AIP | Journal of Mathematical Physics

Positive contraction mappings for classical and quantum Schrödinger systems

Tryphon T. Georgiou and Michele Pavon

Citation: Journal of Mathematical Physics **56**, 033301 (2015); doi: 10.1063/1.4915289
View online: http://dx.doi.org/10.1063/1.4915289
View Table of Contents: http://scitation.aip.org/content/aip/journal/jmp/56/3?ver=pdfcov
Published by the AIP Publishing

JOURNAL OF MATHEMATICAL PHYSICS **56**, 033301 (2015)

Positive contraction mappings for classical and quantum Schrödinger systems

Tryphon T. Georgiou[1,a)] and Michele Pavon[2,b)]
[1]*Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street S.E. Minneapolis, Minnesota 55455, USA*
[2]*Dipartimento di Matematica, Università di Padova, via Trieste 63, 35121 Padova, Italy*

(Received 26 May 2014; accepted 6 March 2015; published online 23 March 2015)

The classical Schrödinger bridge seeks the most likely probability law for a diffusion process, in path space, that matches marginals at two end points in time; the likelihood is quantified by the relative entropy between the sought law and a prior. Jamison proved that the new law is obtained through a multiplicative functional transformation of the prior. This transformation is characterised by an automorphism on the space of endpoints probability measures, which has been studied by Fortet, Beurling, and others. A similar question can be raised for processes evolving in a discrete time and space as well as for processes defined over non-commutative probability spaces. The present paper builds on earlier work by Pavon and Ticozzi and begins by establishing solutions to Schrödinger systems for Markov chains. Our approach is based on the Hilbert metric and shows that the solution to the Schrödinger bridge is provided by the fixed point of a contractive map. We approach, in a similar manner, the steering of a quantum system across a quantum channel. We are able to establish existence of quantum transitions that are multiplicative functional transformations of a given Kraus map for the cases where the marginals are either uniform or pure states. As in the Markov chain case, and for uniform density matrices, the solution of the quantum bridge can be constructed from the fixed point of a certain contractive map. For arbitrary marginal densities, extensive numerical simulations indicate that iteration of a similar map leads to fixed points from which we can construct a quantum bridge. For this general case, however, a proof of convergence remains elusive. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4915289]

I. INTRODUCTION

In 1931, Erwin Schrödinger published a manuscript on "the reversal of the laws of nature" ("Über die Umkehrung der Naturgesetze"). In it, he raised the following "new and unorthodox"[2] question regarding Brownian motion. Suppose that the density of Brownian particles is observed at two points in time, t_0 and t_1, and that the two end-point densities differ from the initial and final marginals of the prior path-space distribution. Schrödinger then asked for the most likely random evolution that the particles have taken so as to reconcile the observed "improbable but still possible outcome." In modern probabilistic language, as observed by Föllmer some fifty years later,[12] Schrödinger was posing (and, to some extent, solving) a problem of large deviations of the empirical distribution. He was considering an abstract setting, although the very foundations of probability theory were still missing! The solution of the large deviations problem requires, in view of Sanov's theorem,[28] solving a maximum entropy problem. Schrödinger's 1931/32 papers were followed soon afterwards by works by Kolmogoroff on "the reversibility of the statistical laws of

a)Electronic address: tryphon@umn.edu. URL: http://www.ece.umn.edu/~georgiou
b)Electronic address: pavon@math.unipd.it. URL: http://www.math.unipd.it/~pavon

nature" ("Zur Umkehrbarkeit der statistischen Naturgesetze") and by Fortet, Beurling, and many others on the mathematical issues that Schrödinger's paper raised.

In the present paper, following Pavon and Ticozzi[25,33] to which we refer for potential applications, we consider discrete-time and discrete-space classical evolutions as well as quantum Markovian evolutions. More precisely, we first consider discrete random vectors with given prior distribution and endpoint marginals. We derive a constructive proof of the existence of multiplicative functional transformations of the prior initial-final time joint distribution that allows connecting the given end-point marginals. The unique solution is in fact the closest law to the prior in a relative entropy sense amongst all probability laws that are in agreement with the two marginals. A key concept is that of the Hilbert metric—this is a metric which is suitable for quantifying distances in homogeneous positive spaces. A similar approach allows a constructive proof for matching uniform marginal density matrices via a multiplicative functional transformation of any given prior Kraus map. In essence, this result extends to the non-commutative case a result of Sinkhorn[31,32] that any strictly positive stochastic matrix can be transformed into a doubly stochastic matrix via a multiplicative functional transformation. Thus, for the quantum case, we establish that any strictly positive Kraus map can be transformed into a doubly stochastic quantum map via a multiplicative functional transformation. Further, extensive simulations have convinced the authors that the approach works in complete generality, i.e., when the specified marginal densities are not necessarily uniform. However, a rigorous proof as well as a variational principle, in analogy with the classical case, is not available at present.

The paper is structured as follows. Section II provides an exposition of the Hilbert metric. The corresponding geometry is key in studying the Schrödinger bridge in the classical and in the quantum case, in Secs. III and IV, respectively. More specifically, Secs. III A-III C explain Schrödinger's bridges for Markov chains. Then, in Sec. III D, the Hilbert metric is used to provide a constructive solution to Schrödinger's bridge problem for Markov chains. Section IV A overviews the formalism of quantum mechanics followed by a description of a quantum analog of the Schrödinger bridge problem in Sec. IV B. Section IV C presents a solution of the quantum Schrödinger bridge for the special case of uniform marginals. The mathematical statement for this special case represents a generalization of a result of Sinkhorn on the existence of doubly stochastic maps to a corresponding quantum probabilistic analog. This result is followed by a discussion and a conjecture about the general quantum Schrödinger problem, namely, that a fixed point of a certain map, which is used to construct doubly stochastic maps, suitably modified, has a fixed point for general marginal density matrices as well.

II. THE HILBERT METRIC

This metric was introduced by David Hilbert in 1895[16] while exploring the foundations of geometry. Earlier special cases, its importance and several subsequent developments are being discussed by Bushell.[7] More recently, the underlying geometry has proven timely on a range of problems in the study of communication and computations over networks (see Ref. 34 and in particular, the work of Sepulchre and collaborators[5,30] on consensus in non-commutative spaces, as well as the references therein) and in quantum information theory.[27] A recent survey on the applications in analysis is Ref. 23. A key result that enables the metric to be used for establishing existence of solutions to various equations was proved by Garrett Birkhoff in 1957.[3] Following Ref. 30, we highlight the basic elements of the theory.

Let $\mathcal{S}$ be a real Banach space and let $\mathcal{K}$ be a closed solid cone in $\mathcal{S}$, i.e., $\mathcal{K}$ is closed with nonempty interior and is such that $\mathcal{K}+\mathcal{K}\subseteq\mathcal{K}$, $\mathcal{K}\cap-\mathcal{K}=\{0\}$ as well as $\lambda\mathcal{K}\subseteq\mathcal{K}$ for all $\lambda\geq 0$. Define the partial order

$$x\preceq y \Leftrightarrow y-x\in\mathcal{K},$$

and for $x,y\in\mathcal{K}\backslash\{0\}$, define

$$M(x,y):=\inf\{\lambda \mid x\preceq\lambda y\},$$
$$m(x,y):=\sup\{\lambda \mid \lambda y\preceq x\}.$$

Then, the Hilbert metric is defined on $\mathcal{K}\backslash\{0\}$ by

$$d_H(x,y) := \log\left(\frac{M(x,y)}{m(x,y)}\right).$$

Strictly speaking, it is a *projective* metric since it remains invariant under scaling by positive constants, i.e., $d_H(x,y) = d_H(\lambda x, y) = d_H(x,\lambda y)$ for any $\lambda > 0$ and, thus, it actually measures distance between rays and not elements.

Birkhoff's theorem, which was originally stated for the linear case and suitably extended by Bushell,[7] provides bounds on the induced gain of positive maps. More specifically, a map $\mathcal{E}$ from $\mathcal{S}$ to $\mathcal{S}$ is said to be *positive* provided it takes the interior of $\mathcal{K}$ into itself, i.e.,

$$\mathcal{E} \ : \ \mathcal{K}\backslash\{0\} \to \mathcal{K}\backslash\{0\}.$$

For such a map, define its *projective diameter*

$$\Delta(\mathcal{E}) := \sup\{d_H(\mathcal{E}(x),\mathcal{E}(y)) \mid x, y \in \mathcal{K}\backslash\{0\}\}$$

and the *contraction ratio*

$$\|\mathcal{E}\|_H := \inf\{\lambda \mid d_H(\mathcal{E}(x),\mathcal{E}(y)) \le \lambda d_H(x,y), \text{ for all } x, y \in \mathcal{K}\backslash\{0\}\}.$$

The Birkhoff-Bushell theorem states the following.

Theorem 1 (Refs. 3, 6, and 7). *Let $\mathcal{E}$ be a positive map as above. If $\mathcal{E}$ is monotone and homogeneous of degree m, i.e., if*

$$x \le y \Rightarrow \mathcal{E}(x) \le \mathcal{E}(y)$$

and

$$\mathcal{E}(\lambda x) = \lambda^m \mathcal{E}(x),$$

then it holds that

$$\|\mathcal{E}\|_H \le m.$$

For the special case where $\mathcal{E}$ is also linear, the (possibly stronger) bound

$$\|\mathcal{E}\|_H = \tanh(\frac{1}{4}\Delta(\mathcal{E}))$$

also holds.

Birkhoff's result provides a far-reaching generalization of the celebrated Perron-Frobenius theorem.[4] Various other applications of the Birkhoff-Bushell result have been developed such as to positive integral operators and to positive definite matrices.[7,23] We will use Theorem 1 to establish existence of solutions to certain equations involving Markovian evolutions on a finite time interval. More specifically, $\mathcal{S}$ will either be $\mathbb{R}^n$ or the space of symmetric/Hermitian $n \times n$ matrices with elements in $\mathbb{R}$ or $\mathbb{C}$, accordingly. Then, $\mathcal{K}$ will either be the positive orthant of vectors with non-negative entries or the cone of non-negative definite matrices, respectively. The former setting will be brought in to give an independent proof of existence of the solution to the Schrödinger bridge problem for Markov chains. The latter will be called in for studying *quantum operations* (Kraus maps) in Sec. IV.

III. SCHRÖDINGER'S PROBLEM FOR DISCRETE RANDOM VECTORS

Following quite closely Schrödinger's original derivation, we give below a rather self-contained presentation for the purpose of later reference and comparison when we deal with the more complex quantum case. First, we discuss general discrete random vectors and then, "time windows" of a Markov chain. A paper dealing with the discrete time, continuous state space setting is Ref. 1. Schrödinger bridges for Markov chains have been discussed in Ref. 25. A nice survey with extensive bibliography for the diffusion case is Ref. 35.

A. Discrete random vectors

Given a finite set $\mathcal{X} = \{1, \dots, N\}$, we are concerned with probability distributions P on "trajectories" $x = (x_0, x_1, \dots, x_T)$ in $\mathcal{X}^{T+1}$. The case of a finite state space $\mathcal{X}$ is chosen for simplicity of exposition. Results extend in a straightforward way to the case of a countable $\mathcal{X}$. We write $\mathcal{P}$ for the simplex of all such distributions. Let us introduce the *coordinate mapping process* $X = \{X(t), 0 \le t \le T\}$ by $X(t)(x) = x_t$. For brevity, we often write $X(t) = x_t$ instead of $X(t)(x) = x_t$. For $P \in \mathcal{P}$, we denote by,

$$p(s, x_s; t, x_t) \coloneqq P(X(t) = x_t \mid X(s) = x_s), \quad 0 \le s < t \le T, \quad x_s, x_t \in \mathcal{X},$$

its transition probabilities. We also use p, but indexed, to denote marginals. Thus,

$$p_t(x_t) \coloneqq P(X(t) = x_t),$$

and similarly, for two-time marginals,

$$p_{st}(x_s, x_t) \coloneqq P(X(s) = x_s, X(t) = x_t).$$

Schrödinger's original formulation was set in continuous time and with continuous state space. The formulation herein represents the case where Schrödinger's problem has undergone "coarse graining" in Boltzmann's style in its phase space and where time has also been discretized. Thus, the *a priori* model is now given by a distribution $P \in \mathcal{P}$ and suppose that in experiments, an initial and a final marginals $\mathbf{p}_0$ and $\mathbf{p}_T$, respectively, have been observed that differ from the marginals p_0 and p_T of the prior distribution P. We denote by $\mathcal{P}(\mathbf{p}_0, \mathbf{p}_T) \subset \mathcal{P}$ the family of distributions having the observed marginals and seek a distribution in $\mathcal{P}(\mathbf{p}_0, \mathbf{p}_T)$ which is close to the given prior P. Large deviation reasoning[28] requires that we employ as "distance" the *relative entropy*.

Definition 1. Let $P, Q \in \mathcal{P}$, that is, they belong to the simplex of probability distributions on $\mathcal{X}^{T+1}$, and let $x = (x_0, x_1, \dots, x_T)$. If $P(x) = 0 \Rightarrow Q(x) = 0$, we say that the *support* of Q is contained in the support of P and write

$$\mathrm{Supp}(Q) \subseteq \mathrm{Supp}(P).$$

The *relative entropy* of Q from P is defined to be

$$\mathbb{D}(Q\|P) = \begin{cases} \sum_{x \in \mathcal{X}^{T+1}} Q(x) \log \dfrac{Q(x)}{P(x)}, & \mathrm{Supp}(Q) \subseteq \mathrm{Supp}(P), \\ +\infty, & \mathrm{Supp}(Q) \not\subseteq \mathrm{Supp}(P), \end{cases} \tag{1}$$

where, by definition, $0 \cdot \log 0 = 0$.

The relative entropy is also known as the *information* or *Kullback-Leibler divergence*. As is well known,[8] $\mathbb{D}(Q\|P) \ge 0$ and $\mathbb{D}(Q\|P) = 0$ if and only if $Q = P$. Given this notion of distance, we seek a probability law $Q^\circ \in \mathcal{P}(\mathbf{p}_0, \mathbf{p}_T)$ which is closest to the prior distribution P in this sense. That is, we seek a solution to the following problem.

Problem 1. Assume that $p(0, \cdot\,; T, \cdot)$ is everywhere positive on its domain. Determine

$$Q^\circ = \mathrm{argmin}\{\mathbb{D}(Q\|P) \mid Q \in \mathcal{P}(\mathbf{p}_0, \mathbf{p}_T)\}.$$

It turns out that if there is at least one Q in $\mathcal{P}(\mathbf{p}_0, \mathbf{p}_T)$ such that $\mathbb{D}(Q\|P) < \infty$, there exists a unique minimizer Q° called *the Schrödinger bridge* from $\mathbf{p}_0$ to $\mathbf{p}_T$ over P. Now let

$$Q_{x_0, x_T} = Q[\,\cdot \mid X(0) = x_0, X(T) = x_T]$$

be the disintegration of Q with respect to the initial and final positions. Then, we have

$$Q(x_0, x_1, \dots, x_T) = Q_{x_0, x_T}(x_0, x_1, \dots, x_T) q_{0T}(x_0, x_T),$$

where we have assumed that q_{0T} is everywhere positive on $\mathcal{X} \times \mathcal{X}$. We get

$$\mathbb{D}(Q\|P) = \sum_{x_0 x_T} q_{0T}(x_0, x_T) \log \frac{q_{0T}(x_0, x_T)}{p_{0T}(x_0, x_T)} + \sum_{x \in \mathcal{X}^{T+1}} Q_{x_0, x_T}(x) \log \frac{Q_{x_0, x_T}(x)}{P_{x_0, x_T}(x)} q_{0T}(x_0, x_T). \tag{2}$$

This is the sum of two non-negative quantities. The second becomes zero if and only if $Q_{x_0,x_T}(x) = P_{x_0,x_T}(x)$ for all $x \in \mathcal{X}^{T+1}$. Thus, $Q^\circ_{x_0,x_T}(x) = P_{x_0,x_T}(x)$. In particular, this shows that when both marginals are delta distributions, the solution is simply obtained from the prior through conditioning. From this point of view, Schrödinger bridges with general marginals appear as a sort of "soft conditioning" of the prior thereby generalizing the Bayesian approach.

As already observed by Schrödinger, Problem 1 now reduces to minimizing

$$\mathbb{D}(q_{0T}\|p_{0T}) = \sum_{x_0x_T} q_{0T}(x_0,x_T)\log\frac{q_{0T}(x_0,x_T)}{p_{0T}(x_0,x_T)} \tag{3}$$

with respect to q_{0T} subject to the (linear) constraints

$$\sum_{x_T} q_{0T}(x_0,x_T) = \mathbf{p}_0(x_0), \quad x_0 \in \mathcal{X}, \tag{4}$$

$$\sum_{x_0} q_{0T}(x_0,x_T) = \mathbf{p}_T(x_T), \quad x_T \in \mathcal{X}. \tag{5}$$

The Lagrangian function has the form

$$\mathcal{L}(q_{0T}) = \sum_{x_0x_T} q_{0T}(x_0,x_T)\log\frac{q_{0T}(x_0,x_T)}{p_{0T}(x_0,x_T)}$$
$$+\sum_{x_0}\lambda(x_0)\left[\sum_{x_T} q_{0T}(x_0,x_T) - \mathbf{p}_0(x_0)\right] + \sum_{x_T}\mu(x_T)\left[\sum_{x_0} q_{0T}(x_0,x_T) - \mathbf{p}_T(x_T)\right].$$

Setting the first variation equal to zero, we get the (sufficient) optimality condition

$$1 + \log q^\circ_{0T}(x_0,x_T) - \log p(0,x_0;T,x_T) - \log p_0(x_0) + \lambda(x_0) + \mu(x_T) = 0,$$

where we have used the expression $p_{0T}(x_0,x_T) = p_0(x_0)p(0,x_0;T,x_T)$. Hence, the ratio $q^\circ_{0T}(x_0,x_T)/p(0,x_0;T,x_T)$ factors into a function of x_0 times a function of x_T; these are denoted $\hat\varphi(x_0)$ and $\varphi(x_T)$, respectively. We can then write the optimal $q^\circ_{0T}(\cdot,\cdot)$ in the form

$$q^\circ_{0T}(x_0,x_T) = \hat\varphi(x_0)p(0,x_0;T,x_T)\varphi(x_T), \tag{6}$$

where φ and $\hat\varphi$ must satisfy

$$\hat\varphi(x_0)\sum_{x_T} p(0,x_0;T,x_T)\varphi(x_T) = \mathbf{p}_0(x_0), \tag{7}$$

$$\varphi(x_T)\sum_{x_0} p(0,x_0;T,x_T)\hat\varphi(x_0) = \mathbf{p}_T(x_T). \tag{8}$$

Let us define $\hat\varphi(0,x_0) = \hat\varphi(x_0)$, $\varphi(T,x_T) = \varphi(x_T)$ and

$$\hat\varphi(T,x_T) = \sum_{x_0} p(0,x_0;T,x_T)\hat\varphi(0,x_0), \quad \varphi(0,x_0) \coloneqq \sum_{x_T} p(0,x_0;T,x_T)\varphi(T,x_T).$$

Then, (7) and (8) can be replaced by the system

$$\hat\varphi(T,x_T) = \sum_{x_0} p(0,x_0;T,x_T)\hat\varphi(0,x_0), \tag{9}$$

$$\varphi(0,x_0) \coloneqq \sum_{x_T} p(0,x_0;T,x_T)\varphi(T,x_T), \tag{10}$$

with the boundary conditions

$$\varphi(0,x_0)\cdot\hat\varphi(0,x_0) = \mathbf{p}_0(x_0), \quad \varphi(T,x_T)\cdot\hat\varphi(T,x_T) = \mathbf{p}_T(x_T), \quad \forall x_0,x_T \in \mathcal{X}. \tag{11}$$

The question of existence and uniqueness of functions $\hat\varphi(x_0)$, $\varphi(x_T)$ satisfying (9)–(11) will be established in Sec. III D. Before we do that, however, we investigate what else can be said about the solution when the prior random vector happens to be a "time window" of a Markov chain.

B. Markovian prior

Consider the special case where $X = \{X(t), 0 \le t \le T\}$ is a window of a Markov chain. Let us introduce, in the language of Doob, the *space-time harmonic* function

$$\varphi(t, x_t) := \sum_{x_T} p(t, x_t; T, x_T)\varphi(x_T), \quad 0 \le t \le T \tag{12}$$

and the *space-time co-harmonic* function

$$\hat{\varphi}(t, x_t) := \sum_{x_0} p(0, x_0; t, x_t)\hat{\varphi}(x_0), \quad 0 \le t \le T. \tag{13}$$

Because of the Markov property, for $0 \le s < t < u \le T$, we have

$$p(s, x_s; u, x_u) = \sum_{x_t} p(s, x_s; t, x_t)p(t, x_t; u, x_u).$$

For compactness, we often write $\pi_{x_t, x_{t+1}}(t)$ and $\pi^{(n)}_{x_t, x_{t+n}}(t)$ instead of $p(t, x_t; t+1, x_{t+1})$ and $p(t, x_t; t+n, x_{t+n})$, respectively. Hence, φ and $\hat{\varphi}$ satisfy the backward and forward equations, respectively,

$$\varphi(t, x_t) = \sum_{x_{t+1}} \pi_{x_t, x_{t+1}}(t)\varphi(t+1, x_{t+1}), \tag{14}$$

$$\hat{\varphi}(t+1, x_{t+1}) = \sum_{x_t} \pi_{x_t, x_{t+1}}(t)\hat{\varphi}(t, x_t). \tag{15}$$

Let q°_t denote the distribution of the Schrödinger bridge at time t. We get

$$q^\circ_t(x_t) = \sum_{x_0}\sum_{x_T} p(0, x_0; t, x_t; T, x_T)q^\circ_{0T}(x_0, x_T) =$$

$$\sum_{x_0}\sum_{x_T} \frac{p(0, x_0; t, x_t)p(t, x_t; T, x_T)}{p(0, x_0; T, x_T)}\hat{\varphi}(x_0)p(0, x_0; T, x_T)\varphi(x_T)$$

$$= \hat{\varphi}(t, x_t) \cdot \varphi(t, x_t). \tag{16}$$

Similarly, one gets

$$q^\circ_{st}(x_s, x_t) = \sum_{x_0}\sum_{x_T} \hat{\varphi}(x_0)p(0, x_0; s, x_s)p(s, x_s; t, x_t)p(t, x_t; T, x_T)\varphi(x_T)$$

which yields, using (12), (13) and (16), the new transition probabilities

$$q^\circ(s, x_s; t, x_t) = \frac{q^\circ_{st}(x_s, x_t)}{q^\circ_s(x_s)} = p(s, x_s; t, x_t)\frac{\varphi(t, x_t)}{\varphi(s, x_s)}. \tag{17}$$

Notice, in particular, that the Schrödinger bridge is also a *Markov chain* which is obtained from the *a priori* model via a suitable *multiplicative functional transformation* just like in the diffusion case.[21] We have therefore established the following result Ref. 25, Theorem 4.1.

Theorem 2. *Assume that $p(0, \cdot; T, \cdot)$ is everywhere positive on $\mathcal{X} \times \mathcal{X}$. Suppose there exist positive functions φ and $\hat{\varphi}$ defined on $[0, T] \times \mathcal{X}$ satisfying for $t \in [0, T-1]$, the system*

$$\varphi(t, x_t) = \sum_{x_{t+1}} \pi_{x_t, x_{t+1}}(t)\varphi(t+1, x_{t+1}), \tag{18}$$

$$\hat{\varphi}(t+1, x_{t+1}) = \sum_{x_t} \pi_{x_t, x_{t+1}}(t)\hat{\varphi}(t, x_t), \tag{19}$$

with the boundary conditions

$$\varphi(0, x_0) \cdot \hat{\varphi}(0, x_0) = \mathbf{p}_0(x_0), \quad \varphi(T, x_T) \cdot \hat{\varphi}(T, x_T) = \mathbf{p}_T(x_T), \quad \forall x_0, x_T \in \mathcal{X}. \tag{20}$$

Then, the Markov distribution Q° in $\mathcal{P}(\mathbf{p}_0, \mathbf{p}_T)$ with one-step transition probabilities

$$\pi^{\circ}_{x_t,x_{t+1}}(t) = \pi_{x_t,x_{t+1}}(t)\frac{\varphi(t+1,x_{t+1})}{\varphi(t,x_t)} \tag{21}$$

is the unique solution of Problem 1.

Notice that, if a pair $\varphi,\hat{\varphi}$ solves (18)–(20), so does the pair $a\varphi, a^{-1}\hat{\varphi}$ for $a > 0$. This arbitrariness in the scaling, however, does not affect transition probabilities (21).

Schrödinger and Kolmogorov were struck by the intrinsic time-reversibility of the solution: swapping $\mathbf{p}_0$ and $\mathbf{p}_T$ leads to a solution bridge which is simply the time reversal of the original one. Moreover, factorization (16) resembles Born's relation between the probability density and the wave function in quantum mechanics $\rho_t(x) = \psi(x,t)\psi^{\dagger}(x,t)$, where †, throughout the paper, denotes conjugation/adjoint in a complex Hilbert space.[29]

At this point, it should be apparent that the bottleneck of the theory of Schrödinger bridges is the existence and uniqueness (up to multiplication by a positive constant) of the pair $\varphi,\hat{\varphi}$ solving Schrödinger system (9)–(11) (in the Markov case, (18)–(20)). Schrödinger thought that existence and uniqueness should be guaranteed in the diffusion case "except possibly for very nasty ρ_0, ρ_T, since the question leading to the pair of equations is so reasonable." This problem turns out to be quite nontrivial and was settled in various degrees of generality by Beurling, Jamison, and Föllmer[2,12,20] establishing the feasibility of the dual optimization problem of Problem 1. There is, however, an alternative approach based on proving convergence of successive approximations by Fortet.[13] As it turns out, this more algorithmic approach had independent counterparts called "iterative fitting algorithms" in the statistical literature on contingency tables.[10] These were later shown to converge to a "minimum discrimination information,"[11,19] namely, to a minimum entropy distance, see also Ref. 9. It will be apparent below that the approach pioneered by Fortet features a very desirable property: while establishing existence and uniqueness for the Schrödinger system, it also provides a computationally efficient *algorithm* to actually *compute* the space-time harmonic function φ and therefore the solution.

Our approach consists in showing that a certain iterative scheme for the Schrödinger system is a contraction mapping on the positive orthant with respect to the projective metric which was introduced in Sec. II. We argue that this is the *natural* metric in which to cast the iteration for these problems both in the classical and in the quantum case. Before we turn to that, however, we show that in the Markov case, the new transition mechanism may be obtained via the solution of a two consecutive times interval problem.

C. Reduction to the one-step bridge problem

Let $\Pi(t) = \left(\pi_{x_t,x_{t+1}}(t)\right)_{x_t,x_{t+1}=1}^{N}$ and $\Pi^{\circ}(t) = \left(\pi^{\circ}_{x_t,x_{t+1}}(t)\right)_{x_t,x_{t+1}=1}^{N}$ be the transition matrices of the prior distribution and of the bridge distribution. It is interesting to express (21) in matrix form. To this end, we introduce the notation

$$\phi(t) = \mathrm{diag}\left(\varphi(t,x_1),\varphi(t,x_2),\ldots,\varphi(t,x_N)\right)$$

for a diagonal matrix formed out of the entries of $\varphi(t,\cdot)$. We now have

$$\Pi^{\circ}(t) = \phi(t)^{-1}\Pi(t)\phi(t+1). \tag{22}$$

Consider now for $n \geq 1$ $\Pi^{(n)}(t) = \left(\pi^{(n)}_{x_t,x_{t+n}}(t)\right)_{x_t,x_{t+n}=1}^{N}$ the n-step transition probabilities matrix. By the Markov property, we have

$$\Pi^{(n)}(t) = \Pi(t)\cdot\Pi(t+1)\cdots\Pi(t+n-1).$$

Being the product of stochastic matrices, $\Pi^{(n)}(t)$ is also *stochastic*. Indeed, the elements of $\Pi^{(n)}(t)$ are non-negative as sum of products of non-negative numbers. Moreover, let $\mathbb{1}^{\dagger} = (1,1,\ldots,1)$. Then, the fact that a matrix Q has rows summing to one can be expressed as $Q\mathbb{1} = \mathbb{1}$. As each $\Pi(t)$ has rows summing to one, we have

$$\begin{aligned}\Pi^{(n)}(t)\mathbb{1} &= \Pi(t)\cdot\Pi(t+1)\cdots\Pi(t+n-1)\mathbb{1}\\ &= \Pi(t)\cdot\Pi(t+1)\cdots\Pi(t+n-2)\mathbb{1}\\ &= \cdots = \Pi(t)\mathbb{1} = \mathbb{1}.\end{aligned}$$

Consider now $\Pi^{\circ(n)}(t)$. By (22), we get

$$\begin{aligned}\Pi^{\circ(n)}(t) &= \Pi^{\circ}(t)\cdot\Pi^{\circ}(t+1)\cdots\Pi^{\circ}(t+n-1)\\ &= \phi(t)^{-1}\Pi^{(n)}(t)\phi(t+n).\end{aligned}$$

Thus, we get the following remarkable generalization of formula (21):

$$\pi^{\circ(n)}_{x_t x_{t+n}}(t) = \pi^{(n)}_{x_t,x_{t+n}}(t)\frac{\varphi(t+n,x_{t+n})}{\varphi(t,x_t)}. \tag{23}$$

Consider now $\Pi^{(T)}(0)$ as the transition matrix of *a prior* "stroboscopic" evolution. Consider also the Schrödinger bridge problem with the same marginals $\mathbf{p}_0$ and $\mathbf{p}_T$ as before at the two *consecutive* times 0 and T. Specializing Theorem 2 to this simple situation, we get that the new transition probabilities are precisely given by (23) with $n = T$, where φ, $\hat{\varphi}$ satisfy (9)–(11), namely,

$$\varphi(0,x_0) = \sum_{x_T}\pi^{(T)}_{x_0,x_T}(0)\varphi(T,x_T), \tag{24}$$

$$\hat{\varphi}(T,x_T) = \sum_{i}\pi^{(T)}_{x_0,x_T}(0)\hat{\varphi}(0,x_0), \tag{25}$$

with the boundary conditions

$$\varphi(0,x_0)\cdot\hat{\varphi}(0,x_0) = \mathbf{p}_0(x_0), \quad \varphi(T,x_T)\cdot\hat{\varphi}(T,x_T) = \mathbf{p}_T(x_T), \quad \forall x_0, x_T \in \mathcal{X}. \tag{26}$$

We see that the solution of Problem 1 yields, as a by-product, the solution of the one-step problem. The converse, however, is also true. Solving (24)-(26) yields the correct terminal value $\varphi(T,\cdot)$ from which φ can be computed at all times through iteration (18). From it, the new transition probabilities are obtained through (21). We already know from Sec. III A that solving the one-step problem suffices to characterize the optimal distribution Q°. The argument above shows that it also permits to obtain the more explicit description, namely, the transition mechanism of Q°, which is desirable in the Markov case.

D. The solution to the one-step bridge problem

Motivated by what we have seen in Secs. III A and III C, we study the one-step bridge problem with possibly non-Markovian prior. The following result establishes existence and uniqueness for system (9)–(11).

Theorem 3. *Given an $N \times N$ stochastic matrix*

$$\Pi = \left[\pi_{x_0,x_T}\right]_{x_0,x_T=1}^{N},$$

with strictly positive entries ($\pi_{x_0,x_T} > 0$) and probability distributions $\mathbf{p}_0$, $\mathbf{p}_T$, there exist four vectors $\varphi(0,x_0)$, $\varphi(T,x_T)$, $\hat{\varphi}(0,x_0)$, $\hat{\varphi}(T,x_T)$, indexed by $x_0, x_T \in \mathcal{X}$, with positive entries such that

$$\varphi(0,x_0) = \sum_{x_T}\pi_{x_0,x_T}\varphi(T,x_T), \tag{27a}$$

$$\hat{\varphi}(T,x_T) = \sum_{x_0}\pi_{x_0,x_T}\hat{\varphi}(0,x_0), \tag{27b}$$

$$\varphi(0,x_0)\hat{\varphi}(0,x_0) = \mathbf{p}_0(x_0), \tag{27c}$$

$$\varphi(T,x_T)\hat{\varphi}(T,x_T) = \mathbf{p}_T(x_T). \tag{27d}$$

The four vectors are unique up to multiplication of $\varphi(0,x_0)$ and $\varphi(T,x_T)$ by the same positive constant and division of $\hat{\varphi}(0,x_0)$ and $\hat{\varphi}(T,x_T)$ by the same constant.

Rather than relying on the results of Beurling and Jamison as in Refs. 25 and 33, we give an independent proof which also yields an effective algorithm. The proof relies on showing that a certain iteration is strictly contractive in the Hilbert metric. We note that Fortet's proof,[13] which we find very difficult to follow, is apparently based on establishing monotonicity of two sequences of functions produced by the iteration.

Lemma 1. Consider the following circular diagram of maps:

$$\begin{array}{ccccc} & & \hat{\varphi}(0,x_0) & \xrightarrow{\mathcal{E}^\dagger} & \hat{\varphi}(T,x_T) \quad = \sum_{x_0} \pi_{x_0,x_T}\hat{\varphi}(0,x_0) \\ \hat{\varphi}(0,x_0) = \dfrac{\mathbf{p}_0(x_0)}{\varphi(0,x_0)} & & \uparrow & & \downarrow \qquad \varphi(T,x_T) = \dfrac{\mathbf{p}_T(x_T)}{\hat{\varphi}(T,x_T)}\,, \\ \sum_{x_N} \pi_{x_0,x_N}\varphi(T,x_T) = & & \varphi(0,x_0) & \xleftarrow{\mathcal{E}} & \varphi(T,x_T) \end{array} \tag{28}$$

where

$$\begin{aligned} \hat{\mathcal{D}}_0 \;:\; \varphi(0,x_0) &\mapsto \hat{\varphi}(0,x_0) = \frac{\mathbf{p}_0(x_0)}{\varphi(0,x_0)}, \\ \mathcal{D}_T \;:\; \hat{\varphi}(T,x_T) &\mapsto \varphi(T,x_T) = \frac{\mathbf{p}_T(x_N)}{\hat{\varphi}(T,x_T)} \end{aligned}$$

represent componentwise division of vectors. Then, the composition

$$\hat{\varphi}(0,x_0) \xrightarrow{\mathcal{E}^\dagger} \hat{\varphi}(T,x_T) \xrightarrow{\mathcal{D}_T} \varphi(T,x_T) \xrightarrow{\mathcal{E}} \varphi(0,x_0) \xrightarrow{\mathcal{D}_0} (\hat{\varphi}(0,x_0))_{\text{next}} \tag{29}$$

is contractive in the Hilbert metric.

Before proceeding with the proof, we provide a note about notation: the map

$$\mathcal{E}^\dagger \;:\; \hat{\varphi}(0,x_0) \mapsto \hat{\varphi}(T,x_T) = \sum_{x_0} \pi_{x_0,x_N}\hat{\varphi}(0,x_0)$$

is the adjoint of the backward evolution

$$\mathcal{E} \;:\; \varphi(T,x_T) \mapsto \varphi(0,x_0) = \sum_{x_N} \pi_{x_0,x_T}\varphi(T,x_T),$$

which is consistent with the standard notation in diffusion processes where the Fokker-Planck (forward) equation involves the adjoint of the *generator* appearing in the backward Kolmogorov equation.

Also notice that the componentwise divisions of $\hat{\mathcal{D}}_0$ and $\mathcal{D}_T$ are well defined. Indeed, even when $\hat{\varphi}(0)$ ($\varphi(T)$) has zero entries, $\hat{\varphi}(T)$ ($\varphi(0)$) has all positive entries since the elements of Π are all positive.

Proof of Lemma 1. The diameter of the range of $\mathcal{E}$ is

$$\begin{aligned} \Delta(\mathcal{E}) &= \sup\{d_H(\mathcal{E}(x),\mathcal{E}(y)) \mid x_i > 0,\ y_i > 0\} \\ &= \sup\{\log\left(\frac{\pi_{ij}\pi_{k,\ell}}{\pi_{i,\ell}\pi_{k,j}}\right) \mid 1 \le i,j,k,\ell \le n\}, \end{aligned}$$

which is finite since all entries $\pi_{i,j}$'s are positive. Birkhoff's theorem (Theorem 1) provides a contraction coefficient for linear positive maps and, in this case, we have

$$\|\mathcal{E}\|_H = \tanh(\frac{1}{4}\Delta(\mathcal{E})) < 1.$$

For the adjoint map $\mathcal{E}^\dagger$, we only need to note that it is homogeneous of degree 1 and therefore, by Birkhoff's theorem,

$$\|\mathcal{E}^\dagger\|_H \le 1.$$

Next, we note that provided $\mathbf{p}_0(x_0)$ and $\mathbf{p}_T(x_N)$ have positive entries, both $\hat{\mathcal{D}}_0$ and $\mathcal{D}_T$ are isometries in the Hilbert metric since inversion and element-wise scaling are both isometries. Indeed, for vectors $[x_i]_{i=1}^N$, $[y_i]_{i=1}^N$, it holds that

$$\begin{aligned} d_H([x_i],[y_i]) &= \log\left((\max_i(x_i/y_i))\frac{1}{\min_i(x_i/y_i)}\right) \\ &= \log\left(\frac{1}{\min_i((x_i)^{-1}/(y_i)^{-1})}\max_i((x_i)^{-1}/(y_i)^{-1})\right) \\ &= d_H([(x_i)^{-1}],[(y_i)^{-1}]) \end{aligned}$$

and

$$\begin{aligned} d_H([\mathbf{p}_i x_i],[\mathbf{p}_i y_i]) &= \log\frac{\max_i((\mathbf{p}_i x_i)/(\mathbf{p}_i y_i))}{\min_i((\mathbf{p}_i x_i)/(\mathbf{p}_i y_i))} \\ &= \log\frac{\max_i(x_i/y_i)}{\min_i(x_i/y_i)} = d_H([x_i],[y_i]). \end{aligned}$$

If $\mathbf{p}_0(x_0)$ and $\mathbf{p}_T(x_N)$ have any zero entries, then $\hat{\mathcal{D}}_0$ and $\mathcal{D}_T$ are in fact contractions. Finally, we observe that

$$\|\hat{\mathcal{D}}_0 \circ \mathcal{E} \circ \mathcal{D}_T \circ \mathcal{E}^\dagger\|_H \le \|\hat{\mathcal{D}}_0\|_H \cdot \|\mathcal{E}\|_H \cdot \|\mathcal{D}_T\|_H \cdot \|\mathcal{E}^\dagger\|_H < 1,$$

where $\circ$ denotes composition. Therefore, the composition $C := \hat{\mathcal{D}}_0 \circ \mathcal{E} \circ \mathcal{D}_T \circ \mathcal{E}^\dagger$ is contractive: namely,

$$d_H(C(x),C(y)) \le \|C\|_H d_H(x,y), \quad 0 \le \|C\|_H < 1.$$

This completes the proof of the lemma. □

It was pointed to us by Idel Martin that the use of the Hilbert metric for this type of problem is in fact central and standard in the nonlinear Frobenius-Perron theory.[22]

Proof of Theorem 3. Since C is contractive in the Hilbert metric, there is a unique positive $\hat{\varphi}(0,\cdot) = [\hat{\varphi}(0,x_0)]$ so that the corresponding ray is invariant under C. That is, in the notation of (29),

$$\begin{aligned} (\hat{\varphi}(0,\cdot))_{\text{next}} &= C(\hat{\varphi}(0,\cdot)) \\ &= \lambda\hat{\varphi}(0,\cdot), \end{aligned}$$

for the composition $C := \hat{\mathcal{D}}_0 \circ \mathcal{E} \circ \mathcal{D}_T \circ \mathcal{E}^\dagger$. From this, we can obtain

$$\begin{aligned} \hat{\varphi}(T,\cdot) &= \mathcal{E}^\dagger(\hat{\varphi}(0,\cdot)), \\ \varphi(0,\cdot) &= \mathcal{E}(\varphi(T,\cdot)), \end{aligned}$$

while

$$\begin{aligned} \hat{\varphi}(T,\cdot)\varphi(T,\cdot) &= \mathbf{p}_T(\cdot) \\ \lambda\hat{\varphi}(0,\cdot)\varphi(0,\cdot) &= \mathbf{p}_0(\cdot). \end{aligned}$$

However, since

$$\begin{aligned} 1 &= \lambda\langle\hat{\varphi}(0,\cdot),\varphi(0,\cdot)\rangle \\ &= \lambda\langle\hat{\varphi}(0,\cdot),\mathcal{E}(\varphi(T,\cdot))\rangle \\ &= \lambda\langle\mathcal{E}^\dagger(\hat{\varphi}(0,\cdot)),\varphi(T,\cdot)\rangle \\ &= \lambda\langle\hat{\varphi}(T,\cdot),\varphi(T,\cdot)\rangle \\ &= \lambda, \end{aligned}$$

we conclude that these satisfy Schrödinger system (27a)–(27d). □

Remark 1. Numerical algorithm: The step above suggests that starting from any positive vector, e.g., $x(0) = \mathbb{1}$, the unique fixed point is given by

$$\hat{\varphi}(0) = \lim_{k\to\infty} x(k), \tag{30}$$

where $x(k+1) = C(x(k))$ *for* $k = 1, 2, \ldots$.

By the variational analysis of Sec. III A, we finally get the following result.

Theorem 4. *Let* $\mathbf{p}_0$ *and* $\mathbf{p}_T$ *be probability distributions on* $\mathcal{X}$ *and assume that* $p(0,\cdot;T,\cdot)$ *is everywhere positive on* $\mathcal{X}\times\mathcal{X}$. *Then, the unique solution to Problem 1 with prior* $P \in \mathcal{P}$ *is given by*

$$Q^\circ(x_0, x_1, \ldots, x_T) = P_{x_0,x_T}(x_0, x_1, \ldots, x_T) q^\circ_{0T}(x_0, x_T),$$

where $q^\circ_{0T}(x_0, x_T)$, *solving the problem for the two consecutive times* $0, T$, *is given by*

$$q^\circ_{0T}(x_0, x_T) = \hat{\varphi}(0, x_0) p(0, x_0; T, x_T) \varphi(T, x_T). \tag{31a}$$

Here, $\hat{\varphi}(0, x_0)$, $\varphi(T, x_T)$ *are as in Theorem 3 with*

$$\Pi = [\pi_{x_0,x_T}]^N_{x_0,x_T=1}, \quad \pi_{x_0,x_T} = p(0, x_0; T, x_T).$$

The corresponding transition probability is given by

$$q^\circ(0, x_0; T, x_T) = p(0, x_0; T, x_T) \frac{\varphi(T, x_T)}{\varphi(0, x_0)}. \tag{32}$$

In fact, when P is Markovian, we also get the new one-step transition probabilities in a similar manner, i.e., in the way described at the end of Sec. III C. Finally, as already observed, extending the results of this section to the case of a countable state space is straightforward.

IV. THE SCHRÖDINGER BRIDGE PROBLEM FOR QUANTUM EVOLUTIONS

We begin with some background on quantum probability and stochastic maps, see, e.g., Refs. 18, 24, and 26. In quantum probability there is neither a notion of a random variable nor of a probability space. Instead, random variables are replaced by Hermitian matrices/operators, referred to as *observables*, while the expected outcome of an experiment is quantified by a suitable functional that provides the *expectation*. Throughout, we restrict our attention to the case of finite-dimensional quantum systems with associated Hilbert space isomorphic to $\mathbb{C}^n$. Here, experiments take values from a finite alphabet and, thereby, the mathematical framework relies on algebras of finite matrices. In this case, the state of the underlying system which defines the expectation is represented by a non-negative definite matrix with trace one which is referred to as a *density matrix*. Notation and basic terminology are explained next, followed by the formulation, the quantum bridge problem, namely, the problem of identifying a quantum channel that is consistent with given quantum states at its two ends and has a specific relation to a given prior quantum channel. For potential applications of quantum bridges and links to topics in quantum information, we refer to Ref. 25, Sec. VII and to Ref. 15 where a surprising connection between the main problem studied in this section and quantum entanglement is explained.

A. Quantum channels

We use $\mathfrak{M} := \{X \in \mathbb{C}^{n\times n}\}$ to denote square matrices, $\mathfrak{H} := \{X \in \mathbb{C}^{n\times n} \mid X = X^\dagger\}$ to denote the set of Hermitian matrices, $\mathfrak{H}_+$ and $\mathfrak{H}_{++}$ the cones of non-negative and positive definite ones, respectively, and $\mathfrak{D} := \{\rho \in \mathfrak{H}_+ \mid \text{trace}\,\rho = 1\}$ the set of density matrices. The latter represents possible "states" of a quantum system. In turn, a state ρ defines an *expectation* functional on observables $X \in \mathfrak{H}$,

$$\mathbb{E}_\rho(X) = \text{trace}(\rho X).$$

The standard model for a quantum experiment as well as for a quantum channel is provided by a linear *trace-preserving completely positive* (TPCP) map between density matrices (which may possibly be of different sizes). This is the quantum counterpart of a Markov evolution and is referred

to as a *Kraus map*. For the finite-dimensional case treated herein, where quantum systems are represented by density matrices and where, for notational simplicity, matrices are all of the same size, a Kraus map assumes the representation

$$\mathcal{E}^\dagger \ : \ \mathfrak{D} \to \mathfrak{D} \ : \ \rho \longrightarrow \sigma = \sum_{i=1}^{n_\mathcal{E}} E_i \rho E_i^\dagger, \tag{33a}$$

with $E_i \in \mathfrak{M}$ such that

$$\sum_{i=1}^{n_\mathcal{E}} E_i^\dagger E_i = I, \tag{33b}$$

see Ref. 26, Chapter 2. Throughout, I denotes the identity matrix. Condition (33b) ensures that the map preserves the trace, i.e., that $\mathrm{trace}(\sigma) = \mathrm{trace}(\rho) = 1$. *Completely positive* refers to the property that, besides the fact that $\mathcal{E}^\dagger(\rho) \geq 0$ for all $\rho \geq 0$ (i.e., being a *positive* map), if $\mathbb{I}_k$ denotes the identity map on $\mathbb{C}^{k\times k}$ and $\otimes$, the tensor product, then $\mathbb{I}_k \otimes \mathcal{E}^\dagger$, is a positive map for all k. The Kraus-representation ((33a) and (33b)) characterizes completely positive maps (see, e.g., Ref. 26, Chapter 2).

Herein, a further condition will be imposed on Kraus maps which we refer to as *positivity improving*. This is the property that

$$\mathcal{E}^\dagger(\rho) > 0 \text{ for any } \rho \in \mathfrak{D}. \tag{34}$$

Thus, positivity improving amounts to having the range of $\mathcal{E}^\dagger$ contained in the interior of $\mathfrak{D}$ or, in other words and in view of (33a), that there is no pair of vectors $w, v \in \mathbb{C}^n$ such that $w^\dagger E_i v = 0$ for all $i \in \{1,\dots,n_\mathcal{E}\}$. A necessary condition for positivity improving is that $n_\mathcal{E} > n$. To see this, assume that $n_\mathcal{E} \leq n$ and take v to be an eigenvector of the matrix pencil $E_1 - \lambda E_2$. Then, $E_1 v$ and $E_2 v$ are linearly dependent and the span of $E_1 v \dots E_n v$ has dimension less than n. This implies that there exists a w such that $w^\dagger E_i v = 0$ for all i. Determining whether a given Kraus map actually has the property of positivity improving, in general, is NP-hard (namely any NP problem can be reduced in polynomial time to it).[14,17]

Our notation in using $\mathcal{E}^\dagger$ for the Kraus map is not standard and differs from that in, e.g., Ref. 26 where $\mathcal{E}$ is used instead. However, our choice maintains consistency with the standard convention throughout probability theory where the (forward) Fokker-Planck equation involves the adjoint of the generator of a Markov semigroup.

A model for quantum measurement amounts to an instantiation of a quantum channel Ref. 26, Chapter 2. Typically, an observable $X \in \mathfrak{H}$ is specified having a decomposition

$$X = \sum_i \lambda_i \chi_i^\dagger \chi_i,$$

with $\chi_i \in \mathfrak{M}$ such that

$$\sum_i \chi_i^\dagger \chi_i = I.$$

A particular case is that of a spectral decomposition of X. When the outcome i is being recorded and the value λ_i registered for a quantum system that is initially in a state ρ, then, after the measurement, the system finds itself in the new state

$$\left(\frac{1}{\mathrm{trace}(\chi_i \rho \chi_i^\dagger)}\right) \chi_i \rho \chi_i^\dagger. \tag{35}$$

The likelihood of this particular outcome associated with recording λ_i is $\mathrm{trace}(\chi_i \rho \chi_i^\dagger)$. Therefore,

$$\begin{aligned} \mathbb{E}_\rho(X) &= \sum_i \lambda_i \,\mathrm{trace}(\chi_i \rho \chi_i^\dagger) \\ &= \mathrm{trace}(\rho X) \end{aligned}$$

in agreement with the earlier statement. On the other hand, if the experiment takes place but no particular outcome is recorded or, equivalently, all possible states in (35) are weighed in by the corresponding probabilities trace$(\chi_i \rho \chi_i^\dagger)$, the new state becomes

$$\sum_i \chi_i \rho \chi_i^\dagger.$$

This is precisely a Kraus map acting on ρ and this type of measurement is referred to as *nonselective*. In the sequel, we will encounter observables $\phi \in \mathfrak{H}$ factored as $\phi = \chi^\dagger \chi$. These represent quantum analogues of the random variables φ in the classical case. The expression $\chi \rho \chi^\dagger$ may be thought of as an "unnormalized" state that results in after a nonzero outcome has been recorded when a measurement performed. The factors of observables ϕ will give rise to the multiplicative functional transformations on Kraus maps sought in addressing the quantum bridge problem. These specify Kraus maps that link a set of two given "marginal" density matrices.

While there are many similarities between quantum channels and Markov evolution, there are also stark differences. A fundamental departure from classical probability arises in that, in general, there is no notion of joint probability between "measurements" at the two ends of a quantum channel. In this, the order with which measurements take place matters.

B. The bridge problem

Consider a reference quantum evolution given by an initial density matrix σ_0 and by a sequence of TPCP maps $\{\mathcal{E}_t^\dagger; 0 \le t \le T-1\}$, each admitting a Kraus representation with matrices $\{E_{t,i}\}$, so that

$$\mathcal{E}_t^\dagger : \sigma_t \mapsto \sigma_{t+1} = \sum_i E_{t,i} \sigma_t E_{t,i}^\dagger, \quad t = 0, 1, \dots, T-1.$$

As usual,

$$\sum_i E_{t,i}^\dagger E_{t,i} = I.$$

Consider also the composition map

$$\mathcal{E}_{0:T}^\dagger := \mathcal{E}_{T-1}^\dagger \circ \cdots \circ \mathcal{E}_1^\dagger \circ \mathcal{E}_0^\dagger.$$

Observe that $\mathcal{E}_{0:T}^\dagger$ is a TPCP map. For instance, it is immediate that $\mathcal{E}_1^\dagger \circ \mathcal{E}_0^\dagger$ admits a (double-indexed) Kraus representation with matrices $\{E_{1,j} E_{0,i}\}$. Hence, as thoroughly argued in the classical case, we only need to consider the one-step situation. Besides the reference quantum evolution, we are also given an initial and a final positive definite density matrices ρ_0 and ρ_T, respectively.

We consider the problem of finding a new quantum evolution $\mathcal{F}_{0:T}^\dagger = \mathcal{F}_{T-1}^\dagger \circ \cdots \circ \mathcal{F}_1^\dagger \circ \mathcal{F}_0^\dagger$, where the $\mathcal{F}_t^\dagger$ are TPCP maps, such that the new evolution is close to the reference one and

$$\mathcal{F}_{0:T}^\dagger(\rho_0) = \rho_T. \tag{36}$$

Providing a mathematical formulation for the selection of $\mathcal{F}_t$'s is nontrivial. In Ref. 25, which deals with the simpler problem where only the initial or final density matrix is prescribed, *quantum paths* were introduced via a family of observables $\{X_t; 0 \le t \le T\}$ with spectral decomposition

$$X_t = \sum_{i_t=1}^{m_t} x_{i_t} \Pi_{i_t}(t).$$

Then, paths are defined as sequences of "events" $\big(\Pi_{i_0}(0), \Pi_{i_1}(1), \dots, \Pi_{i_T}(T)\big)$ and thereby, one can define path-conditioned density evolutions and formulate and solve maximum entropy problems in a somewhat classical-like fashion. However, this approach does not seem to work when both initial and final densities are given, specifically because it is not clear what notion of relative entropy one should use to compare the two Markovian evolutions associated to $\mathcal{E}$ and $\mathcal{F}$. Thus, herein we formulate the quantum Schrödinger bridge problem as that of seeking a suitable multiplicative functional transformation of the prior Kraus evolution so as to meet the marginal conditions. This is explained next.

Given the sequence of TPCP maps $\{\mathcal{E}_t^\dagger$, for $0 \le t \le T-1\}$ and the two marginals ρ_0, ρ_T, we seek a sequence of invertible matrices $\chi_0, \dots, \chi_T$ such that for $t \in \{0, \dots, T-1\}$,

$$\mathcal{F}_t^\dagger(\cdot) = \chi_{t+1}\left(\mathcal{E}_t^\dagger(\chi_t^{-1}(\cdot)\chi_t^{-\dagger})\right)\chi_{t+1}^\dagger \tag{37}$$

is a Kraus map (i.e., totally positive and trace preserving) and the evolution

$$\mathcal{F}_{0:T}^\dagger(\cdot) = \mathcal{F}_{T-1}^\dagger \circ \cdots \circ \mathcal{F}_1^\dagger \circ \mathcal{F}_0^\dagger(\cdot) \tag{38}$$

$$= \chi_T\left(\mathcal{E}_{0:T}^\dagger(\chi_0^{-1}(\cdot)\chi_0^{-\dagger})\right)\chi_T^\dagger \tag{39}$$

is consistent with the given marginals (i.e., (36) holds). Conditions (37) and (39), represent the quantum analogue of a multiplicative functional transformation.

Very much as in the classical case, the solution of the Schrödinger bridge in the multi-step case reduces to solving the one-step bridge problem: given the triple $(\mathcal{E}_{0:T}^\dagger, \rho_0, \rho_T)$, determine invertible matrices χ_0, χ_T such that (39) is Kraus map and satisfies

$$\mathcal{F}_{0:T}^\dagger(\rho_0) = \rho_T. \tag{40}$$

The *unital* property of $\mathcal{F}$, namely, the fact that $\mathcal{F}_{0:T}(I) = I$, when expressed in terms of $\mathcal{E}$ and the factors in (39), implies that

$$\mathcal{E}_{0:T}(\phi_T) = \phi_0 \tag{41}$$

for the Hermitian, positive definite matrices

$$\phi_0 = \chi_0^\dagger\chi_0, \quad \phi_T = \chi_T^\dagger\chi_T. \tag{42}$$

Likewise, in view of (37) and the requirement that $\mathcal{F}_t^\dagger$ is trace preserving, we set

$$\mathcal{E}_t(\phi_{t+1}) = \phi_t, \tag{43}$$

for $t \in \{T-1, \dots, 0\}$ and factor

$$\phi_t = \chi_t^\dagger\chi_t. \tag{44}$$

Conditions (41) and (44) indicate that the ϕ_t's are *space-time harmonic* matrices with respect to $\mathcal{E}_{0:T}^\dagger$ and $\mathcal{E}_t^\dagger$, respectively. On the other hand,

$$\mathcal{E}_{0:T}^\dagger(\underbrace{\chi_0^{-1}(\rho_0)\chi_0^{-\dagger}}_{\hat{\phi}_0}) = \underbrace{\chi_T^{-1}(\rho_T)\chi_T^{-\dagger}}_{\hat{\phi}_T}, \tag{45}$$

where the matrices $\hat{\phi}_t$ in (45) can be thought of as unnormalized density matrices in a dual Heisenberg picture.

The concept of space-time harmonic functions for quantum channels was introduced and studied in Refs. 25 and 33. Thus, following Refs. 25 and 33, when two Kraus maps are related as in ((39), (41), and (42)), we say that $\mathcal{F}^\dagger$ is obtained from $\mathcal{E}^\dagger$ through a *multiplicative functional transformation induced by a space-time harmonic function*. In Ref. 25, it was shown that the solution of the Schrödinger problem, where only the initial or final density is assigned, is indeed given by a multiplicative functional transformation of the prior providing the first such example of a non-commutative counterpart of the classical results. A variational characterisation of the quantum bridge as a critical point of a suitable entropic functional for the present general setting is unknown.

C. Doubly stochastic Kraus maps

The results we present next are for the special case where the two marginal densities are uniform, i.e., when both

$$\rho_0 = \frac{1}{n}I \text{ and } \rho_T = \frac{1}{n}I.$$

This special case is interesting in its own right as, already for the classical case, it represents a well known basic result in the statistics literature (Sinkhorn's theorem).[31,32] Sinkhorn's result states that

for any stochastic matrix $[\pi_{x_0,x_T}]_{x_0,x_T=1}^N$ *having all entries strictly positive*, there exists a unique multiplicative functional transformation

$$\pi^o_{x_0,x_T} = \pi_{x_0,x_T}\frac{\varphi(T,x_T)}{\varphi(0,x_0)}$$

so that π^o is doubly stochastic, that is, it has non-negative elements and satisfies

$$\sum_{x_0}\pi^o_{x_0,x_T} = \sum_{x_T}\pi^o_{x_0,x_T} = 1.$$

In view of Eq. (32), in this classical setting, Sinkhorn's result is a direct corollary of Theorem 3.

The property of a stochastic matrix to have all entries strictly positive corresponds to the positivity improving property (34) of a Kraus map. We proceed to derive the quantum counterpart of Sinkhorn's result.

Theorem 5. *Given a positivity improving Kraus map $\mathcal{E}^\dagger_{0:T}$, i.e., satisfying (34), there exists a pair of observables ϕ_0, $\phi_T \in \mathfrak{H}_{++}$ unique up to multiplication by a positive constant and related as in (41) such that, for any factorization,*

$$\phi_0 = \chi_0^\dagger\chi_0$$
$$\phi_T = \chi_T^\dagger\chi_T,$$

the map $\mathcal{F}^\dagger_{0:T} : \mathfrak{D} \to \mathfrak{D}$ defined by

$$\mathcal{F}^\dagger(\cdot) := \chi_T\left(\mathcal{E}^\dagger_{0:T}(\chi_0^{-1}(\cdot)\chi_0^{-\dagger})\right)\chi_T^\dagger \tag{46}$$

is a doubly stochastic Kraus map, in that $\mathcal{F}(I) = I$ as well as $\mathcal{F}^\dagger(I) = I$.

Theorem 5 follows immediately from the following result that we establish first (it was recently pointed out to us that the result in Theorem 5 has been obtained by Gurvits Ref. 15, Theorem 4.7 using a completely different argument which is based on the existence of certain natural potentials (*locally scalable functionals*) that he constructs for this purpose).

Theorem 6. *Given a Kraus map $\mathcal{E}^\dagger_{0:T}$ satisfying (34), there exist observables ϕ_0, ϕ_T in $\mathfrak{H}_{++}$ unique up to multiplication by a positive constant such that*

$$\mathcal{E}_{0:T}(\phi_T) = \phi_0,$$
$$\mathcal{E}^\dagger_{0:T}(\phi_0^{-1}) = \phi_T^{-1}.$$

Proof of Theorem 6. The claim in Theorem 6 amounts to the existence of a unique fixed point for the following circular diagram of maps:

$$\begin{array}{ccc} \hat\phi_0 & \xrightarrow{\mathcal{E}^\dagger_{0,T}} & \hat\phi_T \\ \hat\phi_0 = \phi_0^{-1}\ \uparrow & & \downarrow\ \phi_T = \hat\phi_T^{-1}, \\ \phi_0 & \xleftarrow{\mathcal{E}_{0,T}} & \phi_T \end{array} \tag{47}$$

Thus, it suffices to show that the composition map

$$C\ :\ (\hat\phi_0)_{\text{starting}} \xrightarrow{\mathcal{E}^\dagger_{0,T}} \hat\phi_T \xrightarrow{(\cdot)^{-1}} \phi_T \xrightarrow{\mathcal{E}_{0,T}} \phi_0 \xrightarrow{(\cdot)^{-1}} (\hat\phi_0)_{\text{next}} \tag{48}$$

from $\mathfrak{H}_{++} \to \mathfrak{H}_{++}$ is contractive in the Hilbert metric. It should be noted that "points" here are defined up to a scaling factor, thus, they in essence represent rays.

Once again, we use Birkhoff's theorem to determine the contraction coefficient. Here, we are working on the positive cone $\mathfrak{H}_+$. This has a nonempty interior $\mathfrak{H}_{++}$ and the partial order defined by non-negative definiteness. Accordingly,

$$\begin{aligned} M(X,Y) &= \inf\{\lambda \mid X \le \lambda Y\} \\ &= \max\{\mathrm{eig}(Y^{-1/2}XY^{-1/2})\} \\ &= \max\{\mathrm{eig}(XY^{-1})\}, \\ m(X,Y) &= \sup\{\lambda \mid \lambda y \le x\} \\ &= \min\{\mathrm{eig}(XY^{-1})\}, \end{aligned}$$

where $\mathrm{eig}(\cdot)$ denotes the "eigenvalues of," and in this case, the Hilbert metric is

$$\begin{aligned} d_H(X,Y) &= d_H(Y^{-1/2}XY^{-1/2}, I) \\ &= \log(\kappa(XY^{-1})), \end{aligned}$$

where $\kappa(\cdot)$ is the "conditioning number" of a $Z \in \mathfrak{M}$,

$$\kappa(Z) = \frac{\max\{\mathrm{eig}(Z)\}}{\min\{\mathrm{eig}(Z)\}}.$$

From the Birkhoff-Bushell theorem, we have that both $\|\mathcal{E}\|_H \le 1$ and $\|\mathcal{E}^\dagger\|_H \le 1$ (from linearity together with monotonicity). Furthermore, the diameter of the range of $\mathcal{E}$,

$$\Delta(\mathcal{E}) = \sup\{d_H(\mathcal{E}(X), \mathcal{E}(Y)) \mid X, Y \in \mathfrak{H}_{++}\},$$

is finite. To see this, first note that

$$\Delta(\mathcal{E}) \le 2\sup\{d_H(\mathcal{E}(X), I) \mid X \in \mathfrak{H}_{++}\},$$

utilizing the metric property and the fact that $\mathcal{E}(I) = I$, and then, note that $d_H(\mathcal{E}(X), I)$ is invariant under scaling of X by a scalar. Therefore, we can restrict our attention to $X \in \mathfrak{D}$ instead, and since $\mathcal{E}(X) > 0$ for all $X \in \mathfrak{D}$, by compactness of $\mathfrak{D}$,

$$\Delta(\mathcal{E}) \le 2\max\{d_H(\mathcal{E}(X), I) \mid X \in \mathfrak{D}\} < \infty.$$

Therefore (see Theorem 1),

$$\|\mathcal{E}\|_H = \tanh\left(\frac{1}{4}\Delta(\mathcal{E})\right) < 1.$$

Finally, we note that the induced Hilbert-gain of the inversion

$$Z \mapsto Z^{-1}$$

is 1 since

$$d_H(X,Y) = d_H(X^{-1}, Y^{-1}).$$

We conclude that with C as in (48), $\|C\|_H < 1$. Again, contractiveness in the Hilbert metric, there exists a unique fixed ray, i.e.,

$$\hat{\phi}_0 = \lambda C(\hat{\phi}_0),$$

for some $\lambda > 0$, and corresponding

$$\begin{aligned} \hat{\phi}_T &= \mathcal{E}^\dagger_{0:T}(\hat{\phi}_0), \\ \phi_0 &= \mathcal{E}_{0:T}(\phi_T), \end{aligned}$$

while

$$\begin{aligned} \phi_T &= \hat{\phi}_T^{-1} \\ \lambda\hat{\phi}_0 &= \phi_0^{-1}. \end{aligned}$$

Since,

$$\begin{aligned} n &= \lambda \operatorname{trace}(\hat{\phi}_0 \phi_0) \\ &= \lambda \operatorname{trace}(\hat{\phi}_0 \mathcal{E}_{0:T}(\phi_T)) \\ &= \lambda \operatorname{trace}(\mathcal{E}^\dagger_{0:T}(\hat{\phi}_0)\phi_T) \\ &= \lambda \operatorname{trace}(\hat{\phi}_T \phi_T) \\ &= \lambda n, \end{aligned}$$

and therefore, $\lambda = 1$. This completes the proof of the theorem. □

Proof of Theorem 5. This is truly a corollary to Theorem 6. Since ϕ_0, ϕ_T in the proof of Theorem 6 are in $\mathfrak{H}_{++}$, we take any factorization

$$\begin{aligned} \phi_0 &= \chi_0^\dagger \chi_0 \\ \phi_T &= \chi_T^\dagger \chi_T. \end{aligned}$$

Then, with $\mathcal{F}^\dagger$ as in (46),

$$\begin{aligned} \mathcal{F}^\dagger_{0:T}(I) &= \chi_T \left[\mathcal{E}^\dagger_{0:T}\left(\chi_0^{-1}(I)\chi_0^{-\dagger}\right)\right] \chi_T^\dagger \\ &= \chi_T \left[\mathcal{E}^\dagger_{0:T}(\phi_0^{-1})\right] \chi_T^\dagger \\ &= \chi_T \left(\phi_T^{-1}\right) \chi_T^\dagger \\ &= I. \end{aligned}$$

Likewise,

$$\begin{aligned} \mathcal{F}_{0:T}(I) &= \chi_0^{-\dagger} \left[\mathcal{E}_{0:T}\left(\chi_T^\dagger(I)\chi_T\right)\right] \chi_0^{-1} \\ &= \chi_0^{-\dagger} \left[\mathcal{E}_{0:T}(\phi_T)\right] \chi_0^{-1} \\ &= \chi_0^{-\dagger} (\phi_0) \chi_0^{-1} \\ &= I. \end{aligned}$$

Thus, $\mathcal{F}^\dagger$ is doubly stochastic as claimed. □

While in the classical case (Sinkhorn's theorem), there is a unique doubly stochastic map obtained from π_{x_0,x_T} via a multiplicative functional transformation, in the quantum case, this is clearly false. Uniqueness in Theorem 5 is claimed for the observables ϕ_0, ϕ_T. Hence, $\mathcal{F}$ is unique modulo corresponding unitary factors; obviously, if $\mathcal{F}^\dagger$ is a Kraus so that $\mathcal{F}(I) = I$ and $\mathcal{F}^\dagger(I) = I$, then for any unitary matrices U_0, U_T,

$$U_T \mathcal{F}^\dagger(U_0(\cdot)U_0^\dagger)U_T^\dagger \tag{49}$$

is also a doubly stochastic Kraus map. It is easy to see that the totality of doubly stochastic Kraus maps that relate to $\mathcal{E}^\dagger$ via a multiplicative transformation are of this form.

D. The quantum bridge for general marginals and a conjecture

Consider now the situation of Sec. IV C with general initial and final density matrices ρ_0 and ρ_T. We are, namely, seeking a quantum bridge, as defined in Sec. IV B, for the triple $(\mathcal{E}^\dagger_{0:T}, \rho_0, \rho_T)$. Theorem 5 admits the following generalization.

Theorem 7. *Given a Kraus map $\mathcal{E}^\dagger_{0:T}$ and two density matrices ρ_0 and ρ_T, suppose there exist observables ϕ_0, ϕ_T, $\hat{\phi}_0$, $\hat{\phi}_T \in \mathfrak{H}_{++}$ solving the Schrödinger system,*

$$\mathcal{E}_{0:T}(\phi_T) = \phi_0, \tag{50}$$

$$\mathcal{E}^\dagger_{0:T}(\hat{\phi}_0) = \hat{\phi}_T, \tag{51}$$

$$\rho_0 = \chi_0 \hat{\phi}_0 \chi_0^\dagger, \tag{52}$$

$$\rho_T = \chi_T \hat{\phi}_T \chi_T^\dagger. \tag{53}$$

Then, for any factorization,

$$\phi_0 = \chi_0^\dagger \chi_0$$
$$\phi_T = \chi_T^\dagger \chi_T,$$

the map $\mathcal{F}_{0:T}^\dagger : \mathfrak{D} \to \mathfrak{D}$ *defined by*

$$\mathcal{F}^\dagger(\cdot) := \chi_T \left(\mathcal{E}_{0:T}^\dagger (\chi_0^{-1}(\cdot)\chi_0^{-\dagger}) \right) \chi_T^\dagger \tag{54}$$

is a quantum bridge for $(\mathcal{E}_{0:T}^\dagger, \rho_0, \rho_T)$*, namely,* $\mathcal{F}(I) = I$ *and* $\mathcal{F}^\dagger(\rho_0) = \rho_T$.

Proof:

$$\mathcal{F}_{0:T}(I) = \chi_0^{-\dagger} \left[\mathcal{E}_{0:T} \left(\chi_T^\dagger(I)\chi_T \right) \right] \chi_0^{-1} = \chi_0^{-\dagger} \left(\mathcal{E}_{0:T}(\phi_T) \right) \chi_0^{-1} = \chi_0^{-\dagger}(\phi_0)\, \chi_0^{-1} = I.$$

Moreover,

$$\mathcal{F}_{0:T}^\dagger(\rho_0) = \chi_T \left[\mathcal{E}_{0:T}^\dagger \left(\chi_0^{-1} \rho_0 \chi_0^{-\dagger} \right) \right] \chi_T^\dagger = \chi_T \left[\mathcal{E}_{0:T}^\dagger(\hat{\phi}_0) \right] \chi_T^\dagger = \chi_T \left(\hat{\phi}_T \right) \chi_T^\dagger = \rho_T.$$

□

The case of initial and final pure states is worthwhile writing out.

Corollary 1. Given a positivity improving Kraus map $\mathcal{E}_{0:T}^\dagger$ *and two pure states*

$$\rho_0 = v_0 v_0^\dagger \text{ and } \rho_T = v_T v_T^\dagger$$

(i.e., v_0, v_T *are unit norm vectors), define*

$$\phi_0 := \mathcal{E}(v_T v_T^\dagger),$$
$$\phi_T := v_T v_T^\dagger$$

and

$$\mathcal{F}^\dagger(\cdot) := \phi_T^{1/2} \mathcal{E}^\dagger(\phi_0^{-1/2}(\cdot)\phi_0^{-1/2})\phi_T^{1/2}$$

(where, clearly, $\phi_T^{1/2} = \phi_T = v_T v_T^\dagger$*). Then,* $\mathcal{F}^\dagger$ *is TPTP and satisfies the marginal conditions*

$$\rho_T = \mathcal{F}^\dagger(\rho_0).$$

Proof: We readily verify that

$$\begin{aligned} \mathcal{F}(I) &= \phi_0^{-1/2} \mathcal{E}(\phi_T) \phi_0^{-1/2} \\ &= \phi_0^{-1/2} \mathcal{E}(v_T v_T^\dagger) \phi_0^{-1/2} \\ &= I. \end{aligned}$$

Next, we observe that

$$\begin{aligned} \mathcal{F}^\dagger(\rho_0) &= \phi_T^{1/2} \mathcal{E}^\dagger(\phi_0^{-1/2}(v_0 v_0^\dagger)\phi_0^{-1/2})\phi_T^{1/2} \\ &= v_T v_T^\dagger \mathcal{E}^\dagger(\phi_0^{-1/2}(v_0 v_0^\dagger)\phi_0^{-1/2}) v_T v_T^\dagger \\ &= v_T v_T^\dagger \\ &= \rho_T. \end{aligned}$$

To see this, consider a representation $\mathcal{E}^\dagger(\cdot) = \sum_i E_i(\cdot)E_i^\dagger$ and note that

$$\begin{aligned} v_T^\dagger \mathcal{E}^\dagger(\phi_0^{-1/2}(v_0 v_0^\dagger)\phi_0^{-1/2})v_T &= \sum_i \left(v_T^\dagger E_i \phi_0^{-1/2} v_0\right)^2 \\ &= v_0^\dagger \phi_0^{-1/2} \mathcal{E}(v_T v_T^\dagger)\phi_0^{-1/2} v_0 \\ &= v_0^\dagger I v_0 \\ &= 1. \end{aligned}$$

□

Thus, precisely as in the classical case, the key challenge is to establish existence and uniqueness for Schrödinger system (50)-(53) and that this is an attractive fixed point of a suitable iterative map. At present, proving the natural generalization of Theorem 6 appears nontrivial. Thus, below, we give the relevant statement as a conjecture since its proof remains elusive.

Conjecture 1. Given a positivity-improving Kraus map $\mathcal{E}_{0:T}^\dagger$, i.e., a Kraus map satisfying (34), and given two density matrices ρ_0 and ρ_T, there exist observables ϕ_0, ϕ_T, $\hat{\phi}_0$, $\hat{\phi}_T$ in $\mathfrak{H}_{++}$ such that

$$\begin{aligned} \mathcal{E}_{0:T}(\phi_T) &= \phi_0, \\ \mathcal{E}_{0:T}^\dagger(\hat{\phi}_0) &= \hat{\phi}_T, \end{aligned}$$

with

$$\rho_0 = \chi_0 \hat{\phi}_0 \chi_0^\dagger, \tag{55a}$$

$$\rho_T = \chi_T \hat{\phi}_T \chi_T^\dagger \tag{55b}$$

and

$$\phi_0 = \chi_0^\dagger \chi_0, \tag{55c}$$

$$\phi_T = \chi_T^\dagger \chi_T \tag{55d}$$

which can be obtained as a fixed point of a suitable map. *In particular, χ_0, χ_T can be taken to be Hermitian, i.e., for $i \in \{0,T\}$, $\chi_i = (\phi_i)^{1/2}$ is the Hermitian square root of ϕ_i, and in this case, the solution to the Schrödinger system can be obtained as a fixed point of a map specified in (56) explained below.*

Equations (55a) and (55c) together represent a non-commutative analogue of the relationship between $\hat{\varphi}(0,x_0), \mathbf{p}_0(x_0), \varphi(0,x_0)$ in

$$\hat{\varphi}(0,x_0) = \frac{\mathbf{p}_0(x_0)}{\varphi(0,x_0)},$$

while Eqs. (55b) and (55d) represent the analogue of

$$\hat{\varphi}(T,x_T) = \frac{\mathbf{p}_T(x_T)}{\varphi(T,x_T)}.$$

By taking $\chi_0 = (\phi_0)^{1/2}$, i.e., the Hermitian square root, clearly,

$$\hat{\phi}_0 = (\phi_0)^{1/2} \rho_0 (\phi_0)^{1/2}.$$

On the other, taking χ_T to be the Hermitian square root of ϕ_T and solving for ϕ_T in terms of $\hat{\phi}_T$ and ρ_T using (55b)-(55d) gives

$$\hat{\phi}_T \mapsto \phi_T = \left(\rho_T^{1/2}\left(\rho_T^{-1/2}\hat{\phi}^{-1}\rho_T^{-1/2}\right)^{1/2}\rho_T^{1/2}\right)^2.$$

Thus, the conjecture claims the validity of the following correspondence:

$$\rho_0 \overset{\chi_0^{-1}(\cdot)\chi_0^{-\dagger}}{\longrightarrow} \hat{\phi}_0 \overset{\mathcal{E}_{0:T}^{\dagger}}{\longrightarrow} \hat{\phi}_T \overset{\chi_T(\cdot)\chi_T^{\dagger}}{\longrightarrow} \rho_T,$$

$$I \overset{\chi_0^{-\dagger}(\cdot)\chi_0^{-1}}{\longleftarrow} \phi_0 = \chi_0^{\dagger}\chi_0 \overset{\mathcal{E}_{0:T}}{\longleftarrow} \phi_T = \chi_T^{\dagger}\chi_T \overset{\chi_T^{\dagger}(\cdot)\chi_T}{\longleftarrow} I,$$

and therefore, that the Kraus map

$$\mathcal{F}_{0:T}^{\dagger}(\cdot) \coloneqq \chi_T \left(\mathcal{E}_{0:T}^{\dagger}(\chi_0^{-1}(\cdot)\chi_0^{-\dagger})\right) \chi_T^{\dagger}$$

solves the one-step quantum Schödinger bridge problem for general marginal density matrices. Extensive simulations have convinced the authors that the composition of maps

$$\hat{\phi}_0 \overset{\mathcal{E}_{0:T}^{\dagger}}{\longrightarrow} \hat{\phi}_T \overset{D_T}{\longrightarrow} \phi_T \overset{\mathcal{E}_{0:T}}{\longrightarrow} \phi_0 \overset{\hat{D}_0}{\longrightarrow} (\hat{\phi}_0)_{\text{next}}, \tag{56}$$

where

$$D_T : \hat{\phi}_T \mapsto \phi_T = \left(\rho_T^{1/2}\left(\rho_T^{-1/2}\hat{\phi}^{-1}\rho_T^{-1/2}\right)^{1/2}\rho_T^{1/2}\right)^2,$$
$$\hat{D}_0 : \phi_0 \mapsto \hat{\phi}_0 = (\phi_0)^{1/2}\rho(\phi_0)^{1/2}$$

have an attractive fixed point. In this, χ_i for $i \in \{0,T\}$ are taken Hermitian for specificity. Simulation shows that the iteration converges to a fixed point for a variety of other normalizations for the factors χ_i of ϕ_i, as well as when the boundary conditions are replaced by $\rho_0 = (\hat{\phi}_0)^{1/2}\phi_0(\hat{\phi}_0)^{1/2}$ and $\rho_T = (\hat{\phi}_T)^{1/2}\phi_T(\hat{\phi}_T)^{1/2}$ and the iteration is modified accordingly.

In the above, $\hat{D}_0$ and D_T are no longer isometries in the Hilbert metric as their commutative analogues and, as a consequence, it is not clear how to generalize the earlier argument in a direct manner. However, as pointed out to us by Idel Martin, the map in (56), suitably normalized by the trace of the image, sends positive definite trace-1 matrices to the same and is continuous, and therefore, existence of a solution to the Schrödinger system is guaranteed by Brouwer's fixed point theorem. Thus, the essence of the conjecture is the uniqueness and attractiveness to a fixed point.

For a bridge with many intermediary time steps, very much as in the commutative classical case, the solution of the one-step bridge via multiplicative functional transformation permits solving the general bridge in a similar manner. More specifically, starting from $\hat{\phi}_0, \phi_T$ that correspond to a fixed point of (56), define for $i \in \{1,\dots,T\}$,

$$\hat{\phi}_i \coloneqq \mathcal{E}_{i-1}^{\dagger}\hat{\phi}_{i-1},$$
$$\phi_{i-1} \coloneqq \mathcal{E}_{i-1}\phi_i,$$
$$\rho_i \coloneqq (\phi_i)^{1/2}\hat{\phi}_i(\phi_i)^{1/2}.$$

The sought sequence of Kraus maps is

$$\mathcal{F}_{i+1}^{\dagger}(\cdot) = (\phi_{i+1})^{1/2}\left(\mathcal{E}_i^{\dagger}((\phi_i)^{-1/2}(\cdot)(\phi_i)^{-1/2})\right)(\phi_{i+1})^{1/2},$$

since, clearly,

$$\mathcal{F}_{0:T}^{\dagger}(\cdot) = (\phi_T)^{1/2}\left(\mathcal{E}_{0:T}^{\dagger}((\phi_0)^{-1/2}(\cdot)(\phi_0)^{-1/2})\right)(\phi_T)^{1/2}$$

and, assuming the validity of the conjecture, satisfies

$$\begin{aligned}\mathcal{F}_{0:T}^{\dagger}(\rho_0) &= (\phi_T)^{1/2}\left(\mathcal{E}_{0:T}^{\dagger}(\hat{\phi}_0)\right)(\phi_T)^{1/2}\\ &= (\phi_T)^{1/2}\hat{\phi}_T(\phi_T)^{1/2}\\ &= \rho_T.\end{aligned}$$

V. EXAMPLES OF DOUBLY STOCHASTIC KRAUS MAPS

Perhaps the simplest nontrivial example of a (self-adjoint) positivity improving doubly stochastic Kraus map is

$$\mathcal{E}^\dagger(\cdot) = E_1(\cdot)E_1^\dagger + E_2(\cdot)E_2^\dagger + E_3(\cdot)E_3^\dagger, \tag{57}$$

with

$$E_1 = \begin{bmatrix} \sqrt{\frac{1}{2}} & 0 \\ 0 & 0 \end{bmatrix}, \; E_2 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\frac{1}{2}} \end{bmatrix}, \; E_3 = \begin{bmatrix} 0 & \sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} & 0 \end{bmatrix}.$$

A second example is

$$E_1 = \begin{bmatrix} \sqrt{\frac{2}{3}} & 0 \\ 0 & 0 \end{bmatrix}, \; E_2 = \begin{bmatrix} \sqrt{\frac{1}{24}} & -\sqrt{\frac{1}{8}} \\ -\sqrt{\frac{1}{8}} & \sqrt{\frac{3}{8}} \end{bmatrix}, \; E_3 = \begin{bmatrix} \sqrt{\frac{1}{24}} & \sqrt{\frac{1}{8}} \\ \sqrt{\frac{1}{8}} & \sqrt{\frac{3}{8}} \end{bmatrix},$$

where all coefficient matrices are again symmetric but they all now of rank one.

In general, neither the symmetry (Hermitian-ness) of the coefficients nor any constraint on the rank is essential. The following example is constructed numerically. For this, we start with

$$E_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} M^{-1/2},$$
$$E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} M^{-1/2},$$
$$E_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} M^{-1/2},$$

where

$$M = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$$

so that $E_1'E_1 + E_2'E_2 + E_3'E_3 = I$. It can be shown that the corresponding Kraus map $\mathcal{E}^\dagger$ is positivity improving, i.e., it satisfies (34). We then compute the fixed point of (48), and this is

$$\phi_0 = \begin{bmatrix} 1.1448 & -0.1350 \\ -0.1350 & 0.8749 \end{bmatrix},$$
$$\phi_1 = \begin{bmatrix} 0.8411 & -0.2362 \\ -0.2362 & 1.3134 \end{bmatrix},$$
$$\hat{\phi}_0 = \begin{bmatrix} 0.8897 & 0.1372 \\ 0.1372 & 1.1642 \end{bmatrix},$$
$$\hat{\phi}_1 = \begin{bmatrix} 1.2521 & 0.2251 \\ 0.2251 & 0.8018 \end{bmatrix}.$$

These space-time harmonics give rise to coefficients

$$F_1 = \begin{bmatrix} 0.5690 & 0.4411 \\ -0.0720 & -0.0558 \end{bmatrix},$$
$$F_2 = \begin{bmatrix} -0.0558 & 0.4411 \\ -0.0720 & 0.5690 \end{bmatrix},$$
$$F_3 = \begin{bmatrix} -0.1441 & 0.5131 \\ 0.8013 & -0.1441 \end{bmatrix}$$

for a corresponding doubly stochastic Kraus map.

When ρ_0 and/or ρ_T are in general different from the identity, extensive numerical experimentation suggests the validity of Conjecture 1 which, in conjunction with Theorem 7, provides solutions of the quantum Schrödinger bridge problem. In particular, for our first example (57) and nonuniform marginals,

$$\rho_0 = \begin{bmatrix} 1/4 & 0 \\ 0 & 3/4 \end{bmatrix} \text{ and } \rho_1 = \begin{bmatrix} 2/3 & 0 \\ 0 & 1/3 \end{bmatrix}$$

by iterating (56), we obtain

$$\phi_0 = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix},$$
$$\phi_1 = \begin{bmatrix} 2/3 & 0 \\ 0 & 1/3 \end{bmatrix},$$
$$\hat{\phi}_0 = \begin{bmatrix} 1/2 & 0 \\ 0 & 3/2 \end{bmatrix},$$
$$\hat{\phi}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and the Kraus map with coefficients

$$F_1 = \begin{bmatrix} \sqrt{2/3} & 0 \\ 0 & 0 \end{bmatrix},$$
$$F_2 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{1/3} \end{bmatrix},$$
$$F_3 = \begin{bmatrix} 0 & \sqrt{2/3} \\ \sqrt{1/3} & 0 \end{bmatrix}.$$

This Kraus map is no longer Hermitian, it is of the form (46) and, as can be readily verified, it satisfies the required condition $\mathcal{F}^{\dagger}(\rho_0) = \rho_1$. Software for numerical experimentation can be found at http://www.ece.umn.edu/~georgiou/papers/schrodinger_bridge/.

VI. CONCLUDING REMARKS

In this paper, we introduced a new approach to studying Schrödinger's systems. In particular, we establish new proofs for existence and uniqueness of solutions. In contrast to earlier treatments, our approach provides a direct computational procedure for obtaining the space-time harmonic function and the corresponding solution of the Schrödinger bridge in finite-dimensions. Space-time harmonics are obtained as fixed points of a certain map. Convergence is established in a suitable *projective geometry* as convergence of *rays* using the Hilbert metric. That is, in the classical case of discrete random vectors that is treated herein, our approach provides a direct new proof of

existence and uniqueness by a contraction mapping principle. Since the approach also provides a computational scheme, it appears to have considerable potential for applications. Indeed, models for stochastic evolution, which include Markov chains, are ubiquitous. The bridge evolution may be viewed as a controlled *steering* problem–a facet that will be thoroughly explored elsewhere. In the case of quantum channels, the solution of an analogous Schrödinger system corresponds to a steering between two given density matrices. We prove convergence of an iterative algorithm in a corresponding projective metric in the case of uniform marginals, i.e., identity matrices, thereby establishing existence of doubly stochastic Kraus maps that can be derived from a given reference map via multiplicative functional transformations. Extensive simulations have convinced the authors of a more general result regarding the general quantum bridge problem which, however, is stated as a conjecture since at present a proof is not available.

ACKNOWLEDGMENTS

The authors are very grateful to Yongxin Chen for a correction in the proof of Theorems 3 and 6 and to Idel Martin for pointing out to us[15] as well as connections to Ref. 22. The research of T. G. was supported in part by grants from the NSF and AFOSR. 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.

[1] Beghi, A., "On the relative entropy of discrete-time Markov processes with given end-point densities," IEEE Trans. Inf. Theory **42**, 1529–1535 (1996).
[2] Beurling, A., "An automorphism of product measures," Ann. Math. **72**, 189–200 (1960).
[3] Birkhoff, G., "Extensions of Jentzsch's theorem," Trans. Am. Math. Soc. **85**, 219–227 (1957).
[4] Birkhoff, G., "Uniformly semi-primitive multiplicative processes," Trans. Am. Math. Soc. **104**, 37–51 (1962).
[5] Bonnabel, S., Astolfi, A., and Sepulchre, R., "Contraction and observer design on cones," in *Decision and Control and European Control Conference (CDC-ECC)* (IEEE, 2011), pp. 12–15.
[6] Bushell, P., "On the projective contraction ratio for positive linear mappings," J. London Math. Soc. **2**, 256–258 (1973).
[7] Bushell, P. J., "Hilbert's metric and positive contraction mappings in a Banach space," Arch. Ration. Mech. Anal. **52**, 330–338 (1973).
[8] Cover, T. M. and Thomas, J. A., *Elements of Information Theory*, 2nd ed. (Wiley, 2006).
[9] Csiszár, I., "I-divergence geometry of probability distributions and minimization problems," Ann. Probab. **3**, 146–158 (1975).
[10] Deming, W. E. and Stephan, F. F., "On a least squares adjustment of a sampled frequency table when the expected marginal totals are known," Ann. Math. Stat. **11**, 427–444 (1940).
[11] Fienberg, S., "An iterative procedure for estimation in contingency tables," Ann. Math. Stat. **41**, 907–917 (1970).
[12] Föllmer, H., "Random fields and diffusion processes," *École d'Ètè de Probabilitès de Saint-Flour XV-XVII*, Lecture Notes in Mathematics Vol. 1362 (Springer, 1988), pp. 102–203.
[13] Fortet, R., "Résolution d'un système d'equations de M. Schrödinger," J. Math. Pure Appl. **IX**, 83–105 (1940).
[14] Gaubert, S. and Qu, Z., "Checking the strict positivity of Kraus maps is NP-hard," preprint arXiv:1402.1429 (2014).
[15] Gurvits, L., "Classical complexity and quantum entanglement," J. Comput. Syst. Sci. **69**, 448–484 (2004).
[16] Hilbert, D., "Über die gerade linie als kürzeste verbindung zweier punkte," Math. Ann. **46**, 91–96 (1895).
[17] Hillar, C. and Lim, L.-H., "Most tensor problems are NP-hard," J. ACM **60**, Article 45 (2013); preprint arXiv:0911.1393 (2009).
[18] Holevo, A. S., *Statistical Structure of Quantum Theory* (Springer, 2001).
[19] Ireland, C. T. and Kullback, S., "Contingency tables with given marginals," Biometrika **55**, 179–188 (1968).
[20] Jamison, B., "Reciprocal processes," Z. Wahrscheinlichkeitstheorie verw. Geb. **30**, 65–86 (1974).
[21] Jamison, B., "The Markov processes of Schrödinger," Z. Wahrscheinlichkeitstheorie verw. Geb. **32**, 323–331 (1975).
[22] Lemmens, B. and Nussbaum, R., *Nonlinear Perron-Frobenius Theory* (Cambridge University Press, 2012), Vol. 189.
[23] Lemmens, B. and Nussbaum, R., "Birkhoff's version of Hilbert's metric and its applications in analysis," in *Handbook of Hilbert geometry*, edited by A. Papadopoulos and M. Troyanov (European Math. Soc. Publ. House, Zuerich, 2014) (to appear); preprint arXiv:1304.7921v1 (2013).
[24] Nielsen, M. A. and Chuang, I. L., *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
[25] Pavon, M. and Ticozzi, F., "Discrete-time classical and quantum Markovian evolutions: Maximum entropy problems on path space," J. Math. Phys. **51**, 042104 (2010).
[26] Petz, D., *Quantum Information Theory and Quantum Statistics* (Springer, 2008).
[27] Reeb, D., Kastoryano, M. J., and Wolf, M. M., "Hilbert's projective metric in quantum information theory," J. Math. Phys. **52**, 082201 (2011).
[28] Sanov, I. S., "On the probability of large deviations of random magnitudes," Mat. Sb. **42**(84), 11–44 (1957) (in Russian).
[29] Schrödinger, E., "Merkwürdige Analogien zur Quantenmechanik, die mir sehr des Hindenkens wert erscheinen" (remarkable analogies to quantum mechanics which appear to me very worth of reflection). Note that Schrödinger never accepted the so-called orthodox theory of measurement in quantum mechanics and was looking for a more classical probabilistic reformulation.

[30] Sepulchre, R., Sarlette, A., and Rouchon, P., "Consensus in non-commutative spaces," *49th IEEE Conference on Decision and Control (CDC)* (IEEE, 2010), pp. 6596–6601.
[31] Sinkhorn, R., "A relationship between arbitrary positive matrices and doubly stochastic matrices," Ann. Math. Stat. **35**, 876–879 (1964).
[32] Sinkhorn, R., "Diagonal equivalence to matrices with prescribed row and column sums," Proc.Am. Math. Soc. **45**, 195–198 (1974).
[33] Ticozzi, F. and Pavon, M., "On time-reversal and space-time harmonic processes for Markovian quantum channels," Quantum Inf. Process. **9**, 551–574 (2010).
[34] Tsitsiklis, J., Bertsekas, D., and Athans, M., "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," IEEE Trans. Autom. Control **31**, 803–812 (1986).
[35] Wakolbinger, "Schrödinger bridges from 1931 to 1991," in *4th Latin American Congress in Probability and Mathematical Statistics, Mexico City 1990*, Contribuciones en probabilidad y estadistica matematica Vol. 3, (1992), pp. 61–79.