Operator Learning for Regime-Switching Black–Cox First-Passage Coupled PDE Systems

Anonymous Author(s)
Affiliation
Address
email

Abstract

Regime-switching first-passage-time structural credit models lead to tightly coupled PDEs with absorbing boundaries. Re-solving these PDEs whenever parameters change becomes a key bottleneck for calibration and stress testing, due to high computational cost. To address this, we propose Reflection DeepONet, a physics-informed neural operator whose trunk embeds reflection-principle survival bases to enforce the absorbing boundary by construction, while its branch predicts spectral regime-mixing coefficients from switching intensities. Training is self-supervised by minimizing PDE residuals, using a boundary-robust Huber loss and hybrid sampling that emphasizes the near-maturity boundary layer with gradual regime coupling. Experiments demonstrate robust, consistent regime sensitivity of term structures across single-name, counterparty-risk, and basket settings, enabling fast amortized scenario analysis across a wide range of parameter values.

1 Introduction

Credit risk modeling fundamentally seeks to characterize default timing and survival probabilities within a mathematically consistent framework[9]. This characterization forms the bedrock for asset pricing, risk management, and strategic credit decisions. While reduced-form models offer mathematical tractability by treating default as an exogenous intensity-driven event [4], structural models provide a more transparent link to economic reality. By endogenizing default as the moment a firm's asset value hits an absorbing boundary—an approach pioneered by Black and Cox [1] and extended to multi-firm correlations by Zhou [15]—first-passage-time (FPT) models offer an intuitive and interpretable mechanism for credit analysis.

Despite their theoretical elegance, extending FPT models to realistic financial environments—such as regime-switching dynamics and multi-firm dependence structures—leads to a rapid escalation in analytical and computational complexity. These extensions give rise to systems of tightly coupled partial differential equations (PDEs) with absorbing boundaries. As demonstrated by Kim et al. [5], solving such systems via traditional finite difference methods is not only computationally expensive but also prone to numerical instabilities near steep boundary gradients. In practical applications such as real-time calibration and large-scale stress testing, where these PDEs must be solved repeatedly over large parameter spaces, classical numerical schemes often become a prohibitive bottleneck.

The emergence of scientific machine learning has introduced a transformative paradigm for addressing these challenges. Physics-Informed Neural Networks (PINNs) [10] embed governing PDEs, boundary conditions, and initial conditions directly into the training objective via automatic differentiation, providing a flexible, mesh-free alternative to classical numerical solvers. PINNs have shown strong performance in high-dimensional financial problems, including complex option pricing and counterparty credit risk modeling [3, 12], and have recently been applied to coupled systems with absorbing barriers, such as barrier option pricing under time-varying interest rates and volatility [2]. Despite

Submitted to 1st 2026AI Co-Scientist Challange Korea. Do not distribute.

these successes, PINNs are fundamentally instance-wise solvers, in the sense that they approximate solutions for a fixed set of model parameters; consequently, changes in volatilities, regime-switching intensities, or default boundaries typically require retraining or costly re-optimization. Moreover, the optimization landscape induced by PDE residual losses becomes particularly challenging in multi-scale systems or problems with sharp absorbing boundaries, leading to slow convergence and sensitivity to hyperparameters [2]. Similar limitations are shared by related neural PDE solvers such as the Deep Galerkin Method (DGM) [11].

These limitations have motivated growing interest in operator learning, which aims to learn the mapping from parameter spaces to solution manifolds for entire families of PDEs, rather than solving individual instances. Representative frameworks include Deep Operator Networks (DeepONet) [8] and Fourier Neural Operators (FNO) [7], which have demonstrated that neural models can efficiently capture parametric solution operators across a wide range of PDE systems. When combined with physical constraints—through physics-informed operator learning approaches [13]—these models achieve improved generalization and data efficiency, alleviating the need for repeated retraining across parameter configurations [14]. This paradigm is particularly well suited to structural credit modeling, where amortized inference enables rapid generation of survival surfaces across regimes and market conditions, replacing costly and repetitive numerical simulations.

Motivated by these considerations, we propose a physics-informed operator learning framework for regime-switching first-passage structural credit models. Our approach embeds analytical survival structure derived from the reflection principle into a neural operator via closed-form marginal survival bases, providing a structure-aware representation of joint survival. As the model learns the full mapping from parameters to survival surfaces, it enables rapid evaluation across parameter regimes and offers an efficient alternative to repeated numerical PDE solvers while preserving financial interpretability.

2 Problem Setting: Regime-Switching First-Passage Survival PDEs

2.1 Regime-switching diffusion and default boundaries

We consider a two-regime market condition process $Y_t \in \{e_B, e_b\}$ (bull B vs. bear b), modeled as a continuous-time Markov chain with constant switching intensities $\lambda_B > 0$ (from B to b) and $\lambda_b > 0$ (from b to B) [5]. The state space is represented by the canonical orthogonal basis vectors $e_B = (1,0)^\top$ and $e_b = (0,1)^\top$ in $\mathbb{R}^2$. Under the risk-neutral measure, the asset value $X_i(t)$ of firm $i = 1, 2$ follows a regime-dependent geometric Brownian motion:

$$dX_i(t) = r^{Y_t} X_i(t)\, dt + \sigma_i^{Y_t} X_i(t)\, dW_i(t), \qquad dW_1(t) dW_2(t) = \rho\, dt, \tag{1}$$

where r^{Y_t} and $\sigma_i^{Y_t}$ denote the instantaneous risk-free interest rate and the asset return volatility, respectively, both of which are stochastic processes modulated by the prevailing market regime Y_t. The stochastic drivers $W_i(t)$ are standard Brownian motions correlated via a constant coefficient $\rho \in [-1, 1]$, capturing the systematic comovement between the two firms. Default is defined endogenously as the first passage time $\tau_i(t)$ at which the asset value $X_i(s)$ breaches a time-dependent solvency barrier $K_i(s)$:

$$\tau_i(t) := \inf\{s \geq t : X_i(s) \leq K_i(s)\}, \tag{2}$$

with the barrier typically specified as an exponential function of the form $K_i(s) = K_i e^{\eta_i s}$, where $K_i > 0$ is a firm-specific constant (the initial solvency barrier level for firm i) and $\eta_i \in \mathbb{R}$ controls the deterministic growth rate of liabilities [5].

2.2 Individual survival PDE system

We define the regime-wise survival probability v for firm i (hereafter we focus on firm 1) as

$$v^j(t, x_1) := \mathbb{P}(\tau_1(t) > T \mid X_1(t) = x_1,\ Y_t = j), \qquad j \in \{B, b\}.$$

Here, B and b denote the bull and bear regimes, respectively. The underlying regime process Y_t formally takes values in the canonical basis $\{e_B, e_b\}$; for notational simplicity, we henceforth write $Y_t = j$ with $j \in \{B, b\}$ instead of $Y_t = e_j$.

Then the vector (v^B, v^b) solves a coupled system of absorbing partial differential equations, subject to the absorbing boundary condition $v^j(t,x) = 0$ for $x \le K_i(t)$ and the terminal condition $v^j(T,x) = \mathbb{I}(x > K_i(T))$, so survival at maturity occurs if and only if the firm value remains above the barrier.

The PDE system is derived via the Feynman–Kac representation for regime-switching diffusions [5, Eq. 2]:

$$\begin{aligned} \frac{\partial v^B}{\partial t} + r^B x_1 \frac{\partial v^B}{\partial x_1} + \frac{1}{2}(\sigma_1^B x_1)^2 \frac{\partial^2 v^B}{\partial x_1^2} + \lambda_B (v^b - v^B) = 0, \\ \frac{\partial v^b}{\partial t} + r^b x_1 \frac{\partial v^b}{\partial x_1} + \frac{1}{2}(\sigma_1^b x_1)^2 \frac{\partial^2 v^b}{\partial x_1^2} + \lambda_b (v^B - v^b) = 0. \end{aligned} \tag{3}$$

Here, r^j and σ_1^j denote the regime-dependent interest rate and volatility, respectively, for $j \in \{B, b\}$.

2.3 Joint survival PDE system

Define the joint regime-wise survival probability for two firms:

$$u^j(t, x_1, x_2) := \mathbb{P}\left(\tau_1(t) > T,\ \tau_2(t) > T \mid X_1(t) = x_1,\ X_2(t) = x_2,\ Y_t = j\right).$$

Then $\left(u^B, u^b\right)$ solves a coupled absorbing PDE system with a cross-derivative term induced by ρ [5, Eq. 1]:

$$\begin{aligned} &\frac{\partial u^B}{\partial t} + r^B \sum_{i=1}^{2} x_i \frac{\partial u^B}{\partial x_i} + \sum_{i=1}^{2} \frac{1}{2}(\sigma_i^B x_i)^2 \frac{\partial^2 u^B}{\partial x_i^2} \\ &\qquad + \rho\, \sigma_1^B \sigma_2^B x_1 x_2 \frac{\partial^2 u^B}{\partial x_1 \partial x_2} + \lambda_B (u^b - u^B) = 0, \\ &\frac{\partial u^b}{\partial t} + r^b \sum_{i=1}^{2} x_i \frac{\partial u^b}{\partial x_i} + \sum_{i=1}^{2} \frac{1}{2}(\sigma_i^b x_i)^2 \frac{\partial^2 u^b}{\partial x_i^2} \\ &\qquad + \rho\, \sigma_1^b \sigma_2^b x_1 x_2 \frac{\partial^2 u^b}{\partial x_1 \partial x_2} + \lambda_b (u^B - u^b) = 0, \end{aligned} \tag{4}$$

2.4 Parameter family and the survival operator

Collect coefficients and regime parameters into a compact parameter vector $\Theta \in \Lambda \subset \mathbb{R}^{d_\Theta}$, e.g.,

$$\Theta := \left(r^B, r^b, \sigma_1^B, \sigma_1^b, \sigma_2^B, \sigma_2^b, \rho, \lambda_B, \lambda_b, \eta_1, \eta_2\right).$$

We define the regime-wise survival operator $\Gamma_{\mathrm{FP}} : \Lambda \to \mathcal{U}$ as:

$$\Gamma_{\mathrm{FP}}(\Theta)(t,\xi) = U^\Theta(t,\xi) := \left(U^{B,\Theta}(t,\xi),\ U^{b,\Theta}(t,\xi)\right)^\top, \tag{5}$$

where $\Theta \in \Lambda$ denotes the model parameter vector and U^Θ denotes the corresponding regime-wise survival function. Here, $\xi = x$ for individual survival and $\xi = (x_1, x_2)$ for joint survival.

3 Method: Reflection DeepONet

We propose the *Reflection DeepONet*, a physics-informed neural operator architecture designed to approximate the solution operator Γ_{FP} for the regime-switching survival PDEs (Eqs. 3–4). Unlike standard finite difference schemes which solve a single instance, our method learns the mapping from regime intensities $\boldsymbol{\lambda} = [\lambda_B, \lambda_b]^\top$ to the full spatio-temporal survival probability manifolds. We embed reflection survival bases in the Trunk and learn regime-mixing coefficients in the Branch, following asymptotics-informed enrichment for residual-based operator learning [6].

3.1 Coordinate system and notation

Time-to-maturity. We work in time-to-maturity coordinates $\tau := T - t$. We denote by $\bar{T} > 0$ the maximum time-to-maturity considered in training/analysis, so $\tau \in [0, \bar{T}]$.

Log-moneyness relative to the default barrier. For each firm i, the (time-dependent) barrier is $K_i(t) = K_i e^{\eta_i t}$ with firm-specific constant $K_i > 0$. We define the barrier-relative log-coordinate

$$z_i := \log\Big(\frac{x_i}{K_i(t)}\Big),$$

so that the default boundary $x_i \le K_i(t)$ corresponds to $z_i \le 0$ (absorbing boundary at $z_i = 0$).

Equivalently, since $t = T - \tau$, we may write $K_i(T-\tau) = K_i e^{\eta_i(T-\tau)}$ and view $z_i = z_i(\tau)$ when working in (τ, z)-coordinates.

Spatial variable. We write $\xi = z \in D_1 := [0, z_{\max}]$ for individual survival and $\xi = (z_1, z_2) \in D_2 \subset \mathbb{R}^2$ for joint survival (both truncated away from infinity). In our experiments we take $\Omega_z = [0, \ln 3]$, so $D_1 = \Omega_z$ and $D_2 = \Omega_z^2$.

Regime indexing. The regime process takes values in $\{e_B, e_b\}$, but we write $Y_t = j$ with $j \in \{B, b\}$ for simplicity, and denote the opposite regime by $\bar{j}$.

Unified survival field. We denote the regime-wise survival field by $U^\Theta(\tau,\xi) := (U^{B,\Theta}(\tau,\xi),\, U^{b,\Theta}(\tau,\xi))^\top$. For clarity, U^Θ corresponds to v^Θ in 1D and u^Θ in 2D in Section 2.

3.2 Operator Architecture and Spectral Expansion

We approximate the regime-wise survival field $U(\tau,\xi;\boldsymbol{\lambda}) = (U^B(\tau,\xi;\boldsymbol{\lambda}),\, U^b(\tau,\xi;\boldsymbol{\lambda}))^\top$ using a modified DeepONet architecture. Here $\xi = z$ for individual survival and $\xi = (z_1, z_2)$ for joint survival (Section 3.1). The model follows a Branch–Trunk synthesis: the Branch encodes the regime-switching intensities $\boldsymbol{\lambda}$ into spectral mixing coefficients, while the Trunk provides a physics-compliant reflection basis over (τ, ξ) that hard-enforces absorbing boundaries. Throughout, all parameters in Θ other than $\boldsymbol{\lambda}$ (e.g., $r^j, \sigma_i^j, \rho, K_i, \eta_i$) are treated as fixed; we learn the operator map $\boldsymbol{\lambda} \mapsto U(\cdot,\cdot;\boldsymbol{\lambda})$.

3.2.1 Branch Network: Parametric Encoder

The Branch network, denoted as $\mathcal{N}(\boldsymbol{\lambda};\theta_{br})$, approximates the mapping from the regime-switching intensity space $\boldsymbol{\lambda} \subset \Lambda$ to the spectral coefficient space. Its role is to determine the mixing weights of the basis functions conditional on the market regime.

We employ a Multi-Layer Perceptron (MLP) with $L = 3$ hidden layers of width $N_h = 256$. To ensure higher-order differentiability required for the PDE operators, we use the Sigmoid Linear Unit (SiLU) activation, denoted by $\mathrm{silu}(\cdot)$:

$$\mathrm{silu}(z) = z \cdot (1 + e^{-z})^{-1}.$$

The mapping from the input intensity vector $\boldsymbol{\lambda} = [\lambda_B, \lambda_b]^\top$ to the coefficient vector $\mathbf{c}(\boldsymbol{\lambda})$ is given by:

$$\begin{aligned} \mathbf{h}_0 &= \boldsymbol{\lambda}, \\ \mathbf{h}_l &= \mathrm{silu}(\mathbf{W}_l \mathbf{h}_{l-1} + \mathbf{b}_l), \quad l = 1, \dots, L. \\ \mathbf{c}(\boldsymbol{\lambda}) &= \mathbf{W}_{\text{out}} \mathbf{h}_L + \mathbf{b}_{\text{out}}. \end{aligned} \tag{6}$$

For individual (1D) survival, $\mathbf{c}(\boldsymbol{\lambda}) \in \mathbb{R}^{2M}$ and is reshaped into $\{c_{j,k}(\boldsymbol{\lambda})\}_{k=1}^M$ for each regime $j \in \{B, b\}$. For joint (2D) survival, $\mathbf{c}(\boldsymbol{\lambda}) \in \mathbb{R}^{2M^2}$ and is reshaped into $\{c_{j,k,\ell}(\boldsymbol{\lambda})\}_{k,\ell=1}^M$.

3.2.2 Trunk Network: Physics-Informed Reflection Basis

Standard neural trunks do not naturally satisfy the absorbing boundary $u(\tau, 0) = 0$. We therefore build a hard-constrained basis using the reflection principle.

Adaptive Volatility Scaling. The network learns a spectrum of M effective volatilities to capture multiscale diffusive behaviors. To distinguish these internal scales from the physical market volatility σ^{Y_t} and the activation function $\sigma(\cdot)$, we denote the k-th effective diffusivity parameter as $\hat{\sigma}_k$. Since the raw learnable parameters $\theta_{\sigma,k}$ reside in an unbounded domain ($\theta_{\sigma,k} \in \mathbb{R}$), direct usage could violate the physical constraint of non-negative volatility. To resolve this, we parametrize $\hat{\sigma}_k$ via a *shifted softplus rectification*:

$$\hat{\sigma}_k = \sigma_{\min} + \ln(1 + \exp(\theta_{\sigma,k}))\,,\; k = 1, \dots, M, \tag{7}$$

Here $\sigma_{\min} > 0$ ensures strict positivity and numerical stability.

This formulation ensures mathematical robustness in two ways. First, the softplus function provides a smooth, differentiable map from $\mathbb{R} \to \mathbb{R}^+$, avoiding the non-differentiable points associated with standard ReLU clipping while ensuring strict positivity. Second, the shift term $\sigma_{\min} = 0.5$ acts as a numerical stabilizer, strictly bounding $\hat{\sigma}_k \geq \sigma_{\min} > 0$ to prevent gradient explosion caused by division-by-zero in the distance metrics. Here, $\theta_{\sigma,k}$ is a learnable scalar embedded in the ansatz architecture, distinct from the branch network inputs $\boldsymbol{\lambda}$. This decoupling allows the model to adaptively scale the coordinate distance metrics for each basis function $k \in \{1, \dots, M\}$ independently of the exogenous regime parameters.

Reflection Basis Formulation. To strictly enforce the structural constraints of the default event, we construct the k-th basis function $\phi_k(\tau, z; \hat{\sigma}_k)$ using the method of images. This formulation satisfies the initial condition $\phi_k(0, z) = 1$ for solvent firms ($z > 0$) and the absorbing boundary condition $\phi_k(\tau, 0) = 0$.

Define the effective drift

$$\mu_k = r - \eta - \frac{1}{2}\hat{\sigma}_k^2, \tag{8}$$

and the distance metrics

$$d_{1,2}^k(\tau, z) = \frac{\pm z + \mu_k \tau}{\hat{\sigma}_k \sqrt{\tau}}, \tag{9}$$

where d_1^k uses $+z$ and d_2^k uses $-z$. The reflection basis is

$$\phi_k(\tau, z; \hat{\sigma}_k) = \begin{cases} 0, & z \leq 0, \\ \Phi(d_1^k) - e^{-\frac{2\mu_k z}{\hat{\sigma}_k^2}}\,\Phi(d_2^k), & z > 0. \end{cases} \tag{10}$$

where $\Phi(\cdot)$ is the standard normal CDF and the exponential term is the reflection factor.

Numerical Enforcement of Hard Constraints. While the analytical basis theoretically satisfies the boundary conditions, we strictly enforce these structural constraints in the implementation to ensure numerical stability.

First, the absorbing boundary condition is enforced via an explicit computational mask on the domain:

$$\phi_k(\tau, z; \hat{\sigma}_k) = 0 \quad \text{for } z \leq 0. \tag{11}$$

Second, to handle the singularity at $\tau = 0$ (where the term $1/\sqrt{\tau}$ in the distance metrics diverges), we override the analytical value with the exact initial condition when τ falls below a numerical tolerance threshold ε:

$$\phi_k(\tau, z; \hat{\sigma}_k) = \mathbb{I}_{\{z>0\}} \quad \text{for } \tau < \varepsilon. \tag{12}$$

In our implementation, we set $\varepsilon = 10^{-4}$ to guarantee robust convergence at the maturity boundary. The effective drift μ_k is determined by the risk-free rate r and the barrier growth rate η: This formulation ensures correct asymptotics: as $\tau \to 0$ for $z > 0$, $d_1^k \to +\infty$ and $d_2^k \to -\infty$, hence $\Phi(d_1^k) \to 1$ and $\Phi(d_2^k) \to 0$, recovering $\phi_k(0, z) = \mathbb{I}_{\{z>0\}}$.

3.2.3 Operator Synthesis: Spectral Expansion in 1D and 2D

The approximate solution is constructed as the inner product of the Branch coefficients and the Trunk reflection basis.

Individual survival (1D). For each regime $j \in \{B, b\}$, we approximate the marginal survival probability $v^j(\tau, z)$ by

$$\hat{v}^j(\tau, z; \boldsymbol{\lambda}) = \sum_{k=1}^{M} c_{j,k}(\boldsymbol{\lambda})\, \phi_k(\tau, z; \hat{\sigma}_k). \tag{13}$$

Joint survival (2D). For each regime $j \in \{B, b\}$, we approximate the joint survival surface $u^j(\tau, z_1, z_2)$ using a full tensor-product expansion:

$$\hat{u}^j(\tau, z_1, z_2; \boldsymbol{\lambda}) = \sum_{k=1}^{M} \sum_{l=1}^{M} c_{j,k,l}(\boldsymbol{\lambda}) \cdot \underbrace{\phi_k^{(1)}(\tau, z_1; \hat{\sigma}_{1,k})}_{\text{Firm 1 Basis}} \cdot \underbrace{\phi_l^{(2)}(\tau, z_2; \hat{\sigma}_{2,l})}_{\text{Firm 2 Basis}}. \tag{14}$$

Here $\phi_k^{(1)}$ and $\phi_l^{(2)}$ are the reflection bases for firm 1 and firm 2, respectively, each equipped with its own learned effective volatility set $\{\hat{\sigma}_{1,k}\}_{k=1}^{M}$ and $\{\hat{\sigma}_{2,l}\}_{l=1}^{M}$. The Branch output $\{c_{j,k,l}(\boldsymbol{\lambda})\}$ forms an interaction tensor that weights cross-terms between the two marginal bases, enabling a non-separable joint surface.

Although each basis factor is one-dimensional, the learned interaction tensor $\{c_{j,k,l}(\boldsymbol{\lambda})\}$ yields a non-separable approximation of the joint surface under the PDE coupling, with the correlation parameter ρ treated as fixed in Eq. 4.

3.3 Robust Training Objective

Let $\boldsymbol{\theta} = \{\theta_{br}, \theta_{\sigma}\}$ denote the set of all learnable network parameters. We optimize $\boldsymbol{\theta}$ by minimizing the residuals of the transformed PDE system using Automatic Differentiation (AD) to compute the exact operators.

3.3.1 PDE Residuals in Log-Moneyness Coordinates

In the transformed (τ, z)-coordinate system, the backward PDE transforms into a forward parabolic equation with constant coefficients. The pointwise residual $\mathcal{R}_j(\boldsymbol{\lambda}, \tau, \mathbf{z})$ for regime j is given by:

$$\mathcal{R}_j := \frac{\partial \hat{u}^j}{\partial \tau} - \mathcal{L}_j^{\text{joint}}[\hat{u}^j] - \lambda_j(\hat{u}^{\bar{j}} - \hat{u}^j) \approx 0, \tag{15}$$

where $\bar{j}$ denotes the complementary regime. Due to the coordinate transformation, the spatial operator $\mathcal{L}_j^{\text{joint}}$ simplifies to a constant-coefficient form:

$$\mathcal{L}_j^{\text{joint}}[u] = \sum_{i=1}^{2} \left(r^j - \frac{1}{2}(\sigma_i^j)^2 \right) \frac{\partial u}{\partial z_i} + \frac{1}{2} \sum_{i,m=1}^{2} \rho_{im} \sigma_i^j \sigma_m^j \frac{\partial^2 u}{\partial z_i \partial z_m}. \tag{16}$$

Note that the variable coefficients (e.g., $x\partial_x$) from the original system have been transformed into constant drift and diffusion terms, which significantly improves the conditioning of the optimization landscape.

3.3.2 Singularity-Robust Loss Function

The PDE residuals exhibit numerical instability near the maturity singularity ($\tau \to 0$). To mitigate gradient explosions, we employ a Huber Loss $\mathcal{H}_\delta$ with threshold $\delta = 0.1$ and a temporal weighting scheme. The total empirical risk $\mathcal{J}(\boldsymbol{\theta})$ over a batch of N_c collocation points is:

$$\mathcal{J}(\boldsymbol{\theta}) = \frac{1}{N_c} \sum_{i=1}^{N_c} \sum_{j \in \{B,b\}} \Big[w_{\text{PDE}}(\tau^{(i)})\, \mathcal{H}_\delta\Big(\mathcal{R}_j(\boldsymbol{\lambda}^{(i)}, \tau^{(i)}, \mathbf{z}^{(i)})\Big) + w_{\text{sum}}\, g(\tau^{(i)})\, \mathcal{H}_\delta\Big(\mathcal{R}_j^{\text{sum}}(\boldsymbol{\lambda}^{(i)})\Big) \Big], \tag{17}$$

where $w_{\text{PDE}}(\tau) := \tanh(5\tau)$ down-weights PDE residuals near $\tau = 0$. To enforce the terminal-scale normalization only in the near-maturity region, we use a cutoff gate

$$g(\tau) := \mathbb{I}_{\{\tau \le \tau_{\text{sum}}\}}, \qquad \tau_{\text{sum}} = 0.1. \tag{18}$$

3.4 Optimization and Discretization

Homotopy Continuation Schedule. To prevent the high-variance regime-coupling terms from destabilizing the learning process, we employ a three-phase curriculum learning schedule over the total 15,000 training epochs:

- **Phase I (Warm-up, Epochs 0–1,000):** We minimize only the terminal-scale normalization penalty to anchor the maturity behavior at $\tau = 0$. The PDE residuals are weighted at zero.
- **Phase II (Ramping, Epochs 1,000–5,000):** We linearly ramp the weights of the PDE residuals and the regime-coupling term from $0 \to 1$. Concretely, the coupling term $\lambda_j(\hat{u}^{\bar{j}} - \hat{u}^j)$ in the residual $\mathcal{R}_j$ is multiplied by a scalar homotopy factor $\alpha(e) = \frac{e-1000}{4000}$, where e is the current epoch.
- **Phase III (Refinement, Epochs 5,000–15,000):** The network trains on the full regime-switching system with constant weights to refine the solution accuracy in the asymptotic region.

Hybrid Domain Sampling. We construct training batches $\mathcal{B}$ using hybrid importance sampling on the time-to-maturity τ:

- **Boundary Layer (20%):** Samples $\tau \sim \mathcal{U}[\varepsilon, 0.1]$ to resolve sharp gradients near the maturity singularity.
- **Asymptotic Region (80%):** Samples $\tau \sim \mathcal{U}[0.1, 10]$ to capture the long-term survival decay.

Spatially, we sample z from the positive domain $\Omega_z = [0, \ln(3.0)]$ (and $(z_1, z_2) \in \Omega_z^2$ for 2D). Since the reflection basis enforces $\hat{u}(\tau, 0) = 0$ by construction, we do not require an explicit boundary-collocation set at $z = 0$; instead, we include near-boundary samples through the bulk distribution.

Pricing Integration. Fair CDS spreads are computed by integrating the learned survival profile against the discount factor $D(\tau) = e^{-r\tau}$. We employ a multi-resolution grid that adapts to the solution curvature: an hourly discretization ($\Delta\tau \approx 10^{-4}$) for $\tau \in [0, 0.02]$ to capture short-term jump-to-default risk, and a daily discretization for $\tau > 0.02$. The fair spread s (in basis points) is derived via the discretized risk-neutral valuation formula:

$$s = 10,000 \times \frac{(1-R)\sum_k D(\tau_k)\Delta\hat{u}_k}{\sum_k D(\tau_k)\bar{u}_k\Delta\tau_k}, \tag{19}$$

where $\Delta\hat{u}_k := \hat{u}(\tau_{k-1}) - \hat{u}(\tau_k)$ is the default probability over $(\tau_{k-1}, \tau_k]$, and $\bar{u}_k := \frac{1}{2}\big(\hat{u}(\tau_{k-1}) + \hat{u}(\tau_k)\big)$ is the interval-average survival probability. For single-name CDS pricing, $\hat{u}$ is taken as the marginal survival estimator (or its regime-mixture, depending on the pricing specification).

4 Numerical Experiments

In this section, we validate the proposed Reflection Deep Operator Network through a series of numerical experiments. We first detail the computational setup and training protocol, demonstrating the method's efficiency. We then present the main results, focusing on the sensitivity of credit spreads and default correlations to regime-switching intensities (λ_B, λ_b). Finally, we discuss the quantitative performance and stability of the learned operator.

4.1 Experimental Setup

Implementation and Computational Environment. The framework was implemented in PyTorch and executed on a Google Cloud Platform instance equipped with a single NVIDIA L4 GPU. A significant advantage of the proposed operator learning approach is its memory efficiency; despite the high dimensionality of the joint default problem, the training pipeline required only 2.5 GB of GPU VRAM. This low resource footprint facilitates substantial batch parallelism ($N_{batch} = 4096$) without the need for high-performance computing clusters.

Network Architecture. The neural architecture is designed to enforce physical plausibility by construction. It comprises two coupled networks:

- **Branch Network:** A Multi-Layer Perceptron (MLP) with dimensions $[2, 256, 256, 128]$ that encodes the regime transition parameters $\boldsymbol{\lambda} = (\lambda_B, \lambda_b)$ into a latent coordinate space.

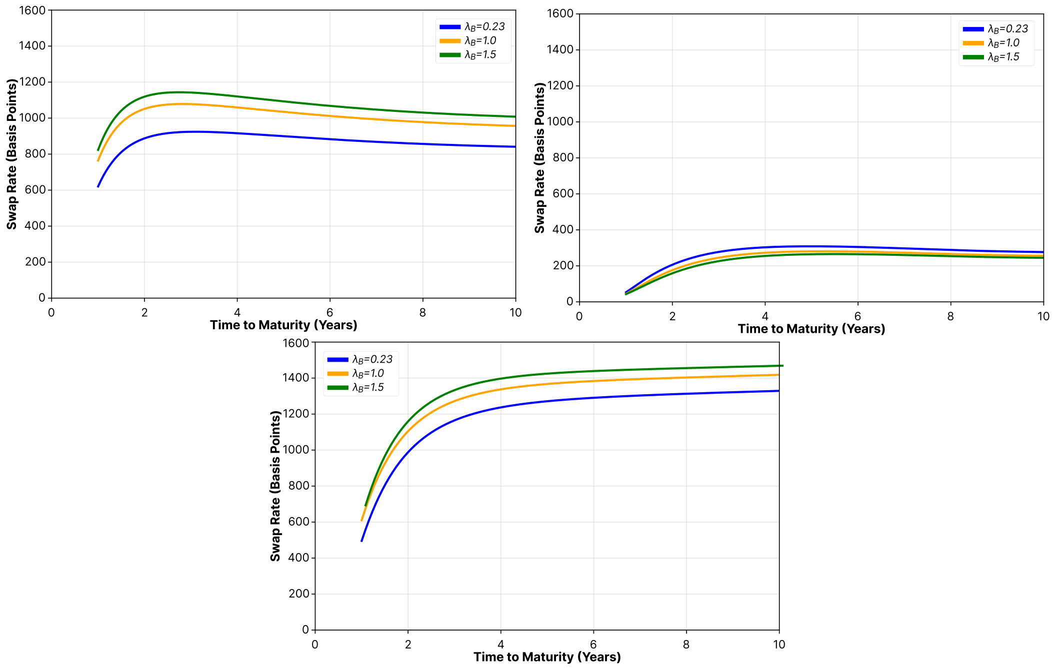

Figure 1: **Regime Sensitivity of Credit Spreads. (Top-Left)** Single-Name CDS with a risk-free counterparty shows monotonic sensitivity to instability λ_B. **(Top-Right)** Introduction of Counterparty Default Risk (CDR) compresses spreads and inverts the sensitivity due to bilateral CVA (the Fragile market curve drops below the Stable market curve). **(Bottom)** First-to-Default (FtD) basket spreads are structurally elevated due to weakest-link exposure.

- **Reflection-Based Trunk:** The trunk provides the reflection basis functions over the spatiotemporal coordinates, and the operator output is formed by a spectral expansion using the Branch coefficients. This construction enforces the absorbing boundary condition by design (in (τ, z) coordinates, the default boundary corresponds to $z = 0$).

Training Protocol. The optimization involved three independent models: two marginal models for the individual firms (Bank and Corporate) and one joint model for the coupled system. We employed a robust curriculum learning strategy over 15,000 epochs:

- **Loss Function:** To mitigate gradient explosions near maturity ($\tau \to 0$), we utilized the Huber loss (Smooth L1) rather than the mean squared error.
- **Optimization:** We used the Adam optimizer with an initial learning rate of 10^{-3}, coupled with a Reduce-on-Plateau scheduler (decay factor 0.5, patience 2000).
- **Sampling Strategy:** Collocation points were generated in time-to-maturity and log-moneyness coordinates (τ, z) for marginal models and (τ, z_1, z_2) for the joint model, with a higher sampling density near maturity ($\tau < 0.1$) to resolve the boundary layer.

4.2 Main Results: Regime Sensitivity Analysis

To validate the Reflection DeepONet's ability to generalize across the parameter manifold, we perform a sensitivity analysis on the regime transition intensities. Specifically, we vary the Bull-to-Bear intensity λ_B to test the market's sensitivity to economic instability, while holding the recovery intensity fixed at $\lambda_b = 1.73$.

Crucially, fixing $\lambda_b = 1.73$ implies an expected Bear market duration of $\tau_{bear} = 1/\lambda_b \approx 7$ months, representing a standard recessionary shock that is severe but temporary.

Configuration Details:

- **Transition Intensity (λ_B):** We examine three distinct stability regimes:
    1. **Stable Bull Market ($\lambda_B = 0.23$):** The economy is robust, with long expected durations of growth.
    2. **Unstable Bull Market ($\lambda_B = 1.0$):** The economy is volatile, with frequent transitions to distress.
    3. **Fragile Bull Market ($\lambda_B = 1.5$):** The economy is highly precarious; the onset of a recession is imminent.
- **Market Parameters:** The risk-free rates are set to $r^B = 5\%$ and $r^b = 2\%$. Asset volatilities increase from $\sigma^B = 20\%$ to $\sigma^b = 50\%$ in the Bear regime.

Figure 1 presents the term structure of fair spreads for three distinct credit instruments generated by the operator under these regimes.

Single-Name CDS (Risk-Free Counterparty). As shown in the **top-left panel**, spreads widen monotonically as the market becomes more fragile ($\lambda_B = 1.5$ vs. $\lambda_B = 0.23$). This reflects the heightened probability of transitioning into the high-volatility Bear regime. The term structure exhibits a characteristic "hump" at intermediate maturities (3–5 years), indicating that regime uncertainty commands the largest risk premium over medium-term horizons before mean-reverting.

Impact of Counterparty Default Risk (CDR). The **top-right panel** reveals a significant structural shift when the protection seller is correlated with the underlying asset. The spread levels are drastically compressed compared to the risk-free case. Notably, the sensitivity to λ_B *inverts*: the "Fragile" market curve drops below the "Stable" market curve. This is because, in a fragile economy, the probability of the insurer defaulting rises simultaneously with the asset's risk. This *Wrong-Way Risk* renders the protection less valuable in the exact states where it is needed most, forcing a discount on the fair premium.

Basket Default Swap (First-to-Default). In the **bottom panel**, the First-to-Default (FtD) spreads are structurally elevated, peaking above 1400 bps. This reflects the "weakest-link" mechanics ($\tau_{\text{FtD}} = \min(\tau_1, \tau_2)$), where risk is effectively additive. While spreads still rise with instability λ_B, the marginal widening is partially dampened by correlation saturation in the Bear regime, where clustered defaults limit the payout to a single event.

Overall, the operator correctly captures the non-linear interplay between regime instability and contract structure, consistent with the theoretical first-passage-time implications reported by Kim et al. [5].

5 Conclusion

We propose a self-supervised, boundary-aware neural operator that amortizes the solution of regime-switching first-passage coupled PDEs over parameter families, replacing repeated numerical solves with fast inference once training is complete. Beyond computational speedups, the operator-learning perspective provides a unified lens for analyzing how regime dynamics and cross-firm dependence propagate through first-passage default mechanisms, facilitating sensitivity analysis, scenario exploration, and stress testing at scale.

Our approach amortizes the solution of regime-switching first-passage coupled PDEs with a self-supervised neural operator, enabling efficient term-structure analysis across parameter families; however, several limitations remain. The regime specification is intentionally stylized, focusing on a two-state Markov chain with constant switching intensities, which may not capture richer time- or state-dependent regime dynamics. Moreover, the diffusion-based structural setting with absorbing default boundaries, while interpretable, may be restrictive in applications where jumps, stochastic volatility, or more flexible dependence structures are important. We also do not yet report a systematic quantitative benchmark against high-accuracy finite-difference (FDM) solutions across broad parameter sweeps.

Future work will therefore extend the regime process to multi-state and time/state-dependent transitions, broaden the structural dynamics beyond pure diffusions, and integrate the operator into calibration workflows that link amortized survival surfaces to market observables and parameter inference.

References

[1] Fischer Black and John C. Cox. Valuing corporate securities: Some effects of bond indenture provisions. *The Journal of Finance*, 31(2):351–367, 1976.

[2] Yu Chen, Xing Lü, Hao Tian, and Rui-Heng Li. Physics-informed neural network for barrier option pricing in coupled financial quantitative systems with varying interest rate and volatility. *Engineering Analysis with Boundary Elements*, 180:106457, 2025. doi: 10.1016/j.enganabound.2025.106457.

[3] Yu Chen, Xing L"u, Hao Tian, and Rui-Heng Li. Physics-informed neural network for barrier option pricing in coupled financial quantitative system with varying interest rate and volatility. *Engineering Analysis with Boundary Elements*, 180:106457, 2025.

[4] Darrell Duffie and Kenneth J. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720, 1999.

[5] In Joon Kim, Jaewoo Jang, and Seung Jin Lee. A first-passage-time model under regime-switching market environment. *Journal of Economic Dynamics and Control*, 32(10):3304–3328, 2008.

[6] Jinsil Lee, Youngjoon Hong, Seungchan Ko, and Jae Yong Lee. Data-free asymptotics-informed operator networks for singularly perturbed pdes. *arXiv preprint arXiv:2512.22006*, 2025.

[7] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations (ICLR)*, 2021.

[8] Lu Lu, Pengzhan Jin, and George E. Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3): 218–229, 2021.

[9] Robert C Merton. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, 29(2):449–470, 1974.

[10] Maziar Raissi, Paris Perdikaris, and George E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.

[11] Justin Sirignano and Konstantinos Spiliopoulos. Dgm: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics*, 375:1339–1364, 2018.

[12] Joel P. Villarino, Álvaro Leitão, and José A. García-Rodríguez. Boundary-safe pinns extension: Application to non-linear parabolic pdes in counterparty credit risk. *arXiv preprint arXiv:2210.02175*, 2022.

[13] Sifan Wang, Hanwen Wang, and Paris Perdikaris. Learning the solution operator of parametric partial differential equations with physics-informed deeponets. *Science Advances*, 7(40): eabi8605, 2021.

[14] Zhiwei Xu, Zongyi Zhang, and George E. Karniadakis. Data-free asymptotics-informed operator networks for singularly perturbed partial differential equations. *Journal of Computational Physics*, 473:111746, 2023.

[15] Chunsheng Zhou. An analysis of default correlations and multiple defaults. *Review of Financial Studies*, 14(2):555–576, 2001.

AI Co-Scientist Challenge Korea Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction propose a physics-informed operator for regime-switching PDEs. Section 3 details the Reflection DeepONet architecture, and Section 4 validates its ability to capture regime sensitivities in credit spreads as claimed.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: In the "Limitation and future work" section, we explicitly discuss the stylized two-state regime assumption, the restriction to diffusion-based structural dynamics, and the current lack of end-to-end market calibration.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: Section 2 clearly defines the stochastic assumptions (regime-switching GBM) and the coupled PDE system. The Reflection Principle construction in Section 3 is derived from standard analytical results for Brownian motion.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: Section 3.3 details the 15,000-epoch curriculum schedule and sampling strategy. Section 4.1 specifies the network architecture, optimizer settings (Adam), and batch size used to generate the results.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While AI Co-Scientist Challenge Korea does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [No]

   Justification: We plan to release the full source code and reproduction scripts upon acceptance of the paper. All experimental details (architecture, hyperparameters, sampling) are fully described in Section 4 to ensure reproducibility in the interim.

   Guidelines:

   - The answer NA means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Justification: Section 4.1 specifies the hyperparameters (Learning Rate 10^{-3}, Batch Size 4096), the Scheduler (Reduce-on-Plateau), and the domain sampling strategy.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [No]

   Justification: The results focus on a regime sensitivity analysis to demonstrate financial plausibility (monotonicity, term structure) rather than aggregated statistical performance metrics. A detailed quantitative benchmark against FDM is described as ongoing work.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Section 4.1 explicitly states the experiments were run on a Google Cloud Platform instance with a single NVIDIA L4 GPU, using approximately 2.5 GB of VRAM.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://nips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research involves mathematical modeling of financial derivatives. We have reviewed the code of ethics and found no conflicts.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: The work contributes to more accurate financial risk management (positive). We acknowledge the general model risk associated with deploying AI in regulated financial sectors (negative/cautionary).

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: The proposed model is a PDE solver for credit risk and does not involve high-risk generative capabilities or personal data.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: We use standard open-source libraries (e.g., PyTorch) which are properly cited.

Guidelines:

- The answer NA means that the paper does not use existing assets.

- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [No]

    Justification: We are not releasing new assets (code, models, or datasets) at the time of submission.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The research does not involve human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The research does not involve human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.