Data-driven Feynman–Kac Discovery with Applications to Prediction and Data Generation

Qi Feng
Department of Mathematics
Florida State University
Tallahassee, FL 32306
qfeng2@fsu.edu

Guang Lin
Department of Mathematics
Purdue University
West Lafayette, IN 47907
guanglin@purdue.edu

Purav Matlia
Department of Computer Science
Purdue University
West Lafayette, IN 47907
pmatlia@purdue.edu

Denny Serdarevic
Department of Mathematics
Florida State University
Tallahassee, FL 32306
ds22ck@fsu.edu

Abstract

In this paper, we propose a novel data-driven framework for discovering probabilistic laws underlying the Feynman–Kac formula. Specifically, we introduce the first stochastic SINDy method formulated under the risk-neutral probability measure to recover the backward stochastic differential equation (BSDE) from a single pair of stock and option trajectories. Unlike existing approaches to identifying stochastic differential equations—which typically require ergodicity—our framework leverages the risk-neutral measure, thereby eliminating the ergodicity assumption and enabling BSDE recovery from limited financial time series data. Using this algorithm, we are able not only to make forward-looking predictions but also to generate new synthetic data paths consistent with the underlying probabilistic law.

1 Introduction

Identifying the governing laws in physical systems is of high interest in various fields, including fluid dynamics [2], plasma dynamics [15], nonlinear optics [7], mesoscale ocean closures [22], and computational chemistry [11], etc. Many of these complex physical laws could be described by PDEs with proper boundary conditions, which are deterministic dynamical systems. For each specific scientific task, the detailed deterministic form of the PDE often remains unknown. In the era of big data, when rich experimental data is available, it gives rise to an opportunity to automatically transform this data into deterministic physical laws. In light of this idea, the data-driven discovery of hidden equations (mostly deterministic equations) has been enabled by the rapid progress in statistics and machine learning. As a toy example, we could select a set of candidate common functions containing derivatives with different orders, and call it the library $\Theta(u) := \{1, u, u_x, u^2, uu_x, u_x^2, \cdots, u_{xx}^2\}$. The discovery of the system/PDE $\mathbf{F}$ can be identified as the following symbolic regression problem,

$$\left(u_t(t_i, x_j)\right)_{i,j} = \begin{pmatrix} 1 & u(t_i, x_j) & u_x(t_i, x_j) & u^2(t_i, x_j) & u_x^2(t_i, x_j) & \cdots & u_{xx}^2(t_i, x_j) \end{pmatrix}_{i,j} \xi, \quad (1)$$

where $\{u(t_i, x_j)\}_{i,j}$ represents the data of the PDE, and t_i (resp. x_j) represents the time (resp. spatial) points. Data driven model discovery for PDE by using physics-informed learning has been studied in [6]. To bridge the deterministic law and probability law discovery in this context, we start with a slight modification of the problem by adding a terminal condition $u(T, x) = g(x)$ instead of the initial condition at $t = 0$, and simply consider the parabolic equation with linear coefficients,

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Generative AI in Finance.

$$\mathbf{F}(\cdot) := u_t + rxu_x + \frac{1}{2}\sigma^2 x^2 u_{xx} - ru + \tilde{f}(t,x) = 0, \tag{2}$$

where σ, r, T are parameters and $\tilde{f}$ is a known function. Due to the famous Feymann-Kac formula [14, 8], the solution $u(t,x)$ (if it exists) can be represented by the following conditional expectation,

$$u(t,x) = \mathbb{E}^{\mathbb{Q}}\Big[e^{-r(T-t)}g(X_T) + \int_t^T e^{-r(s-t)}\tilde{f}(s,X_s)ds \Big| X_t = x\Big], \tag{3}$$

under the probability measure $\mathbb{Q}$, where X_t is the solution of the following forward in time SDE,

$$dX_t = rX_t dt + \sigma X_t dB_t^{\mathbb{Q}}, \quad \text{with initial condition} \quad X_0 = x_0, \tag{4}$$

and B_t is Brownian motion under $\mathbb{Q}$. The dynamic X_t under the market measure $\mathbb{P}$ is unknown. By martingale representation theorem, the solution $u(t,x) := Y_t$ can be further represented by the following backward stochastic differential equation(BSDE), with the pair (Y_t, Z_t) being its solution,

$$Y_t = g(X_T) + \int_t^T [rY_s - \tilde{f}(s,X_s)]ds - \int_t^T Z_s dB_s^{\mathbb{Q}}, \quad \text{with} \quad Y_T = g(X_T). \tag{5}$$

Feynman-Kac formula is fundamental in probability theory, connecting SDE and its corresponding PDE. Feynman-Kac has been generalized to various fields including financial mathematics [3, 13, 16, 20], mathematical physics [1, 17, 12], Quantum mechanics [18, 5, 9], etc. More importantly, Feynman-Kac is paired with the underlying SDE model. The inverse problem has not been studied yet, which is to rediscover the Feynman-Kac formula by linking two different type of data sets collecting from $u(t,x)$ and X_t. In a real world application, the data $\{X_t\}_{0\le t\le T}$ could be collected from the daily stock price, and the solution $u(t,x)$ actually represents the European call option price at time t with terminal time payoff $g(X_T) = \max\{S_T - K, 0\}$, strike price K and maturity T. Our goal is to discover the relation (i.e. BSDE) from a pair of single trajectory (stock, option). The significant difference comparing to deterministic law discovery it that: we are learning PDE with terminal condition instead of initial condition. Compared to data-driven sparse identification of deterministic nonlinear dynamical systems (SINDy) [4], and stochastic dynamical systems [21], our probability law discovery is based on a pair of data. In particular, to identify the analytical form of the generator and diffusion part of the BSDE, we generalize the SINDy algorithm to stochastic SINDy under the risk-neutral probability measure. Once the BSDE is identified, combining with the risk-neutral dynamic of the stock price, we can generate new sample paths for (stock, option) pairs. This paper is organized as below. In Section 2, we present the main algorithm. In Section 3, we present the numerical algorithms for both the Black-Scholes model and real-world financial data.

2 Algorithm

The objective is to discover the differential form of the BSDE in the following general form,

$$Y_t = g(X_T) + \int_t^T f(s,X_s,Y_s,Z_s)ds - \int_t^T Z_s dB_s^{\mathbb{Q}}, \quad \text{with} \quad Y_T = g(X_T). \tag{6}$$

Since the available data are discrete, the discrete-time version is considered instead:

$$\Delta Y_{t_i} = -f(t, X_{t_i}, Y_{t_i}, Z_{t_i})\Delta t_i + Z_{t_i}\Delta B_{t_i}^{\mathbb{Q}}, \tag{7}$$

where Δt_i is the discrete timestep, $\Delta Y_{t_i} = Y_{t_{i+1}} - Y_{t_i}$, and $\Delta B_{t_i}^{\mathbb{Q}} := B_{t_{i+1}}^{\mathbb{Q}} - B_{t_i}^{\mathbb{Q}}$. The dataset consists of a single pair of discrete trajectories $D_u = \{(X_{t_i}, Y_{t_i})\}_{i=0}^{i=N}$ with terminal value $X_{t_N} = X_T$; $Y_{t_N} = Y_T$. To identify the dynamics, we formulate a sparse regression problem which minimizes:

$$\|\Delta Y_{t_i} - \Theta^f \xi^f \Delta t_i - \Theta^Z \xi^Z \Delta B_{t_i}^{\mathbb{Q}}\|^2. \tag{8}$$

The library matrix $\Theta = \Theta(u, X_{t_i})$ consists of two partitions: Θ^f, containing candidate functions for the generator f; and Θ^Z, containing candidate functions Z. Both partitions have sparse vectors of coefficients ξ^f and ξ^Z respectively. The discovered BSDE is then used for online prediction of Y_t and generation of new trajectories consistent with the dynamics.

Step 1 (DNN training): Numerical methods for obtaining partial derivatives are often inapplicable or erroneous when *only a single pair of trajectories* is available. To address this, a Deep Neural Network

(DNN) $\mathcal{N}$ is used to approximate the solution $u(t,x)$ with inputs (t_i, X_{t_i}) and parameters β [6]. Accurate derivatives are obtained by introducing a physics loss and applying automatic differentiation.

Optimization is performed using LBFGS, which updates both β and an auxiliary parameter ζ in the physics loss. The physics loss, computed on collocation points $\mathcal{D}_c$ randomly sampled from the bounds of the observed data, measures the residual $u_t - \Psi\zeta$ where the library matrix $\Psi = \Psi(X_t, u, u_x, u_{xx})$ is reconstructed with updated derivatives after each optimization step. The total loss function is

$$\mathcal{L}(\beta, \zeta; \mathcal{D}_u, \mathcal{D}_c) = \alpha \mathcal{L}_d(\beta; \mathcal{D}_u) + \gamma \mathcal{L}_p(\beta, \zeta; \mathcal{D}_c) + \delta \|\zeta\|_1 \tag{9}$$

where $\mathcal{L}_d$ is the data loss, $\mathcal{L}_p$ is the physics loss, and $\|\zeta\|_1$ is an L1 regularization term. Hyperparameters, α, γ, δ, control the relative weighting of each term. The auxiliary parameter ζ is discarded after training; it serves solely to improve the derivative accuracy by enforcing consistency of $\mathcal{N}$ with the underlying PDE (e.g.: equation 2) constraints.

Step 2.0 (Estimating diffusion coefficients): Construction of the library $\Theta(u, X_{t_i})$ requires the discrete Brownian increments $\Delta B_t^{\mathbb{Q}}$. These increments are obtained by first estimating the diffusion coefficient $\sigma(x)$. A preliminary estimate is computed as $\sigma_{\text{noisy}} = \sqrt{\Delta \langle X_{t_i} \rangle / \Delta t_i}$, which is often noisy due to the discrete sampling of the trajectory. We denote $\Delta\langle X_{t_i}\rangle$ as the increment of the quadratic variation of the stock price X_t. To obtain a smoother and more accurate approximation, the noisy estimate is modeled using SINDy by solving $\sigma_{\text{noisy}}(X_{t_i}) = \Phi(X_{t_i})\nu$ where $\Phi(X_{t_i})$ is a library of candidate functions and ν is the sparse solution vector of coefficients. The resulting diffusion function is discovered by applying SINDy to $\|\hat{\sigma}(X_{t_i}) - \Phi(X_{t_i})\nu\|^2$ similar to (8).

Step 2.1 (Extracting $\mathbb{Q}$ Brownian motion): Once the smoothed diffusion function $\hat{\sigma}(x)$ is obtained, the discrete Brownian increments $\Delta B_{t_i}^{\mathbb{Q}}$ are calculated from the observed trajectory by rearranging the discrete form of the SDE (4): $\Delta B_{t_i}^{\mathbb{Q}} = (\Delta X_{t_i} - r X_{t_i} \Delta t)/\hat{\sigma}(X_{t_i})$. The drift coefficient r is assumed to be a fixed, and corresponds to the risk-neutral probability measure $\mathbb{Q}$. These extracted increments are subsequently used to construct the library matrix Θ.

Step 3 (Stochastic SINDy for BSDE): With the library matrix Θ constructed, the discrete BSDE is identified using stochastic SINDy under probability measure $\mathbb{Q}$. The sparse regression problem in (8) is solved using sparsity-promoting algorithms with regularization, in particular, Sparse Relaxed Regularized Regression (SR3) [23] is employed. The result is a sparse set of coefficients that identify the driver $f = \Theta^f \xi^f$ and $Z = \Theta^Z \xi^Z$. In the current version, we do not provide terminal boundary discovery and we hope to add it in the future.

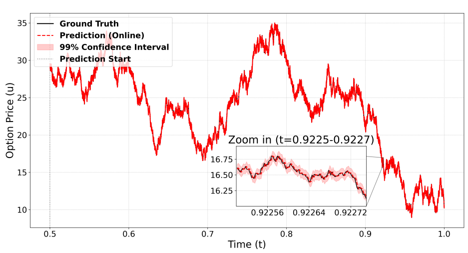
Step 4 (Online Prediction): Online predictions for $Y_{t_{i+1}}$ are performed using a memory window $(X_{t_s}, Y_{t_s})_{s \in [i-\tau, i]}$ of fixed size τ spanning from $t_{i-\tau}$ to t_i. The dynamic of Y_t is backward in time, which is solved backward in time. Different from the deep neural network algorithms [10, 19] with BSDE model parameter given (i.e. f, σ fixed), we are not solving the BSDE as the model is unknown for the future time interval $[t, T]$, but rather take the random form with unknown increment of the Brownian motion $\Delta B_{t_i}^{\mathbb{Q}}$. With the real data, Y_t is known from the market, we simply want to make a prediction for tomorrow's option price. We use the formula (7) with Z_t independent of $\Delta B_{t_i}^{\mathbb{Q}}$. This prediction loop is interleaved with the neural network $\mathcal{N}(t, x; \beta)$ retraining to regenerate partial derivatives and with rediscovery of the dynamics, ensuring consistency with the most recent memory window. We make mean prediction (mimicking $\mathbb{E}_t(Y_{t_{i+1}})$) and confidence intervals based on the desired confidence level by generating independent increments $\Delta B_t^{\mathbb{Q}}$. After each prediction, the memory window is shifted forward by one time step $(X_{t_s}, Y_{t_s})_{s \in [i-\tau+1, i+1]}$ and the dynamic rediscovery and prediction (i.e. **Step 1-4**) is repeated. For data points immediately preceding time skips caused by market opening and closing, the true value of the stock or option after the skip is used instead of a prediction. Our algorithm does not account for market opening and closing times (i.e., non-trading intervals), and we hope to add this feature in the future.

Step 5 (Generation): New trajectories are generated by running multiple simulations of the stock price using learned forward dynamic (4) and using the discovered BSDE (7) by sampling Brownian increments to create trajectories consistent with the learned dynamics with shifted window $(X_{t_s}, Y_{t_s})_{s \in [i-\tau, i]}$ as described in **Step 4**. For online prediction and data generation, our generated option paths are not option prices by solving a given BSDE model [10, 19] backward in time, as the models (both stock and option dynamics) are unknown. We generate the option paths by iteratively updating our models and making future predictions.

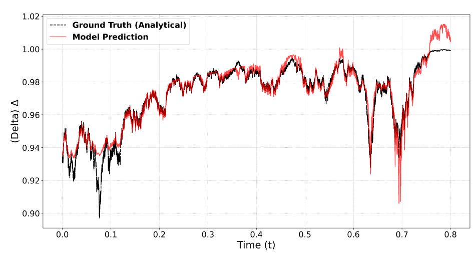
3 Examples

Example 1: Black-Scholes discovery. We use the Black-Sholes model as a benchmark to test our algorithm. For this example, we can compare the analytical solution and its derivatives to our auto-differentiated partial derivatives from our Neural Network in **Step 1**. Following our algorithm, once the BSDE is trained in **Step 3**, given the current market option price Y_t (assumed to be computed from the analytical solution), we can compute the next option price Y_{t+1} given we know the increment of the Brownian motion. We then use this idea to do prediction and generating new data as in **Step 4-5**. To make our prediction and generation consistent with our model and analytical solution, we need to retrain the neural network with the shifting window to update more accurate values of the derivatives of $u(t, x)$. Our numerical results are consistent with our desired algorithm performance. To generate the data pair of stock and option, we simulate the underlying asset price following the geometric Brownian motion, $dS_t = \mu S_t dt + \sigma S_t dB_t^{\mathbb{P}}$, with $\mu = 0.3$, $\sigma = 0.2$. Applying Girsanov theorem with a selected risk-free interest rate $r = 0.1$, we get (4) under $\mathbb{Q}$. Then, the corresponding European call options Y_t are computed using the closed-form formula for the Black-Sholes PDE (i.e. (2) with $\tilde{f} = 0$) with maturity time $T = 1$ and payoff $g(S_T) = \max\{S_T - K, 0\}$. We train the model on $[0, 0.8]$ and test the model for the interval $[0.8, 1]$, with a total number of time steps $N = 5 * 10^5$. By training from **Step1-3**, our stochastic SINDy discovers the BSDE $dY_t = -[1.018(u_t)(dt) + 0.106(S_t)(u_s)(dt) + 0.016(S_t^2)(u_{ss})(dt)] + 0.2(S_t)(u_s)(dB_t^{\mathbb{Q}})$, where we use $(u_t), (u_s), (u_{ss})$ to represent the derivatives computed from the neural network. The numerical results for the Black-Scholes model is presented in Figure [1-(a)(b)(c)(d)].

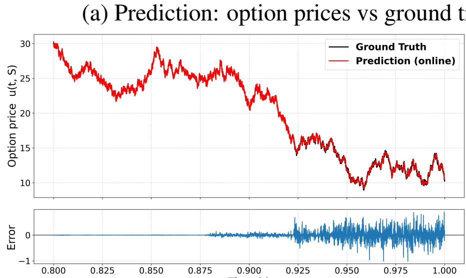
Example 2: Market data Apple. Data was collected from the ticker AAPL from 11/14/2024 up to expiration 07/18/2025 with timer intervals of 1 second at a strike price at $K = 170$. We set $T = 1$ with $N = 3907974$ and we use 80%-20% train-test split. We present the prediction in Figure [1-(e)] and generate stock-option pairs in Figure [1-(f)(g)].



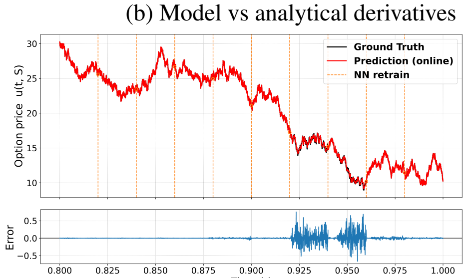
(a) Prediction: option prices vs ground truth



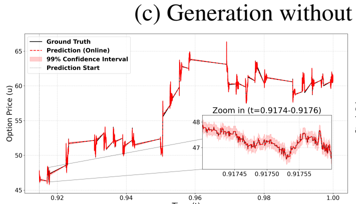
(b) Model vs analytical derivatives



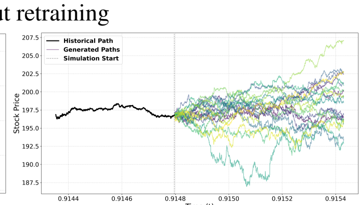
(c) Generation without retraining



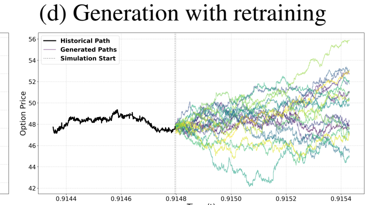
(d) Generation with retraining



(e) Online prediction for Apple



(f) Generation for Apple stock paths



(g) Generation for Apple option paths

Figure 1: Numerical examples for Black-Scholes and Apple stock

4 Conclusion

We propose a stochastic SINDy algorithm for BSDE discovery (i.e., Feynman–Kac representation) under the risk-neutral probability measure, without requiring the ergodicity assumption. To the best of our knowledge, this is the first stochastic SINDy framework developed for backward stochastic differential equations, and the first algorithm to employ generalized stochastic SINDy to uncover probabilistic laws directly from a pair of financial time series—illustrated here with stock and option data. Beyond finance, this new algorithm points to a promising direction for the broader discovery of Feynman–Kac–type probabilistic laws in engineering and physics.

Acknowledgement

Qi Feng is partially supported by the National Science Foundation under grant #DMS-2420029. Denny Serdarevic is partially supported by the National Science Foundation under grant #DMS-2420029 thorough the REU project at FSU. Guang Lin is partially supported by the National Science Foundation (NSF) (DMS-2533878, DMS-2053746, DMS-2134209, ECCS-2328241, CBET-2347401, and OAC-2311848), and DOE Office of Science Advanced Scientific Computing Research program DE-SC0023161, and DOE–Fusion Energy Science, under grant number: DE-SC0024583.

References

[1] Ioannis Anapolitanos and Marcel Griesemer. Multipolarons in a constant magnetic field. In *Annales Henri Poincaré*, volume 15, pages 1037–1059. Springer, 2014.

[2] Zhe Bai, Steven L Brunton, Bingni W Brunton, J Nathan Kutz, Eurika Kaiser, Andreas Spohn, and Bernd R Noack. *Data-driven methods in fluid dynamics: Sparse classification from experimental data*. Springer, 2017.

[3] Tomas Björk. *Arbitrage theory in continuous time*. Oxford university press, 2009.

[4] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, 113(15):3932–3937, 2016.

[5] Michel Caffarel and Pierre Claverie. Development of a pure diffusion quantum monte carlo method using a full generalized feynman–kac formula. i. formalism. *The Journal of chemical physics*, 88(2):1088–1099, 1988.

[6] Zhiping Chen, Yongxin Liu, and Hao Sun. Physics-informed learning of governing equations from scarce data. *Nature Communications*, 12:6136, 2021.

[7] Andrei V Ermolaev, Anastasiia Sheveleva, Goëry Genty, Christophe Finot, and John M Dudley. Data-driven model discovery of ideal four-wave mixing in nonlinear fibre optics. *Scientific Reports*, 12(1):12711, 2022.

[8] Richard Phillips Feynman. Slow electrons in a polar crystal. *Physical Review*, 97(3):660, 1955.

[9] Brian C Hall. *Quantum theory for mathematicians*, volume 267. Springer Science & Business Media, 2013.

[10] Jiequn Han, Arnulf Jentzen, and Weinan E. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018.

[11] Farshad Harirchi, Doohyun Kim, Omar Khalil, Sijia Liu, Paolo Elvati, Mayank Baranwal, Alfred Hero, and Angela Violi. On sparse identification of complex dynamical systems: A study on discovering influential reactions in chemical reaction networks. *Fuel*, 279:118204, 2020.

[12] Benjamin Hinrichs and Oliver Matte. Feynman–kac formula and asymptotic behavior of the minimal energy for the relativistic nelson model in two spatial dimensions. In *Annales Henri Poincaré*, volume 25, pages 2877–2940. Springer, 2024.

[13] Mark S Joshi. *The concepts and practice of mathematical finance*, volume 1. Cambridge University Press, 2003.

[14] Mark Kac. On distributions of certain wiener functionals. *Transactions of the American Mathematical Society*, 65(1):1–13, 1949.

[15] Alan A Kaptanoglu, Christopher Hansen, Jeremy D Lore, Matt Landreman, and Steven L Brunton. Sparse regression for plasma physics. *Physics of Plasmas*, 30(3), 2023.

[16] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. springer, 2014.

[17] Hartmut Löwen. Spectral properties of an optical polaron in a magnetic field. *Journal of mathematical physics*, 29(6):1498–1504, 1988.

[18] Barry Simon. *Functional integration and quantum physics*, volume 86. American Mathematical Soc., 1979.

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

[20] Hanxiao Wang, Jiongmin Yong, and Jianfeng Zhang. Path dependent feynman–kac formula for forward backward stochastic volterra integral equations. In *Annales de l'Institut Henri Poincare (B) Probabilites et statistiques*, volume 58, pages 603–638. Institut Henri Poincaré, 2022.

[21] Mathias Wanner and Igor Mezić. On higher order drift and diffusion estimates for stochastic sindy. *SIAM Journal on Applied Dynamical Systems*, 23(2):1504–1539, 2024.

[22] Laure Zanna and Thomas Bolton. Data-driven equation discovery of ocean mesoscale closures. *Geophysical Research Letters*, 47(17):e2020GL088376, 2020.

[23] Peng Zheng, Travis Askham, Steven L. Brunton, J. Nathan Kutz, and Aleksandr Y. Aravkin. A unified framework for sparse relaxed regularized regression: Sr3. *IEEE Access*, 7:1404–1423, 2019.

NeurIPS 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: Our abstract and introduction is consistent with our results.

   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: We mention the limitations of our algorithm.

   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: [NA]

Justification: We do not prove theorems, we propose new algorithms.

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 the 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: The algorithm and the parameters in the paper are sufficient to reproduce the numerical 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 NeurIPS 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: The code will be released in the final version.

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: We provide a high-level overview of the method and main components (optimizer, train-test splits), low-level implementation details will be released in the final version.

   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: [Yes]

   Justification: We provide confidence intervals and we also plot the errors in some of the numerical examples.

   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: [No]

   Justification: Our method is hardware-independent, requiring only a GPU of reasonable performance.

   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://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We have read the ethics review guidelines and ensured that our paper conforms to it.

   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: [NA]

    Justification:

    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: [NA]

    Justification:

    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: [NA]

    Justification:

    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 will release the assets in the final version. It is not documented here.

    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: [NA]

    Justification:

    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: [NA]

    Justification:

    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.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification:

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.