Hybrid Physics-Informed Neural Networks and Diffusion Models for Option Pricing in Stochastic Markets

Ethan McKeen
Department of Electrical and Computer Engineering
University of British Columbia
Vancouver, BC V6T 1Z4
emckeen1@ece.ubc.ca

Eric Björfors
Department of Electrical and Computer Engineering
University of British Columbia
Vancouver, BC V6T 1Z4
erbj5772@ubc.ca

Abstract

In this project, we propose a novel approach to financial option pricing by combining Physics-Informed Neural Networks (PINNs)[1] with Diffusion Models (DMs) [2]. Traditional methods like the Black-Scholes [3] equation assume constant volatility, which can lead to inaccurate pricing in real-world markets characterized by stochastic volatility and asset price fluctuations. Other models such as the Heston model, have a cha To address this limitation, we propose a hybrid model where the Diffusion Model simulates the stochastic dynamics of asset prices and volatility, while the PINN solves the Black-Scholes PDE, incorporating these stochastic elements. This approach allows us to handle both deterministic and stochastic components especially under volatile market conditions. The model will be validated against real-world market data and compared with existing methods, such as the Black-Scholes formula and Monte Carlo simulations [4].

1 Introduction

1.1 Problem Statement

Option pricing is a fundamental problem in finance as options are in risk management. The traditonal Black-Scholes model has been widely used for this purpose. However, one of its key limitations is the assumption of constant volatility. In real financial markets, volatility is stochastic, which can lead to inaccuracies in option pricing when using the traditional Black-Scholes formula. As a result, more complex models, such as the Heston model, have been developed to capture stochastic volatility, but they come with significant computational costs and are challenging to solve analytically.

1.2 Importance

Addressing the limitations of the Black-Scholes model is crucial for accurately pricing financial derivatives, especially in volatile markets. Mispricing options can lead to significant financial risks, and existing methods either oversimplify or are computationally intensive. By using a hybrid model that combines Diffusion Models (DMs) to simulate stochastic volatility with Physics-Informed Neural

Preprint. Under review.

Networks (PINNs) to solve the underlying Black-Scholes PDE, we aim to improve accuracy while maintaining computational efficiency.

2 Related Work

Physics-Informed Neural Networks (PINNs) [1]: Raissi et al. (2019) proposed PINNs as a framework for solving PDEs by integrating known physical laws into neural networks. PINNs have shown promise in solving high-dimensional PDEs and can be adapted for solving financial PDEs, such as the Black-Scholes equation, in a more flexible manner.

Diffusion Models (DMs) [2]: The paper "Denoising Diffusion Probabilistic Models" by Jonathan Ho, Ajay Jain, and Pieter Abbeel (2020) introduced Diffusion Models as a novel framework for generative modeling. These models involve a two-step process, a forward diffusion step where noise is gradually added to data, and a reverse diffusion step where a neural network is trained to reverse this process, reconstructing data from noise. These models have shown great performance in image-generation, for modelling stochastic processes due to the diffusion process.

Our proposed hybrid model builds on these works by integrating stochastic volatility (using DMs) into the PINN framework, enabling it to handle both the deterministic and stochastic components of option pricing.

3 Model/Method

3.1 Hybrid Model Overview

The proposed hybrid model combines a Diffusion Model with a Physics Informed Neural Network to price options. The Diffusion Model is trained to generate stock paths, simulating the market dynamics. The paths are then used as input into the PINN, which have been to output a option price given real stock paths and physical constraints from the Black Scholes equation (Appendix 6.1). A visualisation of the overview can be seen in Figure 1

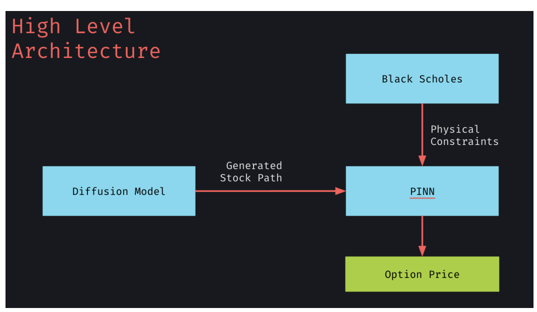

Figure 1: Overview of the Hybrid Model

3.2 Diffusion Model Implementation

The diffusion model is used to generate stochastic stock paths that capture learned market dynamics from the training data. It processes 1D time series data as input and operates over a forward diffusion process of 1000 timesteps. Gaussian noise is incrementally added using a cosine beta schedule, which gradually scales up the noise during the forward process to ensure smooth transitions.

The underlying neural network is a 1D Residual Network (ResNet1D). Its architecture is as follows:

Initial Layer: A 1D convolutional layer with a kernel size of 5, padding of 2, and 64 hidden channels, followed by batch normalization and a SiLU activation function.

Residual Blocks: The network contains 6 residual blocks. Each block consists of two 1D convolutional layers (with kernel size 5 and padding 2), each followed by batch normalization and a SiLU activation. Residual connections ensure efficient gradient flow and help capture hierarchical patterns in the time series data.

Final Layer: A single 1D convolutional layer maps the hidden representation back to one output channel, which predicts the added noise for denoising in the reverse diffusion process.

During training, the Mean Square Error loss is calculated between the predicted and the actual noise. The weights are then adjusted using the Adam optimizer.

During generation, we start with a 1d time series of pure gaussian noise, which is equivalent time step 1000. Using the residual network we predict the noise added from the previous time step and remove the noise. This process continues for all time steps, leading to a fully denoised stock path at time step 0.

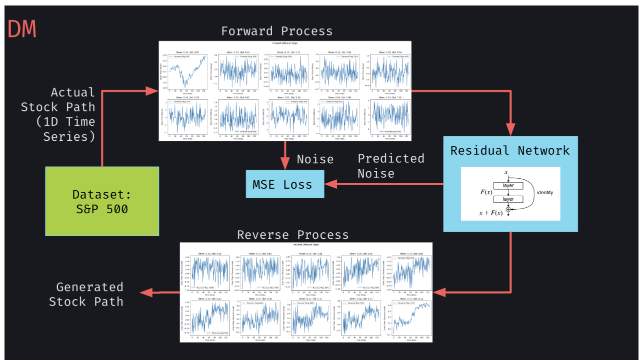

Figure 2: DM Architechture overview

3.3 Physics-Informed Neural Network (PINN):

The PINN solves the Black-Scholes PDE, incorporating the stochastic volatility and asset price paths generated by the DM. The PINN learns the option price by enforcing the physical constraints of the Black-Scholes equation, while accounting for the real-world stochasticity. Considering that the DM generates multiples paths, the PINN is used to predict an option price for each stock path. When evaluating the accuracy of the hybrid model all of the individual option prices as well as the average option price is considered.

3.4 PINN Implementation

To implement the Physics-Informed Neural Network (PINN) for option pricing, we utilize four months of historical stock data for S&P 500 stocks and their corresponding real-world option prices, obtained using the yfinance library. The dataset is then split into an 80-20 ratio to create training and validation sets, ensuring a balanced approach to model evaluation.

The core of the PINN is a feedforward neural network with three hidden layers. This architecture projects the input stock data into lower-dimensional feature spaces, enabling the network to capture important relationships, before reducing these features to a one-dimensional output representing the predicted option price. ReLU activation functions are used in the hidden layers to effectively model non-linear relationships in the data. By leveraging real-world stock data, the model learns general market trends, which enhances its ability to generalize to unseen scenarios and predict option prices accurately. The overall training and loss calculation architecure can be seen below in figure 1.

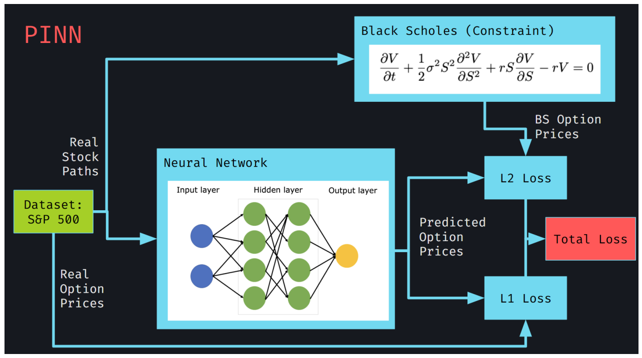

Figure 3: PINN Training Architecture

3.4.1 Black Scholes Implementation

The Black-Scholes model is utilized as a foundational component to inform the PINN's predictions. When predicting option prices using this model, a strike price of 100 is assumed for consistency. Since the model is trained on four months of stock data, the time to maturity for the options is set to four months. A risk-free interest rate of 0.05 is assumed, reflecting typical market conditions.

Volatility, a critical input to the Black-Scholes model, is calculated on a rolling basis for each stock in the dataset using a window of 30 data points. This method acknowledges that volatility is not constant across the dataset and allows for capturing its dynamic nature over time. While the model assumes an initial volatility of 0.2, experiments indicate that significant fluctuations in volatility during the four-month maturity period can decrease the model's performance. This highlights the importance of our models ability to capture a non constant volatility. Fluctuating volatility means greater uncertainty in the price movements of the underlying asset. This increased uncertainty makes it harder to predict future prices accurately.

3.4.2 Constraint and Loss Calculation

The loss function in the PINN model integrates two primary components. The first contribution is derived from the difference between the option price predicted by the Black-Scholes model and the PINN's predicted option price. The second contribution comes from the difference between the real-world option price, obtained using yfinance, and the PINN's predicted price.

For the first component, an L2 loss is used to measure the deviation from the Black-Scholes estimated option price. The L2 loss is well-suited for this task because the Black-Scholes framework is a well-established model, and data outliers are unlikely. The use of L2 loss ensures that large deviations are penalized more heavily, encouraging the PINN to align closely with the Black-Scholes model.

For the second component, an L1 loss is employed to capture the difference between real-world option prices and the PINN's predictions. Real-world data is often noisy and may contain outliers. By using the L1 loss, which is less sensitive to outliers than L2 loss, the model prioritizes learning general market trends over fitting to outliers. This design choice ensures that the model achieves robust predictions that are representative of the broader market rather than overfitting to noisy or anomalous data points.

4 Experiments

4.1 Experiment 1: Diffusion Model Evaluation

The Diffusion Model's performance is evaluated against the dynamics of Geometric Brownian Motion and the Heston Model (Appendix 6.2). These are loosely calibrated using historical stock data, however they are not calibrated to perfection. This should be considered when enterpretating the results. Using the three different models, 100 stock paths are generated and the Mean Absolute Error to the real stock path is calculated to find which model best captures the stock dynamics. This is done for two stocks, AAPL and MSFT, with different behaviour during the examined time period. The

AAPL stock has an upwards trend with high volatility, while the MSFT stock shows more stability with a bit of volatility. We can see that the Diffusion model outperforms both the GBM and Heston Model when the stocks have high volatility, as seen in Table 1 and Figure 4. When the behaviour is more stable, the Diffusion model underperforms both other models, as seen in 5 and Figure 5. This shows the possibilities of using Diffusion Model for stock path generation by just using 1d time series. However, more experimentitaion is required to find a model that can generalise to stable market dynamics as well.

Model	Best MAE
Heston Model	7.2217
Diffusion Model	6.5841
Geometric Brownian Motion (GBM)	16.3607

Table 1: Comparison of Best MAE for AAPL stock

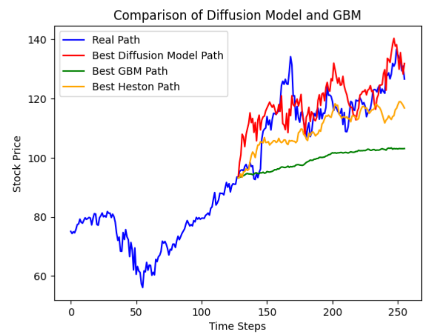

Figure 4: Best AAPL path for DM, GBM and Heston Models

Model	Best MAE
Heston Model	5.4844
Diffusion Model	10.8456
Geometric Brownian Motion (GBM)	4.5484

Table 2: Comparison of Best MAE for MSFT stock

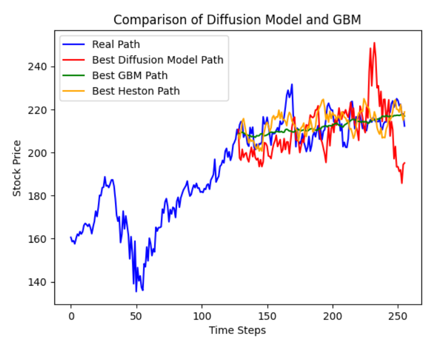

Figure 5: Best MSFT path for DM, GBM and Heston Models

4.2 Experiment 2: Physics Informed Neural Network Evaluation

The performance of the PINN is evaluated against the real option prices (collected using yfinance) and the option price calculated using black scholes equation. The mean absolute error for the models is summarized in table 3.

Model	MAE
Black-Scholes	34.55
PINN	6.92

Table 3: Comparison of MAE for PINN

This experiment showed that high fluctuations in volatility had great impacts on the performance of both models. As seen in figure 6, the stock path has a relativly constant volatility and both models perform very well. In figure 7 the volatility fluctuates much more often and the PINN performs worse than the constant volatility case. However, the PINN performs significantly better than the Black-Scholes prediction, which is expected considering Black-Scholes assumes constant volatility, which means that the PINN is able to adapt to non constant valatility cases far better.

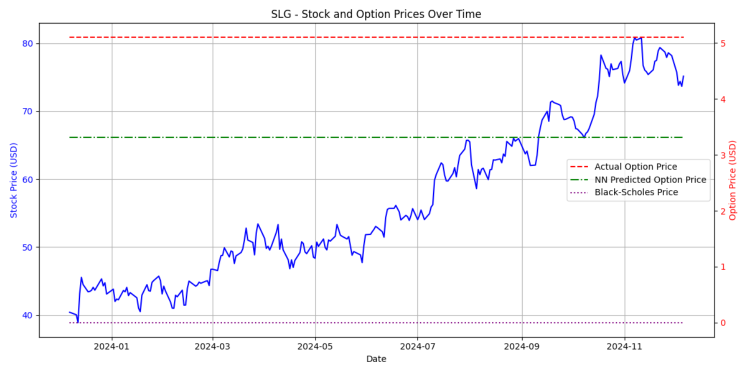

Figure 6: Low Volatility FLuctuation PINN Comparison

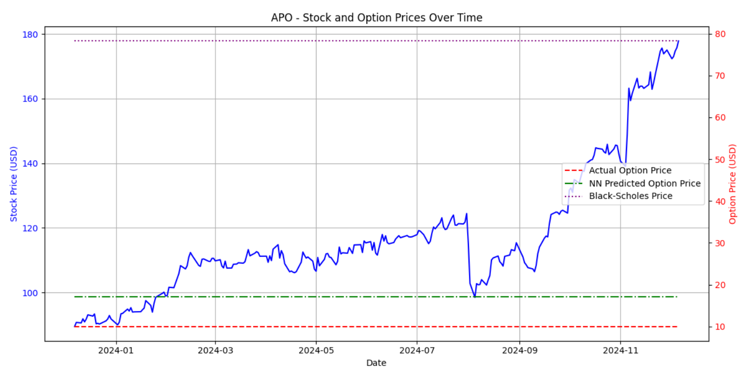

Figure 7: High Volatility Fluctuation PINN Comparison

4.3 Experiment 3: Hybrid Model vs. Heston and GBM

The hybrid model's performance is compared against the Heston and GBM models. As previously mentioned in experiment 1, stock paths can be generated utilizing the DM, Heston, and GBM approach. These stock paths can be fed into the PINN to compare their option price predictions. The hybrid model, which used the DM path to predict the option price, performed very similarly to the Heston model. The GBM model consistently outperformed both and this is likely because of the PINN's sensativity to fluctuations in volatility, as previously discussed. As seen in the experiment 1 figure, GBM stock path is far smoother and has a much more stable volatility. These results can be seen below in figure 8 which shows a test case in which the models made predictions on the AAPL stock.

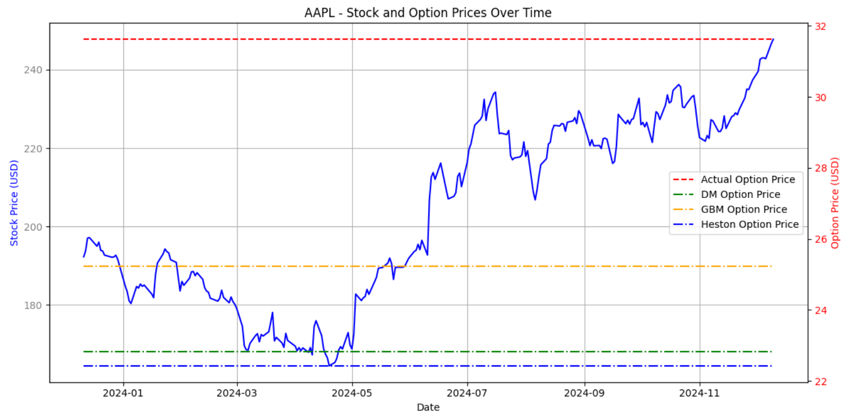

Figure 8: Option Price Predictions for DM, GBM and Heston Models

Model	MAE
Heston Model	9.19
Hybrid Model	8.79
GBM	6.39

Table 4: Comparison of MAE for Option Price Prediction

It is worth noting that the supervised learning contribution to the training of the PINN had no interaction with the DM, Heston, or GBM model. It was trained purely on real stock data so this experiment enforces the claim that both the PINN and DM are able to generate market generalized and well performing results.

4.4 Experiment 4: Ablation Study (PINN Only vs Hybrid Model)

Comparing the PINN alone, predicted with real stock paths, with the hybrid model, where the PINN makes predictions with DM generated path, allows investigation into the value added by the DM in conjunction with the PINN. Also considering a real use case for this model will not have access to future stock data this experiment provides insight into how comparable the option price predictions are when using real stock data versus predicted stock data. The PINN alone tends to perform slightly better than the hybrid model. This could be a result of slight overfitting but both models perform very comparably as seen by table 5. The main reason for this slight disparity in performance is that if a stock path has minimal fluctuations in volatility, these fluctuations tend to increase wit the DM generated paths. As shown in experiment 2, the PINN is sensative to these fluctuations. This behaviour can be seen in figure 9 because, despite the SLG stock having a relatively constant volatility, the PINN noteably alone outperforms the hybrid model.

Model	MAE
PINN Only	6.92
Hybrid Model	8.79

Table 5: Comparison of MAE for Ablation Study

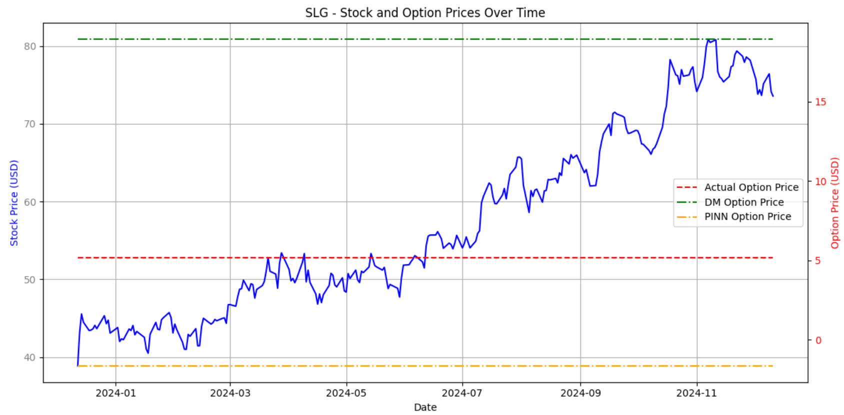

Figure 9: PINN vs Hybrid

5 Conclusion & Future Work

This project aims to improve the accuracy of option pricing by using a hybrid model that combines Diffusion Models (DMs) for simulating stochastic stock price paths with Physics-Informed Neural Networks (PINNs) to solve the Black-Scholes equation. This approach can better capture real-world market conditions, providing more reliable pricing for options, particularly in volatile markets. The results found the hybrid models predictions to be comparable to the Heston method of option price prediction. However, certain aspects like fluctuation in stock volatility did affect performance. Also, the DM struggles to generate stock paths with lower volatility. Future work could try to address this issue and make the DM even more generalizable. The PINN can also be further improved if a dataset of DM generated paths is made allowing PINN to train on real stocks paths and DM stock paths. This will allow the PINN to better recognize behaviours of the DM and gain a wider training dataset to improve its generalizeability.

References

[1] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations,” arXiv:1711.10561 [cs, math, stat], Nov. 2017, Available: https://arxiv.org/abs/1711.10561

[2] J. Ho, A. Jain, and P. Abbeel, “Denoising Diffusion Probabilistic Models,” arXiv:2006.11239 [cs, stat], Dec. 2020, Available: https://arxiv.org/abs/2006.11239

[3] Wikipedia Contributors, “Black–Scholes model,” Wikipedia, Jul. 13, 2019. https://en.wikipedia.org/wiki/Black-Scholes$_m odel$

[4] S. Raychaudhuri, “INTRODUCTION TO MONTE CARLO SIMULATION.” Available: https://www.informs-sim.org/wsc08papers/012.pdf

[5] S. L. Heston, “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options,” Review of Financial Studies, vol. 6, no. 2, pp. 327–343, Apr. 1993, doi: https://doi.org/10.1093/rfs/6.2.327.

[6] “Geometric Brownian motion,” Wikipedia, Dec. 19, 2020. https://en.wikipedia.org/wiki/Geometric$_B rownian_m otion$

6 Appendix

6.1 Black-Scholes Equation:

The classic work by Black and Scholes (1973) [3] introduced a closed-form solution for European options under the assumption of constant volatility. The introduced equation was:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{1}$$

Where:
$V(S, t)$ is the option price.
S is the underlying asset price.
r is the risk-free rate.
σ is the volatility.

The stochasticity comes from the underlying asset S, which is a stochastic variable following the dynamics of geometric brownian motion:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2}$$

Where:
μ is the drift rate, representing the expected return of the asset.
σ is the volatility.
W_t is a Wiener process (Brownian motion) which introduces stochastic behaviour.

While this model is foundational, its assumption of constant volatility has been challenged by real volatility observations in the market.

6.2 Heston Model:

The Heston model (1993) [5] extended the Black-Scholes framework by introducing a stochastic process for volatility as follows:

$$d\sigma_t^2 = \kappa(\theta - \sigma_t^2)dt + \xi\sigma_t dW_t^2 \tag{3}$$

Where:
σ_t is the volatility at time t.
κ is the mean-reversion rate, which controls how quickly volatility returns to a long-term average.
θ is the long-term mean level of variance.
ξ is the volatility of the volatility.
W_t^2 is a second Wiener process, separate from W_t for the asset price.The expensive This model provides more accurate pricing but lacks an analytical solution, requiring computationally expensive numerical methods like Monte Carlo simulations [4].