Schrödinger bridges for discrete-time, classical and quantum Markovian evolutions

Michele Pavon and Francesco Ticozzi

***Abstract*— The theory of *Schrödinger bridges* for diffusion processes is extended to discrete-time Markov chains, and to some problems for quantum discrete-time processes. Taking into account the past-future lack of symmetry of the discrete-time setting, results bear a striking resemblance to the classical ones. In particular, the solution of the path space maximum entropy problems is always obtained from the *prior* model by means of a suitable *multiplicative functional* transformation.**

***Index Terms*— Markov chain, maximum entropy problem, Schrödinger bridge, time reversal evolution, space-time harmonic function, quantum operation.**

I. Introduction

In two remarkable papers [19], [20], published before the very foundations of probability had been laid down, Schrödinger considered, and basically solved, the following abstract probabilistic problem: Suppose a large number N of independent Brownian particles have been observed to have density $\rho_0(x)$ at time t_0 and density $\rho_1(x)$ at some later time t_1. Suppose the latter density considerably differs from what is predicted by the law of large numbers. It is apparent that the particles have been transported in an unlikely way. But of the many unlikely ways in which this could have happened, which one is the most likely? In modern terminology, this is a problem of *large deviations* of the empirical distribution.

Using a coarse graining approach, Schrödinger computed the most likely endpoint distribution under the "prior" transition density of the Brownian motion $p(s,x,t,y)$ and with the prescribed marginals. It turned out that the solution, namely the bridge from ρ_0 to ρ_1 over Brownian motion, has at each time a density q that factors as $q(x,t) = \varphi(x,t)\hat{\varphi}(x,t)$, where φ and $\hat{\varphi}$ are, in the language of Doob, a p-harmonic and a p-coharmonic functions, respectively. The existence and uniqueness of such a pair $(\varphi, \hat{\varphi})$ satisfying the factorization above and the boundary conditions was guessed by Schrödinger on the basis of his intuition. He was later shown to be quite right in various degrees of generality by Fortet [9], Beurling [4], Jamison [11], Föllmer [8].

In our recent paper [18], we have derived corresponding results for discrete-time, classical and quantum Markovian evolutions: The solution process is obtained, in analogy to the diffusion case, via a suitable *multiplicative functional* transformation of the "prior" Markov process, see Theorem 2.3. As in the diffusion case, an abstract result of Beurling and Jamison can be used to prove existence and uniqueness of the solution of the Schrödinger system for finite, irreducible and aperiodic Markov chains, see Corollary 2.5. (The only previous discrete-time paper on this topic is [2], which deals with the continuous state space case and ignores by and large all the delicate points in the construction of the bridge).

This paper is basically a shortened version of [18] without proofs. In Section III, however, we make some ancillary observations on the role of space-time harmonic functions in the problem of feedback steering of a chain without altering the topology of the associated graph.

We then consider quantum channels, namely trace-preserving and completely-positive maps from density matrices to density matrices, representing the natural analogue of Markov chains transition operators. In order to derive corresponding results for these evolutions, we first develop various kinematical results. These concern extending the results on *time-reversal* of the channel, and developing *space-time harmonic processes*. In spite of the obvious difficulties one can expect from the non commutative structure, we are actually able to solve two key maximum entropy problems on path space. Remarkably, in the second case, the solution does not depend on the particular "quantum path" chosen. Moreover, with the appropriate understanding of objects and properties, in both cases it bears a striking similarity to the classical case. For quantum systems, this framework may be useful to attack steering problems [3] and to complement or improve quantum process tomography techinques (see e.g. [15] for a recent review of different methods). Exploring the relations of our framework with the theory of quantum error correction [12] appears to be a particularly promising research direction.

II. Schrödinger bridges for Markov chains

Consider a *Markov chain* $X = \{X(t); t = 0,1,2,\ldots\}$ taking values in the finite or countably infinite set $\mathcal{X}$. Since $\mathcal{X}$ is countable, we can identify $\mathcal{X}$ with a subset of $\mathbb{N}$. Let us introduce the distribution of $X(t)$ given by $p_i(t) = \mathbb{P}(X(t) = i)$ and the *transition probabilities* $p_{ij}(t) := \mathbb{P}(X(t+1) = j | X(t) = i)$. They are connected through

$$p_j(t+1) = \sum_i p_{ij}(t) p_i(t). \tag{1}$$

Let us agree that † always indicates adjoint with respect to the natural inner product. Hence, in the case of matrices,

M. Pavon is with Dipartimento di Matematica Pura ed Applicata, Università di Padova, via Trieste 63, 35131 Padova, Italy email:pavon@math.unipd.it

F. Ticozzi is with Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy, email:ticozzi@dei.unipd.it

Work partially supported by the CPDA080209/08 and QFUTURE research grants of the University of Padova, and by the Department of Information Engineering strategic project QUINTET.

it denotes transposition and, in the complex case below, transposition plus conjugation. We can then rewrite (1) as $p(t+1) = P^\dagger(t)p(t)$, where $p(t)^\dagger = (p_0(t), p_1(t), p_2(t), \dots)$ and $P(t) = (p_{ij}(t))$ is the transition matrix. The latter is *stochastic*, i.e. all elements are nonnegative and rows sum to one. Let us introduce the *n-step transition probabilities*

$$p_{ij}^{(n)}(t) := \mathbb{P}(X(t+n) = j | X(t) = i), \tag{2}$$

and the corresponding (stochastic) matrix $P^{(n)}(t) = (p_{ij}^{(n)}(t))$. Then, by the Markov property,

$$P^{(n)} = P(t+n-1) \cdot P(t+n-2) \cdots P(t). \tag{3}$$

Definition 2.1: A function $h : \mathbb{N} \times \mathcal{X} \to \mathbb{R}$ is called *space-time harmonic* for the transition mechanism $\{P(t); t = 0, 1, \dots\}$ of a chain if, for every $t \geq 0$ and all $i, j \in \mathcal{X}$, it satisfies the *backward equation*

$$h(t,i) = \sum_j p_{ij}(t) h(t+1, j). \tag{4}$$

Space-time harmonic functions, a terminology due to Doob and motivated by the case of diffusion processes, play a central role in constructing Schrödinger bridges. They are closely related to a class of *martingales* that are *instantaneous functions* of $X(t)$, see [5].

Definition 2.2: Let p and q be probability distributions on a finite or countably infinite set. The *Information Divergence* or *Relative Entropy* or *Kullback-Leibler Index* of q from p is

$$\mathbb{D}(p\|q) = \begin{cases} \sum_i p(i) \log \frac{p(i)}{q(i)}, & \text{supp}(p) \subseteq \text{supp}(q), \\ +\infty, & \text{supp}(p) \not\subseteq \text{supp}(q). \end{cases}, \tag{5}$$

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

Let $X = \{X(0), X(1), \dots\}$ be a Markov chain with state space $\mathcal{X}$, transition probabilities $(\pi_{ij}(t))$ and marginal probabilities $\mathbb{P}(X(t) = i) = \pi_i(t)$. Let $\mathbf{\Pi}$ denote the corresponding joint distribution of $\{X(0), X(1), \dots, X(T)\}$ (distributions on $\mathcal{X}^{T+1}$ are always denoted by capital, boldface letters). Let $\mathcal{D}(0, T; p^0, p^1)$ denote the family of Markovian distributions $\mathbf{P}$ on $\mathcal{X}^{T+1}$ that have marginals p^0 at time 0 and p^1 at time T, respectively, and have support contained in the support of $\mathbf{\Pi}$. We consider the following "path space" *Maximum Entropy Problem* (MEP):

$$\text{minimize} \quad \left\{ \mathbb{D}(\mathbf{P}\|\mathbf{\Pi}); \mathbf{P} \in \mathcal{D}(0, T; p^0, p^1) \right\}. \tag{6}$$

Theorem 2.3: Suppose there exists a pair of nonnegative functions $(\varphi, \hat{\varphi})$ satisfying on $[0, T] \times \mathcal{X}$ the system

$$\varphi(t,i) = \sum_j \pi_{ij}(t) \varphi(t+1, j), \tag{7}$$

$$\hat{\varphi}(t+1, j) = \sum_i \pi_{ij}(t) \hat{\varphi}(t, i), \tag{8}$$

and the boundary conditions

$$\varphi(0,i) \cdot \hat{\varphi}(0,i) := p_i^0, \quad \varphi(T,i) \cdot \hat{\varphi}(T,i) := p_i^1, \quad \forall i \in \mathcal{X}. \tag{9}$$

Suppose moreover that $\varphi(t,i) > 0, \ \forall 0 \leq t \leq T, \forall i \in \mathcal{X}$. Then, the Markov distribution $\hat{\mathbf{P}}$ in $\mathcal{D}(0, T; p^0, p^1)$ having transition probabilities

$$\hat{p}_{ij}(t) = \pi_{ij}(t) \frac{\varphi(t+1, j)}{\varphi(t,i)} \tag{10}$$

solves problem (MEP) (6).

If $(\varphi, \hat{\varphi})$ satisfy (7)-(8)-(9), so does the pair $(c\varphi, \frac{1}{c}\hat{\varphi})$ for all $c > 0$. Hence, uniqueness for the Schrödinger system is always intended up to such multiplications. As in the diffusion case, the problem is now reduced to establish, under suitable assumptions, existence and uniqueness for the Schrödinger system (7)-(8)-(9). Existence and uniqueness of the solution to the Schrödinger system (7)-(8)-(9) follows from a very deep result of Beurling [4], suitably extended by Jamison [10, Theorem 3.2].

Theorem 2.4: [18] Let $X = \{X(0), X(1), \dots\}$ be a Markov chain with state space $\mathcal{X}$ and transition probabilities $\pi_{ij}(t)$. Assume

1) p^1 is a distribution on $\mathcal{X}$ with $p_x^1 > 0, \forall x \in \mathcal{X}$;
2) $p(0, x, T, y) > 0, \forall x, y \in \mathcal{X}$.

Then the Schrödinger system (7)-(8)-(9) has a unique solution with $\varphi(t,x) > 0, \ \forall 0 \leq t \leq T, \forall x \in \mathcal{X}$.

In many important applications, the prior transition probabilities do not depend on time. We get the following result for ergodic Markov chains.

Corollary 2.5: Let $\{X(0), X(1), \dots\}$ be a Markov chain with state space $\mathcal{X}$ and transition matrix $\Pi = (\pi_{ij})$. Assume

1) p^1 is a distribution on $\mathcal{X}$ with $p_x^1 > 0, \forall x \in \mathcal{X}$;
2) the matrix Π^T has all positive elements.

Then the Schrödinger system (7)-(8)-(9) has a unique solution with $\varphi(t,x) > 0, \ \forall 0 \leq t \leq T, \forall x \in \mathcal{X}$.

III. Controlling Markov chains via multiplitical functional transformations

Suppose, for simplicity, that $\mathcal{X}$ has cardinality N. Let $\Pi^{(n)} = \Pi^n$ and $\hat{P}^{(n)}(t)$ denote the matrices of old and new n-step transition probabilities, respectively. Then, thanks to a series of cancellations, (10) yields

$$\hat{P}^{(n)}(t) = \text{diag}\left(\varphi(t,1), \varphi(t,2), \dots, \varphi(t,N)\right)^{-1} \Pi^n \times \text{diag}\left(\varphi(t+n,1), \varphi(t+n,2), \dots, \varphi(t+n,N)\right). \tag{11}$$

Thus, we see that also the n-step transition probabilities are obtained from the corresponding "prior" probabilities in a simple way. It is worthwhile to observe that this property is not peculiar of the Schrödinger bridge, but it is shared by any multiplicative functional transformation. Indeed, suppose $h(t,i)$ is positive and satisfies (4) for $t \geq 0$ and $i, j \in \mathcal{X}$, where $p_{ij}(t)$ are the transition probabilities of the uncontrolled chain. Then

$$p_{ij}^h(t) := p_{ij}(t) \frac{h(t+1, j)}{h(t,i)}, \tag{12}$$

provide a new Markov transition mechanism. Indeed, in view of (4),

$$\sum_j p_{ij}^h(t) = \sum_j p_{ij}(t) \frac{h(t+1, j)}{h(t,i)} = 1.$$

This occurs also in the diffusion case where the latter property follows from the Girsanov transformation, see e.g. [8]. It is important to stress that feedback control of a Markov chain through a multiplicative functional transformation does not require to change the topology of the underlying graph. Indeed, as it is apparent from (12), no new edge is required. This seems particularly appealing considering applications where the prior transition matrix is very sparse or building new edges is technologically unfeasible. We finally observe that h may be precomputed in order to obtain, in the light of (11), a desired evolution on a given finite time horizon.

IV. Quantum probabilities, entropy and quantum operations

Consider a finite-dimensional quantum system $\mathcal{Q}$, with associated Hilbert space $\mathcal{H}_{\mathcal{Q}}$ isomorphic to $\mathbb{C}^n$. In the quantum probability formalism, random variables or *observables* for the system are represented by Hermitian matrices $X \in \mathfrak{H}(n)$. They admit a spectral representation $X = \sum_j x_j \Pi_j$, where each real eigenvalue x_j represents the random outcome associated to the quantum *event* corresponding to the orthogonal projection Π_j. The role of the probability distributions is played here by positive-definite, unit-trace matrices $\rho \geq 0, \mathrm{tr}(\rho) = 1$, called *density matrices*. The set $\mathfrak{D}(n)$ of density matrices is convex and has the rank-one orthogonal projections as extreme points. Assume that the density matrix associated to the state of the system is ρ. The probability of measuring x_j, or in general the probability associated to the quantum event Π_j, is $\mathbb{P}_\rho(\Pi_j) = \mathrm{tr}(\Pi_j \rho \Pi_j)$. If the outcome corresponding to an event Π_j has been measured, the density matrix conditioned on the measurement record is $\rho_{|\Pi_j} = \frac{1}{\mathrm{tr}\,\Pi_j \rho \Pi_j} \Pi_j \rho \Pi_j$.

Notice that this implies that if the measurement has occurred, but the outcome has not been recorded, the correct conditional density matrix is: $\rho_{|X} = \sum_j \frac{1}{\mathrm{tr}\,\Pi_j \rho \Pi_j} \Pi_j \rho \Pi_j \cdot \mathbb{P}_\rho(\Pi_j) = \sum_j \Pi_j \rho \Pi_j$, which is in general different from the pre-measurement ρ, in contrast with the classical case. We refer to these "blind" measurement processes as *non-selective* measurements. For any matrix M, the *support* of M, denoted $\mathrm{supp}(M)$, is the orthogonal complement of $\ker(M)$. Given two density matrices ρ, σ, the *quantum relative entropy* is defined by $\mathbb{D}(\rho\|\sigma) = \mathrm{tr}(\rho(\log\rho - \log\sigma))$, if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$, and $+\infty$ otherwise.

As in the classical case, quantum relative entropy has the property of a pseudo-distance (see e.g. [17]): The Klein's Inequality $\mathbb{D}(\rho||\sigma) \geq 0$ holds, equality occurring if and only if $\rho = \sigma$. Moreover, quantum relative entropy is continuous where it is not infinite and it is jointly convex, but not symmetric, in its arguments. A wide class of physically relevant, Markovian transition mechanisms are represented by linear, Trace Preserving and Completely Positive (TPCP) maps from density matrices to density matrices. A TPCP map $\mathcal{E}^\dagger$, in turn, can be represented by a Kraus operator-sum [13], i.e.:

$$\rho_{t+1} = \mathcal{E}^\dagger(\rho_t) = \sum_j M_j \rho_t M_j^\dagger,$$

where the $n \times n$ matrices M_j must satisfy $\sum_j M_j^\dagger M_j = I$ in order for $\mathcal{E}^\dagger$ to be trace preserving. Notice that we employ the adjoint for maps acting on states to be consistent with the classical notation, where the transition matrix $P^\dagger$ acts on probability distributions while P acts on functions, see [16] and [21] for a discussion on the role of duality relations for Markov evolutions. The action of the dynamics on observables can be derived by duality with respect to the Hilbert-Schmidt inner product $\mathrm{tr}(X\mathcal{E}^\dagger(\rho_t)) = \mathrm{tr}(\mathcal{E}(X)\rho_t)$, where $\mathcal{E}(X) = \sum_j M_j^\dagger X M_j$. It follows that if $\mathcal{E}^\dagger(\cdot)$ is trace-preserving, then $\mathcal{E}(\cdot)$ is identity preserving and vice-versa. Consider now a quantum Markov process, generated by ρ_0 and a sequence of TPCP maps $\{\mathcal{E}_t^\dagger\}_{t\in[0,T-1]}$.

Definition 4.1: A sequence of observables $\{Y_t\}_{t\in[0,T-1]}$ is said to be *space-time harmonic* with respect to the family $\{\mathcal{E}_t\}_{t\in[0,T-1]}$ if $Y_t = \mathcal{E}_t(Y_{t+1})$.

As in the classical case, space-time harmonic processes will be shown to play a central role in the solution of maximum entropy problems on path spaces.

V. Time-reversal of quantum operations

Another key ingredient in the study of maximum entropy problems on path space, is, very much like for classical Markov chains, the reverse-time transition mechanism. Define $R_j(\mathcal{E}, \rho_t) = \rho_{t+1}^{-\frac{1}{2}} M_j \rho_t^{\frac{1}{2}}$, and the Kraus map:

$$\mathcal{R}^\dagger_{\mathcal{E},\rho_t}(\cdot) = \sum_j R_j(\mathcal{E}, \rho_t)(\cdot) R_j^\dagger(\mathcal{E}, \rho_t). \tag{13}$$

In [21], it is shown that this map is in fact a quantum operation, that it can be augmented to a trace-preserving quantum operation, and that it is the correct time-reversal for $\mathcal{E}$ with respect to the initial density ρ_t. This is established also in the case $\mathrm{rank}(\rho_{t+1}) < n$, thereby extending the results in [1]. For any ρ and $\mathcal{E}^\dagger$ with Kraus operators $\{M_k\}$, define the map $\mathcal{T}_\rho$ from *quantum operations to quantum operations* $\mathcal{T}_\rho : \mathcal{E}^\dagger \mapsto \mathcal{T}_\rho(\mathcal{E}^\dagger)$, where $\mathcal{T}_\rho(\mathcal{E}^\dagger)$ has Kraus operators $\{\rho^{\frac{1}{2}} M_k^\dagger (\mathcal{E}(\rho))^{-\frac{1}{2}}\}$. The results of [1] show that the action of $\mathcal{T}_\rho$ is independent of the particular Kraus representation of $\mathcal{E}^\dagger$. With this definition, we have that $\mathcal{T}_{\rho_t}(\mathcal{E}^\dagger) = \mathcal{R}^\dagger_{\mathcal{E},\rho_t}$.

Theorem 5.1 ([21]): Let $\mathcal{E}^\dagger$ be a TPCP map. If $\rho_{t+1} = \mathcal{E}^\dagger(\rho_t)$, then for any $\rho_t \in \mathfrak{D}(n)$, $\mathcal{R}^\dagger_{\mathcal{E},\rho_t}(\cdot)$ defined as in (13) is the *time-reversal* of $\mathcal{E}$ for ρ_t, i.e. $\rho_t = \mathcal{T}_{\rho_t}(\mathcal{E}^\dagger)(\rho_{t+1}) = \mathcal{R}^\dagger_{\mathcal{E},\rho_t}(\rho_{t+1})$, and $\mathcal{T}_{\rho_{t+1}}(\mathcal{R}^\dagger_{\mathcal{E},\rho_t})(\sigma_t) = \mathcal{E}^\dagger(\sigma_t)$, for all $\sigma_t \in \mathfrak{D}(\mathcal{H})$ such that $\mathrm{supp}(\sigma_t) \subseteq \mathrm{supp}(\rho_t)$. Moreover, it can be augmented to be TPCP without affecting the above properties. [1]

Remark 5.2: Notice that if ρ_t is full rank, $\mathcal{T}_{\rho_{t+1}} \circ \mathcal{T}_{\rho_t}$ is the the identity map on quantum operations. In general, the time-reversal mechanism is not unique [21], just as in the classical case. While studying error-correction problems , the same $\mathcal{R}^\dagger_{\mathcal{E},\rho}(\cdot)$ has been suggested by Barnum and Knill [1] as a near-optimal correction operator. It has also been

[1] By augmenting a Kraus map $\mathcal{E}$ with Kraus operators $\{M_k\}_{k=1,\ldots,m}$ to a TPCP map, we mean adding a finite number N of Kraus operators $\{M_k\}_{k=m+1,\ldots,m+N}$ such that $\sum_k M_k^\dagger M_k = I$.

proven there that $\mathcal{R}^{\dagger}_{\mathcal{E},\rho}(\cdot)$ is independent of the particular Kraus representation of $\mathcal{E}$.

Given a quantum Markov process, generated by ρ_0 and a sequence of TPCP maps $\{\mathcal{E}^{\dagger}_t\}_{t\in[0,T-1]}$, a sequence of observables $\{Y_t\}_{t\in[0,T-1]}$ is said to be *space-time harmonic in reverse-time* with respect to the family $\{\mathcal{R}_{\mathcal{E}_t,\rho_t}\}_{t\in[0,T-1]}$ if $Y_{t+1} = \mathcal{R}_{\mathcal{E}_t,\rho_t}(Y_t)$, extending Definition 4.1 in analogy with the classical case.

VI. Path space for quantum Markov evolutions

In the quantum case, the definition of a path-space for a Markov process is not obvious. Here, we build up quantum trajectories associating at each time an observable quantity and conditioning the state and the evolution to measurements of such observables. We get results that are in striking analogy with the classical case.

Consider a quantum Markov process for a finite dimensional system $\mathcal{Q}$ with associated Hilbert space $\mathcal{H}_{\mathcal{Q}}$, generated by an initial density matrix σ_0 and a sequence of TPCP maps $\{\mathcal{E}^{\dagger}_t\}_{t\in[0,T-1]}$, with each $\mathcal{E}^{\dagger}_t$ admitting a Kraus representation with matrices $\{M_k(t)\}$. We define a set of possible trajectories, or *quantum paths*, by considering a time-indexed family of observables $\{X_t\}$, $X_t = \sum_{i=1}^{m_t} x_i \Pi_i(t)$, with $t \in [0,T]$. The paths are then all the possible time-ordered sequences of events $(\Pi_{i_0}(0), \Pi_{i_1}(1), \ldots, \Pi_{i_T}(T))$, with $i_t \in [1,\, m_t]$. We can compute the joint probability for a given path with the nested expression:

$$w^{\mathcal{E}}_{(i_0,i_1,\ldots,i_T)}(\sigma_0) = \mathrm{tr}\Big(\Pi_{i_T}(T)\mathcal{E}^{\dagger}_{T-1}(\Pi_{i_{T-1}}(T-1)\ldots \ldots\mathcal{E}^{\dagger}_0(\Pi_{i_0}(0)\sigma_0\Pi_{i_0}(0))\ldots)\Pi_{i_T}(T)\Big).$$

Lemma 6.1: Define the *path-conditioned density matrices* for $t \in [0,T]$ via the relations

$$\begin{aligned}\hat{\sigma}_{\mathcal{E},0} &= \sum_{i_0}\Pi_{i_0}(0)\sigma_0\Pi_{i_0}(0),\\ \hat{\sigma}_{\mathcal{E},t+1} &= \hat{\mathcal{E}}^{\dagger}_t(\hat{\sigma}_{\mathcal{E},t})\\ &= \sum_{i_{t+1}}\Pi_{i_{t+1}}(t+1)\mathcal{E}^{\dagger}_t(\hat{\sigma}_{\mathcal{E},t})\Pi_{i_{t+1}}(t+1),\end{aligned} \quad (14)$$

where $\hat{\mathcal{E}}^{\dagger}_t$ is TPCP and can be represented with double-indexed Kraus operators $\{\Pi_i(t+1)M_k(t)\}$. The *marginal distribution* $w^{\mathcal{E}}_{i_t}(\sigma_0)$ at time $t \in [0,T]$, is then given by:

$$w^{\mathcal{E}}_{i_t}(\sigma_0) = \mathrm{tr}\Big(\Pi_{i_t}(t)\hat{\sigma}_{\mathcal{E},t}\Pi_{i_t}(t)\Big). \quad (15)$$

Remark 6.2: Imposing a (finite) set of possible trajectories by choosing the $\{X_t\}$, we have to condition the density matrix at time t on the past measurements. Unlike classical probability, even "non-selective" conditioning influences the state.

Observe moreover the following fact:

Proposition 6.3: The joint probabilities can be re-written in terms of the time reversal transitions for the *path-conditioned states* as:

$$w^{\mathcal{E}}_{(i_0,i_1,\ldots,i_T)}(\sigma_0) = \mathrm{tr}\Big(\Pi_{i_0}(0)\mathcal{R}^{\dagger}_{\hat{\mathcal{E}}_0,\hat{\sigma}_{\mathcal{E},0}}(\Pi_{i_1}(1)\ldots \mathcal{R}^{\dagger}_{\hat{\mathcal{E}}_{T-1},\hat{\sigma}_{\mathcal{E},T-1}}(\Pi_{i_T}(T)\hat{\sigma}_{\mathcal{E},T}\Pi_{i_T}(T))\ldots)\Pi_{i_0}(0)\Big).$$

This "backward" representation will play a key role in the solution of the maximum entropy problems we discuss in the next Section.

VII. Maximum entropy problems on quantum path spaces

We consider the simpler maximum entropy problems where only the initial or final density matrices are prescribed. The solution to these problems exhibit the same structure of their classical analogues, involving a "symmetrized" multiplicative functional transformation.

Let $\{\mathcal{E}^{\dagger}_t\}$ be a family of TPCP maps generating a quantum Markov process over $[0,T]$ with initial density matrix σ_0. Assume that at time T the density matrix of the system has been found to be $\bar{\rho}_T$, being different from the expected $\sigma_T = \mathcal{E}^{\dagger}_{T-1}\circ\ldots\circ\mathcal{E}^{\dagger}_0(\sigma_0)$. Let $\{X_t\}$ be a time-indexed family of observables defining a path space as above. We constraint *only* X_T to be such that $[X_T, \bar{\rho}_T] = 0$, and it admits a spectral decomposition with rank one $\Pi_j(T)$'s (this is quite natural, since $\bar{\rho}_T$ is given). Let as above $w^{\mathcal{E}}(\sigma_0)$ denote the path-space distribution induced by the initial condition σ_0 and the TPCP transitions $\{\mathcal{E}^{\dagger}_t\}$. For simplicity, in the reminder of the section, the reverse-time quantum operations are assumed to be trace preserving. The general case is simply obtained by augmenting the Kraus operators in order to have a trace preserving transformation, as detailed in Section V. Consider now the

Quantum Maximum Entropy Problem (QMEP1):

$$\text{minimize} \quad \big\{\mathbb{D}(w^{\mathcal{F}}(\rho_0)\|w^{\mathcal{E}}(\sigma_0)); w^{\mathcal{F}}(\rho_0) \in \Omega(\bar{\rho}_T)\big\} \quad (16)$$

with $\Omega(\bar{\rho}_T)$ the set of path space distribution induced by a quantum Markov process generated by a family of TPCP maps $\{\mathcal{F}^{\dagger}_t\}$ and some initial ρ_0 such that their path-conditioned, final density matrix satisfies $\hat{\rho}_{\mathcal{F},T} = \bar{\rho}_T$.

Since we required the $\Pi_i(T)$'s to be rank-one, it follows that for all $i \in [1, m_T]$:

$$\Pi_i(T)\rho_T\Pi_i(T) = \mathrm{tr}\big(\rho_T\Pi_i(T)\big)\Pi_i(T) = w^{\mathcal{F}}_{i_T}(\rho_0)\Pi_i(T).$$

Hence, we can write

$$w^{\mathcal{F}}_{(i_0,i_1,\ldots,i_T)}(\rho_0) = w^{\mathcal{F}}_{(i_0,i_1,\ldots,i_{T-1}|i_T)}\cdot w^{\mathcal{F}}_{i_T}(\rho_0), \quad (17)$$

defining the *conditional probabilities*:

$$w^{\mathcal{F}}_{(i_0,i_1,\ldots,i_{T-1}|i_T)} = \mathrm{tr}\Big(\Pi_{i_0}(0)\mathcal{R}^{\dagger}_{\hat{\mathcal{F}}_0,\hat{\rho}_0}(\Pi_{i_1}(1)\ldots \mathcal{R}^{\dagger}_{\hat{\mathcal{F}}_{T-1},\hat{\rho}_{T-1}}(\Pi_{i_T}(T))\ldots)\Pi_{i_0}(0)\Big).$$

By employing (17) and its equivalent for $w^{\mathcal{E}}_{(i_0,i_1,\ldots,i_T)}(\sigma_0)$, one is able to obtain a convenient relative entropy decomposition that allows to prove the following:

Theorem 7.1: A solution to (QMEP1) (16) is given by the quantum Markov process with path-conditioned final density $\bar{\rho}_T$ at time T and reverse-time transition mechanism equal to that of $\{\hat{\mathcal{E}}_t\}$, namely

$$\mathcal{R}^{\dagger}_{\hat{\mathcal{F}}_t,\hat{\rho}_{\mathcal{F},t}}(\cdot) = \mathcal{R}^{\dagger}_{\hat{\mathcal{E}},\hat{\sigma}_{\mathcal{E},t}}(\cdot), \quad \forall t \in [0,T-1]. \quad (18)$$

Notice that with this optimal choice, the total cost is bounded by the relative entropy of the conditioned final density matrices: $\sum_{i_T} w^{\mathcal{F}}_{i_T}(\rho_0) \log \frac{w^{\mathcal{F}}_{i_T}(\rho_0)}{w^{\mathcal{E}}_{i_T}(\sigma_0)} = \mathbb{D}(\bar\rho_T \| \hat\sigma_{\mathcal{E},T})$.

Let us compute the "forward" quantum operations, which, as in the classical case, will turn out to be time dependent even when the reference process is time-homogeneous. By Theorem 5.1, recalling that the conditioned transition mechanism $\hat{\mathcal{E}}^\dagger_t$ admits a Kraus representation with operators $\Pi_j(t+1)M_k(t)$, see (14), one finds that the Kraus operators of $\mathcal{R}^\dagger_{\hat{\mathcal{E}}_t,\hat\sigma_{\mathcal{E},t}}$ are given by the double-indexed $R_{j,k}(\hat{\mathcal{E}}_t,\hat\sigma_{\mathcal{E},t}) = \hat\sigma^{\frac12}_{\mathcal{E},t} M^\dagger_k(t)\Pi_j(t+1)\hat\sigma^{-\frac12}_{\mathcal{E},t+1}$. Reversing this TPCP map, now with respect to the state $\hat\rho_{\mathcal{F},t+1}$, we get: $F_{j,k}(t) = \hat\rho^{\frac12}_{\mathcal{F},t+1}\hat\sigma^{-\frac12}_{\mathcal{E},t+1}\big(\Pi_j(t+1)M_k(t)\big)\hat\sigma^{\frac12}_{\mathcal{E},t}\hat\rho^{-\frac12}_{\mathcal{F},t}$. which can be consider as a *non-commutative, "symmetrized"* version of a multiplicative functional transformation in the classical case. In fact, define $N_t = \hat\rho^{\frac12}_{\mathcal{F},t}\hat\sigma^{-\frac12}_{\mathcal{E},t}$. Then we have that $Y_t = N^\dagger_t N_t = \hat\sigma^{-\frac12}_{\mathcal{E},t}\hat\rho_{\mathcal{F},t}\hat\sigma^{-\frac12}_{\mathcal{E},t}$ is *space-time harmonic* with respect to the transition $\hat{\mathcal{E}}_t$, completing the analogy to the classical case. We remark that, since every time-reversal can be augmented to be TPCP by Theorem 5.1, one can always complete $\mathcal{R}^\dagger_{\hat{\mathcal{E}}_t,\hat\sigma_{\mathcal{E},t}}(\cdot)$, and then $\mathcal{F}^\dagger_t(\cdot)$, to be TPCP.

Consider now the case where the *initial state* is constrained to be equal to $\bar\rho_0$, different from the a-priori initial condition σ_0. Consider a path-space induced by observables $\{X_t\}$ such that X_0 has non-degenerate spectrum. By arguing as above, we get:

Theorem 7.2: A solution to (QMEP2)

$$\text{minimize} \quad \big\{\mathbb{D}(w^{\mathcal{F}}(\bar\rho_0)\|w^{\mathcal{E}}(\sigma_0)); w^{\mathcal{F}}(\bar\rho_0)\in\Omega(\bar\rho_0)\big\} \quad (19)$$

with $\Omega(\bar\rho_0)$ the set of path space probability distributions induced by a family of TPCP maps $\{\mathcal{F}^\dagger_t\}$ and initial state $\bar\rho_0$, is given by the quantum Markov process with initial density $\bar\rho_0$ and forward transitions:

$$\mathcal{F}_t(\cdot) = \mathcal{E}_t(\cdot), \quad \forall t\in[0,T-1]. \quad (20)$$

Remark 7.3: Altough the QMEP2 problem apparently depends on the choice of the quantum path-space, that is the observables $\{X_t\}_{t\in[0,T]}$, we remark that *its solution does not.* The difference between problems QMEP1 and QMEP2 is given by the fact that in QMEP2 we are concerned with the forward transitions, and we do not need to use the path-conditioned density matrices (14). The classical case does not present this asymmetry since classical non-selective measurements do not alter the state.

The final cost admits a bound similar to that in Problem QMEP1, that can be easily related to the unconditioned states. In fact, using monotonicity of relative entropy with respect to conditioning we get: $\sum_{i_T} w^{\mathcal{F}}_{i_0}(\bar\rho_0)\log\frac{w^{\mathcal{F}}_{i_0}(\bar\rho_0)}{w^{\mathcal{E}}_{i_0}(\sigma_0)} = \mathbb{D}(\hat\rho_0\|\hat\sigma_{\mathcal{E},0}) = \mathbb{D}(\bar{\mathcal{E}}^\dagger(\bar\rho_0)\|\bar{\mathcal{E}}^\dagger(\sigma_0)) \le \mathbb{D}(\bar\rho_0\|\sigma_0)$, with $\bar{\mathcal{E}}^\dagger(\rho) = \sum_i \Pi_i(0)\rho\Pi_i(0)$.

Notice that the operator-sum of the two reverse-time evolutions $\mathcal{R}_{\mathcal{F},\rho_t}, \mathcal{R}_{\mathcal{E},\sigma_t}$ satisfy, under appropriate restriction on the support of ρ_t,σ_t: $R_k(\mathcal{F}_t,\rho_t) = \rho^{\frac12}_t M^\dagger_k(t)\rho^{-\frac12}_{t+1} = \rho^{\frac12}_t\sigma^{-\frac12}_t R_k(\mathcal{E}_t,\sigma_t)\sigma^{\frac12}_{t+1}\rho^{-\frac12}_{t+1}$, which is again as a quantum symmetrized "multiplicative" functional transformation.

VIII. Conclusion and outlook

The classical theory of Schrödinger bridges is connected to a variety of other fascinating topics besides large deviations. First of all, there is Schrödinger's original motivation: He had observed the strong analogy between the time reversibility of the solution bridge and that of quantum mechanics: "Merkwürdige Analogien zur Quantenmechanik, die mir sehr des Hindenkens wert erscheinen". There is, however, another motivation: (Reverse-time) space-harmonic functions are involved in a strong form of the second law, see [21].

The Markov chain Schrödinger bridges appear as a flexible tool to be tested on a variety of applications, given the recent surfacing of the full modeling and computational power of Markov chains, cf. e.g. [5], [14]. For quantum systems, this framework may be useful to attack steering problems [3] and to complement or improve quantum process tomography techinques (see e.g. [15] for a recent review of different methods). Exploring the relations of our framework with the theory of quantum error correction [12] appears to be a particularly promising research direction. The problem of finding the time-reversal of quantum operations or quantum Markov semigroups representing the effect of noisy channels on some quantum code is strictly related to many central problems in quantum information and its realizations. Moreover, our path-space problems appear to be compatible with the general setting proposed in [6] to develop a quantum version of Sanov's theorem for product states. This suggests that our results may play a role in hypothesis testing and large deviation theory for quantum Markov evolution, once more in remarkable analogy with the classical setting.

References

[1] H. Barnum and E. Knill, Reversing quantum dynamics with near-optimal quantum and classical fidelity *J. Math. Phys.* **43**, 2097 (2002).

[2] A. Beghi, On the relative entropy of discrete-time Markov processes with given end-point densities, *IEEE Transactions on Information Theory*, **42** (1996), 1529 - 1535.

[3] A. Beghi, A. Ferrante and M. Pavon, How to steer a quantum system over a Schrödinger bridge, *Quantum Information Processing*, **1** (2002), 183-206.

[4] A. Beurling, An automorphism of product measures, *Ann. Math.* **72** (1960), 189-200.

[5] P. Brémaud, *Markov Chains. Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer-Verlag, New York, 1999.

[6] I. Bjelakovic, J.-D. Deuschel, T. Krüger, R. Seiler, R. Siegmund-Schultze, A. Szkola, A quantum version of Sanov's theorem, *Communications in mathematical physics* **260** (2005), 659-671.

[7] J. L. Doob, A Markov chain theorem, *Probability & Statistics* (The H. Cramér Volume), Wiley, 1959, pp. 50-57.

[8] H. Föllmer, Random fields and diffusion processes, in: *École d'Ètè de Probabilitès de Saint-Flour XV-XVII*, edited by P. L. Hennequin, Lecture Notes in Mathematics, Springer-Verlag, New York, 1988, vol.1362,102-203.

[9] R. Fortet, Résolution d'un système d'equations de M. Schrödinger, *J. Math. Pure Appl.* IX (1940), 83-105.

[10] B. Jamison, Reciprocal processes, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **30** (1974), 65-86.

[11] B. Jamison, The Markov processes of Schrödinger, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **32** (1975), 323-331.

[12] E. Knill and R. Laflamme, A theory of quantum error correcting codes,*Phys. Rev. A*,**55** (1997),900–911.

[13] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory*, Springer Verlag, 1983.
[14] M. Mitzenmacher and E. Upfal, *Probability and computing: Randomized algorithms and probabilistic analysis*, Cambridge University Press, 2005.
[15] M. Mohseni, A. T. Rezakhani and D. A. Lidar, Quantum Process Tomography: Resource Analysis of Different Strategies, *Phys. Rev. A*, **77**, 032322 (2008).
[16] E. Nelson, The adjoint Markov process. *Duke Math. J.*, **25** (1958), 671-690.
[17] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge Univ. Press, Cambridge, UK, 2000.
[18] M. Pavon and F. Ticozzi, Discrete-time classical and quantum Markovian evolutions: Maximum entropy problems on path space, to appear in the *J. Math. Phys.*, 2010, preprint arXiv:math-ph/0811.0933v2.
[19] E. Schrödinger, Über die Umkehrung der Naturgesetze, *Sitzungsberichte der Preuss Akad. Wissen. Berlin, Phys. Math. Klasse* (1931), 144-153.
[20] E. Schrödinger, Sur la théorie relativiste de l'électron et l'interpretation de la mécanique quantique, *Ann. Inst. H. Poincaré* **2**, 269 (1932).
[21] F. Ticozzi and M. Pavon, On time-reversal and space-time harmonic processes for Markovian quantum channels, April 2009, submitted for publication, preprint arXiv:quant-ph/0811.0929v2.