Stochastic-Process Simulators

`generate_ar` — AR(1) Core of the Cascade

Parameters

Symbol	Meaning
ϕ	AR(1) coefficient (`b1` in the script)
σ^2	innovation variance (s_2^2)
x_0	initial value ($X_0^{(2)}$)
α (optional)	intercept if present in the code

1 · Model

$$x_t = \phi\, x_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}\big(0, \sigma^2\big).$$

If the script adds a constant term α, replace the equation with $x_t = \alpha + \phi x_{t-1} + \varepsilon_t$.

2 · Closed-Form Solution

Iterating the recursion:

$$x_t = \phi^{\,t}\, x_0 + \sum_{j=0}^{t-1} \phi^{\,j}\, \varepsilon_{\,t-j}.$$

3 · Mean

$$\mathbb{E}[x_t] = \phi^{\,t} x_0.$$

With an intercept α:

$$\mathbb{E}[x_t] = \phi^{\,t} x_0 + \frac{\alpha}{1-\phi}\big(1-\phi^{\,t}\big) \xrightarrow[t\to\infty]{} \mu = \frac{\alpha}{1-\phi}.$$

4 · Variance

$$\operatorname{Var}(x_t) = \sigma^2 \sum_{j=0}^{t-1} \phi^{\,2j} = \sigma^2\, \frac{1-\phi^{\,2t}}{1-\phi^2}.$$

Unconditional (stationary) variance

$$\boxed{\operatorname{Var}(x_\infty) = \frac{\sigma^2}{1-\phi^2}} \quad \text{for } |\phi| < 1.$$

5 · Autocovariance and Autocorrelation

In the stationary regime ($|\phi| < 1$):

$$\gamma(k) = \mathrm{Cov}(x_t, x_{t+k}) = \frac{\sigma^2}{1-\phi^2}\,\phi^{|k|}, \qquad \rho(k) = \phi^{|k|}.$$

6 · Mean-reversion half-life

Define $k_{1/2}$ so that

$$\Big|\mathbb{E}\big[x_{t+k_{1/2}}\big] - 0\Big| = \tfrac{1}{2}\Big|\mathbb{E}[x_t]\Big|,$$

i.e. $\phi^{k_{1/2}} = \dfrac{1}{2}$. Hence

$$\boxed{k_{1/2} \;=\; -\frac{\ln 2}{\ln|\phi|}}.$$

(Larger $|\phi| \Rightarrow$ slower decay.)

7 · Recap of One-Step Simulation

For each t

$$\begin{aligned} X_t^{(1)} &= b + \varepsilon_t^{(1)}, \\ X_t^{(2)} &= b_1\, X_t^{(1)} + \varepsilon_t^{(2)}, \\ X_t^{(3)} &= b_2\, X_t^{(2)} + \sqrt{|X_t^{(1)}|} + \varepsilon_t^{(3)}, \end{aligned} \qquad \varepsilon_t^{(k)} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, s_k^2).$$

If $|b_1| < 1$ and $|b_2| < 1$ the first two coordinates behave like a linear AR(1) cascade; the third adds a non-linear square-root coupling.

8 · Moments of the Linear Sub-System

8.1 Mean of $X_t^{(1)}$ Because $X_t^{(1)} = b + \varepsilon_t^{(1)}$ is i.i.d.,

$$\mathbb{E}[X_t^{(1)}] = b, \qquad \mathrm{Var}\big[X_t^{(1)}\big] = s_1^2.$$

8.2 Mean and Variance of $X_t^{(2)}$ Conditional on $X_t^{(1)}$, $X_t^{(2)} = b_1 X_t^{(1)} + \varepsilon_t^{(2)}$. Thus

$$\mathbb{E}[X_t^{(2)}] = b_1\, b, \qquad \mathrm{Var}[X_t^{(2)}] = b_1^2 s_1^2 + s_2^2.$$

8.3 Stationarity of $X_t^{(2)}$ Because there is *no* recursion in time, stationarity holds unconditionally—no AR lag exists beyond one step.

9 · Mean of the Non-Linear Third Coordinate

Write $X_t^{(3)} = b_2 X_t^{(2)} + \sqrt{|X_t^{(1)}|} + \varepsilon_t^{(3)}$.

$$\mathbb{E}\big[X_t^{(3)}\big] = b_2\,\mathbb{E}[X_t^{(2)}] + \mathbb{E}\big[\sqrt{|X_t^{(1)}|}\big],$$

because $\varepsilon_t^{(3)}$ has mean 0.

The square-root term has no closed form in elementary functions, but for $X_t^{(1)} \sim \mathcal{N}(b, s_1^2)$ one can use the folded-normal result

$$\mathbb{E}\big[\sqrt{|X_t^{(1)}|}\big] = \sqrt{\frac{s_1}{\sqrt{\pi}}}\,\exp\Big(-\frac{b^2}{2s_1^2}\Big) + b\,\mathrm{erf}\Big(\frac{b}{\sqrt{2}\,s_1}\Big),$$

where erf is the error function.

10 · Covariance Between $X_t^{(2)}$ and $X_t^{(3)}$

Since $X_t^{(2)}$ and $\varepsilon_t^{(3)}$ are independent,

$$\mathrm{Cov}(X_t^{(2)}, X_t^{(3)}) = b_2\ \mathrm{Var}(X_t^{(2)}) + \mathrm{Cov}\big(X_t^{(2)}, \sqrt{|X_t^{(1)}|}\big).$$

The second term is non-zero because both variables depend on $X_t^{(1)}$; it can be evaluated numerically or via series expansion.

11 · Key Take-Aways

- First two coordinates are *linear* in the innovations; third coordinate is non-linear and non-Gaussian.
- No temporal recursion $\Rightarrow$ simulation is $O(n)$ for n observations (all paths independent in time).
- Stationarity requires only $|b_1|, |b_2| < 1$ for the linear moments; the non-linear term always retains finite variance because $\sqrt{|X_t^{(1)}|}$ is square-integrable for Gaussian $X_t^{(1)}$.

`generate_garch` — GARCH(2,0) Conditional Variance

Parameters

Parameter	Role
α_0	long-run variance contribution
α_1, α_2	ARCH weights (positivity ≥ 0)
`n`	length retained after burn-in
`x0`	initial value (discarded)

1 · Model

$$\sigma_t^2 = \alpha_0 + \alpha_1\,x_{t-1}^2 + \alpha_2\,x_{t-2}^2,$$
$$x_t = \sigma_t\,\varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0,1).$$

(Burn in 50 steps before keeping the series.)

2 · Unconditional Variance

Let $v_t = \mathbb{E}[x_t^2] = \mathbb{E}[\sigma_t^2]$. Taking expectations:

$$v_t = \alpha_0 + (\alpha_1 + \alpha_2)\,v_{t-1}.$$

In stationarity ($v_t = v_{t-1} = \bar{v}$):

$$\boxed{\bar{v} = \frac{\alpha_0}{1 - (\alpha_1 + \alpha_2)}} \quad \text{if } \alpha_1 + \alpha_2 < 1.$$

3 · Heavy Tails, Fourth Moment and Kurtosis

The return is a *scale–mixture of normals* $x_t = \sigma_t \varepsilon_t$ with $\varepsilon_t \sim \mathcal{N}(0,1)$ i.i.d. Conditional on σ_t, x_t is Gaussian; unconditionally it exhibits excess kurtosis.

Fourth moment

$$\mathbb{E}[x_t^4] = \mathbb{E}\big[\sigma_t^4\,\varepsilon_t^4\big] = 3\,\mathbb{E}[\sigma_t^4], \qquad \text{because } \mathbb{E}[\varepsilon^4] = 3.$$

To obtain $\mathbb{E}[\sigma_t^4]$, square the variance recursion and take expectations (stationarity assumed):

$$\begin{aligned}\sigma_t^4 &= \Big(\alpha_0 + \alpha_1\,x_{t-1}^2 + \alpha_2\,x_{t-2}^2\Big)^2 \\ &= \alpha_0^2 + 2\alpha_0(\alpha_1+\alpha_2)\,x_{t-1}^2 + (\alpha_1^2 + 2\alpha_1\alpha_2 + \alpha_2^2)\,x_{t-1}^4 + 2\alpha_0\alpha_2\,x_{t-2}^2 - 2\alpha_2(\alpha_1+\alpha_2)\,x_{t-1}^2 x_{t-2}^2.\end{aligned}$$

Taking expectations and using symmetry $\mathbb{E}[x_{t-1}^2 x_{t-2}^2] = \mathbb{E}[x_t^2]^2$,

$$\mathbb{E}[\sigma_t^4] = \alpha_0^2 + 2\alpha_0(\alpha_1+\alpha_2)\,\bar{v} + (\alpha_1^2 + 2\alpha_1\alpha_2 + \alpha_2^2)\,\frac{\mathbb{E}[x_t^4]}{3},$$

where $\bar{v} = \mathbb{E}[x_t^2]$ from Subsec. 2. Solve for $\mathbb{E}[x_t^4]$:

$$\mathbb{E}[x_t^4] = \frac{3\big[\alpha_0^2 + 2\alpha_0(\alpha_1+\alpha_2)\bar{v}\big]}{1 - 3(\alpha_1^2+\alpha_2^2) - 2\alpha_1\alpha_2}, \quad \text{denominator} > 0.$$

Excess kurtosis

$$\boxed{\kappa = \frac{\mathbb{E}[x_t^4]}{\bar{v}^2} = 3\,\frac{\big[1-(\alpha_1+\alpha_2)\big]^2}{1 - 3(\alpha_1^2+\alpha_2^2) - 2\alpha_1\alpha_2}}, \qquad \kappa \downarrow 3 \text{ only when } \alpha_1 = \alpha_2 = 0.$$

As $\alpha_1 + \alpha_2 \to 1$ the denominator approaches zero $\implies$ $\kappa \to \infty$ (very heavy tails).

4 · Dynamics of Squared Returns

Starting from the variance recursion and replacing $x_{t-1}^2 = \sigma_{t-1}^2 \varepsilon_{t-1}^2$, one shows (after algebra and martingale–difference terms)

$$x_t^2 = \alpha_0 + (\alpha_1+\alpha_2)\,x_{t-1}^2 - \alpha_2\,x_{t-2}^2 + \underbrace{\big(\sigma_t^2\varepsilon_t^2 - \sigma_t^2\big)}_{\text{martingale noise}}.$$

Hence the *predictable part* of x_t^2 is an AR(2) with characteristic polynomial $1-(\alpha_1+\alpha_2)z+\alpha_2 z^2 = 0$. Its roots determine the shape of the ACF of x_t^2.

5 · Half-Life of Volatility Shocks

Consider the conditional variance forecast

$$\mathbb{E}[\sigma_{t+h}^2 \mid \mathcal{F}_t] = \alpha_0 + (\alpha_1+\alpha_2)\,\mathbb{E}[x_{t+h-1}^2 \mid \mathcal{F}_t].$$

Iterating the expectation (keep x_{t+h-k}^2 inside each step) gives

$$\mathbb{E}[\sigma_{t+h}^2 \mid \mathcal{F}_t] = \bar{v} + (\alpha_1+\alpha_2)^h\big(\sigma_t^2 - \bar{v}\big), \qquad \bar{v} = \frac{\alpha_0}{1-(\alpha_1+\alpha_2)}.$$

Define the *volatility half-life* $h_{1/2}$ by

$$\left|\mathbb{E}[\sigma^2_{t+h_{1/2}} \mid \mathcal{F}_t] - \bar{v}\right| = \tfrac{1}{2}\left|\sigma_t^2 - \bar{v}\right|.$$

Thus

$$(\alpha_1+\alpha_2)^{h_{1/2}} = \tfrac{1}{2} \implies \boxed{h_{1/2} = -\frac{\ln 2}{\ln(\alpha_1+\alpha_2)}}.$$

It counts how many steps it takes for a variance shock to decay by 50

6 · Positivity & Stability Checklist

1. **Strict positivity** $\alpha_0 > 0$ and $\alpha_1, \alpha_2 \geq 0 \implies \sigma_t^2 > 0$ a.s.
2. **Second-moment stationarity** $\alpha_1 + \alpha_2 < 1$ ensures $\mathbb{E}[x_t^2] < \infty$.
3. **Fourth-moment finiteness** Denominator in the kurtosis formula must be positive: $1 - 3(\alpha_1^2 + \alpha_2^2) - 2\alpha_1\alpha_2 > 0$.
4. **ACF decay of x_t^2** Both roots of $1 - (\alpha_1+\alpha_2)z + \alpha_2 z^2 = 0$ must lie inside the unit circle (automatically true when 1 and 2 hold).

7 · Key Take-Aways

- Captures volatility clustering and heavy tails with only three parameters.
- Stationarity hinge: $\alpha_1 + \alpha_2 < 1$; its proximity to 1 controls persistence and kurtosis.
- Volatility half-life $h = -\ln 2/\ln(\alpha_1+\alpha_2)$ provides an intuitive scale for how long shocks linger.

`generate_ou` — Seasonal Ornstein–Uhlenbeck

Parameters

Parameter	Role
θ	mean-reversion speed
μ	long-run mean
σ	volatility
A	seasonal amplitude
`freq` (f)	seasonal frequency ($1, 4, 12$ = annual, quarterly, monthly)
`dt` (Δt)	time step (default $1/252 \approx 1$ trading day)

1 · Recap of the Simulation Step

For each time increment Δt

$$x_{t+\Delta t} = x_t + \underbrace{\theta(\mu - x_t)\,\Delta t + \sigma\sqrt{\Delta t}\,Z_t}_{\text{“pure” OU core}} + \underbrace{A\sin\!\big(2\pi f\, t\Delta t + \phi\big)\,\Delta t}_{\text{deterministic seasonality}}, \qquad Z_t \sim \mathcal{N}(0,1).$$

If $A = 0$ we recover the classical Ornstein–Uhlenbeck SDE

$$dx_t = \theta(\mu - x_t)\,dt + \sigma\,dW_t,$$

where many texts write κ instead of θ.

2 · Detailed Closed-Form Solution of the Core OU

Integrating-factor derivation

1. Rewrite the SDE as $dx_t + \theta x_t\,dt = \theta\mu\,dt + \sigma\,dW_t$.
2. Multiply by $e^{\theta t}$ (the integrating factor):
$$e^{\theta t}dx_t + \theta e^{\theta t}x_t\,dt = \theta\mu\,e^{\theta t}dt + \sigma\,e^{\theta t}dW_t.$$
3. Recognise the left side as an exact differential: $d\big(e^{\theta t}x_t\big) = \theta\mu\,e^{\theta t}dt + \sigma\,e^{\theta t}dW_t$.
4. Integrate from 0 to t:
$$e^{\theta t}x_t - x_0 = \theta\mu\int_0^t e^{\theta s}ds + \sigma\int_0^t e^{\theta s}dW_s.$$
5. Evaluate the deterministic integral $\displaystyle\int_0^t e^{\theta s}ds = (e^{\theta t}-1)/\theta$.

   Putting it together
$$e^{\theta t}x_t = x_0 + \mu(e^{\theta t}-1) + \sigma\int_0^t e^{\theta s}dW_s, \quad\Longrightarrow\quad \boxed{x_t = \mu + (x_0-\mu)e^{-\theta t} + \sigma\int_0^t e^{-\theta(t-s)}\,dW_s}.$$

3 · Expectation and Variance

Expectation

$$\mathbb{E}[x_t] = \mu + (x_0-\mu)e^{-\theta t},$$

because the stochastic integral has mean 0.

Variance (Itô isometry)

$$\begin{aligned}\mathbb{V}[x_t] &= \sigma^2\,\mathbb{E}\Big[\Big(\int_0^t e^{-\theta(t-s)}dW_s\Big)^2\Big]\\ &= \sigma^2\int_0^t e^{-2\theta(t-s)}\,ds\\ &= \frac{\sigma^2}{2\theta}\big(1-e^{-2\theta t}\big).\end{aligned}$$

Stationary limits

$$x_\infty \sim \mathcal{N}\Big(\mu,\ \sigma^2/(2\theta)\Big).$$

4 · Autocovariance and Autocorrelation

For $0 \le s \le t$:

$$\begin{aligned}\mathrm{Cov}(x_s,x_t) &= \sigma^2\int_0^s e^{-\theta(s-u)}e^{-\theta(t-u)}\,du\\ &= \frac{\sigma^2}{2\theta}\,e^{-\theta(t-s)}\big(1-e^{-2\theta s}\big).\end{aligned}$$

In the *stationary* regime ($s\to\infty$):

$$\boxed{\mathrm{Cov}(x_0,x_\tau) = \frac{\sigma^2}{2\theta}e^{-\theta\tau}, \qquad \rho(\tau) = e^{-\theta\tau}.}$$

5 · Mean-Reversion Half-Life

Define $t_{1/2}$ such that $|\mathbb{E}[x_{t_{1/2}}] - \mu| = \frac{1}{2}|x_0 - \mu|$.

Since $\mathbb{E}[x_t] - \mu = (x_0 - \mu)e^{-\theta t}$, we have $e^{-\theta t_{1/2}} = \frac{1}{2} \implies \boxed{t_{1/2} = \ln 2/\theta}$.

Large $\theta \Rightarrow$ short half-life (fast mean-reversion); small $\theta \Rightarrow$ sluggish adjustment.

6 · Effect of the Seasonal Term

With $A \neq 0$ the mean becomes

$$m(t) = \mu + \frac{A}{2\pi f}\big[1 - \cos(2\pi f t + \phi)\big],$$

so the raw series is non-stationary in mean. If we *detrend* via $\tilde{x}_t := x_t^{\text{seasonal}} - m(t)$, then $\tilde{x}_t$ obeys all stationary results in §§2–5.

7 · Key Take-Aways

- The OU core has constant mean μ, variance $\sigma^2/(2\theta)$, exponential ACF $e^{-\theta\tau}$.
- Exact half-life: $t_{1/2} = \ln 2/\theta$.
- Seasonality adds deterministic non-stationarity; subtract $m(t)$ to isolate the stationary core.
- Euler–Maruyama is accurate for daily `dt`; shrink `dt` or use higher-order schemes if you need more precision.

`generate_bs` — Geometric Brownian Motion (Black–Scholes)

Parameters

Parameter	Role
r	risk-free rate / drift ($\mu = r$ under Q)
σ	volatility (annualised)
`x0`	initial asset level S_0

1 · Continuous-Time Model

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

1.1 Exact Solution (Itô)

1. Divide by S_t: $\dfrac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t$.
2. Integrate $\int_0^t$: $\ln \dfrac{S_t}{S_0} = \big(\mu - \frac{1}{2}\sigma^2\big)t + \sigma W_t$.
3. Exponentiate:

$$\boxed{S_t = S_0\, \exp\Big[\big(\mu - \tfrac{1}{2}\sigma^2\big)t + \sigma W_t\Big]}.$$

1.2 Moments

Because $W_t \sim \mathcal{N}(0,t)$,

$$\begin{aligned} \mathbb{E}[S_t] &= S_0\, e^{\mu t}, \\ \operatorname{Var}[S_t] &= S_0^2\, e^{2\mu t}\big(e^{\sigma^2 t} - 1\big), \\ \ln S_t &\sim \mathcal{N}\big(\ln S_0 + (\mu - \tfrac{1}{2}\sigma^2)t,\ \sigma^2 t\big). \end{aligned}$$

1.3 Key Observations

- **No mean-reversion** — drift scales with level $\Rightarrow$ exponential growth/decay.
- **Constant volatility** — cannot capture clustering or smiles.
- **Analytic tractability** — closed-form option prices via Black–Scholes.
- **Empirical limitation** — fails to match skew & excess kurtosis.

2 · Discrete-Time Simulation (Exact Step)

For step length Δt (e.g. 1/252):

$$\boxed{S_{t+\Delta t} = S_t\ \exp\Big[\big(r - \tfrac{1}{2}\sigma^2\big)\Delta t + \sigma\sqrt{\Delta t}\, Z_t\Big]}, \qquad Z_t \sim \mathcal{N}(0,1).$$

This update reproduces the *exact* log-normal law; no discretisation bias.

2.1 One-Step Conditional Moments

$$\begin{aligned} \mathbb{E}[S_{t+\Delta t} \mid S_t] &= S_t\, e^{r\Delta t}, \\ \operatorname{Var}[S_{t+\Delta t} \mid S_t] &= S_t^2\, e^{2r\Delta t}\big(e^{\sigma^2 \Delta t} - 1\big). \end{aligned}$$

Over n steps ($t = n\Delta t$) these accumulate to the continuous-time expressions.

2.2 Practical Caveats

- **Step size** — larger Δt inflates Monte-Carlo error but *does not* bias the law.
- **Real vs. risk-neutral drift** — set $\mu \neq r$ to obtain real-world (historical) paths.
- **Path-dependent options** — barrier/lookback pay-offs may need sub-day steps for accuracy.

3 · Deeper Insights

3.1 Martingale Property Under Q ($\mu = r$) the discounted process $\tilde{S}_t = e^{-rt} S_t$ satisfies $\mathbb{E}^Q[\tilde{S}_t \mid \mathcal{F}_s] = \tilde{S}_s$ —the fundamental no-arbitrage requirement.

3.2 Log-Return Stationarity

$$\ln \frac{S_{t+\tau}}{S_t} \sim \mathcal{N}\big((\mu - \tfrac{1}{2}\sigma^2)\tau,\ \sigma^2 \tau\big),$$

independent across non-overlapping intervals. Scaling: daily $\sigma_d \Rightarrow$ yearly $\sigma_{yr} = \sigma_d\sqrt{252}$.

3.3 Moment-Generating Function

$$M_S(u) = \mathbb{E}\big[e^{u \ln S_t}\big] = \exp\Big(u \ln S_0 + u(\mu - \tfrac{1}{2}\sigma^2)t + \tfrac{1}{2}u^2\sigma^2 t\Big),$$

hence all moments $\mathbb{E}[S_t^k]$ are finite ($k \in \mathbb{R}$).

4 · Summary

- **Continuous GBM** — solved analytically; moments closed-form.
- **Discrete implementation** — exact exponential step preserves the continuous-time distribution.
- **Use cases** — baseline for option pricing; benchmark for testing stochastic-volatility extensions.

`generate_heston_vol` — CIR Variance Process

Parameters

Parameter	Role
κ	speed of mean-reversion
θ	long-run (target) variance
ξ	volatility of variance ("vol of vol")
`dt` (Δt)	simulation time step (e.g. 1/252)

1 · Continuous-Time Model

$$\boxed{dv_t = \kappa(\theta - v_t)\,dt + \xi\sqrt{v_t}\,dW_t, \qquad v_0 \geq 0}.$$

1.1 Closed-Form Moments

- **Expectation** $\mathbb{E}[v_t] = \theta + (v_0 - \theta)\,e^{-\kappa t}$.
- **Variance**
$$\mathrm{Var}[v_t] = \frac{\xi^2\theta}{2\kappa}\big(1 - e^{-\kappa t}\big)^2 + (v_0 - \theta)\,\frac{\xi^2}{\kappa}\,e^{-\kappa t}\big(1 - e^{-\kappa t}\big).$$
- **Stationary distribution** ($t \to \infty$)
$$v_\infty \sim \Gamma\Big(\tfrac{2\kappa\theta}{\xi^2},\ \tfrac{\xi^2}{2\kappa}\Big), \qquad \mathrm{Var}[v_\infty] = \tfrac{\xi^2\theta}{2\kappa}.$$

1.2 Autocovariance, ACF and Half-Life

$$\mathrm{Cov}(v_s, v_{s+\tau}) = \frac{\xi^2\theta}{2\kappa}\,e^{-\kappa\tau}, \qquad \rho(\tau) = e^{-\kappa\tau}.$$

Half-life of a deviation from θ:

$$e^{-\kappa t_{1/2}} = \tfrac{1}{2} \implies \boxed{t_{1/2} = \frac{\ln 2}{\kappa}.}$$

1.3 Transition Density

$$v_t \sim \frac{\xi^2\,(1-e^{-\kappa t})}{4\kappa}\,\chi^2\Big(d,\ \lambda\Big), \quad d = \tfrac{4\kappa\theta}{\xi^2},\ \lambda = \tfrac{4\kappa\,e^{-\kappa t} v_0}{\xi^2(1-e^{-\kappa t})}.$$

2 · Discrete-Time Simulation (Euler–Maruyama)

$$\boxed{v_{t+\Delta t} = v_t + \kappa(\theta - v_t)\,\Delta t + \xi\sqrt{v_t\,\Delta t}\,Z_t, \qquad Z_t \sim \mathcal{N}(0,1)}$$

2.1 Interpretation

- Drift term $\kappa(\theta - v_t)\Delta t$ pulls v_t toward θ.
- Diffusion term $\xi\sqrt{v_t\,\Delta t}\,Z_t$ mirrors the continuous-time volatility $\xi\sqrt{v_t}$.
- For small Δt, Euler–Maruyama is a good approximation; for larger steps bias and negative draws can appear.

2.2 Practical Caveats

1. **Positivity & Feller condition** $2\kappa\theta \geq \xi^2$ guarantees $v_t > 0$ in continuous time; the discrete scheme does not enforce it.
2. **Alternative schemes** Full-truncation Euler, Quadratic-Exponential (QE), or exact non-central χ^2 sampling remove bias/negativity.
3. **Parameter estimation** MLE via the transition density or method of moments using $\mathbb{E}[v_t]$, $\mathrm{Var}[v_t]$.
4. **Standalone vs. sub-routine** Used alone for interest-rate/volatility modelling, or as the variance driver inside the full Heston model.

3 · Advanced Insights

- Stationary Gamma law implies skewed, heavy-tailed variance.
- Leverage in the Heston price model arises when this v_t is correlated with the asset Brownian motion.
- Half-life $\ln 2/\kappa$ equals that of an OU process — exponential mean-reversion of the first moment.
- **Calibration tip:** rapid reversion in implied vols $\Rightarrow$ choose large κ relative to ξ.

`generate_heston_2` — Alternative Heston Implementation

- Uses a bivariate normal draw each step: $(Z_{1t}, Z_{2t}) \sim \mathcal{N}\Big(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\Big)$.
- Updates S directly (not $\ln S$) but multiplies by exp drift; algebraically equivalent for small Δt.
- Demonstrates two common simulation styles for correlated diffusions.

---

generate_sine — Pure Sinusoid with Random Frequency & Phase

For each feature $k = 1, \ldots, d$:

$$x_j^{(k)} = \sin\bigl(\text{freq} \cdot j + \text{phase}\bigr), \qquad \text{freq} \sim U(0, 0.1),\ \text{phase} \sim U(0, 2\pi).$$

- Afterwards rescale to $[0, 1]$ via $(\text{data} + 1)/2$.
- Deterministic except for randomised frequency/phase — baseline for seasonality or spectral methods.
- Add noise if stochastic periodic data are desired.

---

generate_ar_multi — Gaussian Vector AR(1)

Parameters

Parameter	Role
ϕ	scalar AR coefficient (common to all components)
σ	controls cross-correlation of the noise
`d`	dimensionality
`n`	length

Equations

$$\begin{aligned} \mathbf{x}_t &= \phi\, \mathbf{x}_{t-1} + \varepsilon_t, \\ \varepsilon_t &\sim \mathcal{N}\bigl(\mathbf{0}, \Sigma\bigr), \\ \Sigma &= \sigma\,(\mathbf{1}\mathbf{1}^\top) + (1 - \sigma)\, I_d. \end{aligned}$$

- Stationary if $|\phi| < 1$; unconditional covariance $\text{Var}(\mathbf{x}_t) = \Sigma/(1 - \phi^2)$.
- Noise structure: off-diagonal entries all equal σ, letting you dial from *independent* ($\sigma = 0$) to *perfectly collinear* ($\sigma \to 1$).
- **Use-cases:** multivariate forecasting, PCA on correlated AR data.

General Tips & Caveats

1. **Time step (`dt`) vs. accuracy** — GBM and Heston use the exact exponential solution for prices but Euler for variance; keep `dt` small (daily) or switch to higher-order schemes (Milstein, QE).
2. **Parameter sampling** — all parameters are drawn *independently uniform*; real-world processes often show parameter correlations (e.g. κ and θ in interest-rate models).

3. **Burn-in** — the GARCH routine discards the first 50 observations; consider longer burn-in for slow mean reversion.

4. **Positivity of variance** — CIR/Heston paths can hit zero if the Feller condition fails; restrict parameter ranges or use an exact CIR sampler if zeros are undesirable.

5. **Reproducibility** — all functions rely on NumPy's global RNG; set `np.random.seed()` before calling for deterministic results.