Third Year Review
Numerical Methods for Parametric Option Pricing

Supervisor: Linus Wunderlich

Ivelina Mladenova

Queen Mary University of London, March 2, 2026

Outline

This part

Motivation: What is an option worth?

Option Contract

Option contracts give the holder the **right but not obligation** to buy (or sell) an underlying asset at some point in the future at a predetermined price.

Motivation: What is an option worth?

Option Contract

Option contracts give the holder the **right but not obligation** to buy (or sell) an underlying asset at some point in the future at a predetermined price.

Notation

- S_t: stock price at time t; S_T: stock price at maturity T.
- K: strike price (the predetermined price).
- r: risk-free interest rate (continuously compounded).
- T: option maturity (in years); $\tau = T - t$ time to maturity.
- $(\cdot)^+ := \max(\cdot, 0)$; e.g., call payoff $(S_T - K)^+$.

Motivation: A (European) call option

Option Contract

Option contracts give the holder the **right but not obligation** to buy (or sell) an underlying asset at some point in the future at a predetermined price.

Notation

- S_t: stock price at time t; S_T: stock price at maturity T.
- K: strike price (the predetermined price).
- r: risk-free interest rate (continuously compounded).
- T: option maturity (in years); $\tau = T - t$ time to maturity.
- $(\cdot)^+ := \max(\cdot, 0)$; e.g., call payoff $(S_T - K)^+$.
- A (European) call option pays $(S_T - K)^+$ at maturity T.

The value of these contracts at expiration is easily determined, but what should we pay **today**?

Motivation: The Pricing Idea

A (European) call option pays $(S_T - K)^+ := \max(S_T - K, 0)$ at maturity T.

The Pricing idea: present value of a risk-neutral expectation

$$V(t, S_t) = e^{-r(T-t)}\,\mathbb{E}^{\mathbb{Q}}\big[(S_T - K)^+ \mid \mathcal{F}_t\big].$$

where $\mathbb{Q}$ is the risk-neutral measure. [1]

[1] We use the risk-neutral measure because under no-arbitrage, the discounted price process must be a martingale under $\mathbb{Q}$, ensuring the price today equals the expected discounted future payoff.

Motivation: The Pricing Idea

A (European) call option pays $(S_T - K)^+ := \max(S_T - K, 0)$ at maturity T.

The Pricing idea: present value of a risk-neutral expectation

$$V(t, S_t) = e^{-r(T-t)}\,\mathbb{E}^{\mathbb{Q}}\big[(S_T - K)^+ \mid \mathcal{F}_t\big].$$

where $\mathbb{Q}$ is the risk-neutral measure. [2]

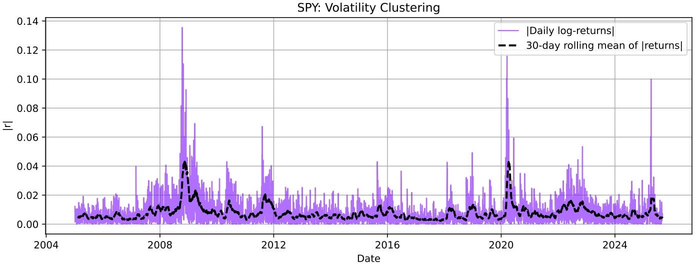
To compute it, we need two ingredients:

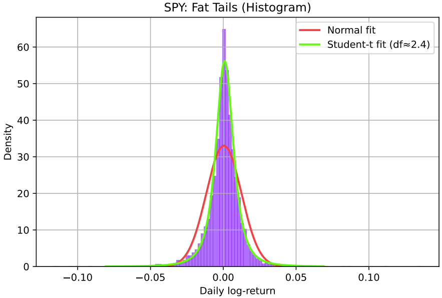
- A **model** for the stock price trajectory S_t, at some time t;
- A **method** to evaluate the expectation.

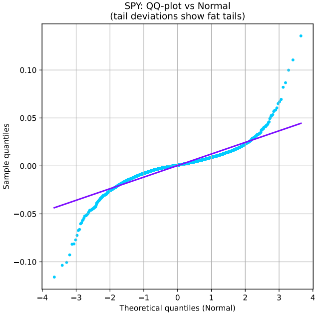
[2] We use the risk-neutral measure because under no-arbitrage, the discounted price process must be a martingale under $\mathbb{Q}$, ensuring the price today equals the expected discounted future payoff.

This part

Stylized Facts: clustering & fat tails (SPY, data from 2005-01-01)





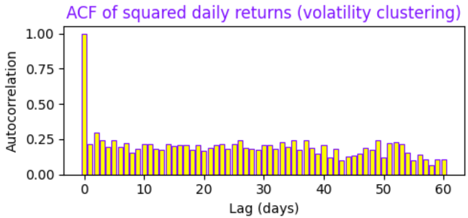


Stylized Facts: What GBM Misses

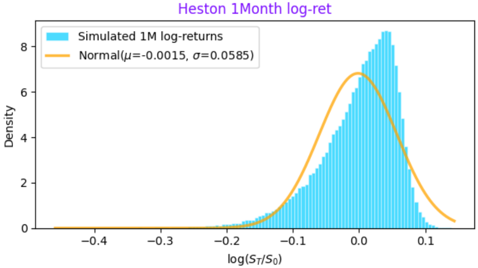
In 2001, Rama Cont released a paper on the empirical observations of stock prices, some of are:

- **Volatility is not constant**: it *clusters* over time.
- **Returns have fat tails** and negative skew.
- **Implied volatility smiles/smirks**: BS needs a different σ for each strike/maturity. See (Gatheral).
- And more, see (Cont).

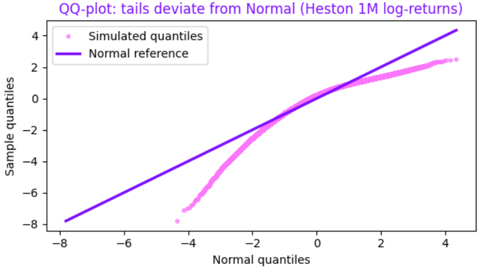
Heston Monte Carlo Simulations - fits stylized facts



ACF of squared returns (volatility clustering)



Histogram vs Normal



QQ-plot (fat tails)

This part

The Heston Model

Dynamics under $\mathbb{Q}$

$$
\begin{aligned}
dS_t &= rS_t\,dt + \sqrt{v_t}\,S_t\,dW_t^{(1)},\\
dv_t &= \kappa(\theta - v_t)\,dt + \sigma\sqrt{v_t}\,dW_t^{(2)},\\
\operatorname{corr}\big(dW_t^{(1)}, dW_t^{(2)}\big) &= \rho.
\end{aligned}
$$

- Mean-reverting variance $v_t \geq 0$ (CIR-type process).
- Captures **volatility clustering** and **implied-volatility smiles**.
- Still *affine* $\Rightarrow$ admits a closed-form *characteristic function* for $\log S_T$.
- See (Heston).

Heston Model: PDE

Domain/terminal: $(t, S, v) \in [0, T) \times (0, \infty) \times [0, \infty), \quad V(T, S, v) = (S - K)^+.$

Boundary conditions (European call)

- $S \to 0$: $V \approx 0$.
- $S \to \infty$: $V \approx S - Ke^{-r\tau}$ (or enforce $V_S \approx 1$).
- $v \to 0$: *degenerate* limit (diffusions drop):

$$\partial_t V + rSV_S + \kappa\theta\, V_v - rV = 0.$$

Parameters & positivity

$$\kappa > 0, \quad \theta > 0, \quad \sigma \geq 0, \quad \rho \in [-1, 1],$$
$$r \geq 0, \quad v_0 \geq 0.$$

Feller condition: $\quad 2\kappa\theta \geq \sigma^2$

No-arbitrage / martingale pricing: for $V(t, S, v) = \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)}\,\Phi(S_T) \mid S_t = S, v_t = v]$ (with payoff Φ), Itô + martingale condition $\Rightarrow$ *Feynman–Kac*

$$\partial_t V + \mathcal{L}V - rV = 0, \qquad V(T, S, v) = \Phi(S)$$

which gives (expanding $\mathcal{L}$) the familiar Heston PDE:

$$\boxed{\partial_t V + \tfrac{1}{2} vS^2 V_{SS} + \rho\sigma S\sqrt{v}\, V_{Sv} + \tfrac{1}{2}\sigma^2 v\, V_{vv} + rSV_S + \kappa(\theta - v)V_v - rV = 0}$$

This part

Section 1: Models & Motivation

- What is an option worth?
- Stylized Facts & Limitations of GBM
- The Heston Stochastic Volatility Model
- The Wishart (Affine) Stochastic Correlation Model

Section 2: Numerical Methods for Option Pricing

Section 3: Results of Numerical Methods

Section 4: Clear Structure of:

Wishart-driven Dynamics

Definition (Dynamics under $\mathbb{Q}$ (La Bua and Marazzina))

Let $S_t \in \mathbb{R}^d$. Then

$$dS_t = \mathrm{diag}(S_t)\Big(r\,\mathbf{1}\,dt + \sqrt{\Sigma_t}\,dW_t\Big),$$

$$d\Sigma_t = \Big(\beta\,Q^\top Q + M\Sigma_t + \Sigma_t M^\top\Big)dt + \sqrt{\Sigma_t}\,dB_t\,Q + Q^\top dB_t^\top\sqrt{\Sigma_t},$$

where $M, Q \in \mathbb{R}^{d\times d}$, $\beta > d-1$ *thm* (Bru), $\Sigma_t \in \mathbb{S}_d^+$, and $\sqrt{\Sigma_t}$ is the principal square root.

Brownian structure:

$dW_t \in \mathbb{R}^d, \quad dB_t \in \mathbb{R}^{d\times d}$ (indep. std. entries); $\quad dW_t = \sqrt{1-\|\rho\|^2}\,dZ_t + dB_t\,\rho,\ \rho \in \mathbb{R}^d,\ \|\rho\| < 1.$

$$d\langle W_i, B_{kj}\rangle_t = \rho_j\,\mathbf{1}_{\{i=k\}}\,dt, \qquad \text{all other cross-variations } = 0.$$

Initial conditions: $S_0 \in \mathbb{R}_+^d,\ \ \Sigma_0 \in \mathbb{S}_d^+$.

Wishart Characteristic Function

Joint CF of log-prices (WASC), (La Bua and Marazzina)

Let $y(t) = \log S(t) \in \mathbb{R}^d$, $\lambda \in \mathbb{C}^d$, and $\tau = T - t$. Then

$$\phi_y^{\mathrm{WASC}}(\lambda, \tau) = \mathbb{E}_t^{\mathbb{Q}}\big[e^{\,i\langle \lambda,\, y(T)\rangle}\big] = \exp\Big(i\langle \lambda, y(t)\rangle + \mathrm{Tr}\big(A_y(\tau;\lambda)\,\Sigma(t)\big) + b_y(\tau;\lambda)\Big),$$

with the deterministic matrix $A_y(\tau;\lambda)$ and the scalar function $b_y(\tau)$:

$$A_y(\tau;\lambda) = A_{22}(\tau)^{-1} A_{21}(\tau), \qquad b_y(\tau;\lambda) = -\tfrac{\beta}{2}\,\mathrm{Tr}\Big(\log A_{22}(\tau) + \tau\big(M + i\,Q^\top \rho\,\lambda^\top\big)\Big).$$

Block matrix exponential:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \exp\left(\tau \begin{pmatrix} M + iQ^\top \rho\,\lambda^\top & -2Q^\top Q \\ -\frac{1}{2}\big(\lambda\lambda^\top + i\,\mathrm{diag}(\lambda)\big) & -(M + iQ^\top \rho\,\lambda^\top)^\top \end{pmatrix}\right).$$

Basket payoff & Fourier transform

Put Payoff for d-assets in log-space

For $K > 0$ and $x = \log s \in \mathbb{R}^d$,

$$P(x) = \left(K - \sum_{i=1}^{d} w_i e^{x_i}\right)^+.$$

Fourier transform of basket put payoff, similar to (Hurd and Zhou)

Let $z = u + i\eta \in \mathbb{C}^d$ with $\eta_i > 0$ for all i and $\sum_{i=1}^d \eta_i > 1$. Then

$$\boxed{\widehat{P}(z) = \int_{\mathbb{R}^d} e^{-\langle \eta, x\rangle} P(x)\, e^{-i\langle u, x\rangle} dx = \left(\prod_{j=1}^{d} w_j^{iz_j}\right) K^{1-i\sum_{j=1}^d z_j} \frac{\prod_{j=1}^{d} \Gamma(-iz_j)}{\Gamma\left(2 - i\sum_{j=1}^{d} z_j\right)}.}$$

Pricing with Fourier

General pricing formula. For maturity $\tau = T - t$ and payoff Φ [3],

$$\Pi(t) \;=\; e^{-r\tau}\,\mathbb{E}_t[\Phi(S_T)] \;=\; \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \widehat{\Phi}_\eta(u+i\eta)\,\phi_y(u+i\eta,\tau)\,du,$$

where

- $\phi_y(u+i\eta,\tau)$ = joint characteristic function of $\log S_{t+\tau}$ (from previous slide),
- $\widehat{\Phi}_\eta(u+i\eta)$ = damped Fourier transform of the payoff (e.g. basket put).

Analyticity strip. Choose $\eta \in \mathbb{R}^d$ such that

$$\eta_i > 0 \quad \text{for all } i, \qquad \sum_{i=1}^{d} \eta_i > 1,$$

ensuring integrability of both $\widehat{\Phi}_\eta$ and $\phi_y(\cdot)$. Initial optimisation algo. from (Samet et al.).

[3] Basket *call on the sum* follows by put-call parity.

Wishart PDE formulation

Option value under $\mathbb{Q}$ (basket payoff Φ):

$$V(t, s, \Sigma) = \mathbb{E}^{\mathbb{Q}}\Big[e^{-r(T-t)}\,\Phi(S_T)\;\Big|\; S_t = s,\ \Sigma_t = \Sigma\Big]\,.$$

Feynman-Kac $\Rightarrow$ PDE:

$$\partial_t V + \mathcal{L}_{y,\Sigma} V - rV = 0, \qquad V(T, s, \Sigma) = \Phi(s),\ \text{log-prices } y = \log s.$$

Infinitesimal generator $\mathcal{L}_{y,\Sigma}$, (Da Fonseca et al.)

$$\mathcal{L}_{y,\Sigma} V = \mathrm{Tr}\big[(\beta Q^\top Q + M\Sigma + \Sigma M^\top)\,D + 2\,\Sigma D Q^\top Q D\big] + \tfrac{1}{2}\nabla_y^\top \Sigma \nabla_y + \nabla_y^\top\Big(r\mathbf{1} - \tfrac{1}{2}\,\mathrm{diag}(\Sigma)\Big).$$

$$+\,2\,\mathrm{Tr}\big[DQ^\top \rho\,\nabla_y \Sigma \nabla_y^\top\big].$$

where

$$D_{i,j} = \frac{\partial}{\partial \Sigma^{ij}}, \quad \nabla_y = \Big(\frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_d}\Big)^\top.$$

Terminal condition: $V(T, s, \Sigma) = \Phi(s)$ (e.g. basket payoff)

Outline

This part

Section 1: Models & Motivation

Section 2: Numerical Methods for Option Pricing

- Overview of Methods
- Neural Networks - the MLP architecture
- PDE Solvers - the Deep Galerkin Method

Section 3: Results of Numerical Methods

Section 4: Clear Structure of:

Numerical Methods for Pricing Options

Monte Carlo

- Simulate (S_t, v_t); discount payoffs; variance reduction for efficiency.

Fourier / Characteristic Function

- Price via risk-neutral characteristic function $\phi(u) = \mathbb{E}^{\mathbb{Q}}[e^{iu \log S_T}]$; invert to get option price (Carr and Madan).
- **COS method**: Fourier-cosine expansion from ϕ; high accuracy with few terms (Fang and Oosterlee).

Numerical Methods for Pricing Options

Monte Carlo

- Simulate (S_t, v_t); discount payoffs; variance reduction for efficiency.

Fourier / Characteristic Function

- Price via risk-neutral characteristic function $\phi(u) = \mathbb{E}^{\mathbb{Q}}[e^{iu \log S_T}]$; invert to get option price (Carr and Madan).
- **COS method**: Fourier-cosine expansion from ϕ; high accuracy with few terms (Fang and Oosterlee).

PDE solvers

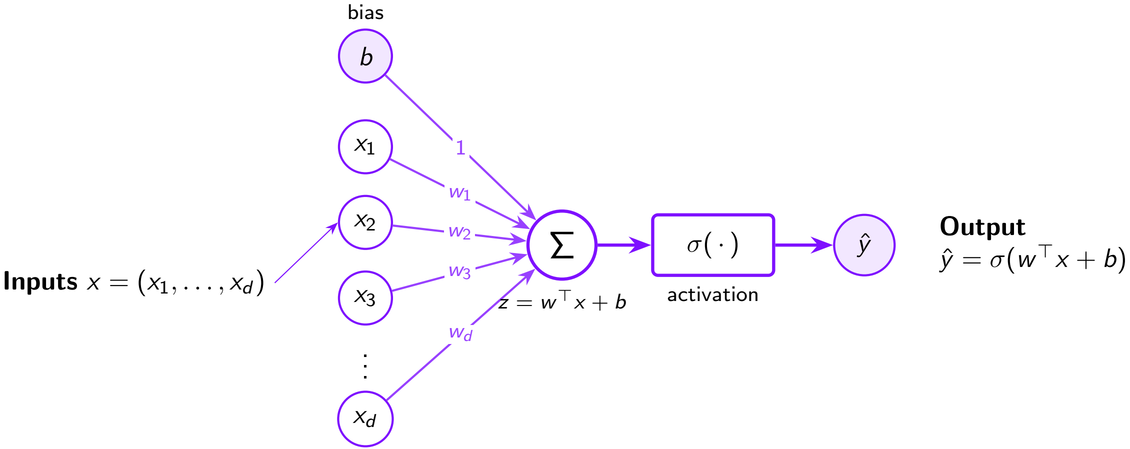
- Finite differences in (t, S, v); careful boundaries and stability.

Deep Learning (DGM)

- Learn $V_\theta(t, S, v)$ by minimizing PDE residual + BC/IC losses; mesh-free, scalable (Sirignano and Spiliopoulos).

This part

Perceptron: the single neuron



Notation: Neural Network parameters: $w = (w_1, \dots, w_d)$ weights, b bias,

Pre-activation $z = w^\top x + b$

σ activation function (e.g. ReLU/Tanh/Sigmoid), single neuron/node output $\hat{y} = \sigma(z)$.

MLP: parameters & supervised loss (quick refresher)

Model (L-layer MLP): map $x \in \mathbb{R}^d \mapsto \hat{y}_\theta$

$$
\begin{aligned}
h^{(0)} &= x, & h^{(\ell)} &= \sigma^{(\ell)}\big(z^{(\ell)}\big), \\
z^{(\ell)} &= W^{(\ell)} h^{(\ell-1)} + b^{(\ell)}, & \ell &= 1, \ldots, L, \\
\hat{y}_\theta &= h^{(L)}.
\end{aligned}
$$

Parameters: $\theta = \{W^{(\ell)}, b^{(\ell)}\}_{\ell=1}^{L}$. Weights: $W^{(\ell)} \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$, biases: $b^{(\ell)} \in \mathbb{R}^{n_\ell}$.

MLP: parameters & supervised loss (quick refresher)

Model (L-layer MLP): map $x \in \mathbb{R}^d \mapsto \hat{y}_\theta$

$$\begin{aligned} h^{(0)} &= x, & h^{(\ell)} &= \sigma^{(\ell)}\big(z^{(\ell)}\big), \\ z^{(\ell)} &= W^{(\ell)} h^{(\ell-1)} + b^{(\ell)}, & \ell &= 1, \ldots, L, \\ \hat{y}_\theta &= h^{(L)}. \end{aligned}$$

Parameters: $\theta = \{W^{(\ell)}, b^{(\ell)}\}_{\ell=1}^{L}$. Weights: $W^{(\ell)} \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$, biases: $b^{(\ell)} \in \mathbb{R}^{n_\ell}$.

Classic Supervised loss (labels available):

$$\mathcal{L}_{\text{sup}}(\theta) = \frac{1}{N} \sum_{i=1}^{N} \ell\big(y_i, \hat{y}_\theta(x_i)\big) \; + \; \lambda\, \Omega(\theta),$$

- Regression (e.g. fit prices): $\ell(y, \hat{y}) = \|y - \hat{y}\|_2^2$ (MSE).
- Regularization: $\Omega(\theta) = \sum_\ell \|W^{(\ell)}\|_F^2$ (weight decay).

Interpretation: We have target labels (x_i, y_i) and train the network to match them.

This part

From supervised to DGM loss

Goal: learn the call/put fair price $V_\theta(t, S, v)$ *without labels* by enforcing the Heston PDE and conditions.

Model: the Deep Galerkin Method (DGM) **has no labels** (Sirignano and Spiliopoulos). Instead it uses:

- **inputs** are time, states and the parameters $(t, S, v, \kappa, \theta, \sigma, \rho, r)$;
- **outputs** the approximated option price $V_\theta(t, S, v)$.

From supervised to DGM loss

Goal: learn the call/put price $V_\theta(t, S, v)$ *without labels* by enforcing the Heston PDE and conditions.

Model: the Deep Galerkin Method (DGM) **has no labels**, see (Sirignano and Spiliopoulos). Instead it uses:

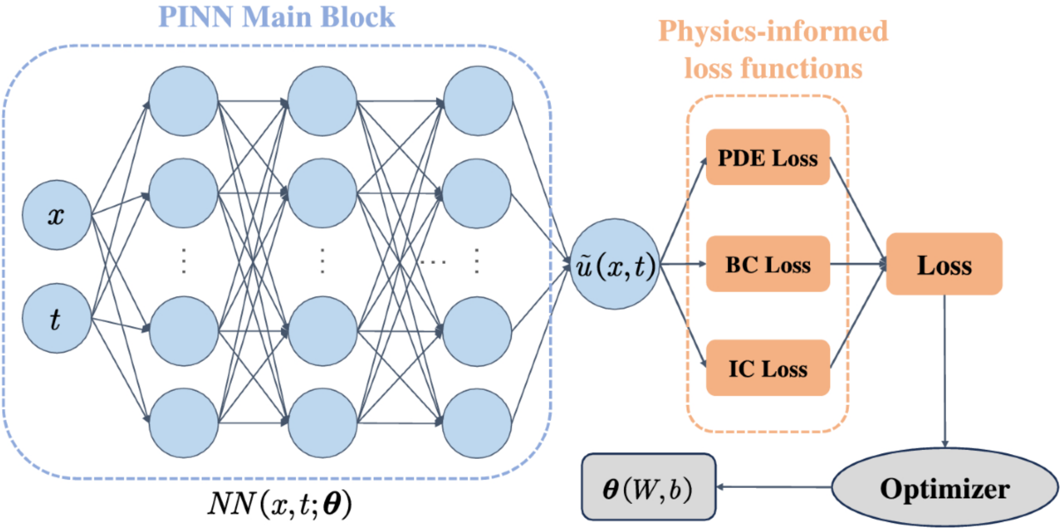
- **inputs** are time, states and the parameters $(t, S, v, \kappa, \theta, \sigma, \rho, r)$;
- **outputs** the approximated option price $V_\theta(t, S, v)$.

Composite ("physics informed") loss:

$$\mathcal{L}_{\mathrm{DGM}}(\theta) = \underbrace{\mathbb{E}_\Omega\big[\|\mathcal{N}[V_\theta]\|^2\big]}_{\text{PDE residual}} + \lambda_{\text{term}} \underbrace{\mathbb{E}_{\partial\Omega_T}\big[\|V_\theta(T, S, v) - (S-K)^+\|^2\big]}_{\text{terminal}} + \lambda_{\text{bdry}} \underbrace{\mathbb{E}_{\partial\Omega}\big[\|\mathcal{B}[V_\theta]\|^2\big]}_{\text{boundaries}},$$

where $\mathcal{N}$ is the Heston operator and $\mathcal{B}$ enforces BCs.

From supervised to DGM loss



Source: (Luo et al.)

Outline

This part

1a. Black Scholes-residual (with $\sigma(v) = \sqrt{v}$): new DGM loss

Decompose: $V = u_{\mathrm{BS}}(x,t,v) + U$, where for each fixed v, u_{BS} solves the Black Scholes PDE in $x = \log S$ with volatility $\sigma(v) = \sqrt{v}$ and terminal $(S-K)^+$.

Source term (apply Heston operator to u_{BS}):

$$\boxed{g(x,t,v) \;=\; \kappa(\theta - v)\, u_v \;+\; \tfrac{1}{2}\sigma^2\, v\, u_{vv} \;+\; \rho\,\sigma\, v\, u_{xv}}$$

(the x-drift/diffusion parts cancel since both have $\frac{1}{2} v\, \partial_{xx}$).

PDE for the correction U:

$$\boxed{\partial_t U + \mathcal{L}_H U - rU = g(x,t,v)}$$

$$\mathcal{L}_H = rS\,\partial_S + \kappa(\theta - v)\,\partial_v + \tfrac{1}{2} vS^2\,\partial_{SS} + \rho\sigma S\sqrt{v}\,\partial_{Sv} + \tfrac{1}{2}\sigma^2 v\,\partial_{vv}.$$

NEW DGM loss:

$$\boxed{\mathcal{L}(\theta) = \mathbb{E}_{\Omega}\big[\|\mathcal{N}_H[U_\theta] - g\|^2\big] + \lambda_T\,\mathbb{E}_{\partial\Omega_T}\big[\|U_\theta\|^2\big] + \lambda_B\,\mathbb{E}_{\partial\Omega}\big[\|\widetilde{\mathcal{B}}[U_\theta]\|^2\big]}$$

1b. Black Scholes-residual (with $\sigma(v) = \sqrt{v}$): Boundaries

Boundaries/terminal for U (European call):

- $t = T$: $U(T, S, v) = 0$ (since $u_{\mathrm{BS}}(T, S, v) = (S - K)^+$).
- $S \to 0$, $S \to \infty$: $U \to 0$ (BS matches call asymptotics).
- $v \to 0$: degenerate first-order boundary for U with source $g|_{v=0}$.

In practice, we don't sample exactly at $v = 0$ since $\sqrt{v}$ and $\partial_v \sqrt{v} = 1/(2\sqrt{v})$ are ill-conditioned at 0. So we clip $0 < v_{\mathrm{min}} \leq v$ to avoid the $\sigma(v) = \sqrt{v}$ singularity.

In fact, we truncate the domain to $v \in [v_{min}, v_{max}]$ with a tiny $v_{min} \sim 10^{-6}$.

This part

Section 1: Models & Motivation

Section 2: Numerical Methods for Option Pricing

Section 3: Results of Numerical Methods

- Neural Network approximation of the Heston Model
- Fourier pricing for Wishart-Put-Options
- Neural network surrogate for Fourier pricing for Wishart-Put-Options

Section 4: Clear Structure of:

Inputs: parameters $(y_t, \Sigma_t, M, Q, \rho, \beta, r, \tau, K, w)$

Algorithm (Fourier pricing with optimal damping)

1. **Characteristic function:** Compute $\log \phi_y(u;\tau)$ via affine Wishart formula
   $\log \phi_y(u;\tau) = i\langle u, y_t\rangle + \mathrm{tr}(A_\tau \Sigma_t) + b_\tau$.
2. **Payoff transform:** For basket put $\Phi(s) = (K - \sum_i w_i s_i)^+$, the damp FT: $\widehat{\Phi}_\eta(u + i\eta)$
3. **Damping parameters (penalised function):** Optimise over $\eta \in \mathbb{R}^2_+$ to stabilise integral:
   $\eta^* = \arg\min_{\eta\in\mathbb{R}^2_+} \left\{ \mathrm{Re}\left[\log \phi_y(u + i\eta;\tau) + \log \widehat{\Phi}(u + i\eta)\right] - \log\left(\min_{\tau\in[0,T]} \det A_{22}(\tau)\right)\right\}\Big|_{u=0}$.
   Note: if $\min_\tau \det A_{22}(\tau) \le 0$ return $+\infty$ (invalid contour).
4. **Fourier inversion:** Evaluate PUT $\Pi(t) = \frac{e^{-r\tau}}{(2\pi)^2}\int_{\mathbb{R}^2} \widehat{\Phi}(u + i\eta^*)\,\phi_y(u + i\eta^*, \tau)\,du$.
5. **Integration:** Compute integral numerically (quadrature, symmetry across quadrants).

Output: option price $\Pi(t)$.

This part

Section 1: Models & Motivation

Section 2: Numerical Methods for Option Pricing

Section 3: Results of Numerical Methods

- Neural Network approximation of the Heston Model
- Fourier pricing for Wishart-Put-Options
- Neural network surrogate for Fourier pricing for Wishart-Put-Options

Section 4: Clear Structure of:

3a. Preliminary results: NN for Fourier Wishart Pricing

Goal: Approximating the Fourier integrand $I(u; y)$ with a separable low-rank model.

Inputs: datasets $\mathcal{D}_{\text{int}} = \{(u^{(i)}, y^{(i)}, I^{(i)})\}$, $\mathcal{D}_{\text{price}} = \{(y^{(m)}, P_{\text{ref}}(y^{(m)}))\}$; K = **rank of decomposition**: number of separable factors; quadrature nodes $\{(u_j, w_j)\}$.

Neural Network Architecture

Build a bilinear low-rank surrogate for the integrand:

$$f(u) = \text{X-Net}(u) \in \mathbb{R}^K, \quad g(y) = \text{Y-Net}(y) \in \mathbb{R}^K,$$

$$\widehat{I}(u, y) = \langle f(u), g(y) \rangle = \sum_{k=1}^{K} f_k(u)\, g_k(y).$$

Price

Then the price integral becomes:

$$F_k \approx \frac{1}{(2\pi)^2} \sum_j w_j f_k(u_j), \quad \widehat{P}(y) = \sum_{k=1}^{K} g_k(y)\, F_k.$$

3b. Preliminary results: Training

Joint Losses:

$$\mathcal{L}_{\text{int}} = \mathrm{MSE}(\widehat{I}(u, y), I(u; y)), \quad \mathcal{L}_{\text{price}} = \mathrm{MSE}(\widehat{P}(y), P_{\text{ref}}(y)),$$
$$\mathcal{L}_{\text{total}} = (1 - \lambda)\, \mathcal{L}_{\text{int}} + \lambda\, \mathcal{L}_{\text{price}}.$$

- $\lambda \in [0, 1]$: trade-off parameter.
  - Small λ: focus on matching integrand samples (local fit).
  - Large λ: focus on matching option prices (global fit).

Optimisation setup

- Optimiser: Adam on NN parameters (θ_X, θ_Y).
- At each epoch:
  - Train on large integrand dataset $\mathcal{D}_{\text{int}}$.
  - Recompute F_k via Gauss-Legendre quadrature for price supervision.

Output: Surrogate balancing local accuracy (integrand) and global accuracy (price).

Summary

- **Classical pricers:** Fourier/COS, PDE/FD, MC all suffer the curse of dimensionality. Applying Fourier methods to a d =2-asset basket is feasible but scales poorly as d increases..
- **DGM / PINNs:** are great mesh-free PDE solvers trained on residual + terminal/boundary losses. But high accuracy is challenging to achieve for more complex high-dimensional problems.
- **BS-localized DGM:** learn a correction U to a BS localisation. More stable; in our runs achieves $\sim 10^{-2}$ relative error despite the large epoch number/training time; and still behind Fourier/COS for 2D Heston Europeans.

Takeaway: Combining *classical numerical methods* with *neural networks* (e.g. accelerating Fourier/COS pricers with NN surrogates achieves $\sim 10^{-1}$ relative error) is a promising direction.

Outline

Section 1: Models & Motivation

Section 2: Numerical Methods for Option Pricing

Section 3: Results of Numerical Methods

Section 4: Clear Structure of:
- Where I am now
- What's in progress
- What's left

This part

Section 1: Models & Motivation

Section 2: Numerical Methods for Option Pricing

Section 3: Results of Numerical Methods

Section 4: Clear Structure of:

- Where I am now
- What's in progress
- What's left

This part

Section 1: Models & Motivation

Section 2: Numerical Methods for Option Pricing

Section 3: Results of Numerical Methods

Section 4: Clear Structure of:

- Where I am now
- What's in progress
- What's left

This part

Section 1: Models & Motivation

Section 2: Numerical Methods for Option Pricing

Section 3: Results of Numerical Methods

Section 4: Clear Structure of:

- Where I am now
- What's in progress
- What's left

References I

- Black, Fischer, and Myron Scholes, "The Pricing of Options and Corporate Liabilities". *Journal of Political Economy*, vol. 81, no. 3, 1973, pp. 637–54.
- Bru, Marie-France, "Wishart processes". *Journal of Theoretical Probability*, vol. 4, no. 4, 1991, pp. 725–51.
- Carr, Peter, and Dilip Madan, "Option Valuation Using the Fast Fourier Transform". *Journal of Computational Finance*, vol. 2, no. 4, 1999, pp. 61–73.
- Cont, Rama, "Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues". *Quantitative Finance*, vol. 1, no. 2, 2001, pp. 223–36.
- Da Fonseca, José, et al., "Option pricing when correlations are stochastic: an analytical framework". *Review of Derivatives Research*, vol. 10, no. 2, 2007, pp. 151–80.
- Fang, Fang, and Cornelis W. Oosterlee, "A Novel Pricing Method for European Options Based on Fourier–Cosine Series Expansions". *SIAM Journal on Scientific Computing*, vol. 31, no. 2, 2008, pp. 826–48.

References II

- Gatheral, Jim, *The Volatility Surface: A Practitioner's Guide*. Wiley, 2006.
- Heston, Steven L., "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options". *Review of Financial Studies*, vol. 6, no. 2, 1993, pp. 327–43.
- Hull, John C., *Options, Futures, and Other Derivatives*. 10th ed., Pearson, 2018.
- Hurd, Thomas R, and Zhuowei Zhou, "A Fourier transform method for spread option pricing". *SIAM Journal on Financial Mathematics*, vol. 1, no. 1, 2010, pp. 142–57.
- La Bua, Gaetano, and Daniele Marazzina, "On the application of Wishart process to the pricing of equity derivatives: the multi-asset case". *Computational Management Science*, vol. 18, no. 2, 2021, pp. 149–76.
- Luo, Kuang, et al., "Physics-informed neural networks for PDE problems: a comprehensive review". *Artificial Intelligence Review*, vol. 58, no. 10, 2025, pp. 1–43.

References III

- Merton, Robert C., "Theory of Rational Option Pricing". *Bell Journal of Economics and Management Science*, vol. 4, no. 1, 1973, pp. 141–83.
- Samet, Michael, et al., "Optimal damping with hierarchical adaptive quadrature for efficient Fourier pricing of multi-asset options in Lévy models". 2022. *arXiv*, arxiv.org/abs/2203.08196.
- Sirignano, Justin, and Konstantinos Spiliopoulos, "DGM: A Deep Learning Algorithm for Solving Partial Differential Equations". *Journal of Computational Physics*, vol. 375, 2018, pp. 1339–64.

Geometric Brownian Motion (GBM)

Model

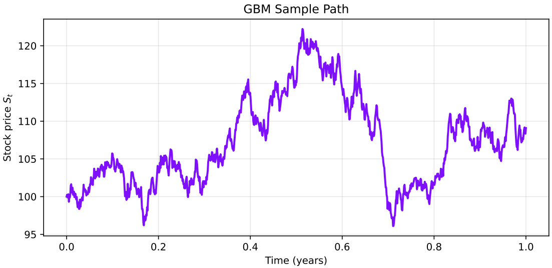
$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t, \quad S_0 > 0.$$

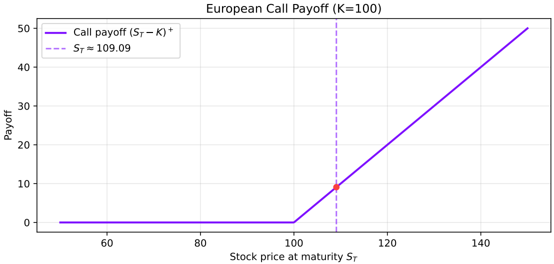
Under the risk-neutral measure $\mathbb{Q}$ (replace μ by r):

$$dS_t = rS_t \, dt + \sigma S_t \, dW_t^{\mathbb{Q}}.$$

- Closed-form solution: $S_t = S_0 \exp\left((r - \frac{1}{2}\sigma^2)t + \sigma W_t^{\mathbb{Q}}\right)$.
- Log-returns are normally distributed, volatility is *constant*.

Geometric Brownian Motion (GBM)





The Method: From GBM to the Black-Scholes PDE (idea only)

Two equivalent routes (sketch), (Hull):

1. **No-arbitrage / replication**: Hedge the option with $\Delta = \partial_S V$; eliminate risk; earn $r \Rightarrow$ PDE.
2. **Feynman-Kac**: Write price as discounted expectation under $\mathbb{Q}$; identify the governing PDE.

Black-Scholes PDE

$$\partial_t V + \frac{1}{2}\sigma^2 S^2 \, \partial_{SS} V + rS \, \partial_S V - rV = 0, \quad V(T, S) = (S - K)^+.$$

Leads to the **Black-Scholes formula** for European calls/puts, (Black and Scholes).

Extra: Heston Boundary Conditions

- As $S \to 0$: for calls, $V \to 0$.
- As $S \to \infty$: for calls, $V \sim S - Ke^{-r\tau}$ (linear growth).
- At $v \to 0$: degenerates to BS with $\sigma = 0$.
- At large v: growth control (e.g., truncation domain for $v \in (0, v_{max}]$).