J Optim Theory Appl (2016) 169:671–691
DOI 10.1007/s10957-015-0803-z

On the Relation Between Optimal Transport and Schrödinger Bridges: A Stochastic Control Viewpoint

Yongxin Chen[1] · **Tryphon T. Georgiou**[1] · **Michele Pavon**[2]

Received: 14 December 2014 / Accepted: 24 August 2015 / Published online: 1 September 2015

Abstract We take a new look at the relation between the optimal transport problem and the Schrödinger bridge problem from a stochastic control perspective. Our aim is to highlight new connections between the two that are richer and deeper than those previously described in the literature. We begin with an elementary derivation of the Benamou–Brenier fluid dynamic version of the optimal transport problem and provide, in parallel, a new fluid dynamic version of the Schrödinger bridge problem. We observe that the latter establishes an important connection with optimal transport without zero-noise limits and solves a question posed by Eric Carlen in 2006. Indeed, the two variational problems differ by a *Fisher information functional*. We motivate and consider a generalization of optimal mass transport in the form of a (fluid dynamic) problem of *optimal transport with prior*. This can be seen as the zero-noise limit of Schrödinger bridges when the prior is any Markovian evolution. We finally specialize to the Gaussian case and derive an explicit computational theory based on matrix Riccati differential equations. A numerical example involving Brownian particles is also provided.

Keywords Optimal transport · Schrödinger bridge · Stochastic control

✉ Michele Pavon
pavon@math.unipd.it

Yongxin Chen
chen2468@umn.edu

Tryphon T. Georgiou
tryphon@umn.edu

1 University of Minnesota, Minneapolis, MN, USA

2 Università di Padova, Padova, Italy

Mathematics Subject Classification 60J60 · 49L20 · 49J20 · 35Q35 · 28A50

1 Introduction

We discuss two problems of very different beginning. Optimal mass transport (OMT) originates in the work of Monge in 1781 [1] and seeks a transport plan that corresponds in an optimal way two distributions of equal total mass. The cost penalizes the distance that mass is transported to ensure exact correspondence. Likewise, data for Erwin Schrödinger's 1931/32 bridge problem [2,3] are again two distributions of equal total mass, in fact, probability distributions. Here, however, these represent densities of diffusive particles at two points in time and the problem seeks the most likely path that establishes a correspondence between the two. A rich relationship between the two problems emerges in the case where the transport cost is quadratic in the distance, and in fact, the problem of OMT emerges as the limit of Schrödinger bridges as the diffusivity tends to zero. The parallel treatment of both problems highlights the time symmetry of both problems and points of contact between stochastic optimal control and information theoretic concepts.

Historically, the modern formulation of OMT is due to Kantorovich [4] and the subject has been the focus of renewed and increased interest because of its relevance in a wide range of fields including economics, physics, engineering, and probability [5–7]. In fact, Kantorovich's contributions and their impact to resource allocation were recognized with the Nobel Prize in Economics in 1975, while in the past twenty years contributions by Ambrosio, Benamou, Brenier, McCann, Cullen, Gangbo, Kinderlehrer, Lott, Otto, Rachev, Ruschendorf, Tannenbaum, Villani, and many others have launched a new fast developing phase; see, e.g., [6–12]. On the other hand, the Schrödinger bridge problem [2,3] has been the subject of strong but intermittent interest by mostly probabilists and mathematical physicists. Early important contributions were due to Fortet, Beurling, Jamison, and Föllmer [13–16]; see [17] for a survey. Schrödinger's original motivation to find a more classical reformulation of quantum mechanics in terms of diffusion processes was, in a sense, accomplished by Nelson [18,19]. Another interesting attempt in this direction was put forward in the eighties by Zambrini [20,21].[1] Renewed interest in Schrödinger bridges was sparked in the past twenty years after a close relationship to stochastic control was recognized [27–29] and a similarly fast developing phase ensued; see the semi-expository paper [30] and [31–34] for other recent contributions.

Besides the intrinsic importance of OMT to the geometry of spaces and the multitude of applications, a significant impetus for some recent work has been the need for effective computation [10,35] which is often challenging. Likewise, excepting special cases [36,37], the computation of the optimal stochastic control for the Schrödinger bridge problem is challenging, as it amounts to two partial differential equations nonlinearly coupled through their boundary values [17]. Only very recently implementable forms have become available for corresponding linear stochastic systems [38–42] and

[1] Over the years, several alternative versions of stochastic mechanics have been proposed by Fényes, Bohm–Vigier, Levy–Krener, Rosenbrock [22–26] to name but a few.

for versions of the problem involving Markov chains and Kraus maps of statistical quantum mechanics [34]; see also [43] which deals with the Schrödinger bridge problem with finite or infinite horizon for a system of nonlinear stochastic oscillators.

The aim of the present paper is to elucidate some of the connections between OMT and Schrödinger bridges, thereby extending both theories. We follow in the footsteps of Léonard [30,33], who investigated their relation, and of Mikami and Thieullen [31,44,45] who employed stochastic control and Schrödinger bridges to solve the optimal transport problem. In particular, we begin with an elementary derivation of the Benamou–Brenier fluid dynamic version of the Monge–Kantorovich problem and provide a parallel time-symmetric fluid dynamic version of the Schrödinger bridge problem. The latter differs from the approach in [30, Section 4] and underscores that an important connection with optimal transport exists even without zero-noise limits. As a side benefit, we note that the particular dynamic formulation of the Schrödinger bridge problem answers a question posed by Carlen in [46, pp. 130–131]. We then formulate a generalization of OMT by introducing a corresponding notion of *prior* and solve this (fluid dynamic) version of *optimal transport with prior*. The formulation allows us to consider zero-noise limits of Schrödinger bridges when the prior is any Markovian evolution. In particular, employing our results in [39], we specialize to the case when the prior evolution is a Gauss–Markov process obtaining explicit results.

The outline of the paper is as follows: In Sect. 2, we derive the Benamou–Brenier version of the OMT problem. In Sect. 3, we formulate the classical Schrödinger bridge problem as a stochastic control problem. In Sect. 4, we give a control time-symmetric formulation of the Schrödinger bridge problem. This leads, in the following Sect. 5, to a new fluid dynamic formulation of the bridge problem. Section 6 is dedicated to the optimal mass transfer problem with prior. In Sect. 7, we investigate the zero-noise limit when the prior is Gaussian. Finally, in Sect. 8, we provide a numerical two-dimensional example of overdamped Brownian particles, where we display the zero-noise limit corresponding to OMT with prior.

2 Optimal Mass Transport as a Stochastic Control Problem

2.1 Monge–Kantorovich Optimal Mass Transport

Given two distributions μ, ν on $\mathbb{R}^n$ having equal total mass, the original formulation of OMT due to Gaspard Monge sought to identify a transport (measurable) map T from $\mathbb{R}^n \to \mathbb{R}^n$ so that the push-forward $T\sharp\mu$ is equal to ν, in the sense that $\nu(\cdot) = \mu(T^{-1}(\cdot))$, while the cost of transportation $\int c(x, T(x))\mu(dx)$ is minimal. Here, $c(x, y)$ represents the transference cost from point x to point y, and for the purposes of the present, it will be $c(x, y) = \frac{1}{2}\|x - y\|^2$.

The dependence of the cost of transportation on T is highly nonlinear which complicated early analyses of the problem. Thus, it was not until Kantorovich's relaxed formulation in 1942 that the Monge's problem received a definitive solution. In this, instead of the transport map one seeks a joint distribution $\Pi(\mu, \nu)$ on the product space $\mathbb{R}^n \times \mathbb{R}^n$, referred to as a "coupling" between μ and ν, so that the marginals along the two coordinate directions coincide with μ and ν, respectively. Thence, one seeks to determine

$$\inf_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^n\times\mathbb{R}^n}\frac{1}{2}\|x-y\|^2\mathrm{d}\pi(x,y). \tag{1}$$

In case an optimal transport map exists, the optimal coupling has support on the graph of this map [6]. Herein, we consider this relaxed Kantorovich formulation. We wish first to give next an elementary derivation of the fact that Problem 1 can be turned into a stochastic control problem as stated in [31, formula(1.6)] and then, to derive an alternative "fluid dynamic" formulation due to Benamou–Brenier. We strive for clarity rather than generality. In particular, we (tacitly) assume throughout the paper that μ does not give mass to sets of dimension $\leq n-1$. Then, by Brenier's theorem [6], there exists a unique optimal transport plan (Kantorovich) induced by a map (Monge) which is the gradient of a convex function.

2.2 A Stochastic Control Formulation

As customary, let us start by observing that

$$\frac{1}{2}\|x-y\|^2=\inf_{x\in\mathcal{X}_{xy}}\int_0^1\frac{1}{2}\|\dot{x}\|^2\mathrm{d}t, \tag{2}$$

where $\mathcal{X}_{xy}$ is the family of $C^1([0,1],\mathbb{R}^n)$ paths with $x(0)=x$ and $x(1)=y$. Let $x^*(t)=(1-t)x+ty$ be the solution of (2), namely the straight line joining x and y. Since $x^*(t)$ is a Euclidean geodesic, any probabilistic average of the lengths of C^1 trajectories starting at x at time 0 and ending in y at time 1 gives necessarily a higher value. Thus, the probability measure on $C^1([0,1],\mathbb{R}^n)$ concentrated on the path $\{x^*(t); 0\leq t\leq 1\}$ solves the problem

$$\inf_{P_{xy}\in\mathbb{D}^1(\delta_x,\delta_y)}\mathbb{E}_{P_{xy}}\left\{\int_0^1\frac{1}{2}\|\dot{x}\|^2\mathrm{d}t\right\}, \tag{3}$$

where $\mathbb{D}^1(\delta_x,\delta_y)$ are the probability measures on $C^1([0,1],\mathbb{R}^n)$ whose initial and final marginals are Dirac measures concentrated at x and y, respectively. Since (3) provides another representation for $\frac{1}{2}\|x-y\|^2$, (1) is equivalent to

$$\inf_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^n\times\mathbb{R}^n}\inf_{P_{xy}\in\mathbb{D}^1(\delta_x,\delta_y)}\mathbb{E}_{P_{xy}}\left\{\int_0^1\frac{1}{2}\|\dot{x}\|^2\mathrm{d}t\right\}\mathrm{d}\pi(x,y). \tag{4}$$

Now observe that if $P_{xy}\in\mathbb{D}^1(\delta_x,\delta_y)$ and $\pi\in\Pi(\mu,\nu)$ then

$$P=\int_{\mathbb{R}^n\times\mathbb{R}^n}P_{xy}\mathrm{d}\pi(x,y)$$

is a probability measure in $\mathbb{D}^1(\mu,\nu)$, namely a measure on $C^1([0,1],\mathbb{R}^n)$ whose one-time marginal at 0 and 1 are specified to be μ and ν, respectively. Conversely, the disintegration of any measure $P\in\mathbb{D}^1(\mu,\nu)$ with respect to the initial and final positions yields $P_{xy} \in \mathbb{D}^1(\delta_x, \delta_y)$ and $\pi \in \Pi(\mu, \nu)$. Thus, the original optimal transport problem is equivalent to

$$\inf_{P \in \mathbb{D}^1(\mu,\nu)} \mathbb{E}_P \left\{ \int_0^1 \frac{1}{2} \|\dot{x}\|^2 \mathrm{d}t \right\}. \tag{5}$$

So far, we have followed [30, pp. 2–3]. Instead of the "particle" picture, we can also consider the hydrodynamic version of (2), namely the optimal control problem

$$\frac{1}{2}\|x - y\|^2 = \inf_{v \in \mathcal{V}_y} \int_0^1 \frac{1}{2} \|v(x^v(t), t)\|^2 \mathrm{d}t,$$
$$\dot{x}^v(t) = v(x^v(t), t), \quad x(0) = x, \tag{6}$$

where the admissible feedback control laws $v(\cdot, \cdot)$ in $\mathcal{V}_y$ are continuous and such that $x^v(1) = y$.

Following the same steps as before, we get that the optimal transport problem is equivalent to the following stochastic control problem with atypical boundary constraints

$$\inf_{v \in \mathcal{V}} \mathbb{E} \left\{ \int_0^1 \frac{1}{2} \|v(x^v(t), t)\|^2 \mathrm{d}t \right\}, \tag{7a}$$
$$\dot{x}^v(t) = v(x^v(t), t), \quad \text{a.s.}, \quad x(0) \sim \mu, \quad x(1) \sim \nu. \tag{7b}$$

Finally, suppose $\mathrm{d}\mu(x) = \rho_0(x)\mathrm{d}x$, $\mathrm{d}\nu(y) = \rho_1(y)\mathrm{d}y$ and $x^v(t) \sim \rho(x, t)\mathrm{d}x$. Then, necessarily, ρ satisfies (weakly) the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) = 0 \tag{8}$$

expressing the conservation of probability mass. Moreover,

$$\mathbb{E} \left\{ \int_0^1 \frac{1}{2} \|v(x^v(t), t)\|^2 \mathrm{d}t \right\} = \int_{\mathbb{R}^n} \int_0^1 \frac{1}{2} \|v(x, t)\|^2 \rho(x, t) \mathrm{d}t \mathrm{d}x.$$

Hence, (7) turns into the celebrated "fluid dynamic" version of the optimal transport problem due to Benamou and Brenier [10]:

$$\inf_{(\rho, v)} \int_{\mathbb{R}^n} \int_0^1 \frac{1}{2} \|v(x, t)\|^2 \rho(x, t) \mathrm{d}t \mathrm{d}x, \tag{9a}$$
$$\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) = 0, \tag{9b}$$
$$\rho(x, 0) = \rho_0(x), \quad \rho(y, 1) = \rho_1(y). \tag{9c}$$

The variational analysis for (7) or, equivalently, for (9) can be carried out in many different ways. For instance, let $\mathcal{P}_{\rho_0 \rho_1}$ be the family of flows of probability densities $\rho = \{\rho(\cdot, t); 0 \le t \le 1\}$ satisfying (9c) and let $\mathcal{V}$ be the family of continuous feedback control laws $v(\cdot, \cdot)$. Consider the unconstrained minimization of the Lagrangian over $\mathcal{P}_{\rho_0 \rho_1} \times \mathcal{V}$

$$\mathcal{L}(\rho, v) = \int_{\mathbb{R}^n} \int_0^1 \left[\frac{1}{2} \|v(x,t)\|^2 \rho(x,t) + \lambda(x,t) \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) \right) \right] \mathrm{d}t \mathrm{d}x, \quad (10)$$

where $\lambda : \mathbb{R}^n \times [0, 1] \to \mathbb{R}$ is a C^1 Lagrange multiplier. Integrating by parts, assuming that limits for $x \to \infty$ are zero, and observing that the boundary values at $t = 0$, $t = 1$ are constant over $\mathcal{P}_{\rho_0\rho_1}$, we are left to minimize

$$\int_{\mathbb{R}^n} \int_0^1 \left[\frac{1}{2} \|v(x,t)\|^2 + \left(-\frac{\partial \lambda}{\partial t} - \nabla \lambda \cdot v \right) \right] \rho(x,t) \mathrm{d}t \mathrm{d}x \quad (11)$$

over $\mathcal{P}_{\rho_0\rho_1} \times \mathcal{V}$. Following a two-step optimization procedure as in, e.g., [47], we consider first pointwise minimization with respect to v for a fixed flow of probability densities $\rho = \{\rho(\cdot, t); 0 \le t \le 1\}$ in $\mathcal{P}_{\rho_0\rho_1}$. Pointwise minimization of the integrand at each time $t \in [0, 1]$ gives that

$$v_\rho^*(x,t) = \nabla \lambda(x,t), \quad (12)$$

which is continuous. Plugging this form of the optimal control into (11) yields

$$J(\rho) = -\int_{\mathbb{R}^n} \int_0^1 \left[\frac{\partial \lambda}{\partial t} + \frac{1}{2} \|\nabla \lambda\|^2 \right] \rho(x,t) \mathrm{d}t \mathrm{d}x. \quad (13)$$

In view of this, if λ satisfies the Hamilton–Jacobi equation

$$\frac{\partial \lambda}{\partial t} + \frac{1}{2} \|\nabla \lambda\|^2 = 0, \quad (14)$$

then $J(\rho)$ is identically zero over $\mathcal{P}_{\rho_0\rho_1}$ and any $\rho \in \mathcal{P}_{\rho_0\rho_1}$ minimizes the Lagrangian (10) together with the feedback control (12). We have therefore established the following [10]:

Proposition 2.1 *Let $\rho^*(x,t)$ with $t \in [0, 1]$ and $x \in \mathbb{R}^n$ satisfy*

$$\frac{\partial \rho^*}{\partial t} + \nabla \cdot (\nabla \psi \rho^*) = 0, \quad \rho^*(x, 0) = \rho_0(x), \quad (15)$$

where ψ is a solution of the Hamilton–Jacobi equation

$$\frac{\partial \psi}{\partial t} + \frac{1}{2} \|\nabla \psi\|^2 = 0 \quad (16)$$

for some boundary condition $\psi(x, 1) = \psi_1(x)$. If $\rho^*(x, 1) = \rho_1(x)$, then the pair (ρ^*, v^*) with $v^*(x,t) = \nabla \psi(x,t)$ is a solution of (9), provided that the boundary conditions in the integration by part step vanish as $\|x\| \to \infty$.

In general, one cannot expect to have a classical solution of (16) and has to be content with a viscosity solution [48]. See [49] for a recent contribution in the case when only samples of ρ_0 and ρ_1 are known.

3 Schrödinger Bridges as a Solution to a Stochastic Control Problem

After some background on finite energy diffusions and the formulation of the Schrödinger bridge problem, we present a stochastic control reformulation.

3.1 Finite Energy Diffusions

We follow [15–17]. Let $\Omega := C([0,1], \mathbb{R}^n)$ denote the family of n-dimensional continuous functions, W_x denote Wiener measure on Ω starting at x at $t = 0$. If, instead of a Dirac measure concentrated at x, we give the volume measure as initial condition, we get the unbounded[2] measure

$$W := \int W_x \, \mathrm{d}x$$

on path space, which is called *stationary Wiener measure* (or, sometimes, reversible Brownian motion). It is a useful tool to introduce the family of distributions $\mathbb{D}$ on Ω which are equivalent to it. By Girsanov's theorem [50], under $Q \in \mathbb{D}$, the coordinate process $x(t,\omega) = \omega(t)$ admits the representations

$$\mathrm{d}x(t) = \beta_+^Q \mathrm{d}t + \mathrm{d}w_+(t), \quad \beta_+^Q \text{ is } \mathcal{F}_t^+ - \text{adapted}, \tag{17}$$

$$\mathrm{d}x(t) = \beta_-^Q \mathrm{d}t + \mathrm{d}w_-(t), \quad \beta_-^Q \text{ is } \mathcal{F}_t^- - \text{adapted}, \tag{18}$$

where $\mathcal{F}_t^+$ and $\mathcal{F}_t^-$ are σ- algebras of events observable up to time t and from time t on, respectively, and w_-, w_+ are standard n-dimensional Wiener processes [51]. Moreover, the forward and the backward drifts β_+^Q, β_-^Q satisfy

$$Q\left[\int_0^1 \|\beta_+^Q\|^2 \mathrm{d}t < \infty\right] = Q\left[\int_0^1 \|\beta_-^Q\|^2 \mathrm{d}t < \infty\right] = 1.$$

For $Q, P \in \mathbb{D}$, the *relative entropy* (*Divergence*, *Kullback–Leibler index*) $H(Q, P)$ of Q with respect to P is

$$H(Q, P) = \mathbb{E}_Q\left[\ln \frac{\mathrm{d}Q}{\mathrm{d}P}\right].$$

It then follows from Girsanov's theorem [51] that

$$H(Q, P) = H(q_0, p_0) + \mathbb{E}_Q\left[\int_0^1 \frac{1}{2}\|\beta_+^Q - \beta_+^P\|^2 \mathrm{d}t\right] \tag{19a}$$

$$= H(q_1, p_1) + \mathbb{E}_Q\left[\int_0^1 \frac{1}{2}\|\beta_-^Q - \beta_-^P\|^2 \mathrm{d}t\right]. \tag{19b}$$

[2] Therefore, W is not a probability measure. Its marginals at each point in time coincide with the Lebesgue measure.

Here q_0, q_1 (p_0, p_1) are the marginal distributions of Q (P) at 0 and 1, respectively. Moreover, β_+^Q and β_-^Q are the forward and the backward drifts of Q, respectively, and similarly for P.

3.2 The Schrödinger Bridge Problem

Now let ρ_0 and ρ_1 be two probability densities. Let $\mathbb{D}(\rho_0, \rho_1)$ denote the set of distributions in $\mathbb{D}$ having the prescribed marginal densities at 0 and 1. Given $P \in \mathbb{D}$, we consider the following problem:

$$\text{minimize} \quad H(Q, P) \quad \text{s.t.} \quad Q \in \mathbb{D}(\rho_0, \rho_1). \tag{20}$$

This optimization solves a large deviations problem in that it seeks the *most likely* evolution between the two given marginals [51,52]. If there is at least one Q in $\mathbb{D}(\rho_0, \rho_1)$ such that $H(Q, P) < \infty$, then, under rather mild assumptions on P, ρ_0 and ρ_1 [30, Proposition2.5], there exists a unique minimizer Q^* in $\mathbb{D}(\rho_0, \rho_1)$, called *the Schrödinger bridge* from ρ_0 to ρ_1 over P. Indeed, let

$$P_{xy} = P\,[\,\cdot \mid x(0) = x, x(1) = y\,]\,, \quad Q_{xy} = Q\,[\,\cdot \mid x(0) = x, x(1) = y\,]$$

be the disintegrations of P and Q with respect to the initial and final positions. Let also μ^P, μ^Q be the joint initial–final time distributions under P and Q, respectively. Then [16]

$$Q^*(\cdot) = \int_{\mathbb{R}^n \times \mathbb{R}^n} P_{xy}(\cdot)\mu^*(\mathrm{d}x\mathrm{d}y),$$

where μ^* is the unique solution of

$$\text{minimize}_{\mu^Q} \quad \int_{\mathbb{R}^n \times \mathbb{R}^n} \left(\ln \frac{d\mu^Q}{d\mu^P}\right) \mu^Q(\mathrm{d}x\mathrm{d}y) \tag{21}$$

subject to the (linear) constraints

$$\mu^Q(\mathrm{d}x \times \mathbb{R}^n) = \rho_0(x)\mathrm{d}x, \quad \mu^Q(\mathbb{R}^n \times \mathrm{d}y) = \rho_1(y)\mathrm{d}y. \tag{22}$$

3.3 A Stochastic Control Formulation

Consider now the case where (the coordinate process under) P is a *Markovian diffusion* with forward drift field $b_+^P(x, t)$ and transition density $p(s, x, t, y)$. The density $\rho(x, t)$ of P is a weak solution of the Fokker–Planck equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left(b_+^P \rho\right) - \frac{1}{2}\Delta\rho = 0. \tag{23}$$

Moreover, forward and backward drifts are related through Nelson's duality relation [18]

$$b_-^P(x,t) = b_+^P(x,t) - \nabla \ln \rho(x,t). \tag{24}$$

In this case, the Schrödinger bridge Q^* is also Markovian [27,28], with forward drift field

$$b_+^{Q^*}(x,t) = b_+^P(x,t) + \nabla \ln \varphi(x,t), \tag{25}$$

where the (everywhere positive) function φ solves, together with another function $\hat{\varphi}$, the system

$$\varphi(x,t) = \int p(t,x,1,y)\varphi(y,1)\mathrm{d}y, \tag{26}$$

$$\hat{\varphi}(x,t) = \int p(0,y,t,x)\hat{\varphi}(y,0)\mathrm{d}y, \tag{27}$$

with boundary conditions

$$\varphi(x,0)\hat{\varphi}(x,0) = \rho_0(x), \quad \varphi(x,1)\hat{\varphi}(x,1) = \rho_1(x).$$

Moreover, the one-time density $\tilde{\rho}$ of Q^* satisfies the factorization

$$\tilde{\rho}(x,t) = \varphi(x,t)\hat{\varphi}(x,t), \ \forall t \in [0,1]. \tag{28}$$

We sketch the derivation of (25). Let $\varphi(x,t)$ be any positive, space–time harmonic function, namely φ satisfies on $\mathbb{R}^n \times [0,1]$

$$\frac{\partial \varphi}{\partial t} + b_+^P \cdot \nabla \varphi + \frac{1}{2}\Delta \varphi = 0. \tag{29}$$

It follows that $\ln \varphi$ satisfies

$$\frac{\partial \ln \varphi}{\partial t} + b_+^P \cdot \nabla \ln \varphi + \frac{1}{2}\Delta \ln \varphi = -\frac{1}{2}\|\nabla \ln \varphi\|^2. \tag{30}$$

Observe that, in view of (19a), problem (20) is equivalent to minimizing over $\mathbb{D}(\rho_0, \rho_1)$ the functional

$$I(Q) = \mathbb{E}_Q\left[\int_0^1 \frac{1}{2}\|\beta_+^Q - b_+^P(x(t),t)\|^2 \mathrm{d}t - \ln \varphi(x(1),1) + \ln \varphi(x(0),0)\right]. \tag{31}$$

This follows from the fact that $H(Q,P)$ and (31) differ by a quantity which is constant over $\mathbb{D}(\rho_0, \rho_1)$. Note that we do not assume that Q is Markov a priori. Under Q, by Ito's rule,

$$d \ln \varphi(x(t),t) = \left[\frac{\partial \ln \varphi}{\partial t} + \beta_+^Q \cdot \nabla \ln \varphi + \frac{1}{2}\Delta \ln \varphi\right](x(t),t)\mathrm{d}t \\ + \nabla \ln \varphi(x(t),t)\mathrm{d}w_+(t).$$

Using this and (30) in (31), we obtain [53]

$$I(Q) = \mathbb{E}_Q\left[\int_0^1 \frac{1}{2}\|\beta_+^Q - b_+^P(x(t),t) - \nabla \ln \varphi(x(t),t)\|^2 \mathrm{d}t\right],$$

where again we assumed that the stochastic integral has zero expectation. Then, the form (25) of the forward drift of Q^* follows. Define

$$\hat{\varphi}(x,t) = \frac{\tilde{\rho}(x,t)}{\varphi(x,t)}.$$

Then, a direct calculation using (29), and the Fokker–Planck equation

$$\frac{\partial \tilde{\rho}}{\partial t} + \nabla \cdot \Big((b_+^P + \nabla \ln \varphi)\tilde{\rho}\Big) - \frac{1}{2}\Delta\tilde{\rho} = 0 \tag{32}$$

yields

$$\frac{\partial \hat{\varphi}}{\partial t} + \nabla \cdot \Big(b_+^P \hat{\varphi}\Big) - \frac{1}{2}\Delta\hat{\varphi} = 0. \tag{33}$$

Thus, $\hat{\varphi}$ is co-harmonic; particularly, it satisfies the original Fokker–Planck equation (23) just like $\rho(x,t)$, the one-time density of the "prior" P.

Suppose we start instead with $\psi(x,t)$, a positive, reverse-time space–time harmonic function, namely ψ, satisfies on $\mathbb{R}^n \times [0,1]$

$$\frac{\partial \psi}{\partial t} + b_-^P \cdot \nabla\psi - \frac{1}{2}\Delta\psi = 0, \tag{34}$$

where $b_-^P(x) = b_+^P(x) - \nabla \ln \rho(x,t)$ is the backward drift of P. Consider now the functional

$$\bar{I}(Q) = \mathbb{E}_Q\left[\int_0^1 \frac{1}{2}\|\beta_-^Q - b_-^P(x(t),t)\|^2 dt + \ln \psi(x(1),1) - \ln \psi(x(0),0)\right].$$

Again, minimizing $\bar{I}(Q)$ over $\mathbb{D}(\rho_0,\rho_1)$ is equivalent to (20). Using Ito's rule, the same "completing the square argument" as before yields that the optimal backward drift is

$$b_-^{Q^*}(x,t) = b_-^P(x,t) - \nabla \ln \psi(x,t). \tag{35}$$

Thus, the solution Q^* appears, in the language of Doob, an h-path process both in the forward and in the backward direction of time. We now identify ψ. By (28), we have

$$\begin{aligned}\tilde{\rho}(x,t) &= \varphi(x,t)\hat{\varphi}(x,t) = \varphi(x,t)\frac{\hat{\varphi}(x,t)}{\rho(x,t)}\rho(x,t)\\ &= \varphi(x,t)\psi(x,t)\rho(x,t), \quad \text{with } \psi(x,t) = \frac{\hat{\varphi}(x,t)}{\rho(x,t)}, \quad \forall t \in [0,1].\end{aligned} \tag{36}$$

Indeed, ψ, being the ratio of two solutions of the original Fokker–Planck (33), is reverse-time space–time harmonic; i.e., it satisfies (34) [54]. This agrees with the following calculation using (35), (24), (36) and (25)

$$\begin{aligned} b_-^{Q^*}(x,t) &= b_-^P(x,t) - \nabla \ln \psi(x,t) = b_+^P(x,t) - \nabla \ln \rho(x,t) - \nabla \ln \psi(x,t) \\ &= b_+^P(x,t) - \nabla \ln \hat{\varphi}(x,t) = b_+^P(x,t) + \nabla \ln \varphi(x,t) \\ &\quad - \nabla \ln \varphi(x,t) - \nabla \ln \hat{\varphi}(x,t) \\ &= b_+^{Q^*}(x,t) - \nabla \ln \tilde{\rho}(x,t), \end{aligned}$$

which is simply (24) for the drifts of Q^*. Formula (36) should be compared to [30, Theorem 3.4].

Finally, there are also conditional versions of these variational problems which are closer to standard stochastic control problems. Consider minimizing

$$\begin{aligned} J(u) &= \mathbb{E}\left[\int_t^1 \frac{1}{2}\|u(\tau)\|^2 d\tau - \ln \varphi_1(x(1))\right], \\ \mathrm{d}x(s) &= \left[b_+^P(x(s),s) + u(x(s),s)\right] ds + dw_+(s), \quad x(t) = x \text{ a.s.} \end{aligned} \tag{37}$$

over feedback controls u such that the differential equation has a weak solution. If $\varphi(x,t)$ solves (29) with terminal condition $\varphi_1(x)$, then, the same argument as before shows that $u^*(x,t) = \nabla \ln \varphi(x,t)$ is optimal and that $S(x,t) = -\ln \varphi(x,t) = \inf_u J(u)$ is the *value function* [55] of the control problem. By (30), the Hamilton–Jacobi–Bellman equation has the form

$$\frac{\partial S}{\partial t} + \inf_u \left[\left(b_+^P + u\right) \cdot \nabla S + \frac{1}{2}\|u\|^2\right] + \frac{1}{2}\Delta S = 0, \quad S(x,1) = -\ln \varphi_1(x).$$

4 A Time-Symmetric Formulation of the Schrödinger Bridge Problem

Inspired by a paper by Nagasawa [56], we proceed to derive a control time-symmetric formulation of the bridge problem. For any $Q \in \mathbb{D}$, define the *current* and *osmotic* drifts

$$v^Q = \frac{\beta_+^Q + \beta_-^Q}{2}, \quad u^Q = \frac{\beta_+^Q - \beta_-^Q}{2}.$$

Then $\beta_+^Q = v^Q + u^Q, \quad \beta_-^Q = v^Q - u^Q$. Observe that

$$\begin{aligned} H(Q,P) =& \frac{1}{2}H(q_0,p_0) + \frac{1}{2}H(q_1,p_1) + \frac{1}{4}\mathbb{E}_Q\left[\int_0^1 \|\beta_+^Q - \beta_+^P\|^2 + \|\beta_-^Q - \beta_-^P\|^2 \mathrm{d}t\right] \\ =& \frac{1}{2}H(q_0,p_0) + \frac{1}{2}H(q_1,p_1) + \frac{1}{2}\mathbb{E}_Q\left[\int_0^1 \|v^Q - v^P\|^2 + \|u^Q - u^P\|^2 \mathrm{d}t\right]. \end{aligned} \tag{38}$$

Since $H(q_0, p_0)$ and $H(q_1, p_1)$ are constant over $\mathbb{D}(\rho_0, \rho_1)$, it follows that the Schrödinger bridge Q^* minimizes the sum of the two incremental kinetic energies. Finally, we consider minimizing over $\mathbb{D}(\rho_0, \rho_1)$ the functional $I_s(Q) = \frac{1}{2}\left[I(Q) + \bar{I}(Q)\right]$. By the previous calculation, this is equivalent to minimizing over $\mathbb{D}(\rho_0, \rho_1)$ the functional

$$\frac{1}{2}\mathbb{E}_Q\left[\int_0^1 \left(\|v^Q - v^P\|^2 + \|u^Q - u^P\|^2\right) \mathrm{d}t - \ln\frac{\varphi}{\psi}(x(1), 1) + \ln\frac{\varphi}{\psi}(x(0), 0)\right].$$

Proposition 4.1 *The following current and osmotic drifts are optimal for the above problem*

$$v^{Q^*}(x,t) = v^P(x,t) + \frac{1}{2}\nabla\ln\frac{\varphi}{\psi}(x,t), \quad u^{Q^*}(x,t) = u^P(x,t) + \frac{1}{2}\nabla\ln(\varphi\psi)(x,t).$$

Proof It suffices to notice that $v^{Q^*}(x,t)$ and $u^{Q^*}(x,t)$ make the functional equal to zero. □

The result of Proposition 4.1 agrees with (25) and (35). A variational analysis with the two controls v^Q and u^Q can be developed along the lines of Pavon [57, Sections III–IV].

5 Fluid Dynamic Formulation of the Schrödinger Bridge Problem

Let us go back to the symmetric representation (38). In the case where the prior measure is $P = W$ stationary Wiener measure, we have $v^W = u^W = 0$.[3] It basically corresponds to the situation where there is no prior information. Since the boundary relative entropies are constant, the problem is equivalent to minimizing

$$\mathbb{E}\left\{\int_0^1 \left[\frac{1}{2}\|v\|^2 + \frac{1}{2}\|u\|^2\right] \mathrm{d}t\right\}$$

over $\mathbb{D}(\rho_0, \rho_1)$. Here we discard the superscript Q for simplification. Let us restrict our search to Markovian processes. By (24), the osmotic drift field u satisfies

$$u(x,t) = \frac{b_+(x,t) - b_-(x,t)}{2} = \frac{1}{2}\nabla\ln\rho(x,t). \tag{39}$$

Moreover, the current drift field v satisfies

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(v\rho) = 0. \tag{40}$$

[3] See [30, pp. 7–8] for a justification of employing unbounded path measures in relative entropy problems.

Thus, the Schrödinger bridge problem becomes

$$\inf_{(\rho,v)} \int_{\mathbb{R}^n} \int_0^1 \left[\frac{1}{2}\|v(x,t)\|^2 + \frac{1}{8}\|\nabla \ln \rho(x,t)\|^2\right] \rho(x,t)\mathrm{d}t\mathrm{d}x, \tag{41a}$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) = 0, \tag{41b}$$

$$\rho(x,0) = \rho_0(x), \quad \rho(y,1) = \rho_1(y), \tag{41c}$$

which should be compared to (9a)–(9b)–(9c). We notice, in particular, that the two functionals differ by a term which is a multiple of the integral over time of the Fisher information functional

$$\int_{\mathbb{R}^n} \|\nabla \ln \rho(x,t)\|^2 \rho(x,t)\mathrm{d}x.$$

This unveils a relation between the two problems without zero-noise limits [33,44]. In [46, p. 131], Carlen, while investigating the connections between OMT and Nelson's stochastic mechanics [18,19] in the spirit of Madelung [58], posed the question of minimizing the *Yasue action* (41a) subject to the continuity equation (41b) for given initial and final marginals (41c). He states that "...the Euler–Lagrange equations for it are not easy to understand." The following result answers Carlen's question.

Proposition 5.1 *Let Q^* be the solution of the Schrödinger bridge problem with $P = W$ stationary Wiener measure as prior and marginals* (41c)*. Then, v^*, its current velocity field, and $\{\rho^*(\cdot,t); 0 \le t \le 1\}$, its flow of one-time marginals, solve Carlen's problem.*

Proof By the above analysis, minimizing relative entropy on path space is equivalent to minimizing the Yasue action (41a) subject to the continuity equation (41b) with the given initial and final marginals. □

Finally, we mention that a fluid dynamic problem concerning swarms of particles diffusing anisotropically with losses has been studied in [59].

6 Optimal Mass Transport with a "Prior"

Considering the relation we have seen between the fluid dynamic versions of the optimal transport problem and the Schrödinger bridge problem, one may wonder whether there exists a formulation of the former which allows for an "a priori" evolution like in the latter. Relative entropy on path space does not work for zero-noise random evolutions as they are singular. Indeed, let P_ϵ and Q_ϵ be the measures on $C([0,1],\mathbb{R}^n)$ equivalent to scaled stationary Wiener measure W_ϵ[4] with forward differentials

$$dx(t) = \beta_+^{P_\epsilon}\mathrm{d}t + \sqrt{\epsilon}\mathrm{d}w_+(t),$$
$$dx(t) = \beta_+^{Q_\epsilon}\mathrm{d}t + \sqrt{\epsilon}\mathrm{d}w_+(t). \tag{42}$$

[4] Measure induced by $\sqrt{\epsilon}w(t)$ on path space Ω with volume measure as initial condition.

Then, one can argue along the same lines as in Sect. 3 that

$$H(Q_\epsilon, P_\epsilon) = H(q_0, p_0) + \mathbb{E}_{Q_\epsilon}\left[\int_0^1 \frac{1}{2\epsilon}\|\beta_+^{Q_\epsilon} - \beta_+^{P_\epsilon}\|^2 \mathrm{d}t\right].$$

As $\epsilon \searrow 0$, the relative entropy becomes infinite unless $Q_\epsilon = P_\epsilon$.[5] To circumvent this difficulty, we start with the following fluid dynamic control problem. Suppose we have two probability densities ρ_0 and ρ_1 and a flow of probability densities $\{\rho(x,t); 0 \le t \le 1\}$ satisfying

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (v\rho) = 0, \tag{43}$$

for some continuous vector field $v(\cdot,\cdot)$. We take (43) as our "prior" evolution and formulate the following problem,

$$\inf_{(\tilde\rho,\tilde v)} \int_{\mathbb{R}^n}\int_0^1 \frac{1}{2}\|\tilde v(x,t) - v(x,t)\|^2 \tilde\rho(x,t)\mathrm{d}t\mathrm{d}x, \tag{44a}$$

$$\frac{\partial \tilde\rho}{\partial t} + \nabla \cdot (\tilde v \tilde\rho) = 0, \tag{44b}$$

$$\tilde\rho(x,0) = \rho_0(x), \quad \tilde\rho(y,1) = \rho_1(y). \tag{44c}$$

Clearly, if the prior flow satisfies $\rho(x,0) = \rho_0(x)$ and $\rho(x,1) = \rho_1(x)$, then it solves the problem and $\tilde v^* = v$. Moreover, the standard optimal transport problem is recovered when the prior evolution is constant, i.e., $v \equiv 0$.

Let us try to provide further motivation to study problem (44). Consider the situation where a previous optimal transport problem (9) has been solved with boundary marginals $\bar\rho_0$ and $\bar\rho_1$ leading to the optimal velocity field $v(x,t)$. Here say $\bar\rho_0$ represent resources being produced to satisfy the demand $\bar\rho_1$. Suppose now new information becomes available showing that the actual resources available are distributed according to ρ_0 and the actual demand is distributed according to ρ_1. As we had already set up a transportation plan according to velocity field v, we seek to solve a new transport problem where the new evolution is close to the one we would have employing the previous velocity field. This is represented in problem (44).

Remark 6.1 The particle version of (44) takes the form of a more familiar OMT problem, namely in the notation of Sect. 2,

$$\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y)\mathrm{d}\pi(x,y), \tag{45}$$

where

$$c(x,y) = \inf_{x \in \mathcal{X}_{xy}} \int_0^1 L(t, x(t), \dot x(t))\mathrm{d}t, \quad L(t,x,\dot x) = \|\dot x - v(x,t)\|^2. \tag{46}$$

[5] This calculation indicates that there may be a limit as $\epsilon \searrow 0$ of $\inf\{\epsilon H(Q_\epsilon, P_\epsilon)\}$ and, hopefully, in suitable sense, of the minimizers. This is indeed the case; see [30,33,44] for a precise statement of limiting results.

The explicit calculation of the function $c(x, y)$ when $v \not\equiv 0$ is nontrivial. Moreover, the zero-noise limit results of [33, Section 3], based on a Large Deviations Principle [52], although very general in other ways, seem to cover here only the case where $c(x, y) = c(x - y)$ strictly convex originating from a Lagrangian $L(t, x, \dot{x}) = c(\dot{x})$. We mention that in our follow-up paper [41], we deal with optimal transport problems where the Lagrangian is not strictly convex with respect to $\dot{x}$. Finally, we feel that our formulation is a most natural one in which to study zero-noise limits of Schrödinger bridges with a general Markovian prior evolution. In the next section, we discuss this problem in the Gaussian case. The proof of the convergence of the path-space measures of the minimizers can be done along the lines of Léonard [33] where Γ-convergence of the bridge minimum problems to the OMT problem is established. This, under suitable assumptions, guarantees convergence of the minimizers.

The variational analysis for (44) can be carried out as in Sect. 2 obtaining the following result:

Proposition 6.1 *If $\tilde{\rho}^*$ satisfying*

$$\frac{\partial \tilde{\rho}^*}{\partial t} + \nabla \cdot [(v + \nabla \psi)\tilde{\rho}^*] = 0, \quad \tilde{\rho}^*(x, 0) = \rho_0(x), \tag{47}$$

where ψ is a solution of the Hamilton–Jacobi equation

$$\frac{\partial \psi}{\partial t} + v \cdot \nabla \psi + \frac{1}{2}\|\nabla \psi\|^2 = 0, \tag{48}$$

is such that $\tilde{\rho}^*(x, 1) = \rho_1(x)$, then the pair $(\tilde{\rho}^*(x, t), \tilde{v}^*(x, t) = v(x, t) + \nabla \psi(x, t))$ is a solution of the problem (44), provided $\psi(x, t)\tilde{v}(x, t)\tilde{\rho}(x, t)$ vanishes as $\|x\| \to \infty$ for each fixed t.

If $v(x, t) = \alpha(t)x$, and both ρ_0 and ρ_1 are Gaussian, then the optimal evolution is given by a linear equation and is therefore given by a Gaussian process as we will study next.

7 The Gaussian Case

In this section, we consider the correspondence between Schrödinger bridges and OMT for the special case where the underlying dynamics are linear and the marginals are normal distributions. To this end, consider the reference evolution

$$\mathrm{d}x(t) = A(t)x(t)\mathrm{d}t + \sqrt{\epsilon}\mathrm{d}w(t), \tag{49}$$

and the two marginals $\rho_0 \sim N(m_0, \Sigma_0)$ and $\rho_1 \sim N(m_1, \Sigma_1)$. In our previous work [39], we derived a “closed form” expression for the corresponding Schrödinger bridge as

$$\mathrm{d}x(t) = (A(t) - \Pi_\epsilon(t))x(t)\mathrm{d}t + m(t)\mathrm{d}t + \sqrt{\epsilon}\mathrm{d}w(t), \tag{50}$$

where $\Pi_\epsilon(t)$ satisfies the matrix Riccati equation

$$\dot{\Pi}_\epsilon(t) + A(t)'\Pi_\epsilon(t) + \Pi_\epsilon(t)A(t) - \Pi_\epsilon(t)^2 = 0 \tag{51}$$

with the boundary condition $\Pi_\epsilon(0)$ given by

$$\frac{\epsilon}{2}\Sigma_0^{-1} + \Phi_{10}'M_{10}^{-1}\Phi_{10} - \Sigma_0^{-1/2}\left(\frac{\epsilon^2}{4}I + \Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Sigma_1 M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2}\right)^{1/2}\Sigma_0^{-1/2},$$

and $m(t)$ satisfies

$$m(t) = \hat{\Phi}(1,t)'\hat{M}(1,0)^{-1}(m_1 - \hat{\Phi}(1,0)m_0). \tag{52}$$

Here $\Phi_{10} := \Phi(1,0)$ is the state transition matrix from 0 to 1 for $A(t)$ and

$$M_{10} := M(1,0) = \int_0^1 \Phi(1,t)\Phi(1,t)'\mathrm{d}t$$

is the controllability Gramian, and $\hat{\Phi}(t,s)$, $\hat{M}(t,s)$ are the state transition matrix and controllability Gramian for $A(t) - \Pi_\epsilon(t)$.

We now consider the zero-noise limit by letting ϵ go to 0. In the case where $A(t) \equiv 0$, the Schrödinger bridge solution process converges to the solution of an OMT problem [30,44]. In general, when $A(t) \not\equiv 0$, by taking $\epsilon = 0$ we obtain

$$\Pi_0(0) = \Phi_{10}'M_{10}^{-1}\Phi_{10} - \Sigma_0^{-1/2}\left(\Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Sigma_1 M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2}\right)^{1/2}\Sigma_0^{-1/2} \tag{53}$$

and a limiting process

$$\mathrm{d}x(t) = (A(t) - \Pi_0(t))x(t)\mathrm{d}t + m(t)\mathrm{d}t, \quad x(0) \sim (m_0, \Sigma_0) \tag{54}$$

with $\Pi_0(t), m(t)$ satisfying (51), (52), and (53). In fact, $\Pi_0(t)$ has the expression

$$\begin{aligned}\Pi_0(t) = &-M(t,0)^{-1}\Phi(t,0)\Big[\Phi_{10}'M_{10}^{-1}\Phi_{10} - \Phi(t,0)'M(t,0)^{-1}\Phi(t,0)\\ &-\Sigma_0^{-1/2}(\Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Sigma_1 M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2})^{1/2}\Sigma_0^{-1/2}\Big]^{-1}\Phi(t,0)'M(t,0)^{-1}\\ &-M(t,0)^{-1}.\end{aligned} \tag{55}$$

It turns out that process (54) yields a solution to problem (44) as stated next.

Proposition 7.1 *Let $\tilde{\rho}(\cdot,t)$ be the probability density of $x(t)$ in (54), and $\tilde{v}(x,t) = (A(t) - \Pi_0(t))x + m(t)$, then the pair $(\tilde{\rho}, \tilde{v})$ is a solution of the problem (44) with prior velocity field $v(x,t) = A(t)x$.*

Proof To show that the pair $(\tilde{\rho}, \tilde{v})$ is a solution, we need to prove i) $\tilde{\rho}$ satisfies the boundary condition $\tilde{\rho}(x,1) = \rho_1(x)$ and ii) $\tilde{v}(x,t) - v(x,t) = \nabla\psi(x,t)$ for some ψ with ψ satisfying the Hamilton–Jacobi equation (48). Here $v(x,t) = A(t)x$ is the drift of the prior process.

We first show that $\tilde{\rho}$ satisfies the boundary condition $\tilde{\rho}(x,1) = \rho_1(x)$. Since the process (54) is a linear diffusion with Gaussian initial condition, $x(t)$ is a Gaussian random vector for all $t \in [0,1]$. Let $\tilde{\rho}(\cdot,t) \sim N(n(t), \Sigma(t))$, then obviously the mean value $n(t)$ is $n(t) = \hat{\Phi}(t,0)m_0 + \int_0^t \hat{\Phi}(t,\tau)m(\tau)d\tau$. We claim that the covariance $\Sigma(t)$ has the explicit expression

$$\Sigma(t) = M(t,0)\Phi(0,t)'\Sigma_0^{-1/2}\Big[(\Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Sigma_1 M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2})^{1/2} - \Sigma_0^{1/2}\Phi_{10}'M_{10}^{-1}\Phi_{10}\Sigma_0^{1/2} + \Sigma_0^{1/2}\Phi(t,0)'M(t,0)^{-1}\Phi(t,0)\Sigma_0^{1/2}\Big]^2 \Sigma_0^{-1/2}\Phi(0,t)M(t,0) \tag{56}$$

for $t \in]0,1]$. This expression is consistent with the initial condition Σ_0 since $\lim_{t\searrow 0}\Sigma(t) = \Sigma_0$. To see that $\Sigma(t)$ is indeed the covariance matrix of $x(t)$, we only need to show that $\Sigma(t)$ satisfies the differential equation

$$\dot{\Sigma}(t) = (A(t) - \Pi_0(t))\Sigma(t) + \Sigma(t)(A(t) - \Pi_0(t))'.$$

This can be verified directly from the expression (56) and (55) after some straightforward but lengthy computations. Now observing that

$$n(1) = \hat{\Phi}(1,0)m_0 + \int_0^1 \hat{\Phi}(1,\tau)m(\tau)\mathrm{d}\tau = m_1$$

by (52) and $\Sigma(1) = \Sigma_1$ from (56), we get that $\tilde{\rho}$ satisfies $\tilde{\rho}(x,1) = \rho_1(x)$. We next show ii). Let $\psi(x,t) = -\frac{1}{2}x'\Pi_0(t)x + m(t)'x + c(t)$ with $c(t) = -\frac{1}{2}\int_0^t m(\tau)'m(\tau)d\tau$. Then, in view of (51) and (52), we establish that

$$\frac{\partial\psi}{\partial t} + v\cdot\nabla\psi + \frac{1}{2}\|\nabla\psi\|^2 = 0.$$

Finally, note that $\psi(x,t)\tilde{v}(x,t)\tilde{\rho}(x,t) \to 0$ as $\|x\| \to \infty$ for any fixed t. □

8 Overdamped Brownian Motion in a Force Field: An Example

Herein, we consider highly overdamped Brownian motion in a force field. In a very strong sense [18, Theorem 10.1], the Smoluchowski model in configuration variables is a good approximation of the full Ornstein–Uhlenbeck model in phase space. We are interested in planar Brownian motion in the quadratic potential $V(x) = \frac{1}{2}(x_1^2 + x_2^2)$. Taking the mass of the particle to be one, the evolution of the Brownian particle is given by the Smoluchowski equation

$$\mathrm{d}x(t) = -\nabla V(x(t))\mathrm{d}t + \sqrt{\epsilon}\,\mathrm{d}w(t), \quad -\nabla V(x) = Ax, \quad A = -3I, \tag{57}$$

Fig. 1 ($\epsilon = 9$) Schrödinger bridge

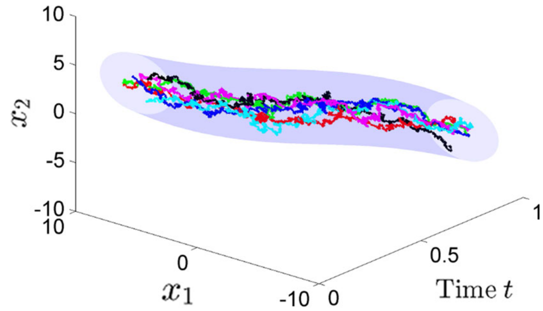


Fig. 2 ($\epsilon = 4$) Schrödinger bridge

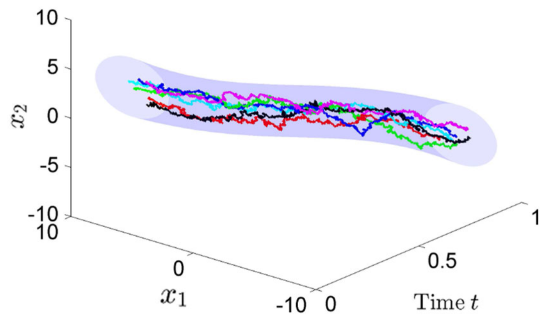
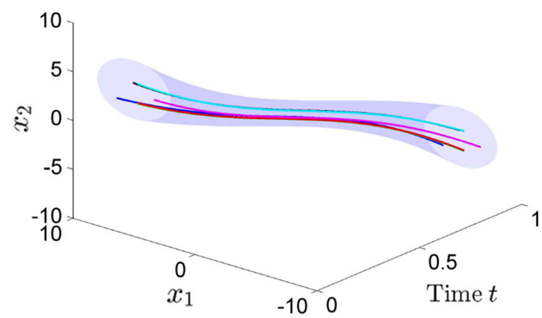


Fig. 3 OMT with prior



where w is a standard, two-dimensional Wiener process. The observed distributions of the particle at the two end-points in time are normal with mean and variance $m_0 = \left[-5,\ -5\right]'$, $\Sigma_0 = I$, and $m_1 = \left[5,\ 5\right]'$, $\Sigma_0 = I$, for $t = 0$ and $t = 1$, respectively. We then seek to interpolate the density of the particle at intermediate points by solving the corresponding Schrödinger bridge problem where (57) plays the role of an a priori evolution.

Figures 1 and 2 depict the flow between the two one-time marginals for the Schrödinger bridge when $\epsilon = 9$ and $\epsilon = 4$, respectively. The transparent tubes represent the "3σ region" where $(x' - m_t')\Sigma_t^{-1}(x - m_t)$ is ≤ 9. Typical sample paths are shown in the figures. As $\epsilon \searrow 0$, the paths of the bridge process resemble those of the corresponding OMT with prior velocity field $v(x, t) = Ax$, which is depicted in Fig. 3. Note that trajectories of OMT without a prior are simply straight lines.

9 Conclusions

Starting from a stochastic control viewpoint, we considerably extended the connections between OMT and Schrödinger bridges. We provided in Sects. 4 and 5 time-symmetric and fluid dynamic formulations of the Schrödinger bridge problem. The latter can then be seen as an OMT dynamic problem with a cost that involves as extra term a Fisher information functional. This formulation directly answers a question posed by Carlen in 2006; see Proposition 5.1. To establish connections between the two problems in a more general setting, we introduced *OMT with prior* which, besides its intrinsic significance, allows direct comparison between OMT and Schrödinger bridges for certain cases with, e.g., possibly non-strictly convex Lagrangian. The stochastic control viewpoint for either OMT or Schrödinger bridge problem, besides unifying the two, allows generalizations of Schrödinger bridges to degenerate and anisotropic diffusions where no maximum entropy problem may be formulated. In Sect. 7, we established a zero-noise limit result and derived explicit formulae for the case (of engineering significance) of Gaussian priors (Proposition 7.1). In Sect. 8, we provided an illustrative example.

Acknowledgments Research partially supported by the NSF under Grant ECCS-1509387, the AFOSR under Grants FA9550-12-1-0319, and FA9550-15-1-0045 and by the University of Padova Research Project CPDA 140897. Part of the research of M.P. was conducted during a stay at the Courant Institute of Mathematical Sciences of the New York University whose hospitality is gratefully acknowledged.

References

1. Monge, G.: Mémoire sur la théorie des déblais et des remblais, De l'Imprimerie Royale (1781)
2. Schrödinger, E.: Über die Umkehrung der Naturgesetze, Sitzungsberichte der Preuss. Phys. Math. Klasse. **10**, 144–153 (1931)
3. Schrödinger, E.: Sur la théorie relativiste de l'electron et l'interpretation de la mécanique quantique. Ann. Inst. H. Poincaré. **2**, 269 (1932)
4. Kantorovich, L.: On the transfer of masses. Dokl. Akad. Nauk. **37**, 227–229 (1942)
5. Rachev, S., Rüschendorf, L.: Mass Transportation Problems: Theory, vol. 1. Springer, New York (1998)
6. Villani, C.: Topics in Optimal Transportation, vol. 58. AMS, Providence (2003)
7. Villani, C.: Optimal Transport: Old and New, vol. 338. Springer, New York (2008)
8. Gangbo, W., McCann, R.J.: The geometry of optimal transportation. Acta Math. **177**(2), 113–161 (1996)
9. Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker–Planck equation. SIAM J. Math. Anal. **29**, 1–17 (1998)
10. Benamou, J., Brenier, Y.: A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. Numerische Mathematik **84**(3), 375–393 (2000)
11. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures. Birkhäuser Verlag, Basel (2008)
12. Ning, L., Georgiou, T.T., Tannenbaum, A.: Matrix-valued Monge–Kantorovich optimal mass transport. IEEE Trans. Autom. Control **60**(2), 373–382 (2015)
13. Fortet, R.: Résolution d'un système d'equations de M. Schrödinger. J. Math. Pure Appl. **9**, 83–105 (1940)
14. Beurling, A.: An automorphism of product measures. Ann. Math. **72**, 189–200 (1960)
15. Jamison, B.: The Markov processes of Schrödinger. Z. Wahrscheinlichkeitstheorie verw. Gebiete **32**, 323–331 (1975)
16. Föllmer, H.: Random fields and diffusion processes. In: Hennequin, P.L. (ed.) Ècole d'Ètè de Probabilitès de Saint-Flour XV–XVII, vol. 1362, pp. 102–203. Springer, New York (1988)

17. Wakolbinger, A.: Schrödinger bridges from 1931 to 1991. Contribuciones en probabilidad y estadistica matematica **3**, 61–79 (1992)
18. Nelson, E.: Dynamical Theories of Brownian Motion. Princeton University Press, Princeton (1967)
19. Nelson, E.: Quantum Fluctuations. Princeton University Press, Princeton (1985)
20. Zambrini, J.C.: Stochastic mechanics according to E. Schrödinger. Phys. Rev. A **33**(3), 1532–1548 (1986)
21. Zambrini, J.C.: Variational processes and stochastic versions of mechanics. J. Math. Phys. **27**(9), 2307–2330 (1986)
22. Fenyes, I.: Eine wahrscheinlichkeitstheoretische Begründung und Interpretation der Quantenmechanik. Z. Physik **132**, 81–106 (1952)
23. Bohm, D.: A suggested interpretation of the quantum theory in terms of "hidden" variables. Phys. Rev. **85**, 166–179 (1952)
24. Bohm, D., Vigier, J.P.: Model of the causal interpretation of quantum theory in terms of a fluid with irregular fluctuations. Phys. Rev. **96**, 208–216 (1954)
25. Levy, B.C., Krener, A.J.: Stochastic mechanics of reciprocal diffusions. J. Math. Phys. **37**, 769 (1996)
26. Rosenbrock, H.H.: Doing quantum mechanics with control theory. IEEE Trans. Autom. Control **54**, 73–77 (2000)
27. Dai Pra, P.: A stochastic control approach to reciprocal diffusion processes. Appl. Math. Optim. **23**(1), 313–329 (1991)
28. Dai Pra, P., Pavon, M.: On the Markov processes of Schrödinger, the Feynman–Kac formula and stochastic control. In: Kaashoek, M.A., van Schuppen, J.H., Ran, A.C.M. (eds.) Realization and Modeling in System Theory, pp. 497–504. Birkhäuser, Boston (1990)
29. Pavon, M., Wakolbinger, A.: On free energy, stochastic control, and Schrödinger processes. In: Di Masi, G.B., Gombani, A., Kurzhanski, A.A. (eds.) Modeling, Estimation and Control of Systems with Uncertainty, pp. 334–348. Birkhäuser, Boston (1991)
30. Léonard, C.: A survey of the Schrödinger problem and some of its connections with optimal transport. Discrete Contin. Dyn. Syst. A **34**(4), 1533–1574 (2014)
31. Mikami, T., Thieullen, M.: Optimal transportation problem by stochastic optimal control. SIAM J. Control Optim. **47**(3), 1127–1139 (2008)
32. Pavon, M., Ticozzi, F.: Discrete-time classical and quantum Markovian evolutions: maximum entropy problems on path space. J. Math. Phys. **51**, 042104–042125 (2010)
33. Léonard, C.: From the Schrödinger problem to the Monge–Kantorovich problem. J. Funct. Anal. **262**, 1879–1920 (2012)
34. Georgiou, T.T., Pavon, M.: Positive contraction mappings for classical and quantum Schrödinger systems. J. Math. Phys. **56**(033301), 1–24 (2015)
35. Angenent, S., Haker, S., Tannenbaum, A.: Minimizing flows for the Monge–Kantorovich problem. SIAM J. Math. Anal. **35**(1), 61–97 (2003)
36. Filliger, R., Hongler, M.O.: Relative entropy and efficiency measure for diffusion-mediated transport processes. J. Phys. A: Math. Gen. **38**, 1247–1250 (2005)
37. Filliger, R., Hongler, M.O., Streit, L.: Connection between an exactly solvable stochastic optimal control problem and a nonlinear reaction-diffusion equation. J. Optim. Theory Appl. **137**, 497–505 (2008)
38. Chen, Y., Georgiou, T.T.: Stochastic bridges of linear systems. Preprint, http://arxiv.org/abs/1407.3421. IEEE Trans. Autom. Control (to appear)
39. Chen, Y., Georgiou, T.T., Pavon, M.: Optimal steering of a linear stochastic system to a final probability distribution, Part I. preprint, http://arxiv.org/abs/1408.2222. IEEE Trans. Autom. Control (to appear)
40. Chen, Y., Georgiou, T.T., Pavon, M.: Optimal steering of a linear stochastic system to a final probability distribution, Part II. Preprint, http://arxiv.org/abs/1410.3447. IEEE Trans. Autom. Control (to appear)
41. Chen, Y., Georgiou, T.T., Pavon, M.: Optimal transport over a linear dynamical system. Preprint, http://arxiv.org/abs/1502.01265 (submitted for publication)
42. Chen, Y., Georgiou, T.T., Pavon, M.: Entropic and displacement interpolation: a computational approach using the Hilbert metric. Preprint, http://arxiv.org/abs/1506.04255 (submitted for publication)
43. Chen, Y., Georgiou, T.T., Pavon, M.: Fast cooling for a system of stochastic oscillators. Preprint, http://arxiv.org/abs/1411.1323v2 (submitted for publication)
44. Mikami, T.: Monge's problem with a quadratic cost by the zero-noise limit of h-path processes. Probab. Theory Relat. Fields **129**, 245–260 (2004)

45. Mikami, T., Thieullen, M.: Duality theorem for the stochastic optimal control problem. Stoch. Proc. Appl. **116**, 1815–1835 (2006)
46. Carlen, E.: Stochastic mechanics: a look back and a look ahead. In: Faris, W.G. (ed.) Diffusion, Quantum Theory and Radically Elementary Mathematics, vol. 47, pp. 117–139. Princeton University Press, Princeton (2006)
47. Kosmol, P., Pavon, M.: Lagrange Lemma and the optimal control of diffusions II: nonlinear Lagrange functionals. Syst. Control Lett. **24**(3), 215–221 (1995)
48. Fleming, W.H., Soner, M.: Controlled Markov Processes and Viscosity Solutions, 2nd edn. Springer, New York (2006)
49. Tabak, E.G., Trigila, G.: Data-driven optimal transport. Commun. Pure. Appl. Math. doi:10.1002/cpa.21588 (2014)
50. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer, New York (1988)
51. Föllmer, H.: Time Reversal on Wiener Space. Springer, New York (1986)
52. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications, 2nd edn. Springer, New York (1998)
53. Wakolbinger, A.: A simplified variational characterization of Schrödinger processes. J. Math. Phys. **30**(12), 2943–2946 (1989)
54. Pavon, M.: Stochastic control and nonequilibrium thermodynamical systems. Appl. Math. Optim. **19**, 187–202 (1989)
55. Fleming, W.H., Rishel, R.W.: Deterministic and Stochastic Optimal Control. Springer, Berlin (1975)
56. Nagasawa, M.: Stochastic variational principle of Schrödinger Processes. In: Cinlar, E., Chung, K.L., Getoo, R.K., Fitzsimmons, P.J., Williams, R.J. (eds.) Seminar on Stochastic Processes, pp. 165–175. Birkhäuser, Boston (1989)
57. Pavon, M.: Hamilton's principle in stochastic mechanics. J. Math. Phys. **36**, 6774–6800 (1995)
58. Madelung, E.: Quantentheorie in hydrodynamischer Form. Z. Physik **40**, 322–326 (1926)
59. Chen, Y., Georgiou, T.T., Pavon, M.: Optimal steering of inertial particles diffusing anisotropically with losses. In: Proceedings of the American Control Conference (2015)