Complex Schrödinger Bridges

Dong-Sig Han **Tolga Birdal**
Department of Computing
Imperial College London

Abstract

Diffusion bridges have emerged as versatile frameworks for facilitating transformations between arbitrary data distributions. However, these approaches are hindered by an immense optimization space and its underlying high-order structures. This paper explores a variant of diffusion bridges for solving real-world problems in complex vector spaces governed by scaling and rotation laws, such as Hamiltonian structures. To address this, we introduce conformal optimal transport (COT), a complex-valued transportation framework designed for physical systems with geometric and dissipative dynamics under a geometric lifting. Building on this, we propose a new class of dynamical diffusion bridges, called complex Schrödinger bridges (CSBs), whose derived algorithms can handle diffusion-based tasks using a conformal geometric prior. For scalability, we propose image-to-image CSB (I^2CSB), a tractable diffusion-bridge technique with an analytic spiral posterior. Finally, we implement a physics-informed CSB algorithm within simulations of quantum electrodynamics, mimicking the logical behavior of quantum logic gates. Our extensive results validate the underlying theory while highlighting advantages of complex-valued diffusion. Our project page is available here.

1 Introduction

Diffusion bridges [37, 70, 87] have evolved into foundational probabilistic models. Among existing formalizations, the Schrödinger bridge (SB) [75] has gained significant traction in generative modeling in machine learning [20, 53, 79]. The framework is intrinsically linked to optimal transportation (OT), stochastic optimal control (SOC), and evolution of physical systems. These connections highlight a wide range of applications, indicating that diffusion bridges are fundamentally governed by the problem formulation and the geometric structures of the underlying spaces. However, designing coordinated priors that can effectively navigate vast optimization landscapes remains a largely open challenge. Embedding physically grounded biases into diffusion bridges provides a principled structural framework for modeling complex real-world dynamics.

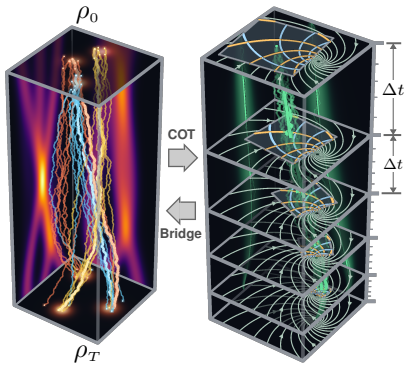

Figure 1: An overview of a complex SB and its associated conformal optimal transport geometry.

The notion of extending SBs requires further elaboration, as the equivalence with dynamic OT implies solid optimization criteria in terms of kinetic energy. Recent studies [53, 54] have introduced *generalized* SBs in a goal-driven manner by treating external state costs as auxiliary terms. In this way, there can be a collection of Hamiltonian functions which describe various particle motions, including scaling, translation, and rotation. This diversity raises a fundamental question:

What principles should govern the real-world dynamics of general diffusion bridges?

Existing SB formulations do not provide a satisfactory answer, as their restriction to linear dynamics on real-valued state spaces prevents them from characterizing rotational, oscillatory, and frequency-

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

dependent behaviors in unified terms. To address this challenge, we introduce an optimization framework based on a phase component endowed with a complex algebraic structure. By leveraging the nature of complex numbers, diffusion bridge models gain the freedom to explore phase trajectories that cannot be represented solely through task-dependent potential energy functions.

In this paper, we study a complex-valued formulation of the Schrödinger bridge problem (SBP). Our complexification of the SBP enables a succinct description of numerical and differential computations using complex operators. Furthermore, the log-Schrödinger potentials are unified using a rotational extension of the Hopf–Cole transform [50]. Fig. 1 presents an overview of the problem formulation; whereas OT priors yield straight paths, our method geometrically transforms SB paths into spiraling arcs. As solutions for this problem, we consider a complex-valued stochastic differential equation (CVSDE) [69] equipped with a McKean–Vlasov structure [57, 84], defined by

$$\mathrm{d}Z_t = i\,\nabla_\alpha H\big(Z_t, \bar{Z}_t, \nabla_{\mathbf{z}}\psi_t, \nabla_{\bar{\mathbf{z}}}\psi_t, \rho_t\big)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t. \tag{1}$$

Here, $H(\mathbf{z}, \bar{\mathbf{z}}, \alpha, \beta, \rho)$ denotes a generalized Hamiltonian function tailored for conjugate variables $\mathbf{z}, \bar{\mathbf{z}} \in \mathbb{C}^d$, paired vector fields $(\alpha, \beta) \coloneqq (\nabla_{\mathbf{z}}\psi, \nabla_{\bar{\mathbf{z}}}\psi)$ associated with the log-potential $\psi(\mathbf{z}, \bar{\mathbf{z}}, t)$, and a density $\rho(\mathbf{z}, \bar{\mathbf{z}}, t)$. Notably, (1) corresponds to a formulation of the Schrödinger equation [75, 85] with an associated Hamiltonian operator $\hat{H}$ on a Hilbert space:

$$i\tfrac{\partial \Psi}{\partial t} = \hat{H}\Psi \qquad \text{and} \qquad i\tfrac{\partial \bar{\Psi}}{\partial t} = -\hat{H}^\dagger\bar{\Psi}, \tag{2}$$

where the factor $i \coloneqq \sqrt{-1}$ plays the role of an orthogonal rotation on the gradient fields within the complex plane, generating phase-coherent trajectories. Under this formulation, the conjugate wavefunctions Ψ and $\bar{\Psi}$ characterize a bosonic coherent state whose physics-inspired mean-field game (MFG) dynamics match the above CVSDE via exponential and logarithmic transformations.

How does this perspective relate to real-valued SB theory? We argue that our extension is a direct geometric consequence of the reference measure under the conformal invariance principle [52, 47, 69]. This introduces an extra geometric degree of freedom motivated by conformal field theory (CFT) [24], with the transport cost defined along conformal geodesics. To this end, we propose the complex Schrödinger bridge (CSB), a novel theoretical framework designed to leverage these underlying structures in real-world physical systems with complex-valued, phase-related dynamics. We demonstrate that complex-valued channels play a role in balancing translational and rotational dynamics. We first derive the general CSB formulation, whose dynamics mirror those of real-valued diffusion bridges while incorporating analytic mapping functions in the complex plane. We then design a tractable class of CSB models tailored for image translation, termed Image-to-Image CSB (I^2CSB). Finally, we validate our approach in simulations of 3D circuit quantum electrodynamics (cQED), demonstrating its efficacy in modeling quantum physical systems.

This paper complements previous studies on diffusion bridges, and our CSB problem presents a novel geometric perspective on classical SBs, fundamentally expanding the capabilities of diffusion bridges for downstream tasks. Our contributions are summarized below.

- Building upon complex analysis and stochastic calculus, we formulate a novel generalization of the classical SB framework. Casting the problem via conformal optimal transport (COT) naturally incorporates the balance between Langevin diffusion dynamics and Hamiltonian rotational motion.
- We develop a comprehensive complex-valued theory and derive algorithms to solve various CSB problems. We propose SB-CVSDEs, complex-valued SDEs for general MFGs using Wirtinger calculus and conjugation algebra. This alleviates the mathematical ambiguity of solving a "two-potential problem": the controllable conjugate potential formally describes the reverse process.
- We validate our theoretical framework and algorithms across medical imaging and quantum physics. Our empirical results demonstrate that our algorithms effectively support both systems relying on implicit priors learned from static images and those governed by explicit physical priors.

2 Preliminaries

Notation. We denote by Z a d-dimensional complex stochastic process $(Z_t)_{0\le t\le T}$, with a map $Z : [0, T] \times \mathcal{Z} \to \mathbb{C}^d$ on the time interval $[0, T]$ and sample space $\mathcal{Z}$. Here, Z_t is a random variable representing the state at time t, and we write $\mathbf{z}_t$ for a realization of Z_t. The complex Brownian motion is denoted by W and is constructed from independent Brownian motions as $(\mathbf{w} + i\hat{\mathbf{w}})/\sqrt{2}$. Let $\Omega = \mathcal{C}([0, T], \mathbb{C}^d)$ denote the space of continuous paths on $[0, T]$, and let $\mathcal{P}(\Omega)$ denote the space of

probability measures on Ω. For a path measure $\mathbb{P} \in \mathcal{P}(\Omega)$ describing the law of Z, we denote by $\mathbb{P}_t$ its time marginal, describing the distribution of Z_t; that is, if $Z \sim \mathbb{P}$, then $Z_t \sim \mathbb{P}_t$. We alternatively write the density as $\rho(\cdot, t)$ to emphasize its interpretation as the limit of an empirical measure. All distributions admit densities on a measurable set $\mathcal{S} \subseteq \mathbb{C}^d$ with respect to a Lebesgue measure, and the KL divergence is defined accordingly and set to $+\infty$ when absolute continuity does not hold. For each SDE, local Lipschitz regularity of partial derivatives of drifts and potentials is assumed.

The Schrödinger bridge problem. Let $\mathbb{Q}_{\mathbf{w}}$ be the reference path measure of a real Wiener process $\mathrm{d}X_t = \sigma \, \mathrm{d}\mathbf{w}_t$, for a diffusion coefficient σ. Among path measures absolutely continuous with respect to $\mathbb{Q}_{\mathbf{w}}$, the SBP seeks a measure that is closest to the reference in terms of the KL divergence [51]:

$$\underset{\mathbb{P}}{\text{minimize}} \ \mathrm{KL}(\mathbb{P} \| \mathbb{Q}_{\mathbf{w}}), \quad \text{such that} \quad \mathbb{P} \in \mathcal{P}\big(\mathcal{C}([0,T], \mathbb{R}^d)\big), \quad \mathbb{P}_0 = \rho_0, \, \mathbb{P}_T = \rho_T, \tag{3}$$

where ρ_0 and ρ_T are the marginal distributions at the beginning and end of the process, respectively. The problem corresponds to a dynamic formulation of entropy-regularized optimal transport (EOT) [83]. In EOT, the optimization yields a unique optimal coupling π^* for the following objective

$$\mathrm{EOT}(\mu, \nu) \coloneqq \inf_{\pi \in \Pi(\mu,\nu)} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|^2 \, \mathrm{d}\pi(\mathbf{x}, \mathbf{y}) + \sigma^2 \mathrm{KL}(\pi \| \mu \otimes \nu), \tag{4}$$

where $\Pi(\mu, \nu)$ represents the set of couplings and $\mu \otimes \nu$ denotes the product of the marginals. In the Benamou–Brenier dynamic formulation [7, 26], one finds that characterizing the optimal $\mathbb{P}$ involves a pair of log-Schrödinger potentials, denoted in this work by $\{\chi_t\}_{t\in[0,T]}$ and $\{\eta_t\}_{t\in[0,T]}$, which together are dual objects associated with the mass-conserving density ρ_t for $\rho_0 = \mu$ and $\rho_T = \nu$.

Geometric vs. dissipative dynamics. Energy-based sampling methods—such as Hamiltonian Monte Carlo [62]—have introduced an auxiliary momentum variable to leverage the underlying symplectic geometry. Following the energy-based formulation, one can accommodate rotational particle dynamics using a broad class of stationary measures, while the corresponding Markov chains remain ergodic and time-reversible [27] up to momentum reversal (or conjugation). Let us consider a diffusion process with phase-space coordinates $(\mathbf{x}_t, \mathbf{v}_t)^\mathsf{T} \in \mathbb{R}^{2d}$ and a density $p(\mathbf{x}, \mathbf{v})$ satisfying $\log p(\mathbf{x}, \mathbf{v}) \propto -E(\mathbf{x}, \mathbf{v})$, combining Hamiltonian dynamics [11] with a dissipative Langevin diffusion

$$\begin{bmatrix} \mathrm{d}\mathbf{x}_t \\ \mathrm{d}\mathbf{v}_t \end{bmatrix} = \gamma \cdot \underbrace{\frac{\sigma^2}{4} \begin{pmatrix} -\nabla_{\mathbf{v}} \log p(\mathbf{x}_t, \mathbf{v}_t) \\ \nabla_{\mathbf{x}} \log p(\mathbf{x}_t, \mathbf{v}_t) \end{pmatrix} \mathrm{d}t}_{\text{Hamiltonian dynamics}} + \underbrace{\frac{\sigma^2}{4} \begin{pmatrix} \nabla_{\mathbf{x}} \log p(\mathbf{x}_t, \mathbf{v}_t) \\ \nabla_{\mathbf{v}} \log p(\mathbf{x}_t, \mathbf{v}_t) \end{pmatrix} \mathrm{d}t + \frac{\sigma}{\sqrt{2}} \begin{pmatrix} \mathrm{d}\mathbf{w}_t \\ \mathrm{d}\hat{\mathbf{w}}_t \end{pmatrix}}_{\text{Langevin diffusion}}, \tag{5}$$

where $\mathbf{w}_t$ and $\hat{\mathbf{w}}_t$ are independent real Wiener processes. The real-valued coefficient $\gamma > 0$ balances oscillatory dynamics and mean-reverting diffusion, providing flexibility to account for an innate symmetry while the process reaches a stationary measure with respect to the given distribution.

Our interest lies in a particular geometric extension (lifting), in which a pair of log-Schrödinger potentials induces energy-based dynamics of the SDE (5). In the special case of Eq. (1), a symmetry arises when H generates unitary evolution, which prevents the formation of a bridge. It is therefore widely accepted that transitions from one steady state to another require nondeterministic mechanisms that break the underlying geometry, such as quantum jumps [92]. On the other hand, decoherence theory in quantum mechanics [74] provides a description purely in terms of wave functions, without jumps or wave function collapse. We focus on this geometric interpretation, particularly on applying conformal geometric approaches inspired by CFT [24, 22, 73] to describe stochastic Hamiltonians.

3 Complex Schrödinger Bridges

This section develops a new bridge framework for the complex plane. We introduce complex operators for constructing SDEs and, drawing on a classical analogy, formulate the COT problem underlying this work. We then discuss a control-based dynamic formulation and computational considerations.

3.1 Complex differential operators and Langevin dynamics

We study a new geometric family of stochastic processes using complex operators. This recasts Eq. (5) as a *complexified* Langevin system [46, 43] for a potential ψ and establishes two distinct connections. The real part yields the probabilistic wave interpretation $\|\Psi_t\|^2 = \rho_t$ with $\Psi = e^{\psi/\gamma}$. The imaginary part is a natural product of diffusion that identifies ψ as an integrated complex Schrödinger potential.

Consider a complex-valued state $\mathbf{z}_t \coloneqq \mathbf{x}_t + i\mathbf{v}_t$. We first introduce the Wirtinger gradient operators, defined with respect to $\mathbf{z}$ and its complex conjugate $\bar{\mathbf{z}}$. These operators permit the construction of a differential calculus for complex variables and the expression of dynamics in complex coordinates.

Definition 1 (Multivariate Wirtinger operators)**.** Let $U \subseteq \mathbb{C}^d$ be an open set and $f : U \to \mathbb{C}$ be a smooth complex-valued function. Define multivariate Wirtinger gradient operators as

$$\begin{bmatrix} \nabla_{\mathbf{z}} f \\ \nabla_{\bar{\mathbf{z}}} f \end{bmatrix} \coloneqq \frac{1}{2}\left(\begin{pmatrix} 1 & -i \\ 1 & +i \end{pmatrix} \otimes \mathbf{I}_d\right)\begin{pmatrix} \nabla_{\mathbf{x}} f \\ \nabla_{\mathbf{v}} f \end{pmatrix}, \tag{6}$$

where $i = \sqrt{-1}$ and $\otimes$ here denotes the Kronecker product (see also Definition 7 in Appendix A).

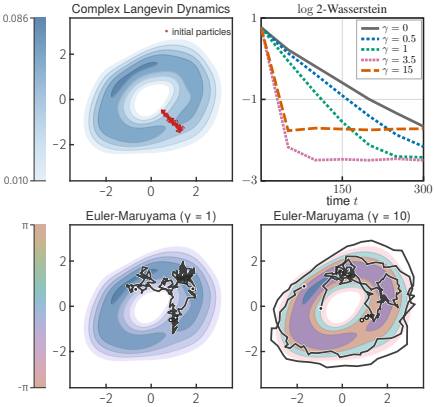

Figure 2: Diffusion models approximate a distribution with different values of γ. Their stochastic trajectories exhibit varying dynamics and coverage of the density.

Using the time-reversibility property [27], we can find a pair of SDEs that describe an equilibrium state in both time directions.

Lemma 1 (Equilibria)**.** *Consider the process $\mathbf{z}_t = \mathbf{x}_t + i\mathbf{v}_t$ defined by* (5) *and its conjugate $\bar{\mathbf{z}}_s$ with reverse time indexed by s. Consider a smooth function $\psi(\mathbf{z}, \bar{\mathbf{z}})$. The laws of two SDEs, denoted by the time-dependent density $\rho(\mathbf{z}, \bar{\mathbf{z}})$, coincide with a stationary measure for the following CVSDEs*

$$\begin{cases} \mathrm{d}Z_t = i\sigma^2 \nabla_{\bar{\mathbf{z}}} \psi(Z_t, \bar{Z}_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t & \text{(7a)} \\ \mathrm{d}\bar{Z}_s = i\sigma^2 \nabla_{\mathbf{z}} \psi(Z_s, \bar{Z}_s)\,\mathrm{d}s + \sigma\,\mathrm{d}\bar{W}_s, & \text{(7b)} \end{cases}$$

if $\psi = \frac{1}{2}(\gamma - i)\log\rho + \text{const.}$ holds for some constant in $\mathbb{C}$[1].

The lemma predicts a family of CVSDE models parameterized by $\gamma \in \mathbb{R}$. Fig. 2 demonstrates how this new degree of freedom in the dynamics affects convergence under different phase structures. This demonstrates that the complex formalization is innately suitable for encoding fundamental geometric priors.

3.2 The conformal optimal transport problem

For an endpoint pair, the additional prior from the Hamiltonian dynamics is given by $b_\gamma(\mathbf{z}, t; \mathbf{z}_T) = -\frac{i\gamma}{T-t}(\mathbf{z} - \mathbf{z}_T)$, which introduces an endpoint-conditioned, proportional rotational drift around $\mathbf{z}_T$. In this setup, we introduce a specific instance of entropy-regularized dynamical OT [89, 15] on the chosen logarithmic covering space and call this problem conformal optimal transport (COT$_\gamma$):

$$\begin{aligned} \mathrm{COT}_\gamma(\rho_0, \rho_T) &: \inf_{\pi \in \Pi(\rho_0, \rho_T)} \iint_{\mathbb{C}^d \times \mathbb{C}^d} \mathcal{C}_\gamma(\mathbf{z}_0, \mathbf{z}_T)\,\mathrm{d}\pi(\mathbf{z}_0, \mathbf{z}_T) + \sigma^2 \mathrm{KL}(\pi \| \rho_0 \otimes \rho_T), \\ \mathcal{C}_\gamma(\mathbf{z}_0, \mathbf{z}_T) &\coloneqq \inf_{\substack{u(0)=\mathbf{z}_0 \\ u(T)=\mathbf{z}_T}} \int_0^T L_\gamma(u(\tau), \dot{u}(\tau), \tau)\ \mathrm{d}\tau, \quad L_\gamma(\mathbf{z}, \mathbf{p}, t; \mathbf{z}_T) = \left\| \mathbf{p} - b_\gamma(\mathbf{z}, t; \mathbf{z}_T) \right\|^2 \end{aligned} \tag{8}$$

where the inner problem minimizes the integrated quadratic control cost over paths $u \in AC([0, T], \mathbb{C}^d)$. Minimizing the energy relative to b_γ yields the logarithmic spiral $u^*(t) = \mathbf{z}_T + (\mathbf{z}_0 - \mathbf{z}_T)(1 - t/T)^{c_\gamma}$, with γ controlling its rotation without changing the minimum control cost. If Novikov's condition holds under $\mathbb{Q}_W$ on each interval $[0, t]$ with $t < T$, applying Girsanov's theorem [67] to a pushforward measure $\mathbb{P}$, namely $\mathbb{P} : \mathrm{d}Z_t = i\mathsf{A}_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$, yields

$$\log \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}_W}\bigg|_{\mathcal{F}_t} = 2\Re\left[\int_0^t (\bar{\vartheta}_\tau)^\mathsf{T}\,\mathrm{d}W_\tau\right] - \int_0^t \left\|\vartheta_\tau\right\|^2 \mathrm{d}\tau, \qquad \vartheta_t = \frac{i\mathsf{A}_t}{\sigma} \tag{9}$$

where $\Re[\cdot]$ denotes the real part and $\mathcal{F}_t$ is the natural filtration up to time t. We can think of a pairwise mapping $(\mathcal{H}_\gamma \mathbf{z})_t = \mathbf{z}_T + \frac{h_\gamma(\log(1-t/T))}{h_0(\log(1-t/T))}(\mathbf{z}_t - \mathbf{z}_T)$ where $h_\gamma(\ell) \coloneqq e^{c_\gamma \ell}\ \ \forall \ell \in \mathbb{C}$ and $c_\gamma = 1 + i\gamma$. For the preimage $\mathbb{P} = (\mathcal{H}_{-\gamma})_\# \mathbb{P}^\gamma$, we obtain $\mathrm{KL}(\mathbb{P}^\gamma \| \mathbb{Q}_W^\gamma) = \mathrm{KL}(\mathbb{P} \| \mathbb{Q}_W)$, which subsequently admits a distinct transformation using the reference transition density and a separate time scale and metric.

The rotational factor studied in Eq. (8) has a distinct geometric interpretation as a coordinatewise *conformal* transformation. The conformal mapping preserves angular magnitudes and orientations between curves, carrying essential geometric structure. According to Lévy's characterization of

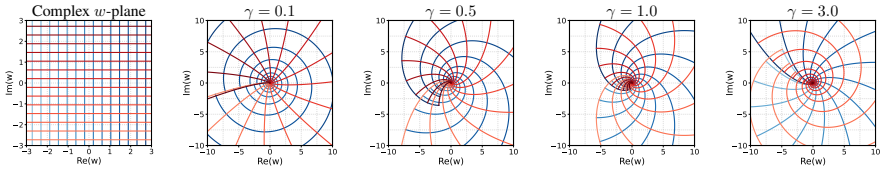

Figure 3: Conformal mappings under $h_\gamma(w)$ for varying coefficients. A Cartesian grid (normalized w-plane) is mapped biholomorphically onto a curvilinear grid, preserving right angles.

complex Brownian motion (BM) [52, 69], the pushforward $\mathbb{Q}_W^\gamma = (\mathcal{H}_\gamma)_\#(\mathbb{Q}_W)$ admits, in each uniquely normalized spiral coordinate under endpoint conditioning, a conformal local martingale that corresponds to a planar BM under its own quadratic-variation time change. Consequently, this allows us to formulate a KL objective:

$$\underset{\mathbb{P}_{0,T}^\gamma = \rho_{0,T}}{\text{minimize}}\, \mathrm{KL}(\mathbb{P}^\gamma \| \mathbb{Q}_W^\gamma), \quad \text{s.t.} \quad \mathbb{Q}_W^\gamma = (\mathcal{H}_\gamma)_\#(\mathbb{Q}_W). \tag{10}$$

As shown in Fig. 3, for small γ, the mapping is predominantly characterized by scaling, whereas rotation dominates as γ increases. As a whole, $\mathcal{H}_\gamma$ is a quasiconformal homeomorphism that generates a family of boundary-constrained paths (see Appendix B). Indeed, each time slice shows that the spiral connecting the endpoint pair exhibits self-similarity; this gives equivalent SB problems when the reference measure, admissible measures, and path space are transformed consistently.

3.3 Complex Schrödinger bridges in a mean field formulation

We identify CVSDEs that address general CSB problems. To proceed, let f be an analytic function that serves as a *base* drift for the diffusion bridge. By adopting an *ansatz* for the generalized PDE system that retains the γ-rotational prior for Eq. (10), we formulate the paired CVSDEs

$$\mathrm{d}Z_t = i\big(f(Z_t,t) + \sigma^2 \nabla_{\bar{\mathbf{z}}} \psi(Z_t, \bar{Z}_t, t)\big)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t \tag{11a}$$

$$\mathrm{d}\bar{Z}_s = i\big(\bar{f}(Z_s, T\text{-}s) + \sigma^2 \nabla_{\mathbf{z}} \psi(Z_s, \bar{Z}_s, T\text{-}s)\big)\,\mathrm{d}s + \sigma\,\mathrm{d}\bar{W}_s, \tag{11b}$$

in a symmetric form with respect to alternative forward and reverse time indices t and s. The following proposition verifies the connection between an SB model and Eq. (2), while Φ evolves under a twisted Hamiltonian function in heat-type PDEs on a measurable subset of $\mathbb{C}^d$.

Proposition 1 (CSB-PDEs)**.** *Suppose that a function* $H(\cdot,\cdot,\nabla_{\mathbf{z}}\psi, \nabla_{\bar{\mathbf{z}}}\psi, \rho) = \sigma^2\langle \nabla_{\mathbf{z}}\psi, \nabla_{\mathbf{z}}\bar{\psi}\rangle + \nabla_{\mathbf{z}}\psi^\mathsf{T} f - \nabla_{\bar{\mathbf{z}}}\psi^\mathsf{T}\bar{f} + \frac{i\gamma\sigma^2}{2}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}$ *acts as a Hamiltonian in Eq.* (24)*. Then,* $\Phi = \exp(i\psi/c_\gamma)$ *satisfies*

$$\begin{cases} \frac{\partial \Phi(\mathbf{z},\bar{\mathbf{z}},t)}{\partial t} = -i\nabla_{\mathbf{z}}\Phi^\mathsf{T} f + i\nabla_{\bar{\mathbf{z}}}\Phi^\mathsf{T}\bar{f} - \frac{\sigma^2}{4}c_\gamma \Delta\Phi + \frac{i\gamma\sigma^2}{2c_\gamma}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\Phi + c_\gamma^{-1}F\Phi, \\ \frac{\partial \hat{\Phi}(\mathbf{z},\bar{\mathbf{z}},t)}{\partial t} = -i\nabla_{\mathbf{z}}\cdot(\hat{\Phi}f) + i\nabla_{\bar{\mathbf{z}}}\cdot(\hat{\Phi}\bar{f}) + \frac{\sigma^2}{4}c_\gamma\Delta\hat{\Phi} - \frac{i\gamma\sigma^2}{2c_\gamma}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\hat{\Phi} - c_\gamma^{-1}F\hat{\Phi}, \end{cases} \quad \textit{s.t.} \quad \begin{matrix} \Phi_0\hat{\Phi}_0 = \rho_0 \\ \Phi_T\hat{\Phi}_T = \rho_T \end{matrix}$$

where Δ *denotes the Laplacian operator,* $\hat{\Phi} = \rho/\Phi$*, and* $F(\cdot,\cdot,\rho(\cdot,\cdot,t))$ *denotes a real-valued function of state-dependent penalties and distributional interaction costs.*

The complex Hopf—Cole transform $(\Phi, \hat{\Phi})$ preserves mass under dynamics governed by the Hamiltonian and a heat term $c_\gamma\Delta\psi$. This allow us to derive forward and backward SDEs (FBSDEs) with mean field costs akin to the SB-FBSDEs [13], yet generalized to diffusion models of the COT problem.

Theorem 1 (General SB-CVSDEs)**.** *Suppose that* $Y_t := \chi(\mathbf{z}_t, \bar{\mathbf{z}}_t, t)$ *and* $\widehat{Y}_t := -\eta(\mathbf{z}_t, \bar{\mathbf{z}}_t, t)$*, with analytic* f *satisfying* $\nabla_{\mathbf{z}}\cdot f_t = \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}_t$*. For* $\Phi = e^{i\psi/c_\gamma}$ *and* $\hat{\Phi} = e^{i\eta/c_\gamma}$*, the CSB-PDEs give forward and backward SDEs with respect to Eq.* (11a)*,* $t \in [0, T)$*:*

$$\textit{SB CVSDEs: wrt (11a)} \quad \begin{cases} \mathrm{d}Z_t = i\big(f_t + \sigma^2\nabla_{\bar{\mathbf{z}}}\psi_t(Z_t,\bar{Z}_t,t)\big)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, & \text{(12a)} \\ \mathrm{d}Y_t = -i\big(\|v_t\|^2 + F_t + \gamma\sigma\nabla_{\mathbf{z}}\cdot\check{v}_t - \frac{i\gamma}{2}\|\hat{v}_t + \check{v}_t\|^2\big)\,\mathrm{d}t + v_t^\mathsf{T}\,\mathrm{d}W_t + \hat{v}_t^\mathsf{T}\,\mathrm{d}\bar{W}_t, & \text{(12b)} \\ \mathrm{d}\widehat{Y}_t = -i\big(\|v_t\|^2 + F_t - \gamma\sigma\nabla_{\mathbf{z}}\cdot\hat{v}_t - (1+\frac{i\gamma}{2})\|\hat{v}_t + \check{v}_t\|^2\big)\,\mathrm{d}t - & \text{(12c)} \\ \qquad \big(\sigma\nabla_{\mathbf{z}}\cdot\check{v}_t + \sigma\nabla_{\bar{\mathbf{z}}}\cdot\bar{v}_t\big)\,\mathrm{d}t - \bar{v}_t^\mathsf{T}\,\mathrm{d}W_t - \check{v}_t^\mathsf{T}\,\mathrm{d}\bar{W}_t, \end{cases}$$

where $\psi = \chi + \gamma\log\rho$*,* $\chi + \eta = -i\log\rho$*, and* $(v_t, \hat{v}_t, \check{v}_t, \bar{v}_t)$ *denote the Wirtinger gradients*

$$v_t := \sigma\nabla_{\mathbf{z}}\chi(\mathbf{z}_t,\bar{\mathbf{z}}_t,t), \quad \hat{v}_t := \sigma\nabla_{\bar{\mathbf{z}}}\chi(\mathbf{z}_t,\bar{\mathbf{z}}_t,t), \quad \check{v}_t := \sigma\nabla_{\bar{\mathbf{z}}}\eta(\mathbf{z}_t,\bar{\mathbf{z}}_t,t), \quad \bar{v}_t := \sigma\nabla_{\mathbf{z}}\eta(\mathbf{z}_t,\bar{\mathbf{z}}_t,t).$$

Conversely, the reversed time evolution for $s = T - t$ *with respect to Eq.* (11b) *follows symmetrically (Appendix C.3). The law of* Z_t *matches that of* $\bar{Z}_s$ *under complex conjugation.*

Theorem 1 characterizes the general dynamics, where the potential induces a drift that follows the transformed reference along its gradient. For actual learning with nonlinear neural networks, one may employ an alternating optimization scheme based on changes of measure in opposite directions, or iterative proportional fitting (IPF) [28, 66]. In its imaginary channel, a derived algorithm updates parameters using bidirectional errors, with the optimum achieved when both fitted path laws agree.

Theorem 2 (Iterative proportional fitting)**.** *Consider the CSB problem without external costs. For parameters (θ, ϕ), let the IPF estimates be defined using backward propagated changes of measure*

$$\mathcal{E}^{+}_{T-\tau}(\phi;\theta) = i\int_0^\tau \mathbb{E}\Big[\|\check{v}^{\phi}_s\|^2 - \gamma\sigma\nabla_{\mathbf{z}}\cdot\check{v}^{\phi}_s - \big(1+\tfrac{i\gamma}{2}\big)\|v^{\theta}_s+\bar{v}^{\phi}_s\|^2 - i\sigma\big(\nabla_{\mathbf{z}}\cdot\hat{v}^{\theta}_s + \nabla_{\bar{\mathbf{z}}}\cdot v^{\theta}_s\big)\Big]\mathrm{d}s,$$

$$\mathcal{E}^{-}_{\tau}(\theta;\phi) = -i\int_0^\tau \mathbb{E}\Big[\|v^{\theta}_t\|^2 - \gamma\sigma\nabla_{\mathbf{z}}\cdot\hat{v}^{\theta}_t - \big(1+\tfrac{i\gamma}{2}\big)\|\hat{v}^{\theta}_t+\check{v}^{\phi}_t\|^2 - i\sigma\big(\nabla_{\mathbf{z}}\cdot\check{v}^{\phi}_t + \nabla_{\bar{\mathbf{z}}}\cdot\bar{v}^{\phi}_t\big)\Big]\,\mathrm{d}t,$$

where the signs $+$ and $-$ denote integration with respect to s and t, respectively. The signed imaginary parts of the accumulated errors agree with the corresponding KL functionals up to additive constants,

$$\mathrm{KL}(\mathbb{Q}_\phi\|\,\mathbb{P}_\theta) = -\Im\big[\mathcal{E}^{+}_0(\phi;\theta)\big] + C_\phi \qquad \textit{and} \qquad \mathrm{KL}(\mathbb{P}_\theta\|\,\mathbb{Q}_\phi) = \Im\big[\mathcal{E}^{-}_T(\theta;\phi)\big] + C_\theta,$$

where $\mathbb{P}_\theta$ and $\mathbb{Q}_\phi$ are the path measures obtained from CVSDEs (11a) *and* (11b)*, respectively. Here, C_ϕ is independent of θ, and C_θ is independent of ϕ, derived in Appendix C.4.*

The IPF theorem suggests that the proposed family of complex-valued bridges can be trained.

4 Algorithmic Considerations

We derive two primary training approaches from our theoretical framework. The first is a general CSB that utilizes IPF updates derived from SB-CVSDEs, whereas the second is a conditional CSB derived from COT. Both derived algorithms are distinguished by their capabilities and scalability.

Algorithm 1 General CSB: biproportional projection.

Input: networks $(Y_\theta, \widehat{Y}_\phi, \hat{v}_\theta, \bar{v}_\phi)$ and $\lambda_{\text{IPF}}, \lambda_{\text{TD}} \in \mathbb{R}^+$.
while (θ, ϕ) have not converged **do**
  for $k = 1$ **to** K **do** ▷ Forward KL projection
    $\mathcal{L}^{+}(\phi;\phi',\theta') \leftarrow \lambda_{\text{IPF}}\mathcal{L}^{+}_{\text{IPF}} + \lambda_{\text{TD}}\mathcal{L}^{+}_{\text{TD}} + \mathcal{L}^{+}_0 + \mathcal{L}^{+}_{\text{Drift}}$
    Update ϕ with gradient $\hat{g}^{+}_\phi \approx \mathbb{E}[\nabla_\phi \mathcal{L}^{+}_\phi]$.
    $\phi' \leftarrow \texttt{EMA}(\phi, \phi')$
  end for
  for $k = 1$ **to** K **do** ▷ Reverse projection
    $\mathcal{L}^{-}(\theta;\theta',\phi') \leftarrow \lambda_{\text{IPF}}\mathcal{L}^{-}_{\text{IPF}} + \lambda_{\text{TD}}\mathcal{L}^{-}_{\text{TD}} + \mathcal{L}^{-}_{T} + \mathcal{L}^{-}_{\text{Drift}}$
    Update θ with gradient $\hat{g}^{-}_\theta \approx \nabla_\theta \mathcal{L}^{-}_\theta$.
    $\theta' \leftarrow \texttt{EMA}(\theta, \theta')$
  end for
end while
return trained potentials and drifts $(Y_\theta, \widehat{Y}_\phi, \hat{v}_\theta, \bar{v}_\phi)$.

Algorithm 2 Conditional CSB.

Input: a model ϵ_θ, $p_0(\cdot)$, and $p_T(\cdot|\mathbf{x}_0)$.
repeat
  $\mathbf{z}_0 \leftarrow \mathbf{x}_0 + i\mathbf{v}_0$, $\mathbf{z}_T \leftarrow \mathbf{x}_T + i\mathbf{v}_T$
  $\mathbf{z}_t \sim q(\,\cdot\,|\,\mathbf{z}_0, \mathbf{z}_T)$ according to (15).
  Update ϵ_θ using the gradient of (16).
until convergence

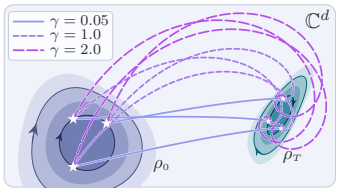

Figure 4: Pairwise posterior flows.

4.1 Training general SB-CVSDEs via neural networks

Algorithm 1 presents a straightforward realization of the general theory. To solve the SB-CVSDEs, we parameterize independent potentials $(Y^\theta_t, \widehat{Y}^\phi_s)$ [53]. As a direct consequence of Theorem 2, we propose IPF losses, namely $\mathcal{L}^{+}_{\text{IPF}}(\phi)$ and $\mathcal{L}^{-}_{\text{IPF}}(\theta)$, to optimize likelihoods of the path measures with respect to KL projections. As boundary conditions, we apply score losses for marginals using the log-density estimate $i(Y^\theta - \widehat{Y}^\phi)$: $\mathcal{L}^{+}_0(\phi) = \mathcal{L}_{\text{Score}}(\phi;\rho_0)$ and $\mathcal{L}^{-}_T(\theta) = \mathcal{L}_{\text{Score}}(\theta;\rho_T)$. For each time direction, the algorithm calculates the temporal differences (TDs) for small intervals $(\delta t, \delta s)$ as

$$\mathrm{TD}^{+}_s(\phi;\theta') \approx b(Y^{\theta'}_s)\delta s + (v^{\theta'}_s)^\intercal \delta W_s + (\hat{v}^{\theta'}_s)^\intercal \delta\bar{W}_s, \tag{13a}$$

$$\mathrm{TD}^{-}_t(\theta;\phi') \approx b(\widehat{Y}^{\phi'}_t)\delta t - (\bar{v}^{\phi'}_t)^\intercal \delta W_t - (\check{v}^{\phi'}_t)^\intercal \delta\bar{W}_t, \tag{13b}$$

where $b(Y^{\theta'}_s)$ and $b(\widehat{Y}^{\phi'}_t)$ denote the target Itô drift coefficients in Eq. (43c) and Eq. (12c), respectively, and $\delta W_t \sim \mathcal{CN}(0, \delta t\,\mathbf{I}_d)$ and $\delta W_s \sim \mathcal{CN}(0, \delta s\,\mathbf{I}_d)$ denote the standard approximation of complex

Wiener differentials, with barred increments given by complex conjugation. Suppose that R_s^+ and R_t^- are single-step or multistep return estimates [78] of Y_θ and $\widehat{Y}_\phi$ constructed from $(\mathrm{TD}_s^+, \mathrm{TD}_t^-)$, respectively, including the ground-truth boundary or bootstrap values. We propose regression losses

$$\mathcal{L}_{\mathrm{TD}}^+(\phi;\phi',\theta') = \mathbb{E}\big\|\widehat{Y}_\phi(\mathbf{z}_s,\bar{\mathbf{z}}_s,T-s) - \bar{G}_s^+\big\|, \qquad \mathcal{L}_{\mathrm{TD}}^-(\theta;\theta',\phi') = \mathbb{E}\big\|Y_\theta(\mathbf{z}_t,\bar{\mathbf{z}}_t,t) - G_t^-\big\|.$$

Using our Hopf–Cole relation $\widehat{Y} = Y + i\log\rho$ in the corresponding parameterization, we set $\bar{G}_s^+ = R_s^+ + i\log\rho_s$ and $G_t^- = R_t^- - i\log\rho_t$, with the density evaluated at the return's starting state and time. Jointly minimizing the IPF, TD, and marginal score losses yields a unique solution to SB-CVSDEs, as established in Theorem 1. We adopt exponential moving averages (EMAs) for the potential targets for additional stability. Additionally, we minimize the drift-matching losses $\mathcal{L}_{\mathrm{Drift}}^{\pm}$ to enable fast SDE sampling and efficient iterative updates in practice [53].

4.2 Analytic posterior for high-dimensional image translation and matching

Algorithm 2 demonstrates a conditional, image-to-image CSB (I^{2}CSB) that can be solved analytically from known boundary pairs using conformal geodesics. For a fixed endpoint pair $(\mathbf{z}_0, \mathbf{z}_T)$ and the posterior density [55] $q(\bar{\mathbf{z}}_t \mid \bar{\mathbf{z}}_T, \bar{\mathbf{z}}_0)$, the reverse-drift target of our CSB at $0 < t < T$ is given by

$$i\hat{\sigma}_s^2\nabla_{\mathbf{z}}\psi_\theta(\mathbf{z}_t,\bar{\mathbf{z}}_t,t|\mathbf{z}_0) = -\bar{c}_\gamma\hat{\sigma}_s^2\nabla_{\mathbf{z}_t}\log p_{\mathrm{ref}}(\mathbf{z}_T, T \mid \mathbf{z}_t, t) + \hat{\sigma}_s^2\nabla_{\mathbf{z}_t}\log q(\bar{\mathbf{z}}_t \mid \bar{\mathbf{z}}_T, \bar{\mathbf{z}}_0), \tag{14}$$

where $s = T - t$ and p_{ref} denotes the transition density of the drift-free reference Brownian motion. Without diffusion, the trajectory follows the logarithmic spiral predicted from h_γ. We utilize the following proposition for computing an analytic posterior using the conjugate conformal geodesic.

Proposition 2 (Analytic posterior)**.** *Let $\gamma \in \mathbb{R}$, and let $\sigma_t > 0$ be scheduled continuously on $[0, T]$. Given a complex-valued boundary pair $(\bar{\mathbf{z}}_0, \bar{\mathbf{z}}_T) \in \mathbb{C}^d \times \mathbb{C}^d$ and a time-dependent diffusion coefficient, the COT$_\gamma$ posterior admits an analytic form based on a conjugate conformal geodesic:*

$$q(\bar{\mathbf{z}}_s|\bar{\mathbf{z}}_T,\bar{\mathbf{z}}_0) = \mathcal{CN}(\bar{\mathbf{z}}_s|\bar{\mu}_s,\Sigma_s), \quad \bar{\mu}_s = \bar{\mathbf{z}}_T + (\bar{\mathbf{z}}_0 - \bar{\mathbf{z}}_T)\left(\frac{\hat{\omega}_s^2}{\hat{\omega}_s^2+\omega_s^2}\right)^{\bar{c}_\gamma}, \ \Sigma_s = \left(\frac{\omega_s^2\hat{\omega}_s^2}{\omega_s^2+\hat{\omega}_s^2}\right)\mathbf{I}_d, \tag{15}$$

where $\hat{\omega}_s^2 = \int_0^s \hat{\sigma}_\tau^2\,\mathrm{d}\tau$, $\omega_s^2 = \int_s^T \hat{\sigma}_\tau^2\,\mathrm{d}\tau$ with $\hat{\sigma}_\tau^2 = \sigma_{T-\tau}^2$, and $\bar{c}_\gamma = 1 - i\gamma$.

Fig. 4 shows that, with the rotational drift controlled by γ, the diffusion model will exhibit spiraling motion in the plane, for example by accelerating early and then rotating around the true samples. This is oftentimes useful in practice since we can reuse most of the score-based learning structure. For our approach, we consider a loss function

$$\mathcal{L}(\theta) = \mathbb{E}\left\|\mathtt{cpx}\big(\epsilon_\theta(\mathbf{x}_t,\mathbf{v}_t,T\text{-}s)\big) - \mathtt{conj}\left(\frac{\bar{\mathbf{z}}_s - \bar{\mathbf{z}}_0}{\hat{\omega}_s}\right)\right\|^2, \tag{16}$$

where $\mathtt{cpx}(\cdot)$ denotes the basic operation of merging each corresponding pair of channels $\mathbf{x}_j$ and $\mathbf{v}_j$ into a single complex channel $\mathbf{z}_j$ and $\mathtt{conj}(\cdot)$ denotes complex conjugation within each channel. With the model trained to predict the exact expected value, we perform DDPM-style generation [34, 55] with $\mathcal{CN}(\bar{\mathbf{z}}_n \mid \bar{\mu}_{n|n+1}, \Sigma_{n|n+1})$ for $1 \le n \le N-2$ on a proper discretization into N steps, which often improves robustness on image tasks. We defer the implementation details to Section D.

5 Experimental Results

We delineate our objectives as follows. ① We aimed to validate our theory across a wide range of domains in terms of generality, scalability, and the incorporation of explicit priors regarding complex values. ② We sought to analyze the derived algorithms and thoroughly understand their strengths and weaknesses compared with those of conventional real-valued models. ③ We aimed to apply our method to data from realistic complex-valued datasets to observe the effectiveness of complex-valued data processing. To achieve these goals, we evaluated our algorithm across diverse problem settings, including classical control, image translation, and applications with complex-valued data.

5.1 General CSB for two-dimensional control tasks.

We trained MLPs with the general CSB algorithm for four two-dimensional (2D) control benchmarks involving non-Gaussian initial marginals or a subtask of obstacle navigation. The mean-field particles were sampled from task-specific distributions and assigned random momentum drawn from a scaled normal distribution. We trained full potential and drift neural networks according to Algorithm 1.

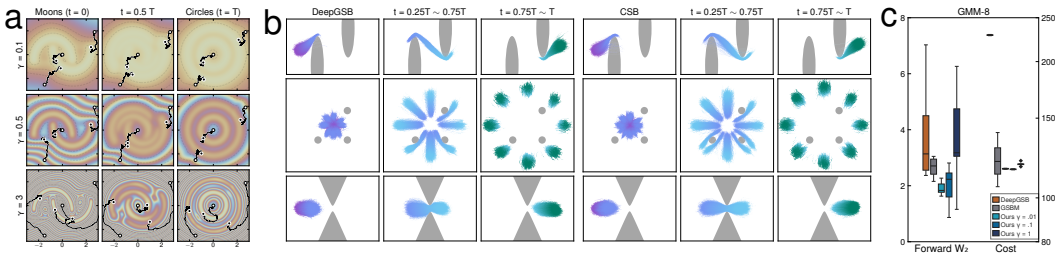

Figure 5: Results on two-dimensional control problems. (a) Visualization of the potential function and Langevin dynamics for 3 seconds in the Moons-Circles bridge. (b) Learned CSB dynamics for S-Tunnel, GMM-8, and V-Neck. (c) Statistics on marginal fitting and total costs (5 different seeds).

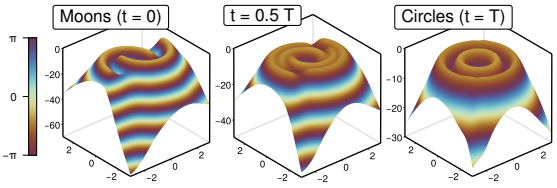

Figure 6: Learned complex-valued log-potential functions at three distinct time points ($\gamma = 0.5$).

In Fig. 5 (a), our method successfully learned the Moons-Circles bridge, where color stripes depict energy levels across the phase structure. Fig. 6 further displays the amplitude of $\frac{1}{2}c_\gamma \log \rho_t$ on 3D plots at zero momentum. For time indices $t \in \{0, 0.5T, T\}$, we ran Langevin SDEs at each checkpoint. As γ increases, the trajectories follow the internal energy landscape rather than arbitrary paths, demonstrating the correctness of our claim.

Fig. 5 (b) and (c) show qualitative and quantitative results on the S-Tunnel, GMM-8, and V-Neck benchmarks, where mean-field particles must avoid fixed obstacles while maintaining a diffusion bridge. We compared against two baselines: DeepGSB on the equivalent state space ($\mathbb{R}^{2d}$), and the standard GSBM algorithm ($\mathbb{R}^d$) using trainable splines for SB matching. Across all tested values of γ, CSB achieved lower mean endpoint W_2 than DeepGSB on all three benchmarks. CSB was competitive with GSBM on S-Tunnel and achieved the lowest mean endpoint 2-Wasserstein distance on GMM-8, where both baselines had their largest endpoint errors. CSB also achieved the lowest mean total trajectory cost in all three 2D control environments.



Figure 7: Image-to-image translation results on RGB datasets. Left: unpaired MNIST generation. Middle: CIFAR-10 and ImageNet256 inpainting (25%). Right: CIFAR-10 inpainting FID scores.

5.2 Image-to-image translation

Paired and unpaired image generation. Using I^{2}CSB derived from Algorithm 2, we conducted unpaired and paired image-to-image translation initialized with Gaussian noise and corrupted inputs, respectively, alongside a Gaussian momentum prior. Fig. 7 demonstrates that our analytic posterior maintains superior stability across a wide hyperparameter window. Consistent with the trend observed for general CSB, the FID scores on CIFAR-10 also exhibited U-shaped dependence on the hyperparameter γ (see Table 9 in the appendix for the MNIST case). Because FID reflects domain discrepancy on the latent space under Gaussian assumptions, this empirical minimum confirms that per-step tuning of γ enables more faithful diffusion trajectories covering target modalities.

ImageNet256. We extended our method to the ImageNet256 inpainting tasks, featuring higher dimensionality and various corruption methods. We fine-tuned the pretrained *FreeForm1020*, *FreeForm2030*, and *Center* checkpoints from [55]. We replaced the training posterior with one based on the conformal geodesic. After 45K training iterations for each $\gamma \in \{10^{-4}, 10^{-3}, \ldots, 10^{1}\}$, the I^{2}CSB algorithm was able to recover the performance of I^{2}SB in terms of the FID-10K score [72]. This verifies that our method is applicable to diffusion bridge workflows involving RGB images.

Table 1: Magnitude FID and SSIM (structural similarity index measure) metrics on the paired 64×64 synthetic brain MRI validation dataset.

Method	FID ↓	SSIM ↑
DSBM [76]	23.1730	0.6126
SB Flow [19]	21.4061	0.6312
I^2SB [55]	15.2626	0.6597
Ours (best)	**12.0135**	**0.7139**

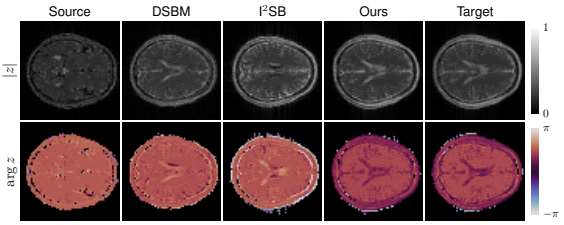

Figure 8: Complex-valued MRI generation: magnitude (top) and phase (bottom) representations.

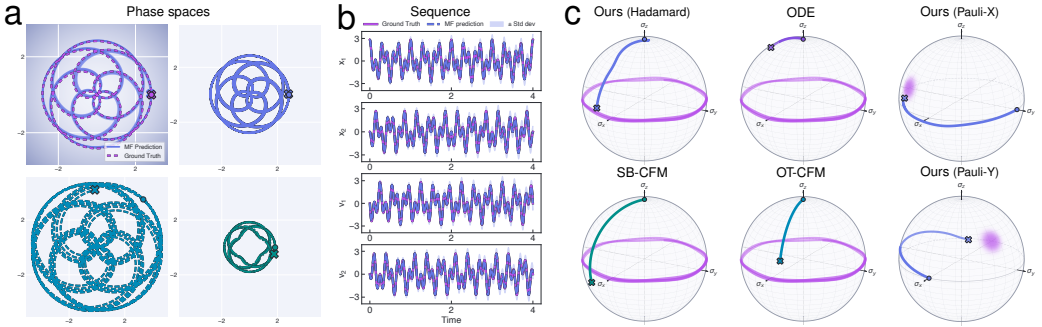

Figure 9: Complex SB with a synthetic cQED dataset. (a) Physical simulation and prediction for a harmonic oscillator with fluctuation noise. (b) For a long time horizon, our SB representation was fitted to the optimal state. (c) 3D qubit logic gate simulation (averaging over 100 particles).

MRI. We evaluated our method on simulated brain MRI slices at 64×64 resolution, with 24 mm source–target separation. Our evaluation in Table 1 confirmed that our method achieved high-fidelity translation. Fig. 8 compares translation results across DSBM [76], I^2SB, and I^2CSB, illustrating qualitative improvements in both magnitude and phase reconstruction. Quantitative performance gains in FID and magnitude structural similarity index measure (SSIM) are summarized in Table 1, where Ours (best) uses $\gamma = 0.1$; full results are reported in Table 10 of Appendix E. This confirms that our method holds a distinct advantage in handling explicit complex-valued data.

5.3 Circuit quantum electrodynamics simulation

Closed systems. We used a quantum physics simulator [58] to acquire trajectories of a harmonic oscillator in a system with $d = 2$ modes. In this simulation, the system dynamics are governed by a Hamiltonian featuring local cavity detuning and a hopping interaction between the two modes:

$$\hat{H}_{\text{osc}} = D(\hat{a}_1^\dagger \hat{a}_1) + D(\hat{a}_2^\dagger \hat{a}_2) + J(\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_1 \hat{a}_2^\dagger), \tag{17}$$

where $\hat{a}_j$ and $\hat{a}_j^\dagger$ are the bosonic annihilation and creation operators for each mode $j \in \{1, 2\}$; D denotes a detuning parameter; J denotes inter-mode coupling strength. At each step, we extracted expectation values for the position and momentum quadratures as an observation with added fluctuation noise, using the quadrature definitions $\hat{x}_j = (\hat{a}_j + \hat{a}_j^\dagger)/2$ and $\hat{p}_j = (\hat{a}_j - \hat{a}_j^\dagger)/(2i)$.

Fig. 9 (a) and Fig. 9 (b) show that particles' averaged trajectories (blue) given by the Wirtinger gradient of a potential neural network with 4 hidden layers successfully tracked the target trajectory (violet) for a long time period. The teal and green curves in Fig. 9 (a) show generalization to unseen high- and low-energy trajectories with different starting points; the model successfully generated harmonic motion. These results verify our hypothesis regarding a bosonic coherent state.

3D qubit gate transformation. We designed a physics-informed OT setup in this simulation problem domain. We expanded Eq. (17) to approximate the dynamics of a hybrid tripartite cQED system. We designed a three-mode harmonic cavity coupled to a three-mode Josephson junction (JJ), which in turn interacted with a two-level qubit [61]. The system evolution is governed by a total Hamiltonian formulated in the rotating frame: $\hat{H} = \hat{H}_{\text{osc}} + \hat{H}_{\text{jj}} + \hat{H}_{\text{qubit}} + \hat{H}_{\text{int}}$. We structured the simulations as source-to-target mapping tasks. Each algorithm was provided with learnable physics heads, *i.e.* learnable differentials for the JJ operator and ground-truth qubit projection operators.

For benchmarks, we considered a neural ODE [14] model with RK4 integration and conditional flow matching (CFM) [82] methods based on SB and OT solutions. When training the CFM variants and our model, we configured the methods to use mini-batch OT with cost matrices based on geodesic priors [82]. In particular, we utilized the conformal geodesic calculation to approximate ground-truth pairs. Fig. 9 (c) shows that, for an unseen energy level (doubling of the expected photon number from 1.0 to 2.0), our model successfully transformed one qubit state into another, mimicking a gate operation through the transport of particles in the oscillator. As shown in Table 2, our approach outperformed the benchmarks in seven of the nine settings for cQED at unseen energies.

Table 2: Qubit OT validation at higher, unseen energies for Hadamard and Pauli gates (5 runs).

Hadamard	L_2 (↓)	EMD (↓)	SWD (↓)
ODE (RK4)	$1.885_{\pm 0.20}$	$3.702_{\pm 0.61}$	$1.306_{\pm 0.25}$
OT-CFM	$1.910_{\pm 0.06}$	$3.653_{\pm 0.18}$	$1.094_{\pm 0.08}$
SB-CFM	$1.888_{\pm 0.063}$	$3.621_{\pm 0.24}$	$1.164_{\pm 0.08}$
Ours	$\mathbf{0.998_{\pm 0.61}}$	$\mathbf{1.378_{\pm 1.18}}$	$\mathbf{0.601_{\pm 0.33}}$
Pauli-X	L_2 (↓)	EMD (↓)	SWD (↓)
ODE (RK4)	$1.250_{\pm 0.45}$	$1.850_{\pm 1.41}$	$0.753_{\pm 0.28}$
OT-CFM	$\mathbf{0.557_{\pm 0.54}}$	$0.610_{\pm 0.90}$	$\mathbf{0.359_{\pm 0.318}}$
SB-CFM	$1.544_{\pm 0.107}$	$2.404_{\pm 0.36}$	$0.710_{\pm 0.40}$
Ours	$0.584_{\pm 0.25}$	$\mathbf{0.538_{\pm 0.43}}$	$0.393_{\pm 0.167}$
Pauli-Y	L_2 (↓)	EMD (↓)	SWD (↓)
ODE (RK4)	$1.232_{\pm 0.66}$	$1.863_{\pm 1.34}$	$0.738_{\pm 0.38}$
OT-CFM	$1.039_{\pm 0.58}$	$1.315_{\pm 0.11}$	$0.709_{\pm 0.52}$
SB-CFM	$1.213_{\pm 0.61}$	$1.641_{\pm 1.08}$	$0.671_{\pm 0.40}$
Ours	$\mathbf{0.741_{\pm 0.25}}$	$\mathbf{0.838_{\pm 0.35}}$	$\mathbf{0.601_{\pm 0.33}}$

6 Related Work

Diffusion bridges [88, 20] have gained attention in recent years, with connections to score and flow matching [55, 76]. SBs were subsequently applied to jump measures [90, 21, 91], second-order dynamics [80], and structured data [36], as well as extended to Riemannian manifolds [81]. Concurrently, Noble et al. [64] studied reference processes modified by Feynman–Kac potentials, while Mahmood and Akhtar [56] applied SBs to rotating $\mathrm{SO}(2)$ geometries to control angular distributions. In contrast, we explore complex-valued dynamics through complex stochastic calculus and conformal geometry [24, 22], providing a new analytic posterior for downstream tasks.

Our work also connects to complex optimal transport [68], a foundational problem in complex geometry. Through generalized dual structures [65, 33], diffusion formulations attain a high level of abstraction and have been applied to topological structures [86]. Studies on complex-valued neural networks [31, 49, 32] have considered both holomorphic and non-holomorphic network architectures and their backpropagation. Our Schrödinger potentials are complex-valued functions of $\mathbf{z}$ and $\bar{\mathbf{z}}$ that serve as dual parameters of distributions and need not be holomorphic.

The construction of complex-valued fractional Brownian motion [42] and recent fractional diffusion bridge models [63] suggest extensions to complex-valued bridges with non-Markovian dynamics, which may be useful for modeling chaotic systems. Quantum Schrödinger bridges (QSBs) [60] connect density matrices using Kraus maps in a noncommutative operator setting. Furthermore, numerical methods for the Lindblad equation explicitly preserve complete positivity and trace conservation [3]. These works address density matrix evolution within quantum field theory, whereas our cQED experiments investigate diffusion bridge trajectories for bosonic states.

7 Conclusion

In this paper, we presented complex Schrödinger bridges, a novel diffusion bridge formulation for solving the SBP in complex-valued domains. The proposed framework successfully modeled stochastic Hamiltonian physics and medical imaging, and also was able to perform existing diffusion-based generative workflows with improved stability. Our findings show that navigating complex spaces effectively requires structured geometric priors. We anticipate that our theoretical and empirical results will contribute to the broader field of AI for science (AI4Science).

Limitations and future work. Our work has the potential to broaden algorithmic capabilities for solving various OT problems. Nevertheless, the use of complex-valued arguments might not always be the optimal choice, especially if the domain lacks a general energy-based structure. Additionally, complex-valued neural architectures based on non-holomorphic functions remain underexplored, and incorporating them might have further consolidated our findings. Conformality itself does not guarantee the existence of bridges in the complex plane. The underlying complex structure becomes rigid as the dimension grows; thus, the mappings for OT must eventually be quasiconformal. In future work, we will extend this line of research to OT on general manifolds, which provide the geometric formulation of complex structures [8], as well as to the matrix OT problem [16].

Acknowledgments

The authors would like to thank the anonymous reviewers for their helpful discussion and comments. T. Birdal was supported by a UKRI Future Leaders Fellowship [grant number MR/Y018818/1]. The authors acknowledge support from the UK AI Research Resource (AIRR Isambard AI) through grant 0251-4584-0945-1 - TopoFound.

References

[1] Ahlfors, L. V. (2006). *Lectures on quasiconformal mappings*, volume 38. American Mathematical Society.

[2] Albergo, M., Boffi, N. M., and Vanden-Eijnden, E. (2025). Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 26(209):1–80.

[3] Appelö, D. and Cheng, Y. (2025). Kraus is king: High-order completely positive and trace preserving (CPTP) low rank method for the Lindblad master equation. *Journal of Computational Physics*, 534:114036.

[4] Arnol'd, V. I. (1989). *Mathematical methods of classical mechanics*, volume 60. Springer.

[5] Ash, R. B. (2014). *Complex variables*. Academic Press.

[6] Astala, K., Iwaniec, T., Prause, I., and Saksman, E. (2015). Bilipschitz and quasiconformal rotation, stretching and multifractal spectra. *Publications mathématiques de l'IHÉS*, 121:113–154.

[7] Benamou, J.-D. and Brenier, Y. (2000). A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393.

[8] Berman, R. J. and Nyström, D. W. (2014). Complex optimal transport and the pluripotential theory of Kähler-Ricci solitons. *arXiv preprint arXiv:1401.8264*.

[9] Brown, J. W. and Churchill, R. V. (2009). *Complex variables and applications*. McGraw-Hill,.

[10] Buckdahn, R., Li, J., Peng, S., and Rainer, C. (2017). Mean-field stochastic differential equations and associated PDEs. *The Annals of Probability*, 45(2):824 – 878.

[11] Burylko, O., Mielke, A., Wolfrum, M., and Yanchuk, S. (2018). Coexistence of hamiltonian-like and dissipative dynamics in rings of coupled phase oscillators with skew-symmetric coupling. *SIAM Journal on Applied Dynamical Systems*, 17(3):2076–2105.

[12] Castillo-Passi, C., Coronado, R., Varela-Mattatall, G., Alberola-López, C., Botnar, R., and Irarrazaval, P. (2023). KomaMRI.jl: An open-source framework for general MRI simulations with GPU acceleration. *Magnetic resonance in medicine*, 90(1):329–342.

[13] Chen, T., Liu, G., and Theodorou, E. A. (2022). Likelihood training of Schrödinger bridge using forward-backward SDEs theory. In *10th International Conference on Learning Representations*.

[14] Chen, T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. (2018). Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, pages 6572–6583.

[15] Chen, Y., Georgiou, T. T., and Pavon, M. (2016). Optimal transport over a linear dynamical system. *IEEE Transactions on Automatic Control*, 62(5):2137–2152.

[16] Chen, Y., Georgiou, T. T., and Tannenbaum, A. (2017). Matrix optimal mass transport: A quantum mechanical approach. *IEEE Transactions on Automatic Control*, 63(8):2612–2619.

[17] Cole, J. D. (1951). On a quasi-linear parabolic equation occurring in aerodynamics. *Quarterly of applied mathematics*, 9(3):225–236.

[18] Conway, J. B. (2012). *Functions of one complex variable II*. Springer Science & Business Media.

[19] De Bortoli, V., Korshunova, I., Mnih, A., and Doucet, A. (2024). Schrödinger bridge flow for unpaired data translation. *Advances in Neural Information Processing Systems*, 37:103384–103441.

[20] De Bortoli, V., Thornton, J., Heng, J., and Doucet, A. (2021). Diffusion Schrödinger bridge with applications to score-based generative modeling. *Advances in Neural Information Processing Systems*, 34:17695–17709.

[21] De Marco, S., Pham, H., and Zanni, D. (2026). Schrödinger bridges with jumps for time series generation. *arXiv preprint arXiv:2602.20011*.

[22] Del Campo, A. and Takayanagi, T. (2020). Decoherence in conformal field theory. *Journal of High Energy Physics*, 2020(2):170.

[23] Dunajski, M. and Kryński, W. (2021). Variational principles for conformal geodesics. *Letters in Mathematical Physics*, 111(5):127.

[24] Francesco, P., Mathieu, P., and Sénéchal, D. (2012). *Conformal field theory*. Springer Science & Business Media.

[25] Freeman, N. (2015). Itô calculus and complex Brownian motion. *Lecture Notes. University of Sheffield*.

[26] Gigli, N. and Tamanini, L. (2020). Benamou-Brenier and duality formulas for the entropic cost on $RCD^*(K, N)$ spaces. *Probability Theory and Related Fields*, 176(1-2):1–34.

[27] Girolami, M. and Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(2):123–214.

[28] Gramer, E. (2000). Probability measures with given marginals and conditionals: I-projections and conditional iterative proportional fitting. *Statistics & Risk Modeling*, 18(3):311–330.

[29] Greene, R. E., Kim, K.-T., and Krantz, S. G. (2011). *The geometry of complex domains*. Springer Science & Business Media.

[30] Greydanus, S., Dzamba, M., and Yosinski, J. (2019). Hamiltonian neural networks. *Advances in neural information processing systems*, 32.

[31] Guberman, N. (2016). On complex valued convolutional neural networks. *arXiv preprint arXiv:1602.09046*.

[32] Hammad, M. (2024). Comprehensive survey of complex-valued neural networks: Insights into backpropagation and activation functions. *arXiv preprint arXiv:2407.19258*.

[33] Han, D.-S., Kim, J., Yoo, H. B., and Zhang, B.-T. (2025). Variational online mirror descent for robust learning in Schrödinger bridge. *Transactions on Machine Learning Research*.

[34] Ho, J., Jain, A., and Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851.

[35] Hopf, E. (1950). The partial differential equation ut+ uux= μxx. *Communications on Pure and Applied Mathematics*, 3(3):201–230.

[36] Howard, S., Potaptchik, P., and Deligiannidis, G. (2026). Schrödinger bridge matching for tree-structured costs and entropic Wasserstein barycentres. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

[37] Huang, Y. and Birdal, T. (2026). HOG-Diff: Higher-order guided diffusion for graph generation. In *The Fourteenth International Conference on Learning Representations*.

[38] Hudson, R. L. and Parthasarathy, K. R. (1984). Quantum Ito's formula and stochastic evolutions. *Communications in mathematical physics*, 93(3):301–323.

[39] Hyvärinen, A. and Dayan, P. (2005). Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(4).

[40] Itô, K. (1951). *On stochastic differential equations*. American Mathematical Society.

[41] Itô, K. (1952). Complex multiple Wiener integral. In *Japanese journal of mathematics: transactions and abstracts*, volume 22, pages 63–86. The Mathematical Society of Japan.

[42] Jumarie, G. (1998). A new approach to complex-valued fractional Brownian motion via rotating white noise. *Chaos, Solitons & Fractals*, 9(6):881–893.

[43] Koor, K., Qiu, Y., Kwek, L. C., and Rebentrost, P. (2023). A short tutorial on Wirtinger calculus with applications in quantum information. *arXiv preprint arXiv:2312.04858*.

[44] Koskela, P. (1998). Old and new on the quasihyperbolic metric. In *Quasiconformal Mappings and Analysis: A Collection of Papers Honoring FW Gehring*, pages 205–219. Springer.

[45] Krantz, S. G. (1999). *Handbook of complex variables*. Springer.

[46] Kreutz-Delgado, K. (2009). The complex gradient operator and the CR-calculus. *arXiv preprint arXiv:0906.4835*.

[47] Lawler, G. F. (2008). *Conformally invariant processes in the plane*, volume 114. American Mathematical Society.

[48] Le Gall, J.-F. (2016). *Brownian motion, martingales, and stochastic calculus*, volume 274. Springer.

[49] Lee, C., Hasegawa, H., and Gao, S. (2022). Complex-valued neural networks: A comprehensive survey. *IEEE/CAA Journal of Automatica Sinica*, 9(8):1406–1426.

[50] Léger, F. and Li, W. (2021). Hopf–Cole transformation via generalized Schrödinger bridge problem. *Journal of Differential Equations*, 274:788–827.

[51] Léonard, C. (2013). A survey of the Schrödinger problem and some of its connections with optimal transport. *arXiv preprint arXiv:1308.0215*.

[52] Lévy, P. (1965). *Processus stochastiques et mouvement Brownien*. Gauthier-Villars, 2 edition.

[53] Liu, G.-H., Chen, T., So, O., and Theodorou, E. (2022). Deep generalized Schrödinger bridge. In *Advances in Neural Information Processing Systems*, pages 9374–9388.

[54] Liu, G.-H., Lipman, Y., Nickel, M., Karrer, B., Theodorou, E., and Chen, R. T. Q. (2024). Generalized Schrödinger bridge matching. In *The Twelfth International Conference on Learning Representations*.

[55] Liu, G.-H., Vahdat, A., Huang, D.-A., Theodorou, E. A., Nie, W., and Anandkumar, A. (2023). I^2SB: Image-to-image Schrödinger bridge. In *International Conference on Machine Learning*, pages 22042–22062. PMLR.

[56] Mahmood, H. and Akhtar, A. (2026). A geometric solution of the Schrödinger bridge problem on SO(2) via stochastic optimal control. In *2026 European Control Conference (ECC)*, pages 2283–2288. IEEE.

[57] McKean Jr, H. P. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences*, 56(6):1907–1911.

[58] Mercurio, A., Huang, Y.-T., Cai, L.-X., Chen, Y.-N., Savona, V., and Nori, F. (2025). QuantumToolbox.jl: An efficient Julia framework for simulating open quantum systems. *Quantum*, 9:1866.

[59] Moon, P. and Spencer, D. E. (2012). *Field theory handbook: Including coordinate systems, differential equations and their solutions*. Springer.

[60] Movilla Miangolarra, O., Sabbagh, R., and Georgiou, T. T. (2025). Quantum Schrödinger bridges: Large deviations and time-symmetric ensembles. *Physical Review A*, 112(1):012202.

[61] Nakamura, Y., Pashkin, Y. A., and Tsai, J. (1999). Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature*, 398(6730):786–788.

[62] Neal, R. M. et al. (2011). MCMC using Hamiltonian dynamics. *Handbook of Markov chain Monte Carlo*, 2(11):2.

[63] Nobis, G., Springenberg, M., Belova, A., Daems, R., Knochenhauer, C., Opper, M., Birdal, T., and Samek, W. (2025). Fractional diffusion bridge models. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

[64] Noble, M., Scheid, M., Janati, Y., Moulines, E., and Durmus, A. (2026). Twisted Schrödinger bridge matching. *arXiv preprint arXiv:2607.16987*.

[65] Nutz, M. (2021). Introduction to entropic optimal transport. *Lecture notes, Columbia University*.

[66] Nutz, M. and Wiesel, J. (2023). Stability of Schrödinger potentials and convergence of Sinkhorn's algorithm. *The Annals of Probability*, 51(2):699–722.

[67] Øksendal, B. (2003). *Stochastic differential equations: An introduction with applications*. Springer.

[68] Peyré, G., Cuturi, M., et al. (2019). Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607.

[69] Revuz, D. and Yor, M. (2013). *Continuous martingales and Brownian motion*. Springer Science & Business Media.

[70] Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov processes, and martingales*, volume 2. Cambridge university press.

[71] Rudin, W. (1987). *Real and complex analysis*. McGraw-Hill, Inc.

[72] Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D., and Norouzi, M. (2022). Palette: Image-to-image diffusion models. In *ACM SIGGRAPH 2022 conference proceedings*, pages 1–10.

[73] Sajnok, K. and Matuszewski, M. (2025). Near-equilibrium propagation training in nonlinear wave systems. *arXiv preprint arXiv:2510.16084*.

[74] Schlosshauer, M. (2007). *Decoherence and the quantum-to-classical transition*. Springer.

[75] Schrödinger, E. (1926). Quantisierung als Eigenwertproblem. *Annalen der physik*, 385(13):437–490.

[76] Shi, Y., De Bortoli, V., Campbell, A., and Doucet, A. (2023). Diffusion Schrödinger bridge matching. In *Advances in Neural Information Processing Systems*, volume 36, pages 62183–62223.

[77] Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. (2021). Score-based generative modeling through stochastic differential equations. In *9th International Conference on Learning Representations*.

[78] Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. The MIT Press, second edition.

[79] Tang, S. (2026). Foundations of Schrödinger bridges for generative modeling. *arXiv preprint arXiv:2603.18992*.

[80] Tang, S., Zhang, Y., and Chatterjee, P. (2025). Entangled Schrödinger bridge matching. In *NeurIPS 2025 Workshop on Structured Probabilistic Inference & Generative Modeling*.

[81] Thornton, J., Hutchinson, M., Mathieu, E., De Bortoli, V., Teh, Y. W., and Doucet, A. (2022). Riemannian diffusion Schrödinger bridge. *arXiv preprint arXiv:2207.03024*.

[82] Tong, A., Fatras, K., Malkin, N., Huguet, G., Zhang, Y., Rector-Brooks, J., Wolf, G., and Bengio, Y. (2023). Improving and generalizing flow-based generative models with minibatch optimal transport. *arXiv preprint arXiv:2302.00482*.

[83] Villani, C. (2009). *Optimal transport: Old and new*, volume 338. Springer.

[84] Vlasov, A. A. (1968). The vibrational properties of an electron gas. *Soviet Physics Uspekhi*, 10(6):721–733.

[85] Wiseman, H. M. and Milburn, G. J. (2009). *Quantum measurement and control*. Cambridge university press.

[86] Yang, M. (2025). Topological Schrödinger bridge matching. In *The Thirteenth International Conference on Learning Representations*.

[87] Yue, C., Peng, Z., Ma, J., Du, S., Wei, P., and Zhang, D. (2024). Image restoration through generalized Ornstein-Uhlenbeck bridge. In *International Conference on Machine Learning*, pages 58068–58089. PMLR.

[88] Zhou, L., Lou, A., Khanna, S., and Ermon, S. (2023). Denoising diffusion bridge models. *arXiv preprint arXiv:2309.16948*.

[89] Zhu, C. and Müller, H.-G. (2026). Geodesic optimal transport regression. *Biometrika*, 113(1):asaf086.

[90] Zlotchevski, A. and Chen, L. (2024). Schrödinger bridge problem for jump diffusions. *arXiv preprint arXiv:2411.13765*.

[91] Zlotchevski, A. and Chen, L. (2025). The Schrödinger bridge problem for jump diffusions with regime switching. *arXiv preprint arXiv:2511.06079*.

[92] Zoller, P., Marte, M., and Walls, D. (1987). Quantum jumps in atomic systems. *Physical Review A*, 35(1):198.

Appendices for
Complex Schrödinger Bridges

Abbreviation and Notation

Abbreviation	Expansion
SB	Schrödinger Bridge
SBP	Schrödinger Bridge Problem
CSB	Complex Schrödinger Bridge
I^2CSB	Image-to-Image CSB
BM	Brownian motion
OT	Optimal Transport
EOT	Entropy-regularized OT
COT	Conformal OT
SOC	Stochastic Optimal Control
SDE	Stochastic Differential Equation
CVSDE	Complex-Valued SDE
FBSDE	Forward–Backward SDE
PDE	Partial Differential Equation
MFG	Mean-Field Game
CFT	Conformal Field Theory
OU	Ornstein–Uhlenbeck
KL	Kullback–Leibler
IPF	Iterative Proportional Fitting
TD	Temporal Difference
FID	Fréchet Inception Distance
RK4	Fourth-order Runge–Kutta
CFM	Conditional Flow Matching
EMD	Earth Mover's Distance
cQED	circuit Quantum Electrodynamics
JJ	Josephson Junction

Notation	Usage
t, s	forward and reverse time indices
$Z_t, \bar{Z}_s$	conjugate stochastic states
$\mathbf{z}_t, \bar{\mathbf{z}}_s$	complex state realizations
$\mathbf{x}_t, \mathbf{v}_t$	position and momentum
$W_t, \bar{W}_s$	complex Brownian motions
$\mathbf{w}_t, \hat{\mathbf{w}}_t$	real Brownian motions
$\mathbb{C}^d, \mathbb{R}^{2d}$	complex and real state spaces
$\mathcal{Z}, \mathcal{S}$	sample space, state domain
Ω	continuous-path space
$\mathcal{P}(\Omega)$	path-measure space
$\mathbb{P}, \mathbb{P}_t$	path measure and marginal
$\mathbb{Q}_{\mathbf{w}}, \mathbb{Q}_W$	reference measures
$\mathbb{Q}^\gamma$	conformal reference measure
ρ_t	time density function
μ, ν	endpoint marginals (ρ_0, ρ_T)
$\pi \in \Pi(\mu, \nu)$	coupling and coupling set
σ_t	diffusion coefficient
h_γ	conformal map
$c_\gamma, \hat{c}_\gamma$	conformal coefficients
L_γ	conformal geodesic length
$H, \hat{H}$	MF Hamiltonian and its operator
ψ, φ	log-Schrödinger potentials
$\Psi, \bar{\Psi}$	conjugate wavefunctions
$\Phi, \hat{\Phi}$	complex Hopf–Cole transform
$\nabla_{\mathbf{z}}, \nabla_{\bar{\mathbf{z}}}$	Wirtinger gradients
f, F	base drift, MF cost
$Y_t, \widehat{Y}_t$	CVSDE log-potentials
χ, η	parameterized potentials
$v_t, \hat{v}_t, \check{v}_t, \bar{v}_t$	scaled potential gradients
$\mathcal{E}^+_{T-\tau}, \mathcal{E}^-_\tau$	accumulated IPF errors
$\mathcal{CN}$	complex normal distribution
$\bar{\mu}_s, \Sigma_s$	posterior mean and covariance
$\omega_s^2, \hat{\omega}_s^2$	integrated diffusion variances

A Complex Differential Operators

The complex number system [18, 5] consists of elements of the form $a + ib$, where $a, b \in \mathbb{R}$ and i denotes the imaginary unit. This field exhibits profound mathematical properties that distinguish it from the real numbers. In this section, we establish the foundational mathematical background required to formulate complex-valued differential equations and develop the Hopf–Cole transformation.

A.1 Background on complex differentiation

The concept of differentiation in the complex plane requires more stringent analytical conditions than its real-valued counterpart; that is, the existence of a complex derivative in a neighborhood is a remarkably strong condition. We define differentiability and holomorphic functions as follows.

Definition 2 (Complex differentiability)**.** Let $U \subset \mathbb{C}$ be an open set. A function $f : U \to \mathbb{C}$ is said to be (complex) differentiable at a point $z_0 \in U$ if the limit

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} \tag{18}$$

exists regardless of the path along which z approaches z_0. This limit is denoted by $f'(z_0) = \left.\frac{\mathrm{d}f}{\mathrm{d}z}\right|_{z=z_0}$ and is called the complex derivative of f at z_0.

Definition 3 (Holomorphic functions)**.** Let $\mathcal{S} \subseteq U$ be a nonempty open set. A function f is called holomorphic (or analytic) on $\mathcal{S}$ if f is complex differentiable at every point $z_0 \in \mathcal{S}$.

Holomorphic functions are central objects in complex analysis [71] due to their analytic properties, such as a Taylor series representation around every point where the function is locally differentiable. Furthermore, complex differentiability is characterized by a fundamental system of partial differential equations known as the Cauchy–Riemann equations.

Lemma 2 (Cauchy–Riemann equations)**.** *Let $f : \mathbb{C} \to \mathbb{C}$ be defined as $f(z) = u(x, y) + iv(x, y)$, where $z = x + iy$ and $u, v \in \mathbb{R}$. A necessary condition for f to be analytic in a region $\mathcal{S}$ is that u and v satisfy the Cauchy–Riemann equations [45, 9] throughout $\mathcal{S}$:*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{19}$$

Moreover, if these partial derivatives are continuous in $\mathcal{S}$, then the Cauchy–Riemann equations serve as a sufficient condition for $f(z)$ to be analytic in $\mathcal{S}$.

In physics, Eq. (19) is known as Hamilton's equation [4], if we consider z as an element of a 1-dimensional combined position-momentum space. Lastly, we discuss integration of holomorphic functions.

Definition 4 (Cauchy integral theorem)**.** Let f be a holomorphic function on a simply connected open region $\mathbf{D}$. Then, for any piecewise smooth, closed contour C contained entirely within $\mathbf{D}$, the contour integral of $f(z)$ is zero:

$$\oint_C f(z)\,\mathrm{d}z = 0. \tag{20}$$

A direct consequence of this theorem is that the value of an analytic function at any interior point of a region is completely determined by its values on the boundary.

Lemma 3 (Cauchy integral formula)**.** *Let f be holomorphic on an open set containing a closed contour C and its interior. For an arbitrary z_0 inside C, the function evaluates to*

$$f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0}\,\mathrm{d}z, \tag{21}$$

where the contour of integration is traversed in the positive (counterclockwise) direction.

Corollary 1. *Under the same conditions, the n-th derivative of f at z_0 is given by*

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{n+1}}\,\mathrm{d}z.$$

We further explore holomorphic functions, focusing specifically on their global properties across the complex plane. To establish the necessary foundation, we first define entire functions.

Definition 5 (Entire functions)**.** An entire function is a complex-valued function that is holomorphic at all finite points in the whole complex plane.

An entire function is also called an integral function. Such functions are differentiable everywhere, and each has a power-series representation with an infinite radius of convergence. We next define boundedness of a function.

Definition 6 (Boundedness)**.** A function f is bounded if $\exists M \in \mathbb{R}^+ : |f(z)| < M$.

Liouville's theorem establishes that every bounded entire function must be constant. Formally:

Lemma 4 (Liouville's theorem on holomorphicity)**.**

$$f : \mathbb{C} \to \mathbb{C} \textit{ holomorphic } / \exists M \in \mathbb{R}^+ : |f(z)| < M,\ \forall z \in \mathbb{C} \Rightarrow f = c,\ c \in \mathbb{C}.$$

Equivalently, any non-constant holomorphic function on $\mathbb{C}$ must have an unbounded image. While mathematically elegant, this presents an unsatisfactory limitation from a practical standpoint, particularly in machine learning architectures. The unbounded behavior of purely analytic functions makes them notoriously difficult to train, typically requiring highly sophisticated architectures and specialized activation functions to effectively exploit their phase structure. Fortunately, our framework does not demand strict holomorphicity, as we opt to replace standard differential operators with Wirtinger derivatives. We observe that non-holomorphic complex functions are as important as holomorphic functions, especially when solving parabolic-type PDEs.

A.2 Wirtinger calculus

To describe complex PDEs and CVSDEs involving complex functions consistently and clearly, we use the following definition of the Wirtinger derivatives [46, 43].

Definition 7 (Wirtinger derivatives)**.** Let $f : U \to \mathbb{C}$ be a 2-dimensional smooth function (when viewed as a map $f : U \to \mathbb{R}^2$). Let us consider the case of $f(z) = r(x, y) + i\kappa(x, y)$ with $(x, y) \mapsto x + iy = z$. Define the derivatives with respect to the variable z and its conjugate $\bar{z}$ as

$$\frac{\partial}{\partial z} \coloneqq \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \qquad \text{and} \qquad \frac{\partial}{\partial \bar{z}} \coloneqq \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right),$$

where $\frac{\partial f}{\partial x} = \frac{\partial r}{\partial x} + i\frac{\partial \kappa}{\partial x}$ and $\frac{\partial f}{\partial y} = \frac{\partial r}{\partial y} + i\frac{\partial \kappa}{\partial y}$. If f is holomorphic, the derivative $\frac{\partial f}{\partial z}$ agrees with the complex derivative.

By this definition, the operators $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial \bar{z}}$ are often referred to as the holomorphic and antiholomorphic derivatives, respectively, since the Cauchy–Riemann equations express the condition that the antiholomorphic derivative is zero.

Remark 1 (Cauchy–Riemann condition)**.** $\frac{\partial f}{\partial \bar{z}} = 0$ iff f is *holomorphic*.

The Wirtinger derivative operators, $\frac{\partial}{\partial z}$ and $\frac{\partial}{\partial \bar{z}}$, are inherently $\mathbb{C}$-linear. This property ensures that standard Wirtinger calculus respects the complex structure of the domain by coupling the real and imaginary axes rather than treating them as independent real variables.

Remark 2 ($\mathbb{C}$-linearity)**.** For any complex-valued smooth function f and any complex constant $c \in \mathbb{C}$ (such as the imaginary unit i), the operators satisfy:

$$\frac{\partial}{\partial z}(cf) = c\frac{\partial f}{\partial z}, \qquad \text{and} \qquad \frac{\partial}{\partial \bar{z}}(cf) = c\frac{\partial f}{\partial \bar{z}}.$$

By this linearity, the complex conjugate of the gradient equals the gradient of the conjugate.

Remark 3 (Conjugation)**.** By linearity, it holds that $\frac{\partial \bar{f}}{\partial \bar{z}} = \overline{\left(\frac{\partial f}{\partial z}\right)}$.

For a complex number $z = re^{i\vartheta} \in \mathbb{C}$, one can naturally write a complex logarithm as

$$\log|z| + i\arg(z), \quad \arg(z) = \{\vartheta \mid \vartheta = \mathrm{Arg}(z) + 2\pi k,\ k \in \mathbb{Z}\},$$

where $|z|$ is the magnitude of the complex number and equals $|r|$; $\arg(z)$ is the set of all possible angles for z, obtained by periodic shifts of the principal angular value $\mathrm{Arg}(z) \in (-\pi, \pi]$. The multivalued complex logarithm can be visualized as an infinite helicoid embedded in $\mathbb{R}^2$. To obtain an appropriate single-valued function in $\mathbb{C}$, we can restrict the angle and denote the resulting function by

$$\mathrm{Log}(z) = \log|z| + i\,\mathrm{Arg}(z).$$

Instead of restricting ϑ to a single interval, let it range freely over $\mathbb{R}$. One can then take the Riemann surface (the helicoid) as the domain of the function and define $\log : \mathcal{R} \to \mathbb{C}^*$ by

$$\log(z) \coloneqq \log|r| + i\vartheta. \tag{22}$$

In the definition (22), the function $\log(\cdot)$ is single-valued and holomorphic on this surface, with well-defined behavior of its derivatives on the manifold.

To the best of our knowledge, this work has introduced a novel approach to the complexification of the Hopf–Cole transform [35, 17, 50], in which the potentials have two different interpretations through the introduction of the Hamiltonian dynamics coefficient γ and the imaginary unit "i" respectively:

$$\Psi = e^{\psi/\gamma} \ \text{ and } \ \Phi = e^{i\chi} \qquad \Longrightarrow \qquad \nabla_{\mathbf{z}}\Psi = \frac{\Psi}{\gamma}\nabla_{\mathbf{z}}\psi \ \text{ and } \ \nabla_{\bar{\mathbf{z}}}\Phi = i\Phi\nabla_{\bar{\mathbf{z}}}\chi,$$

where Ψ corresponds to the wave interpretation in the limit leading to the Schrödinger equation considered in Remark 4, whereas Φ is primarily used in the CSB-PDEs for deriving forward and backward SDEs. For the above reasons regarding the complex logarithmic function, we define the relationship using the exponential function as an extension of the Madelung wave function [50]. We treat ψ as the complex logarithm of Ψ constructed from χ and c_γ. The Wirtinger derivatives work

consistently with these logarithmic potentials, akin to the log-derivative trick for real-valued functions. Lastly, in Wirtinger calculus, the standard Laplacian is expressed in the following interchangeable ways using dagger and dot product notation

$$\frac{1}{4}\Delta = -\nabla_{\mathbf{z}}^{\dagger}\nabla_{\mathbf{z}} = \nabla_{\bar{\mathbf{z}}}\cdot\nabla_{\mathbf{z}} = \nabla_{\mathbf{z}}\cdot\nabla_{\bar{\mathbf{z}}} = -\nabla_{\bar{\mathbf{z}}}^{\dagger}\nabla_{\bar{\mathbf{z}}} \tag{23}$$

where the dagger denotes the formal adjoint $\nabla_{\mathbf{z}}^{\dagger} = -\nabla_{\bar{\mathbf{z}}}^{\mathsf{T}}$ and $\nabla_{\bar{\mathbf{z}}}^{\dagger} = -\nabla_{\mathbf{z}}^{\mathsf{T}}$. In this work, these properties of complex operators are leveraged to derive the CSB formulations succinctly. Note that a semantically equivalent real-valued formulation requires writing at least four times as many operations without the introduction of an analogous number system. Beyond providing a concise exposition, deriving the formulations via complex operators significantly reduces human error during implementation on machine learning hardware optimized for fast complex-number computation.

A.3 Complex Hopf–Cole transform

For a Hamiltonian $H(\cdot,\cdot,\nabla_{\mathbf{z}}\psi,\nabla_{\bar{\mathbf{z}}}\psi)$, we assume that the general dynamic formulation is described with a generalized Schrödinger-type system for $t\in[0,T]$. This is derived by the Feynman–Kac formula in the real-valued problem. In our complex setting, we propose to write the system in the following alternative form:

$$\begin{cases} i\partial_t\psi = H(\mathbf{z},\bar{\mathbf{z}},\nabla_{\mathbf{z}}\psi,\nabla_{\bar{\mathbf{z}}}\psi,\rho) - \frac{i}{4}c_\gamma\Delta(\sigma^2\psi), \\ \partial_t\rho = i\nabla_{\bar{\mathbf{z}}}^{\dagger}\big(\nabla_\alpha H(\mathbf{z},\bar{\mathbf{z}},\nabla_{\mathbf{z}}\psi,\nabla_{\bar{\mathbf{z}}}\psi,\rho)\,\rho\big) - i\nabla_{\mathbf{z}}^{\dagger}\big(\overline{\nabla_\alpha H(\mathbf{z},\bar{\mathbf{z}},\nabla_{\mathbf{z}}\psi,\nabla_{\bar{\mathbf{z}}}\psi,\rho)}\,\rho\big) + \frac{1}{4}\Delta(\sigma^2\rho), \end{cases} \tag{24}$$

where the diffusion coefficient $\sigma=\sigma(t)$ is spatially constant. We take $\rho>0$ and impose $\psi+\bar{\psi} = \gamma\log\rho$ and $\nabla_{\mathbf{z}}\cdot f = \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}$. The PDEs can be viewed as a blending of two optimality conditions of diffusion bridges; the first row indicates a Hamilton–Jacobi (HJ) equation, and the second row indicates a Fokker–Planck (FP) equation [10, 53]. The parameter $\gamma\in\mathbb{R}$ enters through $c_\gamma = 1+i\gamma$, the logarithmic derivative $h_\gamma'(\ell)/h_\gamma(\ell)$ of the spiral map, and through the density-dependent term in the Hamiltonian below. We use the complex Hopf–Cole transform as follows

$$\Phi\big(z(t),\bar{z}(t),t\big) = \exp\Big(\frac{i}{c_\gamma}\big(\chi\big(z(t),\bar{z}(t),t\big)+\gamma\log\rho\big)\Big),\;\; z(t)=\mathbf{z}_t, \tag{25a}$$

$$\hat{\Phi}\big(z(t),\bar{z}(t),t\big) = \rho\exp\Big(\frac{i}{c_\gamma}\bar{\chi}\big(z(t),\bar{z}(t),t\big)\Big). \tag{25b}$$

where $\psi=\chi+\gamma\log\rho$, $\bar{\chi}=-\psi$ and $\Phi\hat{\Phi}=\rho$. Apparently, the construction of $\hat{\Phi}$ here is a relaxation of complex conjugation that preserves the measure, and the density $\rho(\mathbf{z},\bar{\mathbf{z}},t)$ is recovered from the product $\Phi\hat{\Phi}$. Note that, for potentials $Y=e^{i\psi/c}$ and $\hat{Y}=\rho e^{i\bar{\chi}/c}$, applying the multivariate calculus yields

$$\begin{aligned}
\nabla Y &= e^{i\psi/c}\big\{ic^{-1}\nabla\psi\big\},\\
\nabla\hat{Y} &= \hat{Y}\big\{ic^{-1}\nabla\bar{\chi}+\nabla\log\rho\big\},\\
\nabla^{\dagger}\nabla Y &= e^{i\psi/c}\big\{c^{-2}(\bar{\nabla}\psi)^{\mathsf{T}}(\nabla\psi)+ic^{-1}\nabla^{\dagger}\nabla\psi\big\},\\
\nabla^{\dagger}\nabla\hat{Y} &= \hat{Y}\big\{-c^{-2}(i\bar{\nabla}\bar{\chi}+c\bar{\nabla}\log\rho)^{\mathsf{T}}(i\nabla\bar{\chi}+c\nabla\log\rho)+ic^{-1}\nabla^{\dagger}\nabla\bar{\chi}+\nabla^{\dagger}\nabla\log\rho\big\},
\end{aligned}$$

and the equations hold for the gradient operators appearing in the PDE: $(\nabla,\bar{\nabla}) = (\nabla_{\mathbf{z}},\nabla_{\bar{\mathbf{z}}})$ and $(\nabla,\bar{\nabla}) = (\nabla_{\bar{\mathbf{z}}},\nabla_{\mathbf{z}})$; the first two identities also hold for $\nabla=\partial_t$. Hence, one shall carry out the following derivations regarding a twisted Hamiltonian,

$$H(\cdot,\cdot,\nabla_{\mathbf{z}}\psi,\nabla_{\bar{\mathbf{z}}}\psi,\rho) \coloneqq \sigma^2\big\langle\nabla_{\mathbf{z}}\psi,\nabla_{\mathbf{z}}\bar{\psi}\big\rangle + (\nabla_{\mathbf{z}}\psi)^{\mathsf{T}}f - (\nabla_{\bar{\mathbf{z}}}\psi)^{\mathsf{T}}\bar{f} + \frac{i\gamma\sigma^2}{2}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}, \tag{26}$$

where $\langle\alpha,\beta\rangle = \alpha^{\mathsf{T}}\bar{\beta}$ denotes the standard Hermitian inner product. Substituting this Hamiltonian into the expression for the forward time derivative of Φ, we get

$$\begin{aligned}
\frac{\partial\Phi}{\partial t} = \frac{i\Phi}{c_\gamma}\frac{\partial\psi}{\partial t} &= \frac{\Phi}{c_\gamma}\bigg(\sigma^2(\nabla_{\mathbf{z}}\psi)^{\mathsf{T}}\nabla_{\bar{\mathbf{z}}}\psi + (\nabla_{\mathbf{z}}\psi)^{\mathsf{T}}f - (\nabla_{\bar{\mathbf{z}}}\psi)^{\mathsf{T}}\bar{f} - \frac{i\sigma^2}{4}c_\gamma\Delta\psi + \frac{i\gamma\sigma^2}{2}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\bigg)\\
&= -i\nabla_{\mathbf{z}}\Phi^{\mathsf{T}}f + i\nabla_{\bar{\mathbf{z}}}\Phi^{\mathsf{T}}\bar{f} - \frac{\sigma^2}{4}c_\gamma\Delta\Phi + \frac{i\gamma\sigma^2}{2c_\gamma}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\Phi,
\end{aligned}$$

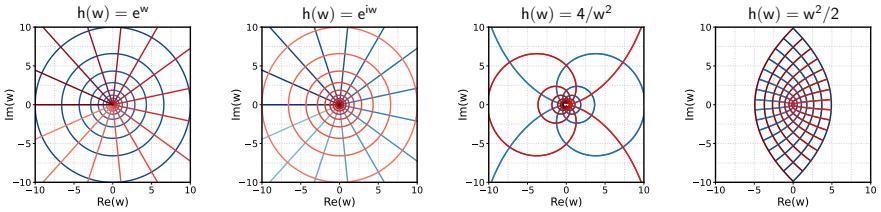

Figure 10: Conformal mappings in one complex dimension $\mathbb{C}^1$ from Moon and Spencer [59].

and using $\nabla_\alpha H = f + \sigma^2 \nabla_{\bar{\mathbf{z}}} \psi$ in the density equation in Eq. (24) with $\psi + \bar{\psi} = \gamma \log \rho$, we get

$$\begin{aligned}
\frac{\partial \hat{\Phi}}{\partial t} &= \frac{1}{\Phi}\left(\frac{\partial \rho}{\partial t} - \hat{\Phi}\frac{\partial \Phi}{\partial t}\right) \\
&= \frac{1}{\Phi}\left\{-i\nabla_{\mathbf{z}} \cdot (\rho f) + i\nabla_{\bar{\mathbf{z}}} \cdot (\rho \bar{f}) + \frac{\sigma^2}{4} c_\gamma \big(\Phi \Delta \hat{\Phi} - \hat{\Phi}\Delta\Phi\big) - \hat{\Phi}\frac{\partial \Phi}{\partial t}\right\} \\
&= -i\nabla_{\mathbf{z}} \cdot (\hat{\Phi} f) + i\nabla_{\bar{\mathbf{z}}} \cdot (\hat{\Phi} \bar{f}) + \frac{\sigma^2}{4} c_\gamma \Delta \hat{\Phi} - \frac{i\gamma\sigma^2}{2c_\gamma} \frac{\Delta\sqrt{\rho}}{\sqrt{\rho}} \hat{\Phi}.
\end{aligned}$$

These equations give the corresponding CSB-PDEs. Furthermore, the Schrödinger equation Eq. (2) naturally guides us to the time-reversed version Eq. (24). Using the conjugation of our Hermitian function $H^*(\mathbf{z}, \bar{\mathbf{z}}, \nabla_{\mathbf{z}}\psi, \nabla_{\bar{\mathbf{z}}}\psi) = \overline{H(\bar{\mathbf{z}}, \mathbf{z}, \nabla_{\bar{\mathbf{z}}}\psi, \nabla_{\mathbf{z}}\psi)}$, we obtain alternative McKean–Vlasov CVSDE

$$\mathrm{d}Z_s = i\,\nabla_\alpha H\big(Z_s, \bar{Z}_s, \nabla_{\mathbf{z}}\bar{\psi}_s, \nabla_{\bar{\mathbf{z}}}\bar{\psi}_s, \rho_s\big)\,\mathrm{d}s + \sigma\,\mathrm{d}W_s. \tag{27}$$

Undoubtedly, one can apply the corresponding Hopf–Cole transform akin to Eq. (25) with the given real-valued density $\rho = \bar{\rho}$. Therefore, the symmetric CVSDEs Eq. (11) are both mean-field SDEs and distinct optimal control solutions under time reversibility. In addition, we leave the following remark that relates our McKean–Vlasov-type CVSDEs to the Schrödinger equation Eq. (2).

Remark 4 (Schrödinger equation)**.** Consider a one-parameter family $\{(\gamma_\tau, \sigma_\tau)\}_{\tau \geq 0}$ satisfying $\sigma_\tau \to 0$ and $\gamma_\tau \sigma_\tau \to 1$ as $\tau \to \infty$. Under the hypotheses of the SDE existence and uniqueness theorem [67], the rotated potential $\Psi = e^{\psi/\gamma_\tau}$ converges to the harmonic oscillator Eq. (2) if and only if $\rho_0 = \rho_T$.

Consequently, we argue that solving the CSB problem sits between modelling a pure rotational Hamiltonian of the Schrödinger equation and learning for higher-order optimal control subject to an energy function that comes from an affine Hamiltonian prior.

B Conformal Mappings and Complex-Valued Stochastic Processes

This appendix discusses relevant topics regarding conformal mapping and stochastic processes.

Conformal mappings. The conformal mapping preserves angular magnitudes and orientations between curves, carrying essential geometric structure underlying sample spaces. Because conformal mappings preserve angles locally, two straight lines in a plane may bend in such a way as to preserve the angles at which they intersect. The notion of conformality is closely related to the holomorphic function and the Cauchy–Riemann equation where a nonconstant holomorphic function h is guaranteed to be locally conformal in $\mathbb{C}^1$.

Lemma 5 (Riemann mapping theorem; Theorem 14.8 of Rudin [71])**.** *Every nonempty simply connected domain $\mathcal{Z} \subsetneq \mathbb{C}$ admits a biholomorphism onto the open unit disc $\mathbb{D} \coloneqq \{z \in \mathbb{C} \mid |z| < 1\}$, i.e., a bijective holomorphic map with a holomorphic inverse.*

Conformal mappings have been widely accepted in real two-dimensional ($d = 1$) field problems [59] and offer considerable flexibility in this setting. As various example functions in Fig. 10 demonstrate, the images of perpendicular Cartesian grid lines still meet at right angles. For a chosen origin $z_0 \in \mathcal{Z}$, the conditions $h(z_0) = 0$ and $h'(z_0) > 0$ uniquely determine the Riemann map, as the following Schwarz lemma statement shows.

Lemma 6 (Schwarz; Theorem 12.2 of Rudin [71]). *Let $h : \mathbb{D} \to \mathbb{D}$ be holomorphic with $h(0) = 0$. Then $|h(z)| \leq |z|$ for all $z \in \mathbb{D}$ and $|h'(0)| \leq 1$. If $|h(z)| = |z|$ for some nonzero $z \in \mathbb{D}$ or $|h'(0)| = 1$, then $h(z) = az$ for a constant $a \in \mathbb{C}$ with $|a| = 1$.*

Under the Schwarz lemma, simply $h'(0) = 1$ implies $h(z) = z$ in one complex dimension. As a higher-dimensional parallel, Cartan's uniqueness theorem extends this rigidity to bounded domains in $\mathbb{C}^d$. In multidimensional settings, simply connected domains need not be biholomorphically equivalent, and biholomorphisms need not preserve Euclidean angles. In the following statements, $\mathbf{0}$ denotes a fixed point chosen as the origin and $\mathbf{J}_f$ denotes the holomorphic Jacobian matrix of f, explicitly defined with the Wirtinger operator as $\mathbf{J}_f \coloneqq [\nabla_{\mathbf{z}}^{\mathsf{T}} f_1, \nabla_{\mathbf{z}}^{\mathsf{T}} f_2, \cdots, \nabla_{\mathbf{z}}^{\mathsf{T}} f_d]^{\mathsf{T}}$.

Lemma 7 (Cartan; Theorem 1.3.1 of [29]). *Let $\mathcal{Z} \subset \mathbb{C}^d$ be bounded and $f : \mathcal{Z} \to \mathcal{Z}$ be a holomorphic function. If $f(\mathbf{0}) = \mathbf{0}$ and $\mathbf{J}_f(\mathbf{0}) = \mathbf{I}_d$ at some $\mathbf{0} \in \mathcal{Z}$, then $f = \mathrm{id}_{\mathcal{Z}}$ holds.*

Corollary 2. *If $\mathcal{Z}$ is a bounded domain in $\mathbb{C}^d$ and $p \in \mathcal{Z}$, then the mapping $f \mapsto \mathbf{J}_f(p)$ is an injective homomorphism of the isotropy group $\mathcal{Q}_p \coloneqq \{f \in \mathrm{Aut}(\mathcal{Z}) : f(p) = p\}$ into $GL(d, \mathbb{C})$.*

Corollary 3. *The injective homomorphism $f \mapsto \mathbf{J}_f(p)$ of $\mathcal{Q}_p$ onto $d\mathcal{Q}_p \coloneqq \{\mathbf{J}_f(p) : f \in \mathcal{Q}_p\}$ is a homeomorphism of $\mathcal{Q}_p$ (in the compact-open topology) onto a compact subgroup of $GL(d, \mathbb{C})$.*

The uniqueness theorem and associated corollaries provided by Greene et al. [29] imply that every automorphism fixing $\mathbf{0}$ ($h(\mathbf{0}) = \mathbf{0}$) is uniquely determined by its Jacobian at $\mathbf{0}$. They also imply that there is a Hermitian metric for which $\mathcal{Q}_p$ acts isometrically on the tangent space at p such that the corresponding map $f \mapsto \mathbf{J}_f(p)$ is also a homeomorphism onto its image, and that the matrix group $\mathbf{J}_f(p)$ is compact. Consequently, admitting a common conformal chart additionally requires $\mathbf{J}_h(\mathbf{z})\mathbf{J}_h^{\dagger}(\mathbf{z}) = \Omega(\mathbf{z})^2\mathbf{I}_d$ with a real field Ω.

Conformal martingales. Conformality plays a central role in formulating diffusions under generalized geometric priors, as it allows us to analyze diffusion bridge processes across both the planar and transformed domains in a unified manner. To characterize these processes, we first introduce the following definition of martingales.

Definition 8 (Martingales). Let $(\mathcal{Z}, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\in I})$ for an ordered time set $I \subset [0, \infty)$ be a filtered probability space. A real-valued stochastic process $\{X_t\}_{t\in I}$ is a *martingale* with respect to the filtration $\{\mathcal{F}_t\}_{t\in I}$ if ① X_t is $\mathcal{F}_t$-measurable for all $t \in I$, ② $\mathbb{E}\|X_t\| < \infty$ for all $t \in I$, and ③ for all $\tau, t \in I$ with $\tau \leq t$, the conditional expectation satisfies $\mathbb{E}[X_t|\mathcal{F}_\tau] = X_\tau$, P-almost surely.

A martingale is a process whose conditional expectation at any future time, given all past information, equals the process's current value. Notably, score-based diffusion models rely not only on the properties of BM, but also on Girsanov's theorem to govern changes of probability measure. Central to our framework is the following result on the conformal invariance of complex BM. Under proper normalization of time, W_t is a proper complex Wiener process satisfying $\mathbb{E}[\mathrm{d}W_t\, \mathrm{d}W_t^{\dagger}] = \mathbf{I}_d\, \mathrm{d}t$ and $\mathbb{E}[\mathrm{d}W_t\, \mathrm{d}W_t^{\mathsf{T}}] = 0$.

Proposition 3 (Conformal invariance; Theorem V.2.5 of Revuz and Yor [69]). *Let h be an entire and non-constant function, and consider a complex Wiener process, i.e., $W_t = (\mathbf{w}_t + i\hat{\mathbf{w}}_t)/\sqrt{2}$ with independent standard real BMs $\mathbf{w}$ and $\hat{\mathbf{w}}$. Then, $\mathfrak{B}_t = h(W_t)$ is a standard BM with a time change. More precisely, there exist complex BMs W and $\tilde{\mathfrak{B}}$ on the probability space such that*

$$h(W_t) = h(W_0) + \tilde{\mathfrak{B}}_{\langle Z,Z\rangle_t}, \tag{28}$$

where $\langle Z, Z\rangle_t = \int_0^t \|h'(W_\tau)\|^2\, \mathrm{d}\tau$; $\langle Z, Z\rangle_\infty = \infty$ almost surely, with the clock denoted by $\langle Z, Z\rangle_t$.

Note that the monotonically increasing term $\langle Z, Z\rangle_t$ serves as an alternative time scale for the transformed trajectories. Therefore, mappings preserve continuous conformal local martingales $(X + iY)$ up to domain exit, with $\langle X, X\rangle = \langle Y, Y\rangle$ and $\langle X, Y\rangle = 0$; additional integrability conditions are required for true martingales.

Quasiconformal mappings. Due to the geometrical rigidity, every conformal mapping for $d \geq 2$ reduces to a Möbius transformation, a composition of basic operations—namely translations, scalings, rotations, reflections, and inversions. In this work, we resort to coordinatewise normalized operations, which are more flexible than global conformal mapping and are sufficient to support generative modeling and simulation under Hamiltonian dynamics. As a result, we relax the requirement to that of quasiconformal mappings [1], orientation-preserving homeomorphisms between open sets.

Definition 9 (Quasiconformal mappings)**.** Suppose that there is a smooth function $f : \mathcal{Z} \to \mathcal{Z}'$ where $\mathcal{Z}, \mathcal{Z}' \in \mathbb{C}^1$. Then f is quasiconformal if its partial derivatives satisfy the Beltrami equation

$$\frac{\partial f}{\partial \bar{z}} = \mu(z) \frac{\partial f}{\partial z} \tag{29}$$

for some complex-valued Lebesgue measurable function μ satisfying $|\mu|_\infty < 1$.

The definition describes a 1-dimensional complex map whose local distortion is prescribed by a complex-valued function μ, called the Beltrami coefficient. The Beltrami equation is a first-order partial differential equation that can be viewed as a generalization of the Cauchy–Riemann equations.[1]

Quasiconformal mappings admit a natural geometric interpretation related to quasihyperbolic planes. Suppose we equip $\mathcal{Z}$ with the Riemannian metric

$$(\mathrm{d}s)^2 = \Omega(z)^2 \left|\mathrm{d}z + \mu(z)\, \mathrm{d}\bar{z}\right|^2 \tag{30}$$

where the conformal factor $\Omega(z) > 0$ is an arbitrary smooth function that sets the local length scale. Then a mapping h satisfies (29) if and only if it is conformal from $(\mathcal{Z}, \mathrm{d}s)$ to the target domain $\mathcal{Z}'$ endowed with the standard Euclidean metric; accordingly, h is said to be μ-conformal.

A quasiconformal map permits directional stretching with a uniformly bounded ratio: at any regular point of differentiability, its real differential maps infinitesimal circles to ellipses with an axis ratio bounded by the maximal dilatation $K = \frac{1+k}{1-k}$, $k = |\mu|_\infty$. By the measurable Riemann mapping theorem (the quasiconformal analogue of the Riemann mapping theorem), for every measurable Beltrami coefficient μ on $\mathbb{C}$ satisfying $|\mu|_\infty < 1$, there exists a quasiconformal homeomorphism that solves the Beltrami equation almost everywhere. This indicates even a discontinuous measurable field of admissible local distortions can be realized by a global map and such a homeomorphism is uniquely determined under the normalization.

Based on these details, this work considers a path-normalized quasiconformal homeomorphism $H_\gamma : C([0,T];\mathbb{C}^d) \to C([0,T];\mathbb{C}^d)$ in the uniform topology, where we take conformal spatial slices into consideration at fixed $t < T$ and a fixed endpoint:

$$(\mathcal{H}_\gamma \mathbf{z})_t = \mathbf{z}_T + \frac{h_\gamma(\log(1-t/T))}{h_0(\log(1-t/T))}(\mathbf{z}_t - \mathbf{z}_T), \quad (\mathcal{H}_\gamma \mathbf{z})_T = \mathbf{z}_T. \tag{31}$$

Apart from standard treatment using a quasihyperbolic metric [44, 6], our construction is conditioned on the respective endpoint $\mathbf{z}_T$ due to the marginal constraint $\mathbb{P}_0 = \mathbb{P}_0^\gamma = \rho_0$ and $\mathbb{P}_T = \mathbb{P}_T^\gamma = \rho_T$. Still, it is a continuous, bijective function between two topological spaces that has a continuous inverse function, with properties such as $\mathcal{H}_\gamma^{-1} = \mathcal{H}_{-\gamma}$, $\mathcal{H}_0 = \mathrm{id}$, and $\mathcal{H}_{\gamma_1} \circ \mathcal{H}_{\gamma_2} = \mathcal{H}_{\gamma_1+\gamma_2}$ for $\gamma_1, \gamma_2 \in \mathbb{R}$.

Fig. 11 illustrates the self-similar geometry of COT in (8). Since general conformality alone does not guarantee global invertibility, equivalence of path measures, or preservation of an arbitrary transport objective, we focus on this explicit similarity of our instance-based quasiconformal spiral mapping, which contains no reflection or folding. The time-dependent similarity in the figure describes its rotating contraction. Let the terminal time be $T = 1$, and let $\mathbf{z} \in \mathbb{C}^d \simeq \mathbb{R}^{2d}$. Conditional on a fixed endpoint $\mathbf{z}_T$, our construction with $\gamma > 0$ uses $\Upsilon_t^\gamma(\mathbf{z}) = (\mathcal{H}_\gamma \mathbf{z}; \mathbf{z}_T)_t$, $\Upsilon_t^\gamma(\mathbf{z}) \coloneqq \mathbf{z}_T + (1-t)^{1+i\gamma}(\mathbf{z} - \mathbf{z}_T)$, with $\Upsilon_1^\gamma(\mathbf{z}) = \mathbf{z}_T$ by continuity. Since $(1-t)^{1+i\gamma} = (1-t)e^{i\gamma \log(1-t)}$, the map is the identity at $t = 0$ and a rotation followed by a strict contraction toward $\mathbf{z}_T$ for $0 < t < 1$; for every fixed $t < 1$, it is a global conformal similarity of $\mathbb{C}^d \simeq \mathbb{R}^{2d}$. For $\mathbf{z} \neq \mathbf{z}_T$, the spiral satisfies $\|\Upsilon_t^\gamma(\mathbf{z}) - \mathbf{z}_T\| = (1-t)\|\mathbf{z} - \mathbf{z}_T\|$. Its distance from the endpoint decreases strictly, so the trajectory cannot self-intersect before reaching $\mathbf{z}_T$. Because the transformation fixes the boundary states, consistently transforming the reference measure, admissible measures, and path space yields equivalent SB problems. Likewise, for every fixed $t < 1$ and a common fixed endpoint $\mathbf{z}_T$, applying Υ_t^γ to both arguments of a quadratic transport cost multiplies all squared distances by $(1-t)^2$, leaving the minimizing coupling unchanged.

Let $\mathbb{P}^{\gamma,*} = (\mathcal{H}_\gamma)_\# \mathbb{P}$ be the optimal path law in Eq. (10), with terminal marginal ρ_T. For fixed $t < T$ and terminal point $\mathbf{z}_T$, write

$$\mu_{\gamma,t}^{\delta(\mathbf{z}_T)} \coloneqq \frac{\partial_{\bar{z}}(\mathcal{H}_\gamma)_t}{\partial_z(\mathcal{H}_\gamma)_t} = \frac{0}{(1-t/T)^{i\gamma}} = 0, \qquad \mu_{\gamma,t}^{\rho_T}(\mathrm{d}\mathbf{z}) \coloneqq (\mathbb{P}^{\gamma,*})_t(\mathrm{d}\mathbf{z}) = \rho_t^*(\mathbf{z})\ \mathrm{d}^{2d}\mathbf{z}.$$

[1] Plugging in $\mu = 0$ recovers the Cauchy–Riemann equations and also holomorphicity.

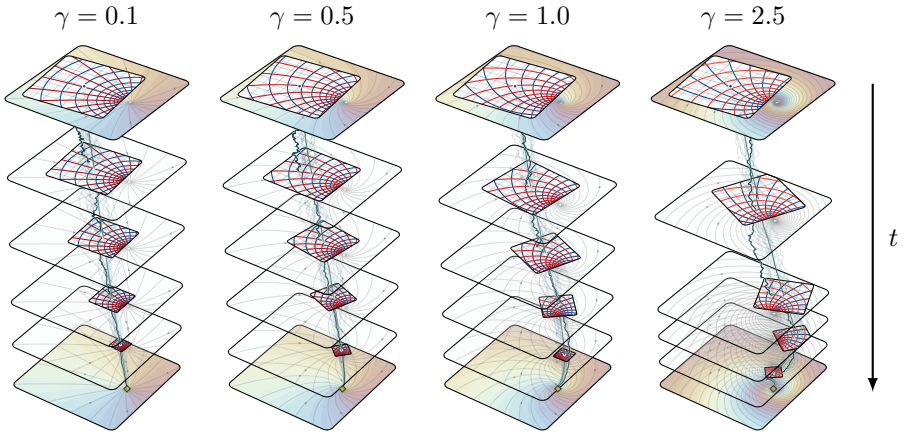

Figure 11: Conformal optimal transport geometry for $\gamma = \{0.1, 0.5, 1.0, 2.5\}$. Time increases downward. Increasing γ produces greater angular winding through $e^{i\gamma \log(1-t)}$, while the radial contraction remains $1 - t$, preserving the self-similarity of the transported geometry.

Here $\delta(\mathbf{z}_T)$ denotes the Dirac measure at $\mathbf{z}_T$. The first quantity is the Beltrami coefficient in one complex spatial coordinate. Its vanishing shows that each fixed-time spatial slice is conformal for every $\gamma \in \mathbb{R}$. The second quantity is the time-t marginal of the optimal path law, with density ρ_t^*.

Itô formulas. We recall the real and complex Itô lemmas for stochastic differential equations.

Lemma 8 (Itô's formula [40])**.** *Let $\mathbf{x}_t$ be the solution to a real-valued SDE:*

$$\mathrm{d}X_t = f(X_t, t)\,\mathrm{d}t + g(X_t, t)\,\mathrm{d}W_t. \tag{32}$$

Then, the stochastic process $u(\mathbf{x}_t, t)$, where $u \in C^{2,1}(\mathbb{R}^d, [0, T])$, is also an Itô process satisfying

$$\begin{aligned}\mathrm{d}u(X_t, t) = \frac{\partial u}{\partial t}\,\mathrm{d}t + \Big\{\nabla_{\mathbf{x}} u(X_t, t)^{\mathsf{T}} f(X_t, t) + \\ \frac{1}{2}\operatorname{Tr}\big[gg^{T}(X_t, t)\operatorname{Hess} u(X_t, t)\big]\Big\}\,\mathrm{d}t + \big\{\nabla_{\mathbf{x}} u(X_t, t)^{\mathsf{T}} g(X_t, t)\big\}\,\mathrm{d}W_t\end{aligned} \tag{33}$$

where $\nabla_{\mathbf{x}}$ is the gradient operator with respect to $\mathbf{x}_t$.

The complex Itô lemma (also known as the quantum Itô formula in the literature) [41, 38, 25] describes the progression of complex variations. In the one-dimensional case, the formula gives the differential $\mathrm{d}\varphi(z_t, \bar{z}_t, t)$ through a second-order Taylor expansion:

$$\mathrm{d}\varphi = \frac{\partial \varphi}{\partial t}\,\mathrm{d}t + \frac{\partial \varphi}{\partial z}\,\mathrm{d}Z_t + \frac{\partial \varphi}{\partial \bar{z}}\,\mathrm{d}\bar{Z}_t + \frac{1}{2}\left(\frac{\partial^2 \varphi}{\partial z^2}(\mathrm{d}Z_t)^2 + 2\frac{\partial^2 \varphi}{\partial z \partial \bar{z}}\,\mathrm{d}Z_t\,\mathrm{d}\bar{Z}_t + \frac{\partial^2 \varphi}{\partial \bar{z}^2}(\mathrm{d}\bar{Z}_t)^2\right)$$

The second-order terms (the quadratic variations) constitute the main difference between stochastic calculus and standard calculus. To evaluate $(\mathrm{d}Z_t)^2$, $(\mathrm{d}\bar{Z}_t)^2$, and $\mathrm{d}Z_t\,\mathrm{d}\bar{Z}_t$, we expand them using standard real-valued Itô multiplication rules (where $\mathrm{d}t^2 = 0$, $\mathrm{d}t\,\mathrm{d}W_t = 0$, and $\mathrm{d}W_t^2 = \mathrm{d}t$) as follows:

$$\begin{aligned}(\mathrm{d}Z_t)^2 &= (\mathrm{d}X_t)^2 - (\mathrm{d}Y_t)^2 + 2i\,\mathrm{d}\langle X, Y\rangle_t \\ (\mathrm{d}\bar{Z}_t)^2 &= (\mathrm{d}X_t)^2 - (\mathrm{d}Y_t)^2 - 2i\,\mathrm{d}\langle X, Y\rangle_t \\ (\mathrm{d}Z_t)(\mathrm{d}\bar{Z}_t) &= (\mathrm{d}X_t)^2 + (\mathrm{d}Y_t)^2\end{aligned}$$

In the simple case of Brownian motion we can replace the quadratic Itô correction term with the Laplacian by the following remark.

Remark 5 (Complex quadratic variations)**.** Let $W_t = (\mathbf{w}_t + i\hat{\mathbf{w}}_t)/\sqrt{2}$ be a standard complex Brownian motion constructed from two independent real Brownian motions $\mathbf{w}_t$ and $\hat{\mathbf{w}}_t$. Then $\mathrm{d}\langle W, W\rangle_t = 0$. Let Z_t be a stochastic process satisfying $\mathrm{d}Z_t = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$. A well-known consequence of their independence and equal variance is that the complex quadratic variation vanishes [48]

$$(\mathrm{d}Z_t)^2 = (\mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t)^2 = 0,$$

and $(\mathrm{d}\bar{Z}_t)^2 = 0$ also holds.

As a natural consequence of Remark 5, we state a multivariate complex Itô formula herein.

Lemma 9 (Complex Itô formula [41])**.** *Let Z_t be the solution to a CVSDE:*

$$\mathrm{d}Z_t = \mu(Z_t, \bar{Z}_t, t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$$

where W_t is the complex Wiener process and μ and σ are sufficiently smooth. Then, the stochastic process defined by an arbitrary time-dependent smooth function $\varphi(\mathbf{z}_t, \bar{\mathbf{z}}_t, t)$ satisfies the SDE:

$$\begin{aligned}\mathrm{d}\varphi(Z_t, \bar{Z}_t, t) =& \left\{\frac{\partial\varphi}{\partial t} + \nabla_{\mathbf{z}}\varphi(Z_t, \bar{Z}_t, t)^{\mathsf{T}}\mu(Z_t, \bar{Z}_t, t) + \nabla_{\bar{\mathbf{z}}}\varphi(Z_t, \bar{Z}_t, t)^{\mathsf{T}}\bar{\mu}(\bar{Z}_t, Z_t, t)\right\}\mathrm{d}t + \\ & \left\{\sigma^2\nabla_{\bar{\mathbf{z}}}\cdot\nabla_{\mathbf{z}}\varphi(Z_t, \bar{Z}_t, t)\right\}\mathrm{d}t + \sigma\nabla_{\mathbf{z}}\varphi(Z_t, \bar{Z}_t, t)^{\mathsf{T}}\,\mathrm{d}W_t + \sigma\nabla_{\bar{\mathbf{z}}}\varphi(Z_t, \bar{Z}_t, t)^{\mathsf{T}}\,\mathrm{d}\bar{W}_t\end{aligned} \tag{34}$$

where $(\nabla_{\mathbf{z}}, \nabla_{\bar{\mathbf{z}}})$ are multivariate Wirtinger gradient operators with respect to $(\mathbf{z}, \bar{\mathbf{z}})$.

The time-dependent rotation in $\mathcal{H}_\gamma$ modifies the drift and preserves the quadratic variation. In our general bridge setting, we explicitly use the definitions of both potentials and their mixing operator, with the resulting drift derived from the corrected complex Itô formula.

C Proofs

We introduce a complete, multivariate derivation of SB-CVSDEs, which provide an explicit complexified formulation of dynamic Schrödinger bridge models. The derivation is motivated by relevant topics in statistical physics, differential equations and controls [57, 84], as well as recent advances in score-based diffusion models and Schrödinger bridges [77, 13, 53].

C.1 Proof of Lemma 1 (complex Langevin dynamics)

With $\mathbf{z} = \mathbf{x} + i\mathbf{v}$, let us write the scores, the real gradients of the log-density, as follows:

$$s_{\mathbf{x}}(\mathbf{x}, \mathbf{v}) = \nabla_{\mathbf{x}}\log p(\mathbf{x}, \mathbf{v}) \qquad \text{and} \qquad s_{\mathbf{v}}(\mathbf{x}, \mathbf{v}) = \nabla_{\mathbf{v}}\log p(\mathbf{x}, \mathbf{v}).$$

This naturally yields

$$s_{\mathbf{z}}(\mathbf{z}, \bar{\mathbf{z}}) = \nabla_{\mathbf{z}}\log\rho(\mathbf{z}, \bar{\mathbf{z}}) = \frac{s_{\mathbf{x}}(\mathbf{x}, \mathbf{v})}{2} - i\frac{s_{\mathbf{v}}(\mathbf{x}, \mathbf{v})}{2} \tag{35}$$

One can naturally define their Wirtinger calculus counterparts as

$$s_{\bar{\mathbf{z}}}(\mathbf{z}, \bar{\mathbf{z}}) = \nabla_{\bar{\mathbf{z}}}\log\rho(\mathbf{z}, \bar{\mathbf{z}}) = \frac{s_{\mathbf{x}}(\mathbf{x}, \mathbf{v})}{2} + i\frac{s_{\mathbf{v}}(\mathbf{x}, \mathbf{v})}{2} \tag{36}$$

Then, the suggested forward drift coefficient is written with the Wirtinger gradients (36) as

$$i\sigma^2\nabla_{\bar{\mathbf{z}}}\psi = i\sigma^2\frac{\gamma s_{\bar{\mathbf{z}}} - i s_{\bar{\mathbf{z}}}}{2} = i\sigma^2\left(\gamma\frac{s_{\mathbf{x}} + i s_{\mathbf{v}}}{4} - i\frac{s_{\mathbf{x}} + i s_{\mathbf{v}}}{4}\right) = \sigma^2\frac{s_{\mathbf{x}} - \gamma s_{\mathbf{v}}}{4} + i\sigma^2\frac{\gamma s_{\mathbf{x}} + s_{\mathbf{v}}}{4},$$

with omitted inputs $(\mathbf{z}, \bar{\mathbf{z}})$. Therefore, the SDE for $[\mathbf{x}_t, \mathbf{v}_t]^{\mathsf{T}}$ takes the form

$$\begin{bmatrix}\mathrm{d}\mathbf{x}_t \\ \mathrm{d}\mathbf{v}_t\end{bmatrix} = \Bigg(\underbrace{\frac{\sigma^2}{4}\begin{bmatrix}\nabla_{\mathbf{x}}\log p(\mathbf{x}_t, \mathbf{v}_t) \\ \nabla_{\mathbf{v}}\log p(\mathbf{x}_t, \mathbf{v}_t)\end{bmatrix}}_{\text{Contraction}} + \underbrace{\frac{\sigma^2}{4}\gamma\begin{bmatrix}-\nabla_{\mathbf{v}}\log p(\mathbf{x}_t, \mathbf{v}_t) \\ \nabla_{\mathbf{x}}\log p(\mathbf{x}_t, \mathbf{v}_t)\end{bmatrix}}_{\text{Forward dynamics}}\Bigg)\,\mathrm{d}t + \underbrace{\frac{\sigma}{\sqrt{2}}\begin{bmatrix}\mathrm{d}\mathbf{w}_t \\ \mathrm{d}\hat{\mathbf{w}}_t\end{bmatrix}}_{\text{Diffusion}}.$$

For each $\mathbf{x}$ and $\mathbf{v}$, this recovers Langevin dynamics using a real-valued score function $\log p(\mathbf{x}, \mathbf{v}) \propto -E(\mathbf{x}, \mathbf{v})$ for some energy $E(\cdot, \cdot)$.

Conversely, for $[\mathbf{x}_t, -\mathbf{v}_t]^{\mathsf{T}}$ and $s_{\mathbf{z}} = \nabla_{\mathbf{z}}\log\rho$, applying (35) yields

$$\begin{aligned}i\sigma^2\nabla_{\mathbf{z}}\psi &= i\sigma^2\frac{\gamma s_{\mathbf{z}} - i s_{\mathbf{z}}}{2} = i\sigma^2\left(\frac{\gamma s_{\mathbf{x}} - s_{\mathbf{v}}}{4} - i\frac{s_{\mathbf{x}} + \gamma s_{\mathbf{v}}}{4}\right) \\ &= \sigma^2\frac{s_{\mathbf{x}} + \gamma s_{\mathbf{v}}}{4} + i\sigma^2\frac{\gamma s_{\mathbf{x}} - s_{\mathbf{v}}}{4} = \sigma^2\frac{\overleftarrow{s}_{\mathbf{x}} - \gamma\overleftarrow{s}_{-\mathbf{v}}}{4} + i\sigma^2\frac{\gamma\overleftarrow{s}_{\mathbf{x}} + \overleftarrow{s}_{-\mathbf{v}}}{4},\end{aligned}$$

where $\overleftarrow{s}_{\mathbf{x}} = \nabla_{\mathbf{x}}\log\overleftarrow{p}(\mathbf{x}, -\mathbf{v})$ and $\overleftarrow{s}_{-\mathbf{v}} = \nabla_{(-\mathbf{v})}\log\overleftarrow{p}(\mathbf{x}, -\mathbf{v}) = -s_{\mathbf{v}}$ denote the score functions of the reversed time process. Thus, this yields

$$\begin{bmatrix}\mathrm{d}\mathbf{x}_t \\ -\,\mathrm{d}\mathbf{v}_t\end{bmatrix} = \left(\frac{\sigma^2}{4}\begin{bmatrix}\nabla_{\mathbf{x}}\log\overleftarrow{p}(\mathbf{x}_t, -\mathbf{v}_t) \\ \nabla_{(-\mathbf{v})}\log\overleftarrow{p}(\mathbf{x}_t, -\mathbf{v}_t)\end{bmatrix} + \frac{\sigma^2}{4}\gamma\begin{bmatrix}-\nabla_{(-\mathbf{v})}\log\overleftarrow{p}(\mathbf{x}_t, -\mathbf{v}_t) \\ \nabla_{\mathbf{x}}\log\overleftarrow{p}(\mathbf{x}_t, -\mathbf{v}_t)\end{bmatrix}\right)\mathrm{d}s + \frac{\sigma}{\sqrt{2}}\begin{bmatrix}\mathrm{d}\mathbf{w}_s \\ \mathrm{d}\hat{\mathbf{w}}_s\end{bmatrix}$$

The proof is completed by considering ergodicity and the time reversibility of Hamiltonian Markov chains [27]. □

C.2 Proof of Proposition 1 (CSB-PDEs)

For a time-dependent smooth complex function ψ under the conditions of Appendix A.3, consider the MF Hamiltonian:

$$H(\mathbf{z}, \bar{\mathbf{z}}, \nabla_{\mathbf{z}}\psi, \nabla_{\bar{\mathbf{z}}}\psi, \rho, \sigma) = \sigma^2\big\langle \nabla_{\mathbf{z}}\psi, \nabla_{\mathbf{z}}\bar{\psi}\big\rangle + (\nabla_{\mathbf{z}}\psi)^{\mathsf{T}} f - (\nabla_{\bar{\mathbf{z}}}\psi)^{\mathsf{T}}\bar{f} + \frac{i\gamma\sigma^2}{2}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}. \tag{37}$$

The potentials $\Phi, \hat{\Phi} \in C^{2,1}(\mathbb{C}^d, [0,T])$ are parameterized via the complex Hopf–Cole transform (25) in Appendix A.3 for a positive, real-valued density ρ_t satisfying $\rho_t = \Phi(\cdot,\cdot,t)\hat{\Phi}(\cdot,\cdot,t)$. Applying the results from Appendix A.3, with F added to the HJ equation, we derive $(\frac{\partial\Phi}{\partial t}, \frac{\partial\hat{\Phi}}{\partial t})$. Substituting both equations, we obtain the following PDEs:

$$\begin{cases} \dfrac{\partial\Phi(\mathbf{z},\bar{\mathbf{z}},t)}{\partial t} = -i\nabla_{\mathbf{z}}\Phi^{\mathsf{T}} f + i\nabla_{\bar{\mathbf{z}}}\Phi^{\mathsf{T}}\bar{f} - \dfrac{\sigma^2}{4}c_\gamma\Delta\Phi + \dfrac{i\gamma\sigma^2}{2c_\gamma}\dfrac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\Phi + c_\gamma^{-1}F\Phi, & \text{(38a)} \\ \dfrac{\partial\hat{\Phi}(\mathbf{z},\bar{\mathbf{z}},t)}{\partial t} = -i\nabla_{\mathbf{z}}\cdot(\hat{\Phi}f) + i\nabla_{\bar{\mathbf{z}}}\cdot(\hat{\Phi}\bar{f}) + \dfrac{\sigma^2}{4}c_\gamma\Delta\hat{\Phi} - \dfrac{i\gamma\sigma^2}{2c_\gamma}\dfrac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\hat{\Phi} - c_\gamma^{-1}F\hat{\Phi} & \text{(38b)} \end{cases}$$

which hold for $t \in [0,T]$, where Δ denotes the Laplacian and F is a real-valued functional modeling the mean-field interaction. Furthermore, $\hat{\Phi}$ is defined by ρ/Φ, and the boundary conditions $(\Phi\hat{\Phi})(\cdot,\cdot,0) = \rho_0$ and $(\Phi\hat{\Phi})(\cdot,\cdot,T) = \rho_T$ are imposed. This completes the proof. □

For the reverse Hopf–Cole transform, denote the conjugate of the backward potential in (11b) by $\varphi := \bar{\eta} + \gamma\log\rho = -\eta$. In the reversed time coordinate $s = T - t \in [0,T]$, the Hopf–Cole factors are

$$\Phi'(\mathbf{z}_s, \bar{\mathbf{z}}_s, s) = \rho\exp\Big(-\frac{i}{c_\gamma}\bar{\varphi}(\mathbf{z}_s, \bar{\mathbf{z}}_s, T-s)\Big), \tag{39a}$$

$$\hat{\Phi}'(\mathbf{z}_s, \bar{\mathbf{z}}_s, s) = \exp\Big(\frac{i}{c_\gamma}\bar{\varphi}(\mathbf{z}_s, \bar{\mathbf{z}}_s, T-s)\Big), \tag{39b}$$

so $\Phi'\hat{\Phi}' = \rho$. The primed factors denote the Hopf–Cole pair in (38) reindexed by $t = T - s$, with all coefficients evaluated at $t = T - s$. For $F = 0$, applying $\partial_s = -\partial_t$ and the drift-divergence condition constructed in Appendix A.3 gives us a system of reversed-time PDEs

$$\frac{\partial\hat{\Phi}'}{\partial s} = -i\nabla_{\bar{\mathbf{z}}}\hat{\Phi}'^{\mathsf{T}}\bar{f} + i\nabla_{\mathbf{z}}\hat{\Phi}'^{\mathsf{T}} f - \frac{\sigma^2}{4}c_\gamma\Delta\hat{\Phi}' + \frac{i\gamma\sigma^2}{2c_\gamma}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\hat{\Phi}', \tag{40a}$$

$$\frac{\partial\Phi'}{\partial s} = -i\nabla_{\bar{\mathbf{z}}}\cdot(\Phi'\bar{f}) + i\nabla_{\mathbf{z}}\cdot(\Phi' f) + \frac{\sigma^2}{4}c_\gamma\Delta\Phi' - \frac{i\gamma\sigma^2}{2c_\gamma}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\Phi'. \tag{40b}$$

This yields the time-reversed system as the following corollary.

Corollary 4. *For the MF Hamiltonian Eq.* (37)*, applying the Hopf–Cole transform Eq.* (39) *in the reversed time coordinate yields the following PDEs:*

$$\begin{cases} \dfrac{\partial\hat{\Phi}'(\mathbf{z},\bar{\mathbf{z}},s)}{\partial s} = -i\nabla_{\bar{\mathbf{z}}}\hat{\Phi}'^{\mathsf{T}}\bar{f} + i\nabla_{\mathbf{z}}\hat{\Phi}'^{\mathsf{T}} f - \dfrac{\sigma^2}{4}c_\gamma\Delta\hat{\Phi}' + \dfrac{i\gamma\sigma^2}{2c_\gamma}\dfrac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\hat{\Phi}' + c_\gamma^{-1}F\hat{\Phi}', & \text{(41a)} \\ \dfrac{\partial\Phi'(\mathbf{z},\bar{\mathbf{z}},s)}{\partial s} = -i\nabla_{\bar{\mathbf{z}}}\cdot(\Phi'\bar{f}) + i\nabla_{\mathbf{z}}\cdot(\Phi' f) + \dfrac{\sigma^2}{4}c_\gamma\Delta\Phi' - \dfrac{i\gamma\sigma^2}{2c_\gamma}\dfrac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\Phi' - c_\gamma^{-1}F\Phi', & \text{(41b)} \end{cases}$$

where the operator $\nabla_{\bar{\mathbf{z}}}$ here denotes the Wirtinger gradient of $\hat{\Phi}'(\mathbf{z}_s, \bar{\mathbf{z}}_s, s)$ according to Definition 1, i.e. differentiation with respect to $\bar{\mathbf{z}}_s$.

Utilizing two PDE systems of Proposition 1 and Corollary 4, we can now derive SB-CVSDEs in the following proof of Theorem 1.

C.3 Proof of Theorem 1 (SB-CVSDEs)

We first state an expanded version of the SB theorem for the reverse time coordinate s.

Theorem 1 (SB-CVSDEs). *Suppose that $(Y_t, \widehat{Y}_t)$ are the potentials with $Y_t \coloneqq \chi(\mathbf{z}_t, \bar{\mathbf{z}}_t, t)$ and $\widehat{Y}_t \coloneqq -\eta(\mathbf{z}_t, \bar{\mathbf{z}}_t, t)$. Let the base drift f be analytic and satisfy $\nabla_{\mathbf{z}} \cdot f_t = \nabla_{\bar{\mathbf{z}}} \cdot \bar{f}_t$. Then, the CVSDEs corresponding to the CSB-PDE in Proposition 1 for the evolution of Eq.* (11a) *are given by*

$$\underset{\substack{\textit{CVSDEs}: \\ \textit{wrt}\ (11a)}}{\textit{SB}} \begin{cases} \mathrm{d}Z_t = i\big(f_t + \sigma^2 \nabla_{\bar{\mathbf{z}}} \psi_t\big)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, & (42a) \\ \mathrm{d}Y_t = -i\big(\|v_t\|^2 + F_t + \gamma\sigma\nabla_{\mathbf{z}} \cdot \check{v}_t - \frac{i\gamma}{2}\|\hat{v}_t + \check{v}_t\|^2\big)\,\mathrm{d}t + v_t^{\mathsf{T}}\,\mathrm{d}W_t + \hat{v}_t^{\mathsf{T}}\,\mathrm{d}\bar{W}_t, & (42b) \\ \mathrm{d}\widehat{Y}_t = -i\big(\|v_t\|^2 + F_t - \gamma\sigma\nabla_{\mathbf{z}} \cdot \hat{v}_t - (1+\frac{i\gamma}{2})\|\hat{v}_t + \check{v}_t\|^2\big)\,\mathrm{d}t - & (42c) \\ \qquad \big(\sigma\nabla_{\mathbf{z}} \cdot \check{v}_t + \sigma\nabla_{\bar{\mathbf{z}}} \cdot \bar{v}_t\big)\,\mathrm{d}t - \bar{v}_t^{\mathsf{T}}\,\mathrm{d}W_t - \check{v}_t^{\mathsf{T}}\,\mathrm{d}\bar{W}_t, \end{cases}$$

where $\psi = \chi + \gamma \log \rho$, $\Phi = e^{i\psi/c_\gamma}$, $\hat{\Phi} = \rho e^{i\bar{\chi}/c_\gamma} = e^{i\eta/c_\gamma}$ and $\Phi\hat{\Phi} = \rho$; $(v_t, \hat{v}_t, \check{v}_t, \bar{v}_t)$ denote the Wirtinger gradients for χ and η

$$v_t \coloneqq \sigma\nabla_{\mathbf{z}}\chi(\mathbf{z}_t, \bar{\mathbf{z}}_t, t),\ \ \hat{v}_t \coloneqq \sigma\nabla_{\bar{\mathbf{z}}}\chi(\mathbf{z}_t, \bar{\mathbf{z}}_t, t),\ \ \check{v}_t \coloneqq \sigma\nabla_{\bar{\mathbf{z}}}\eta(\mathbf{z}_t, \bar{\mathbf{z}}_t, t),\ \ \bar{v}_t \coloneqq \sigma\nabla_{\mathbf{z}}\eta(\mathbf{z}_t, \bar{\mathbf{z}}_t, t).$$

Conversely, conjugating (11b) *gives the evolution of Z_s in increasing reverse time $s = T - t$. The CSB-PDEs in Corollary 4 then yield*

$$\underset{\substack{\textit{CVSDEs}: \\ \textit{wrt}\ (11b)}}{\textit{SB}} \begin{cases} \mathrm{d}Z_s = i\big(-f_s - \sigma^2 \nabla_{\bar{\mathbf{z}}} \varphi_s\big)\,\mathrm{d}s + \sigma\,\mathrm{d}W_s, & (43a) \\ \mathrm{d}\widehat{Y}_s = i\big(\|\check{v}_s\|^2 + F_s + \gamma\sigma\nabla_{\bar{\mathbf{z}}} \cdot v_s - \frac{i\gamma}{2}\|v_s + \bar{v}_s\|^2\big)\,\mathrm{d}s - \check{v}_s^{\mathsf{T}}\,\mathrm{d}\bar{W}_s - \bar{v}_s^{\mathsf{T}}\,\mathrm{d}W_s, & (43b) \\ \mathrm{d}Y_s = i\big(\|\check{v}_s\|^2 + F_s - \gamma\sigma\nabla_{\mathbf{z}} \cdot \check{v}_s - (1+\frac{i\gamma}{2})\|v_s + \bar{v}_s\|^2\big)\,\mathrm{d}s + & (43c) \\ \qquad \big(\sigma\nabla_{\mathbf{z}} \cdot \hat{v}_s + \sigma\nabla_{\bar{\mathbf{z}}} \cdot v_s\big)\,\mathrm{d}s + v_s^{\mathsf{T}}\,\mathrm{d}W_s + \hat{v}_s^{\mathsf{T}}\,\mathrm{d}\bar{W}_s, \end{cases}$$

where $\varphi = \bar{\eta} + \gamma \log \rho = -\eta$, $\hat{\Phi}' = e^{i\bar{\varphi}/c_\gamma}$, $\Phi' = \rho e^{i\bar{\eta}/c_\gamma} = e^{i\chi/c_\gamma}$ and $\Phi'\hat{\Phi}' = \rho$. Here $(Y_s, \widehat{Y}_s) \coloneqq (\chi, -\eta)$, and the Wirtinger gradients are

$$v_s = \sigma\nabla_{\mathbf{z}}\chi(\mathbf{z}_s, \bar{\mathbf{z}}_s, T\text{-}s), \hat{v}_s = \sigma\nabla_{\bar{\mathbf{z}}}\chi(\mathbf{z}_s, \bar{\mathbf{z}}_s, T\text{-}s), \check{v}_s = \sigma\nabla_{\bar{\mathbf{z}}}\eta(\mathbf{z}_s, \bar{\mathbf{z}}_s, T\text{-}s), \bar{v}_s = \sigma\nabla_{\mathbf{z}}\eta(\mathbf{z}_s, \bar{\mathbf{z}}_s, T\text{-}s).$$

When the forward and reverse potentials parameterize the same Hopf–Cole pair and the reverse process starts from ρ_T, the law of Z_s in (43a) *is ρ_{T-s}.*

Proof. To apply the complex Itô formula in Lemma 9, let us first apply the complex Hopf–Cole transform Eq. (25) [35, 17, 50] as observed in Proposition 1, and let $\chi, \eta \in C^{2,1}(\mathbb{C}^d, [0, T])$ be functions that represent the variables $Y = \chi$ and $\widehat{Y} = -\eta$ with the relation $\widehat{Y}_t = Y_t + i \log \rho_t$, using compatible logarithm branches. The last equality of each combined formula below uses the stated equality of the base-drift divergences. We first verify that the heat-type CSB-PDEs simplify the second-order term of the complex Itô formula.

Next, apply the complex Itô formula (Lemma 9) to the forward drift (42a)

$$\begin{aligned} \mathrm{d}\chi =& \frac{\partial\chi}{\partial t}\,\mathrm{d}t + \left[i(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}\big(f + \sigma^2\nabla_{\bar{\mathbf{z}}}\psi\big) - i(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\big(\overline{f + \sigma^2\nabla_{\bar{\mathbf{z}}}\psi}\big) + \frac{\sigma^2}{4}\Delta\chi\right]\mathrm{d}t + \\ & \sigma(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}\,\mathrm{d}W_t + \sigma(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\,\mathrm{d}\bar{W}_t, \end{aligned} \tag{44}$$

where the above expression uses simplified notation, such as ψ for ψ_t. Note that, from PDEs (38a) and (38b) and the Hopf–Cole relations above, which give $\chi + \eta = -i \log \rho$ and $\partial_t \log \rho = \Phi^{-1}\partial_t\Phi + \hat{\Phi}^{-1}\partial_t\hat{\Phi}$, the general expression for $\frac{\partial\chi}{\partial t}$ is derived as

$$\begin{aligned} \frac{\partial\chi}{\partial t} &= -\frac{ic_\gamma}{\Phi}\left(-i\nabla_{\mathbf{z}}\Phi^{\mathsf{T}}f + i\nabla_{\bar{\mathbf{z}}}\Phi^{\mathsf{T}}\bar{f} - \frac{\sigma^2}{4}c_\gamma\Delta\Phi + \frac{i\gamma\sigma^2}{2c_\gamma}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\Phi + c_\gamma^{-1}F\Phi\right) - \gamma\frac{\partial}{\partial t}\log\rho \\ &= -i(\nabla_{\mathbf{z}}\psi)^{\mathsf{T}}f + i(\nabla_{\bar{\mathbf{z}}}\psi)^{\mathsf{T}}\bar{f} + \sigma^2\frac{ic_\gamma^2}{\Phi}\nabla_{\bar{\mathbf{z}}} \cdot \left(\frac{i}{c_\gamma}\Phi\nabla_{\mathbf{z}}\psi\right) + \frac{\gamma\sigma^2}{2}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}} - iF - \gamma\frac{\partial}{\partial t}\log\rho \\ &= -i(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}f + i(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\bar{f} + i\gamma\big(\nabla_{\mathbf{z}} \cdot f - \nabla_{\bar{\mathbf{z}}} \cdot \bar{f}\big) \\ &\quad - i\sigma^2\nabla_{\bar{\mathbf{z}}}\chi^{\mathsf{T}}\nabla_{\mathbf{z}}\chi - \frac{\sigma^2}{4}\Delta(\chi + i\gamma\eta) - \frac{\gamma}{2}\|\sigma\nabla_{\bar{\mathbf{z}}}(\chi + \eta)\|^2 - iF. \end{aligned}$$

Combining the time derivative with the whole Itô formula and using the conjugation and linearity properties, we obtain

$$\begin{aligned}
\mathrm{d}\chi &= -i\Big(-\sigma^2\nabla_{\mathbf{z}}\chi^{\mathsf{T}}\nabla_{\bar{\mathbf{z}}}\psi + \sigma^2\nabla_{\bar{\mathbf{z}}}\chi^{\mathsf{T}}\overline{\nabla_{\bar{\mathbf{z}}}\psi} + \sigma^2\nabla_{\bar{\mathbf{z}}}\chi^{\mathsf{T}}\nabla_{\mathbf{z}}\chi + F + \frac{\gamma\sigma^2}{4}\Delta\eta \\
&\quad - \frac{i\gamma}{2}\|\sigma\nabla_{\bar{\mathbf{z}}}(\chi+\eta)\|^2\Big)\,\mathrm{d}t + i\gamma\big(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}\big)\,\mathrm{d}t + \sigma(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}\,\mathrm{d}W_t + \sigma(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\,\mathrm{d}\bar{W}_t \\
&= -i\Big(\|\sigma\nabla_{\mathbf{z}}\chi\|^2 + F + \frac{\gamma\sigma^2}{4}\Delta\eta - \frac{i\gamma}{2}\|\sigma\nabla_{\bar{\mathbf{z}}}(\chi+\eta)\|^2\Big)\,\mathrm{d}t + \\
&\quad i\gamma\big(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}\big)\,\mathrm{d}t + \sigma(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}\,\mathrm{d}W_t + \sigma(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\,\mathrm{d}\bar{W}_t \\
&= -i\big(\|v_t\|^2 + F_t + \gamma\sigma\nabla_{\mathbf{z}}\cdot\check{v}_t - \frac{i\gamma}{2}\|\hat{v}_t+\check{v}_t\|^2\big)\,\mathrm{d}t + v_t^{\mathsf{T}}\,\mathrm{d}W_t + \hat{v}_t^{\mathsf{T}}\,\mathrm{d}\bar{W}_t.
\end{aligned}$$

This implies that ψ and χ are the associated mathematical objects in forward SDE integration. Applying the complex Itô lemma to η, we obtain

$$\begin{aligned}
\mathrm{d}\eta =& \frac{\partial\eta}{\partial t}\,\mathrm{d}t + \Big[i(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}\big(f + \sigma^2\nabla_{\bar{\mathbf{z}}}\psi\big) - i(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\big(\overline{f + \sigma^2\nabla_{\bar{\mathbf{z}}}\psi}\big) + \frac{\sigma^2}{4}\Delta\eta\Big]\,\mathrm{d}t + \\
&\sigma(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}\,\mathrm{d}W_t + \sigma(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\,\mathrm{d}\bar{W}_t,
\end{aligned} \tag{45}$$

and for PDE (38b) the general expression for $\frac{\partial\eta}{\partial t}$ is derived as

$$\begin{aligned}
\frac{\partial\eta}{\partial t} &= -\frac{ic_\gamma}{\hat{\Phi}}\left(-i\nabla_{\mathbf{z}}\cdot\big(\hat{\Phi}f\big) + i\nabla_{\bar{\mathbf{z}}}\cdot\big(\hat{\Phi}\bar{f}\big) + \frac{\sigma^2}{4}c_\gamma\Delta\hat{\Phi} - \frac{i\gamma\sigma^2}{2c_\gamma}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\hat{\Phi} - c_\gamma^{-1}F\hat{\Phi}\right) \\
&= -i(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}f - c_\gamma\nabla_{\mathbf{z}}\cdot f + i(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\bar{f} + c_\gamma\nabla_{\bar{\mathbf{z}}}\cdot\bar{f} \\
&\quad - \sigma^2\frac{ic_\gamma^2}{\hat{\Phi}}\nabla_{\mathbf{z}}\cdot\left(\frac{i}{c_\gamma}\hat{\Phi}\nabla_{\bar{\mathbf{z}}}\eta\right) - \frac{\gamma\sigma^2}{2}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}} + iF \\
&= -i(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}f + i(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\bar{f} - c_\gamma\big(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}\big) \\
&\quad + i\sigma^2\nabla_{\bar{\mathbf{z}}}\eta^{\mathsf{T}}\nabla_{\mathbf{z}}\eta + \frac{\sigma^2}{4}\Delta(\eta - i\gamma\chi) - \frac{\gamma}{2}\|\sigma\nabla_{\bar{\mathbf{z}}}(\chi+\eta)\|^2 + iF.
\end{aligned}$$

Combining the above formulas yields

$$\begin{aligned}
-\,\mathrm{d}\eta &= -i\Big(\sigma^2\nabla_{\mathbf{z}}\eta^{\mathsf{T}}\nabla_{\bar{\mathbf{z}}}\psi - \sigma^2\nabla_{\bar{\mathbf{z}}}\eta^{\mathsf{T}}\overline{\nabla_{\bar{\mathbf{z}}}\psi} + \sigma^2\nabla_{\bar{\mathbf{z}}}\eta^{\mathsf{T}}\nabla_{\mathbf{z}}\eta \\
&\qquad + F - \frac{\gamma\sigma^2}{4}\Delta\chi + \frac{i\gamma}{2}\|\sigma\nabla_{\bar{\mathbf{z}}}(\chi+\eta)\|^2\Big)\,\mathrm{d}t + \\
&\qquad \big(c_\gamma(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}) - \frac{\sigma^2}{2}\Delta\eta\big)\,\mathrm{d}t - \sigma(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}\,\mathrm{d}W_t - \sigma(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\,\mathrm{d}\bar{W}_t \\
&= -i\Big(\|\sigma\nabla_{\mathbf{z}}\chi\|^2 + F - \frac{\gamma\sigma^2}{4}\Delta\chi - \big(1+\frac{i\gamma}{2}\big)\|\sigma\nabla_{\bar{\mathbf{z}}}(\chi+\eta)\|^2\Big)\,\mathrm{d}t + \\
&\qquad \big(c_\gamma(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}) - \frac{\sigma^2}{2}\Delta\eta\big)\,\mathrm{d}t - \sigma(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}\,\mathrm{d}W_t - \sigma(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\,\mathrm{d}\bar{W}_t \\
&= -i\big(\|v_t\|^2 + F_t - \gamma\sigma\nabla_{\mathbf{z}}\cdot\hat{v}_t - \big(1+\frac{i\gamma}{2}\big)\|\hat{v}_t+\check{v}_t\|^2\big)\,\mathrm{d}t - \\
&\qquad \sigma\big(\nabla_{\mathbf{z}}\cdot\check{v}_t + \nabla_{\bar{\mathbf{z}}}\cdot\bar{v}_t\big)\,\mathrm{d}t - \bar{v}_t^{\mathsf{T}}\,\mathrm{d}W_t - \check{v}_t^{\mathsf{T}}\,\mathrm{d}\bar{W}_t.
\end{aligned}$$

At corresponding times $t = T - s$, the matched Hopf–Cole factors satisfy $\Phi' = \Phi$ and $\hat{\Phi}' = \hat{\Phi}$. Comparing their exponents in (25) and (39) on compatible logarithm branches gives

$$(\chi,\eta) = \begin{cases} (-\bar{\psi},\bar{\varphi}) & \text{in the forward parameterization,} \\ (\psi,-\varphi) & \text{in the backward parameterization.} \end{cases}$$

Conjugating the forward identity $\chi + \eta = -i\log\rho$ gives $\varphi = \psi + i\log\rho$. Taking its $\nabla_{\bar{\mathbf{z}}}$ gradient gives the drift of Z_s in (43a) directly from (42a),

$$i\big(-f - \sigma^2\nabla_{\bar{\mathbf{z}}}\varphi\big) = -i\big(f + \sigma^2\nabla_{\bar{\mathbf{z}}}\psi\big) + \sigma^2\nabla_{\bar{\mathbf{z}}}\log\rho.$$

The reverse process initialized with the forward terminal law ρ_T therefore has marginal law ρ_{T-s}.

For independently fitted potentials, let f, b denote the forward and backward parameterizations and ρ the marginal density of the fixed sampling law. Under $\mathbb{P}_\theta$, $\eta_{\mathrm{f}}^{\phi} - \eta_{\mathrm{b}}^{\phi} = \gamma \log \rho = -2\Re\chi_{\mathrm{f}}^{\theta}$, while under $\mathbb{Q}_\phi$, $\chi_{\mathrm{b}}^{\theta} - \chi_{\mathrm{f}}^{\theta} = \gamma \log \rho = -2\Re\eta_{\mathrm{b}}^{\phi}$. Taking Wirtinger gradients gives

$$\overline{v_{\mathrm{f}}^{\theta}} - \check{v}_{\mathrm{b}}^{\phi} = \begin{cases} -\hat{v}_{\mathrm{f}}^{\theta} - \check{v}_{\mathrm{f}}^{\phi} & \text{under } \mathbb{P}_\theta, \\ v_{\mathrm{b}}^{\theta} + \bar{v}_{\mathrm{b}}^{\phi} & \text{under } \mathbb{Q}_\phi. \end{cases}$$

Multiplying by $-i\sigma$ gives the drift sum in $\mathbf{z}$-coordinates. Each shift adds $\frac{\gamma\sigma}{4}\Delta \log \rho \in \mathbb{R}$ to the corresponding divergence, preserving its imaginary part.

Below we suppress the backward subscripts on χ and η, and ∂_s differentiates the time-reindexed potentials. Applying the complex Itô lemma to η using (43a), we obtain

$$\begin{aligned} \mathrm{d}\eta = \frac{\partial \eta}{\partial s}\,\mathrm{d}s + \Big[i(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\big(\,\bar{f} + \sigma^2\nabla_{\mathbf{z}}\bar{\varphi}\,\big) - i(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}\big(\,f + \sigma^2\nabla_{\bar{\mathbf{z}}}\varphi\,\big) + \frac{\sigma^2}{4}\Delta\eta \Big]\,\mathrm{d}s \\ + \sigma(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\,\mathrm{d}\bar{W}_s + \sigma(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}\,\mathrm{d}W_s. \end{aligned} \tag{46}$$

From PDEs (41a) and (41b) in Corollary 4 and the Hopf–Cole relations above, which give $\chi + \eta = -i\log\rho$ and $\partial_s \log\rho = (\Phi')^{-1}\partial_s\Phi' + (\hat{\Phi}')^{-1}\partial_s\hat{\Phi}'$, the general expression for $\frac{\partial\eta}{\partial s}$ is

$$\begin{aligned} \frac{\partial\eta}{\partial s} &= -\frac{ic_\gamma}{\hat{\Phi}'}\left(-i\nabla_{\bar{\mathbf{z}}}\hat{\Phi}'^{\mathsf{T}}\bar{f} + i\nabla_{\mathbf{z}}\hat{\Phi}'^{\mathsf{T}}f - \frac{\sigma^2}{4}c_\gamma\Delta\hat{\Phi}' + \frac{i\gamma\sigma^2}{2c_\gamma}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\hat{\Phi}' + c_\gamma^{-1}F\hat{\Phi}'\right) - \gamma\frac{\partial}{\partial s}\log\rho \\ &= i(\nabla_{\mathbf{z}}\bar{\varphi})^{\mathsf{T}}f - i(\nabla_{\bar{\mathbf{z}}}\bar{\varphi})^{\mathsf{T}}\bar{f} + \sigma^2\frac{ic_\gamma^2}{\hat{\Phi}'}\nabla_{\bar{\mathbf{z}}}\cdot\left(\frac{i}{c_\gamma}\hat{\Phi}'\nabla_{\mathbf{z}}\bar{\varphi}\right) + \frac{\gamma\sigma^2}{2}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}} - iF - \gamma\frac{\partial}{\partial s}\log\rho \\ &= i(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}f - i(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\bar{f} - i\gamma\big(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}\big) \\ &\quad - i\sigma^2\nabla_{\bar{\mathbf{z}}}\eta^{\mathsf{T}}\nabla_{\mathbf{z}}\eta - \frac{\sigma^2}{4}\Delta(\eta + i\gamma\chi) - \frac{\gamma}{2}\|\sigma\nabla_{\bar{\mathbf{z}}}(\chi+\eta)\|^2 - iF. \end{aligned}$$

By repeating a similar process, we obtain

$$\begin{aligned} \mathrm{d}\eta &= -i\Big(-\sigma^2\nabla_{\bar{\mathbf{z}}}\eta^{\mathsf{T}}\nabla_{\mathbf{z}}\bar{\varphi} + \sigma^2\nabla_{\mathbf{z}}\eta^{\mathsf{T}}\nabla_{\bar{\mathbf{z}}}\varphi + \sigma^2\nabla_{\bar{\mathbf{z}}}\eta^{\mathsf{T}}\nabla_{\mathbf{z}}\eta + F + \frac{\gamma\sigma^2}{4}\Delta\chi \\ &\quad - \frac{i\gamma}{2}\|\sigma\nabla_{\bar{\mathbf{z}}}(\chi+\eta)\|^2\Big)\,\mathrm{d}s - i\gamma\big(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}\big)\,\mathrm{d}s + \sigma(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\,\mathrm{d}\bar{W}_s + \sigma(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}\,\mathrm{d}W_s \\ &= -i\Big(\|\sigma\nabla_{\bar{\mathbf{z}}}\eta\|^2 + F + \frac{\gamma\sigma^2}{4}\Delta\chi - \frac{i\gamma}{2}\|\sigma\nabla_{\mathbf{z}}(\chi+\eta)\|^2\Big)\,\mathrm{d}s \\ &\quad - i\gamma\big(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}\big)\,\mathrm{d}s + \sigma(\nabla_{\bar{\mathbf{z}}}\eta)^{\mathsf{T}}\,\mathrm{d}\bar{W}_s + \sigma(\nabla_{\mathbf{z}}\eta)^{\mathsf{T}}\,\mathrm{d}W_s \\ &= -i\big(\|\check{v}_s\|^2 + F_s + \gamma\sigma\nabla_{\bar{\mathbf{z}}}\cdot v_s - \frac{i\gamma}{2}\|v_s + \bar{v}_s\|^2\big)\,\mathrm{d}s + \check{v}_s^{\mathsf{T}}\,\mathrm{d}\bar{W}_s + \bar{v}_s^{\mathsf{T}}\,\mathrm{d}W_s. \end{aligned}$$

Since $\widehat{Y}_s = -\eta$, we have $\mathrm{d}\widehat{Y}_s = -\,\mathrm{d}\eta$.

Lastly, applying the complex Itô lemma to χ using (43a), we obtain

$$\begin{aligned} \mathrm{d}\chi = \frac{\partial\chi}{\partial s}\,\mathrm{d}s + \Big[i(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\big(\,\bar{f} + \sigma^2\nabla_{\mathbf{z}}\bar{\varphi}\,\big) - i(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}\big(\,f + \sigma^2\nabla_{\bar{\mathbf{z}}}\varphi\,\big) + \frac{\sigma^2}{4}\Delta\chi \Big]\,\mathrm{d}s \\ + \sigma(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\,\mathrm{d}\bar{W}_s + \sigma(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}\,\mathrm{d}W_s. \end{aligned} \tag{47}$$

From PDE (41b), the general expression for $\frac{\partial\chi}{\partial s}$ is derived as

$$\begin{aligned} \frac{\partial\chi}{\partial s} &= -\frac{ic_\gamma}{\Phi'}\left(-i\nabla_{\bar{\mathbf{z}}}\cdot\big(\,\Phi'\bar{f}\,\big) + i\nabla_{\mathbf{z}}\cdot\big(\,\Phi' f\,\big) + \frac{\sigma^2}{4}c_\gamma\Delta\Phi' - \frac{i\gamma\sigma^2}{2c_\gamma}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}\Phi' - c_\gamma^{-1}F\Phi'\right) \\ &= i(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}f + c_\gamma\nabla_{\mathbf{z}}\cdot f - i(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\bar{f} - c_\gamma\nabla_{\bar{\mathbf{z}}}\cdot\bar{f} \\ &\quad - \sigma^2\frac{ic_\gamma^2}{\Phi'}\nabla_{\mathbf{z}}\cdot\left(\frac{i}{c_\gamma}\Phi'\nabla_{\bar{\mathbf{z}}}\chi\right) - \frac{\gamma\sigma^2}{2}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}} + iF \\ &= i(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}f - i(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\bar{f} + c_\gamma\big(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}\big) \\ &\quad + i\sigma^2\nabla_{\bar{\mathbf{z}}}\chi^{\mathsf{T}}\nabla_{\mathbf{z}}\chi + \frac{\sigma^2}{4}\Delta(\chi - i\gamma\eta) - \frac{\gamma}{2}\|\sigma\nabla_{\mathbf{z}}(\chi+\eta)\|^2 + iF. \end{aligned}$$

With the same procedure as for Eq. (42c), combining the above formulas yields

$$
\begin{aligned}
\mathrm{d}\chi = -i\Big(&-\sigma^2\nabla_{\bar{\mathbf{z}}}\chi^{\mathsf{T}}\nabla_{\mathbf{z}}\bar{\varphi} + \sigma^2\nabla_{\mathbf{z}}\chi^{\mathsf{T}}\nabla_{\bar{\mathbf{z}}}\varphi - \sigma^2\nabla_{\bar{\mathbf{z}}}\chi^{\mathsf{T}}\nabla_{\mathbf{z}}\chi \\
&- F + \frac{\gamma\sigma^2}{4}\Delta\eta - \frac{i\gamma}{2}\|\sigma\nabla_{\mathbf{z}}(\chi+\eta)\|^2\Big)\,\mathrm{d}s + \\
&\big(c_\gamma(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}) + \frac{\sigma^2}{2}\Delta\chi\big)\,\mathrm{d}s + \sigma(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\,\mathrm{d}\bar{W}_s + \sigma(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}\,\mathrm{d}W_s \\
= -i\Big(&-\|\sigma\nabla_{\bar{\mathbf{z}}}\eta\|^2 - F + \frac{\gamma\sigma^2}{4}\Delta\eta + \big(1+\frac{i\gamma}{2}\big)\|\sigma\nabla_{\mathbf{z}}(\chi+\eta)\|^2\Big)\,\mathrm{d}s + \\
&\big(c_\gamma(\nabla_{\mathbf{z}}\cdot f - \nabla_{\bar{\mathbf{z}}}\cdot\bar{f}) + \frac{\sigma^2}{2}\Delta\chi\big)\,\mathrm{d}s + \sigma(\nabla_{\bar{\mathbf{z}}}\chi)^{\mathsf{T}}\,\mathrm{d}\bar{W}_s + \sigma(\nabla_{\mathbf{z}}\chi)^{\mathsf{T}}\,\mathrm{d}W_s \\
= -i\big(&-\|\check{v}_s\|^2 - F_s + \gamma\sigma\nabla_{\mathbf{z}}\cdot\check{v}_s + \big(1+\frac{i\gamma}{2}\big)\|v_s+\bar{v}_s\|^2\big)\,\mathrm{d}s + \\
&\sigma\big(\nabla_{\mathbf{z}}\cdot\hat{v}_s + \nabla_{\bar{\mathbf{z}}}\cdot v_s\big)\,\mathrm{d}s + \hat{v}_s^{\mathsf{T}}\,\mathrm{d}\bar{W}_s + v_s^{\mathsf{T}}\,\mathrm{d}W_s,
\end{aligned}
$$

which proves Eq. (43c). □

C.4 Proof of Theorem 2 (iterative proportional fitting)

We assume the Hopf–Cole reality condition for each law's own marginal density, with $\mathbb{P}_{\theta,0} = \rho_0$ and $\mathbb{Q}_{\phi,T} = \rho_T$, and the regularity and integrability required for Girsanov's theorem, time reversal, and the log-density Itô formulas. We establish both KL representations. Fix θ and take expectations under $\mathbb{P}_\theta$, with t-indexed fields in the forward parameterization. Using Theorem 1 with $F = 0$, the parameterization identities in its proof, the change of measure (9), and the time-reversal calculation of Liu et al. [53, Proposition 9], we obtain

$$
\begin{aligned}
\mathrm{KL}(\mathbb{P}_\theta\|\,\mathbb{Q}_\phi) = \mathbb{E}_{\mathbb{P}_\theta}\big[\log\rho_0(Z_0) - \log\rho_T(Z_T)\big] \\
+ \int_0^T \mathbb{E}_{\mathbb{P}_\theta}\Big[\big\|\hat{v}_t^\theta + \check{v}_t^\phi\big\|^2 - 2\sigma\Im\big\{\nabla_{\mathbf{z}}\cdot\check{v}_t^\phi\big\}\Big]\mathrm{d}t.
\end{aligned}
$$

Indeed, the real diffusion covariance is $(\sigma^2/2)\mathbf{I}_{2d}$. Applying Itô's formula to the log marginal density of $\mathbb{Q}_\phi$ cancels the score terms in the squared drift difference. Its boundary contribution cancels the ϕ-dependent initial log-density in Girsanov's initial density ratio, leaving logarithms of the fixed boundary densities ρ_0 and ρ_T. Under the assumed integrability, the stochastic integrals have zero expectation, and the base-drift condition gives $\Im\nabla_{\mathbf{z}}\cdot f = 0$.

Since $\nabla_{\mathbf{z}}\cdot\check{v}_t^\phi = \nabla_{\bar{\mathbf{z}}}\cdot\bar{v}_t^\phi$, taking the imaginary part of $\mathcal{E}_T^-$ gives

$$
\Im\big[\mathcal{E}_T^-(\theta;\phi)\big] = \int_0^T \mathbb{E}_{\mathbb{P}_\theta}\Big[\big\|\hat{v}_t^\theta + \check{v}_t^\phi\big\|^2 - 2\sigma\Im\big\{\nabla_{\mathbf{z}}\cdot\check{v}_t^\phi\big\}\Big]\mathrm{d}t - \int_0^T \mathbb{E}_{\mathbb{P}_\theta}\Big[\|v_t^\theta\|^2 - \gamma\sigma\Re\big\{\nabla_{\mathbf{z}}\cdot\hat{v}_t^\theta\big\}\Big]\mathrm{d}t.
$$

The endpoint expectation and the integral subtracted in the preceding display are independent of ϕ. Denoting their sum by C_θ, we have $\mathrm{KL}(\mathbb{P}_\theta\|\,\mathbb{Q}_\phi) = \Im[\mathcal{E}_T^-(\theta;\phi)] + C_\theta$. Thus this projection updates ϕ while keeping θ and the sampling law $\mathbb{P}_\theta$ fixed.

For the reverse IPF, take expectations under $\mathbb{Q}_\phi$, with s-indexed fields in the backward parameterization at $t = T - s$ and the path law expressed in the same $\mathbf{z}$-coordinates. The same argument gives

$$
\begin{aligned}
\mathrm{KL}(\mathbb{Q}_\phi\|\,\mathbb{P}_\theta) = \mathbb{E}_{\mathbb{Q}_\phi}\big[\log\rho_T(Z_T) - \log\rho_0(Z_0)\big] \\
+ \int_0^T \mathbb{E}_{\mathbb{Q}_\phi}\Big[\big\|v_s^\theta + \bar{v}_s^\phi\big\|^2 - 2\sigma\Im\big\{\nabla_{\bar{\mathbf{z}}}\cdot v_s^\theta\big\}\Big]\mathrm{d}s.
\end{aligned}
$$

Using $\nabla_{\mathbf{z}}\cdot\hat{v}_s^\theta = \nabla_{\bar{\mathbf{z}}}\cdot v_s^\theta$ in the definition of $\mathcal{E}_0^+$ gives

$$
\begin{aligned}
-\Im\big[\mathcal{E}_0^+(\phi;\theta)\big] = \int_0^T \mathbb{E}_{\mathbb{Q}_\phi}\Big[\big\|v_s^\theta + \bar{v}_s^\phi\big\|^2 - 2\sigma\Im\big\{\nabla_{\bar{\mathbf{z}}}\cdot v_s^\theta\big\}\Big]\mathrm{d}s \\
- \int_0^T \mathbb{E}_{\mathbb{Q}_\phi}\Big[\|\check{v}_s^\phi\|^2 - \gamma\sigma\Re\big\{\nabla_{\mathbf{z}}\cdot\check{v}_s^\phi\big\}\Big]\mathrm{d}s.
\end{aligned}
$$

The endpoint expectation and the integral subtracted in the preceding display are independent of θ. Denoting their sum by C_ϕ, we have $\mathrm{KL}(\mathbb{Q}_\phi \| \mathbb{P}_\theta) = -\Im[\mathcal{E}_0^+(\phi;\theta)] + C_\phi$. Thus this projection updates θ while keeping ϕ and the sampling law $\mathbb{Q}_\phi$ fixed.

To match the active-parameter labels of the training losses, let θ' and ϕ' denote frozen parameter copies and define

$$\begin{aligned}\mathcal{L}^+_{\mathrm{IPF}}(\phi;\theta') &:= \Im[\mathcal{E}_T^-(\theta';\phi)], && \text{with expectations under } \mathbb{P}_{\theta'},\\ \mathcal{L}^-_{\mathrm{IPF}}(\theta;\phi') &:= -\Im[\mathcal{E}_0^+(\phi';\theta)], && \text{with expectations under } \mathbb{Q}_{\phi'}.\end{aligned}$$

Here the signs on $\mathcal{E}^\pm$ specify the integration coordinates (s for $+$ and t for $-$), whereas the signs on $\mathcal{L}^\pm_{\mathrm{IPF}}$ follow the active-parameter labels of the training losses. Furthermore, suppose that the change to the backward parameterization and the reality condition for $\mathbb{Q}_{\phi'}$ give

$$\varphi_{\phi'} = \widehat{Y}_s^{\phi'}, \qquad \psi_\theta = \gamma \log \rho - \overline{Y_s^\theta} = 2\Re \widehat{Y}_s^{\phi'} - \overline{Y_s^\theta},$$

and also $Y_s^\theta - \widehat{Y}_s^{\phi'} = \overline{\varphi_{\phi'} - \psi_\theta}$ and $\Im \Delta Y_s^\theta = \Im \Delta \psi_\theta$. Substituting these identities into the reverse KL integrand gives the Hyvärinen score-matching objective [39]

$$\arg\min_\theta \mathcal{L}^-_{\mathrm{IPF}}(\theta;\phi') = \arg\min_\theta \int_0^T \mathbb{E}_{\mathbb{Q}_{\phi'}}\Big[\big\|\sigma\nabla_{\bar{\mathbf{z}}}\psi_\theta - \sigma\nabla_{\bar{\mathbf{z}}}\varphi_{\phi'}\big\|^2 - \frac{1}{2}\sigma^2 \Im\Delta\psi_\theta\Big]\,\mathrm{d}s, \tag{48}$$

considering $\varphi_{\phi'}$ as a fixed target. □

C.5 Proof of Proposition 2 (analytic posterior)

For the variance-time generalization of COT_γ in Eq. (54), consider the endpoint-conditioned CVSDE

$$\mathrm{d}Z_t = -c_\gamma \frac{\sigma_t^2}{\int_t^T \sigma_\tau^2\,\mathrm{d}\tau}(Z_t - \mathbf{z}_T)\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t, \qquad Z_0 = \mathbf{z}_0, \quad 0 \le t < T. \tag{49}$$

The posterior formula holds for $0 < s < T$, while the endpoint laws degenerate to point masses at $\bar{\mathbf{z}}_T$ for $s = 0$ and at $\bar{\mathbf{z}}_0$ for $s = T$.

For the complex exponential function $h_\gamma(\ell) = e^{c_\gamma \ell}$, the parametrization $\ell = \log(1 - t/T)$ yields a logarithmic spiral traversed at constant Euclidean speed. Consider first the COT_γ subproblem with constant σ_t for absolutely continuous functions:

$$\inf_{\substack{u(0)=\mathbf{z}_0\\ u(T)=\mathbf{z}_T}} \int_0^T L_\gamma\big(u(t), \dot{u}(t), t\big)\,\mathrm{d}t, \qquad u \in AC([0,T], \mathbb{C}^d). \tag{50}$$

For paths of finite action, the transformed path $(1 - t/T)^{-i\gamma}(u(t) - \mathbf{z}_T)$ has a square-integrable derivative and extends absolutely continuously to $t = T$. We apply the Cauchy–Schwarz inequality to the objective against the constant function of value 1. Combining this with the fundamental theorem of calculus and the endpoint conditions [23]

$$(1 - t/T)^{-i\gamma}(u(t) - \mathbf{z}_T)\big|_{t=0} = \mathbf{z}_0 - \mathbf{z}_T, \qquad \lim_{t\uparrow T}(1 - t/T)^{-i\gamma}(u(t) - \mathbf{z}_T) = \mathbf{0},$$

we obtain the lower bound

$$\begin{aligned}\int_0^T L_\gamma\big(u(t), \dot{u}(t), t\big)\,\mathrm{d}t &= \int_0^T \Big\|\frac{\mathrm{d}}{\mathrm{d}t}\Big[(1 - t/T)^{-i\gamma}(u(t) - \mathbf{z}_T)\Big]\Big\|^2\,\mathrm{d}t\\ &\ge \frac{1}{T}\Big\|\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\Big[(1 - t/T)^{-i\gamma}(u(t) - \mathbf{z}_T)\Big]\,\mathrm{d}t\Big\|^2 = \frac{\|\mathbf{z}_0 - \mathbf{z}_T\|^2}{T}.\end{aligned}$$

For an optimal forward path $\vec{u} \in AC([0,T], \mathbb{C}^d)$, the optimality condition

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[(1 - t/T)^{-i\gamma}(\vec{u}(t) - \mathbf{z}_T)\Big] = \frac{\mathbf{z}_T - \mathbf{z}_0}{T} \tag{51}$$

holds almost everywhere. Recall that for $a \in \mathbb{R}^+$ and $b \in \mathbb{C}$, complex conjugation satisfies

$$\overline{a^b} = \overline{e^{b \log a}} = e^{\bar{b} \log a} = a^{\bar{b}}.$$

Combining the endpoint conditions with the optimality condition in Eq. (51), we obtain the minimizing forward path and its conjugate time reversal in the form

$$\begin{cases} \vec{u}(t) = \mathbf{z}_T + (\mathbf{z}_0 - \mathbf{z}_T)(1 - t/T)^{c_\gamma}, & 0 \le t < T, \quad (52a) \\ \overleftarrow{u}(s) = \bar{\mathbf{z}}_T + (\bar{\mathbf{z}}_0 - \bar{\mathbf{z}}_T)(s/T)^{\bar{c}_\gamma}, & 0 < s \le T. \quad (52b) \end{cases}$$

These paths extend continuously to $\vec{u}(T) = \mathbf{z}_T$ and $\overleftarrow{u}(0) = \bar{\mathbf{z}}_T$, with $\overleftarrow{u}(T) = \bar{\mathbf{z}}_0$ and $\vec{u}(t) = \overleftarrow{u}(T-t)$. In particular, $\|\dot{\vec{u}}(t)\| = \sqrt{1+\gamma^2}\|\mathbf{z}_0 - \mathbf{z}_T\|/T$, so the forward path is absolutely continuous and has constant speed.

Suppose that σ_t is scheduled in a time-dependent manner. We use reverse time for the diffusion coefficient σ and define the accumulated and remaining variances as

$$\hat{\omega}_s^2 = \int_0^s \hat{\sigma}_\tau^2 \, \mathrm{d}\tau, \qquad \omega_s^2 \coloneqq \int_s^T \hat{\sigma}_\tau^2 \, \mathrm{d}\tau, \qquad 0 \le s \le T.$$

Using these accumulated variances, we restate the objective following the diffusion scheduling argument of Liu et al. [55].

Definition 10 (Generalization of Eq. (8))**.** We set

$$r_s \coloneqq \frac{\hat{\omega}_s^2}{\hat{\omega}_s^2 + \omega_s^2} = \frac{\hat{\omega}_s^2}{\hat{\omega}_T^2}, \qquad k_s = r_s^{\bar{c}_\gamma}, \qquad 0 \le s \le T, \tag{53}$$

where $k_0 = 0$ is defined by continuity and $\mathbb{Q}_W$ denotes the reference law with zero drift and diffusion coefficient σ_t. Replacing $1 - t/T$ by r_{T-t}, we generalize the KL objective and redefine the rotation and its induced drift as

$$\begin{aligned} &\underset{\mathbb{P}^\gamma_{0,T} = \rho_{0,T}}{\text{minimize}} \, \mathrm{KL}(\mathbb{P}^\gamma \| \mathbb{Q}_W^\gamma), \quad \text{s.t.} \quad \mathbb{Q}_W^\gamma = (\mathcal{H}_\gamma)_\#(\mathbb{Q}_W), \\ &\qquad (\mathcal{H}_\gamma \mathbf{z})_t \coloneqq \mathbf{z}_T + \frac{h_\gamma(\log r_{T-t})}{h_0(\log r_{T-t})}(\mathbf{z}_t - \mathbf{z}_T), \\ &b_\gamma(\mathbf{z}, t; \mathbf{z}_T) \coloneqq i\gamma \frac{\mathrm{d}}{\mathrm{d}t} \log r_{T-t} \, (\mathbf{z} - \mathbf{z}_T) = -\frac{i\gamma\sigma_t^2}{\hat{\omega}_{T-t}^2}(\mathbf{z} - \mathbf{z}_T). \end{aligned} \tag{54}$$

These formulas hold for $0 \le t < T$, with $(\mathcal{H}_\gamma u)_T = \mathbf{z}_T$ defined by continuity.

Applying the Cauchy–Schwarz inequality with weight σ_t^2 to the variance-time action yields

$$\begin{aligned} \int_0^T \frac{\|\dot{u}(t) - b_\gamma(u(t), t; \mathbf{z}_T)\|^2}{\sigma_t^2} \, \mathrm{d}t &= \int_0^T \frac{1}{\sigma_t^2} \left\| \frac{\mathrm{d}}{\mathrm{d}t} \left[r_{T-t}^{-i\gamma}(u(t) - \mathbf{z}_T) \right] \right\|^2 \mathrm{d}t \\ &\ge \frac{1}{\hat{\omega}_T^2} \left\| \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \left[r_{T-t}^{-i\gamma}(u(t) - \mathbf{z}_T) \right] \mathrm{d}t \right\|^2 = \frac{\|\mathbf{z}_0 - \mathbf{z}_T\|^2}{\hat{\omega}_T^2}. \end{aligned}$$

Equality holds when the transformed derivative equals $(\mathbf{z}_T - \mathbf{z}_0)\sigma_t^2/\hat{\omega}_T^2$. Given the endpoint conditions, the minimizing forward path is $\mathbf{z}_T + (\mathbf{z}_0 - \mathbf{z}_T) r_{T-t}^{c_\gamma}$, and its conjugate time reversal takes the form

$$\overleftarrow{u}(s) = \bar{\mathbf{z}}_T + (\bar{\mathbf{z}}_0 - \bar{\mathbf{z}}_T) k_s. \tag{55}$$

For a constant $\sigma_t = \sigma$, the variance-time action reduces to the action in Eq. (50) divided by σ^2, and the minimizing paths recover Eq. (52).

When the minimum in (54) is finite, the optimizer and the reference process have the same bridge for almost every endpoint pair under the optimal coupling [51, Proposition 2.3]. Since $\mathcal{H}_\gamma$ leaves both endpoints fixed, applying it to the ordinary Gaussian bridge gives this shared bridge. To obtain the posterior distribution, let $Z_t^{(0)}$ denote the ordinary Gaussian bridge obtained from Eq. (49) with $\gamma = 0$. Gaussian conditioning, as in Liu et al. [55, Proposition 3.3], gives

$$\bar{Z}_s^{(0)} \mid \bar{\mathbf{z}}_T, \bar{\mathbf{z}}_0 \sim \mathcal{CN}\left(\bar{\mathbf{z}}_T + (\bar{\mathbf{z}}_0 - \bar{\mathbf{z}}_T) r_s, \frac{\omega_s^2 \hat{\omega}_s^2}{\omega_s^2 + \hat{\omega}_s^2} \mathbf{I}_d \right). \tag{56}$$

Applying the rotation in Eq. (54) to this bridge yields

$$\begin{aligned} Z_t &= \mathbf{z}_T + r_{T-t}^{i\gamma}\big(Z_t^{(0)} - \mathbf{z}_T\big), \\ \bar{Z}_s &= \bar{\mathbf{z}}_T + r_s^{-i\gamma}\big(\bar{Z}_s^{(0)} - \bar{\mathbf{z}}_T\big). \end{aligned} \tag{57}$$

Using Itô's formula [67] and $\frac{\mathrm{d}}{\mathrm{d}t}\log r_{T-t} = -\sigma_t^2/\hat{\omega}_{T-t}^2$, we obtain

$$\mathrm{d}Z_t = -c_\gamma \frac{\sigma_t^2}{\hat{\omega}_{T-t}^2}(Z_t - \mathbf{z}_T)\,\mathrm{d}t + \sigma_t r_{T-t}^{i\gamma}\,\mathrm{d}W_t^{(0)}.$$

Since $|r_{T-t}^{i\gamma}| = 1$, the stochastic integral $W_t = \int_0^t r_{T-\tau}^{i\gamma}\,\mathrm{d}W_\tau^{(0)}$ is a standard proper complex Wiener process. This verifies Eq. (49); uniqueness on each compact subinterval of $[0, T)$ identifies its solution with Eq. (57). Moreover, $\|Z_t - \mathbf{z}_T\| = \|Z_t^{(0)} - \mathbf{z}_T\| \to 0$ almost surely as $t \uparrow T$, which establishes the continuous endpoint extension.

The transformation in Eq. (57) is affine and preserves the isotropic variance of the Gaussian noise. The posterior therefore remains proper complex Gaussian, with mean and covariance

$$\bar{\mu}_s = \bar{\mathbf{z}}_T + r_s^{-i\gamma}\big(\mathbb{E}\big[\bar{Z}_s^{(0)}\big] - \bar{\mathbf{z}}_T\big) = \bar{\mathbf{z}}_T + (\bar{\mathbf{z}}_0 - \bar{\mathbf{z}}_T)k_s,$$
$$\Sigma_s = |r_s^{-i\gamma}|^2 \frac{\omega_s^2\hat{\omega}_s^2}{\omega_s^2 + \hat{\omega}_s^2}\mathbf{I}_d = \frac{\omega_s^2\hat{\omega}_s^2}{\omega_s^2 + \hat{\omega}_s^2}\mathbf{I}_d.$$

This proves Eq. (15). Furthermore, expanding and rearranging the terms yields the marginal mean

$$\bar{\mu}_s = k_s\bar{\mathbf{z}}_0 + (1 - k_s)\bar{\mathbf{z}}_T.$$

For the discrete updates, let $0 = t_0 < \cdots < t_N = T$ with $N \geq 2$, set $s_n = T - t_n$, and evaluate r_n, k_n, ω_n^2, $\hat{\omega}_n^2$, $\bar{\mu}_n$, and Σ_n at s_n. Write $\bar{\mathbf{z}}_n$ for a realization of $\bar{Z}_{s_n}$, so that $\bar{\mathbf{z}}_N = \bar{\mathbf{z}}_T$ and $\bar{\mathbf{z}}_0$ remains the initial forward endpoint. All subsequent laws are conditional on this boundary pair. To sample from the posterior, we run the reverse process from step $n+1$ to step n.

For the ordinary bridge, Gaussian conditioning gives [55, Appendix A]

$$\mathbb{E}\big[\bar{Z}_{s_n}^{(0)} \mid \bar{Z}_{s_{n+1}}^{(0)}\big] = \frac{\hat{\omega}_n^2 - \hat{\omega}_{n+1}^2}{\omega_{n+1}^2}\bar{\mathbf{z}}_0 + \frac{\omega_n^2}{\omega_{n+1}^2}\bar{Z}_{s_{n+1}}^{(0)},$$
$$\mathrm{Cov}\big(\bar{Z}_{s_n}^{(0)} \mid \bar{Z}_{s_{n+1}}^{(0)}\big) = \frac{\omega_n^2(\hat{\omega}_n^2 - \hat{\omega}_{n+1}^2)}{\omega_{n+1}^2}\mathbf{I}_d.$$

For $1 \leq n \leq N-2$, we substitute the inverse rotation

$$\bar{\mathbf{z}}_{n+1}^{(0)} = \bar{\mathbf{z}}_N + r_{n+1}^{i\gamma}(\bar{\mathbf{z}}_{n+1} - \bar{\mathbf{z}}_N)$$

and apply the rotation at step n. Rearranging the resulting expression yields the discrete posterior transition and its moments:

$$q(\bar{\mathbf{z}}_n \mid \bar{\mathbf{z}}_{n+1}, \bar{\mathbf{z}}_0, \bar{\mathbf{z}}_N) = \mathcal{CN}(\bar{\mathbf{z}}_n \mid \bar{\mu}_{n|n+1}, \Sigma_{n|n+1}), \tag{58a}$$
$$\bar{\mu}_{n|n+1} = \bar{\mu}_n + \frac{\omega_n^2}{\omega_{n+1}^2}\left(\frac{r_n}{r_{n+1}}\right)^{-i\gamma}(\bar{\mathbf{z}}_{n+1} - \bar{\mu}_{n+1}), \tag{58b}$$
$$\Sigma_{n|n+1} = \frac{\omega_n^2(\hat{\omega}_n^2 - \hat{\omega}_{n+1}^2)}{\omega_{n+1}^2}\mathbf{I}_d. \tag{58c}$$

At the initial reverse step $n = N-1$, the state is sampled directly from Eq. (15), since $r_N = 0$ and the endpoint is fixed. At the final step $n = 0$, the distribution reduces to the point mass at $\bar{\mathbf{z}}_0$.

Taking the expectation of Eq. (58b) and substituting the inductive hypothesis $\mathbb{E}[\bar{Z}_{s_{n+1}}] = \bar{\mu}_{n+1}$ yields

$$\mathbb{E}[\bar{Z}_{s_n}] = \bar{\mu}_n + \frac{\omega_n^2}{\omega_{n+1}^2}\left(\frac{r_n}{r_{n+1}}\right)^{-i\gamma}\big(\mathbb{E}[\bar{Z}_{s_{n+1}}] - \bar{\mu}_{n+1}\big) = \bar{\mu}_n.$$

For the covariance, the rotation again has unit modulus. Using $\hat{\omega}_n^2 - \hat{\omega}_{n+1}^2 = \omega_{n+1}^2 - \omega_n^2$ and the fact that $\omega_n^2 + \hat{\omega}_n^2$ is constant, we have

$$\Sigma_{n|n+1} + \left(\frac{\omega_n^2}{\omega_{n+1}^2}\right)^2\Sigma_{n+1} = \left[\frac{\omega_n^2(\hat{\omega}_n^2 - \hat{\omega}_{n+1}^2)}{\omega_{n+1}^2} + \frac{\omega_n^4\hat{\omega}_{n+1}^2}{\omega_{n+1}^2(\omega_{n+1}^2 + \hat{\omega}_{n+1}^2)}\right]\mathbf{I}_d$$
$$= \frac{\omega_n^2\hat{\omega}_n^2}{\omega_n^2 + \hat{\omega}_n^2}\mathbf{I}_d = \Sigma_n.$$

These match the exact marginal mean and covariance in Eq. (15). Since Gaussian conditional distributions with affine means and constant covariances compose to a Gaussian, the reverse step guarantees by induction that the sampler reproduces the marginal distribution in Eq. (15) at each step. If $\gamma = 0$, the rotation vanishes and Eq. (58) reduces to the ordinary I^{2}SB posterior. □

D Implementation Details

D.1 Mean-field simulation

Across all experiments, we employed the AdamW optimizer with a decay factor of 10^{-2}. SDE integration was performed via the Euler–Maruyama scheme with step size $\Delta t = 0.005$, where the total time T varied depending on the tasks [54]. For the marginal distribution on momenta, we used scaled normal distributions $\mathbf{v}_0 \sim \mathcal{N}(\mathbf{0}, \varsigma^2 \mathbf{I}_d)$, setting the scale larger than the local modal dispersion of the position distributions. Specifically, we used $\varsigma = 0.5$ for moons-circles (see Fig. 12 for trajectories), $\varsigma = 2.0$ for GMM-8, $\varsigma = 1.5$ for S-Tunnel, and $\varsigma = 1.0$ for V-Neck.

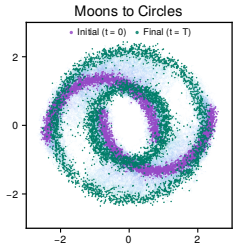

Figure 12: Moons to circles.

TD Learning. The temporal difference estimates $\mathrm{TD}_s^+(\phi;\theta')$ and $\mathrm{TD}_t^-(\theta;\phi')$ are approximated as

$$i\big(\|\check{v}_s^{\phi}\|^2 - \gamma\sigma\nabla_{\mathbf{z}}\cdot\check{v}_s^{\phi} - \big(1+\tfrac{i\gamma}{2}\big)\|v_s^{\theta'}+\bar{v}_s^{\phi}\|^2 - i\sigma\nabla_{\mathbf{z}}\cdot\hat{v}_s^{\theta'} - i\sigma\nabla_{\bar{\mathbf{z}}}\cdot v_s^{\theta'}\big)\delta s + (v_s^{\theta'})^{\intercal}\delta W_s + (\hat{v}_s^{\theta'})^{\intercal}\delta\bar{W}_s$$
$$-i\big(\|v_t^{\theta}\|^2 - \gamma\sigma\nabla_{\mathbf{z}}\cdot\hat{v}_t^{\theta} - \big(1+\tfrac{i\gamma}{2}\big)\|\hat{v}_t^{\theta}+\check{v}_t^{\phi'}\|^2 - i\sigma\nabla_{\mathbf{z}}\cdot\check{v}_t^{\phi'} - i\sigma\nabla_{\bar{\mathbf{z}}}\cdot\bar{v}_t^{\phi'}\big)\delta t - (\bar{v}_t^{\phi'})^{\intercal}\delta W_t - (\check{v}_t^{\phi'})^{\intercal}\delta\bar{W}_t,$$

where F_s and F_t can be added to the respective expressions for nonzero external MF costs [53]. For base drifts, we keep the value constant or zero because the divergence terms produce side effects in TD learning. Raw returns are constructed from one or more TD steps, incorporating the appropriate boundary or bootstrap values. Applying the Hopf–Cole relation $\widehat{Y} = Y + i\log\rho$ in the corresponding parameterization, we transform these returns into regression targets: $\bar{G}_s^+ = R_s^+ + i\log\rho_s$ and $G_t^- = R_t^- - i\log\rho_t$, with the density evaluated at the return's starting state and time.

For each starting point, let $N \geq 1$ denote the number of remaining steps to the boundary in the respective time direction, with $0 \leq \lambda \leq 1$. We define $R_{s,n}^+$ as the value of Y_θ at the sampled endpoint $s + n\delta s$ minus the sum of the intervening TD^+ increments, and define $R_{t,n}^-$ analogously from $\widehat{Y}_\phi$, $t + n\delta t$, and TD^-. The positive endpoint corresponds to physical time $T - s - n\delta s$. Applying the same conversion yields $\bar{G}_{s,n}^+ = R_{s,n}^+ + i\log\rho_s$ and $G_{t,n}^- = R_{t,n}^- - i\log\rho_t$. Following the undiscounted finite-horizon episodic λ return formulation of $\mathrm{TD}(\lambda)$ [78], we obtain

$$\bar{G}_s^{+,\lambda}(\phi;\theta) = (1-\lambda)\sum_{n=1}^{N-1}\lambda^{n-1}\bar{G}_{s,n}^+(\phi;\theta) + \lambda^{N-1}\bar{G}_{s,N}^+(\phi;\theta) \tag{59a}$$

$$G_t^{-,\lambda}(\theta;\phi) = (1-\lambda)\sum_{n=1}^{N-1}\lambda^{n-1}G_{t,n}^-(\theta;\phi) + \lambda^{N-1}G_{t,N}^-(\theta;\phi) \tag{59b}$$

Higher λ was often the preferred choice due to the interplay of multiple losses and fast syncing with the EMA estimates. We set $\lambda = 1$ for comparison with benchmarks.

Network architecture. For the 2D control problems, we adopted the potential and drift neural network architectures from Liu et al. [53], augmenting both input and output layers with additional momentum coordinates. We also evaluated fully complex-valued networks with holomorphic activation functions; however, the holomorphy constraint ① conflicted with our core assumptions, and ② provided no empirical benefits to justify the added computational complexity and memory overhead. This observation likewise held for the subsequent image generation workflows.

D.2 Image-to-image translation

RGB. For MNIST and CIFAR-10 data, we deployed a simplified custom NCSN_{++} architecture [77][2] featuring two residual blocks. We adopted a symmetric parabolic diffusion schedule $\sigma_t = 4 \cdot 0.012 \cdot t(1-t) + 10^{-5}$ for both cases, and set the initial momentum standard deviation to 0.01, except for the CIFAR-10 inpainting task, for which we used 0.003. For ImageNet256, modeling and evaluation relied on the official I^{2}SB repository[3]. Across all settings, the network architectures remained virtually identical, varying only in their input and output channel dimensions.

[2] https://github.com/yang-song/score_sde
[3] https://github.com/NVlabs/I2SB

Table 3: Fitting a harmonic oscillator trajectory from noisy quantum observations (5 seeds).

Method	$\xi = 10^{-3}$	$\xi = 10^{-2}$	$\xi = 10^{-1}$
NODE [14]	4027144	4197223	234321
HNN [30]	$8.37_{\pm 0.55}$	$7.79_{\pm 1.52}$	$9.35_{\pm 0.77}$
Neural SDE	$1.39_{\pm 2.47}$	$1.54_{\pm 3.19}$	$2.92_{\pm 2.63}$
Ours (potential net)	$0.16_{\pm 0.10}$	$0.35_{\pm 0.33}$	$1.61_{\pm 1.40}$

MRI. We simulated complex-valued brain MRI slices at 64×64 resolution using the open-source package `KomaMRI.jl` [12]. Given a source-target separation of 24 mm between fixed regions of the brain, the objective was to reconstruct the correct amplitude and colorization of a subject's brain image from a source image. In total, we generated 24 synthetic subjects whose characteristics were controlled through anatomical and tissue-property perturbations, yielding 10,000 training pairs (20 subjects) for the SB models and 2,000 validation pairs (4 subjects). For this synthetic benchmark, we employed the standard NCSN_{++} architecture which has a larger total parameter count than the model in the MNIST and CIFAR experiments.

D.3 Circuit quantum electrodynamics simulation

To simulate a closed system, we employed the open-source package `QuantumToolbox.jl` [58] to numerically integrate the Schrödinger equation for a harmonic oscillator with $d = 2$ modes with high precision. The system dynamics are governed by a Hamiltonian incorporating local cavity detuning and inter-mode hopping:

$$\hat{H}_{\text{osc}} = D(\hat{a}_1^\dagger \hat{a}_1) + D(\hat{a}_2^\dagger \hat{a}_2) + J(\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_1 \hat{a}_2^\dagger)$$

where $\hat{a}_j$ and $\hat{a}_j^\dagger$ denote the bosonic annihilation and creation operators for mode $j \in \{1, 2\}$, respectively. We set the detuning parameter to $D = 10\pi$, and the inter-mode coupling strength to $J = 4\pi$. The system was initialized in a pure tensor product state representing a coherent state with an intensity of 4.0 in the first mode and a vacuum state in the second. Deterministic unitary evolution was computed over a total duration $T = 4$. At each step, we extracted the exact expectation values for the position and momentum quadratures, defined respectively as $\hat{x}_i = (\hat{a}_i + \hat{a}_i^\dagger)/2$ and $\hat{p}_i = (\hat{a}_i - \hat{a}_i^\dagger)/(2i)$, which served as our main observation inputs. For data generation, the injected Gaussian noise was non-stationary and explicitly scaled by the instantaneous quantum standard deviation of each observable, *e.g.*, $\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$. We added Gaussian noise proportional to these quantum fluctuations, with a noise scalar ξ.

We tested tracking these time evolutions with Wirtinger gradients of a potential network and compared the potential network's accuracy over 5 seconds of simulated time with that of the Hamiltonian neural net [30], a Neural ODE [14] model and a neural SDE (NFE $= 1,600$). The neural SDE model was trained using Itô drift, whereas the other models were trained with the mean drift obtained by averaging both drifts. For training, we used $5,000$ episodes, and we evaluated using an average over $1,000$ particles. Table 3 shows that estimating drift with Wirtinger gradients gave us solid results in predicting the next step in a system governed by the Schrödinger equation. Even though the neural SDE got closer as ξ diminished, it failed to capture the structure of the dynamics, as seen in Fig. 9 (b).

3D qubit gate transform. Building on the harmonic oscillator benchmarks, we extended our framework to simulate the dynamics of a tripartite hybrid quantum system using `QuantumToolbox.jl`. Specifically, we modeled a three-mode harmonic cavity coupled to a three-mode Josephson junction, which in turn interacted with a two-level qubit [61]. Dynamics are governed by the rotating-frame Hamiltonian:

$$\hat{H} = \hat{H}_{\text{osc}} + \hat{H}_{\text{jj}} + \hat{H}_{\text{qubit}} + \hat{H}_{\text{int}}.$$

We framed these experiments as source-to-target state mapping tasks. The system was initialized with the cavity in coherent states spanning multiple intensities (mean photon numbers in $\{0.5, 1.0, 2.0\}$), the JJ in the vacuum state, and the qubit in canonical states on the Bloch sphere (*e.g.*, north pole, equator). For each configuration, deterministic unitary trajectories were integrated over a duration of $T = 5.0$ across 2,001 time steps. Target states were simulated analogously after first transforming the initial qubit state through ideal unitary gates. Models were trained on 5,000 episodes and evaluated

across 1,000 trajectories in a fully unpaired setting. As a baseline, we trained an MLP on the complete oscillator data using batchwise cost matching, incorporating no physical priors beyond the projection layer. For OT-CFM and SB-CFM [82], we provided an equivalent data budget, a trainable Kerr-nonlinear Hamiltonian for the Josephson junction, and exact cost matrices for target matching. Assuming a flat space, we trained the network with the conditional CSB algorithm, using an OT solver for batchwise couplings where the distances were calculated from conformal geodesics.

D.4 Computational costs

Throughout the experiments, we controlled the internal neural network architectures (*e.g.*, maintaining an identical number of hidden layers) to ensure a fair comparison. Under this setup, the only difference in parameter count lies in the input and output projection layers, which were adjusted to match the dimensionalities of the respective input/output specifications. This discussion concerns comparisons with fixed internal architectures; the two-qubit experiment below also examines a change in architecture. We clarify the computational costs separately for the general CSB formulation and the conditional (including Image-to-Image) CSB algorithms below:

- **General CSB**
  - **Forward pass and particle forward CVSDE.** While particle simulation based on the complex Itô lemma double-counts operations over the real and imaginary components, this overhead is negligible compared to the forward pass through the non-linear neural network. Consequently, trajectory generation at inference remains highly efficient once the model is trained.
  - **Potential backward CVSDE.** Due to the nature of automatic differentiation (vector-Jacobian products), temporal difference computations for the backward CVSDE double the gradient evaluations, while exact Laplacian calculations add non-trivial overhead. To mitigate the cost of second-order differentiation, gradient caching with neural networks [53] can be applied.
- **Conditional CSB**
  - **Forward and backward updates.** In image-based setups, input/output projection differences are practically negligible. Crucially, the costly backward potential updates are replaced with analytic posterior representations, bypassing explicit Laplacian evaluations and drastically reducing sampling latency to match that of standard real-valued diffusion bridge baselines.

E Additional Experimental Results

E.1 Empirical validation of the analytic posterior and hyperparameter selection

Choice of the hyperparameter γ. The benefits of incorporating Hamiltonian dynamics into stochastic processes are well established [62], and our empirical results generally confirm that our approach is consistently effective across datasets. For its validation, we conducted a grid search based on validation performance and empirically observed that a higher value of γ is particularly effective when our theoretical assumptions explicitly hold and nonconservative biases between samples are minimal. To precisely validate this argument, let us define a stochastic interpolant [2] $z_t \sim \mathbb{P}_t^\gamma$ using the analytic posterior curve parameterized by γ. The γ-interpolant is defined as

$$z_t = z_1 + (z_0 - z_1)(1-t)^{1+i\gamma} + 2\sigma\sqrt{t(1-t)}\,\hat{\xi}, \qquad t \in [0,1], \tag{60}$$

where $\hat{\xi}$ denotes the interpolation noise. Now, using this construction, we considered the remaining area under the curve (rAUC), which measures how close the path measure stays to the target. For instance, the 2-Wasserstein rAUC (W_2-rAUC) metric is defined as

$$W_2\text{-rAUC}(\tau) = \frac{1}{T}\int_\tau^T W_2(\mathbb{P}_t^\gamma, \mathbb{Q}_T)\,\mathrm{d}t, \tag{61}$$

where $\mathbb{Q}_T$ denotes the target marginal, and $T = 1$ for the normalized time interval used above. We approximated the integration by partitioning the time span into discrete intervals.

In Table 4, we report the 2-Wasserstein rAUC metric and the squared MMD metric (MMD^2-rAUC) for a representative instance in the cQED simulation task. We observed that $\gamma = 0$ gives the lowest W_2-rAUC and MMD^2-rAUC for $\tau \in \{0.94, 0.95, 0.97\}$ among the tested values. At $\tau = 0.99$, $\gamma = 30$ gives the lowest W_2-rAUC, while $\gamma = 0$ and $\gamma = 30$ tie for the lowest MMD^2-rAUC at the

Table 4: 2-Wasserstein saturation of the analytic posterior for oscillator states of the Hadamard operation. Bold and underlined values denote the best and second-best results.

γ	W_2-rAUC(.94) ↓	MMD^2-rAUC(.94) ↓	W_2-rAUC(.95) ↓	MMD^2-rAUC(.95) ↓	W_2-rAUC(.97) ↓	MMD^2-rAUC(.97) ↓	W_2-rAUC(.99) ↓	MMD^2-rAUC(.99) ↓
0	**0.202734**	**0.051280**	**0.154899**	**0.039132**	**0.063955**	**0.015614**	0.007683	**0.001424**
1	0.203424	0.051459	0.155543	0.039302	0.064253	0.015706	0.007710	0.001433
3	0.203286	0.051405	0.155257	0.039217	0.064030	0.015637	0.007709	0.001433
5	0.203372	0.051429	0.155256	0.039224	0.064083	0.015658	0.007723	0.001437
10	0.203020	0.051384	0.155370	0.039272	0.064292	0.015720	0.007733	0.001441
30	0.203512	0.051452	0.155503	0.039270	0.064053	0.015648	**0.007682**	**0.001424**

Table 5: CIFAR-10 posterior FID scores for different values of γ.

γ	FID at 0.9	FID at 0.95	FID at 0.99	FID at 0.995
0	31.316002	2.152040	0.081899	0.094650
0.0001	31.315914	2.152060	0.081906	0.094652
0.001	**31.315853**	**2.151930**	0.081904	0.094637
0.01	31.321283	2.152645	**0.081816**	**0.094604**
0.1	31.520388	2.184362	0.081928	0.094841
1	47.182999	5.063271	0.094534	0.100584
3	97.487869	23.163961	0.109886	0.110451
9	208.948288	75.324211	0.225691	0.114714

reported precision. We also report FID scores (noise-to-image) in Table 5 with similar tendencies and different scales. As a result, the optimal choice of γ depends largely on data characteristics and the underlying relationship between marginal distributions.

E.2 Two-dimensional control benchmarks

Since $\mathbb{C}^d \cong \mathbb{R}^{2d}$, there is no mathematical distinction between using the space $\mathbb{C}^d$ and using an unconstrained model in $\mathbb{R}^{2d}$ per se.[4] Therefore, the ablation study was conducted by augmenting the state space of DeepGSB [53] to $\mathbb{R}^{2d}$, achieving the same dimensionality as CSB but without the Hamiltonian prior. As shown in Tables 6 and 7, CSB shows substantial improvement over the DeepGSB formulation while being competitive with GSBM [54] in terms of endpoint matching. Notably, CSB obtained the lowest total trajectory costs (both kinetic and potential) on all three tasks, verifying our claim about higher-order dynamics for diffusion bridges.

Table 6: 2-Wasserstein 2D control benchmark results (forward direction, 5 different seeds).

Method	GMM-8	S-Tunnel	V-Neck
DeepGSB ($\mathbb{R}^{2d}$)	3.92 ± 2.15	0.56 ± 0.04	0.81 ± 0.03
GSBM ($\mathbb{R}^d$)	2.65 ± 0.41	$\mathbf{0.17 \pm 0.03}$	$\mathbf{0.09 \pm 0.01}$
CSB ($\gamma = 0.001$, $\mathbb{C}^d$)	1.92 ± 0.83	0.21 ± 0.07	0.61 ± 0.03
CSB ($\gamma = 0.01$, $\mathbb{C}^d$)	$\mathbf{1.91 \pm 0.25}$	0.18 ± 0.05	0.63 ± 0.04
CSB ($\gamma = 0.1$, $\mathbb{C}^d$)	1.99 ± 0.77	0.20 ± 0.08	0.65 ± 0.04
CSB ($\gamma = 1$, $\mathbb{C}^d$)	3.67 ± 1.92	0.23 ± 0.06	0.52 ± 0.19

Table 7: Total trajectory cost calculated from a sum of kinetic energy and obstacle costs (5 runs).

Method	GMM-8	S-Tunnel	V-Neck
DeepGSB ($\mathbb{R}^{2d}$)	228.86 ± 0.61	510.16 ± 4.93	12.00 ± 1.87
GSBM ($\mathbb{R}^d$)	121.48 ± 14.40	484.98 ± 2.30	144.72 ± 0.37
CSB ($\gamma = 0.001$, $\mathbb{C}^d$)	116.06 ± 0.32	$\mathbf{295.66 \pm 36.64}$	8.08 ± 0.68
CSB ($\gamma = 0.01$, $\mathbb{C}^d$)	115.97 ± 0.23	300.97 ± 46.00	8.14 ± 0.56
CSB ($\gamma = 0.1$, $\mathbb{C}^d$)	$\mathbf{115.64 \pm 0.23}$	303.22 ± 41.43	8.10 ± 0.73
CSB ($\gamma = 1$, $\mathbb{C}^d$)	118.80 ± 1.28	306.32 ± 45.83	$\mathbf{4.00 \pm 2.78}$

E.3 Two-qubit gate experiments

Building upon our single-qubit experiment setups, we expanded our evaluation to two-qubit gates by comparing a 4-layer MLP against a 3-block residual network, representing a roughly fourfold increase in raw trainable parameters (total 800K). Notably, we applied weak pairing to faithfully replicate the output of each gate's truth table, and this indicates a single neural network handled the four input cases in each two-qubit gate's truth table. Then, we measured *qubit*

Table 8: Qubit fidelity of CNOT and SWAP 2-qubit logic gates.

	CNOT	SWAP
OT-CFM	0.9993	0.9448
SB-CFM	0.9062	0.9991
Ours ($\gamma = 1$)	**0.9994**	**0.9998**

[4] Nonetheless, this work introduces a compact presentation of the dynamics for a coherent geometric prior and an inductive bias for dealing with coupled position-momentum and magnitude-phase dynamics with useful complex operations.

fidelity, a measure of accuracy and reliability that quantifies how closely an actual quantum state or operation aligns with an ideal target state or operation. The results are summarized in Table 8. In both quantum logic gates (CNOT and SWAP operations), our approach showed stable performance and near-perfect transformation fidelity in conditional SB learning and also demonstrated scalability. The result demonstrates that our OT application for 3D qubits is scalable to larger cQED applications.

E.4 RGB image translation

In Table 9, the FID trends for unpaired MNIST generation mirror those observed for CIFAR-10 in Fig. 7. Because this behavior recurred across image benchmarks, we conclude that our assumption regarding Hamiltonian structures holds broadly, even for static visual scenes. Figs. 13 and 14 show qualitative image-to-image translation results, illustrating how each dataset-specific model configuration contributes to FID improvements. On MNIST, gains are driven by increased mode diversity, whereas on higher-dimensional images, improvements stem primarily from contextual fidelity, followed by sharper object contours.

Table 9: MNIST FID.

Method	FID (↓)
I^2SB	3.7557
Ours ($\gamma = 0.0001$)	2.7138
Ours ($\gamma = 0.001$)	2.881
Ours ($\gamma = 0.01$)	**2.4791**
Ours ($\gamma = 0.1$)	<u>2.5536</u>
Ours ($\gamma = 1.0$)	17.0554

E.5 Brain MRI image translation

MRI preserves both the amplitude and the phase of an oscillating signal; complex numbers are a convenient way to retain both pieces of information. We conducted experiments on a complex-valued brain MRI dataset created with the open-source package `KomaMRI.jl` [12] and comprising 24 simulated subjects generated through controlled anatomical and tissue-property perturbations. As shown in Table 10, rotational dynamics improved the transport of complex MRI images across both real and imaginary signal components, consistent with our theoretical assumptions. Furthermore, all four tested values of γ achieved lower FID and higher magnitude SSIM than I^2SB, with $\gamma = 0.1$ yielding the best performance across both metrics.

For qualitative evaluation on the MRI dataset, Fig. 15 displays image generation results across three distinct subjects for DSBM [76], SB Flow [19], I^2SB [55], and our method. In this comparison, CSB recovered plausible target images in terms of shape, magnitude, and angle. We observed that the improvements in image quality were more pronounced when handling explicit complex-valued data than when preprocessing RGB images with zero-centered momenta. Combined with the quantitative evaluation on the validation set in Table 10, these results on the complex-valued dataset demonstrated the broader applicability of our theoretical framework to scientific data.

Table 10: Magnitude FID and SSIM metrics on the paired 64×64 synthetic brain MRI validation dataset for unseen subjects (source-target slice separation: 24 mm).

Method	FID ↓	SSIM ↑
DSBM	23.1730	0.6126
SB Flow	21.4061	0.6312
I^2SB	15.2626	0.6597
Ours ($\gamma = 0.001$)	12.0784	0.7130
Ours ($\gamma = 0.01$)	12.4734	0.7104
Ours ($\gamma = 0.1$)	**12.0135**	**0.7139**
Ours ($\gamma = 1.0$)	12.8268	0.7129

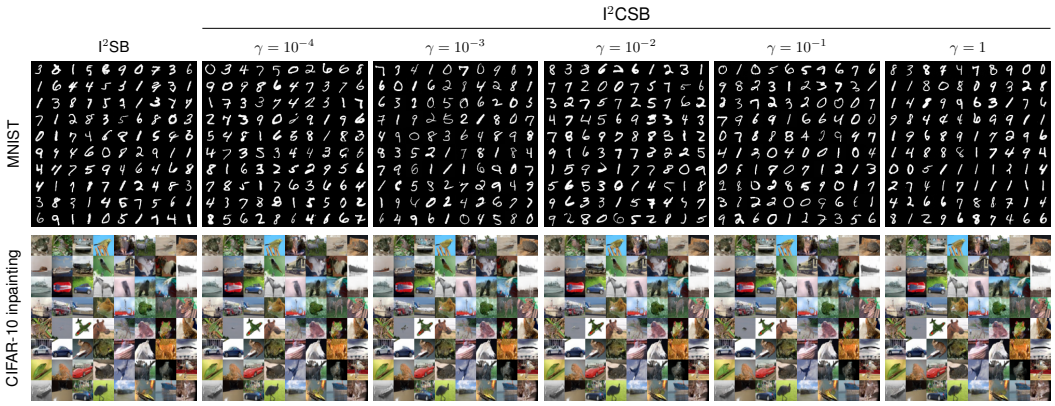

Figure 13: MNIST and CIFAR-10 image-to-image translation



Figure 14: Image recovery from three different types of corruption on the ImageNet256 validation dataset.

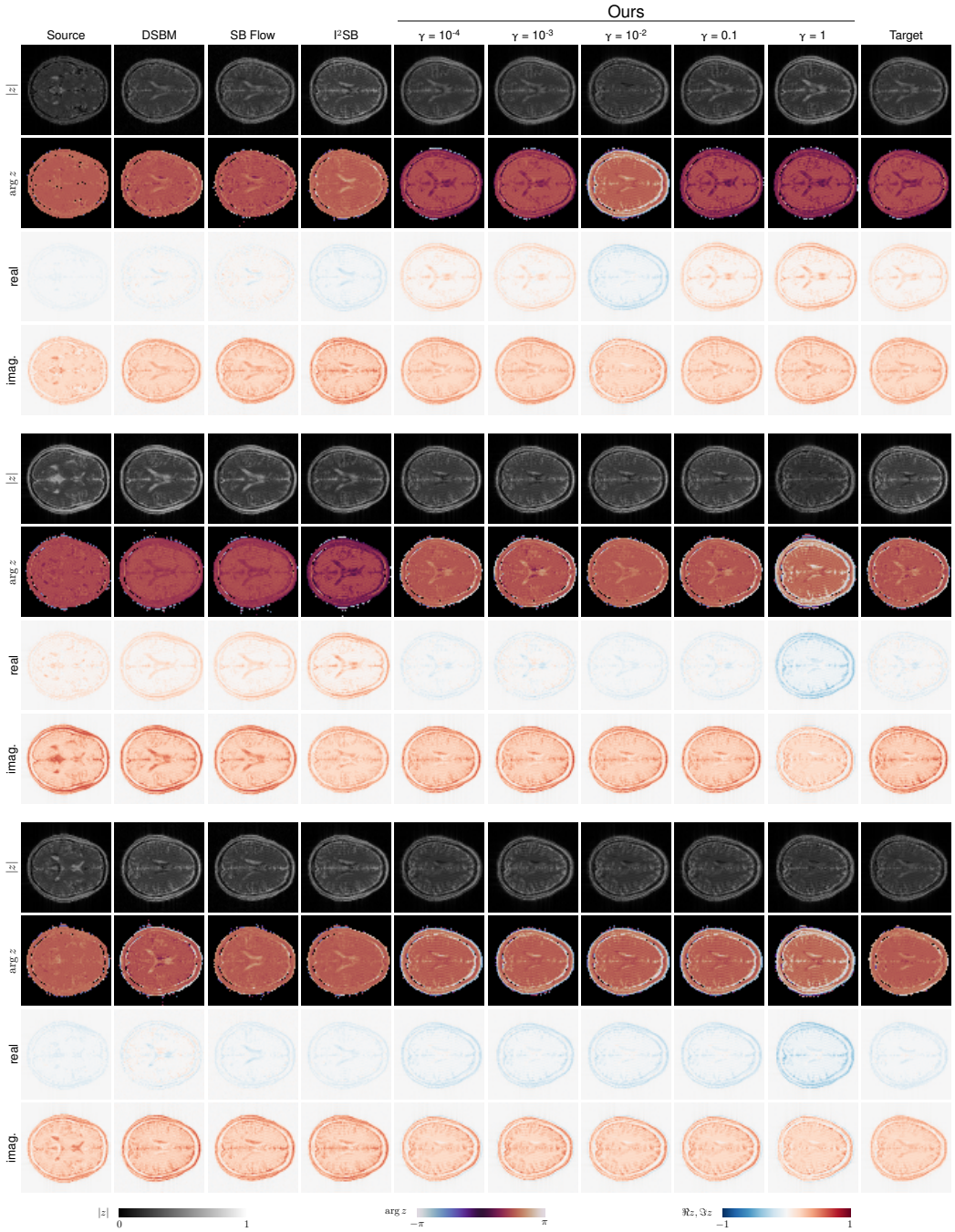

Figure 15: Complex MRI slice-transport results for fixed validation examples. Each block shows magnitude, phase, real, and imaginary components with shared display ranges.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: Our claims made in the abstract and the introduction are justified through theoretical and experimental results. Theoretically, we investigated problem settings, a general algorithm as well as justifications for a simplified algorithm for dealing with high-dimensional data. Empirically, we solve a wide range of problems to consolidate our theory.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: See Section 7.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: See Appendices A-C.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: See Appendix D and the official homepage.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: See `https://circle-group.github.io/research/CSB`.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: Appendix D contains implementation details and Appendix E discusses hyperparameter settings for each task.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: We correctly identified how many runs we have repeated in each caption.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

Justification: Some portion of experiments are conducted with proprietary resources as disclosed in the Acknowledgments section. Nevertheless, the overwhelming majority of experiments firmly remain reproducible using the appendices, additional experiments and code.

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: See Section 7 for potential benefits to scientific research. Errors in generated samples could lead to misleading conclusions in downstream applications, which may fall outside the scope of this methodological study.

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [No]

    Justification: We cite and credit the original sources of the assets used, but do not explicitly list all license terms.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: See Appendix D.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]