A new approach to robust transportation over networks

Yongxin Chen, Tryphon Georgiou, Michele Pavon and Allen Tannenbaum

***Abstract*— We consider transporting a mass on a given directed graph. The mass is initially concentrated on certain nodes and needs to be transported in a certain time period to another set of nodes, typically disjoint from the first. We seek a transport plan which is *robust* with respect to links and nodes failure. In order to achieve that, we need our mass to spread over time on all available routes between source and sink nodes as much as the topology of the graph allows. The scheduling consists in selecting transition probabilities for a Markovian evolution which is designed to be consistent with given initial and final marginal distributions. In order to construct such a transportation plan, we set up a maximum entropy problem (*Schrödinger bridge problem*) for probability laws on paths which can be viewed as an atypical stochastic control problem. To achieve robustness, we employ a *prior* distribution on paths which allocates equal probability to paths of equal length connecting any two nodes namely the *Ruelle-Bowen random walker* ($\mathfrak{M}_{\rm RB}$). The latter is also shown to be itself the time-homogeneous solution of a maximum entropy problem for (unnormalized) measures on paths. Since the optimal transport plan is computed via a Schrödinger bridge like problem, for which an efficient iterative algorithm is available [1], our approach appears also computationally attractive. We provide a few examples which illustrate the effectiveness of this method. While in this paper, we only consider strongly connected graphs, in a forthcoming journal paper [2], we show that our approach can be readily extended to not strongly connected graphs and weighted graphs. In the latter case, this strategy may be used to design a transportation plan which effectively compromises between robustness and other criteria such as *cost*.**

I. Introduction

Transport over networks is attracting increasing interest in the literature due to its relevance in a variety of classical and modern applications that include power transmission, traffic, communication networks, financial transactions, biological systems and so on [3]–[6]. Furthermore, the topic relates to a host of other questions pertaining to the connectivity of graphs and the relative significance of their nodes as in the Google PageRank problem [7] and the study of interaction between genes in biological networks [8].

Y. Chen is with the Department of Mechanical Engineering, University of Minnesota, Minneapolis, Minnesota MN 55455, USA; chen2468@umn.edu

T.T. Georgiou is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minnesota MN 55455, USA; tryphon@umn.edu

M. Pavon is with the Dipartimento di Matematica, Università di Padova, via Trieste 63, 35121 Padova, Italy; pavon@math.unipd.it

A. Tannenbaum is with the Department of Computer Science, Stony Brook University, Stony Brook, NY 11794: allen.tannenbaum@stonybrook.edu

Supported in part by the NSF under Grant ECCS-1509387, the AFOSR under Grants FA9550-12-1-0319 and FA9550-15-1-0045, the NIH under Grants P41-RR-013218, P41-EB-015902 and 1U24CA18092401A1. This work was also supported by the Vincentine Hermes-Luh Chair, and by the University of Padova Research Project CPDA 140897.

Our starting point is an important insight on the relation between the topological structure of a network and the entropy rate of a random walker on the graph [9], [10]. As it turns out, there is a unique way to specify transition probabilities in such a way that all paths of equal length joining any two nodes have equal probability. The corresponding Ruelle-Bowen-Parry measure $\mathfrak{M}_{\rm RB}$ on paths maximizes the entropy rate of a random walker, and this is a characteristic of the network. So far, the use of this concept has been to assign significance to each node in relation to the corresponding occupancy stationary distribution (*centrality measures*). Our interest in the measure $\mathfrak{M}_{\rm RB}$ has a different motivation. In the attempt to design a robust transportation plan featuring a mass which spreads on all available routes, we employ $\mathfrak{M}_{\rm RB}$ as a prior in a maximum entropy problem. Since the Ruelle-Bowen random walk provides a natural notion of "uniform" prior which gives equal importance to all paths, the resulting transportation flow that is selected to agree with specified initial and final marginals tends indeed to spread across all available paths as much as possible given the topological structure of the network. Thereby, such a flow leads to relatively low probability of conflict and congestion, and ensures a certain degree of inherent robustness of the transport plan.

The paper is outlined as follows. In Section II, we present the solution to a general Schrödinger bridge problem (SBP), where the prior measure is not necessarily a probability measure, as a straightforward extension of the results in [1], [11]. Section III is devoted to solutions of the SBP with equal initial and final marginals which have a *time-invariant* transition mechanism so that they admit invariant measures. We establish the surprising result (Theorem 3.1) that there is only one such bridge. This measure on paths can be constructed generalizing a classical result by Parry [12]. In Section IV, considering the special case of a prior transition given by the adjacency matrix, we describe the most important features of the Ruelle-Bowen random walker along the lines of [9]. We observe that this measure $\mathfrak{M}_{\rm RB}$ on trajectories can be viewed as a solution to a "time-homogeneous" Schrödinger bridge problem where the prior transition mechanism is given by the adjacency matrix. Section V describes our procedure to produce a robust transportation plan over a given strongly connected network: We take the Ruelle-Bowen distribution $\mathfrak{M}_{\rm RB}$ as prior in a Schrödinger bridge problem with prescribed initial and final marginals. We also prove that the optimal transportation can also be obtained in one step by taking the adjacency matrix as prior transition mechanism (Proposition 5.2). Finally, in Section VI we illustrate our approach on a simple graph.

II. Maximum entropy problems on paths

We discuss a generalization of the discrete Schrödinger Bridge problem (SBP) considered in [1], [11], where the "prior" is not necessarily a probability law and mass is not necessarily preserved during the evolution. We consider a finite state space $\mathcal{X} = \{1, \dots, n\}$ e.g., the nodes of a network, over a time-indexing set $\mathcal{T} = \{0, 1, \dots, N\}$. The goal is to determine a probability distribution P on the paths $\mathcal{X}^{N+1}$ in a way such that it matches the specified marginal distributions $\nu_0(\cdot)$ and $\nu_N(\cdot)$ and the resulting random evolution is closest to the "prior" in a suitable sense.

The prior law is induced by the Markovian evolution

$$\mu_{t+1}(x_{t+1}) = \sum_{x_t \in \mathcal{X}} \mu_t(x_t) m_{x_t x_{t+1}} \tag{1}$$

for nonnegative distributions $\mu_t(\cdot)$ over $\mathcal{X}$ with $t \in \mathcal{T}$. Throughout, we assume that $m_{ij} \geq 0$ for all indices $i, j \in \mathcal{X}$ and for simplicity, for the most part, that the matrix

$$M = [m_{ij}]_{i,j=1}^n$$

does not depend on t. In this case, we will often assume that all entries of M^N are positive. The rows of the transition matrix M do not necessarily sum up to one, so that the "total transported mass" is not necessarily preserved. This is the case, in particular, of a Markov chain with "creation" and "killing". It also occurs when M simply encodes the topological structure of a directed network with m_{ij} being zero or one, depending whether a certain transition is allowed. The evolution (1), together with measure $\mu_0(\cdot)$, which we assume positive on $\mathcal{X}$, i.e.,

$$\mu_0(x) > 0 \text{ for all } x \in \mathcal{X}, \tag{2}$$

induces a measure $\mathfrak{M}$ on $\mathcal{X}^{N+1}$ as follows. It assigns to a path $x = (x_0, x_1, \dots, x_N) \in \mathcal{X}^{N+1}$ the value

$$\mathfrak{M}(x_0, x_1, \dots, x_N) = \mu_0(x_0) m_{x_0 x_1} \cdots m_{x_{N-1} x_N}, \tag{3}$$

and gives rise to a flow of *one-time marginals*

$$\mu_t(x_t) = \sum_{x_{\ell \neq t}} \mathfrak{M}(x_0, x_1, \dots, x_N), \quad t \in \mathcal{T}.$$

The "prior" distribution $\mathfrak{M}$ on the space of paths may be at odds with a pair of specified marginals ν_0 and ν_N in that one or possibly both,

$$\mu_0(x_0) \neq \nu_0(x_0), \quad \mu_N(x_N) \neq \nu_N(x_N).$$

We denote by $\mathcal{P}(\nu_0, \nu_N)$ the family of probability distributions on $\mathcal{X}^{N+1}$ having the prescribed marginals. We seek a distribution in this set which is closest to the prior $\mathfrak{M}$ in a suitable entropic sense.

Recall that, for P and Q probability distributions, the *Kullback-Leibler distance* (divergence, relative entropy) $\mathbb{D}(P\|Q)$ is nonnegative and equal to zero if and only if $P = Q$ [13]. It can be extended to positive measures that are not probability distributions. Naturally, while the value of $\mathbb{D}(P\|Q)$ may turn out negative due to miss-match of scaling, the relative entropy is always jointly convex. We view the prior $\mathfrak{M}$ (specified by M and μ_0) in a similar manner, and consider the Schrödinger bridge problem:

Problem 2.1: Determine

$$\mathfrak{M}^*[\nu_0, \nu_N] = \operatorname{argmin}\{\mathbb{D}(P\|\mathfrak{M}) \mid P \in \mathcal{P}(\nu_0, \nu_N)\}. \tag{4}$$

Provided all entries of M^N are positive, the problem has a solution, which is unique due to strict convexity. This is stated next.

Theorem 2.2: Assume that M^N has all positive elements. Then there exist nonnegative functions $\varphi(\cdot)$ and $\hat{\varphi}(\cdot)$ on $[0, N] \times \mathcal{X}$ satisfying, for $t \in [0, N-1]$, the system

$$\varphi(t, i) = \sum_j m_{ij} \varphi(t+1, j), \tag{5a}$$

$$\hat{\varphi}(t+1, j) = \sum_i m_{ij} \hat{\varphi}(t, i) \tag{5b}$$

with the boundary conditions

$$\varphi(0, x_0)\hat{\varphi}(0, x_0) = \nu_0(x_0) \tag{5c}$$

$$\varphi(N, x_N)\hat{\varphi}(N, x_N) = \nu_N(x_N), \tag{5d}$$

for all $x_0, x_N \in \mathcal{X}$. Moreover, the solution $\mathfrak{M}^*[\nu_0, \nu_N]$ to Problem 2.1 is unique and obtained by

$$\mathfrak{M}^*(x_0, \dots, x_N) = \nu_0(x_0) \pi_{x_0 x_1}(0) \cdots \pi_{x_{N-1} x_N}(N-1),$$

where[1]

$$\pi_{ij}(t) := m_{ij} \frac{\varphi(t+1, j)}{\varphi(t, i)}. \tag{6}$$

Equation (6) specifies *one-step transition probabilities* that are well defined.

Proof: The argument in [11, Theorem 4.1] and [1, Section III] applies verbatim to this setting which is slightly more general in that M does not prescribe a probability kernel. The system (5a-5d) is known as a *Schrödinger system*. The existence of solution is shown in [1, Section III] by establishing that the composition

$$\hat{\varphi}(0, x_0) \overset{(M^T)^N}{\longrightarrow} \hat{\varphi}(N, x_N) \overset{(5d)}{\longrightarrow} \varphi(N, x_N) \longrightarrow \dots$$

$$\dots \overset{M^N}{\longrightarrow} \varphi(0, x_0) \overset{(5c)}{\longrightarrow} (\hat{\varphi}(0, x_0))_{\text{next}} \tag{7}$$

is contractive in the Hilbert metric [14]–[17]. The fact that $\pi_{ij}(t)$ in (6) satisfy $\sum_j \pi_{ij}(t) = 1$ follows from (5a). ■

The factors φ and $\hat{\varphi}$ are unique up to multiplication of φ by a positive constant and division of $\hat{\varphi}$ by the same constant. The statement of the theorem is analogous to results for the classical Schrödinger system (5) of diffusions that have been established by Fortet, Beurling, Jamison and Föllmer [18]–[21]. The requirement for M^N to have positive entries can be slightly relaxed and replaced by a suitable condition that guarantees existence of solution for the particular ν_0 and

[1] Here we use the convention that $0/0 = 0$.

ν_N. The case when M is time varying can also be readily established along the lines of [11, Theorem 4.1] and [1, Theorem 2].

Finally, to simplify the notations, let $\varphi(t)$ and $\hat{\varphi}(t)$ denote the column vectors with components $\varphi(t,i)$ and $\hat{\varphi}(t,i)$, respectively, with $i \in \mathcal{X}$. In matricial form, (5a), (5b) and (6) read

$$\varphi(t) = M\varphi(t+1), \quad \hat{\varphi}(t+1) = M^T\hat{\varphi}(t), \tag{8a}$$

and

$$\Pi(t) = [\pi_{ij}(t)] = \operatorname{diag}(\varphi(t))^{-1} M \operatorname{diag}(\varphi(t+1)). \tag{8b}$$

III. TIME-HOMOGENEOUS BRIDGES

In this section, we consider the case of Schrödinger bridge problems when the marginals are idential, namely, $\nu_0 = \nu_N = \nu$. In particular, we are interested in the case when the solution of the SBP corresponds to a time-homogeneous Markov evolution.

Since for the nonnegative matrix M we have M^N with only positive elements, by the celebrated Perron-Frobenius Theorem (see [22]), M has a unique positive eigenvalue λ_M and it is equal to the spectral radius. Let ϕ and $\hat{\phi}$ be the corresponding right and left eigenvectors, then both of them have only positive components. We normalize ϕ and $\hat{\phi}$ so that

$$\sum_{x\in\mathcal{X}} \phi(x)\hat{\phi}(x) = 1.$$

This leads to a special probability distribution

$$\bar{\nu}(x) = \phi(x)\hat{\phi}(x). \tag{9}$$

It turns out that $\bar{\nu}$ is the only probability measure such that the associated SBP has a time-homogeneous solution; we shall name it the *time-homogeneous bridge* associated with M. This is summarized in the following theorem. For reasons of space, we only discuss fully connected graphs referring to [2] for the general case.

Theorem 3.1: Let M be a nonnegative matrix such that M^N has only positive elements, and $\mathfrak{M}$ the measure on $\mathcal{X}^{N+1}$ given by (3) with μ_0 satisfying (2). Then the solution to the Schrödinger bridge problem

$$\mathfrak{M}^*[\bar{\nu}] = \operatorname{argmin}\{\mathbb{D}(P\|\mathfrak{M}) | P \in \mathcal{P}(\bar{\nu},\bar{\nu})\}, \tag{10}$$

where $\bar{\nu}$ is as in (9), has the time-invariant transition matrix

$$\bar{\Pi} = \lambda_M^{-1} \operatorname{diag}(\phi)^{-1} M \operatorname{diag}(\phi) \tag{11}$$

and invariant measure $\bar{\nu}$. Conversely, suppose $N > 1$ and that M has only positive elements. Given probability measure ν, suppose that the transition matrix of $\mathfrak{M}^*[\nu]$ does not depend on time. Then $\nu = \bar{\nu}$.

Proof: Since ϕ and $\hat{\phi}$ are the right and left eigenvectors of M associated with eigenvalue λ_M, the nonnegative functions φ and $\hat{\varphi}$ defined by

$$\varphi(t) = \lambda_M^t \phi, \quad \hat{\varphi}(t) = \lambda_M^{-t}\hat{\phi}$$

satisfy the Schrödinger system (5). By Theorem 2.2, the solution $\mathfrak{M}^*[\bar{\nu}]$ of the Schrödinger bridge problem (10) then has the transition matrix (see (6))

$$\begin{aligned}\bar{\Pi} &= \operatorname{diag}(\varphi(0))^{-1} M \operatorname{diag}(\varphi(1)) \\ &= \lambda_M^{-1} \operatorname{diag}(\phi)^{-1} M \operatorname{diag}(\phi),\end{aligned}$$

which is exactly (11). Moreover, since

$$\bar{\Pi}^T\bar{\nu} = \lambda_M^{-1} \operatorname{diag}(\phi) M^T \hat{\phi} = \bar{\nu},$$

it follows that $\bar{\nu}$ is the corresponding invariant measure. Converserly, suppose $N > 1$ and the transition matrix Π_ν of $\mathfrak{M}^*(\nu)$ is time invariant. Let $\varphi_\nu(t) = M\varphi_\nu(t+1)$ be the space-time harmonic function associated to the minimizer $\mathfrak{M}^*(\nu)$. Consider times $t = N-2, N-1, N$. By (8) and the time invariance of Π_ν, we must have

$$\begin{aligned}\Pi_\nu &= \operatorname{diag}(\varphi_\nu(N-2))^{-1} M \operatorname{diag}(\varphi_\nu(N-1)) \\ &= \operatorname{diag}(\varphi_\nu(N-1))^{-1} M \operatorname{diag}(\varphi_\nu(N)).\end{aligned}$$

It follows that

$$M = D_\nu(N-1) M D_\nu(N)^{-1},$$

where the $D_\nu(t) = \operatorname{diag}(\varphi_\nu(t))\operatorname{diag}(\varphi_\nu(t-1))^{-1} = \operatorname{diag}(d_1^\nu(t),\ldots,d_n^\nu(t))$ are still diagonal. Hence,

$$m_{ij} = e_i^T M e_j = d_i(N-1) m_{ij} d_j(N)^{-1}, \quad \forall i,j.$$

Varying j for a fixed i, since $m_{ij} \neq 0$, we get that $D(N)$ is a scalar matrix, say λI, not depending on t and $\varphi(N)$ is a right eigenvector of M. By the Perron-Frobenius Theorem, it follows that $\varphi(N)$ corresponds to λ_M. It readily follows that $\hat{\varphi}(0)$ is an eigenvector of M^T with positive components corresponding to the same eigenvalue λ_M. By (5c)-(5d), ν is equal to $\bar{\nu}$. This completes the proof. ∎

As we shall see in the next section, when M is the adjacency matrix of a strongly connected, directed graph, the associated time-homogenous bridge turns out to be the Ruelle-Bowen measure $\mathfrak{M}_{\rm RB}$ [9, Section III]. This probability measure has a number of useful properties, in particular it gives the same probability to paths of the same length between any two given nodes. All of these are discussed in Section IV.

IV. THE RUELLE-BOWENS RANDOM WALK

In this section we explain the Ruelle-Bowen (RB) random walk and some of its properties. We follow closely Delvenne and Libert [9]. The RB random walk amounts to a Markovian evolution on a directed graph that assigns equal probabilities to all paths of equal length between any two nodes. The motivation of [9] was to assign a natural invariant probability to nodes based on relations that are encoded by a graph, and thereby determine a *centrality measure*, akin to Google Page ranking, yet more robust and discriminating. Our motivation is quite different. The RB random walk provides a uniform distribution on paths. Therefore, it represents a natural distribution to serve as prior in the SBP in order to achieve a maximum spreading of the mass transported over the available paths. In this section, besides reviewing basics on the RB random walk, we show that the RB distribution is itself a solution to the SBP.

We consider a strongly connected aperiodic, directed graph

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}).$$

The idea in Google Page ranking the nodes is based on a random walk where a jump takes place from one node to any of its neighbors with equal probability. The alternative proposed in [9] is an *entropy ranking*, based on the stationary distribution of the RB random walk [12], [23]. The transition mechanism is such that it induces a uniform distribution on *paths* of equal length joining any two nodes. This distribution is characterized as the one maximizing the *entropy rate* [13] for the random walker. Let us briefly recall the relevant concept. The Shannon entropy for paths of length t is at most

$$\log |\{\text{paths of length } t\}|.$$

Hence, the entropy rate is bounded by the *topological entropy rate*

$$H_{\mathcal{G}} = \limsup_{t\to\infty} (\log |\{\text{paths of length } t\}|/t).$$

Here $|\{\cdot\}|$ denotes the cardinality of a set. Notice that $H_{\mathcal{G}}$ only depends on the graph $\mathcal{G}$ and not on the probability distribution on paths. More specifically, if A denotes the adjacency matrix of the graph, then the number of paths of length t is the sum of all the entries of A^t. Thus, it follows that $H_{\mathcal{G}}$ is the logarithm of the spectral radius of A, namely the maximum of the absolute values of the eigenvalues of A, namely,

$$H_{\mathcal{G}} = \log(\lambda_A). \tag{12}$$

We next construct the Rulle-Bowen random walk. Since A is the adjacency matrix of a strongly connected aperiodic graph, it satisfies that A^N has only positive elements for some $N > 0$. By Perron-Frobenius Theorem, the spectral radius λ_A is an eigenvalue of A, and the associated left and right eigenvectors[2] u and v, i.e.,

$$A^T u = \lambda_A u, \quad Av = \lambda_A v \tag{13}$$

have only positive components. We further normalize u and v so that

$$\langle u, v\rangle := \sum_i u_i v_i = 1.$$

As in Section III, it is readily seen that their componentwise multiplication

$$\mu_{RB}(i) = u_i v_i \tag{14}$$

defines a probability distribution which is invariant under the transition matrix

$$R = [r_{ij}], \quad r_{ij} = \frac{v_j}{\lambda_A v_i} a_{ij}, \tag{15}$$

that is,

$$R^T \mu_{RB} = \mu_{RB}. \tag{16}$$

The transition matrix R in (15) together with the stationary measure μ_{RB} in (14), define the *Ruelle-Bowen path measure*

$$\mathfrak{M}_{\rm RB}(x_0, x_1, \ldots, x_N) := \mu_{RB}(x_0) r_{x_0 x_1} \cdots r_{x_{N-1} x_N}. \tag{17}$$

Proposition 4.1: The Ruelle-Bowen measure $\mathfrak{M}_{\rm RB}$ (17) assigns probability $\lambda_A^{-t} u_i v_j$ to any path of length t from node i to node j.

Proof: Starting from the stationary distribution (14), and in view of (15), the probability of a path ij is

$$u_i v_i \left(\frac{1}{\lambda_A} v_i^{-1} v_j\right) = \frac{1}{\lambda_A} u_i v_j,$$

assuming that node j is accessible from node i in one step. Likewise, the probability of the path ijk is

$$u_i v_i \left(\frac{1}{\lambda_A} v_i^{-1} v_j\right)\left(\frac{1}{\lambda_A} v_j^{-1} v_k\right) = \frac{1}{\lambda_A^2} u_i v_k$$

independent of the intermediate state j, and so on. Thus, the claim follows. ■

Thus, the Ruelle-Bowen measure $\mathfrak{M}_{\rm RB}$ has the striking property that it induces a uniform probability measure on paths of equal length between any two given nodes. We quote from [9] "Since the number of paths of length t is of the order of λ_A^t (up to a factor) the distribution on paths of fixed length is uniform up to a factor (which does not depend on t). Hence the Shannon entropy of paths of length t grows as $t \log \lambda_A$, up to an additive constant. The entropy rate of this distribution is thus $\log \lambda_A$ which is optimal" by the expression for $H_{\mathcal{G}}$ in (12).

The construction of the Ruelle-Bowen distribution is obvious a special case of the measure $\bar{\nu}$ in (9) in Section III, in the special case when M is the adjacency matrix A of a graph. Therefore, the RB measure is the solution of the particular SBP when the "prior" transition mechanism is given by the adjacency matrix! This observation is apparently new and beautifully links the topological entropy rate to a maximum entropy problem on path space. This is summarized as follows.

Proposition 4.2: Let A be the adjacency matrix of a strongly connected graph aperiodic $\mathcal{G}$, $\mathfrak{M}$ the nonnegative measure on $\mathcal{X}^{N+1}$ given by (3) with $M = A$ and μ_0 satisfying (2). Then, the Ruelle-Bowen measure $\mathfrak{M}_{\rm RB}$ (17) solves the SBP (4) with marginals $\nu_0 = \nu_N = \mu$.

V. Robust transport over networks

Once again we consider a strongly connected aperiodic, directed graph $\mathcal{G}$ with n vertices and seek to transport a unit mass from initial distribution ν_0 to terminal distribution ν_N in at most N steps. We identify node 1 as a *source* and node n as a *sink*. The task is formalized by setting an initial marginal distribution $\nu_0(x) = \delta_{1x}(x)$ Kronecker's

[2] We are now conforming to notation in [9] for ease of comparison. Hence we use u and v rather than $\hat{\phi}$ and ϕ.

delta. Similarly, the final distribution is $\nu_N(x) = \delta_{nx}(x)$. Generally, we seek a transportation plan which is *robust* and avoids *congestion* as much as the topology of the graph permits. This latter feature of the transportation plan will be achieved in this section indirectly, without explicitly bringing into the picture the capacity of each edge. With these two key specifications in mind, we like to control the flux so that the initial mass *spreads as much as possible* on the feasible paths joining vertices 1 and n in N steps before reconvening at time N in vertex n. We shall achieve this by constructing a suitable Markovian transition mechanism. As we want to allow for the possibility that all or part of the mass reaches node n at some time less than N, we always include a loop in node n so that our adjacency matrix A always has $a_{nn} = 1$. As explained in Section IV that $\mathfrak{M}_{\rm RB}$ gives equal probability to paths joining two specific vertices, it is natural to use it as a prior in the SBP with marginals δ_{1x}, δ_{nx} so as to achieve the spreading of the probability mass on the feasible paths joining the source with the sink. Thus, we consider the following.

Problem 5.1: Determine

$$\mathfrak{M}^*[\delta_{1x}, \delta_{nx}] = \operatorname{argmin}\{\mathbb{D}(P\|\mathfrak{M}_{\rm RB}) | P \in \mathcal{P}(\delta_{1x}, \delta_{nx})\}.$$

By Theorem 2.2, the optimal, time varying transition matrix $\Pi^*(t)$ of the above problem is given, recalling the notations in (8), by

$$\Pi^*(t) = \operatorname{diag}(\varphi(t))^{-1} R \operatorname{diag}(\varphi(t+1)), \tag{18}$$

where

$$\varphi(t) = R\varphi(t+1),\ \hat\varphi(t+1) = R^T\hat\varphi(t),$$

with the boundary conditions

$$\varphi(0,x)\hat\varphi(0,x) = \delta_{1x}(x),\ \varphi(N,x)\hat\varphi(N,x) = \delta_{nx}(x) \tag{19}$$

for all $x \in \mathcal{X}$. In view of (15), if we define

$$\varphi_v(t) := \lambda_A^{-t} \operatorname{diag}(v)\varphi(t), \quad \hat\varphi_v(t) := \lambda_A^t \operatorname{diag}(v)^{-1}\hat\varphi(t),$$

then we have

$$\varphi_v(t) = A\varphi_v(t{+}1),\ \hat\varphi_v(t{+}1) = A^T\hat\varphi_v(t),\ t = 0, \dots, N{-}1.$$

Moreover,

$$\varphi_v(t,x)\hat\varphi_v(t,x) = \varphi(t,x)\hat\varphi(t,x),\ t = 0, \dots, N-1,\ x \in \mathcal{X}.$$

Here, again, A is the adjacency matrix of $\mathcal{G}$ and v is the right eigenvector corresponding to the spectral radius λ_A.

The above analysis provides another interesting way to express $\mathfrak{M}^*[\delta_{1x}, \delta_{nx}]$; it also solves the Schrödinger bridge problem with the same marginals δ_{1x} and δ_{nx} while different prior transition matrix A, the adjacency matrix. Thus, we can replace the two-step procedure by a single bridge problem. This is summarized in the following proposition.

Proposition 5.2: Let A be the adjacency matrix of a strongly connected aperiodic graph $\mathcal{G}$, $\mathfrak{M}$ the nonnegative measure on $\mathcal{X}^{N+1}$ given by (3) with $M = A$ and μ_0 satisfying (2). Then, the solution $\mathfrak{M}^*[\delta_{1x}, \delta_{nx}]$ of Problem 5.1 also solves the Schrödinger bridge problem

$$\min\{\mathbb{D}(P\|\mathfrak{M}) | P \in \mathcal{P}(\delta_{1x}, \delta_{nx})\}. \tag{20}$$

The iterative algorithm of [1, Section III] can now be based on (20) to efficiently compute the transition matrix of the optimal robust transport plan $\mathfrak{M}^*[\delta_{1x}, \delta_{nx}]$.

Remark 5.3: Finally, observe that if A^N has also zero elements, the robust transport described in this section may still be feasible provided there is at least one path of length N joining node 1 with node n, i.e., $(A^N)_{1n} > 0$.

As we discussed in the beginning of this section, the intuition to use $\mathfrak{M}_{RB}$ as a prior is to achieve the spreading of the probability on all the feasible paths connecting the source and the sink. It turns out that this is indeed the case; the solution $\mathfrak{M}^*[\delta_{1x}, \delta_{nx}]$ of Problem 5.1 assigns equal probability to all the feasible paths of lengths N joining the source 1 with the sink n. Too see this, by (18), the probability of the optimal transport plan $\mathfrak{M}^*[\delta_{1x}, \delta_{nx}]$ assigns on path $x = (x_0, x_1, \dots, x_N)$ is

$$\begin{aligned}\mathfrak{M}^*[\delta_{1x}, \delta_{nx}](x) &= \delta_{1x}(x_0)\prod_{t=1}^{N-1} r_{x_t x_{t+1}} \frac{\varphi(t+1, x_{t+1})}{\varphi(t, x_t)}\\ &= \delta_{1x}(x_0)\frac{\varphi_v(N, x_N)}{\varphi_v(0, x_0)}\prod_{t=1}^{N-1} a_{x_t x_{t+1}}.\end{aligned}$$

Indeed, $\prod_{t=1}^{N-1} a_{x_t x_{t+1}} = 1$ for a feasible path and 0 otherwise. Moreover, $\delta_{1x}(x_0)\frac{\varphi_v(N,x_N)}{\varphi_v(0,x_0)}$ depends only on the boundary points x_0, x_N. We conclude that $\mathfrak{M}^*[\delta_{1x}, \delta_{nx}]$ assigns equal probability to all the feasible paths. Finally, there are $(A^N)_{1n}$ feasible paths of length N connecting nodes 1 and n. Thus we have established the following result.

Proposition 5.4: $\mathfrak{M}^*[\delta_{1x}, \delta_{nx}]$ assigns probability $1/(A^N)_{1n}$ to each of all the feasible paths of length N connecting node 1 with node n.

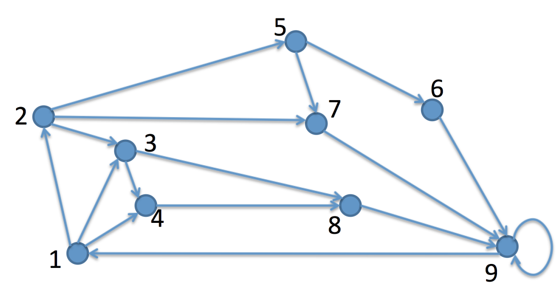

Fig. 1: Network topology

VI. Examples

We study Problem 5.1 for a simple academic example to illustrate our method. Consider the graph in Figure 1 with the following adjacency matrix

$$A = \begin{bmatrix} 0&1&1&1&0&0&0&0&0\\ 0&0&1&0&1&0&1&0&0\\ 0&0&0&1&0&0&0&1&0\\ 0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&1&1&0&0\\ 0&0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&0&1\\ 1&0&0&0&0&0&0&0&1 \end{bmatrix}.$$

We seek to transport a unit mass from node 1 to node 9 in $N = 3$ and 4 steps. We add a self loop at node 9, i.e., $a_{99} = 1$, to allow for transport paths with different step sizes.

The shortest path from node 1 to 9 is of length 3 and there are three such paths, which are $1-2-7-9$, $1-3-8-9$ and $1-4-8-9$. If we want to transport the mass with minimum number of steps, we may end up using one of these three paths. This is not so robust. On the other hand, if we apply the Schrödinger bridge framework with the RB measure $\mathfrak{M}_{\rm RB}$ as the prior, then we get a transport plan with equal probabilities using all these three paths. The evolution of mass distribution is given by

$$\begin{bmatrix} 1&0&0&0&0&0&0&0&0\\ 0&1/3&1/3&1/3&0&0&0&0&0\\ 0&0&0&0&0&0&1/3&2/3&0\\ 0&0&0&0&0&0&0&0&1 \end{bmatrix},$$

where the four rows of the matrix show the mass distribution at time step $t = 0, 1, 2, 3$ respectively. As we can see, the mass spreads out first and then goes to node 9. When we allow for more steps $N = 4$, the mass spreads even more before reassembling at node 9, as shown below

$$\begin{bmatrix} 1&0&0&0&0&0&0&0&0\\ 0&4/7&2/7&1/7&0&0&0&0&0\\ 0&0&1/7&1/7&2/7&0&1/7&2/7&0\\ 0&0&0&0&0&1/7&1/7&2/7&3/7\\ 0&0&0&0&0&0&0&0&1 \end{bmatrix}.$$

VII. Concluding remarks & future directions

In this paper, we have proposed a novel approach to design a robust transportation plan on a given strongly connected aperiodic, directed graph. It is based on a sort of generalized maximum entropy problem (Schrödinger bridge problem) for measures on paths of the given network. Taking as prior measure the Ruelle-Bowen law $\mathfrak{M}_{\rm RB}$, the solution naturally tends to spread the mass on all available routes joining the source and the sink. Hence, the resulting transport appears robust with respect to links/nodes failure.

This approach can be adapted to graphs that are not strongly connected, as well as to weighted graphs [2]. In the latter case, it can be used to effectively compromise between robustness and cost or achieving robustness taking capacity of the links into consideration. Since the optimal transport plan is computed via a Schrödinger bridge like problem, for which an efficient iterative algorithm is available, our approach appears also computationally attractive. Also, this approach relates to a number of fast developing and fascinating nearby areas such as robustness of graphs [3], [8], [24], [25], optimal mass transport [26] and discrete curvature [27], [28]. All these will be investigated in our future work.

References

[1] T. T. Georgiou and M. Pavon, "Positive contraction mappings for classical and quantum schrödinger systems," *Journal of Mathematical Physics*, vol. 56, no. 3, p. 033301, 2015.

[2] Y. Chen, T. T. Georgiou, M. Pavon, and A. Tannenbaum, "Robust transport over networks," *under preparation*, 2016.

[3] D. S. Callaway, M. E. Newman, S. H. Strogatz, and D. J. Watts, "Network robustness and fragility: Percolation on random graphs," *Physical review letters*, vol. 85, no. 25, p. 5468, 2000.

[4] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[5] D. M. Scott, D. C. Novak, L. Aultman-Hall, and F. Guo, "Network robustness index: a new method for identifying critical links and evaluating the performance of transportation networks," *Journal of Transport Geography*, vol. 14, no. 3, pp. 215–227, 2006.

[6] G. Cabanes, E. van Wilgenburg, M. Beekman, and T. Latty, "Ants build transportation networks that optimize cost and efficiency at the expense of robustness," *Behavioral Ecology*, vol. 26, no. 1, pp. 223–231, 2014.

[7] S. Brin and L. Page, "Reprint of: The anatomy of a large-scale hypertextual web search engine," *Computer networks*, vol. 56, no. 18, pp. 3825–3833, 2012.

[8] R. Sandhu, T. Georgiou, E. Reznik, L. Zhu, I. Kolesov, Y. Senbabaoglu, and A. Tannenbaum, "Graph curvature for differentiating cancer networks," *Scientific reports*, vol. 5, 2015.

[9] J.-C. Delvenne and A.-S. Libert, "Centrality measures and thermodynamic formalism for complex networks," *Physical Review E*, vol. 83, no. 4, p. 046117, 2011.

[10] H. Kitano, "Towards a theory of biological robustness," *Molecular systems biology*, vol. 3, no. 1, p. 137, 2007.

[11] M. Pavon and F. Ticozzi, "Discrete-time classical and quantum markovian evolutions: Maximum entropy problems on path space," *Journal of Mathematical Physics*, vol. 51, no. 4, p. 042104, 2010.

[12] W. Parry, "Intrinsic markov chains," *Transactions of the American Mathematical Society*, vol. 112, no. 1, pp. 55–66, 1964.

[13] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[14] B. Lemmens and R. Nussbaum, *Nonlinear Perron-Frobenius Theory*. Cambridge University Press, 2012, no. 189.

[15] G. Birkhoff, "Extensions of jentzsch's theorem," *Transactions of the American Mathematical Society*, vol. 85, no. 1, pp. 219–227, 1957.

[16] P. Bushell, "On the projective contraction ratio for positive linear mappings," *Journal of the London Mathematical Society*, vol. 2, no. 2, pp. 256–258, 1973.

[17] P. J. Bushell, "Hilbert's metric and positive contraction mappings in a Banach space," *Archive for Rational Mechanics and Analysis*, vol. 52, no. 4, pp. 330–338, 1973.

[18] R. Fortet, "Résolution d'un système d'équations de M. Schrödinger," *J. Math. Pures Appl.*, vol. 83, no. 9, 1940.

[19] A. Beurling, "An automorphism of product measures," *The Annals of Mathematics*, vol. 72, no. 1, pp. 189–200, 1960.

[20] B. Jamison, "Reciprocal processes," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 30, pp. 65–86, 1974.

[21] H. Föllmer, "Random fields and diffusion processes," in *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*. Springer, 1988, pp. 101–203.

[22] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 2012.

[23] D. Ruelle, *Thermodynamic formalism: the mathematical structure of equilibrium statistical mechanics*. Cambridge University Press, 2004.

[24] A. Jamakovic and S. Uhlig, "On the relationship between the algebraic connectivity and graph's robustness to node and link failures," in *Next Generation Internet Networks, 3rd EuroNGI Conference on*. IEEE, 2007, pp. 96–102.

[25] C. Wang, E. Jonckheere, and R. Banirazi, "Wireless network capacity versus ollivier-ricci curvature under heat-diffusion (hd) protocol," in *American Control Conference (ACC), 2014*. IEEE, 2014, pp. 3536–3541.

[26] C. Villani, *Optimal transport: old and new*. Springer, 2008, vol. 338.

[27] Y. Ollivier, "Ricci curvature of markov chains on metric spaces," *Journal of Functional Analysis*, vol. 256, no. 3, pp. 810–864, 2009.

[28] ——, "A survey of ricci curvature for metric spaces and markov chains," *Probabilistic approach to geometry*, vol. 57, pp. 343–381, 2010.