A Taxonomy of Generalized Diffusion Processes for Generative Modeling

Overview of Hierarchical Classification

We organize diffusion processes into four top-level categories:

1. **A. Continuous-State Diffusions:**
   (a) A1. Continuous Gaussian Diffusions (OU, Langevin)
   (b) A2. Jump and Lévy-Driven Diffusions (Lévy-OU, NIG-OU)
   (c) A3. Memory-Driven Diffusions (Fractional Brownian Motion)
2. **B. Constrained-State Diffusions:**
   (a) B1. Bounded-Domain Diffusions (Jacobi, Dirichlet)
   (b) B2. Manifold-Valued Diffusions (Sphere, Lie Groups)
3. **C. Structured-State Diffusions:**
   (a) C1. Matrix-Valued Diffusions (Wishart)
   (b) C2. Interacting/Measure-Valued Diffusions (McKean–Vlasov, Fleming–Viot)
4. **D. Hybrid Processes:**
   (a) D1. Stochastic Volatility and PDMP (Heston, Piecewise-Deterministic MP)
   (b) D2. Discrete-State Processes (Birth–Death)

Contents

1 Unified Principles of Diffusion Processes **2**

2 A. Continuous-State Diffusions **3**
2.1 A1. Continuous Gaussian Diffusions . . . 3
2.1.1 Ornstein–Uhlenbeck (OU) Process . . . 3
2.1.2 Overdamped Langevin Diffusion . . . 4
2.2 A2. Jump and Lévy-Driven Diffusions . . . 4
2.2.1 Lévy-Driven OU . . . 4
2.2.2 Normal–Inverse Gaussian OU . . . 4
2.3 A3. Memory-Driven Diffusions . . . 5
2.3.1 Fractional Brownian Motion . . . 5

3 B. Constrained-State Diffusions **5**
3.1 B1. Bounded-Domain Diffusions . . . 5
3.1.1 Jacobi Process . . . 5
3.1.2 Dirichlet Diffusion . . . 6
3.2 B2. Manifold-Valued Diffusions . . . 6
3.2.1 Brownian Motion on the Sphere . . . 6
3.2.2 Diffusions on Lie Groups . . . 6

4 C. Structured-State Diffusions **7**
4.1 C1. Matrix-Valued Diffusions 7
4.1.1 Wishart Process 7
4.2 C2. Interacting/Measure-Valued Diffusions 8
4.2.1 McKean–Vlasov Diffusion 8
4.2.2 Fleming–Viot Process 8

5 D. Hybrid Processes **8**
5.1 D1. Stochastic Volatility & PDMP 9
5.1.1 Heston Model 9
5.1.2 Piecewise-Deterministic Markov Process (PDMP) 9
5.2 D2. Discrete-State Processes 10
5.2.1 Birth–Death Process 10

1 Unified Principles of Diffusion Processes

Before diving into the four major classes (A–D), it is helpful to step back and see the common thread that runs through every diffusion-based generative model:

1. Forward Corruption & Stationary Target Each model defines a *forward* process that progressively perturbs or "corrupts" the data distribution toward a tractable stationary law:

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t \quad \Longrightarrow \quad X_t \xrightarrow[t\to\infty]{} q_\infty,$$

where q_∞ is Gaussian (A1), jump-driven (A2), bounded (B1), manifold (B2), matrix (C1), measure (C2), hybrid (D1), or discrete (D2).

2. Master/Fokker–Planck Equation The density $q(x,t)$ evolves according to a (possibly nonlocal) PDE:

$$\partial_t q = -\nabla\cdot(b\,q) + \tfrac{1}{2}\nabla\cdot\big(a\,\nabla q\big) + \int [q(x+z) - q(x)]\,\nu(\mathrm{d}z) - (\text{boundary/jump terms}),$$

which unifies continuous, jump, and boundary-constrained dynamics.

3. Time Reversal & Score-Based Denoising By reversing time in the forward process (Haussmann–Pardoux and its jump/boundary analogues), one obtains a *reverse SDE* or Markov kernel of the form

$$\mathrm{d}X_t = \big[b(X_t) - a\,\nabla_x \ln q(X_t,t)\big]\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}\bar{W}_t + \dots$$

which in practice is implemented via a *score network* estimating $\nabla \ln q$.

4. Classification by Noise & State Structure

- **A. Continuous-State:** • A1: Gaussian noise only. • A2: Gaussian drift + Lévy jumps. • A3: Fractional kernels (memory).
- **B. Constrained-State:** • B1: Bounded domains (intervals, simplices). • B2: Manifolds (spheres, Lie groups).
- **C. Structured-State:** • C1: Matrices (Wishart). • C2: Measures/interacting particles (McKean–Vlasov, Fleming–Viot).
- **D. Hybrid:** • D1: Diffusion + deterministic flows/jumps (Heston, PDMP). • D2: Purely discrete jumps (Birth–Death).

5. Bottom-Line Logic Regardless of the noise type or state space, the *core* of diffusion-based generative modeling is:

1. *Define* a forward corruption process with known or tractable equilibrium q_∞.
2. *Compute* or approximate the score $\nabla_x \ln q(x,t)$ along this process.
3. *Simulate* the learned reverse dynamics to transform samples of q_∞ back to data-like samples.

This unified framework underlies all sections A–D, differing only in the specifics of b, σ, jump terms, and state constraints."'

2 A. Continuous-State Diffusions

General Form and Common Properties Continuous-state diffusions on $\mathbb{R}^d$ can be written in the unified SDE form

$$\mathrm{d}X_t = b(X_t,t)\,\mathrm{d}t + \sigma(X_t,t)\,\mathrm{d}W_t + \int_{\mathcal{Z}} Z(X_{t^-},z)\,\tilde{N}(\mathrm{d}t,\mathrm{d}z),$$

where:

- $b(x,t)$ is the drift vector field,
- $\sigma(x,t)$ the diffusion matrix (Gaussian noise),
- W_t a Brownian motion,
- $\tilde{N}$ a compensated Poisson measure with jump amplitude $Z(x,z)$.

The Fokker–Planck (or Master) equation governs the density $q(x,t)$, and if reversible, the process satisfies detailed balance w.r.t. its stationary law $q_\infty(x)$.

2.1 A1. Continuous Gaussian Diffusions

2.1.1 Ornstein–Uhlenbeck (OU) Process

Intuition A linear restoring force $-\lambda X_t$ plus isotropic Gaussian noise yields exponential "forgetting" of initial data.

Forward Dynamics

$$\mathrm{d}X_t = -\lambda X_t\,\mathrm{d}t + \sqrt{2\lambda}\,\mathrm{d}W_t, \quad X_0 \sim p_{\mathrm{data}}.$$

Its transition kernel is

$$q(x,t \mid x_0,0) = \mathcal{N}\big(e^{-\lambda t}x_0,\,(1-e^{-2\lambda t})I\big).$$

Reverse Dynamics

$$\mathrm{d}X_t = \Big[-\lambda X_t + 2\lambda\,\nabla_x \ln q(X_t,t)\Big]\mathrm{d}t + \sqrt{2\lambda}\,\mathrm{d}\bar{W}_t.$$

Proof of the Reverse SDE. The forward Fokker–Planck equation is

$$\partial_t q = -\nabla\cdot(b\,q) + \tfrac{1}{2}\nabla\cdot\big(a\,\nabla q\big), \quad b(x) = -\lambda x,\ a = 2\lambda I.$$

By Haussmann–Pardoux time-reversal, the backward drift is

$$b_{\mathrm{rev}}(x,t) = b(x) - a\,\nabla_x \ln q(x,t) = -\lambda x - 2\lambda\,\nabla_x \ln q(x,t).$$

Rewriting in forward time yields the stated reverse-time SDE. □

Pros/Cons

Pros: Closed-form forward/backward; stable; exact score.

Cons: Gaussian only; unimodal; unbounded support.

2.1.2 Overdamped Langevin Diffusion

Intuition Samples from a target density $p(x) \propto e^{-E(x)}$ by following gradient ascent on $\ln p(x)$ plus thermal noise.

Forward and Reverse Dynamics

$$\mathrm{d}X_t = \nabla \ln p(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t.$$

This SDE is self-adjoint under $p(x)$, hence reversible.

Proof of Reversibility. The generator is

$$\mathcal{L}f = \langle \nabla \ln p, \nabla f \rangle + \Delta f.$$

Detailed balance requires $p(x)\mathcal{L} = \mathcal{L}^*(p\,\cdot)$, which holds because

$$\int g\,\mathcal{L}f\,p = \int f\,\mathcal{L}g\,p$$

for smooth f, g, making the forward and backward SDE identical. □

Pros/Cons

Pros: Flexible target densities; direct sampling mechanism.

Cons: Potentially slow mixing; no built-in noise schedule; requires small integration steps.

2.2 A2. Jump and Lévy-Driven Diffusions

2.2.1 Lévy-Driven OU

Intuition Adds jump noise to capture heavy-tailed and discontinuous behavior.

Forward Dynamics

$$\mathrm{d}X_t = -\lambda X_t\,\mathrm{d}t + \mathrm{d}L_t, \quad \mathbb{E}[e^{iuL_t}] = e^{-t\Psi(u)}.$$

Master Equation

$$\partial_t q(x,t) = \lambda \nabla\cdot(xq) + \int_{\mathbb{R}^d} \big[q(x+z,t) - q(x,t)\big]\,\nu(\mathrm{d}z).$$

Reverse Dynamics

$$\mathrm{d}X_t = \big[-\lambda X_t + \nabla_x \ln q(X_t,t)\big]\mathrm{d}t + \mathrm{d}\bar{L}_t.$$

Pros/Cons

Pros: Captures jumps, heavy tails, multimodality.

Cons: Score intractable; requires characteristic-function methods; sampling is costly.

2.2.2 Normal–Inverse Gaussian OU

A special case of Lévy-OU driven by NIG noise, yielding skewed heavy-tailed equilibrium; reverse formulation parallels the general case.

2.3 A3. Memory-Driven Diffusions

2.3.1 Fractional Brownian Motion

Intuition Imparts long-range dependence via the Hurst parameter $H \neq \frac{1}{2}$, modeling persistent or anti-persistent behavior.

Representation

$$X_t = \int_0^t K_H(t,s)\,\mathrm{d}W_s,$$

with kernel K_H giving covariance $\mathbb{E}[X_t X_s] \propto |t-s|^{2H}$. Non-Markovian: reverse requires the entire trajectory.

Pros/Cons

Pros: Captures realistic temporal correlations and roughness.

Cons: No local SDE; inference must handle full path history.

3 B. Constrained-State Diffusions

General Form and Common Properties Constrained-state diffusions evolve on a domain $D \subseteq \mathbb{R}^d$ (e.g. an interval, simplex, or manifold) and must respect boundary or manifold constraints. In general they satisfy

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t + \mathrm{d}K_t,$$

where:

- $b(x)$ is a drift field tangent to D,
- $\sigma(x)$ is a diffusion coefficient vanishing or projecting at the boundary,
- K_t is a reflection or local-time term enforcing $X_t \in D$.

The density $q(x,t)$ solves a Fokker–Planck equation with zero probability flux at ∂D. Reversibility (if it holds) imposes detailed-balance conditions under the stationary law $q_\infty(x)$.

3.1 B1. Bounded-Domain Diffusions

3.1.1 Jacobi Process

Intuition Generates Beta(α,β)-distributed samples on $(0,1)$. The noise amplitude $\sqrt{X(1-X)}$ vanishes at the boundaries, preventing exit.

Forward Dynamics

$$\mathrm{d}X_t = \kappa\big(\alpha - (\alpha+\beta)X_t\big)\,\mathrm{d}t + \sigma\sqrt{X_t(1-X_t)}\,\mathrm{d}W_t, \quad X_t \in (0,1).$$

Fokker–Planck Equation

$$\partial_t q = -\partial_x\Big[\kappa(\alpha-(\alpha+\beta)x)\,q\Big] + \tfrac{1}{2}\,\partial_{xx}\Big[\sigma^2 x(1-x)\,q\Big],$$

with no-flux boundary conditions at $x = 0, 1$.

Reverse Dynamics By time-reversal of diffusions with state-dependent diffusion:

$$\mathrm{d}X_t = \Big[\kappa(\alpha-(\alpha+\beta)X_t) - \sigma^2\,\partial_x \ln f_{\mathrm{Beta}}(X_t)\Big]\mathrm{d}t + \sigma\sqrt{X_t(1-X_t)}\,\mathrm{d}\bar{W}_t.$$

Pros/Cons

Pros: Analytic stationary; natural for proportions.

Cons: Score diverges near boundaries; numerical schemes must respect reflecting behavior.

3.1.2 Dirichlet Diffusion

Intuition Models compositional data on the simplex Δ^{K-1}, maintaining non-negativity and unit-sum.

Forward Dynamics For $X_t = (X_t^1, \dots, X_t^K) \in \Delta^{K-1}$,

$$\mathrm{d}X_t^i = \sum_{j=1}^{K} a_{ij}\big(\pi_j - X_t^i\big)\,\mathrm{d}t + \sum_{j=1}^{K} \sqrt{X_t^i(\delta_{ij} - X_t^j)}\,\mathrm{d}W_t^{ij},$$

with parameters $\{a_{ij}\}$ chosen so the stationary law is Dirichlet(π).

Fokker–Planck The generator is the Dirichlet operator on the simplex; boundary conditions ensure mass stays within Δ^{K-1}.

Reverse Dynamics

$$\mathrm{d}X_t^i = \Big[\sum_j a_{ij}(\pi_j - X_t^i) - \sum_j \partial_{X^i} \ln q(X_t)\, X_t^i(\delta_{ij} - X_t^j)\Big]\mathrm{d}t + \sum_j \sqrt{X_t^i(\delta_{ij} - X_t^j)}\,\mathrm{d}\bar{W}_t^{ij}.$$

Pros/Cons

Pros: Maintains simplex constraints; analytic stationary.

Cons: Complex geometry; scoring requires manifold derivatives.

3.2 B2. Manifold-Valued Diffusions

3.2.1 Brownian Motion on the Sphere

Intuition Uniform exploration of the sphere S^{d-1} preserving curvature constraints.

Dynamics Embedded in $\mathbb{R}^d$:

$$\mathrm{d}X_t = (I - X_t X_t^T)\,\mathrm{d}W_t \;-\; \tfrac{d-1}{2}\,X_t\,\mathrm{d}t, \quad X_t \in S^{d-1}.$$

Generator and Reversibility The generator is the Laplace–Beltrami operator on S^{d-1}, self-adjoint under the uniform measure. Forward and backward SDE coincide.

Pros/Cons

Pros: Intrinsic to manifold; closed-form; reversible.

Cons: Requires projection operations; no global coordinate chart.

3.2.2 Diffusions on Lie Groups

Intuition Generalizes manifold diffusions to compact Lie groups G, sampling according to Haar measure.

Dynamics Using Stratonovich SDE in the Lie algebra $\mathfrak{g}$:

$$\mathrm{d}g_t = g_t \circ \mathrm{d}W_t^{\mathfrak{g}}, \quad g_t \in G.$$

Generator and Reverse The generator is the Casimir (bi-invariant Laplacian) on G, reversible under Haar measure; forward and reverse dynamics coincide in Stratonovich form.

Pros/Cons

Pros: Respects group structure; analytic stationary.

Cons: Algebraically involved; requires exponential/log maps for implementation.

4 C. Structured-State Diffusions

General Form and Common Properties Structured-state diffusions evolve on spaces whose elements are not plain vectors in $\mathbb{R}^d$, but rather structured objects such as matrices or probability measures. A generic formulation is

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \Sigma(X_t)\,\mathrm{d}W_t + \mathcal{J}(X_t)\,\mathrm{d}J_t,$$

where

- b is a drift mapping that preserves the structure (e.g. keeps matrices positive-definite, or measures normalized),
- $\Sigma\,\mathrm{d}W_t$ represents Gaussian perturbations consistent with the structure,
- $\mathcal{J}\,\mathrm{d}J_t$ may add jumps or resets in structured form (e.g. cloning/branching for measures).

The corresponding forward equation is a matrix- or measure-valued Fokker–Planck (or master) equation. Reversibility, when it holds, is characterized by a detailed-balance condition under the target law (e.g. Wishart for matrices, Dirichlet–Fleming–Viot for measures).

4.1 C1. Matrix-Valued Diffusions

4.1.1 Wishart Process

Intuition Generates random covariance matrices $X_t \succ 0$; used in Bayesian statistics and random matrix theory.

Forward Dynamics

$$\mathrm{d}X_t = \kappa\big(M - X_t\big)\,\mathrm{d}t \;+\; \sqrt{X_t}\,\mathrm{d}W_t \;+\; \mathrm{d}W_t^T\,\sqrt{X_t},$$

where W_t is a matrix of independent Brownian motions and $\sqrt{X_t}$ the unique PSD square-root.

Matrix Fokker–Planck The density $q(X,t)$ on the space of PSD matrices satisfies

$$\partial_t q = -\nabla_X \cdot \big(\kappa(M - X)\,q\big) + \tfrac{1}{2}\,\nabla_X \cdot \big(X\,\nabla_X q + (\nabla_X q)\,X\big),$$

where ∇_X denotes the matrix-gradient (Fréchet derivative).

Reverse Dynamics By time-reversal of matrix diffusions, the backward drift becomes

$$b_{\mathrm{rev}}(X) = \kappa(M - X) - \big[X\,\nabla_X \ln q(X) + (\nabla_X \ln q(X))\,X\big].$$

Hence

$$\mathrm{d}X_t = b_{\mathrm{rev}}(X_t)\,\mathrm{d}t + \sqrt{X_t}\,\mathrm{d}\bar{W}_t + \mathrm{d}\bar{W}_t^T\,\sqrt{X_t}.$$

Pros/Cons

Pros: Ensures $X_t \succ 0$; closed-form forward kernel; interpretable as covariance evolution.

Cons: High-dimensional state; computing $\nabla_X \ln q$ (matrix score) is costly; eigen-decompositions required.

4.2 C2. Interacting/Measure-Valued Diffusions

4.2.1 McKean–Vlasov Diffusion

Intuition Models a large population of particles whose individual drift depends on the empirical law μ_t, capturing mean-field interactions.

Forward Dynamics For each particle X_t:

$$\mathrm{d}X_t = b\big(X_t, \mu_t\big)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \quad \mu_t = \mathcal{L}(X_t).$$

Nonlinear Fokker–Planck The law μ_t evolves by the McKean–Vlasov equation:

$$\partial_t \mu_t + \nabla_x \cdot \Big(b(x, \mu_t)\,\mu_t\Big) = \tfrac{\sigma^2}{2}\,\Delta\mu_t.$$

Reverse Dynamics Reversal requires functional derivatives:

$$b_{\mathrm{rev}}(x, \mu) = b(x, \mu) - \sigma^2\,\nabla_x \frac{\delta}{\delta\mu} \ln\mu(x),$$

leading to an infinite-dimensional backward SDE on path-space.

Pros/Cons

Pros: Captures collective phenomena (synchronization, flocking).

Cons: Infinite-dimensional; requires Fréchet/functional derivatives of the density.

4.2.2 Fleming–Viot Process

Intuition A measure-valued diffusion modeling gene-genealogies with random reproduction and mutation.

Forward Dynamics The state is a probability measure X_t; its evolution is governed by a stochastic partial differential equation (SPDE) combining diffusion, branching, and resampling.

Reverse Dynamics Time-reversal lives on the space of probability measures (Wasserstein manifold) and requires advanced stochastic calculus in infinite dimensions.

Pros/Cons

Pros: Rich model for evolving distributions; includes branching.

Cons: Infinite-dimensional SPDE; very complex reverse formulation.

5 D. Hybrid Processes

General Form and Common Properties Hybrid processes combine deterministic flows, continuous diffusions, and discrete jumps. A general formulation on state space E is:

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t + \int_{\mathcal{Z}} J(X_{t^-}, z)\,N(\mathrm{d}t, \mathrm{d}z),$$

where:

- b is the deterministic drift (ODE part),
- $\sigma\,\mathrm{d}W_t$ the Gaussian diffusion,

- N a Poisson random measure with jump kernel J,
- Jumps at random times reset or shift the state.

The forward evolution of the law satisfies a piecewise-deterministic Fokker–Planck equation mixed with a jump term. Reverse-time dynamics require adjusting both the drift and the jump intensities via detailed balance.

5.1 D1. Stochastic Volatility & PDMP

5.1.1 Heston Model

Intuition Captures joint dynamics of an asset S_t and its stochastic variance V_t, combining continuous diffusion in both variables.

Forward Dynamics

$$
\begin{aligned}
\mathrm{d}S_t &= S_t\sqrt{V_t}\,\mathrm{d}W_t^S,\\
\mathrm{d}V_t &= \kappa(\theta - V_t)\,\mathrm{d}t + \xi\sqrt{V_t}\,\mathrm{d}W_t^V,
\end{aligned}
$$

with $\mathrm{d}W^S\,\mathrm{d}W^V = \rho\,\mathrm{d}t$.

Generator & Fokker–Planck The joint density $q(s,v,t)$ satisfies

$$
\partial_t q = -\partial_s\big(\tfrac{1}{2}sv\,q\big) - \partial_v\big(\kappa(\theta - v)\,q\big) + \tfrac{1}{2}\partial_{ss}\big(s^2 v\,q\big) + \tfrac{1}{2}\partial_{vv}\big(\xi^2 v\,q\big) + \rho\,\partial_{sv}\big(\xi sv\,q\big).
$$

Reverse Dynamics By time-reversal of multidimensional diffusions:

$$
\begin{aligned}
\mathrm{d}S_t &= \big[\tfrac{1}{2}S_tV_t + \partial_s \ln q(S_t,V_t)\,S_t^2V_t + \rho\,\partial_v \ln q(S_t,V_t)\,\xi S_tV_t\big]\mathrm{d}t + \sqrt{V_t}\,S_t\,\mathrm{d}\bar{W}_t^S,\\
\mathrm{d}V_t &= \big[\kappa(\theta - V_t) + \partial_v \ln q(S_t,V_t)\,\xi^2V_t + \rho\,\partial_s \ln q(S_t,V_t)\,\xi S_tV_t\big]\mathrm{d}t + \xi\sqrt{V_t}\,\mathrm{d}\bar{W}_t^V.
\end{aligned}
$$

Pros/Cons

Pros: Captures volatility clustering; well-studied in finance.

Cons: Requires joint score estimation in 2D; correlated Brownian motions complicate training.

5.1.2 Piecewise-Deterministic Markov Process (PDMP)

Intuition Follows deterministic ODE flows punctuated by random jumps (e.g. queueing systems, network failures).

Forward Dynamics

$$
\begin{cases}
\dot{X}_t = f(X_t), & \text{between jumps},\\
X_{t^+} = R(X_{t^-}, Z), & \text{when a jump occurs at rate } \lambda(X_{t^-}).
\end{cases}
$$

Jump times follow an inhomogeneous Poisson process with intensity $\lambda(X)$, and R is the reset map using random mark Z.

Master Equation The law μ_t satisfies

$$
\partial_t\mu_t + \nabla\cdot\big(f(x)\,\mu_t\big) = -\lambda(x)\,\mu_t + \int \lambda(y)\,\mu_t(\mathrm{d}y)\,Q(y,\mathrm{d}x),
$$

where $Q(y,\mathrm{d}x)$ is the distribution of post-jump states.

Reverse Dynamics Reverse requires modified flow and jump intensity:

$$\begin{cases} \dot{X}_t = f(X_t) - \sigma_J(X_t)\,\nabla_x \ln q(X_t), \\ \lambda_{\rm rev}(x) = \lambda(x)\,\frac{q(x)}{q_{\rm pre}(x)}, \quad R_{\rm rev}(x,z) \text{ derived from detailed balance,} \end{cases}$$

where q and $q_{\rm pre}$ are pre- and post-jump densities.

Pros/Cons

Pros: Models hybrid dynamics; exact likelihood when jumps known.

Cons: Reverse-time jump intensities and maps are complex; path-dependence.

5.2 D2. Discrete-State Processes

5.2.1 Birth–Death Process

Intuition Counts evolve with births at rate λ_k and deaths at rate μ_k.

Forward Master Equation

$$\dot{p}_k = \lambda_{k-1}\,p_{k-1} + \mu_{k+1}\,p_{k+1} - (\lambda_k + \mu_k)\,p_k.$$

Stationary Law & Detailed Balance If π_k satisfies $\pi_{k-1}\lambda_{k-1} = \pi_k\mu_k$, then $\pi_k \propto \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}$.

Reverse Dynamics Birth and death rates swap:

$$\lambda_k^{\rm rev} = \mu_{k+1}\,\frac{\pi_{k+1}}{\pi_k}, \quad \mu_k^{\rm rev} = \lambda_{k-1}\,\frac{\pi_{k-1}}{\pi_k}.$$

Pros/Cons

Pros: Exact discrete dynamics; analytic stationary.

Cons: State-space explosion; discrete-time learning required.

Conclusion

We have detailed intuition, equations, derivations, and trade-offs for a broad spectrum of diffusion processes across domains.

References

[1] Anderson, B. D. O. (1982). "Time-reversal of linear stochastic systems." *SIAM Journal on Control and Optimization*, 20(5), 745–757.