A COMPLETE DECOMPOSITION OF STOCHASTIC DIFFERENTIAL EQUATIONS

Samuel Duffield
Normal Computing
sam@normalcomputing.com

ABSTRACT

We show that any stochastic differential equation with prescribed time-dependent marginal distributions admits a decomposition into three components: a unique scalar field governing marginal evolution, a symmetric positive-semidefinite diffusion matrix field and a skew-symmetric matrix field.

1 INTRODUCTION

Stochastic differential equations (SDEs) play a fundamental role across numerous scientific and engineering disciplines, providing a mathematical framework for modelling systems subject to random fluctuations. In particular, SDEs have emerged as a powerful tool in machine learning, driving advances in generative modelling through diffusion processes (Ho et al., 2020; Song et al., 2020; Karras et al., 2022) and enabling efficient sampling algorithms via Markov chain Monte Carlo (Horowitz, 1991; Ma et al., 2015; Duffield et al., 2024).

the probability measure over entire trajectories $x_{[0,T]} = \{x_t : t \in [0,T]\}$ of the stochastic process, and the temporal marginal distributions $p(x,t)$, which describe the probability density of the process at each fixed time t. While the path distribution captures temporal correlations and the full joint statistics of the process across time, the marginal distributions $p(x,t)$ only specify the instantaneous state distribution at each moment. In many applications, the marginals $p(x,t)$ are of primary interest, as they are directly observable, computationally more tractable, and sufficient for characterizing equilibrium states or generative model outputs. An SDE uniquely determines both, but prescribing only the marginals $p(x,t)$ leaves substantial freedom in the path structure, as many different SDEs can share the same time-dependent marginals while exhibiting different temporal dependencies.

In this paper, we show that any SDE with prescribed temporal marginal distributions $p(x,t)$ can be decomposed into three components:

- a unique scalar field $\phi(x,t) : \mathbb{R}^d \times [0,T] \to \mathbb{R}$ governing marginal evolution,
- a symmetric positive-semidefinite matrix field $D(x,t) : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d\times d}$,
- a skew-symmetric matrix field $Q(x,t) : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d\times d}$,

and that any SDE with prescribed temporal marginal distributions must conform to this decomposition. Theorem 1 states the decomposition formally and Theorem 2 states existence and uniqueness of the scalar field $\phi(x,t)$. Proofs are provided in the Appendix.

2 A COMPLETE DECOMPOSITION OF SDES

We now present our complete characterization of SDEs with prescribed temporal marginal distributions. The decomposition separates the drift into three interpretable components: a scalar field $\phi(x,t)$ that governs the evolution of the marginals as well as a symmetric positive-semidefinite diffusion matrix $D(x,t)$ and a skew-symmetric matrix $Q(x,t)$ that are probability preserving. The scalar field $\phi(x,t)$ is linked to $p(x,t)$ through a Poisson equation, while $D(x,t)$ and $Q(x,t)$ leave the marginals invariant.

Theorem 1. An SDE has temporal marginal distributions $p(x,t)$ (with mild assumptions equation 10) if and only if it has the following form

$$\begin{aligned} dx &= \phi(x,t)\nabla_x \log p(x,t)dt + \nabla_x\phi(x,t)dt \\ &\quad + [D(x,t)+Q(x,t)]\nabla_x \log p(x,t)dt + \nabla_x \cdot [D(x,t)+Q(x,t)]dt \qquad (1) \\ &\quad + \sqrt{2D(x,t)}dw, \\ \partial_t p(x,t) &= -\Delta_x[\phi(x,t)p(x,t)], \qquad (2) \end{aligned}$$

for positive-semidefinite $D(x,t) = D(x,t)^\top$ and skew-symmetric $Q(x,t) = -Q(x,t)^\top$.

While the diffusion and skew-symmetric components $D(x,t)$ and $Q(x,t)$ can be chosen freely, the scalar field $\phi(x,t)$ is uniquely determined by the marginal distributions. This ensures that $\phi(x,t)$ carries all the information about how the distribution evolves over time.

Theorem 2. For given temporal marginal distributions $p(x,t)$ there exists a unique scalar field $\phi(x,t)$ in the decomposition equation 1 such that $\lim_{|x|\to\infty}[\phi(x,t)p(x,t)] = 0$.

3 Related Work

The most closely related work is the complete recipe for autonomous SDEs and stationary distributions (Ma et al., 2015). This result states that an ergodic autonomous SDE (i.e. an SDE with coefficients independent of time) has a prescribed stationary distribution $\pi(x)$ if and only if it has the following form

$$dx = [D(x)+Q(x)]\nabla_x \log \pi(x)dt + \nabla_x \cdot [D(x)+Q(x)]dt + \sqrt{2D(x)}dw, \qquad (3)$$

As we will see in Section 4.1, this result can be viewed as a special case of our decomposition equation 1. The completeness of this result is shown using a Fourier analysis of the Fokker-Planck equation. A simpler approach using a Helmholtz decomposition (which is also used in our proof of Theorem 1) is provided in Da Costa & Pavliotis (2023) in the context of entropy production. A related decomposition of ergodic autonomous SDEs into symmetric and skew-symmetric components is also provided in Ao (2004).

In the context of time-reversed generative differential equations (Anderson, 1982; Haussmann & Pardoux, 1986; Song et al., 2020), it is well known that multiple SDEs (and ODEs) can match the same temporal marginal distributions. Most notably in Karras et al. (2022) it was shown that Langevin diffusions with different noise levels can be added without modifying the temporal marginal distributions. Our decomposition provides a complete characterization of all such SDEs. This is expanded on in Section 4.2 and Section 4.3.

The problem of constructing stochastic processes with prescribed marginal distributions is closely related to the Schrödinger bridge problem and, more broadly, to optimal transport. Given initial and terminal marginals, the Schrödinger bridge selects a diffusion process that interpolates between them while minimizing relative entropy with respect to a reference Brownian motion (Jamison, 1974; Föllmer, 1988; Léonard, 2014). Similarly, there has been work on finding vector fields that solve Liouville's equation (Heng, 2016) for a given transport path defined by $p(x,t)$. In contrast to this variational perspective, which singles out one distinguished process, our result provides a complete characterization of *all* SDEs consistent with a given family of time-dependent marginals.

4 Some Corollaries

4.1 Invariant SDEs and Stationary Distributions

We now show that our Theorem 1 generalizes the autonomous complete recipe from Ma et al. (2015).

For an ergodic invariant SDE with stationary distribution $\pi(x)$ we have $\partial_t p(x,t) = 0$. This gives Poisson equation equation 2 $\Delta_x[\phi(x,t)p(x,t)] = 0$, to which $\phi(x,t) = 0$ is a trivial bounded solution and by Theorem 2 is also unique.

With $\phi(x,t) = 0$ and $p(x,t) = \pi(x)$ then equation 1 becomes

$$dx = [D(x,t) + Q(x,t)]\nabla_x \log \pi(x)dt + \nabla_x \cdot [D(x,t) + Q(x,t)]dt + \sqrt{2D(x,t)}dw.$$

Interestingly, this generalises the form in equation 3 from Ma et al. (2015) as it highlights that the SDE is still invariant for $\pi(x)$ despite $D(x,t)$ and $Q(x,t)$ potentially depending on t and therefore not autonomous.

4.2 SDE Matching

Suppose we are given an SDE

$$dx = f(x,t)dt + \sqrt{2\Sigma(x,t)}dw, \tag{4}$$

but not necessarily knowledge of the marginals $p(x,t)$.

By Theorem 1 the following SDE has the same marginals $p(x,t)$ for any choice of skew-symmetric $Q(x,t)$ and positive-semidefinite $D(x,t)$

$$\begin{aligned} dx =& f(x,t)dt + [D(x,t) - \Sigma(x,t) + Q(x,t)]\nabla_x \log p(x,t)dt \\ &+ \nabla_x \cdot [D(x,t) - \Sigma(x,t) + Q(x,t)]dt + \sqrt{2D(x,t)}dw, \end{aligned} \tag{5}$$

We can see that equation 5 matches equation 4 exactly with $D(x,t) = \Sigma(x,t)$ and $Q(x,t) = 0$ and then by Theorem 1 all SDEs of the form equation 5 have the same marginals for all skew-symmetric $Q(x,t)$ and positive-semidefinite $D(x,t)$.

Note that equation 5 without Q and D terms has the same marginals but does not necessarily match equation 1 without Q and D terms (i.e. the ODE induced by the unique ϕ) since an alternate Q matrix may already be embedded in the ODE.

4.3 Time Reversal

It is well established (Anderson, 1982; Haussmann & Pardoux, 1986; Cattiaux et al., 2023; Kim, 2025) that an SDE equation 4 can be time-reversed to give the following SDE in $y_s = x_{T-s}$

$$dy = -\bar{f}(y,s)ds + 2\nabla_y \cdot \bar{\Sigma}(y,s)ds + 2\bar{\Sigma}(y,s)\nabla_y \log \bar{p}(y,s)ds + \sqrt{2\bar{\Sigma}(y,s)}d\bar{w}(s), \tag{6}$$

where $\bar{f}(y,s) = f(y, T-s)$, $\bar{\Sigma}(y,s) = \Sigma(y, T-s)$, $\bar{p}(y,s) = p(y, T-s)$ and $d\bar{w}$ is a time-reversed Brownian motion.

The SDE equation 6 is a strict time-reversal in the sense that it matches the path distribution $p(\bar{y}_{[0,T]}) = p(x_{[0,T]})$ of the original SDE equation 4 (here $\bar{y}_s = y_{T-s}$).

In practice (Song et al., 2020), we are often happy with a weaker time-reversal which only matches the temporal marginal distributions $p(y,s) = p(x, T-s)$ of the original SDE equation 4. Our Theorem 1 and Corollary 4.2 applied to equation 6 provide a complete characterization of all such time-reversals

$$\begin{aligned} dy =& -\bar{f}(y,s)ds + [\bar{D}(y,s) + \bar{\Sigma}(y,s) + \bar{Q}(y,s)]\nabla_y \log \bar{p}(y,s)ds \\ &+ \nabla_y \cdot [\bar{D}(y,s) + \bar{\Sigma}(y,s) + \bar{Q}(y,s)]ds + \sqrt{2\bar{D}(y,s)}d\bar{w}(s), \end{aligned}$$

for any choice of skew-symmetric $\bar{Q}(y,s)$ and positive-semidefinite $\bar{D}(y,s)$.

4.4 Generative Denoising Models

A dominant strategy in generative models is to train a parameterised ODE or SDE as the time-reversal of a specified linear noising SDE with $p(x,0)$ an empirical data distribution (Song et al., 2020; Huang et al., 2021; Karras et al., 2022; Duffield et al., 2025).

Following Karras et al. (2022), this noising process can be written as a linear SDE

$$dx = f(t)xdt + g(t)dw, \tag{7}$$

or equivalently as mollified Gaussian conditionals

$$p(x_t \mid x_0) = \mathcal{N}(x_t \mid s(t)x_0, s(t)^2\sigma(t)^2\mathbb{I}). \tag{8}$$

See Karras et al. (2022) for details including how to convert between $(f(t), g(t))$ and $(s(t), \sigma(t))$.

Also provided in Karras et al. (2022) is a recipe for constructing weak time-reversals

$$dy = -[\partial_s\bar{\sigma}(s)]\bar{\sigma}(s)\nabla_y \log \bar{p}(y,s)ds + \bar{\beta}(s)\bar{\sigma}(s)^2\nabla_y \log \bar{p}(y,s)ds + \sqrt{2\bar{\beta}(s)}\bar{\sigma}(s)d\bar{w},$$

for $\bar{\sigma}(s) = \sigma(T-s)$, $\bar{p}(y,s) = p(y, T-s)$ and any choice of $\bar{\beta}(s) \geq 0$.

Our decomposition elucidates that this recipe is incomplete and a complete recipe of all valid weak time-reversals is given by

$$\begin{aligned} dy = &- [\partial_s\bar{\sigma}(s)]\bar{\sigma}(s)\nabla_y \log \bar{p}(y,s)ds \\ &+ [\bar{D}(y,s) + \bar{Q}(y,s)]\nabla_y \log \bar{p}(y,s)ds + \nabla_y \cdot [\bar{D}(y,s) + \bar{Q}(y,s)]ds + \sqrt{2\bar{D}(y,s)}d\bar{w}(s), \end{aligned} \tag{9}$$

for any choice of skew-symmetric $\bar{Q}(y,s)$ and positive-semidefinite $\bar{D}(y,s)$.

The form of equation 9 also exactly matches the form in Theorem 1 and thus we can conclude that

$$\phi(y,s) = -[\partial_s\bar{\sigma}(s)]\bar{\sigma}(s),$$

is the unique scalar field that satisfies the Poisson equation equation 2 for the weak time-reversal of the noising process (7-8).

5 Discussion

We have shown that prescribing the full family of temporal marginals $p(x,t)$ essentially fixes one and only one component of the dynamics: the scalar field $\phi(x,t)$ defined by the Poisson equation $\partial_t p(x,t) = -\Delta_x[\phi(x,t)p(x,t)]$ (under mild regularity and decay conditions). All remaining freedom in the path law consistent with the same marginals is captured by a symmetric positive-semidefinite diffusion field $D(x,t)$ and a skew-symmetric field $Q(x,t)$, which are probability-preserving in the sense that they do not affect the instantaneous evolution of the marginals.

Given the generality and completeness of the decomposition, we expect it to motivate many organic research directions; we highlight a few representative avenues. First, the decomposition suggests new optimization and control formulations: one can view D and Q as controllable fields that reshape temporal correlations while keeping $p(x,t)$ fixed, opening the door to objectives for their selection such as minimal entropy production (Da Costa & Pavliotis, 2023), accelerated convergence, or variance reduction for sampling and simulation (Chak et al., 2023). Second, in generative modelling, the characterization provides a complete description of the freedom in weak time-reversals and clarifies how additional noise, non-reversible flows, and state-dependent diffusion can be introduced without affecting the learned marginal path. Specifically, equation 9 with $Q = D = 0$ is often referred to as the probability flow ODE (Song et al., 2020; Karras et al., 2022) but our formulation points to infinitely many marginal-preserving ODEs dictated by the choice of Q. Third, the autonomous recipe Ma et al. (2015) has been used to devise SDEs that act on an extended space but still preserve the stationary distribution on a subspace as in Hamiltonian Monte Carlo (Betancourt, 2017) and underdamped Langevin dynamics (Horowitz, 1991; Chak et al., 2023) or via interacting particles (Leimkuhler et al., 2018; Duncan et al., 2023). Our introduced decomposition enables similar approaches in non-autonomous settings.

Overall, the introduced decomposition provides a simple and general organizing principle for SDEs with prescribed marginals: ϕ is the unique signature of marginal evolution, while D and Q parameterise all remaining dynamical freedom.

Acknowledgments

We thank Maxwell Aifer, Adrien Corenflos, Gavin Crooks, Lars Holdijk, Samuel Power and Max Welling for discussions and helpful feedback on the manuscript. We thank the Advanced Research and Invention Agency's (ARIA) Scaling Compute programme for funding this work.

References

Brian DO Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.

Ping Ao. Potential in stochastic differential equations: novel construction. *Journal of Physics A: Mathematical and General*, 37(3):L25, 2004.

Michael Betancourt. A conceptual introduction to Hamiltonian Monte Carlo. *arXiv preprint arXiv:1701.02434*, 2017.

Patrick Cattiaux, Giovanni Conforti, Ivan Gentil, and Christian Léonard. Time reversal of diffusion processes under a finite entropy condition. In *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, volume 59, pp. 1844–1881. Institut Henri Poincaré, 2023.

Martin Chak, Nikolas Kantas, Tony Lelièvre, and Grigorios A Pavliotis. Optimal friction matrix for underdamped Langevin sampling. *ESAIM: Mathematical Modelling and Numerical Analysis*, 57 (6):3335–3371, 2023.

Lancelot Da Costa and Grigorios A Pavliotis. The entropy production of stationary diffusions. *Journal of Physics A: Mathematical and Theoretical*, 56(36):365001, 2023.

Samuel Duffield, Kaelan Donatella, Johnathan Chiu, Phoebe Klett, and Daniel Simpson. Scalable Bayesian learning with posteriors. *arXiv preprint arXiv:2406.00104*, 2024.

Samuel Duffield, Maxwell Aifer, Denis Melanson, Zach Belateche, and Patrick J Coles. Lattice random walk discretisations of stochastic differential equations. *arXiv preprint arXiv:2508.20883*, 2025.

Andrew Duncan, Nikolas Nüsken, and Lukasz Szpruch. On the geometry of Stein variational gradient descent. *Journal of Machine Learning Research*, 24(56):1–39, 2023.

Gerald B Folland. *Introduction to Partial Differential Equations*, volume 102. Princeton University Press, 1995.

Hans Föllmer. Random fields and diffusion processes. *Lecture Notes in Mathematics*, 1362:101–203, 1988.

Erhard Glötzl and Oliver Richters. Helmholtz decomposition and potential functions for n-dimensional analytic vector fields. *Journal of Mathematical Analysis and Applications*, 525(2): 127138, 2023.

Ulrich G Haussmann and Etienne Pardoux. Time reversal of diffusions. *The Annals of Probability*, pp. 1188–1205, 1986.

Jeremy Heng. *On the use of transport and optimal control methods for Monte Carlo simulation*. PhD thesis, University of Oxford, 2016.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

Alan M Horowitz. A generalized guided Monte Carlo algorithm. *Physics Letters B*, 268(2):247–252, 1991.

Chin-Wei Huang, Jae Hyun Lim, and Aaron C Courville. A variational perspective on diffusion-based generative models and score matching. *Advances in Neural Information Processing Systems*, 34:22863–22876, 2021.

Benton Jamison. Reciprocal processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 30(1):65–86, 1974.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in Neural Information Processing Systems*, 35:26565–26577, 2022.

Ji-Ha Kim. Deriving reverse-time stochastic differential equations (SDEs). Feb 2025. URL https://jiha-kim.github.io/posts/deriving-reverse-time-stochastic-differential-equations-sdes/.

Benedict Leimkuhler, Charles Matthews, and Jonathan Weare. Ensemble preconditioning for Markov chain Monte Carlo simulation. *Statistics and Computing*, 28(2):277–290, 2018.

Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete & Continuous Dynamical Systems - A*, 34(4):1533–1574, 2014.

Yi-An Ma, Tianqi Chen, and Emily Fox. A complete recipe for stochastic gradient MCMC. *Advances in Neural Information Processing Systems*, 28, 2015.

Edward Nelson. A proof of Liouville's theorem. In *Proc. AMS*, volume 12, pp. 995, 1961.

Grigorios A Pavliotis. Stochastic processes and applications. *Texts in Applied Mathematics*, 60, 2014.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Elias M Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970.

A NOTATION

- For a vector field $v : \mathbb{R}^d \to \mathbb{R}^d$, we write $\nabla \cdot v(x) := \sum_{i=1}^d \partial_{x_i} v_i(x)$.
- For a matrix field $A : \mathbb{R}^d \to \mathbb{R}^{d \times d}$, we define the matrix divergence by $[\nabla \cdot A(x)]_i := \sum_{j=1}^d \partial_{x_j} A_{ij}(x)$, with $\nabla \cdot A : \mathbb{R}^d \to \mathbb{R}^d$.
- For a scalar field $g : \mathbb{R}^d \to \mathbb{R}$, the Laplacian is $\Delta g(x) := \nabla \cdot \nabla g(x) = \sum_{i=1}^d \partial_{x_i}^2 g(x)$.
- For $k \in \mathbb{N}$, we write $C^k(\mathbb{R}^d)$ for the space of k-times continuously differentiable functions $f : \mathbb{R}^d \to \mathbb{R}$.
- For $1 \leq p < \infty$, we write $L^p(\mathbb{R}^d)$ for the space of measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ such that $\|f\|_{L^p} := (\int_{\mathbb{R}^d} |f(x)|^p \, dx)^{1/p}$ is finite.
- A scalar field $f : \mathbb{R}^d \to \mathbb{R}$ is said to be *bounded* if there exists $M > 0$ such that $|f(x)| \leq M$ for all $x \in \mathbb{R}^d$.
- A function $f : \mathbb{R}^d \to \mathbb{R}^n$ is said to *decay sufficiently at infinity* if there exists $\varepsilon > 0$ such that $\lim_{|x| \to \infty} |x|^{d+\varepsilon} |f(x)| = 0$.
- We say an SDE is *autonomous* if its coefficients are time-independent, i.e. it has the form $dx = b(x)dt + \sqrt{2D(x)}dw$. We say an autonomous SDE is *ergodic* if it has a unique stationary distribution.
- We say an SDE (which is not necessarily autonomous) is *invariant* if $\partial_t p(x, t) = 0$ for $x \sim \pi(x)$ for some distribution π. Again the SDE is *ergodic* if $\partial_t p(x, t) = 0$ for a unique π.
- For a complex number $z = a + bi \in \mathbb{C}$ with $a, b \in \mathbb{R}$ we denote the complex conjugate $\bar{z} = a - bi$.
- We denote the Fourier transform of a (potentially complex) scalar field $f : \mathbb{R}^d \to \mathbb{C}$ as $\text{FT}[f](\xi) = \int e^{-ix^\top \xi} f(x) dx$ with $\text{FT}[f] : \mathbb{R}^d \to \mathbb{C}$.
- We denote the inverse Fourier transform of a (potentially complex) scalar field $g : \mathbb{R}^d \to \mathbb{C}$ as $\text{IFT}[g](x) = \frac{1}{(2\pi)^d} \int e^{ix^\top \xi} g(\xi) d\xi$ with $\text{IFT}[g] : \mathbb{R}^d \to \mathbb{C}$.

B Assumptions

Throughout we make the following mild assumptions on $p(x,t)$:

$$p(x,t) > 0 \quad \forall x, t, \tag{10a}$$
$$p(\cdot,t) \in C^2(\mathbb{R}^d) \quad \forall t, \tag{10b}$$
$$p(x,\cdot) \in C^1(\mathbb{R}) \quad \forall x, \tag{10c}$$
$$\int p(x,t)dx = 1 \quad \forall t, \tag{10d}$$
$$p(\cdot,t), \nabla_x p(\cdot,t) \text{ and } \partial_t p(\cdot,t) \text{ decay sufficiently at infinity } \forall t. \tag{10e}$$

We also note that the decay sufficiently at infinity condition implies $p(\cdot,t), \nabla_x p(\cdot,t), \partial_t p(\cdot,t) \in L^1(\mathbb{R}^d)$. This follows since if $\lim_{|x|\to\infty} |x|^{d+\varepsilon}|f(x)| = 0$ for some $\varepsilon > 0$, then $|f(x)| \le C|x|^{-d-\varepsilon}$ for $|x| \ge R$, hence

$$\int_{\mathbb{R}^d} |f(x)|\,dx \le \int_{|x|\le R} |f(x)|\,dx + C\int_{|x|\ge R} |x|^{-d-\varepsilon}\,dx < \infty,$$

so $f \in L^1(\mathbb{R}^d)$.

These assumptions are commonly satisfied in practice and sufficient to justify to ensure validity of the Fokker-Planck equation and suitable integration by parts, see Section 4.1 in Pavliotis (2014).

For example these assumptions are satisfied for distributions of the form $p(x,t) \propto \exp(-U(x,t))$ where $U(x,t)$ is super-logarithmic in the tails.

C Proof of Theorem 1

We restate Theorem 1 for the reader's ease.

Theorem 1. An SDE has temporal marginal distributions $p(x,t)$ (with mild assumptions equation 10) if and only if it has the following form

$$\begin{aligned} dx &= \phi(x,t)\nabla_x \log p(x,t)dt + \nabla_x\phi(x,t)dt \\ &\quad + [D(x,t)+Q(x,t)]\nabla_x \log p(x,t)dt + \nabla_x \cdot [D(x,t)+Q(x,t)]dt \\ &\quad + \sqrt{2D(x,t)}dw, \end{aligned} \tag{11}$$
$$\partial_t p(x,t) = -\Delta_x[\phi(x,t)p(x,t)], \tag{12}$$

for symmetric positive-semidefinite $D(x,t) = D(x,t)^\top$ and skew-symmetric $Q(x,t) = -Q(x,t)^\top$.

Proof: equation 1 $\implies p(x,t)$. We need to show that the Fokker-Planck equation applied to equation 11 reduces to the Poisson equation equation 12.

The full Fokker-Planck equation (see also equation 14) reads

$$\begin{aligned} \partial_t p(x,t) = -\nabla_x \cdot \big[&[\phi(x,t)\nabla_x \log p(x,t) + \nabla_x\phi(x,t)]p(x,t) \\ &+ [[D(x,t)+Q(x,t)]\nabla_x \log p(x,t)]p(x,t) + [\nabla_x \cdot [D(x,t)+Q(x,t)]]p(x,t) \\ &- \nabla_x \cdot [D(x,t)p(x,t)]\big]. \end{aligned}$$

Following the result in Ma et al. (2015) for autonomous SDEs we can show the terms containing Q and D evaluate to zero.

By linearity of the divergence operator $\nabla\cdot$ we can consider just the Q terms

$$\begin{aligned} Q \text{ terms} &= -\nabla_x \cdot \big[Q(x,t)[\nabla_x \log p(x,t)]p(x,t) + [\nabla_x \cdot Q(x,t)]p(x,t)\big], \\ &= -\nabla_x \cdot \nabla_x \cdot [Q(x,t)p(x,t)] \\ &= 0. \end{aligned}$$

Where the second line uses the (reverse) product rule and the third line the skew-symmetry of Q.

Similarly the D terms

$$\begin{aligned} D \text{ terms} &= -\nabla_x \cdot \left[D(x,t)[\nabla_x \log p(x,t)]p(x,t) + [\nabla_x \cdot D(x,t)]p(x,t) - \nabla_x \cdot [D(x,t)p(x,t)]\right], \\ &= -\nabla_x \cdot [\nabla_x \cdot D(x,t)p(x,t) - \nabla_x \cdot [D(x,t)p(x,t)]] \\ &= 0. \end{aligned}$$

This leaves the Fokker-Planck equation as

$$\begin{aligned} \partial_t p(x,t) &= -\nabla_x \cdot [[\phi(x,t)\nabla_x \log p(x,t) + \nabla_x \phi(x,t)]p(x,t)], \\ &= -\nabla_x \cdot [\nabla_x [\phi(x,t)p(x,t)]], \\ &= -\Delta_x [\phi(x,t)p(x,t)], \end{aligned}$$

which is exactly the Poisson equation in equation 12. □

Proof: equation 11 $\Longleftarrow$ $p(x,t)$. Now assume we are provided an SDE

$$dx = b(x,t)dt + \sqrt{2D(x,t)}dw, \tag{13}$$

which has temporal marginal distributions $p(x,t)$. We now are left to show that there exists a scalar field $\phi(x,t)$ and skew-symmetric matrix field $Q(x,t)$ such that

$$b(x,t) = \phi(x,t)\nabla_x \log p(x,t) + \nabla_x \phi(x,t) + [D(x,t) + Q(x,t)]\nabla_x \log p(x,t) + \nabla_x \cdot [D(x,t) + Q(x,t)].$$

The Fokker-Planck equation for equation 13 takes the form

$$\begin{aligned} \partial_t p(x,t) &= -\nabla_x \cdot J(x,t), \\ J(x,t) &= b(x,t)p(x,t) - \nabla_x \cdot [D(x,t)p(x,t)]. \end{aligned} \tag{14}$$

Given assumptions 10 on p and bounded b, D the vector field $J(x,t)$ decays sufficiently at infinity and therefore the Helmholtz decomposition Glötzl & Richters (2023) states that it decomposes into conservative and divergence-free components

$$\begin{aligned} J(x,t) &= \nabla_x [\phi(x,t)p(x,t)] + c(x,t)p(x,t), \\ \nabla_x \cdot [c(x,t)p(x,t)] &= 0. \end{aligned} \tag{15}$$

By 3.10 in Glötzl & Richters (2023) we have the vector field $c(x,t)$ satisfies

$$\nabla_x \cdot [c(x,t)p(x,t)] = 0 \iff c(x,t) = Q(x,t)\nabla_x \log p(x,t) + \nabla \cdot Q(x,t),$$

for a skew-symmetric $Q(x,t) = -Q(x,t)^\top$ and $c(x,t)p(x,t) = \nabla_x \cdot [Q(x,t)p(x,t)]$.

Combining the above we get

$$\begin{aligned} & b(x,t)p(x,t) = \nabla_x [\phi(x,t)p(x,t)] + \nabla_x \cdot [Q(x,t)p(x,t)] + \nabla_x \cdot [D(x,t)p(x,t)], \\ \implies & b(x,t) = \phi(x,t)\nabla_x \log p(x,t) + \nabla_x \phi(x,t) \\ & \qquad + [D(x,t) + Q(x,t)]\nabla_x \log p(x,t) + \nabla_x \cdot [D(x,t) + Q(x,t)]. \end{aligned}$$

Assuming $p(x,t) > 0$ everywhere. This now matches equation 11.

Finally, the Poisson equation $equation$ 12 is direct from equation 15

$$\begin{aligned} \partial_t p(x,t) &= -\nabla_x \cdot \nabla_x [\phi(x,t)p(x,t)] - \nabla_x \cdot [c(x,t)p(x,t)], \\ &= -\Delta_x [\phi(x,t)p(x,t)]. \end{aligned}$$

□

D PROOF OF THEOREM 2

We restate Theorem 2 for the reader's ease.

Theorem 2. For given temporal marginal distributions $p(x,t)$ there exists a unique scalar field $\phi(x,t)$ in the decomposition equation 1 such that $\lim_{|x|\to\infty}[\phi(x,t)p(x,t)] = 0$.

We provide two proofs of the existence of the scalar field ϕ solving the Poisson equation equation 2 for given distributions $p(x,t)$ satisfying assumptions equation 10. In both cases we assume $d \geq 3$, for a thorough treatment of solutions to Poisson equations including $d = 1, 2$ see Folland (1995).

We then prove uniqueness as an application of the Liouville theorem for harmonic functions.

We therefore note that by the uniqueness of ϕ the resulting ϕs from the two existence constructions must be equivalent.

D.1 EXISTENCE

D.1.1 VIA THE FUNDAMENTAL SOLUTION OF THE LAPLACIAN

Proof: Existence. We seek a solution ϕ to the Poisson equation

$$\begin{aligned} \partial_t p(x,t) &= -\Delta_x u(x,t), \\ u(x,t) &= \phi(x,t)p(x,t). \end{aligned}$$

Define the fundamental solution of the Laplacian in $\mathbb{R}^d$ (see e.g. Chapter 2 in Folland (1995)) as:

$$\Phi(x) := \frac{|x|^{2-d}}{d(d-2)\omega_d}, \qquad d \geq 3,$$

where ω_d denotes the volume of the unit ball in $\mathbb{R}^d$. For integrable $f \in C^2(\mathbb{R}^d)$ we have that the function $g(x) = \int_{\mathbb{R}^d} \Phi(x-y)f(y)dy$ is well-defined and locally integrable and

$$\Delta_x \int_{\mathbb{R}^d} \Phi(x-y)f(y)\,dy = -f(x).$$

So we can set

$$u(x,t) = \int_{\mathbb{R}^d} \Phi(x-y)\partial_t p(y,t)dy.$$

Noting that our assumptions equation 10 imply integrability of $\partial_t p(x,t)$ and $\int \partial_t p(x,t)dx = 0$. Further that $\lim_{|x|\to\infty} u(x,t) = 0$ is also implied through the definition of Φ and $\partial_t p$ decaying sufficiently at infinity.

This gives $\Delta_x u(x,t) = -\partial_t p(x,t)$. Thus we can set $\phi(x,t) = u(x,t)/p(x,t)$ as $p(x,t) > 0$ everywhere.

□

D.1.2 VIA FOURIER TRANSFORM

Proof: Existence. We seek a solution ϕ to the Poisson equation

$$\begin{aligned} \partial_t p(x,t) &= -\Delta_x u(x,t), \\ u(x,t) &= \phi(x,t)p(x,t). \end{aligned}$$

Since $\partial_t p$ decays sufficiently at infinity we have the relation between Laplacian and Fourier transform (e.g. Chapter V in Stein (1970) or Chapter 2-C in Folland (1995))

$$\mathrm{FT}(\Delta_x u)(\xi,t) = -|\xi|^2 \mathrm{FT}(u)(\xi,t).$$

Thus

$$\begin{aligned} \mathrm{FT}(u)(\xi,t) &= -|\xi|^{-2}\mathrm{FT}(\Delta_x u)(\xi,t), \\ &= |\xi|^{-2}\mathrm{FT}(\partial_t p)(\xi,t) := v(\xi,t), \\ u(x,t) &= \mathrm{IFT}[v](x,t). \end{aligned}$$

Where our assumptions equation 10 imply $\int \partial_t p(x,t)dx = 0 \;\; \forall t$ which in turn implies $\mathrm{FT}(\partial_t p)(0,t) = 0$ and therefore $\mathrm{FT}(u)(0,t) = 0$ is well defined for $\xi = 0$, further $\partial_t p$ is bounded and continuous so $v(\xi,t)$ is locally integrable near $\xi = 0$ for $d \geq 3$.

Since $\partial_t p(\cdot,t)$ is real-valued, we have the Hermitian symmetry $\mathrm{FT}(\partial_t p)(-\xi,t) = \overline{\mathrm{FT}(\partial_t p)(\xi,t)}$; as $|\xi|^{-2}$ is real and even, it follows that $v(-\xi,t) = \overline{v(\xi,t)}$, and hence $u(x,t) = \mathrm{IFT}[v](x,t)$ is real-valued.

□

D.2 UNIQUENESS

Proof: Uniqueness. Suppose $\phi_1(\cdot,t)$ and $\phi_2(\cdot,t)$ both satisfy equation 2, with $u_i(x,t) := \phi_i(x,t)p(x,t)$. Then

$$\Delta_x u_1(x,t) = -\partial_t p(x,t) = \Delta_x u_2(x,t),$$

so $w(x,t) := u_1(x,t) - u_2(x,t)$ satisfies $\Delta_x w(x,t) = 0$. Thus $w(\cdot,t)$ is harmonic on $\mathbb{R}^d$. We assume each $u_i(\cdot,t)$ vanish at infinity and therefore

$$\lim_{|x|\to\infty} w(x,t) = 0.$$

By the Liouville theorem for harmonic functions Nelson (1961), any harmonic function on $\mathbb{R}^d$ that is bounded—and in particular one that is continuous and vanishes at infinity—must be constant, hence identically zero $w(x,t) = 0\ \forall x,t$. Therefore $u_1(x,t) = u_2(x,t)$ for all x.

Since we assume $p(x,t) > 0$ everywhere, it follows that $\phi_1(x,t) = \phi_2(x,t)$ for all x,t. □