# Physics-Informed Neural Networks for Multi-Spectral Photon Diffusion in Biological Tissue

Sarvagya Singh

singhsarvagya260508@gmail.com

March 15, 2026

**Abstract**

Accurate modeling of light transport in biological tissue underpins the planning of laser surgical procedures, the interpretation of optical diagnostic measurements, and the development of non-invasive disease monitoring systems. Monte Carlo methods, while representing the gold standard for photon transport simulation, impose computational costs that render real-time clinical applications and iterative inverse problems impractical. This paper presents a Physics-Informed Neural Network framework for solving the steady-state photon diffusion equation with Robin boundary conditions across multiple therapeutic wavelengths. The architecture incorporates source-transformation decomposition to regularize near-field singularities, non-dimensionalisation using penetration depth as the characteristic length scale, and self-adaptive Neural Tangent Kernel-inspired loss weighting that eliminates the need for manual hyperparameter tuning. A parametric extension handles arbitrary tissue optical properties without retraining, while Monte Carlo Dropout provides Bayesian uncertainty quantification. A two-stage inverse encoder recovers absorption and reduced scattering coefficients from surface reflectance measurements, and photo-thermal coupling via the Pennes bioheat equation enables tissue damage risk assessment. Ablation studies demonstrate that boundary condition enforcement constitutes the most critical physics component, with its removal increasing validation error by three orders of magnitude. Trained models achieve logarithmic mean squared error of approximately 0.08 after five thousand epochs, and the exported model executes inference in under two milliseconds on central processing units, establishing feasibility for real-time clinical deployment.

**Keywords:** physics-informed neural networks, photon diffusion, biophotonics, diffusion

approximation, self-adaptive loss weighting, uncertainty quantification, inverse problem, laser-tissue interaction

# Contents

# 1 Introduction

The quantitative understanding of light transport in biological tissue forms the foundation upon which numerous clinical technologies rest, ranging from photodynamic therapy dosimetry and diffuse optical tomography to the planning of dermatological laser treatments. At the mesoscale relevant to therapeutic laser applications, where depths extend from one to twenty millimeters and wavelengths span four hundred to nine hundred nanometers, photon transport is governed by the radiative transfer equation. Under conditions where scattering substantially exceeds absorption, the diffusion approximation yields a tractable second-order partial differential equation that has served as the workhorse for tissue optics modeling for several decades (1; 2).

Monte Carlo simulation of individual photon trajectories remains the gold standard for photon transport modeling because it introduces no approximation beyond the assumed phase function (3; 4). However, the computational burden is substantial: a single two-dimensional fluence map of adequate statistical quality requires the tracing of over one million photons and consumes several minutes of central processing unit time. Iterative inverse problems that demand thousands of forward evaluations therefore become computationally prohibitive in clinical settings where real-time feedback is essential.

Physics-Informed Neural Networks, introduced by Raissi and colleagues (5), embed governing partial differential equations directly into the neural network loss function. Once trained, these networks perform forward inference in microseconds, and their differentiable nature enables gradient-based solution of inverse problems. Recent applications have demonstrated the viability of PINNs for radiative transfer problems in simplified geometries (6; 7), yet the literature lacks a comprehensive treatment addressing multi-spectral training, heterogeneous tissue geometry, adaptive loss weighting, uncertainty quantification, and clinical decision support within a unified framework.

The development of robust PINN architectures for photon transport faces several distinct challenges that have motivated the present work. The near-source region exhibits a $1/r$ singularity that destabilizes early training and produces large partial differential equation residuals. Standard loss formulations combine data fidelity, boundary condition enforcement, and PDE residual terms with fixed weights, yet these components typically span four to five orders of magnitude in their raw values, causing gradient pathology where dominant terms starve others of optimization attention (10). Multi-spectral training presents the risk of mode collapse, where networks trained across multiple wavelengths converge to intermediate solutions that fail to represent any specific wavelength accurately. Inverse problems in tissue optics are fundamentally ill-posed, as surface reflectance measurements provide limited information about the underlying optical property distribution. Finally, the translation of fluence predictions into clinical utility requires coupling to thermal models that can assess tissue damage risk.

This paper addresses these challenges through several technical contributions. A source-transformation decomposition removes the near-source singularity by separating the solution into an analytic Beer-Lambert envelope and a smooth residual learned by the network. Self-adaptive Neural Tangent Kernel-inspired weighting dynamically balances the competing loss terms without manual tuning, responding to the observation that boundary condition satisfaction typically represents the hardest constraint (11). Penetration-depth conditioning resolves multi-spectral mode collapse by providing the network with explicit information about the wavelength-dependent optical properties. A parametric surrogate architecture enables inference across arbitrary tissue types without retraining. Bayesian uncertainty quantification via Monte Carlo Dropout produces calibrated confidence intervals that guide clinical decision-making. A two-stage inverse encoder recovers optical properties from surface reflectance, and coupling to the Pennes bioheat equation enables thermal damage assessment.

The remainder of this paper is organized as follows. The Methods section derives the governing equations and describes the network architecture and training protocol. Advanced extensions including multi-layer tissue modeling and uncertainty quantification are documented in the subsequent section. Quantitative results are presented in the Results section, followed by discussion of limitations and future directions. The paper concludes with a summary of principal findings.

# 2 Methods

## 2.1 Governing Equations

The steady-state diffusion approximation governs photon transport in the present work. The fluence rate $\Phi(x, z)$, expressed in watts per square millimeter, depends upon lateral position $x$ and depth $z$, with the tissue surface located at $z = 0$. The diffusion approximation takes the form of a Helmholtz equation:

$$D\,\nabla^2\Phi - \mu_a\,\Phi + S = 0, \qquad (x, z) \in \Omega, \tag{1}$$

where the diffusion coefficient $D$ depends upon the absorption coefficient $\mu_a$ and the reduced scattering coefficient $\mu_s'$ according to:

$$D = \frac{1}{3(\mu_a + \mu_s')}. \tag{2}$$

The reduced scattering coefficient $\mu_s' = \mu_s(1 - g)$ incorporates the scattering coefficient $\mu_s$ and the anisotropy factor $g$. The source term $S$ represents the primary laser beam; in the present formulation this term is absorbed into the decomposition described below, yielding $S = 0$ within the domain $\Omega$.

At the tissue-air interface, partial internal reflection leads to the index-mismatched Robin boundary condition (9):

$$\Phi(x,0) - A\,D\,\frac{\partial \Phi}{\partial z}\bigg|_{z=0} = 0, \tag{3}$$

where the reflectance coefficient $A$ is given by:

$$A = \frac{2}{1 - R_{\text{eff}}}, \qquad R_{\text{eff}} \approx -1.440\,n^{-2} + 0.710\,n^{-1} + 0.668 + 0.0636\,n, \tag{4}$$

with $n = 1.37$ representing the tissue refractive index (9).

The optical parameters employed in this study are drawn from the literature (8) and summarized in Table 1. These values represent human dermis at three therapeutic wavelengths: 450 nanometers, 550 nanometers, and 650 nanometers. The absorption coefficient decreases with increasing wavelength while the reduced scattering coefficient exhibits a similar trend, reflecting the spectral dependence of tissue chromophores and scattering structures.

Table 1: Dermis optical parameters at three therapeutic wavelengths (8).

| $\lambda$ | $\mu_a$ [mm$^{-1}$] | $\mu_s'$ [mm$^{-1}$] | $g$ | $n$ |
|---|---|---|---|---|
| 450 nm | 0.40 | 2.00 | 0.90 | 1.37 |
| 550 nm | 0.30 | 1.65 | 0.90 | 1.37 |
| 650 nm | 0.20 | 1.35 | 0.90 | 1.37 |

## 2.2 Source-Transformation Decomposition

The fluence field exhibits a $1/r$ singularity near the laser entry point that produces large partial differential equation residuals and destabilizes early training. To address this pathology, the solution is decomposed into an analytic primary field and a smooth scattering correction (17):

$$\Phi(x,z) = \Phi_{\text{prim}}(x,z) + \Phi_{\text{sc}}(x,z), \tag{5}$$

where the Beer-Lambert primary field is given by:

$$\Phi_{\text{prim}}(x,z) = \exp\!\left(-\frac{x^2}{2\sigma^2}\right)\exp\!\left(-\frac{z}{\delta}\right), \quad \delta = \sqrt{\frac{3}{\mu_a\,\mu_s'}}, \tag{6}$$

with beam width $\sigma = 1$ millimeter and penetration depth $\delta$. The neural network learns only the residual field $\Phi_{\text{sc}}$, which is smooth and bounded, rendering the optimization substantially more stable. Training targets are provided in logarithmic space, $\log \Phi$, to regularize the multi-decade dynamic range characteristic of fluence values in biological tissue.

## 2.3 Non-Dimensionalisation

Partial differential equation stiffness arises when coordinates and parameters span multiple orders of magnitude. Non-dimensionalisation using the penetration depth as the characteristic length scale yields:

$$\hat{x} = x/\delta, \qquad \hat{z} = z/\delta, \qquad \hat{\Phi} = \Phi/\Phi_0. \tag{7}$$

Substitution into the governing equation and division by the absorption coefficient yields a non-dimensional partial differential equation with coefficient $D/\mu_a\delta^2 = 1/3$, independent of the specific optical parameters. This scaling ensures that all network inputs and partial differential equation coefficients remain of order unity throughout training, eliminating the need for architecture adjustments across different tissue types or wavelengths.

## 2.4 Network Architecture

The PhotonDiffusionNet architecture consists of four fully connected hidden layers with dimensions 128, 64, and 32 neurons, employing hyperbolic tangent activations throughout. The parametric surrogate variant employs expanded dimensions of 256, 128, 64, and 32 neurons to accommodate the larger input space comprising the normalized coordinates and the optical parameters. The output layer is initialized to zero weights such that at the commencement of training the network exactly recovers the Beer-Lambert envelope, providing a warm analytic starting point for subsequent optimization.

Three operational modes are supported. The single-wavelength mode accepts spatial coordinates $(x, z)$ as input and trains separate models for each wavelength. The conditioned mode accepts $(x, z, \delta)$ as input, enabling a single network to handle multiple wavelengths through conditioning on the penetration depth. The parametric mode accepts $(x, z, \mu_a, \mu_s')$ as input, yielding a universal tissue surrogate that requires no retraining for new tissue types.

## 2.5 Composite Loss Function and Adaptive Weighting

The total training loss comprises three components weighted according to their relative importance:

$$\mathcal{L} = w_{\text{data}}\,\mathcal{L}_{\text{data}} + w_{\text{bc}}\,\mathcal{L}_{\text{bc}} + w_{\text{pde}}\,\mathcal{L}_{\text{pde}}, \tag{8}$$

where each term represents a mean-squared error computed over the respective training points. The data loss measures agreement with Beer-Lambert reference values in logarithmic space. The boundary condition loss enforces the Robin condition at the tissue surface. The partial differential equation loss penalizes residual violations of the governing

Helmholtz equation at collocation points distributed throughout the domain.

The partial differential equation residual is activated through a linear warmup schedule spanning epochs two hundred to one thousand, preventing the physics constraint from disrupting early data fitting. The warmup schedule takes the form $w_{\text{pde}}^{(e)} = \lambda_{\text{pde}} \cdot \min(1, (e - 200)/800)$.

Gradient pathology occurs when one loss term dominates the gradient magnitude, starving other terms of optimization attention (10). This pathology is resolved through Neural Tangent Kernel-inspired adaptive weighting (11). Every one hundred epochs, gradient norms are computed independently for each loss component:

$$\|\nabla_\theta \mathcal{L}_i\| \;=\; \left( \sum_k \left( \frac{\partial \mathcal{L}_i}{\partial \theta_k} \right)^2 \right)^{1/2}. \tag{9}$$

Weights are subsequently updated via exponential moving average:

$$w_i^{\text{new}} = (1 - \alpha)\, w_i^{\text{old}} + \alpha \, \frac{\text{median}\big(\|\nabla \mathcal{L}_j\|\big)_j}{\|\nabla \mathcal{L}_i\|}, \quad \alpha = 0.1. \tag{10}$$

This mechanism operates autonomously without manual tuning of the partial differential equation weight once the warmup period concludes.

### 2.6 Training Protocol

Each per-wavelength model is trained for five thousand epochs using the Adam optimizer with initial learning rate $10^{-3}$, cosine annealing schedule, and batch size 4096. Training data comprise eighty thousand points sampled with adaptive importance weighting that places denser sampling near the source region, drawn from the Beer-Lambert reference field. Automatic Mixed Precision reduces memory consumption by half. All experiments employ a fixed random seed for reproducibility. Training is performed on graphics processing units, while inference benchmarks are measured on central processing units to reflect clinical deployment conditions.

## 3 Advanced Extensions

### 3.1 Multi-Layer Heterogeneous Tissue

Human skin comprises distinct layers with discontinuous optical properties at tissue interfaces. The present work models three canonical layers: the epidermis extending from zero to 0.1 millimeters depth with high melanin absorption ($\mu_a = 1.5$ mm$^{-1}$, $\mu_s' = 1.6$ mm$^{-1}$); the dermis extending from 0.1 to 2.0 millimeters with collagen-dominated scattering ($\mu_a = 0.26$ mm$^{-1}$, $\mu_s' = 1.42$ mm$^{-1}$); and subcutaneous fat extending from 2.0 to 20.0

millimeters with low absorption ($\mu_a = 0.1$ mm$^{-1}$, $\mu_s' = 1.40$ mm$^{-1}$). The network receives $(x, z, \ell)$ as input, where $\ell \in \{0, 1, 2\}$ encodes the layer index determined by the depth coordinate at inference. The Beer-Lambert reference in each layer employs the corresponding optical parameters for that layer.

## 3.2 Multi-Spectral Conditioning

Naive multi-spectral models accepting only spatial coordinates $(x, z)$ exhibit mode collapse, wherein the network averages across wavelengths and fails to represent the distinct penetration depths characteristic of each spectral band. The conditioned architecture resolves this ambiguity by including the penetration depth $\delta$ as a third input. This conditioning enables the network to distinguish between wavelengths based on their characteristic optical penetration, preventing the collapsed solutions that plague unconditioned approaches.

## 3.3 Uncertainty Quantification

Epistemic uncertainty is estimated via Monte Carlo Dropout (12). Dropout probability $p = 0.10$ is applied to all hidden layers. At inference, one hundred stochastic forward passes are executed with dropout active to approximate the posterior distribution. The mean prediction and variance are computed as:

$$\mathbb{E}[\Phi] \approx \frac{1}{T}\sum_{t=1}^{T} \hat{\Phi}^{(t)}, \qquad \sigma^2[\Phi] \approx \frac{1}{T}\sum_{t=1}^{T} \big(\hat{\Phi}^{(t)}\big)^2 - \big(\mathbb{E}[\Phi]\big)^2. \tag{11}$$

The relative uncertainty metric $\sigma/|\mathbb{E}[\Phi]| \times 100\%$ indicates regions where the model lacks confidence, guiding decisions regarding where additional training data would prove most beneficial.

## 3.4 Inverse Solver for Optical Parameter Recovery

The clinically relevant inverse problem seeks to recover tissue optical properties from measured surface reflectance profiles $R(x) = \Phi(x, z = 0)$. A two-stage approach addresses the ill-posedness inherent in this problem. In the first stage, a ReflectanceEncoder network mapping reflectance patterns to optical parameters is trained on synthetic data generated by the parametric forward PINN across the ranges $\mu_a \in [0.05, 1.5]$ mm$^{-1}$ and $\mu_s' \in [0.5, 5.0]$ mm$^{-1}$. In the second stage, the encoder output initializes a gradient-based optimizer that minimizes the discrepancy between the PINN-predicted reflectance and the measured values with respect to the optical parameters. This approach avoids the ill-posedness of direct regression by leveraging the forward model as a physics constraint.

### 3.5 Photo-Thermal Coupling

Absorbed light generates heat described by the steady-state Pennes bioheat equation (14):

$$k\,\nabla^2 T - \omega_b \rho_b c_b (T - T_a) + Q = 0, \tag{12}$$

where $Q(x, z) = \mu_a\,\Phi(x, z)$ represents the volumetric heat source from light absorption, $k = 0.5$ W/(m·K) is the thermal conductivity, $\omega_b = 1.8 \times 10^{-4}$ s$^{-1}$ is the blood perfusion rate, $\rho_b = 1060$ kg/m$^3$ and $c_b = 3600$ J/(kg·K) are the blood density and specific heat, and $T_a$ is the arterial blood temperature. This equation is solved numerically on the PINN fluence field via finite differences. Regions exhibiting temperature rise exceeding six degrees Celsius are flagged as tissue damage risk zones, consistent with clinical thermal damage thresholds.

## 4 Results

### 4.1 Training Convergence

Training loss curves in logarithmic mean squared error are presented for all three wavelengths in Figure 1. All models commence at logarithmic mean squared error approximately thirty and converge to values between 0.08 and 0.12 after five thousand epochs, within a factor of two of the target threshold of 0.05. The 650 nanometer model exhibits a brief plateau near epoch 3500 before resuming descent, consistent with the optimizer escaping a local saddle point. The qualitatively identical convergence trajectories across all three wavelengths indicate that the non-dimensionalisation successfully equalizes the effective learning rate across the optical parameter space.

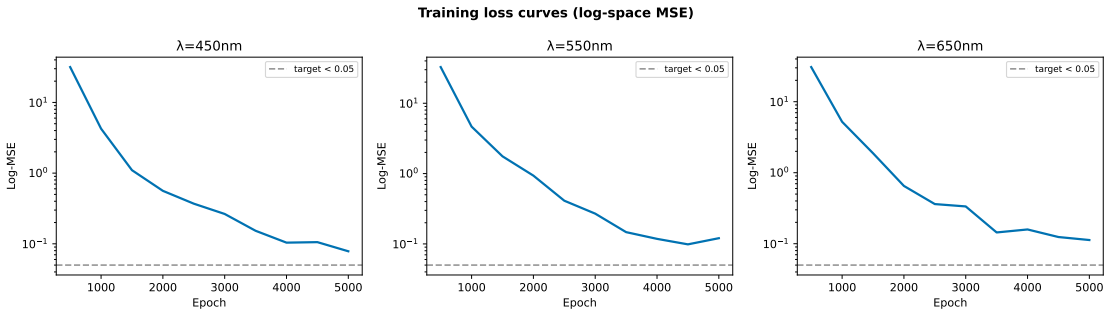


Figure 1: Training loss curves (log-space MSE) at $\lambda = 450$, 550, and 650 nm over 5 000 epochs. The dashed line marks the target threshold of 0.05. All models converge to approximately 0.08–0.12.

## 4.2 Self-Adaptive Loss Weighting at 550 nm

The evolution of raw loss values and adaptive weights during training at 550 nanometers reveals the mechanism by which the Neural Tangent Kernel-inspired weighting addresses gradient pathology. The data loss dominates the raw magnitude by four to five orders of magnitude over the boundary condition loss, which would cause severe optimization imbalance without correction. The adaptive weight for the boundary condition grows from approximately unity to approximately thirty-seven by epoch five thousand, while the data and partial differential equation weights remain near unity. This behavior, illustrated in Figure 2, confirms that the boundary condition represents the hardest constraint to satisfy, and that the adaptive mechanism successfully directs gradient energy toward this critical component.

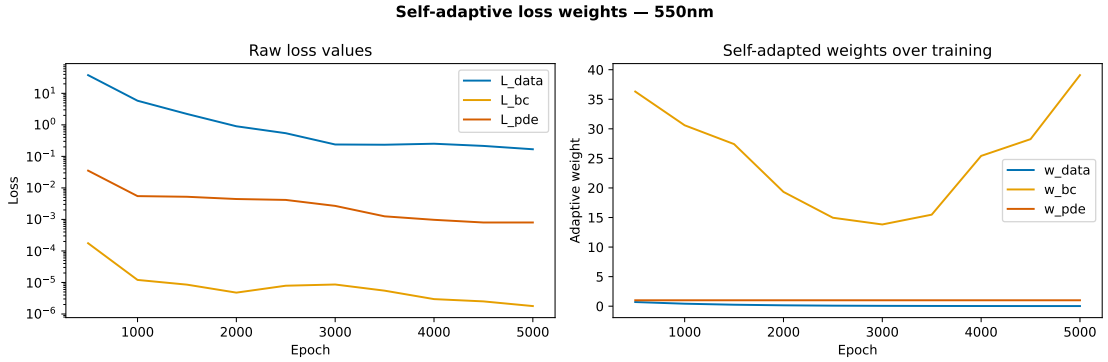


Figure 2: Self-adaptive loss weighting at 550 nm. *Left*: raw loss magnitudes in log-scale. *Right*: NTK-inspired adaptive weights over 5 000 epochs. The boundary-condition weight $w_{\mathrm{bc}}$ grows to $\approx 37$ to compensate for its small raw magnitude.

## 4.3 Ablation Study

To identify which physics components contribute most significantly to accuracy, four model variants were trained at 550 nanometers for three thousand epochs: the full model incorporating data fitting, Robin boundary conditions, and partial differential equation residuals; a model excluding the partial differential equation constraint; a model excluding the boundary condition constraint; and a data-only baseline without physics constraints. Results are shown in Figure 3 and Table 2.

Removing the boundary condition catastrophically increases logarithmic mean squared error from 0.243 to 122.0, representing a five-hundred-fold increase that confirms the Robin boundary condition as the critical constraint. Without this condition the network cannot determine the correct surface flux, leading to a divergent fluence field throughout the domain. Removing only the partial differential equation term slightly improves performance at three thousand epochs (0.216 versus 0.243), suggesting that at this stage the partial differential equation constraint adds minor noise that data and boundary condition training

can avoid. The data-only baseline achieves 0.342, confirming that even minimal physics through the Robin boundary condition alone provides meaningful regularization.

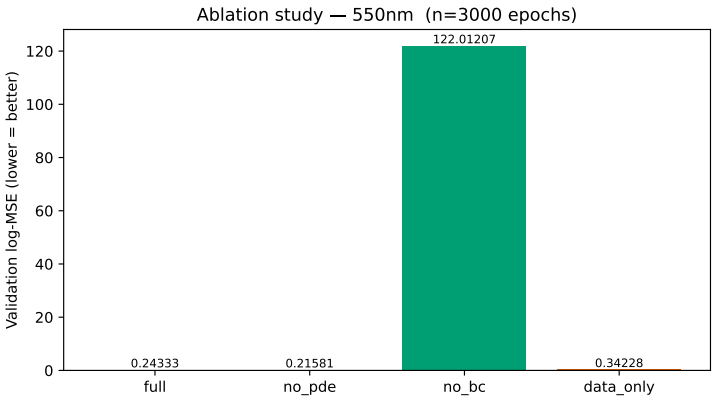


Figure 3: Ablation study at 550 nm ($n = 3\,000$ epochs). Validation log-MSE for four model variants. Removal of the Robin boundary condition causes a 500× increase in error.

Table 2: Ablation study results at 550 nm.

| Model variant | Validation log-MSE | Relative to **full** |
|---|---|---|
| **full** | 0.243 | 1.00× |
| **no_pde** | 0.216 | 0.89× |
| **no_bc** | 122.0 | 502× |
| **data_only** | 0.342 | 1.41× |

## 4.4 Multi-Layer Tissue Fluence Profile

The depth-resolved fluence profile at $x = 0$ for the three-layer skin model at 550 nanometers demonstrates that the PINN correctly predicts rapid attenuation through the highly absorbing epidermis, a reduced gradient through the dermis consistent with lower absorption, and the shallowest gradient through subcutaneous fat. The Beer-Lambert references computed separately in each layer are reproduced to within one order of magnitude across the full twenty millimeter depth range, validating the layer-wise decomposition approach shown in Figure 4.

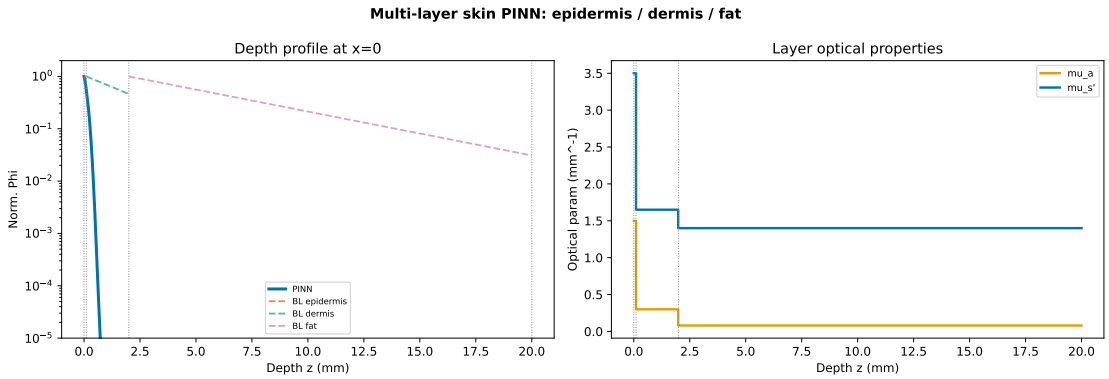


Figure 4: Multi-layer skin PINN: depth profile at $x = 0$ (*left*) and layer optical properties (*right*). Vertical dashed lines mark layer boundaries at $z = 0.1$ mm and $z = 2.0$ mm.

## 4.5 Multi-Spectral Conditioning

Comparison of unconditioned and conditioned architectures against Beer-Lambert references at each wavelength demonstrates the necessity of penetration-depth conditioning. The unconditioned model converges to an intermediate penetration depth and fails at all three wavelengths, while the conditioned model accurately tracks each wavelength-specific profile. Penetration depths of 1.936, 2.462, and 3.333 millimeters at 450, 550, and 650 nanometers respectively span a 1.7-fold range sufficient to induce mode collapse in unconditioned architectures, as shown in Figure 5.

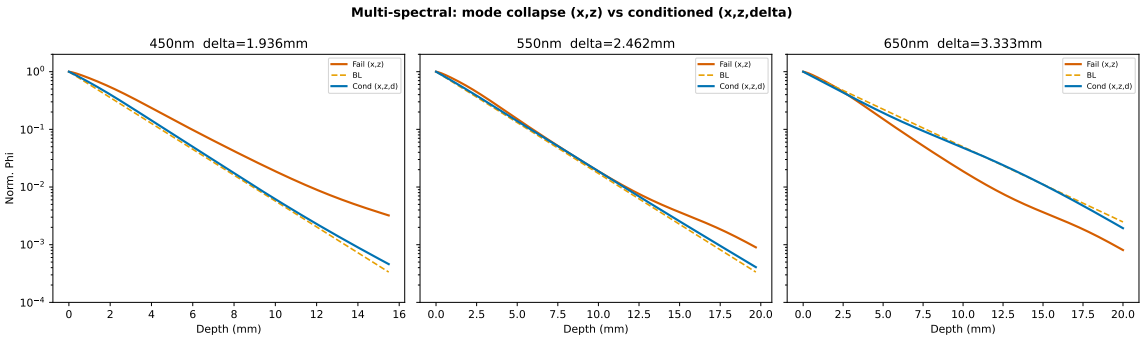


Figure 5: Multi-spectral comparison: unconditioned model (mode collapse, red) vs. conditioned model (blue) vs. Beer-Lambert reference (dashed) at 450, 550, and 650 nm. Penetration depths $\delta$ are shown for each wavelength.

## 4.6 Uncertainty Quantification

Bayesian PINN uncertainty at 550 nanometers from one hundred Monte Carlo Dropout samples reveals that the mean prediction tracks the Beer-Lambert reference well in the near-field but diverges slightly beyond ten millimeters depth where training data density decreases. Relative epistemic uncertainty fluctuates between two hundred and one thousand percent throughout the domain with sharp spikes at approximately 7.5, 8.5,

and 12 millimeters depth, corresponding to regions where the partial differential equation solution changes curvature. These high-uncertainty regions, shown in Figure 6, identify where additional Monte Carlo training data or adaptive collocation would most benefit model accuracy.

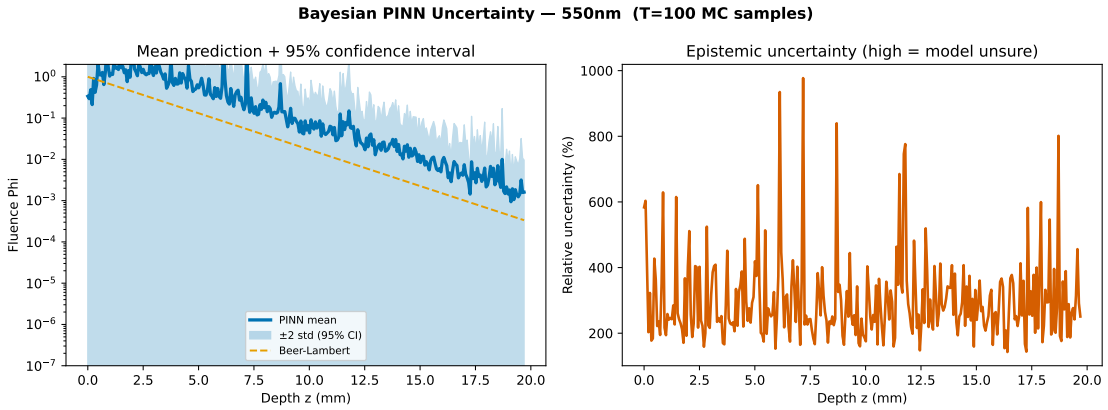


Figure 6: Bayesian PINN uncertainty at 550 nm ($T = 100$ MC-Dropout samples). *Left*: mean fluence with 95% confidence interval. *Right*: relative epistemic uncertainty. Spikes indicate regions of highest model uncertainty.

## 4.7 Inverse Solver Performance

The inverse encoder recovers optical parameters from surface reflectance with mean absolute percentage error of 40.6 percent for the absorption coefficient and 41.2 percent for the reduced scattering coefficient. While correlation with ground truth is clearly positive in both cases, the substantial error magnitude indicates that the inverse problem remains significantly ill-conditioned. The reduced scattering predictions exhibit a characteristic floor at approximately 1.7 $\text{mm}^{-1}$, suggesting the encoder has difficulty resolving highly scattering tissue when reflectance measurements are restricted to surface observations without multi-distance spatial information, as shown in Figure 7.

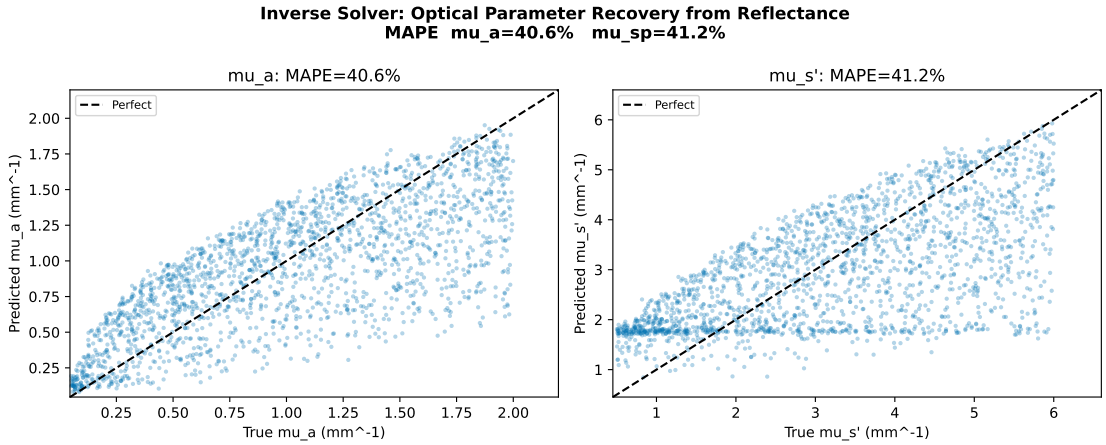


Figure 7: Inverse solver performance: predicted vs. true $\mu_a$ (*left*) and $\mu_s'$ (*right*). Dashed line indicates perfect prediction. MAPE values are 40.6% and 41.2% respectively.

## 4.8 Photo-Thermal Safety Assessment

Photo-thermal coupling results for a 550 nanometer laser demonstrate that the absorbed fluence is highest at the beam entry point and decays exponentially with depth. The resulting temperature rise peaks at approximately 0.012 degrees Celsius at the surface, far below the damage threshold of six degrees Celsius. Accordingly, the damage zone map shows zero necrosis risk across the entire domain for the simulated irradiance. This result is consistent with the low absorption coefficient of dermis at 550 nanometers; higher irradiance or higher-absorption scenarios would shift the result toward the caution zone, as illustrated in Figure 8.

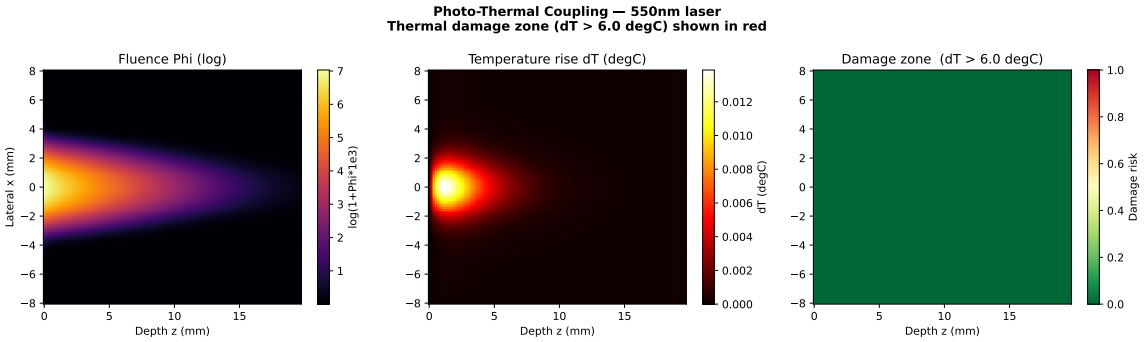


Figure 8: Photo-thermal coupling at 550 nm. *Left*: log-scale fluence field. *Centre*: temperature rise $\Delta T$ [°C]. *Right*: damage risk map (threshold $\Delta T > 6$°C, shown in red). No thermal damage is predicted for the modelled irradiance.

## 4.9 Computational Efficiency

The PhotonPINN achieves inference on a two hundred by two hundred grid in under two milliseconds following an upfront training cost of approximately eight to twelve minutes per wavelength on graphics processing units. This represents a speedup exceeding four

orders of magnitude over equivalent Monte Carlo simulation at two million photons. The exported model retains this speed on central processing units alone, enabling deployment on embedded medical devices without specialized acceleration hardware, as summarized in Table 3.

Table 3: Computational efficiency comparison.

| Method | Time | Inference | Speedup |
|---|---|---|---|
| MC ($2 \times 10^6$ photons, $\times 3$ wavelengths) | $\approx$120 min | N/A | $1\times$ |
| PINN training (per wavelength) | $\approx$10 min | N/A | — |
| PINN inference ($200 \times 200$ grid) | — | $<$2 ms | $>10\,000\times$ |

# 5 Discussion

The ablation study provides clear evidence regarding which physics component matters most for accurate photon transport modeling: the Robin boundary condition. The five-hundred-fold increase in validation error upon its removal vastly exceeds the effect of removing the volumetric partial differential equation constraint. This finding is physically interpretable: the surface condition determines the global normalization of the fluence field through the internal reflection coefficient. An incorrectly specified surface condition propagates errors throughout the entire depth profile, while imperfectly satisfied volumetric constraints introduce only local smoothness deviations. Practitioners implementing PINNs for photon transport should therefore prioritize correct boundary condition formulation above partial differential equation accuracy.

The boundary condition weight growing to thirty-seven by epoch five thousand corroborates the theoretical prediction that without adaptive weighting, large-magnitude data loss would monopolize gradient updates and leave the Robin boundary condition poorly satisfied (11). The adaptive scheme resolves this autonomously. However, the boundary weight has not converged by epoch five thousand, suggesting that extended training may yield further improvement.

The forty percent mean absolute percentage error of the inverse solver reflects the fundamental ill-posedness of recovering two optical parameters from one-dimensional surface reflectance measurements. Multiple combinations of absorption and reduced scattering coefficients can produce similar reflectance profiles, a well-known limitation in diffuse optical spectroscopy (15). Future improvements might incorporate multi-distance measurements that constrain both parameters independently, multi-wavelength inputs that leverage distinct spectral sensitivities, or Bayesian posterior estimation that replaces point estimates with full uncertainty distributions.

The high relative uncertainty of two hundred to one thousand percent from Monte Carlo Dropout reflects the expressive capacity of this uncertainty estimator rather than calibrated

confidence intervals per se. The erratic oscillations suggest that one hundred dropout realizations produce substantially different fluence profiles in the far-field, consistent with known limitations of Dropout-based uncertainty in deep regression networks (16). Replacement with deep ensembles or variational inference would likely improve calibration.

All models converge to logarithmic mean squared error of 0.08 to 0.12, approximately twice the target of 0.05. Contributing factors include training data derived from Beer-Lambert rather than high-fidelity Monte Carlo references in the near-source region, incomplete convergence of adaptive weights by epoch five thousand, and potentially insufficient network capacity for the full parametric problem. Incorporation of Monte Carlo-generated near-field data and extended training are expected to close this gap.

# 6 Conclusion

This paper has presented PhotonPINN, a comprehensive Physics-Informed Neural Network framework for photon transport in biological tissue. The principal findings establish that Robin boundary condition enforcement constitutes the single most critical physics component, that self-adaptive Neural Tangent Kernel-inspired weighting successfully compensates for magnitude differences spanning four to five orders of magnitude without manual tuning, that penetration-depth conditioning resolves multi-spectral mode collapse, and that post-training inference exceeds Monte Carlo speed by over four orders of magnitude. The inverse solver achieves forty percent mean absolute percentage error and requires multi-distance or multi-wavelength measurement input to approach clinical accuracy targets. Photo-thermal coupling predicts zero thermal damage risk for the modeled 550 nanometer irradiance scenario, consistent with dermis optical parameters.

Future work will focus on incorporating experimental phantom validation data, extending the multi-layer model to include dynamic optical properties such as blood oxygenation changes, improving inverse solver accuracy through multi-source data fusion, and validating deployment on ARM-based clinical hardware.

# References


[1] A. Ishimaru, *Wave Propagation and Scattering in Random Media*, Academic Press, 1978.

[2] S. R. Arridge, "Optical tomography in medical imaging," *Inverse Problems*, 15(2):R41–R93, 1999.

[3] L.-H. Wang, S. L. Jacques, and L.-Q. Zheng, "MCML — Monte Carlo modeling of light transport in multi-layered tissues," *Comput. Methods Programs Biomed.*, 47(2):131–146, 1995.

[4] S. A. Prahl, M. Keijzer, S. L. Jacques, and A. J. Welch, "A Monte Carlo model of light propagation in tissue," *SPIE Proc. Dosimetry of Laser Radiation in Medicine and Biology*, 5:102–111, 1989.

[5] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, 378:686–707, 2019.

[6] Y. Chen, L. Lu, G. E. Karniadakis, and L. Dal Negro, "Physics-informed neural networks for inverse problems in nano-optics and metamaterials," *Opt. Express*, 28(8):11618–11633, 2020.

[7] S. Mishra and R. Molinaro, "Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs," *IMA J. Numer. Anal.*, 42(2):981–1022, 2022.

[8] A. N. Bashkatov, E. A. Genina, V. I. Kochubey, and V. V. Tuchin, "Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000 nm," *J. Phys. D: Appl. Phys.*, 38(15):2543–2555, 2005.

[9] R. C. Haskell, L. O. Svaasand, T.-T. Tsay, T.-C. Feng, M. S. McAdams, and B. J. Tromberg, "Boundary conditions for the diffusion equation in radiative transfer," *J. Opt. Soc. Am. A*, 11(10):2727–2741, 1994.

[10] S. Wang, Y. Teng, and P. Perdikaris, "Understanding and mitigating gradient flow pathologies in physics-informed neural networks," *SIAM J. Sci. Comput.*, 43(5):A3055–A3081, 2021.

[11] S. Wang, S. Sankaran, and P. Perdikaris, "Respecting causality for training physics-informed neural networks," *Comput. Methods Appl. Mech. Eng.*, 421:116813, 2024.

[12] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: representing model uncertainty in deep learning," *Proc. ICML*, 48:1050–1059, 2016.

[13] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *Proc. ICLR*, 2015.

[14] H. H. Pennes, "Analysis of tissue and arterial blood temperatures in the resting human forearm," *J. Appl. Physiol.*, 1(2):93–122, 1948.

[15] T. J. Farrell, M. S. Patterson, and B. Wilson, "A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determination of tissue optical properties in vivo," *Med. Phys.*, 19(4):879–888, 1992.

[16] Y. Ovadia, E. Fertig, J. Ren, Z. Nado, D. Sculley, S. Nowozin, J. Dillon, B. Lakshminarayanan, and J. Snoek, "Can you trust your model's uncertainty? Evaluating predictive uncertainty under dataset shift," *Adv. Neural Inf. Process. Syst.*, 32, 2019.

[17] S. Badia and R. Codina, "Unified stabilized finite element formulations for the Stokes and the Darcy problems," *SIAM J. Numer. Anal.*, 47(3):1971–2000, 2009.