometry," *2D Materials*, vol. 10, no. 2, p. 025 025, 2023.

[150] J. M. Bass, M. Ballester, S. M. Fernández, A. K. Katsaggelos, E. Márquez, and F. Willomitzer, "An angular spectrum approach to inverse synthesis for the characterization of optical and geometrical properties of semiconductor thin films," *arXiv preprint arXiv:2407.06535*, 2024.

[151] J. M. Bass, M. Ballester, S. M. Fernández, A. K. Katsaggelos, E. Márquez, and F. Willomitzer, "Characterizing optical and geometrical properties of semiconductor

thin films with a split-step angular spectrum approach to inverse synthesis," in *Photonic Instrumentation Engineering XII*, SPIE, vol. 13373, 2025, pp. 65–68.

[152] J. M. Bass, M. Ballester, S. M. Fernández, A. K. Katsaggelos, E. Márquez, and F. Willomitzer, "Increasing the precision of transmission spectroscopy by optimization of thin film surface shapes," in *Laser Science*, Optica Publishing Group, 2024, JTu4A–42.

[153] A. Crook, "The reflection and transmission of light by any system of parallel isotropic films," *JOSA*, vol. 38, no. 11, pp. 954–964, 1948.

[154] F. Abelès, "La détermination de l'indice et de l'épaisseur des couches minces transparentes," *J. Phys. Radium*, vol. 11, no. 7, pp. 310–314, 1950.

[155] F. Abeles, "Facteurs de réflexion et de transmission des couches métalliques très minces: Méthode nouvelle pour déterminer leurs indices et leurs épaisseurs," *Rev. Opt., Theor. Instrum.*, vol. 32, pp. 257–268, 1953.

[156] F. Abeles, "Optical properties of thin absorbing films," *JOSA*, vol. 47, no. 6, pp. 473–482, 1957.

[157] P. W. Mullen, "Optical properties of thin solid films," *Nature*, vol. 177, 1956.

[158] O. S. Heavens, *Optical Properties of Thin Solid Films*. Courier Corporation, 1991.

[159] J. Nestell and R. Christy, "Derivation of optical constants of metals from thin-film measurements at oblique incidence," *Applied Optics*, vol. 11, no. 3, pp. 643–651, 1972.

[160] J. J. R. Pérez, "Nuevos métodos de caracterización óptica de semiconductores basados en medidas espectroscópicas de reflexión," Ph.D. dissertation, Universidad de Cádiz, 1997.

[161] M. Cesaria, A. P. Caricato, and M. Martino, "Realistic reflectance spectrum of thin films covering a transparent optically thick substrate," *Applied Physics Letters*, vol. 105, no. 3, 2014.

[162] E. Potapov and A. Rakov, "Determination of the dispersion of the optical constants of a thin absorbing film on a transparent or weakly absorbing substrate," *Journal of Applied Spectroscopy*, vol. 14, no. 1, pp. 110–113, 1971.

[163] L. Vriens and W. Rippens, "Optical constants of absorbing thin solid films on a substrate," *Applied Optics*, vol. 22, no. 24, pp. 4105–4110, 1983.

[164] R. Swanepoel, "Transmission and reflection of an absorbing thin film on an absorbing substrate," *South African Journal of Physics*, vol. 12, no. 4, pp. 148–156, 1989.

[165] E. Kotlikov and G. Tereshchenko, "Study of optical constants of films used for the synthesis of broad-band antireflection coatings," *Optics and Spectroscopy*, vol. 82, no. 4, pp. 603–609, 1997.

[166] W. E. Vargas and D. Castro, "Closed equation for the normal incidence reflectance of thin films on absorbing substrates," *Applied Optics*, vol. 46, no. 4, pp. 502–505, 2007.

[167] E. Nichelatti, "Complex refractive index of a slab from reflectance and transmittance: Analytical solution," *Journal of Optics A: Pure and Applied Optics*, vol. 4, no. 4, pp. 400–403, 2002.

[168] A. Barybin and V. Shapovalov, "Substrate effect on the optical reflectance and transmittance of thin-film structures," *International Journal of Optics*, vol. 2010, no. 1, p. 137 572, 2010.

[169] E. Kotlikov, V. Prokashev, V. Ivanov, and A. Tropin, "Thickness uniformity of films deposited on rotating substrates," *Journal of Optical Technology*, vol. 76, no. 2, pp. 100–103, 2009.

[170] D. Nečas, I. Ohlídal, and D. Franta, "The reflectance of non-uniform thin films," *Journal of Optics A: Pure and Applied Optics*, vol. 11, no. 4, p. 045 202, 2009.

[171] D. Franta, D. Nečas, L. Zajíčková, I. Ohlídal, and J. Stuchlík, "Advanced modeling for optical characterization of amorphous hydrogenated silicon films," *Thin Solid Films*, vol. 541, pp. 12–16, 2013.

[172] J. Baek, D. Kovar, J. W. Keto, and M. F. Becker, "Determination of properties of wedged, nonuniformly thick, and absorbing thin films by using a new numerical method," *Applied Optics*, vol. 45, no. 7, pp. 1627–1639, 2006.

[173] J. Prentice, "Coherent, partially coherent and incoherent light absorption in thin-film multilayer structures," *Journal of Physics D: Applied Physics*, vol. 33, no. 24, p. 3139, 2000.

[174] C. C. Katsidis and D. I. Siapkas, "General transfer-matrix method for optical multilayer systems with coherent, partially coherent, and incoherent interference," *Applied Optics*, vol. 41, no. 19, pp. 3978–3987, 2002.

[175] H. Kwak, I. Jung, D. Kim, *et al.*, "Resonant-mode engineering for additive reflective structural colors with high brightness and high color purity," *Scientific Reports*, vol. 14, no. 1, p. 13 694, 2024.

[176] S. Ju, C. Kang, D. Kim, *et al.*, "Transmissive structural color filters employing multicavity for high efficiency and high color purity," in *Conference on Lasers and Electro-Optics/Pacific Rim*, Optica Publishing Group, 2024, P2_106.

[177] A. Lamminpää, S. Nevas, F. Manoocheri, and E. Ikonen, "Characterization of thin films based on reflectance and transmittance measurements at oblique angles of incidence," *Applied Optics*, vol. 45, no. 7, pp. 1392–1396, 2006.

[178] P. Bousquet, F. Flory, and P. Roche, "Scattering from multilayer thin films: Theory and experiment," *JOSA*, vol. 71, no. 9, pp. 1115–1123, 1981.

[179] S. Schröder, M. Trost, M. Garrick, *et al.*, "Origins of light scattering from thin film coatings," *Thin Solid Films*, vol. 592, pp. 248–255, 2015.

[180] A. Poruba, A. Fejfar, Z. Remeš, *et al.*, "Optical absorption and light scattering in microcrystalline silicon thin films and solar cells," *Journal of Applied Physics*, vol. 88, no. 1, pp. 148–160, 2000.

[181] J. J. Ruiz-Pérez and E. M. Navarro, "Optical transmittance for strongly-wedge-shaped semiconductor films: Appearance of envelope-crossover points in amorphous as-based chalcogenide materials," *Coatings*, vol. 10, no. 11, p. 1063, 2020.

[182] M. Ballester, A. Márquez, C. García-Vázquez, *et al.*, "Energy-band-structure calculation by below-band-gap spectrophotometry in thin layers of non-crystalline semiconductors: A case study of unhydrogenated a-si," *Journal of Non-Crystalline Solids*, vol. 594, p. 121 803, 2022.

[183] M. Ballester, M. García, A. P. Márquez, *et al.*, "Application of the holomorphic tauc-lorentz-urbach function to extract the optical constants of amorphous semiconductor thin films," *Coatings*, vol. 12, no. 10, p. 1549, 2022.

[184] P. Yeh and M. Hendry, "Optical waves in layered media," *Physics Today*, vol. 43, no. 1, p. 77, 1990.

[185] G. New, *Introduction to Nonlinear Optics*. Cambridge University Press, 2011.

[186] A. Stadler, "Analyzing uv/vis/nir spectra—correct and efficient parameter extraction," *IEEE Sensors Journal*, vol. 10, no. 12, pp. 1921–1931, 2010.

[187] L. Wahab, H. Zayed, and A. A. El-Galil, "Study of structural and optical properties of $cd_{1-x}zn_x$se thin films," *Thin Solid Films*, vol. 520, no. 16, pp. 5195–5199, 2012.

[188] E. Centurioni, "Generalized matrix method for calculation of internal light energy flux in mixed coherent and incoherent multilayers," *Applied Optics*, vol. 44, no. 35, pp. 7532–7539, 2005.

[189] C. Akcay, P. Parrein, and J. P. Rolland, "Estimation of longitudinal resolution in optical coherence imaging," *Applied Optics*, vol. 41, no. 25, pp. 5256–5262, 2002.

[190] M. J. Ablowitz and A. S. Fokas, *Complex Variables: Introduction and Applications*. Cambridge University Press, 2003.

[191] *Mathematica, version 14.0*, Wolfram Research, Inc., Champaign, IL, 2024.

[192] K. Atkinson, *An Introduction to Numerical Analysis*. John Wiley & Sons, 1991.

[193] I. Grebenstikov, A. Vlassov, B. Neporent, and N. Suykovskaya, *Decreasing of the Light Reflection on the Glass Surface*. Moscow: Gos. Izd. Teh-Teor. Lit., 1946.

[194] S. Humphrey, "Direct calculation of the optical constants for a thin film using a midpoint envelope," *Applied Optics*, vol. 46, no. 21, pp. 4660–4666, 2007.

[195] J. Luňáček, P. Hlubina, and M. Luňáčková, "Simple method for determination of the thickness of a nonabsorbing thin film using spectral reflectance measurement," *Applied Optics*, vol. 48, no. 5, pp. 985–989, 2009.

[196] E. Fortunato, P. Barquinha, and R. Martins, "Oxide semiconductor thin-film transistors: A review of recent advances," *Advanced materials*, vol. 24, no. 22, pp. 2945–2986, 2012.

[197] N. A. Al-Ahmadi, "Metal oxide semiconductor-based schottky diodes: A review of recent advances," *Materials Research Express*, vol. 7, no. 3, p. 032 001, 2020.

[198] L. Sang, M. Liao, and M. Sumiya, "A comprehensive review of semiconductor ultraviolet photodetectors: From thin film to one-dimensional nanostructures," *Sensors*, vol. 13, no. 8, pp. 10 482–10 518, 2013.

[199] V. H. López-Lugo, M. García-Hipólito, A. Rodríguez-Gómez, and J. C. Alonso-Huitrón, "Fabrication of li-doped nio thin films by ultrasonic spray pyrolysis and its application in light-emitting diodes," *Nanomaterials*, vol. 13, no. 1, p. 197, 2023.

[200] M.-J. Spijkman, K. Myny, E. C. Smits, P. Heremans, P. W. Blom, and D. M. De Leeuw, "Dual-gate thin-film transistors, integrated circuits and sensors," *Advanced Materials*, vol. 23, no. 29, pp. 3231–3242, 2011.

[201] J. Henry and J. Livingstone, "Thin-film amorphous silicon position-sensitive detectors," *Advanced materials*, vol. 13, no. 12-13, pp. 1022–1026, 2001.

[202] Y. S. Rim, "Review of metal oxide semiconductors-based thin-film transistors for point-of-care sensor applications," *Journal of Information Display*, vol. 21, no. 4, pp. 203–210, 2020.

[203] A. Shah, R. Platz, and H. Keppner, "Thin-film silicon solar cells: A review and selected trends," *Solar energy materials and solar cells*, vol. 38, no. 1-4, pp. 501–520, 1995.

[204] E. Maiorano, S. Gianvittorio, M. Lanzi, D. Tonelli, H. Pick, and A. Lesch, "Print-light-synthesis of gold thin film electrodes for electrochemical sensing," *Advanced Materials Technologies*, vol. 8, no. 12, p. 2 202 039, 2023.

[205] M. G. Toudeshkchoui, N. Rabiee, M. Rabiee, *et al.*, "Microfluidic devices with gold thin film channels for chemical and biomedical applications: A review," *Biomedical Microdevices*, vol. 21, pp. 1–17, 2019.

[206] M. M. Singh, G. Vijaya, M. Krupashankara, B. Sridhara, and T. Shridhar, "Deposition and characterization of aluminium thin film coatings using dc magnetron sputtering process," *Materials Today: Proceedings*, vol. 5, no. 1, pp. 2696–2704, 2018.

[207] W. He and C. Ye, "Flexible transparent conductive films on the basis of ag nanowires: Design and applications: A review," *Journal of Materials Science & Technology*, vol. 31, no. 6, pp. 581–588, 2015.

[208] J. P. Silva, K. C. Sekhar, H. Pan, J. L. MacManus-Driscoll, and M. Pereira, "Advances in dielectric thin films for energy storage applications, revealing the promise of group iv binary oxides," *ACS Energy Letters*, vol. 6, no. 6, pp. 2208–2217, 2021.

[209] R. Politanskyi, M. Vistak, G. Barylo, and A. Andrushchak, "Simulation of anti-reflecting dielectric films by the interference matrix method," *Optical Materials*, vol. 102, p. 109 782, 2020.

[210] M. Kumar, N. Kumari, P. K. Rao, V. Karar, S. Ramagopal, and A. L. Sharma, "Multilayer dielectric thin-film optical filters for beam folding applications," in *International Conference on Optics and Photonics 2015*, SPIE, vol. 9654, 2015, pp. 90–94.

[211] J. Benedikt, V. Raballand, A. Yanguas-Gil, K. Focke, and A. Von Keudell, "Thin film deposition by means of atmospheric pressure microplasma jet," *Plasma physics and controlled fusion*, vol. 49, no. 12B, B419, 2007.

[212] F. Massines, C. Sarra-Bournet, F. Fanelli, N. Naudé, and N. Gherardi, "Atmospheric pressure low temperature direct plasma technology: Status and challenges for thin film deposition," *Plasma Processes and Polymers*, vol. 9, no. 11-12, pp. 1041–1073, 2012.

[213] A. Tumuluri, K. L. Naidu, and K. J. Raju, "Band gap determination using tauc's plot for linbo3 thin films," *Int. J. ChemTech Res*, vol. 6, no. 6, pp. 3353–3356, 2014.

[214] J. Guerra, A. Tejada, J. Töfflinger, R. Grieseler, and L. Korte, "Band-fluctuations model for the fundamental absorption of crystalline and amorphous semiconductors: A dimensionless joint density of states analysis," *Journal of Physics D: Applied Physics*, vol. 52, no. 10, p. 105 303, 2019.

[215] M. Bhowmick, H. Xi, and B. Ullrich, "Optical bandgap definition via a modified form of urbach's rule," *Materials*, vol. 14, no. 7, p. 1639, 2021.

[216] C. Kaiser, O. J. Sandberg, N. Zarrabi, W. Li, P. Meredith, and A. Armin, "A universal urbach rule for disordered organic semiconductors," *Nature Communications*, vol. 12, no. 1, p. 3988, 2021.

[217] N. Ueno and S. Kera, "Electron spectroscopy of functional organic thin films: Deep insights into valence electronic structure in relation to charge transport property," *Progress in Surface Science*, vol. 83, no. 10-12, pp. 490–557, 2008.

[218] J. Park, Y. J. Cho, W. Chegal, J. Lee, Y.-S. Jang, and J. Jin, "A review of thin-film thickness measurements using optical methods," *International Journal of Precision Engineering and Manufacturing*, vol. 25, no. 8, pp. 1725–1737, 2024.

[219] H. Fujiwara, *Spectroscopic Ellipsometry: Principles and Applications*. John Wiley & Sons, 2007.

[220] R. Swanepoel, "Determining refractive index and thickness of thin films from wavelength measurements only," *JOSA A*, vol. 2, no. 8, pp. 1339–1343, 1985.

[221] M. Nowak, "Determination of optical constants and average thickness of inhomogeneous-rough thin films using spectral dependence of optical transmittance," *Thin Solid Films*, vol. 254, no. 1-2, pp. 200–210, 1995.

[222] X. Wu, F. Lai, L. Lin, *et al.*, "Optical inhomogeneity of zns films deposited by thermal evaporation," *Applied surface science*, vol. 254, no. 20, pp. 6455–6460, 2008.

[223] Y. Jin, B. Song, C. Lin, *et al.*, "Extension of the swanepoel method for obtaining the refractive index of chalcogenide thin films accurately at an arbitrary wavenumber," *Optics Express*, vol. 25, no. 25, pp. 31 273–31 280, 2017.

[224] Z. Hu, Z. Huang, Z. Lai, G. Wang, and J. Chu, "Optical properties of amorphous pbzr$_x$ti$_{1-x}$o$_3$ ($x = 0.52$) thin films prepared by rf magnetron sputtering," *Thin Solid Films*, vol. 437, no. 1-2, pp. 223–229, 2003.

[225] Z. E. Hebal, N. Sengouga, A. Meftah, and A. Djelloul, "Development of a software "hebaloptics" to extract optical properties and thickness of thin films from a single transmittance measurement," *Optical and Quantum Electronics*, vol. 56, no. 10, p. 1660, 2024.

[226] D. Poelman and P. Smet, "Methods for the determination of the optical constants of thin films from single transmission measurements: A critical review," *JOURNAL OF PHYSICS D-APPLIED PHYSICS*, vol. 36, pp. 1850–1857, Aug. 2003.

[227] A. Ferlauto, G. Ferreira, J. M. Pearce, *et al.*, "Analytical model for the optical functions of amorphous semiconductors from the near-infrared to ultraviolet: Applications in thin film photovoltaics," *Journal of Applied Physics*, vol. 92, no. 5, pp. 2424–2436, 2002.

[228] L. V. Rodríguez-de Marcos and J. I. Larruquert, "Analytic optical-constant model derived from tauc-lorentz and urbach tail," *Optics express*, vol. 24, no. 25, pp. 28 561–28 572, 2016.

[229] J. I. Larruquert and L. V. Rodríguez-de Marcos, "Procedure to convert optical-constant models into analytic," *Thin Solid Films*, vol. 664, pp. 52–59, 2018.

[230] J. I. Larruquert and L. V. Rodríguez-de Marcos, "Corrigendum to procedure to convert optical-constant models into analytic thin solid films (2018) 52-59," *Thin Solid Films*, vol. 733, p. 138 816, 2021.

[231] D. Franta, D. Nečas, and I. Ohlídal, "Universal dispersion model for characterization of optical thin films over a wide spectral range: Application to hafnia," *Applied Optics*, vol. 54, no. 31, pp. 9108–9119, 2015.

[232] D. Franta, D. Nečas, A. Giglia, P. Franta, and I. Ohlídal, "Universal dispersion model for characterization of optical thin films over wide spectral range: Application to magnesium fluoride," *Applied Surface Science*, vol. 421, pp. 424–429, 2017.

[233] D. Franta and M.-G. Mureșan, "Wide spectral range optical characterization of yttrium aluminum garnet (yag) single crystal by the universal dispersion model," *Optical Materials Express*, vol. 11, no. 12, pp. 3930–3945, 2021.

[234] D. Franta, J. Vohánka, J. Dvořák, *et al.*, "Optical characterization of gadolinium fluoride films using universal dispersion model," *Coatings*, vol. 13, no. 2, p. 218, 2023.

[235] M. McClain, A. Feldman, D. Kahaner, and X. Ying, "An algorithm and computer program for the calculation of envelope curves," *Computers in Physics*, vol. 5, no. 1, pp. 45–48, 1991.

[236] M. H. Saleh, N. M. Ershaidat, M. J. A. Ahmad, B. N. Bulos, and M. M. A.-G. Jafar, "Evaluation of spectral dispersion of optical constants of a-se films from their normal-incidence transmittance spectra using swanepoel algebraic envelope approach," *Optical Review*, vol. 24, pp. 260–277, 2017.

[237] B. S. Richards, "Optical characterization of sputtered silicon thin films for photovoltaic application," Ph.D. dissertation, MSc. Thesis, 1998.

[238] T. Güngör and B. Saka, "Calculation of the optical constants of a thin layer upon a transparent substrate from the reflection spectrum using a genetic algorithm," *Thin Solid Films*, vol. 467, no. 1-2, pp. 319–325, 2004.

[239] D. Minkov and R. Swanepoel, "Computerization of the optical characterization of a thin dielectric film," *Optical Engineering*, vol. 32, no. 12, pp. 3333–3337, 1993.

[240] D. Minkov and R. Swanepoel, "Computer drawing of the envelopes of spectra with interference," in *Thin Films for Optical Systems*, SPIE, vol. 1782, 1993, pp. 212–220.

[241] G. Gavrilov, D. Minkov, E. Marquez, and S. Ruano, "Advanced computer drawing envelopes of transmittance spectra of thin film specimens," *spectrum*, vol. 5, p. 6, 2016.

[242] D. A. Minkov, G. M. Gavrilov, E. Marquez, S. M. F. Ruano, and A. V. Stoynova, "Development of algorithm for computer drawing envelopes of interference reflectance spectra for thin film specimens," *Optik*, vol. 132, pp. 320–328, 2017.

[243] M. Salah, A. Abdelnaeim, S. Makhlolf, A. EL-Taher, and E. Shaaban, "Optical parameters of varies thickness of bismuth (iii) iodide thin films for photovoltaic and nonlinear applications," *Int. J. Thin. Film. Sci. Tec*, vol. 11, no. 3, pp. 275–282, 2022.

[244] V. Matějec, J. Pedliková, I. Barton, J. Zavadil, and P. Kostka, "Optical properties of as2s3 layers deposited from solutions," *Journal of Non-Crystalline Solids*, vol. 431, pp. 47–51, 2016.

[245] D. J. Borah and A. Mostako, "Investigation on dispersion parameters of molybdenum oxide thin films via wemple–didomenico (wdd) single oscillator model," *Applied Physics A*, vol. 126, no. 10, p. 818, 2020.

[246] B. Ullrich, A. K. Singh, M. Bhowmick, *et al.*, "Photoluminescence lineshape of zno," *AIP Advances*, vol. 4, no. 12, 2014.

[247] E. Marquez, J. Gonzalez-Leal, R. Prieto-Alcon, R. Jimenez-Garay, and M. Vlcek, "On the photo-and thermally-induced darkening phenomena in as40 s40 se20 amorphous chalcogenide thin films," *Journal of Physics D: Applied Physics*, vol. 32, no. 24, p. 3128, 1999.

[248] K. Tanaka, "Optical properties and photoinduced changes in amorphous ass films," *Thin solid films*, vol. 66, no. 3, pp. 271–279, 1980.

[249] M. Foldyna, K. Postava, J. Bouchala, J. Pistora, and T. Yamaguchi, "Model dielectric functional of amorphous materials including urbach tail," in *Microwave and Optical Technology 2003*, SPIE, vol. 5445, 2004, pp. 301–305.

[250] N. Malkova, L. Poslavsky, M. Di, D. Hu, and Q. Zhao, "New dispersion model for band gap tracking," *Thin Solid Films*, vol. 595, pp. 32–35, 2015.

[251] E. Garoudja, R. Amrani, W. Filali, *et al.*, "Artificial bee colony algorithm: A novel strategy for optical constants and thin film thickness extraction using only optical transmittance spectra for photovoltaic applications," *Optik*, vol. 241, p. 167 030, 2021.

[252] P. Virtanen, R. Gommers, T. E. Oliphant, *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[253] J. Gonzalez-Leal, M. Stuchlik, M. Vlcek, R. Jimenez-Garay, and E. Marquez, "Influence of the deposition technique on the structural and optical properties of amorphous ass films," *Applied surface science*, vol. 246, no. 4, pp. 348–355, 2005.

[254] R. P. Canale and S. C. Chapra, *Numerical methods for engineers*. Mcgraw-hill Education-Europe, 2014.

[255] Y. Yuan, S. Liu, J. Zhang, Y. Zhang, C. Dong, and L. Lin, "Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 814–81 409, 2018.

[256] K. Zhang, W. Zuo, and L. Zhang, "Deep plug-and-play super-resolution for arbitrary blur kernels," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2019.

[257] J. Cai, H. Zeng, H. Yong, Z. Cao, and L. Zhang, "Toward real-world single image super-resolution: A new benchmark and a new model," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2019.

[258] S. Lopez-Tapia, J. Mateos, R. Molina, and A. K. Katsaggelos, "Learning moore-penrose based residuals for robust non-blind image deconvolution," *Digital Signal Processing*, vol. 142, p. 104 193, 2023, ISSN: 1051-2004. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1051200423002889`.

[259] A. Ganjoo and R. Golovchak, "Computer program parav for calculating optical constants of thin films and bulk materials: Case study of amorphous semiconductors," *Journal of Optoelectronics and Advanced Materials*, vol. 10, no. 6, p. 1328, 2008.

[260] J. Watt, R. Borhani, and A. K. Katsaggelos, *Machine learning refined: Foundations, algorithms, and applications*. Cambridge University Press, 2020.

[261] D. Minkov, G. Gavrilov, G. Angelov, *et al.*, "Optimisation of the envelope method for characterisation of optical thin film on substrate specimens from their normal incidence transmittance spectrum," *Thin Solid Films*, vol. 645, pp. 370–378, 2018.

[262] G. Cody, T. Tiedje, B. Abeles, B. Brooks, and Y. Goldstein, "Disorder and the optical-absorption edge of hydrogenated amorphous silicon," *Physical Review Letters*, vol. 47, no. 20, p. 1480, 1981.

[263] S. Wemple, "Optical oscillator strengths and excitation energies in solids, liquids, and molecules," *The Journal of Chemical Physics*, vol. 67, no. 5, pp. 2151–2168, 1977.

[264] Z. Yue, H. Yong, Q. Zhao, L. Zhang, D. Meng, and K.-Y. K. Wong, "Deep variational network toward blind image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 11, pp. 7011–7026, 2024.

[265] E. Haller, "Germanium: From its discovery to sige devices," *Materials science in semiconductor processing*, vol. 9, no. 4-5, pp. 408–422, 2006.

[266] K. Hayashi, I. Park, K. Dotsu, *et al.*, "Radiation effects on the silicon semiconductor detectors for the astro–h mission," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 699, pp. 225–229, 2013.

[267] P. N. Luke, M. Amman, C. Tindall, and J. S. Lee, "Recent developments in semiconductor gamma-ray detectors," *Journal of radioanalytical and nuclear chemistry*, vol. 264, no. 1, pp. 145–153, 2005.

[268] J. Ponpon and M. Sieskind, "Recent advances in γ-and x-ray spectrometry by means of mercuric iodide detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 380, no. 1-2, pp. 173–178, 1996.

[269] J. Liu and Y. Zhang, "Growth of lead iodide single crystals used for nuclear radiation detection of gamma-rays," *Crystal Research and Technology*, vol. 52, no. 3, p. 1 600 370, 2017.

[270] H. Kim, L. Cirignano, A. Churilov, *et al.*, "Developing larger tlbr detectors—detector performance," *IEEE Transactions on Nuclear Science*, vol. 56, no. 3, pp. 819–823, 2009.

[271] K. Ikegami, H. Nishizawa, K. Takashima, *et al.*, "Czt semiconductor radiation sensor for high energy gamma rays," *Radiation effects and defects in solids*, vol. 146, no. 1-4, pp. 161–173, 1998.

[272] E. Kalemci, J. L. Matteson, R. T. Skelton, P. L. Hink, and K. R. Slavis, "Model calculations of the response of czt strip detectors," in *Hard X-Ray, Gamma-Ray, and Neutron Detector Physics*, SPIE, vol. 3768, 1999, pp. 360–373.

[273] Y. Du, J. LeBlanc, G. E. Possin, B. D. Yanoff, and S. Bogdanovich, "Temporal response of czt detectors under intense irradiation," in *2002 IEEE Nuclear Science Symposium Conference Record*, IEEE, vol. 1, 2002, pp. 480–484.

[274] M. Picone, A. Glière, and P. Massé, "A three-dimensional model of cdznte gamma-ray spectrometer," *Nuclear Instruments and Methods in Physics Research Section A: Ac-*

celerators, Spectrometers, Detectors and Associated Equipment, vol. 504, no. 1-3, pp. 313–316, 2003.

[275] F. Mathy, A. Glière, E. G. d'Aillon, *et al.*, "A three-dimensional model of cdznte gamma-ray detector and its experimental validation," *IEEE Transactions on Nuclear Science*, vol. 51, no. 5, pp. 2419–2426, 2004.

[276] J. Fink, H. Krueger, P. Lodomez, and N. Wermes, "Characterization of charge collection in cdte and czt using the transient current technique," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 560, no. 2, pp. 435–443, 2006.

[277] P. Guerra, A. Santos, and D. Darambara, "Development of a simplified simulation model for performance characterization of a pixellated cdznte multimodality imaging system," *Physics in Medicine & Biology*, vol. 53, no. 4, p. 1099, 2008.

[278] L. Chen and Y.-X. Wei, "Monte carlo simulations of the snm spectra for czt and nai spectrometers," *Applied Radiation and Isotopes*, vol. 66, no. 8, pp. 1146–1150, 2008.

[279] Y. Zhu, S. E. Anderson, and Z. He, "Sub-pixel position sensing for pixelated, 3-d position sensitive, wide band-gap, semiconductor, gamma-ray detectors," *IEEE Transactions on Nuclear Science*, vol. 58, no. 3, pp. 1400–1409, 2011.

[280] M. E. Myronakis and D. G. Darambara, "Monte carlo investigation of charge-transport effects on energy resolution and detection efficiency of pixelated czt detectors for spect/pet applications," *Medical physics*, vol. 38, no. 1, pp. 455–467, 2011.

[281] G. Montémont, S. Lux, O. Monnet, S. Stanchina, and L. Verger, "Studying spatial resolution of czt detectors using sub-pixel positioning for spect," *IEEE Transactions on Nuclear Science*, vol. 61, no. 5, pp. 2559–2566, 2014.

[282] M. Bettelli and N. S. Amadè, "A first principles method to simulate the spectral response of cdznte-based x-and gamma-ray detectors," in *Radiation Detection Systems*, CRC Press, 2021, pp. 33–60.

[283] I. Abt, F. Fischer, F. Hagemann, *et al.*, "Simulation of semiconductor detectors in 3d with solidstatedetectors. jl," *Journal of Instrumentation*, vol. 16, no. 08, P08007, 2021.

[284] M. Ballester, J. Kaspar, F. Massanes, S. Banerjee, A. H. Vija, and A. K. Katsaggelos, "Modeling and simulation of charge-induced signals in photon-counting czt detectors for medical imaging applications," *arXiv preprint arXiv:2405.13168*, 2024.

[285] M. Ballester, S. Banerjee, M. Rodrigues, J. Kaspar, A. H. Vija, and A. K. Katsaggelos, "Materials and defects characterization of cdznte sensors using the inverse synthesis method," in *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, IEEE, 2022, pp. 1–2.

[286] S. Banerjee, M. Rodrigues, M. Ballester, A. H. Vija, and A. K. Katsaggelos, "Machine learning approaches in room temperature semiconductor detectors," in *X-ray Photon Processing Detectors: Space, Industrial, and Medical applications*, Springer, 2023, pp. 67–94.

[287] S. Banerjee, M. Rodrigues, M. Ballester, A. H. Vija, and A. Katsaggelos, "Identifying defects without a priori knowledge in a room-temperature semiconductor detector using physics inspired machine learning model," *Sensors*, vol. 24, no. 1, p. 92, 2023.

[288] S. Banerjee, M. Rodrigues, M. Ballester, A. H. Vija, and A. K. Katsaggelos, "Learning-based physical models of room-temperature semiconductor detectors with reduced data," *Scientific reports*, vol. 13, no. 1, p. 168, 2023.

[289] S. Banerjee, M. Rodrigues, M. Ballester, A. H. Vija, and A. K. Katsaggelos, "A physics based machine learning model to characterize room temperature semiconductor detectors in 3d," *Scientific reports*, vol. 14, no. 1, p. 7803, 2024.

[290] H. Barrett, J. Eskin, and H. Barber, "Charge transport in arrays of semiconductor gamma-ray detectors," *Physical Review Letters*, vol. 75, no. 1, p. 156, 1995.

[291] Y. Cui, A. Bolotnikov, G. Camarda, A. Hossain, G. Yang, and R. James, "Czt virtual frisch-grid detector: Principles and applications," in *2009 IEEE Long Island Systems, Applications and Technology Conference*, IEEE, 2009, pp. 1–5.

[292] W. Shockley, "Currents to conductors induced by a moving point charge," *Journal of applied physics*, vol. 9, no. 10, pp. 635–636, 1938.

[293] S. Ramo, "Currents induced by electron motion," *Proceedings of the IRE*, vol. 27, no. 9, pp. 584–585, 1939.

[294] Z. He, "Review of the shockley–ramo theorem and its application in semiconductor gamma-ray detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 463, no. 1-2, pp. 250–267, 2001.

[295] M. Brigida, C. Favuzzi, P. Fusco, *et al.*, "A new monte carlo code for full simulation of silicon strip detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 533, no. 3, pp. 322–343, 2004.

[296] E. Gaubas, T. Ceponis, V. Kalesinskas, J. Pavlov, and J. Vysniauskas, "Simulations of operation dynamics of different type gan particle sensors," *Sensors*, vol. 15, no. 3, pp. 5429–5473, 2015.

[297] Y. Eisen and A. Shor, "Cdte and cdznte materials for room-temperature x-ray and gamma ray detectors," *Journal of crystal growth*, vol. 184, pp. 1302–1312, 1998.

[298] M. Zanichelli, A. Santi, M. Pavesi, and A. Zappettini, "Charge collection in semi-insulator radiation detectors in the presence of a linear decreasing electric field," *Journal of Physics D: Applied Physics*, vol. 46, no. 36, p. 365 103, 2013.

[299] S. G. Fritz and P. M. Shikhaliev, "Czt detectors used in different irradiation geometries: Simulations and experimental results," *Medical physics*, vol. 36, no. 4, pp. 1098–1108, 2009.

[300] A. L. Washington, L. C. Teague, M. C. Duff, A. Burger, M. Groza, and V. Buliga, "The effect of various detector geometries on the performance of czt using one crystal," *Journal of electronic materials*, vol. 40, pp. 1744–1748, 2011.

[301] H. W. Yao, R. J. Anderson, and R. B. James, "Optical characterization of the internal electric field distribution under bias of cdznte radiation detectors," in *Hard X-Ray and Gamma-Ray Detector Physics, Optics, and Applications*, SPIE, vol. 3115, 1997, pp. 62–68.

[302] M. Pavesi, A. Santi, M. Bettelli, A. Zappettini, and M. Zanichelli, "Electric field reconstruction and transport parameter evaluation in czt x-ray detectors," *IEEE Transactions on Nuclear Science*, vol. 64, no. 10, pp. 2706–2712, 2017.

[303] Š. Uxa, E. Belas, R. Grill, P. Praus, and R. B. James, "Determination of electric-field profile in cdte and cdznte detectors using transient-current technique," *IEEE Transactions on Nuclear Science*, vol. 59, no. 5, pp. 2402–2408, 2012.

[304] M. Jung, J. Morel, P. Fougeres, M. Hage-Ali, and P. Siffert, "A new method for evaluation of transport properties in cdte and czt detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 428, no. 1, pp. 45–57, 1999.

[305] C. Buurma, S. Krishnamurthy, and S. Sivananthan, "Shockley-read-hall lifetimes in cdte," *Journal of Applied Physics*, vol. 116, no. 1, 2014.

[306] T. Höhr, A. Schenk, and W. Fichtner, "Revised shockley–read–hall lifetimes for quantum transport modeling," *Journal of applied physics*, vol. 95, no. 9, pp. 4875–4882, 2004.

[307] S. Kasap, K. O. Ramaswami, M. Kabir, and R. Johanson, "Corrections to the hecht collection efficiency in photoconductive detectors under large signals: Non-uniform electric field due to drifting and trapped unipolar carriers," *Journal of Physics D: Applied Physics*, vol. 52, no. 13, p. 135 104, 2019.

[308] K. O. Ramaswami, "Monte carlo simulation of drifting charge carriers in photoconductive integrating detectors," Ph.D. dissertation, University of Saskatchewan Saskatoon, 2019.

[309] E. Gatti, A. Longoni, P. Rehak, and M. Sampietro, "Dynamics of electrons in drift detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 253, no. 3, pp. 393–399, 1987.

[310] M. Benoit and L. Hamel, "Simulation of charge collection processes in semiconductor cdznte γ-ray detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 606, no. 3, pp. 508–516, 2009.

[311] A. Bolotnikov, G. Camarda, G. Carini, Y. Cui, L. Li, and R. James, "Cumulative effects of te precipitates in cdznte radiation detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 571, no. 3, pp. 687–698, 2007.

[312] A. Bolotnikov, G. Camarda, Y. Cui, *et al.*, "Characterization and evaluation of extended defects in czt crystals for gamma-ray detectors," *Journal of crystal growth*, vol. 379, pp. 46–56, 2013.

[313] K. Spartiotis, A. Leppänen, T. Pantsar, *et al.*, "A photon counting cdte gamma-and x-ray camera," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 550, no. 1-2, pp. 267–277, 2005.

[314] E. Gatti and P. Rehak, "Review of semiconductor drift detectors," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 541, no. 1-2, pp. 47–60, 2005.

[315] C. Sundberg, M. Persson, J. J. Wikner, and M. Danielsson, "1-μm spatial resolution in silicon photon-counting ct detectors," *Journal of Medical Imaging*, vol. 8, no. 6, pp. 063 501–063 501, 2021.

[316] B. J. Van Zeghbroeck, *Principles of semiconductor devices*. Bart Van Zeghbroeck University of Colorado, CO, 2011.

[317] J. E. Marsden and A. Tromba, *Vector calculus*. Macmillan, 2003.

[318] D. J. Griffiths, *Introduction to electrodynamics*. Cambridge University Press, 2023.

[319] J. C. Strikwerda, *Finite difference schemes and partial differential equations*. SIAM, 2004.

[320] J. K. Blitzstein and J. Hwang, *Introduction to probability*. Chapman and Hall/CRC, 2019.

[321] S. M. Ross, S. M. Ross, S. M. Ross, and S. M. Ross, *A first course in probability*. Macmillan New York, 1976, vol. 2.

[322] P. Virtanen, R. Gommers, T. E. Oliphant, *et al.*, "Scipy 1.0: Fundamental algorithms for scientific computing in python," *Nature methods*, vol. 17, no. 3, pp. 261–272, 2020.

[323] K. Iniewski, "Czt detector technology for medical imaging," *Journal of Instrumentation*, vol. 9, no. 11, p. C11001, 2014.

[324] S.-D. Chun, S.-H. Park, D. H. Lee, *et al.*, "Property of a czt semiconductor detector for radionuclide identification," *Journal of Nuclear Science and Technology*, vol. 45, no. sup5, pp. 421–424, 2008.

[325] A. E. Bolotnikov, S. O. Babalola, G. S. Camarda, *et al.*, "Extended defects in cdznte radiation detectors," *IEEE Transactions on Nuclear Science*, vol. 56, no. 4, pp. 1775–1783, 2009.

[326] U. N. Roy, G. S. Camarda, Y. Cui, *et al.*, "Evaluation of cdzntese as a high-quality gamma-ray spectroscopic material with better compositional homogeneity and reduced defects," *Scientific reports*, vol. 9, no. 1, p. 7303, 2019.

[327] A. Bolotnikov, G. Camarda, G. Wright, and R. James, "Factors limiting the performance of cdznte detectors," *IEEE transactions on nuclear science*, vol. 52, no. 3, pp. 589–598, 2005.

[328] A. Makeev, M. Rodrigues, G.-C. Wang, and S. J. Glick, "Modeling czt/cdte x-ray photon-counting detectors," in *Medical Imaging 2015: Physics of Medical Imaging*, SPIE, vol. 9412, 2015, pp. 1194–1202.

[329] E. Kamieniecki, "Effect of charge trapping on effective carrier lifetime in compound semiconductors: High resistivity cdznte," *Journal of Applied Physics*, vol. 116, no. 19, p. 193 702, 2014.

[330] W. Zimmerman, "Experimental verification of the shockley–read–hall recombination theory in silicon," *Electronics Letters*, vol. 9, no. 16, pp. 378–379, 1973.

[331] A. Buttacavoli, F. Principato, G. Gerardi, *et al.*, "Incomplete charge collection at inter-pixel gap in low-and high-flux cadmium zinc telluride pixel detectors," *Sensors*, vol. 22, no. 4, p. 1441, 2022.

[332] D. S. Bale and C. Szeles, "Nature of polarization in wide-bandgap semiconductor detectors under high-flux irradiation: Application to semi-insulating cd 1- x zn x te," *Physical Review B*, vol. 77, no. 3, p. 035 205, 2008.

[333] S. Banerjee, M. Rodrigues, A. H. Vija, and A. K. Katsaggelos, "A learning-based physical model of charge transport in room-temperature semiconductor detectors," *IEEE Transactions on Nuclear Science*, vol. 69, no. 1, pp. 2–16, 2021.

[334] S. Sun, Z. Cao, H. Zhu, and J. Zhao, "A survey of optimization methods from a machine learning perspective," *IEEE transactions on cybernetics*, vol. 50, no. 8, pp. 3668–3681, 2019.

[335] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[336] S. Leng, M. Bruesewitz, S. Tao, *et al.*, "Photon-counting detector ct: System design and clinical applications of an emerging technology," *Radiographics*, vol. 39, no. 3, pp. 729–743, 2019.

[337] T. Takahashi and S. Watanabe, "Recent progress in cdte and cdznte detectors," *IEEE Transactions on nuclear science*, vol. 48, no. 4, pp. 950–959, 2001.

[338] D. R. Monismith and B. E. Mayfield, "Slime mold as a model for numerical optimization," in *2008 IEEE swarm intelligence symposium*, IEEE, 2008, pp. 1–8.

[339] N. Van Thieu and S. Mirjalili, "Mealpy: An open-source library for latest meta-heuristic algorithms in python," *Journal of Systems Architecture*, vol. 139, p. 102 871, 2023.

[340] S. Banerjee, M. Rodrigues, M. Ballester, A. H. Vija, and A. K. Katsaggelos, "Machine learning approaches for cdznte/cdte radiation detectors," in *Deep Learning for Advanced X-ray Detection and Imaging Applications*, Springer, 2025, pp. 143–176.

[341] S. Banerjee, M. Rodrigues, A. H. Vija, and A. K. Katsaggelos, "Event reconstruction in radiation detectors using convolutional neural networks," in *2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, IEEE, 2021, pp. 1–3.

[342] J. H. Bappy, C. Simons, L. Nataraj, B. Manjunath, and A. K. Roy-Chowdhury, "Hybrid lstm and encoder–decoder architecture for detection of image forgeries," *IEEE Transactions on Image Processing*, vol. 28, no. 7, pp. 3286–3300, 2019.

[343] M. Takeda and S. Kobayashi, "Lateral aberration measurements with a digital talbot interferometer," *Applied Optics*, vol. 23, no. 11, pp. 1760–1764, 1984.

[344] N. Verrier and M. Atlan, "Off-axis digital hologram reconstruction: Some practical considerations," *Applied optics*, vol. 50, no. 34, H136–H146, 2011.

[345] J. W. Goodman, "Introduction to fourier optics. 3rd," *Roberts and Company Publishers*, vol. 3, 2005.

[346] J. W. Goodman, *Speckle phenomena in optics: theory and applications*. Roberts and Company Publishers, 2007.

[347] R. Shanmugamani, *Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras*. Packt Publishing Ltd, 2018, pp. 129–156.

[348] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, no. 02, pp. 107–116, 1998.

[349] F. Xu, H. Ma, J. Sun, R. Wu, X. Liu, and Y. Kong, "Lstm multi-modal unet for brain tumor segmentation," in *2019 IEEE 4th international conference on image, vision and computing (ICIVC)*, IEEE, 2019, pp. 236–240.

[350] R. Shanmugamani, *Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras*. Packt Publishing Ltd, 2018, pp. 139–140.

[351] W. Jakob, *Mitsuba renderer*, http://www.mitsuba-renderer.org, 2010.

Appendix A

Supplementary Material for Single-shot Interferometric Camera

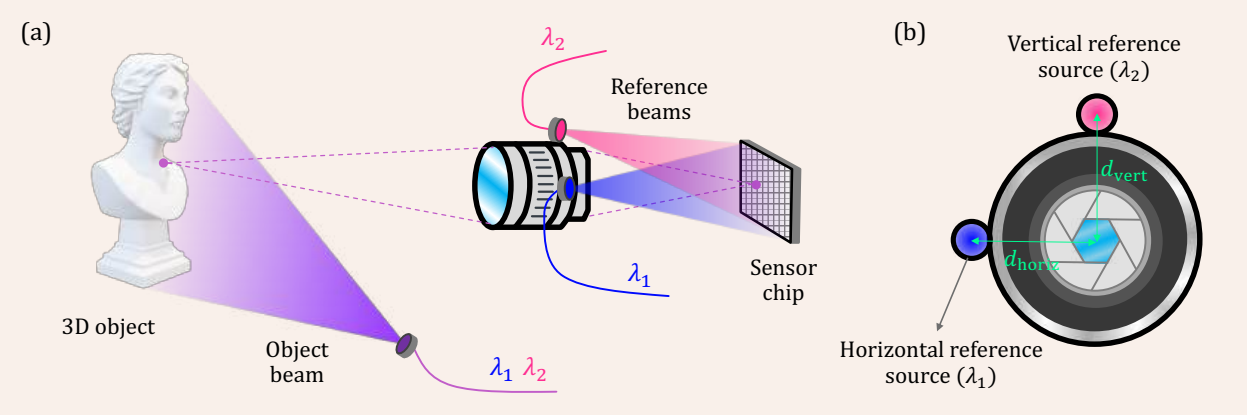


Figure A.1: Setup diagram for the novel camera based on the synthetic wavelength concept. (a) The 3D object is illuminated with an object beam that contains light waves at wavelengths λ_1 and λ_2. The field reflected in the object will pass through an imaging system and reach the camera sensor chip. In addition, two reference beams directly illuminate the sensor. (b) The frontal view of the optical system: Looking at the objective from the sensor location. The horizontal reference uses wavelength λ_1 and vertical reference λ_2. We control the angle at which the reference beams reach the sensor by modifying the distances d_{hor} and d_{vert}.

A.1 Retrieving the optical fields

In this supplementary section, we will describe in detail how to retrieve the optical fields $E(\lambda_1)$ and $E(\lambda_2)$ from a single captured image $I(x, y)$. In the Supp. Sec. A.1.1, we will

provide some mathematical background, defining the object and reference beams that arrive at the sensor. Then, Supp. Sec. A.1.2 will explain the Fourier-based demodulation algorithm that recovers the optical fields from the single-shot image.

A.1.1 Mathematical background

As mentioned in section 3.1 of the main manuscript, our setup has two reference sources that point directly towards the sensor array. One reference beam at λ_1 encloses an angle θ_1 with the horizontal x-axis, while the other reference beam at λ_2 encloses an angle θ_2 with the vertical y-axis of the detector. In addition, there is an object source that emits light beams at λ_1 and λ_2 to illuminate the object uniformly. The schematic of our setup is shown in Fig. A.1, and our full setup including diagram of the illumination engine is shown in Fig. A.2.

The light fields that scatter off the object and arrive at the sensor plane (at $z = z_0$) are represented as $E(\lambda_1) \equiv E_1(x, y, t)$ and $E(\lambda_2) \equiv E_2(x, y, t)$. We will employ the notation on the right-hand side of these equations to explicitly show that the optical fields are time and space dependent.

Object fields and reference fields:

The reference beams propagate as spherical waves from their point sources (the end of the optical fibers) toward the sensor chip. The distance from the reference sources to the sensor location is relatively long (about 38 cm) compared to the sensor size (less than 1 cm). Therefore, one can reasonably assume that the reference beam wavefront is almost planar when arrives at the sensor location. We have modeled these beams as planar waves hitting the sensor (at $z = z_0$) with an offset angle, as shown in the Fig. A.1. The reference

electric fields are then described as

$$\begin{aligned} E_1^{\text{ref}}(x,y,t) &= \frac{1}{2}|E_1^{\text{ref}}|e^{i(\omega_1 t - \boldsymbol{k_1}\cdot\boldsymbol{r}+\sigma_1)} + \text{cc}, \\ E_2^{\text{ref}}(x,y,t) &= \frac{1}{2}|E_2^{\text{ref}}|e^{i(\omega_2 t - \boldsymbol{k_2}\cdot\boldsymbol{r}+\sigma_2)} + \text{cc}. \end{aligned} \tag{A.1}$$

Here cc denotes the complex conjugate, which ensures that the reference fields are "real" in nature. This common complex notation [185] is employed to ease future calculations. The field $E_1^{\text{ref}}(x,y,t)$ denotes the horizontal reference field at λ_1 and has the approximately constant amplitude $|E_1^{\text{ref}}|$ at the sensor location. Moreover, $\omega_1 = 2\pi\, c/\lambda_1$ is the angular frequency (c being the speed of light), and $\mathbf{k_1} = (k_1^x, k_1^y, k_1^z)$ is the so-called "propagation vector" that indicates the three-dimensional propagation direction of the planar wave [69]. The constant shift σ_1 denotes the initial phase. It is known that $|\mathbf{k_1}| = k_1 = 2\pi/\lambda_1$ and $\mathbf{k_1}\cdot\mathbf{r} = (k_1^x, k_1^y, k_1^z)\cdot(x, y, z = z_0)$ is the usual dot product. To simplify the formulas in this study, we assume that both the object and reference beams arriving at the sensor have nearly aligned linear polarizations. This assumption enables us to adopt a simplified scalar wave approach.

The angle between the reference beams and the optical axis (as shown in Fig. A.1) can be represented as θ_1 and θ_2. Using these angles, the propagation vectors can be expressed as $\mathbf{k_1} = (k_1\sin(\theta_1), 0, k_1\cos(\theta_1))$ and $\mathbf{k_2} = (0, k_2\sin(\theta_2), k_2\cos(\theta_2))$. The reference beams from Eq. A.1 can then be rewritten as:

$$\begin{aligned} E_1^{\text{ref}}(x,y,t) &= \frac{1}{2}|E_1^{\text{ref}}|e^{i(\omega_1 t - k_1 x\sin\theta_1 - k_1 z_0\cos\theta_1+\sigma_1)} + cc \\ E_2^{\text{ref}}(x,y,t) &= \frac{1}{2}|E_2^{\text{ref}}|e^{i(\omega_2 t - k_2 y\sin\theta_2 - k_2 z_0\sin\theta_2+\sigma_2)} + cc \end{aligned} \tag{A.2}$$

It should be noted that $(k_1\cos(\theta_1)z_0 + \sigma_1)$ and

$(k_2 \cos(\theta_2) z_0 + \sigma_2)$ are constant shift phases that do not affect the *phase difference* at different positions (x, y). Therefore, these terms can be included within their random initial phases σ_1 and σ_2. Following a similar approach, we can express the two object beams at λ_1 and λ_2 with the following general beam expressions:

$$\begin{aligned} E_1(x, y, t) &= \frac{1}{2}|E_1(x, y)|e^{i(\omega_1 t - \phi_1(x,y))} + cc \\ E_2(x, y, t) &= \frac{1}{2}|E_2(x, y)|e^{i(\omega_2 t - \phi_2(x,y))} + cc \end{aligned} \tag{A.3}$$

The unknown phasemaps we aim to obtain are represented by the functions $\phi_1(x, y)$ and $\phi_2(x, y)$. For optically rough surfaces, these phasemaps exhibit a fully developed speckle pattern. $\phi_1(x, y)$ and $\phi_2(x, y)$ are rapidly-changing functions influenced by the random microscopic depth variations in the object's surface. The also unknown amplitudes of these phasemaps are denoted by $|E_1(x, y)|$ and $|E_2(x, y)|$.

Intensity at the sensor location:

The sensor encounters four incoming electric fields: two associated with the object beam (E_1 and E_2), and two with the reference beams (E_1^{ref} and E_2^{ref}). The intensity field at the sensor's position can be represented as the time-averaged magnitude of the Poynting vector resulting from the superposition of these four beams,

$$I(x, y) = \langle |\mathbf{S}| \rangle_T \propto \langle |E_1^{\text{ref}} + E_1^{\text{ref}} + E_1 + E_2|^2 \rangle_T \tag{A.4}$$

Here, T refers to the sensor's exposure time, $\langle \cdot \rangle_T$ signifies time-averaging, and $\mathbf{S}$ is the Poynting vector. During the free-space propagation, the norm of the Poynting vector

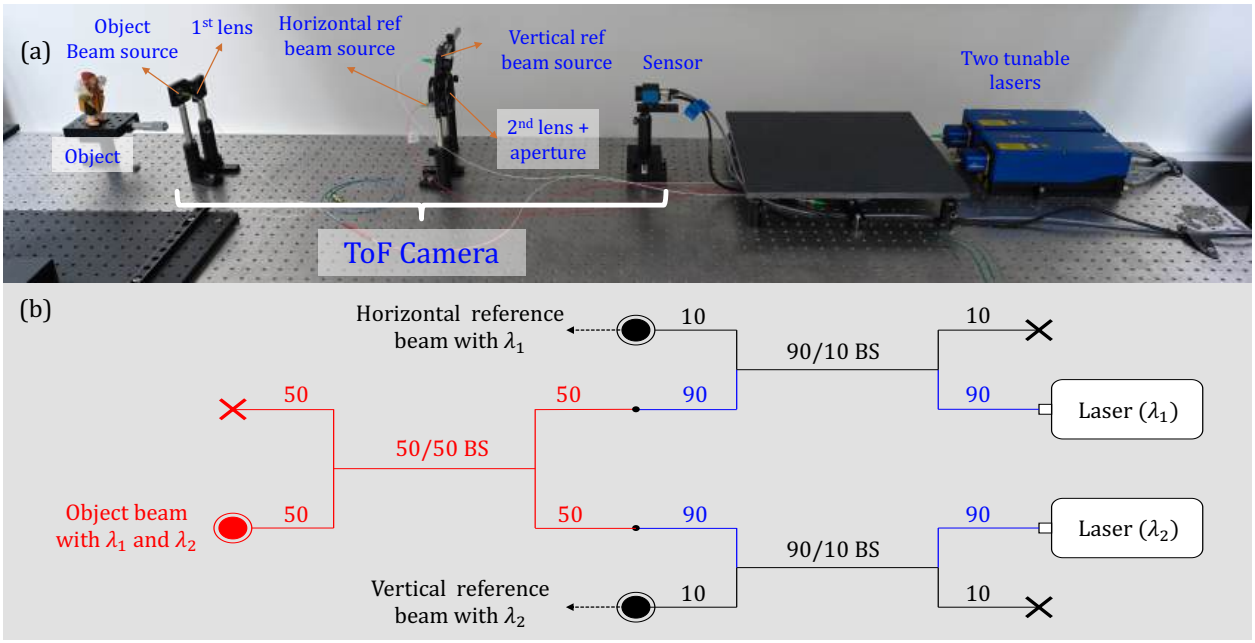


Figure A.2: (a) Photograph of our particular setup. Two identical tunable lasers will emit light in the NIR range at around 850 nm. We use optical fibers and beam splitters (BS) to direct the beams flexibly. We employ a 4f-system to image the object into the sensor array. (b) Diagram of the illumination engine.

becomes $|\mathbf{S}| = |\mathbf{E}_{\text{all}} \times \mathbf{B}_{\text{all}}| = |\mathbf{E}_{\text{all}}|^2/c$, where $\mathbf{E}_{\text{all}}$ denotes the superposition of the fields. As mentioned, if we assume assume that all beams reaching the sensor has roughly the same polarization direction (for instance, when a polarizer is placed in front of the sensor), then the wave superposition reduces to simple addition of the scalar waves, as shown in Eq. A.4. Absent the polarizer, we must consider the object beams as unpolarized, given their reflection from the object's rough surface. At any given time, each electric field then points in a random direction within the plane perpendicular to their propagation vector. These vectors can be decomposed into their component axes for superposition, and only the components that align will superpose. Even though our analysis simplifies the situation by considering aligned beams with linear polarization, a similar mathematical treatment applies to this more complex scenario.

When we calculate the squared norm from Eq. A.4, we find several interferences terms coming from the superposition of the different waves at wavelengths λ_1 and λ_2. However, some of these terms corresponds to time-dependent interference fringes that oscillate too fast to be temporally resolved with a conventional sensor. During the integration over the sensor's exposure time, rapidly oscillating fringes are averaged out. These oscillating terms are encountered when superposing waves with slightly different wavelengths, such as in the pairings of (E_1,E_2), (E_1^{ref},E_2^{ref}), (E_1^{ref},E_2), and (E_1,E_2^{ref}). Further details about these time-varying fringes and an exploration of the interferences between the object beams, specifically the (E_1,E_2) case, are provided in Supp. section A.2.

On the other hand, the interference between waves of identical wavelengths, such as the pairs (E_1^{ref},E_1) and (E_2^{ref},E_2), produces time-invariant fringes that the detector can discern. Taking these factors into account, we can rewrite the intensity field from Eq. A.4 using Eqs. A.2-A.3 as follows:

$$\begin{aligned} I(x,y) = &|E_1^{\text{ref}}|^2 + |E_2^{\text{ref}}|^2 + |E_1(x,y)|^2 + |E_2(x,y)|^2 \\ &+|E_1^{\text{ref}}||E_1(x,y)|\cos(\phi_1(x,y) - k_1 x\sin(\theta_1)) \\ &+|E_2^{\text{ref}}||E_2(x,y)|\cos(\phi_2(x,y) - k_2 y\sin(\theta_2)) \end{aligned} \tag{A.5}$$

To simplify Eq. A.5, we first establish the DC term as

$$a(x,y) = |E_1^{\text{ref}}|^2 + |E_2^{\text{ref}}|^2 + |E_1(x,y)|^2 + |E_2(x,y)|^2 \tag{A.6}$$

Then, the amplitudes of the static interference fringes will be denoted as follows:

$$\begin{aligned} b_1(x,y) &= |E_1^{\text{ref}}||E_1(x,y)| \\ b_2(x,y) &= |E_2^{\text{ref}}||E_2(x,y)| \end{aligned} \tag{A.7}$$

Finally, we establish $k_i \sin(\theta_i) = 2\pi f_i$, where f_i are the so-called spatial carrier frequencies, for $i \in \{1, 2\}$. Subsequently, we reformulate Eq. A.5 using Eqs. A.6-A.7 and finally obtain

$$\begin{aligned} I(x,y) = a(x,y) + b_1(x,y)\cos(\phi_1(x,y) + 2\pi f_1 x) \\ + b_2(x,y)\cos(\phi_2(x,y) + 2\pi f_2 y) \end{aligned} \tag{A.8}$$

Eq. A.8 encompasses all information about the experimental image $I(x, y)$ obtained with our single-shot camera. The image $I(x, y)$ displays a speckle pattern superimposed with crossed interference fringes. The vertical fringes, with a spatial carrier frequency f_1 along the x-axis, result from the interference between the object beam $E_1(x, y)$ and the horizontal reference beam (at λ_1). Similarly, the horizontal fringes, with a carrier frequency f_2 along the y-axis, originate from the interference of the object beam $E_2(x, y)$ and the vertical reference beam (at λ_2). Fig. A.3 shows a real captured image with two zoom-in sections. The first zoom-in window (Fig. A.3 b) showcases the speckle pattern generated by the object beams, while the second zoom-in window (Fig. A.3 c) clearly exhibits the cross-like static fringes due to the interference between the object and reference beams.

Our subsequent objective is to extract the phase values $\phi_1(x, y)$ and $\phi_2(x, y)$ that appear in Eq. A.8 using only the known real 2D numerical array $I(x, y)$. The demodulation algorithm for this purpose is detailed below, in Supp. section A.1.2.

A.1.2 Demodulation algorithm

Takeda *et al.* [57] pioneered a method for retrieving a single phasemap using a Fourier-based demodulation algorithm. This technique has since been applied in off-axis holography, interferometry, and fringe projection profilometry [15], [17], [41], [61], [343], [344]. Drawing inspiration from these methods, we developed an algorithm tailored for computer vision applications, extending the capability to retrieve two phasemaps. Following the argument from [57], we define the following variables:

$$\begin{aligned} c_1(x,y) &= 1/2 \cdot b_1(x,y)e^{i\phi_1(x,y)} \\ c_2(x,y) &= 1/2 \cdot b_2(x,y)e^{i\phi_2(x,y)} \end{aligned} \tag{A.9}$$

Then, we rewrite the image intensity formula from Eq. A.8 using the definition of the complex cosine:

$$\begin{aligned} I(x,y) = a(x,y) + c_1(x,y)e^{i(2\pi f_1 x)} + c_1^*(x,y)e^{-i(2\pi f_1 x)} \\ + c_2(x,y)e^{i(2\pi f_2 y)} + c_2^*(x,y)e^{-i(2\pi f_2 y)} \end{aligned} \tag{A.10}$$

Where * denotes the complex conjugate. The 2D Fourier transform of the captured image $I(x,y)$ has then the following mathematical expression:

$$\begin{aligned} \mathscr{F}[I](f_x,f_y) = A(f_x,f_y) + C_1(f_x - f_1, f_y) \\ + C_1^*(f_x + f_1, f_y) + C_2(f_x, f_y - f_2) + C_2^*(f_x, f_y + f_2) \end{aligned} \tag{A.11}$$

The capital letters in Eq. A.11 represent the Fourier Transform of their respective functions. By leveraging the linearity and shifting properties of the Fourier Transform

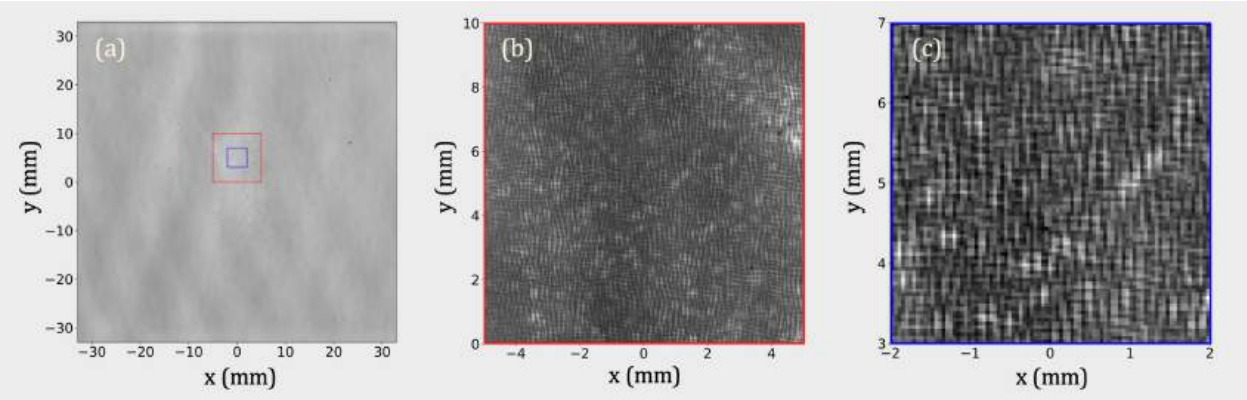


Figure A.3: (a) Image captured with our camera sensor. The spatial FoV of the camera is around 66mm × 66mm and the pixel resolution of the captured images is 1448×1448 pixels. We have cropped a window in the image at the central location pixels so that one can clearly see (b) the speckles (red window) and (c) the crossed vertical and horizontal fringes (blue window). The distance between fringes is about 3 pixels, near the sampling limit.

[345], we are able to derive Eq. A.11 from Eq. A.10. It should be noted that each of the functions (A,C_1,C_1^*,C_2,C_2^*) corresponds to a specific circular spectral region, as depicted in Fig. A.4a. As detailed in section 4 of the main manuscript, we have meticulously adjusted the aperture size within the experimental setup to ensure no overlap between these regions. Consequently, we can effectively isolate each region in the Fourier Transform $\mathscr{F}[I]$ of our experimental data for further analysis.

In line with the methodology detailed in section 2.2 of the main manuscript, we proceed to isolate the right-side spectral region that corresponds to the term $C_1(f_x - f_1, f_y)$. Initially, we estimate f_1, the spatial carrier frequency, as the center of the right-side spectral region. Subsequently, by shifting the spectrum using the estimated carrier frequency, we position f_1 at the center of the Fourier Transform (refer to Fig. A.4b). Following this, we apply a Hanning or Gaussian filter to effectively eliminate any unrelated signals (refer to Fig. A.4c). This sequence of steps results in successful isolation of the horizontal (right-hand side) spectral region, denoted as $\mathscr{F}_{\text{hor}}[I(x,y)] = C_1(f_x, f_y)$. Applying the in-

verse Fourier transform to this isolated region, we obtain:

$$\mathscr{F}^{-1}\{\mathscr{F}_{\text{hor}}[I(x,y)]\} = c_1(x,y) = 1/2 \cdot b_1(x,y)e^{i\phi_1(x,y)} \tag{A.12}$$

Following the definition provided in Eq. A.7, we know that $b_1(x,y) = |E_1^{\text{ref}}||E_1(x,y)|$. This leads us to derive Eq. 3 of the main manuscript:

$$\mathscr{F}^{-1}\{\mathscr{F}_{\text{hor}}[I(x,y)]\} \propto |E_1(x,y)|e^{i\phi_1(x,y)} = E_1(x,y) \tag{A.13}$$

In this manner, we effectively recover the speckle object field $E_1(x,y) \equiv E(\lambda_1)$. Applying a similar approach to isolate the vertical region $C_2(f_x, f_y - f_2)$ yields $E(\lambda_2)$.

Figs. A.5 a-d present the absolute amplitudes and the phases of each optical field at each optical wavelength. The speckle pattern is clearly visible in the amplitude images (refer to Fig. A.5 a,c), while the randomized phases are shown in Fig. A.5 b,d.

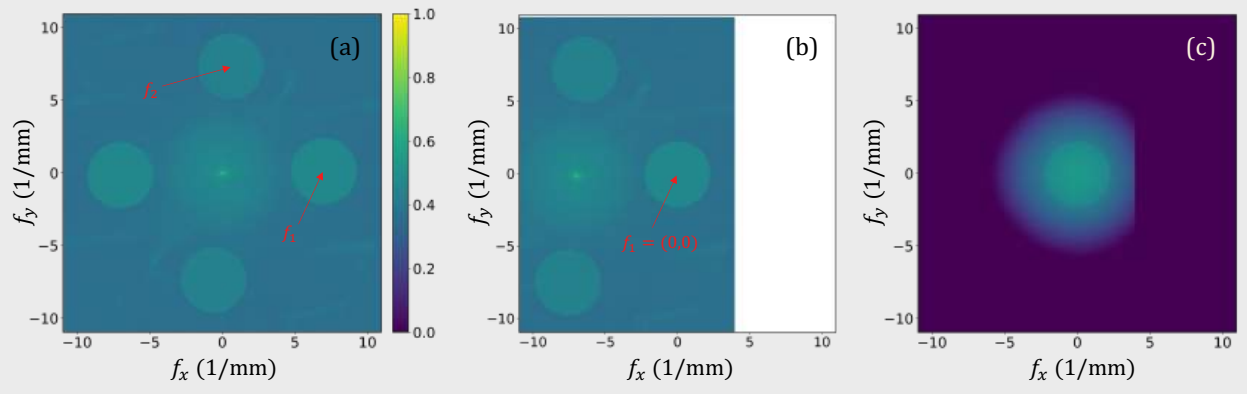


Figure A.4: (a) Fourier transform of the captured image $I(x,y)$. (b) Shifting the Fourier spectrum so that the carrier frequency f_1 corresponding to the horizontal reference beam is displaced to the center. (c) Applying a Gaussian filter to the shifted Fourier spectrum to uniformly remove any signals outside the region.

A.2 Synthetic Wavelength Interferometry

So far, we have recovered the object optical fields from Eq. A.3, finding their respective amplitude and phases. However, the optical phasemaps are randomized due to the speckle noise and do not provide the depth information of the macroscopic object. To extract the 3D shape of the object, we propose the following *thought experiment*: Consider the two object beams now arriving at a "virtual" sensor capable of resolving the rapidly oscillating optical waves. As the object beams $E_1(x, y, t)$ and $E_2(x, y, t)$ have slightly different wavelengths, their superposition $E_{\text{obj}}(x, y) = E_1(x, y, t) + E_2(x, y, t)$ generates a *beat*

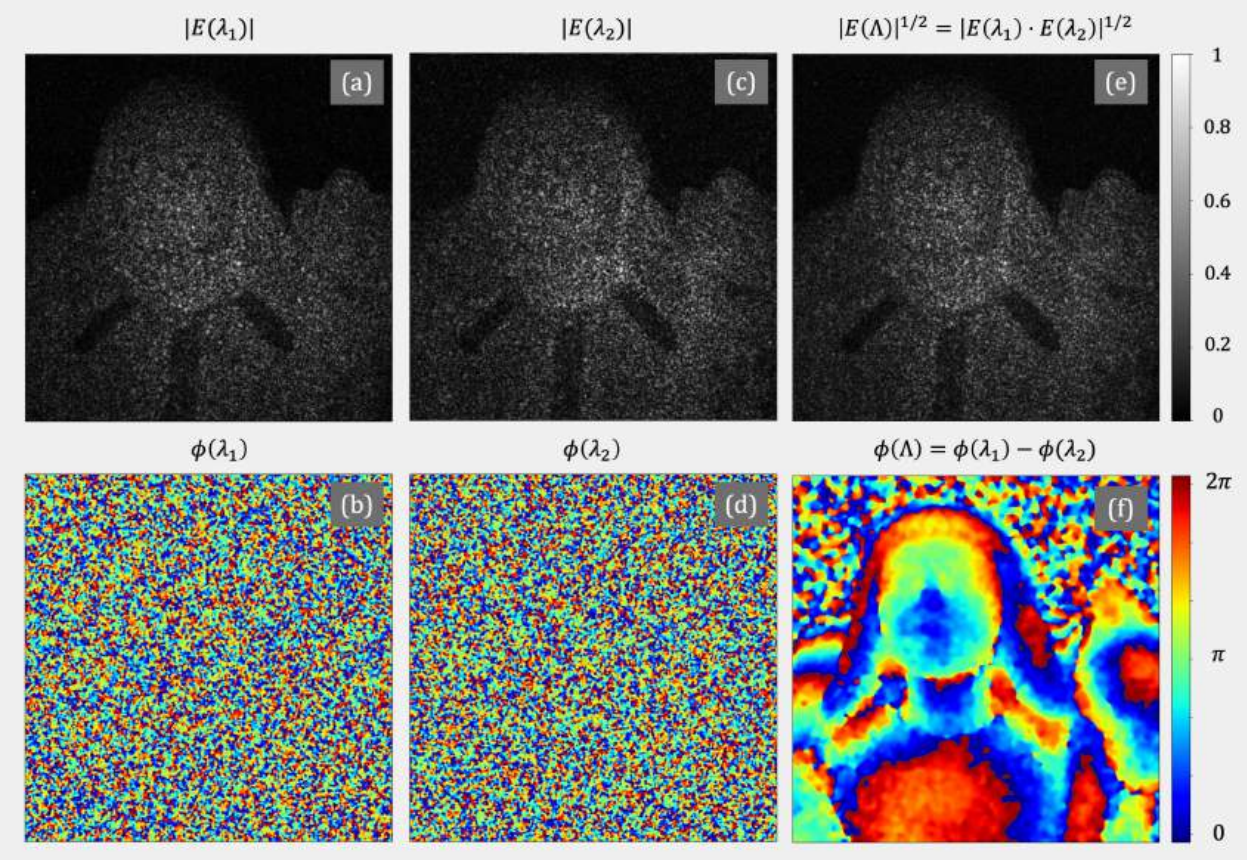


Figure A.5: (a) Amplitude and (b) phase of the object beam reflected by the object with wavelength λ_1. Similarly, (c) and (d) represent the amplitude and phase of the object beam reflected by the object with wavelength λ_2. Then, (e) and (f) show the calculated synthetic wave, Λ.

wave, a phenomenon well-known in optical physics [69]. The time-dependent intensity of this *simulated* beat at the ideal sensor can be determined as follows:

$$I_{\text{obj}}^{\text{virtual}}(x,y,t) \propto |E_1(x,y,t) + E_2(x,y,t)|^2 = \\ = |E_1(x,y)|^2 + |E_1(x,y)|^2 + E_{\Delta\omega}(x,y,t) + E_{\Sigma\omega}(x,y,t) \tag{A.14}$$

Considering that our virtual camera can capture the light energy at each particular instance, we no longer need to consider the time-average (as done in Eq. A.14). The first two terms of the beat intensity equation correspond to a DC term, representing static background illumination. The latter two terms relate to interferences and generate time-dependent fringes. Specifically, $E_{\Sigma\omega}(x,y,t)$ comprises the waves with high temporal frequencies, corresponding to the beat's carrier frequencies $\omega_1 + \omega_2$, $2\omega_1$, and $2\omega_2$. Conversely, $E_{\Delta\omega}(x,y)$ denotes to the slowly oscillating envelope with beat frequency $|\omega_1 - \omega_2|$ [69]. By substituting Eq. A.3 into Eq. A.14, we express the beat envelope explicitly as follows:

$$E_{\Delta\omega}(x,y,t) = |E_1(x,y)| \cdot |E_2(x,y)| \cdot \\ \exp\big((\omega_1 - \omega_2)t + \phi_1(x,y) - \phi_2(x,y)\big) + cc \tag{A.15}$$

We can determine the *beat wavelength* of this envelope wave using the following relationships:

$$\Lambda = \frac{2\pi c}{\Delta\omega} = \frac{2\pi c}{|\omega_1 - \omega_2|} = \frac{2\pi c}{|\frac{2\pi c}{\lambda_1} - \frac{2\pi c}{\lambda_2}|} = \frac{\lambda_1\lambda_2}{|\lambda_1 - \lambda_2|} \tag{A.16}$$

So far, we have worked with our designed *thought experiment*. However, it should be emphasized that, in the real single-shot experiment, the two object waves indeed superpose and form these time-varying interference patterns. Unfortunately, conventional

sensors cannot resolve $E_{\Delta\omega}(x, y, t)$ or $E_{\Sigma\omega}(x, y, t)$. To give some rough estimation, $\lambda_1 = 850$ nm and $\lambda_2 = 850.02$ nm lead to $\Sigma\omega \gg \Delta\omega = 1.3 \times 10^9$ 1/s, while conventional cameras have a presumably exposure time of around 10^{-3} seconds. Therefore, when we integrate Eq. A.14 over the exposure time of a *real* sensor, the average measurable intensity of the time-varying fringes blurs out.

Please note that in the double-shot method (see section 3.2 of the main manuscript), we recover the optical fields $\mathbf{E_1}(x, y, t)$ and $\mathbf{E_2}(x, y, t)$ sequentially. In that scenario, the object fields are produced at different time and do not superpose in the actual experiment, being the beat wave purely simulated. In either case, beat envelope is commonly denoted as $E(\Lambda) \equiv E_{\Delta\omega}(x, y, t)$ and called the *synthetic wave*. Since $E(\Lambda)$ changes over time, we should compute it at a fixed time (for example, at $t = 0$). We can then rewrite Eq. A.15 in a compact form as

$$E(\Lambda) = |E(\lambda_1)||E(\lambda_2)| \cos(\phi(\lambda_1) - \phi(\lambda_2)) \tag{A.17}$$

Where $\phi(\Lambda) = \phi(\lambda_1) - \phi(\lambda_2)$ is referred to as the *synthetic phasemap*. For the previous example with $\Delta\omega = 1.3 \times 10^9$ Hz, we would get a synthetic wavelength $\Lambda \approx 3.6$ mm, though we could tune our lasers for different λ_1 and λ_2 to get multiple synthetic wavelengths. The importance of our *thought experiment* and the computational construction of the *synthetic wave* is that this synthetic wavelength is now larger than the object surface roughness, and therefore:

While the optical waves are susceptible to microscopic optical path length differences generated by the rough object surface, the synthetic wavelength is larger and less sensitive, providing us with a phasemap that is not "speckled," i.e., not subject to random phase variations.

Please note that $E_{\Sigma\omega}(x, y, t)$ instead do not carry valuable information about the

macroscopic object surface variation, as the corresponding wavelengths are roughly similar to the optical wavelengths, i.e., in the nanometer range: $\lambda_{12} = \lambda_1\lambda_2/|\lambda_1+\lambda_2|$, $\lambda_{11} = \lambda_1/2$ and $\lambda_{22} = \lambda_2/2$. We shall neglect these terms in the following for our virtual beat wave.

An intuitive idea of the above explanation is provided in section 2.1 of the main manuscript. Fig. A.5f shows the synthetic phasemap, which encodes the depth information of the object surface via Eq. 1 of the main manuscript. Moreover, Fig. A.5e shows the square-root amplitude $|E(\Lambda)|^{1/2}$ of the synthetic wave. We plot the square-root amplitude to preserve the physical units of the electric field (W/m) and to see the clear correlation between $|E(\lambda_1)|$ and $|E(\lambda_2)|$ in Fig. A.5. If two speckle points are correlated in $|E(\lambda_1)|$ and $|E(\lambda_2)|$, then we will see a bright point in $|E(\Lambda)|^{1/2} = |E(\lambda_1) \cdot E(\lambda_2)|^{1/2}$. However, this bright point will dim as the correlation between both fields is lost. In our experiments, λ_1 is only slightly different from λ_2 (with less than 0.02 nm of difference). Therefore, we observe a high spectral correlation between the speckles of $|E(\lambda_1)|$ and $|E(\lambda_1)|$. For our current experiments, we typically use $\Lambda = 1, 3, 10, 15, 30, 45, 50$ mm (which are clearly larger than the surface roughness of our objects).

Observation: There exists an alternate way to derive the (synthetic) interference term $E(\Lambda)$ from Eq. A.17, which involves the *wave mixing* formula. This method is discussed in ref. [55] and aligns with Chapter 1 from [185]. In this context, $E(\lambda_1)$ and $E(\lambda_2)$ denote complex waves. The synthetic wave can then be expressed by the following *wave mixing* term:

$$E(\Lambda) = \text{Real}\big\{E(\lambda_1) \cdot E(\lambda_2)^*\big\} \tag{A.18}$$

While we can employ both complementary theories to derive the synthetic beat wave computationally, it is crucial to emphasize that the wave mixing process does not

happen in the actual single-shot experiment. In real-world settings, wave mixing may occur when two (or more) optical waves propagate through non-linear crystals, thereby generating another *actual* light wave at different frequencies, such as the second harmonic. In the single-shot experiment, $E(\Lambda)$ corresponds to the *envelope* of the beat wave between $E(\lambda_1)$ and $E(\lambda_2)$, which oscillates too rapidly to be discerned by our sensors. In other words, $E(\Lambda)$ is not an *actual* electromagnetic wave with wavelength Λ, but rather represents the amplitude modulation of two interfering waves.

A.3 Limitations of the double-shot method

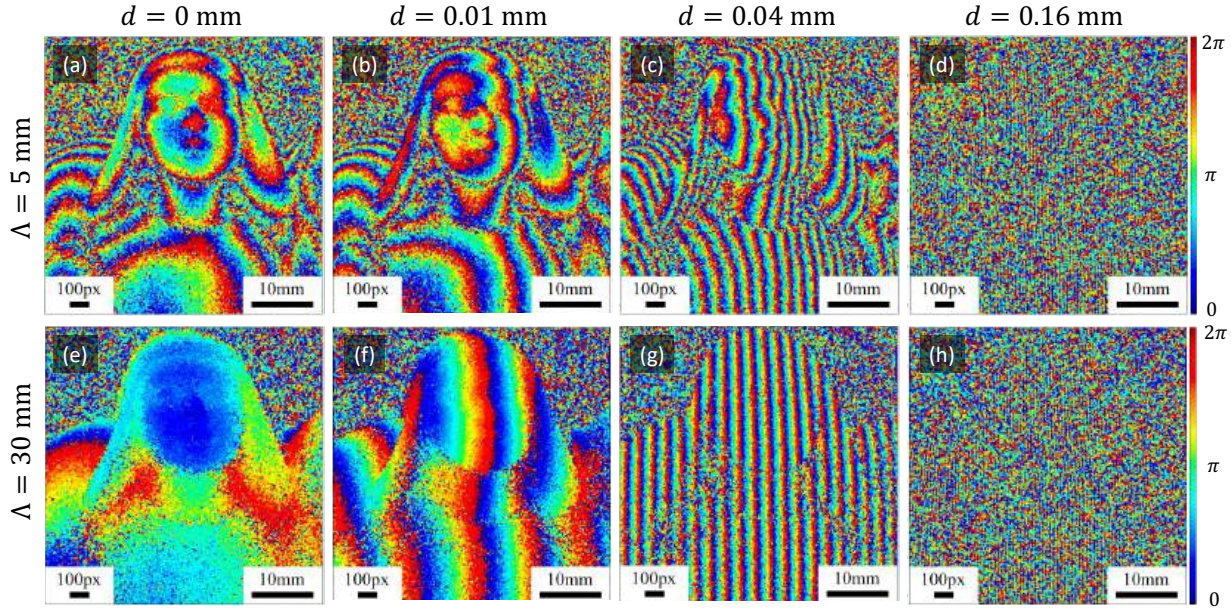


Figure A.6: Limitations of the double-shot method: The top row displays an analysis for a synthetic wavelength $\Lambda = 5$ mm with translations d of (a) 0, (b) 0.01, (c) 0.04, and (d) 0.16 mm. Similarly, the bottom row presents an analysis for $\Lambda = 30$ mm with translations of (e) 0, (f) 0.01, (g) 0.04, and (h) 0.16 mm.

In this section, we explain in detail why and how the double-shot method generally fails to capture moving objects, motivating the need for our proposed single-shot camera. In our explanations above, we assumed that the object remains static during the

two sequential images taken with the double-shot method. As mentioned in the main manuscript, the speckle patterns $E(\lambda_1)$ and $E(\lambda_2)$ would start to decorrelate if the object moves between captures. The loss in correlation leads to a degeneration of the synthetic phase map. Ultimately, a total decorrelation would lead to a meaningless synthetic field $\phi(\Lambda)$ that no longer provides the 3D information of the object. As described in [26], [55], the effect can be compared with the formation of "synthetic speckle".

How fast the sequentially captured fields $E(\lambda_1)$ and $E(\lambda_2)$ decorrelate during object motion depends the difference $|\lambda_1 - \lambda_2|$ between both wavelengths (spectral decorrelation), the surface roughness, and the resulting isoplanatic angle spanned with object surface and sensor geometry and described by the angular memory effect [55], [346].

For typically measured objects and setup configurations shown in this paper, it is important to estimate the "critical object movement speed", i.e., the smallest object movement speed when the double-shot method starts to fail. To estimate this speed, we have conducted the following experiment:

The double-shot experiment described in section 3.B of the main manuscript serves as a basis for our new experiment. The experiment is conducted using a single tunable laser, which first generates the illumination beam at wavelength λ_1 and later at λ_2. The wavelength-switching process of the tunable laser leads to some inherent time delays, resulting in a few seconds between two sequential captures. As we do not want to take the laser tuning time into account for our consideration, we perform a "stop-and-go" measurement of the object. In other words, we acquire $E(\lambda_1)$, and then move the object a small specific distance d before acquiring $E(\lambda_2)$. We use a high-precision translation stage to move the object to the exact positions. By this approach, we can approximate how a measurement with continuous object motion would look like where both fields $E(\lambda_1)$ and

$E(\lambda_2)$ would be directly taken after each other (without having to wait for the laser).

Fig. A.6 presents the recovered synthetic phasemaps for various displacements ($d =$ 0.0, 0.01, 0.04, and 0.16 mm) at two distinct wavelengths. The top and bottom rows correspond to the reconstructions at $\Lambda = 5$ mm and $\Lambda = 30$ mm, respectively. One can clearly observe the performance degradation when the displacement d increases. Compensating for the undesired "fringing" effects observed in the phasemaps is a complicated challenge that falls outside the scope of this project. For instance, it is important to clarify that a mere shift of the captured optical fields (in order to account for object displacement) is insufficient to avoid the artifacts. This is because the speckle pattern experiences further shift and decorrelation in relation to the moving object frame. Specifically, this occurs as the object is illuminated under different angles for each sequential shot.

We could optimistically suppose that it is *somehow* possible to eliminate this artifact and retrieve the true 3D object information. Nevertheless, it can be seen that beyond the limit $d \geq 0.16$ mm, either the correlation between the speckle patterns is lost, or the fringe frequency becomes too high (resulting in undersampled fringes). Considering this distance as the operational limit for our double-shot camera and assuming a standard camera frame rate of 60 fps, the maximum speed of a moving object allowed by this system would be less than 1 cm/s (or 0.4 inches/s). This severe constraint highlights the need for the proposed single-shot method.

A.4 Non-Line-of-Sight experiment

Section 3.4 of the main manuscript shows a first single-shot Non-Light-of-Sight (NLoS) experiment. To recap briefly, we used the object source (the end of an optical fiber that simultaneously emits light at λ_1 and λ_2) as a self-illuminating single-point object. This

object was hidden behind a 220 grit ground glass diffuser. Direct imaging of the object with the camera was not possible due to the fairly strong scattering (see Fig. 7 of the main manuscript). We focused the camera on the rough surface of the diffuser and captured a single-shot image $I(x, y)$. Using our proposed method (section 2.2 of the main manuscript), we recovered the synthetic field $E(\Lambda = 1\text{mm})$, which can be now interpreted as a *"synthetic hologram"* of the object. This is because $E(\Lambda = 1\text{mm})$ is not the field distribution directly on the object surface but rather the propagated field (since the camera was focused at the diffuser location and the object was placed further away). Similar to [55], we can now use the Angular Spectrum Method [345] to backpropagate the field into the hidden volume to obtain a reconstruction of the object.

As it was done in [55], we now want to determine the theoretical resolution of our object reconstruction by calculating the diffraction limit of our system at the synthetic wavelength. According to the Abbe diffraction limit (see [69] and [55]), the minimal laterally resolvable distance in the hidden volume is given by $\delta x = \frac{\Lambda}{2\text{NA}}$. In our case, the aperture is defined by the observable region on the diffuser (with a diameter of about $D = 66$ mm). With the object standoff distance of $z = 300$ mm, the numerical aperture calculates to $\text{NA} = n\sin(\theta) \approx \tan(\theta) = \frac{D/2}{z} \approx 33\text{mm}/300\text{mm} = 0.11$ and hence, the theoretical lateral resolution to $\delta x = \frac{\Lambda}{2\text{NA}} \approx \frac{1\text{mm}}{2 \cdot 0.11} \approx 4.55\text{mm}$. It can be seen in Fig. 7c of the main manuscript that this theoretical calculation roughly matches the radius of the reconstructed object (about $\delta_x^{\text{exp}} \approx 5$ mm radius), meaning that our shown NLoS measurement reached diffraction-limited performance at the synthetic wavelength.

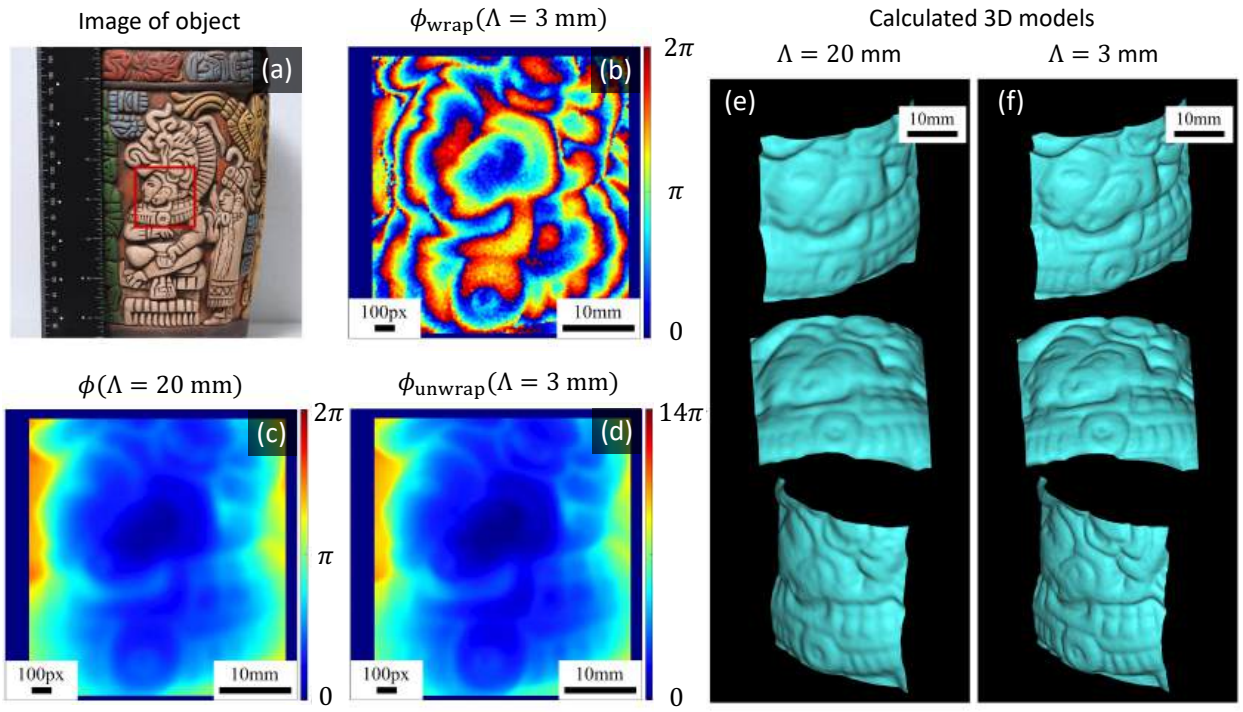


Figure A.7: Double-shot acquisition mode. The measured object is a clay pot approximately 20 cm in height. The camera FoV was cropped to about 40×40 mm^2: (a) Image of the small object with a window (red) showing the FoV. (b) Wrapped phasemap $\phi(\Lambda = 3\text{mm})$. (c) Acquired synthetic phasemap $\phi(\Lambda = 20\text{mm})$. (d) Phasemap $\phi^{\text{unwrap}}(\Lambda = 3\text{mm})$, unwrapped with our multi-frequency unwrapping algorithm. (e) 3D model of the object calculated from $\phi(\Lambda = 20\text{mm})$, and (f) from unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 3\text{mm})$.

A.5 Analysis and further results

A.5.1 Results using the multifrequency unwrapping algorithm

In the section 2.2 and 2.3 of the main manuscript, we show the results of the double-shot and single-shot methods using our multifrequency unwrapping algorithm (see Figs. 3-4 in the main manuscript). In this section, we show additional results to demonstrate the versatility of our approach.

For the double-shot method, we imaged a small clay pot with fine 3D reliefs (see

Fig. A.7c). Fig. A.7a displays an unwrapped synthetic phasemap of the object acquired at a large synthetic wavelength of $\Lambda = 20$ mm. Fig. A.7b shows a wrapped synthetic phasemap acquired at $\Lambda = 3$ mm.

Using our multifrequency unwrapping algorithm, one can find the unwrapped synthetic phasemap $\phi^{\text{unwrap}}(\Lambda = 3\text{mm})$ using $\phi(\Lambda = 20\text{mm})$ as the guidance phasemap. Figs. A.7 e-f show the object's 3D models, which are calculated from $\phi(\Lambda = 20\text{mm})$ and $\phi^{\text{unwrap}}(\Lambda = 3\text{mm})$ using Eq. 1 of the main manuscript. As can be seen, our method is able to impressively resolve the fine imprinted details of the clay pot in 3D.

In addition, we show similar results for the single-shot method to image a small bust figure (see Fig. A.8c). In this case, the large wavelength is $\Lambda = 50$ mm, and the smaller one, $\Lambda = 5$ mm.

A.5.2 Learning-based single-shot unwrapping algorithm

To exploit the main feature of our novel camera, we need to unwrap the synthetic phasemaps from a single shot image. Conventional single-shot unwrapping algorithms (such as those predefined in Matlab or Python) fail to unwrap our phasemaps due to their geometric complexity and the relatively large noise level. For that reason, we have built a Deep-Learning-based single-shot unwrapping algorithm. Compared to other state-of-the-art phase unwrapping procedures [51], [52], our approach is specifically tailored to our data structure.

As briefly discussed in the main manuscript, our approach defines a Convolutional Neural Network (CNN) with an encoder-decoder structure [77], [260], [347]. Inspired by the work from Perera *et al.* [75], we also added a Long Short-Term Memory (LSTM) module into our network. The LSTM modules are commonly used in Recurrent Neural Networks

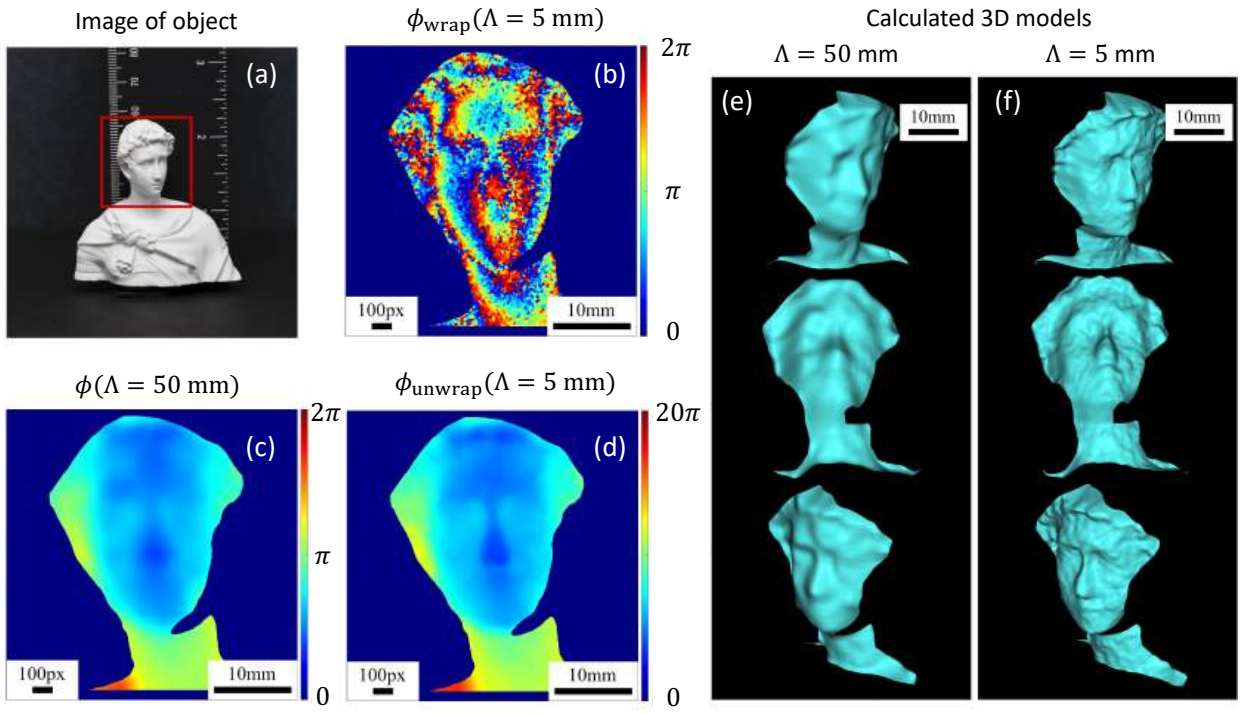


Figure A.8: Single-shot acquisition mode. The measured object is a plaster bust of approximately 6 cm in height. (a) Image of the small object with a window (red) showing the FoV. (b) Wrapped phasemap $\phi(\Lambda = 5\text{mm})$. (c) Acquired synthetic phasemap $\phi(\Lambda = 50\text{mm})$. (d) Phasemap $\phi^{\text{unwrap}}(\Lambda = 5\text{mm})$, unwrapped with our multi-frequency unwrapping algorithm. (e) 3D model of the object calculated from $\phi(\Lambda = 50\text{mm})$, and (f) from the unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 5\text{mm})$.

to solve the so-called "vanishing gradient" problem [78], [348]. Nevertheless, it is also becoming a powerful tool in CNNs to learn the spatial relations between local features [342], [349]. In phase unwrapping problems, these spatial relations between local features are essential to identify the phase jumps.

The LSTM module was added between the encoder and decoder networks: It takes in the local features from the encoded image and provides their local spatial relations to the decoder network. The structure of the network is shown in Fig. A.9. One can observe that each layer from the encoder is concatenated with a corresponding decoder layer [347], [350]. This concatenation provides information to the decoder that was oth-

erwise lost during downsampling (or encoding phase). The network predicts more accurately if we combine the extracted features (from the encoder) with the original values (given through the concatenation).

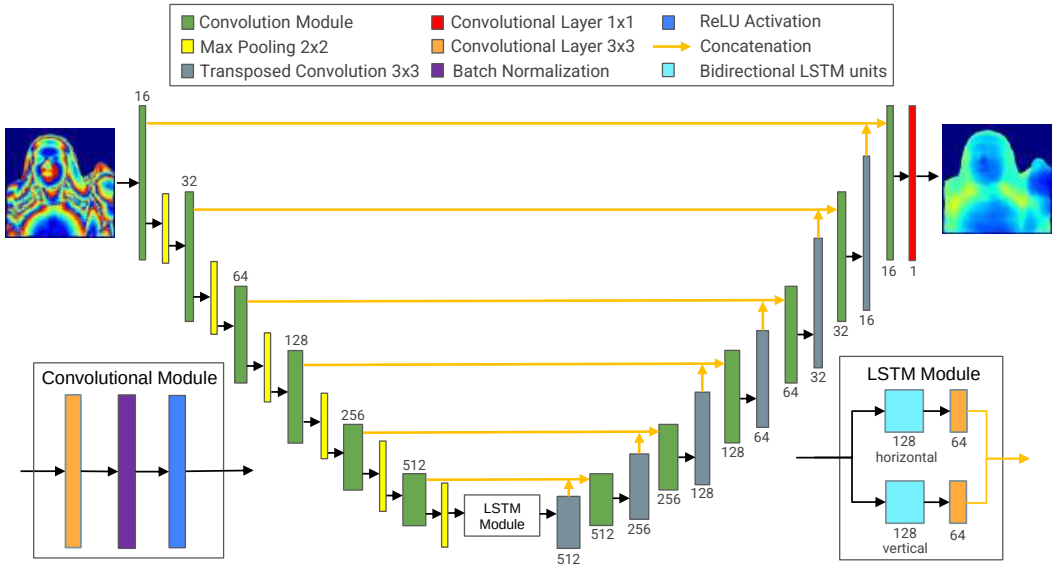


Figure A.9: Architecture of the Deep Learning algorithm.

The training dataset employs both experimental and simulated data. For the experimental part, we use wrapped synthetic phasemaps directly obtained from our camera. To build the input-output pairs consisting of wrapped and unwrapped phasemaps, we use our multifrequency unwrapping algorithm.

For the simulated data, we first created realistic 3D models of common objects using *Mitsuba 3* [351]. Given the depth information $D(x, y) = z$ from *Mitsuba*, one can quickly generate unwrapped phasemaps $\phi^{\text{unwrap}}(x, y, \Lambda)$ for different synthetic wavelengths Λ using the Eq. 1 of the main manuscript. Eventually, phase wrapping is performed using the *modulus* operation. To have more realistic simulated data, we have introduced a certain

noise level. We have added Gaussian noise $(\mu = 0, \sigma = \pi/6)$ to each wrapped phasemap pixel that randomly reduces or increases the pixel value.

We used 22 wrapped-unwrapped phasemap pairs retrieved from real experiments at relatively small synthetic wavelengths (around 3-10 mm). We increased this experimental dataset to 66 by including rigid transformations of the phasemaps (reflection, rotation, and translations). In addition, we simulated 21 wrapped-unwrapped phasemap pairs using the Mitsuba renderer *Mitsuba*. Implementing also rigid transformations of the simulated phasemaps, different noise levels, and different synthetic wavelengths, we increased the simulated dataset size up to 1512. The training dataset (with real and simulated data) had 1578 data. Fig. A.10 shows one of the simulated pairs of data from the training dataset: A wrapped phasemap that serves as input and its associated nicely unwrapped phasemap corresponding to its output.

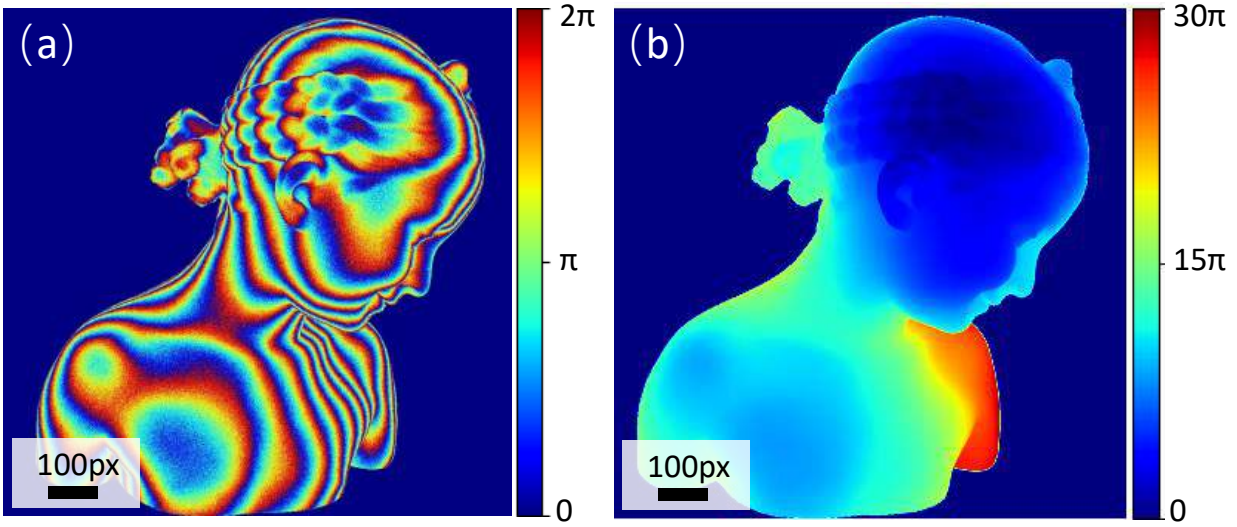


Figure A.10: Simulated data for the training of the Deep Learning unwrapping algorithm. We employed *Mitsuba* [351] to simulate realistic 3D objects. (a) Input: Wrapped noisy phasemap. (b) Output: Unwrapped phasemap.

The loss function for our network measures the mean squared error (MSE) between the ground truth unwrapped phasemap and the predicted phasemap. During the training

process, the parameters of the CNN are optimized so that the model produces accurate unwrapped output phasemaps out of the wrapped input phasemaps. Once the model is trained, we test the model's performance using the wrapped phasemaps of new objects. We emphasize again that the unwrapping model has neither seen the test data phasemaps before nor any other measurements of the test objects (e.g., phasemaps at different wavelengths, measurements of the same object taken from a different perspective, etc.).

To improve the network's accuracy, we should also mention that we divided the training according to different synthetic wavelengths (relatively small wavelengths between 3-10 mm, or rather large wavelengths around 30-45 mm). Besides the result provided in the main manuscript in Fig. 7, we provide in Fig. A.11 another result for the measurement of a bust.

A.5.3 Smoothing parameters for shown images

For the sake of transparency and reproducibility, we provide below the smoothing parameters we have employed for our results (both in the main manuscript and the Supp. Mat.).

- **Fig. 3 (a)** of the main manuscript. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 45\text{mm})$ with $\sigma = 4$ and a kernel size $K = 20 \times 20$. Then, we employed the median filter on the phasemap, with a kernel size $K = 21 \times 21$ pixels. Finally, we used a strong Gaussian filter on the phasemap with $\sigma = 30$ and $K = 121 \times 121$ pixels.

- **Fig. 3 (d)** of the main manuscript. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 10\text{mm})$ with $\sigma = 4$ and $K = 20 \times 20$ pixels. Then, we employed the median filter on the phasemap with $K = 41 \times 41$ pixels. Finally, we

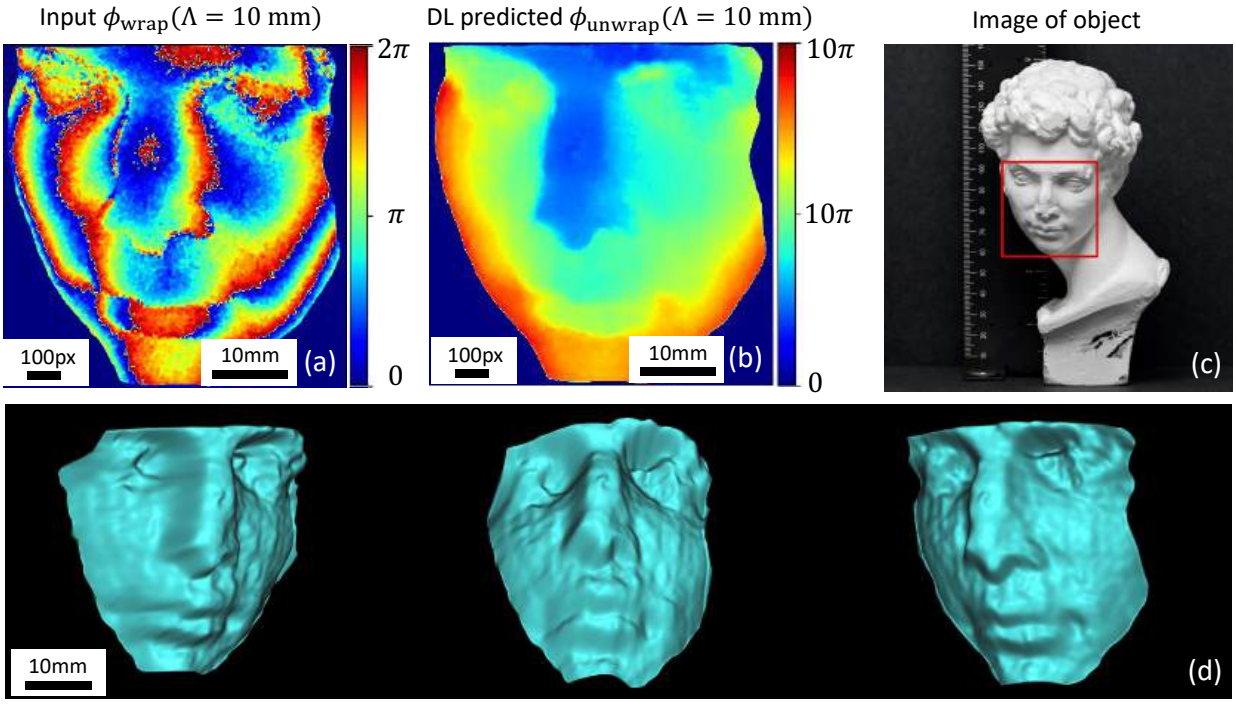


Figure A.11: Single-shot deep learning-based phase unwrapping method. (a) Wrapped phasemap $\phi(\Lambda = 10\text{mm})$. (b) Unwrapped phasemap predicted by the neural network approach. (c) Image of the small object with a window (red) showing FoV. (d) 3D model of the object calculated from unwrapped phasemap $\phi^{\text{unwrap}}(\Lambda = 10\text{mm})$.

used the Gaussian filter on the unwrapped phasemap with $\sigma = 9$ and $K = 37 \times 37$ pixels.

- **Fig. 4 (a)** of the main manuscript. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 50\text{mm})$ with $\sigma = 4$ and $K = 20 \times 20$ pixels. Then, we employed the median filter on the phasemap with $K = 21 \times 21$ pixels. Finally, we used the Gaussian filter on the phasemap with $\sigma = 40$ and $K = 161 \times 161$ pixels.

- **Fig. 4 (d)** of the main manuscript. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 10\text{mm})$ with $\sigma = 10$ and $K = 10 \times 10$ pixels. Then, we employed the median filter on the phasemap with $K = 41 \times 41$ pixels. Finally, we used the Gaussian filter on the unwrapped phasemap with $\sigma = 9$ and $K = 37 \times 37$

pixels.

- **Fig. 5 (b) and (d)** of the main manuscript. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 30\text{mm})$ with $\sigma = 10$ and $K = 60 \times 60$ pixels. Then, we employed the median filter on the phasemap with $K = 21 \times 21$ pixels. We additionally applied a bilateral filter using the predefined Matlab function *imbilatfilt* using 1 degree of smoothing a $\sigma = 20$. Finally, we used the Gaussian filter on the phasemap with $\sigma = 20$ and $K = 81 \times 81$ pixels.

- **Fig. A.7 (a)**. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 20\text{mm})$ with $\sigma = 4$ and $K = 20 \times 20$ pixels. Then, we employed the median filter on the phasemap with $K = 21 \times 21$ pixels. Finally, we used the Gaussian filter on the phasemap with $\sigma = 20$ and $K = 81 \times 81$ pixels.

- **Fig. A.7 (d)**. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 3\text{mm})$ with $\sigma = 4$ and $K = 20 \times 20$ pixels. Then, we employed the median filter on the phasemap with $K = 41 \times 41$ pixels. Finally, we used the Gaussian filter on the unwrapped phasemap with $\sigma = 9$ and $K = 37 \times 37$ pixels.

- **Supplementary Fig. A.8 (a)**. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 50\text{mm})$ with $\sigma = 4$ and $K = 20 \times 20$ pixels. Then, we employed the median filter on the phasemap with $K = 21 \times 21$ pixels. Finally, we used the Gaussian filter on the phasemap with $\sigma = 40$ and $K = 161 \times 161$ pixels.

- **Supplementary Fig. A.8 (d)**. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 5\text{mm})$ with $\sigma = 2$ and $K = 10 \times 10$ pixels. Then, we employed the median filter on the phasemap with $K = 41 \times 41$ pixels. Finally, we used the Gaussian filter on the unwrapped phasemap with $\sigma = 9$ and $K = 37 \times 37$ pixels.

- **Figure 6b of the paper and Fig. A.11b**. We first applied a 2D Gaussian filter on the complex synthetic field $E(\Lambda = 10\text{mm})$ with $\sigma = 7$ and $K = 40 \times 40$ pixels. Then, we employed the median filter on the unwrapped phasemap with $K = 21 \times 21$ pixels. Finally, we used the Gaussian filter on the unwrapped phasemap with $\sigma = 10$ and $K = 41 \times 41$ pixels.

Appendix B

Supplementary Material for Thin Solid Films Spectroscopy

B.1 Swanepoel transmission and reflection coefficients

Following the notation from Swanepoel [164], we define the following coefficients for Section 3.1.4 (see next page):

$$\mathbf{r}_{ij} = r_{ij} + ir'_{ij}$$

$$r_{ij} = \frac{(n_i^2 - n_j^2) + (\kappa_i^2 - \kappa_j^2)}{(n_i + n_j)^2 + (\kappa_i + \kappa_j)^2}$$

$$r'_{ij} = \frac{2(n_i\kappa_j - n_j\kappa_i)}{(n_i + n_j)^2 + (\kappa_i + \kappa_j)^2}$$

$$R_{ij} = |\mathbf{r}_{ij}|^2 = r_{ij}^2 + r_{ij}'^2$$

$$\begin{aligned}
a &= 1 + R_{01}R_{12}x_1^2 + R_{01}R_{23}x_1^2x_2^2 + R_{12}R_{23}x_2^2 + \\
&\quad 2(r_{01}r_{12}(1 + R_{23}x_2^2) - r'_{01}r'_{12}(1 - R_{23}x_2^2))x_1\cos(\phi_1) + \\
&\quad 2(r_{01}r'_{12}(1 - R_{23}x_2^2) + r'_{01}r_{12}(1 + R_{23}x_2^2))x_1\sin(\phi_1) \\
e &= R_{01} + R_{12}x_1^2 + R_{23}x_1^2x_2^2 + R_{01}R_{12}R_{23}x_2^2 + \\
&\quad 2(r_{01}r_{12}(1 + R_{23}x_2^2) + r'_{01}r'_{12}(1 - R_{23}x_2^2))x_1\cos(\phi_1) + \\
&\quad 2(r_{01}r'_{12}(1 - R_{23}x_2^2) - r'_{01}r_{12}(1 + R_{23}x_2^2))x_1\sin(\phi_1) \\
b &= r_{12}r_{23}(1 + R_{01}x_1^2) - r'_{12}r'_{23}(1 - R_{01}x_1^2)x_2 + \\
&\quad (r_{01}r_{23}(1 + R_{12}) - r'_{01}r'_{23}(1 - R_{12}))x_1x_2\cos(\phi_1) + \\
&\quad (r_{01}r'_{23}(1 - R_{12}) + r'_{01}r_{23}(1 + R_{12}))x_1x_2\sin(\phi_1) \\
f &= r_{12}r_{23}(x_1^2 + R_{01})x_2 + r'_{12}r'_{23}(x_1^2 - R_{01})x_2 + \\
&\quad (r_{01}r_{23}(1 + R_{12}) + r'_{01}r'_{23}(1 - R_{12}))x_1x_2\cos(\phi_1) + \\
&\quad (r_{01}r'_{23}(1 - R_{12}) - r'_{01}r_{23}(1 + R_{12}))x_1x_2\sin(\phi_1) \\
c &= (r_{12}r'_{23}(1 + R_{01}x_1^2) + r'_{12}r_{23}(1 - R_{01}x_1^2))x_2 + \\
&\quad (r_{01}r'_{23}(1 + R_{12}) + r'_{01}r_{23}(1 - R_{12}))x_1x_2\cos(\phi_1) - \\
&\quad (r_{01}r_{23}(1 - R_{12}) - r'_{01}r'_{23}(1 + R_{12}))x_1x_2\sin(\phi_1) \\
g &= r_{12}r'_{23}(x_1^2 + R_{01})x_2 - r'_{12}r_{23}(x_1^2 - R_{01})x_2 + \\
&\quad (r_{01}r'_{23}(1 + R_{12}) - r'_{01}r_{23}(1 - R_{12}))x_1x_2\cos(\phi_1) - \\
&\quad (r_{01}r_{23}(1 - R_{12}) + r'_{01}r'_{23}(1 + R_{12}))x_1x_2\sin(\phi_1) \\
u &= 1 + R_{01}R_{12}x_1^2 - R_{01}R_{23}x_1^2x_2^2 - R_{12}R_{23}x_2^2 + \\
&\quad 2(r_{01}r_{12}(1 - R_{23}x_2^2) - r'_{01}r'_{12}(1 + R_{23}x_2^2))x_1\cos(\phi_1) + \\
&\quad 2(r_{01}r'_{12}(1 + R_{23}x_2^2) + r'_{01}r_{12}(1 - R_{23}x_2^2))x_1\sin(\phi_1) \\
w &= 1 + R_{01}R_{12}x_1^2 + 2(r_{01}r_{12} - r'_{01}r'_{12})x_1\cos(\phi_1) + \\
&\quad 2(r_{01}r'_{12} + r'_{01}r_{12})x_1\sin(\phi_1) \\
h &= ((1 + r_{01})^2 + r_{01}'^2)((1 + r_{12})^2 + r_{12}'^2)((1 + r_{23})^2 + r_{23}'^2)x_1x_2 \\
a'_1 &= 1 + R_{01}R_{12}x_1^2 - R_{01}R_{23}x_1^2x_2^2 - R_{12}R_{23}x_2^2 \\
b'_1 &= \left(r_{01}r_{12}(1 - R_{23}x_2^2) - r'_{01}r'_{12}(1 + R_{23}x_2^2)\right)x_1 \\
c'_1 &= \left(r_{01}r'_{12}(1 + R_{23}x_2^2) + r'_{01}r_{12}(1 - R_{23}x_2^2)\right)x_1
\end{aligned}$$

B.2 Substrate absorption model for simulations

In Section 3.1.4, we explain the light propagation through our sample of interest: a thin film deposited over a substrate. It should be noted that in real samples, the absorption of the glass substrate varies with wavelength. For instance, the common Borofloat33 (a type of borosilicate substrate) exhibits significant absorption in the UV range (below 350 nm) and the NIR range (above 2100 nm) but is nearly transparent in the middle of the visible spectrum. Despite this quasi-transparency, subtle variations in transmission and reflection still occur. Researchers often limit their analysis to regions of evident transparency, though completely eliminating absorption effects is not feasible in practice (e.g., see the slight variations in substrate absorption lines in Fig. 3 of [183]). In our study, we use a simulated substrate with a constant weak absorption across the spectrum to provide a generalized analysis of errors in transmission and reflection formulas. This approach ensures that our findings are applicable to any type of substrate, as real substrates have diverse absorption curves.

B.3 High frequency noise in spectrophotometric measures

In Section 3.1.5, we analyze the exact formula for uniform films. There, we argue that the superposition of infinite waves with a phase delay range exceeding 2π will effectively average out. Note that the phase delay change $\Delta\delta_2$ depends not only on the bandwidth but also on the central wavelength. As the central wavelength increases, the change in the phase delay decreases, making interference effects from the substrate more likely to appear in the sample's transmittance and reflectance. This can result in subtle noise at longer wavelengths. Most modern spectrophotometers [103] can mitigate this high-frequency noise, as well as photometric noise, using smoothing techniques, such as the

Savitzky-Golay filter.

Appendix C

Supplementary Material for Gamma-ray radiation detector

C.1 Solving diffusion-repulsion equation numerically

To solve the partial differential equation from Eq. 4.6 for the total charge $Q(R,t)$, we employ a classical approach, using the Implicit Finite Difference Method (IFDM) to discretize the equation and then solving the resulting non-linear system with the Newton-Raphson method.

The equation is discretized in both space and time. We use a uniform grid in the radial coordinate $R \in (0, 400]$ µm, where the radial domain is divided into $N = 400$ points. The time range $t \in (0, 1]$ µs is discretized in $T = 200$ time steps.

The first and second spatial derivative is approximated using first-order central finite differences for the interior points and one-sided differences at the boundaries. There-

fore, at the inner points, we would have:

$$\partial_{RR}Q \approx \frac{Q_{i+1} - 2Q_i + Q_{i-1}}{\Delta R^2} \tag{C.1a}$$

$$\partial_R Q \approx \frac{Q_{i+1} - Q_{i-1}}{2\Delta R} \tag{C.1b}$$

Considering these spatial derivatives, we discretize Eq. 4.6 using an implicit scheme as:

$$\frac{Q_i^{t+1} - Q_i^t}{\Delta t} = D\left(\partial_{RR}Q^{t+1} - \frac{2}{R_i}\partial_R Q^{t+1}\right) - \frac{\mu Q_i^{t+1}}{4\pi\epsilon R_i^2}\,\partial_R Q^{t+1} \tag{C.2}$$

The non-linear system of equations defined in for different time and spatial steps is then solved using the Newton-Raphson method. Considering the update of the new time step $t+1$, we define the function F for different spatial steps i as:

$$\begin{aligned} F[i] = &-Q_i^{t+1} + Q_i^t - \Delta t\, D\left(\partial_{RR}Q_i^{t+1} - \frac{2}{R_i}\partial_R Q_i^{t+1}\right) + \dots \\ &+ \Delta t \frac{\mu Q_i^{t+1}}{4\pi\epsilon R_i^2}\,\partial_R Q_i^{t+1} \end{aligned} \tag{C.3a}$$

The Jacobian matrix is zero everywhere except near the diagonal, specifically at the space steps $i+1$, i, and $i-1$, where it is defined as:

$$J[i, i-1] = -\Delta t D\left(\frac{1}{\Delta R^2} + \frac{1}{R_i\,\Delta R}\right) - \Delta t\,\frac{\mu Q_i^{t+1}}{4\pi\epsilon R_i^2}\frac{1}{2\,\Delta R}$$

$$J[i, i+1] = -\Delta t D\left(\frac{1}{\Delta R^2} - \frac{1}{R_i\,\Delta R}\right) + \Delta t\,\frac{\mu Q_i^{t+1}}{4\pi\epsilon R_i^2}\frac{1}{2\,\Delta R}$$

$$J[i, i] = 1 + \Delta t\, D\,\frac{2}{\Delta R^2} + \Delta t\,\frac{\mu}{4\pi\epsilon R_i^2}\partial_R Q_i^{t+1}$$

We can then update the total charge at the next time step $t+1$ for every spatial step by iterating the Newton-Raphson algorithm during k rounds:

$$[Q^{t+1}]_{k+1} \leftarrow [Q^{t+1}]_k - (J^{-1}F)\big([Q^{t+1}]_k\big) \tag{C.5}$$

We repeat the iterations until the update during two consecutive iterations k and $k+1$ is below the tolerance threshold 10^{-9}.

C.2 Generalization of the diffusion-repulsion algorithm

We present here the parameter found with our correlation study from Section 4.1.5.4 to generalize our diffusion-repulsion algorithm to semiconductor materials with different properties under diverse injection energies:

- $a_1 = -8.80 \times 10^{-6}$, $b_1 = 5.08 \times 10^{-7}$, $c_1 = 1.61$
- $d_1 = -2.23 \times 1-^{-7}$, $e_1 = -2.45 \times 10^{-9}$, $f_1 = 10^{-4}$
- $a_2 = 9.75 \times 10^{-7}$, $b_2 = -2.45 \times 10^{-9}$, $c_2 = 0.95$
- $d_2 = 1.47 \times 1-^{-8}$, $e_2 = -9.11 \times 10^{-10}$, $f_2 = 0$
- $a_3 = 3.34$, $b_3 = 6.19$, $c_3 = 80.01$, $d_3 = -0.17$, $e_3 = 2.36$

Vita

Manuel Ballester Matito was born in Cadiz, Spain. He earned a B.S. in Mathematics with a focus on Mathematical Engineering from the University of Cadiz (UCA) in 2018. He then pursued an elite M.S. in Advanced Optical Technologies at Friedrich-Alexander University (FAU), Germany, with a double major in Physics of Light and Computational Optics.

In 2020, he joined the Department of Computer Science at Northwestern University as a predoctoral researcher and formally entered the Ph.D. program in 2021, following the interfaces track (graphics and interactive media). His research centers on computational optics and sensing, particularly using physics-based machine learning and advanced optimization techniques. He earned an M.S. in Computer Science in 2023 as part of the doctoral program.

His work spans applications such as interferometric 3D imaging, spectroscopic analyses of thin film materials, gamma-ray detector modeling, and physics-informed neural networks for astrophysical imaging. Throughout his Ph.D. path, he has collaborated with Siemens Healthineers, The Art Institute of Chicago, the University of Arizona, the Eastern Switzerland University of Applied Sciences, and several international institutions. His Ph.D. research was supported by Siemens Healthineers and advised by a multidisciplinary team across machine learning, physics, and imaging.