Hamiltonian process on finite graphs

Jianbo Cui, Shu Liu, Haomin Zhou

Hong Kong Polytechnic University
Georgia Institute of Technology
Georgia Institute of Technology
jianbo.cui@polyu.edu.hk
sliu459@gatech.edu
hmzhou@math.gatech.edu

March 2022

Table of Contents

Motivation

1. Curiosity.

[1] J. Maas, Gradient flows of the entropy for finite Markov chains, J. Funct. Anal. 261 (8) (2011)

[2] S. Chow, W. Huang, Y. Li, H. Zhou, Fokker-Planck equations for a free energy functional or Markov process on a graph, Arch. Ration. Mech. Anal. 203 (3) (2012)

Motivation

1. Curiosity.
2. The notion of *gradient flow on graph* has been investigated extensively using optimal transport theory[1] [2]; Whether the concept of Hamiltonian process on graph exists or not?

[1] J. Maas, Gradient flows of the entropy for finite Markov chains, J. Funct. Anal. 261 (8) (2011)

[2] S. Chow, W. Huang, Y. Li, H. Zhou, Fokker-Planck equations for a free energy functional or Markov process on a graph, Arch. Ration. Mech. Anal. 203 (3) (2012)

Motivation

3 Recent developments on **discrete** *optimal transport (OT) problem*[1], *Schrödinger equations (SE)* [2] as well as *Schrödinger Bridge Problem (SBP)*[3] [4] have demonstrated Hamiltonian principles on graph. Can we unify them and establish a general framework for Hamiltonian process on graph?

---

[1] W. Gangbo, W. Li, C. Mou, Geodesics of minimal length in the set of probability measures on graphs, ESAIM Control Optim. Calc. Var. 25 (2019) 78

[2] S. Chow, W. Li, H. Zhou, A discrete Schrödinger equation via optimal transport on graphs, J. Funct. Anal. 276 (8) (2019)

[3] C. Léonard, A survey of the Schrödinger problem and some of its connections with optimal transport, Discrete Contin. Dyn. Syst. 34 (4) (2014)

[4] C. Léonard, Lazy random walks and optimal transport on graphs, Ann. Probab. 44 (3) (2016)

Table of Contents

Background: Hamiltonian system in probability space

Let us start from

- The dynamical version of OT problem

$$\min_{v} \left\{ \int_0^1 \mathbb{E}[L(\boldsymbol{X}_t, v(\boldsymbol{X}_t, t))] \, dt \right\}, \tag{1}$$
$$\dot{\boldsymbol{X}}_t = v(\boldsymbol{X}_t, t), \ \boldsymbol{X}_0 \sim \rho_a, \ \boldsymbol{X}_1 \sim \rho_b,$$

Background: Hamiltonian system in probability space

Let us start from

- The dynamical version of OT problem

$$\min_{v} \left\{ \int_0^1 \mathbb{E}[L(\boldsymbol{X}_t, v(\boldsymbol{X}_t, t))] \, dt \right\}, \tag{1}$$
$$\dot{\boldsymbol{X}}_t = v(\boldsymbol{X}_t, t), \ \boldsymbol{X}_0 \sim \rho_a, \ \boldsymbol{X}_1 \sim \rho_b,$$

it is equivalent to

- The optimal control problem on $\mathcal{P}(\mathbb{R}^d)$

$$\min_{\rho, v} \left\{ \int_0^1 \int_{\mathbb{R}^d} L(x, v(x,t)) \, \rho(x,t) \, dxdt \right\}, \tag{2}$$
$$\frac{\partial \rho(x,t)}{\partial t} + \nabla \cdot (\rho(x,t) v(x,t)) = 0, \ \rho(\cdot, 0) = \rho_a, \ \rho(\cdot, 1) = \rho_b.$$

Background: Hamiltonian system in probability space

- Solution to (1) leads to Hamiltonian system on $\mathcal{T}^*\mathbb{R}^d$

$$\dot{\boldsymbol{X}}_t = \frac{\partial}{\partial p} H(\boldsymbol{X}_t, \boldsymbol{p}_t), \; \boldsymbol{X}_0 \sim \rho_a$$
$$\dot{\boldsymbol{p}}_t = -\frac{\partial}{\partial x} H(\boldsymbol{X}_t, \boldsymbol{p}_t). \text{ choose } \boldsymbol{p}_0 = \boldsymbol{p}_0(\boldsymbol{X}_0), \text{ s.t. } \boldsymbol{X}_1 \sim \rho_b. \tag{3}$$

Here we define the Hamiltonian $H(x,p) = \sup_v \{p \cdot v - L(x,v)\}$.

Background: Hamiltonian system in probability space

- Solution to (1) leads to Hamiltonian system on $\mathcal{T}^*\mathbb{R}^d$

$$\dot{\boldsymbol{X}}_t = \frac{\partial}{\partial p} H(\boldsymbol{X}_t, \boldsymbol{p}_t), \ \boldsymbol{X}_0 \sim \rho_a$$
$$\dot{\boldsymbol{p}}_t = -\frac{\partial}{\partial x} H(\boldsymbol{X}_t, \boldsymbol{p}_t). \text{ choose } \boldsymbol{p}_0 = \boldsymbol{p}_0(\boldsymbol{X}_0), \text{ s.t. } \boldsymbol{X}_1 \sim \rho_b. \tag{3}$$

- Solution to (2) leads to **Hamiltonian system** on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$

$$\partial_t \rho(x,t) + \nabla \cdot (\rho(x,t) \frac{\partial H}{\partial p}(x, \nabla S(x,t))) = 0, \rho(\cdot, 0) = \rho_a$$
$$\partial_t S(x,t) + H(x, \nabla S(x,t)) = 0. \text{ choose } S(\cdot,0) \text{ s.t. } \rho(\cdot,1) = \rho_b. \tag{4}$$

Here we define the Hamiltonian $H(x,p) = \sup_v \{p \cdot v - L(x,v)\}$.

Background: Hamiltonian system in probability space

Q: How can one treat (4) as Hamiltonian system on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$?

[1] S. Chow, W. Li, H. Zhou, Wasserstein Hamiltonian flows, J. Differ. Equ. 268 (3) (2020)

Background: Hamiltonian system in probability space

Q: How can one treat (4) as Hamiltonian system on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$?

A: We claim that (4) is the Hamiltonian flow[1] of

$$\mathscr{H}(\rho, S) = \int_{\mathbb{R}^d} H(x, \nabla S(x))\rho(x)\, dx$$

on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$ with respect to certain symplectic form ω.

[1] S. Chow, W. Li, H. Zhou, Wasserstein Hamiltonian flows, J. Differ. Equ. 268 (3) (2020)

Background: Hamiltonian system in probability space

We define the *cotangent bundle (phase space)* of $\mathcal{P}(\mathbb{R}^d)$ as

$$\mathcal{T}^*\mathcal{P}(\mathbb{R}^d) = \left\{ (\rho, S) \,\middle|\, \rho \in \mathcal{P}(\mathbb{R}^d),\ S \in L^1(\rho)/\sim \right\},$$

where we denote $S_1 \sim S_2$ if $S_1(x) = S_2(x) + \mathrm{Const}.$

Background: Hamiltonian system in probability space

We define the *cotangent bundle (phase space)* of $\mathcal{P}(\mathbb{R}^d)$ as

$$\mathcal{T}^*\mathcal{P}(\mathbb{R}^d) = \left\{ (\rho, S) \;\middle|\; \rho \in \mathcal{P}(\mathbb{R}^d),\ S \in L^1(\rho)/\sim \right\},$$

where we denote $S_1 \sim S_2$ if $S_1(x) = S_2(x) + \mathrm{Const}.$
We define the *symplectic form* ω on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$ as

$$\omega((\dot{\rho}_1, \dot{S}_1), (\dot{\rho}_2, \dot{S}_2)) = \int_{\mathbb{R}^d} \dot{\rho}_1 \dot{S}_2 - \dot{\rho}_2 \dot{S}_1 \, dx,$$

for any two tangent vectors $(\dot{\rho}_1, \dot{S}_1), (\dot{\rho}_2, \dot{S}_2)$ at $(\rho, S) \in \mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$.

Background: Hamiltonian system in probability space

By definition of ω, one can derive the Hamiltonian flow of $\mathscr{H}(\rho, S)$ on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$ as

$$\partial_t \rho = \frac{\delta}{\delta S}\mathscr{H}(\rho, S), \quad \partial_t S = -\frac{\delta}{\delta \rho}\mathscr{H}(\rho, S). \tag{5}$$

Background: Hamiltonian system in probability space

By definition of ω, one can derive the Hamiltonian flow of $\mathscr{H}(\rho, S)$ on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$ as

$$\partial_t \rho = \frac{\delta}{\delta S}\mathscr{H}(\rho, S), \quad \partial_t S = -\frac{\delta}{\delta \rho}\mathscr{H}(\rho, S). \tag{5}$$

Here $\frac{\delta}{\delta\rho}, \frac{\delta}{\delta S}$ denotes the L^2 variation w.r.t. ρ, S. We calculate

$$\frac{\delta}{\delta S}\mathscr{H}(\rho, S) = -\nabla \cdot (\rho \frac{\partial H(x, \nabla S)}{\partial p}) \quad \frac{\delta}{\delta \rho}\mathscr{H}(\rho, S) = H(x, \nabla S) \tag{6}$$

Background: Hamiltonian system in probability space

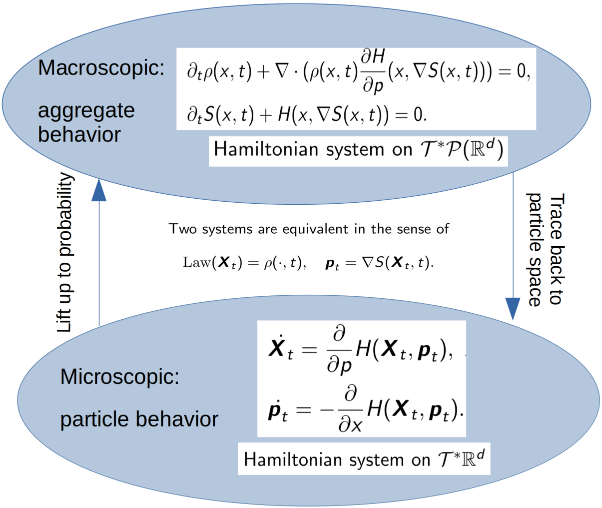
By definition of ω, one can derive the Hamiltonian flow of $\mathscr{H}(\rho, S)$ on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$ as

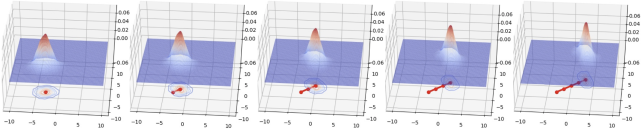
$$\partial_t \rho = \frac{\delta}{\delta S}\mathscr{H}(\rho, S), \quad \partial_t S = -\frac{\delta}{\delta \rho}\mathscr{H}(\rho, S). \tag{5}$$

Here $\frac{\delta}{\delta\rho}, \frac{\delta}{\delta S}$ denotes the L^2 variation w.r.t. ρ, S. We calculate

$$\frac{\delta}{\delta S}\mathscr{H}(\rho, S) = -\nabla \cdot (\rho \frac{\partial H(x, \nabla S)}{\partial p}) \quad \frac{\delta}{\delta \rho}\mathscr{H}(\rho, S) = H(x, \nabla S) \tag{6}$$

Plug (6) in (5), we recover the PDE system (4). This verifies our previous claim.

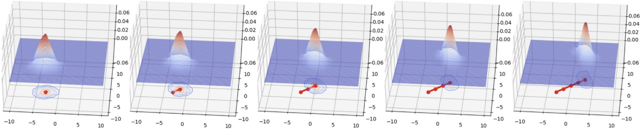

Macroscopic:
aggregate behavior
$\partial_t \rho(x,t) + \nabla \cdot \left(\rho(x,t) \frac{\partial H}{\partial p}(x, \nabla S(x,t))\right) = 0,$
$\partial_t S(x,t) + H(x, \nabla S(x,t)) = 0.$
Hamiltonian system on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$
Lift up to probability
Trace back to particle space
Two systems are equivalent in the sense of
$\mathrm{Law}(\boldsymbol{X}_t) = \rho(\cdot, t), \quad \boldsymbol{p}_t = \nabla S(\boldsymbol{X}_t, t).$
Microscopic:
particle behavior
$\dot{\boldsymbol{X}}_t = \frac{\partial}{\partial p} H(\boldsymbol{X}_t, \boldsymbol{p}_t),$
$\dot{\boldsymbol{p}}_t = -\frac{\partial}{\partial x} H(\boldsymbol{X}_t, \boldsymbol{p}_t).$
Hamiltonian system on $\mathcal{T}^*\mathbb{R}^d$



(a) t=0 (b) t=0.25 (c) t=0.5 (d) t=0.75 (e) t=1

Figure: [1] Hamiltonian system on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^2)$ as solution to dynamical OT

[1] S. Liu, S. Ma, Y. Chen, H. Zha, and H. Zhou, “Learning high dimensional wwsserstein geodesics,” arXiv preprint arXiv:2102.02992, 2021

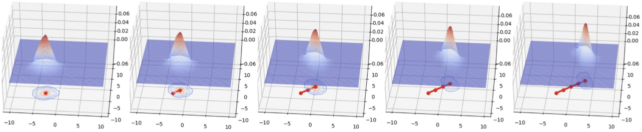


(a) t=0 (b) t=0.25 (c) t=0.5 (d) t=0.75 (e) t=1

Figure: [1] Hamiltonian system on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^2)$ as solution to dynamical OT

Set $L(x,v) = \frac{|v|^2}{2}$, $H(x,p) = \frac{|p|^2}{2}$, $\mathscr{H}(\rho, S) = \int_{\mathbb{R}^d} \frac{1}{2}|\nabla S(x)|^2 \rho(x)\, dx$.

[1] S. Liu, S. Ma, Y. Chen, H. Zha, and H. Zhou, "Learning high dimensional wwsserstein geodesics," arXiv preprint arXiv:2102.02992, 2021



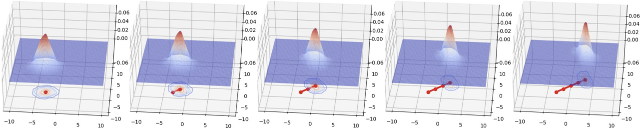
(a) t=0 (b) t=0.25 (c) t=0.5 (d) t=0.75 (e) t=1

Figure: [1] Hamiltonian system on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^2)$ as solution to dynamical OT

Set $L(x,v) = \frac{|v|^2}{2}$, $H(x,p) = \frac{|p|^2}{2}$, $\mathscr{H}(\rho, S) = \int_{\mathbb{R}^d} \frac{1}{2}|\nabla S(x)|^2 \rho(x)\, dx$.

- On $\mathcal{T}^*\mathbb{R}^d$, (red trajectory as $\{\boldsymbol{X}_t\}$), $\dot{\boldsymbol{X}}_t = \boldsymbol{p}_t, \quad \dot{\boldsymbol{p}}_t = 0$.

[1] S. Liu, S. Ma, Y. Chen, H. Zha, and H. Zhou, "Learning high dimensional wwsserstein geodesics," arXiv preprint arXiv:2102.02992, 2021



(a) t=0 (b) t=0.25 (c) t=0.5 (d) t=0.75 (e) t=1

Figure: [1] Hamiltonian system on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^2)$ as solution to dynamical OT

Set $L(x,v) = \frac{|v|^2}{2}$, $H(x,p) = \frac{|p|^2}{2}$, $\mathscr{H}(\rho, S) = \int_{\mathbb{R}^d} \frac{1}{2}|\nabla S(x)|^2 \rho(x)\, dx$.

- On $\mathcal{T}^*\mathbb{R}^d$, (red trajectory as $\{\boldsymbol{X}_t\}$), $\dot{\boldsymbol{X}}_t = \boldsymbol{p}_t, \quad \dot{\boldsymbol{p}}_t = 0.$
- On $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$ (evolving density as $\{\rho_t\}$)

$$\partial_t \rho(x,t) + \nabla \cdot (\rho \nabla S(x,t)) = 0,\ \partial_t S(x,t) + \frac{1}{2}|\nabla S(x,t)|^2 = 0.$$

[1] S. Liu, S. Ma, Y. Chen, H. Zha, and H. Zhou, "Learning high dimensional wwsserstein geodesics," arXiv preprint arXiv:2102.02992, 2021

Table of Contents

A motivating example: optimal transport on graph G

Our logic:

A motivating example: optimal transport on graph G

Our logic:

Consider an optimal control problem (actually an OT problem) on $\mathcal{P}(G)$;

A motivating example: optimal transport on graph G

Our logic:

Consider an optimal control problem (actually an OT problem) on $\mathcal{P}(G)$;

$$\Downarrow$$

A motivating example: optimal transport on graph G

Our logic:

Consider an optimal control problem (actually an OT problem) on $\mathcal{P}(G)$;

$$\Downarrow$$

The critical point of such problem naturally forms a Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$;

A motivating example: optimal transport on graph G

Our logic:

Consider an optimal control problem (actually an OT problem) on $\mathcal{P}(G)$;

$$\Downarrow$$

The critical point of such problem naturally forms a Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$;

$$\Downarrow$$

A motivating example: optimal transport on graph G

Our logic:

Consider an optimal control problem (actually an OT problem) on $\mathcal{P}(G)$;

$$\Downarrow$$

The critical point of such problem naturally forms a Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$;

$$\Downarrow$$

Find a random process on G that realizes the Hamiltonian system;

A motivating example: optimal transport on graph G

Our logic:

Consider an optimal control problem (actually an OT problem) on $\mathcal{P}(G)$;

$$\Downarrow$$

The critical point of such problem naturally forms a Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$;

$$\Downarrow$$

Find a random process on G that realizes the Hamiltonian system;

$$\Downarrow$$

A motivating example: optimal transport on graph G

Our logic:

Consider an optimal control problem (actually an OT problem) on $\mathcal{P}(G)$;

$$\Downarrow$$

The critical point of such problem naturally forms a Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$;

$$\Downarrow$$

Find a random process on G that realizes the Hamiltonian system;

$$\Downarrow$$

Such random process is our desired definition.

Notations & general setting

Notations & general setting

- Consider a undirected graph $G(V, E)$ with N vertices $V = \{1, 2, ..., N\}$ and edge set $E \subset \{(i,j) | i \neq j, \ i,j \in V, \ (i,j), \ (j,i) \text{ denote the same edge}\}$.

Notations & general setting

- Consider a undirected graph $G(V,E)$ with N vertices $V = \{1,2,...,N\}$ and edge set $E \subset \{(i,j) | i \neq j,\ i,j \in V,\ (i,j),\ (j,i) \text{ denote the same edge}\}$.
- We define

$$\mathcal{P}(G) = \{(\rho_i)_{i=1}^N \mid \rho_i \geq 0,\ \sum_{i=1}^N \rho_i = 1\}.$$

Notations & general setting

- Consider a undirected graph $G(V,E)$ with N vertices $V = \{1, 2, ..., N\}$ and edge set $E \subset \{(i,j) | i \neq j,\ i,j \in V,\ (i,j),\ (j,i) \text{ denote the same edge}\}$.
- We define

$$\mathcal{P}(G) = \{(\rho_i)_{i=1}^N \mid \rho_i \geq 0,\ \sum_{i=1}^N \rho_i = 1\}.$$

- And the neighbouring set of vertex j as

$$N(j) = \left\{ l \in V \mid (j,l) \in E \right\}.$$

Notations & general setting

- Consider a undirected graph $G(V,E)$ with N vertices $V=\{1,2,...,N\}$ and edge set $E\subset\{(i,j)|i\neq j,\ i,j\in V,\ (i,j),\ (j,i)\text{ denote the same edge}\}$.
- We define

$$\mathcal{P}(G)=\{(\rho_i)_{i=1}^N \mid \rho_i\geq 0,\ \sum_{i=1}^N \rho_i=1\}.$$

- And the neighbouring set of vertex j as

$$N(j)=\left\{l\in V \mid (j,l)\in E\right\}.$$

- For any $(j,l)\in E$, define the weight function $\theta_{jl}(\rho)=\theta_{lj}(\rho)$ as certain kind of average of density ρ_j,ρ_l, i.e., $\min\{\rho_j,\rho_l\}\leq\theta_{jl}(\rho)\leq\max\{\rho_j,\rho_l\}$.

OT problem on G

OT problem on G

We mimic (2), and consider the OT problem on graph G,

$$\min_{\rho,v}\left\{\int_0^1 \langle v,v\rangle_{\theta(\rho)}dt\right\}, \tag{7}$$
$$\partial_t\rho+\mathrm{div}_G^{\theta}(\rho v)=0,\ \rho(\cdot,0)=\rho_a,\ \rho(\cdot,1)=\rho_b.$$

One should require v to be **skew-symmetric**, $v^{\top}=-v$, this guarantees mass conservation $\sum_{i=1}^N \rho_i=1$.
We define

$$\langle v,v\rangle_{\theta(\rho)}=\frac{1}{2}\sum_{(j,l)\in E}\theta_{jl}(\rho)v_{jl}^2,\quad (\mathrm{div}_G^{\theta}(\rho v))_j=-\sum_{l\in N(j)}\theta_{jl}(\rho)v_{jl}.$$

$$\min_{\rho,v}\left\{\int_0^1\int_{\mathbb{R}^d}\frac{1}{2}|v(x,t)|^2\ \rho(x,t)\ dxdt\right\}, \tag{2}$$
$$\frac{\partial\rho(x,t)}{\partial t}+\nabla\cdot(\rho(x,t)v(x,t))=0,\ \rho(\cdot,0)=\rho_a,\ \rho(\cdot,1)=\rho_b.$$

Critical point of (7) as Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$

Critical point of (7) as Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$

Notice that (7) is a constrained optimization problem, Lagrange multiplier method and KKT condition leads the system of (ρ, S)

$$\begin{cases} \frac{d\rho_i}{dt} + \mathrm{div}_G^\theta(\rho \nabla_G S) = 0, \\ \frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)} \frac{\partial \theta_{ij}(\rho)}{\partial \rho_i}(S_i - S_j)^2 = 0. \end{cases} \tag{8}$$

$\nabla_G S = (S_i - S_j)_{ij}$, and $(\mathrm{div}_G^\theta(\rho \nabla_G S))_i = \sum_{j\in N(i)} \theta_{ij}(\rho)(S_j - S_i)$.

Critical point of (7) as Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$

Notice that (7) is a constrained optimization problem, Lagrange multiplier method and KKT condition leads the system of (ρ, S)

$$\begin{cases} \frac{d\rho_i}{dt} + \mathrm{div}_G^{\theta}(\rho \nabla_G S) = 0, \\ \frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)} \frac{\partial \theta_{ij}(\rho)}{\partial \rho_i}(S_i - S_j)^2 = 0. \end{cases} \tag{8}$$

$\nabla_G S = (S_i - S_j)_{ij}$, and $(\mathrm{div}_G^{\theta}(\rho \nabla_G S))_i = \sum_{j\in N(i)} \theta_{ij}(\rho)(S_j - S_i)$.

Consider the Hamiltonian $\mathscr{H}(\rho, S) = \frac{1}{4}\sum_{(i,j)\in E} \theta_{ij}(\rho)(S_i - S_j)^2$,

Critical point of (7) as Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$

Notice that (7) is a constrained optimization problem, Lagrange multiplier method and KKT condition leads the system of (ρ, S)

$$\begin{cases} \frac{d\rho_i}{dt} + \mathrm{div}_G^\theta(\rho \nabla_G S) = 0, \\ \frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)} \frac{\partial \theta_{ij}(\rho)}{\partial \rho_i}(S_i - S_j)^2 = 0. \end{cases} \tag{8}$$

$\nabla_G S = (S_i - S_j)_{ij}$, and $(\mathrm{div}_G^\theta(\rho\nabla_G S))_i = \sum_{j\in N(i)} \theta_{ij}(\rho)(S_j - S_i)$.

Consider the Hamiltonian $\mathscr{H}(\rho, S) = \frac{1}{4}\sum_{(i,j)\in E} \theta_{ij}(\rho)(S_i - S_j)^2$, set symplectic matrix $\Omega = \begin{pmatrix} & I_N \\ -I_N & \end{pmatrix}$, one can verify the Hamiltonian flow of $\mathscr{H}(\rho, S)$ w.r.t. Ω on $\mathcal{T}^*\mathcal{P}(G)$ is exactly (8).

Relate (8) to certain Markov process on G

Relate (8) to certain Markov process on G

Recall the first continuity equation in (8)

$$\frac{d\rho_i}{dt} + \sum_{j \in N(i)} \theta_{ij}(\rho)(S_j - S_i) = 0,$$

Relate (8) to certain Markov process on G

Recall the first continuity equation in (8)

$$\frac{d\rho_i}{dt} + \sum_{j \in N(i)} \theta_{ij}(\rho)(S_j - S_i) = 0,$$

To figure out the random process behind it, we recast this equation in the form of *Master (Chapman–Kolmogorov) equation*

$$\frac{d\rho}{dt} = \rho Q, \tag{9}$$

where (we assume $\rho_i > 0$ for all $i \in V$)

$$Q_{ji}(t) = 1_{\{(i,j)\in E\}} \frac{\theta_{ij}(\rho(t))}{\rho_j(t)}(S_i(t) - S_j(t)), \ j \neq i$$

$$Q_{ii}(t) = -\sum_{j \in N(i)} Q_{ij}(t) = -\sum_{j \in N(i)} \frac{\theta_{ji}(\rho(t))}{\rho_i(t)}(S_j(t) - S_i(t))$$

Relate (8) to certain Markov process on G

[1] V.N. Kolokoltsov, Nonlinear Markov Processes and Kinetic Equations, Cambridge Tracts in Mathematics, vol. 182, Cambridge University Press, Cambridge, 2010

Relate (8) to certain Markov process on G

The matrix Q is called the *transition rate matrix*, the Master equation (9) corresponds to a Markov process if Q satisfies

$$Q_{ii} \leq 0, \quad Q_{ij} \geq 0, \ \text{for any } j \neq i, \quad \sum_{j=1}^{N} Q_{ij} = 0. \tag{10}$$

[1] V.N. Kolokoltsov, Nonlinear Markov Processes and Kinetic Equations, Cambridge Tracts in Mathematics, vol. 182, Cambridge University Press, Cambridge, 2010

Relate (8) **to certain Markov process on** G

The matrix Q is called the *transition rate matrix*, the Master equation (9) corresponds to a Markov process if Q satisfies

$$Q_{ii} \leq 0, \quad Q_{ij} \geq 0, \text{ for any } j \neq i, \quad \sum_{j=1}^{N} Q_{ij} = 0. \qquad (10)$$

Assume (10) holds, then (9) corresponds to a **time-inhomogeneous Markov process**[1]. To be more specific, (9) corresponds to a **nonlinear Markov processes**[1] whose transition rate matrix Q depends not only on the current state i but also on the current distribution ρ.

[1] V.N. Kolokoltsov, Nonlinear Markov Processes and Kinetic Equations, Cambridge Tracts in Mathematics, vol. 182, Cambridge University Press, Cambridge, 2010

Relate (8) to certain Markov process on G

Relate (8) to certain Markov process on G

To summarize, we associate the Hamiltonian system (8) with a nonlinear Markov process

$$\frac{d\rho}{dt} = \rho Q(S, \rho),$$

with transition rate matrix $Q(S, \rho)$ depending on density ρ and potential S,

Relate (8) to certain Markov process on G

To summarize, we associate the Hamiltonian system (8) with a nonlinear Markov process

$$\frac{d\rho}{dt} = \rho Q(S, \rho),$$

with transition rate matrix $Q(S, \rho)$ depending on density ρ and potential S, such process is second order in the sense that time-dependent S further solves the discrete Hamilton-Jacobi equation

$$\frac{dS_i}{dt} + \frac{1}{2} \sum_{j \in N(i)} \frac{\partial \theta_{ij}(\rho)}{\partial \rho_i} (S_i - S_j)^2 = 0.$$

Table of Contents

Hamiltonian process on a finite graph

Definition (Hamiltonian process on G)

A stochastic process $\{X_t\}$ is called a **Hamiltonian process on the graph** G if

1. The density ρ_t of X_t satisfies the following generalized Master equation,

$$\frac{d\rho_t}{dt} = \rho_t Q(S_t, \rho_t, t),$$

with

$$Q_{ij}(S, \rho, t) = 1_{(i,j)\in E} f_{ji}(S_j - S_i, \rho, t), \; Q_{ii}(S, \rho, t) = -\sum_{j\in N(i)} Q_{ij}(S, \rho, t).$$

And $f_{ji} : \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^+ \cup \{0\} \to \mathbb{R}^+ \cup \{0\}$ (guarantees (10)), is a real-valued measurable function which is piece-wise continuous in the first component.

2. The density ρ and the potential S form a Hamiltonian system on the cotangent bundle $\mathcal{T}^*\mathcal{P}(G)$ of the density space $\mathcal{P}(G)$.

Theorem (Exact form of the Hamiltonian)

Suppose that the stochastic process $\{X_t\}_{t\geq 0}$ with density $\{\rho_t\}_{t\geq 0}$ and potential $\{S_t\}_{t\geq 0}$ forms a Hamiltonian process on the graph G. In addition assume that F_{ij} is the antiderivative of f_{ij}. Then the Hamiltonian always possesses the form

$$\mathscr{H}(\rho, S) = \sum_{i\in V} \sum_{j\in N(i)} \rho_i F_{ji}(S_j - S_i, \rho, t) + \mathcal{V}(\rho, t), \tag{11}$$

where $\mathcal{V}$ is a function depending ρ and t. Moreover, the Hamiltonian system on $\mathcal{T}^*\mathcal{P}(G)$ is

$$\frac{\partial}{\partial t}\rho_i(t) = \sum_{j\in N(i)} f_{ij}(S_i - S_j, \rho, t)\rho_j - f_{ji}(S_j - S_i, \rho, t)\rho_i,$$

$$\frac{\partial}{\partial t}S_i(t) = -\sum_{j\in N(i)} \left(F_{ji}(S_j - S_i, \rho, t) + \rho_j \frac{\partial}{\partial \rho_i} F_{ji}(S_j - S_i, \rho, t)\right) - \frac{\partial}{\partial \rho_i}\mathcal{V}(\rho, t).$$

Proposition (Properties of Hamiltonian process)

Assume that a stochastic process X_t on G is a Hamiltonian process. Then it holds that

Proposition (Properties of Hamiltonian process)

Assume that a stochastic process X_t on G is a Hamiltonian process. Then it holds that

1. *(preservation of symplectic form) the symplectic structure on $\mathcal{T}^*\mathcal{P}(G)$ is preserved, i.e.,*

$$\omega_{g(\rho,S)}(g'(\rho,S)\xi, g'(\rho,S)\eta) = \omega_{(\rho,S)}(\xi,\eta),$$

where ω denotes the symplectic form on $\mathcal{T}^*\mathcal{P}(G)$, $\xi,\eta \in \mathcal{T}_{(\rho,S)}(\mathcal{T}^*P(G))$ and $g'(\rho,S)$ is the Jacobi matrix of the Hamiltonian flow g on $\mathcal{T}^*\mathcal{P}(G)$;

Proposition (Properties of Hamiltonian process)

Assume that a stochastic process X_t on G is a Hamiltonian process. Then it holds that

1. *(preservation of symplectic form) the symplectic structure on $\mathcal{T}^*\mathcal{P}(G)$ is preserved, i.e.,*

$$\omega_{g(\rho,S)}(g'(\rho,S)\xi, g'(\rho,S)\eta) = \omega_{(\rho,S)}(\xi,\eta),$$

   where ω denotes the symplectic form on $\mathcal{T}^*\mathcal{P}(G)$, $\xi,\eta \in \mathcal{T}_{(\rho,S)}(\mathcal{T}^*P(G))$ and $g'(\rho,S)$ is the Jacobi matrix of the Hamiltonian flow g on $\mathcal{T}^*\mathcal{P}(G)$;

2. *(conservation of energy) $\mathscr{H}(t) = \mathscr{H}(0)$, if the Hamiltonian $\mathscr{H}$ is independent of t;*

Proposition (Properties of Hamiltonian process)

Assume that a stochastic process X_t on G is a Hamiltonian process. Then it holds that

1. *(preservation of symplectic form) the symplectic structure on $\mathcal{T}^*\mathcal{P}(G)$ is preserved, i.e.,*

$$\omega_{g(\rho,S)}(g'(\rho,S)\xi, g'(\rho,S)\eta) = \omega_{(\rho,S)}(\xi,\eta),$$

   where ω denotes the symplectic form on $\mathcal{T}^*\mathcal{P}(G)$, $\xi,\eta \in \mathcal{T}_{(\rho,S)}(\mathcal{T}^*P(G))$ and $g'(\rho,S)$ is the Jacobi matrix of the Hamiltonian flow g on $\mathcal{T}^*\mathcal{P}(G)$;

2. *(conservation of energy) $\mathscr{H}(t) = \mathscr{H}(0)$, if the Hamiltonian $\mathscr{H}$ is independent of t;*
3. *(conservation of mass) X_t preserves mass, i.e., $\sum_{i=1}^N \rho_i(t) = \sum_{i=1}^N \rho_i(0)$.*

Particle-level properties of Hamiltonian process

Q: We have argued that a Hamiltonian system on $\mathcal{P}(G)$ leads the Hamiltonian process $\{X_t\}$ on G, can we endow any Hamiltonian property such as energy conservation to such process on G?

Particle-level properties of Hamiltonian process

Q: We have argued that a Hamiltonian system on $\mathcal{P}(G)$ leads the Hamiltonian process $\{X_t\}$ on G, can we endow any Hamiltonian property such as energy conservation to such process on G?

A: **Yes**, but in the sense of the **expectation of energy**.

Particle-level properties of Hamiltonian process

Q: We have argued that a Hamiltonian system on $\mathcal{P}(G)$ leads the Hamiltonian process $\{X_t\}$ on G, can we endow any Hamiltonian property such as energy conservation to such process on G?

A: **Yes**, but in the sense of the **expectation of energy**.
Consider the Hamiltonian with specific form (separable + linear potential)

$$\mathscr{H}(\rho, S) = \sum_{i \in V} \sum_{j \in N(i)} \rho_j F_{ji}(S_j - S_i) + \sum_{i \in V} \rho_i V_i.$$

Particle-level properties of Hamiltonian process

Q: We have argued that a Hamiltonian system on $\mathcal{P}(G)$ leads the Hamiltonian process $\{X_t\}$ on G, can we endow any Hamiltonian property such as energy conservation to such process on G?

A: **Yes**, but in the sense of the **expectation of energy**.
Consider the Hamiltonian with specific form (separable + linear potential)

$$\mathscr{H}(\rho, S) = \sum_{i\in V}\sum_{j\in N(i)} \rho_j F_{ji}(S_j - S_i) + \sum_{i\in V} \rho_i V_i.$$

Suppose that $\{X(t)\}$ is associated with the Hamiltonian $\mathscr{H}$. Then one can verify the expectation of energy $\mathbb{E}[H(X(t), S(t))]$ with

$$H(X(t), S(t)) = \sum_{j\in N(X(t))} F_{jX(t)}(S_j(t) - S_{X(t)}(t)) + V_{X(t)}.$$

remains constant as time t evolves.

Table of Contents

Example 1: Optimal transport on graph

Recall the Hamiltonian system (8) derived for OT problem on graph (7) as

$$\frac{d\rho_i}{dt}+\sum_{j\in N(i)}\theta_{ij}(\rho)(S_j-S_i)=0,\quad \frac{dS_i}{dt}+\frac{1}{2}\sum_{j\in N(i)}\frac{\partial\theta_{ij}(\rho)}{\partial\rho_i}(S_i-S_j)^2=0. \tag{12}$$

Example 1: Optimal transport on graph

Recall the Hamiltonian system (8) derived for OT problem on graph (7) as

$$\frac{d\rho_i}{dt}+\sum_{j\in N(i)}\theta_{ij}(\rho)(S_j-S_i)=0,\quad \frac{dS_i}{dt}+\frac{1}{2}\sum_{j\in N(i)}\frac{\partial\theta_{ij}(\rho)}{\partial\rho_i}(S_i-S_j)^2=0. \tag{12}$$

To guarantee the existence of Markov process, we recall (10).

Example 1: Optimal transport on graph

Recall the Hamiltonian system (8) derived for OT problem on graph (7) as

$$\frac{d\rho_i}{dt}+\sum_{j\in N(i)}\theta_{ij}(\rho)(S_j-S_i)=0,\quad \frac{dS_i}{dt}+\frac{1}{2}\sum_{j\in N(i)}\frac{\partial\theta_{ij}(\rho)}{\partial\rho_i}(S_i-S_j)^2=0. \tag{12}$$

To guarantee the existence of Markov process, we recall (10). Common choices of θ such as

- (Arithmetic mean) $\theta_{ij}^A(\rho)=\frac{\rho_i+\rho_j}{2}$;
- (Geometric mean) $\theta_{ij}^G(\rho)=\sqrt{\rho_i\rho_j}$;

will not satisfy (10).

Example 1: Optimal transport on graph

Recall the Hamiltonian system (8) derived for OT problem on graph (7) as

$$\frac{d\rho_i}{dt}+\sum_{j\in N(i)}\theta_{ij}(\rho)(S_j-S_i)=0,\quad \frac{dS_i}{dt}+\frac{1}{2}\sum_{j\in N(i)}\frac{\partial\theta_{ij}(\rho)}{\partial\rho_i}(S_i-S_j)^2=0. \tag{12}$$

To guarantee the existence of Markov process, we recall (10). Common choices of θ such as

- (Arithmetic mean) $\theta_{ij}^A(\rho)=\frac{\rho_i+\rho_j}{2}$;
- (Geometric mean) $\theta_{ij}^G(\rho)=\sqrt{\rho_i\rho_j}$;

will not satisfy (10).

But fortunately, we have a feasible choice:

- (Upwind choice) $\theta_{ij}^U(\rho)=\begin{cases}\rho_j & S_j<S_i\\ \rho_i & S_i<S_j\end{cases}$

Under the Upwind choice, (12) becomes

$$\frac{d\rho_i}{dt} = \sum_{j\in N(i)} \rho_j(S_j - S_i)^- - \rho_i(S_j - S_i)^+, \quad \frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)} ((S_j - S_i)^+)^2 = 0. \tag{13}$$

Here we define the positive part of $x \in \mathbb{R}$ as $(x)^+ = \max\{x, 0\}$, and the negative part as $(x)^- = \max\{-x, 0\}$.

Under the Upwind choice, (12) becomes

$$\frac{d\rho_i}{dt} = \sum_{j\in N(i)} \rho_j(S_j - S_i)^- - \rho_i(S_j - S_i)^+, \quad \frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)} ((S_j - S_i)^+)^2 = 0. \tag{13}$$

Here we define the positive part of $x \in \mathbb{R}$ as $(x)^+ = \max\{x, 0\}$, and the negative part as $(x)^- = \max\{-x, 0\}$.
If we write the first equation as Master equation $d_t\rho_t = \rho_t Q$, then

$$Q_{ji}(S,\rho,t) = 1_{(i,j)\in E}(S_j - S_i)^- = 1_{(i,j)\in E}(S_i - S_j)^+, \; j \neq i$$
$$Q_{ii}(S,\rho,t) = -\sum_{j=1}^N Q_{ij}(t) = -\sum_{j\in N(i)} (S_j - S_i)^+.$$

Under the Upwind choice, (12) becomes

$$\frac{d\rho_i}{dt} = \sum_{j\in N(i)} \rho_j(S_j - S_i)^- - \rho_i(S_j - S_i)^+, \quad \frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)} ((S_j - S_i)^+)^2 = 0. \tag{13}$$

Here we define the positive part of $x \in \mathbb{R}$ as $(x)^+ = \max\{x, 0\}$, and the negative part as $(x)^- = \max\{-x, 0\}$.
If we write the first equation as Master equation $d_t\rho_t = \rho_t Q$, then

$$Q_{ji}(S, \rho, t) = 1_{(i,j)\in E}(S_j - S_i)^- = 1_{(i,j)\in E}(S_i - S_j)^+, \ j \neq i$$
$$Q_{ii}(S, \rho, t) = -\sum_{j=1}^{N} Q_{ij}(t) = -\sum_{j\in N(i)} (S_j - S_i)^+.$$

We verify $Q_{ji}(S, \rho, t) = f_{ij}(S_i - S_j, \rho, t) = 1_{(i,j)\in E}(S_i - S_j)^+ \geq 0$.
Thus (10) is guaranteed and there exists a Markov process associated to Hamiltonian system (13).

We can also verify that $F_{ij}(S_i - S_j, \rho, t) = \frac{1}{2} 1_{(i,j)\in E}((S_i - S_j)^+)^2$, and the Hamiltonian

$$\mathscr{H}(\rho, S) = \sum_{i\in V} \sum_{j\in N(i)} \rho_i F_{ji}(S_j - S_i, \rho, t) = \sum_{i\in V} \sum_{j\in N(i)} \frac{1}{2} \rho_i ((S_j - S_i)^+)^2.$$

We can also verify that $F_{ij}(S_i - S_j, \rho, t) = \frac{1}{2} 1_{(i,j)\in E}((S_i - S_j)^+)^2$, and the Hamiltonian

$$\mathscr{H}(\rho, S) = \sum_{i\in V} \sum_{j\in N(i)} \rho_i F_{ji}(S_j - S_i, \rho, t) = \sum_{i\in V} \sum_{j\in N(i)} \frac{1}{2} \rho_i ((S_j - S_i)^+)^2.$$

Furthermore, the expectation of energy

$$\mathbb{E}_{X_t} \left[\sum_{j\in N(X(t))} \frac{1}{2} ((S_j(t) - S_{X(t)}(t))^+)^2 \right]$$

of the Hamiltonian process $\{X_t\}$ will be conserved for $t \geq 0$.

Example 2: Schrödinger Bridge Problem (SBP) on graph

Background of SBP[1] on $\mathbb{R}^d$

[1] C. Léonard, A survey of the Schrödinger problem and some of its connections with optimal transport, Discrete Contin. Dyn. Syst. 34 (4) (2014)

Example 2: Schrödinger Bridge Problem (SBP) on graph

Background of SBP[1] on $\mathbb{R}^d$

- R: the reference path measure of Brownian motion on $(\mathbb{R}^d)^{[0,1]}$;

[1] C. Léonard, A survey of the Schrödinger problem and some of its connections with optimal transport, Discrete Contin. Dyn. Syst. 34 (4) (2014)

Example 2: Schrödinger Bridge Problem (SBP) on graph

Background of SBP[1] on $\mathbb{R}^d$

- R: the reference path measure of Brownian motion on $(\mathbb{R}^d)^{[0,1]}$;
- P: the path measure of certain stochastic process on $(\mathbb{R}^d)^{[0,1]}$ with fixed marginals at $t = 0, 1$, $P_0 = \rho_a$, $P_1 = \rho_b$.

[1] C. Léonard, A survey of the Schrödinger problem and some of its connections with optimal transport, Discrete Contin. Dyn. Syst. 34 (4) (2014)

Example 2: Schrödinger Bridge Problem (SBP) on graph

Background of SBP[1] on $\mathbb{R}^d$

- R: the reference path measure of Brownian motion on $(\mathbb{R}^d)^{[0,1]}$;
- P: the path measure of certain stochastic process on $(\mathbb{R}^d)^{[0,1]}$ with fixed marginals at $t = 0, 1$, $P_0 = \rho_a$, $P_1 = \rho_b$.
- **SBP**: minimize the relative entropy between P and R

$$\min_P \left\{ \mathcal{H}(P|R) = \int_{(\mathbb{R}^d)^{[0,1]}} \log\left(\frac{dP}{dR}\right) dP \right\}, \; P_0 = \rho_a, P_1 = \rho_b. \tag{14}$$

[1] C. Léonard, A survey of the Schrödinger problem and some of its connections with optimal transport, Discrete Contin. Dyn. Syst. 34 (4) (2014)

Two equivalent formulations of SBP on $\mathbb{R}^d$

Two equivalent formulations of SBP on $\mathbb{R}^d$

- (14) can be reduced to an optimal control problem on $\mathcal{P}(\mathbb{R}^d)$

$$\min\{\mathcal{H}(P|R) : P_0 = \rho_a, P_1 = \rho_b\} - \mathcal{H}(\rho_a|Leb) \tag{15}$$

$$= \min_{\rho, v} \left\{ \int_0^1 \int_{\mathbb{R}^d} \frac{|v_t|^2}{2} \rho_t \, dxdt : (\partial_t - \frac{\Delta}{2})\rho_t + \nabla \cdot (v_t \rho_t) = 0, \right.$$
$$\left. P_0 = \rho_a, P_1 = \rho_b \right\}$$

Two equivalent formulations of SBP on $\mathbb{R}^d$

- (14) can be reduced to an optimal control problem on $\mathcal{P}(\mathbb{R}^d)$

$$\min\{\mathcal{H}(P|R) : P_0 = \rho_a, P_1 = \rho_b\} - \mathcal{H}(\rho_a|Leb) \tag{15}$$
$$= \min_{\rho,v}\Big\{\int_0^1\int_{\mathbb{R}^d}\frac{|v_t|^2}{2}\rho_t\, dxdt : (\partial_t - \frac{\Delta}{2})\rho_t + \nabla\cdot(v_t\rho_t) = 0,$$
$$P_0 = \rho_a, P_1 = \rho_b\Big\}$$

- By replacing $\tilde{v}_t = v_t - \nabla \log \rho_t$, (15) can be casted as

$$\min_{\rho,\tilde{v}}\Big\{\int_0^1\int_{\mathbb{R}^d}\frac{|\tilde{v}_t|^2}{2}\rho_t dx + \frac{1}{8}\mathcal{I}(\rho_t)dt : {\substack{\partial_t\rho_t+\nabla\cdot(\tilde{v}_t\rho_t)=0,\\ P_0=\rho_a, P_1=\rho_b}}\Big\} \tag{16}$$

Here $\mathcal{I}(\rho) = \int |\nabla \log \rho|^2\rho\, dx$ is the *Fisher Information* of ρ.

Optimal solutions to SBP on $\mathbb{R}^d$ as Hamiltonian systems

Optimal solutions to SBP on $\mathbb{R}^d$ as Hamiltonian systems

- optimal solution of (15) leads to

$$
\begin{aligned}
&(\partial_t - \frac{\Delta}{2})\rho + \nabla \cdot (\nabla \phi \rho) = 0, \ \rho(0) = \rho_0, \\
&(\partial_t + \frac{\Delta}{2})\phi + \frac{1}{2}|\nabla \phi|^2 = 0, \ \phi(1) = \log(g_1),
\end{aligned}
$$

it is the Hamiltonian flow of $\mathscr{H}(\rho, \phi) = \int \frac{1}{2}|\nabla \phi|^2 \rho - \nabla \rho \cdot \nabla \phi \, dx$.

Optimal solutions to SBP on $\mathbb{R}^d$ as Hamiltonian systems

- optimal solution of (15) leads to

$$(\partial_t - \frac{\Delta}{2})\rho + \nabla \cdot (\nabla\phi\rho) = 0,\ \rho(0) = \rho_0,$$
$$(\partial_t + \frac{\Delta}{2})\phi + \frac{1}{2}|\nabla\phi|^2 = 0,\ \phi(1) = \log(g_1),$$

  it is the Hamiltonian flow of $\mathscr{H}(\rho,\phi) = \int \frac{1}{2}|\nabla\phi|^2\rho - \nabla\rho\cdot\nabla\phi\ dx$.

- optimal solution of (16) leads to

$$\partial_t\rho + \nabla\cdot(\rho\nabla S) = 0,\ \rho(0) = \rho_0,$$
$$\partial_t S + \frac{1}{2}|\nabla S|^2 = \frac{1}{8}\frac{\delta}{\delta\rho}\mathcal{I}(\rho),\ \phi(1) = \log(g_1),$$

  it is the Hamiltonian flow of $\mathscr{H}(\rho,S) = \int \frac{1}{2}|\nabla S|^2\rho\ dx - \frac{1}{8}\mathcal{I}(\rho)$.

Optimal solutions to SBP on $\mathbb{R}^d$ as Hamiltonian systems

- optimal solution of (15) leads to

$$(\partial_t - \frac{\Delta}{2})\rho + \nabla\cdot(\nabla\phi\rho) = 0,\ \rho(0) = \rho_0,$$
$$(\partial_t + \frac{\Delta}{2})\phi + \frac{1}{2}|\nabla\phi|^2 = 0,\ \phi(1) = \log(g_1),$$

  it is the Hamiltonian flow of $\mathscr{H}(\rho,\phi) = \int \frac{1}{2}|\nabla\phi|^2\rho - \nabla\rho\cdot\nabla\phi\ dx$.
- optimal solution of (16) leads to

$$\partial_t\rho + \nabla\cdot(\rho\nabla S) = 0,\ \rho(0) = \rho_0,$$
$$\partial_t S + \frac{1}{2}|\nabla S|^2 = \frac{1}{8}\frac{\delta}{\delta\rho}\mathcal{I}(\rho),\ \phi(1) = \log(g_1),$$

  it is the Hamiltonian flow of $\mathscr{H}(\rho,S) = \int \frac{1}{2}|\nabla S|^2\rho\ dx - \frac{1}{8}\mathcal{I}(\rho)$.
- (ρ,ϕ) and (ρ,S) are related via the symplectic transform τ on $\mathcal{T}^*\mathcal{P}(\mathbb{R}^d)$, i.e., $(\rho,S) = \tau(\rho,\phi) = (\rho,\phi-\log\rho)$.

SBP on graph

Two ways to discretize SBP:

1. Discretize from (14);

SBP on graph

Two ways to discretize SBP:

1. Discretize from (14);
2. Discretize from (16).

SBP on graph

Two ways to discretize SBP:

1. Discretize from (14);
2. Discretize from (16).

In our research, they will lead to **different** Hamiltonian processes on G even though (14) and (16) are equivalent in continuous space $\mathbb{R}^d$.

First way of discretization:

Based on entropy-minimization formulation (14),

- Consider R as the reference path measure on $G^{[0,1]}$ whose marginal $\{\widetilde{\rho}_t\}$ solves $d_t\widetilde{\rho}_i = \sum_{j\in N(i)} m_{ji}^t\widetilde{\rho}_j - m_{ij}^t\widetilde{\rho}_i$.

[1] C. Léonard, Girsanov theory under a finite entropy condition, in: Séminaire de Probabilités XLIV, in: Lecture Notes in Math., vol. 2046, Springer, Heidelberg, 2012

First way of discretization:

Based on entropy-minimization formulation (14),

- Consider R as the reference path measure on $G^{[0,1]}$ whose marginal $\{\widetilde{\rho}_t\}$ solves $d_t\widetilde{\rho}_i = \sum_{j\in N(i)} m^t_{ji}\widetilde{\rho}_j - m^t_{ij}\widetilde{\rho}_i$.
- Consider P as the path measure of certain stochastic process on $G^{[0,1]}$ whose marginal $\rho_0 = \rho_a$, $\rho_1 = \rho_b$ are fixed, and $\{\rho_t\}$ solves $d_t\rho_i = \sum_{j\in N(i)} \widehat{m}^t_{ji}\rho_j - \widehat{m}^t_{ij}\rho_i$.

[1] C. Léonard, Girsanov theory under a finite entropy condition, in: Séminaire de Probabilités XLIV, in: Lecture Notes in Math., vol. 2046, Springer, Heidelberg, 2012

First way of discretization:

Based on entropy-minimization formulation (14),

- Consider R as the reference path measure on $G^{[0,1]}$ whose marginal $\{\widetilde{\rho}_t\}$ solves $d_t\widetilde{\rho}_i = \sum_{j\in N(i)} m_{ji}^t\widetilde{\rho}_j - m_{ij}^t\widetilde{\rho}_i$.
- Consider P as the path measure of certain stochastic process on $G^{[0,1]}$ whose marginal $\rho_0 = \rho_a$, $\rho_1 = \rho_b$ are fixed, and $\{\rho_t\}$ solves $d_t\rho_i = \sum_{j\in N(i)} \widehat{m}_{ji}^t\rho_j - \widehat{m}_{ij}^t\rho_i$.
- One can compute the relative entropy $\mathcal{H}(P|R)$ as[1]

$$\mathcal{H}(P|R) = \int_0^1 \sum_{i\in V} \rho(i,t) \sum_{j\in N(i)} \left(\frac{\widehat{m}_{ij}^t}{m_{ij}^t} \log\left(\frac{\widehat{m}_{ij}^t}{m_{ij}^t}\right) - \frac{\widehat{m}_{ij}^t}{m_{ij}^t} + 1 \right) m_{ij}^t \, dt.$$

[1] C. Léonard, Girsanov theory under a finite entropy condition, in: Séminaire de Probabilités XLIV, in: Lecture Notes in Math., vol. 2046, Springer, Heidelberg, 2012

First way of discretization:

Then the Schrödinger Bridge Problem on G is formulated as

$$\min_{\widehat{m}^t \geq 0} \left\{ \int_0^1 \sum_{i \in V} \rho(i,t) \sum_{j \in N(i)} \left(\frac{\widehat{m}_{ij}^t}{m_{ij}^t} \log\left(\frac{\widehat{m}_{ij}^t}{m_{ij}^t}\right) - \frac{\widehat{m}_{ij}^t}{m_{ij}^t} + 1 \right) m_{ij}^t \, dt \right\} \tag{17}$$

$$\text{subject to:} \quad \frac{d}{dt}\rho(i,t) = \sum_{j \in N(i)} \widehat{m}_{ji}^t \rho_j - \widehat{m}_{ij}^t \rho_i \quad \rho(\cdot,0) = \rho_a, \ \rho(\cdot,1) = \rho_b.$$

First way of discretization:

Then the Schrödinger Bridge Problem on G is formulated as

$$\min_{\widehat{m}^t \geq 0} \left\{ \int_0^1 \sum_{i \in V} \rho(i,t) \sum_{j \in N(i)} \left(\frac{\widehat{m}_{ij}^t}{m_{ij}^t} \log\left(\frac{\widehat{m}_{ij}^t}{m_{ij}^t}\right) - \frac{\widehat{m}_{ij}^t}{m_{ij}^t} + 1 \right) m_{ij}^t \, dt \right\} \tag{17}$$

$$\text{subject to:} \quad \frac{d}{dt}\rho(i,t) = \sum_{j \in N(i)} \widehat{m}_{ji}^t \rho_j - \widehat{m}_{ij}^t \rho_i \quad \rho(\cdot,0) = \rho_a, \;\; \rho(\cdot,1) = \rho_b.$$

By introducing Lagrange multiplier S, the KKT condition yields to the system

$$\begin{aligned} \frac{d}{dt}\rho_i &= \sum_{j \in N(i)} -e^{S_j - S_i} m_{ij}^t \rho_i + e^{S_i - S_j} m_{ji}^t \rho_j, \\ \frac{d}{dt} S_i &= -\sum_{j \in N(i)} (e^{S_j - S_i} - 1) m_{ij}^t. \end{aligned} \tag{18}$$

First way of discretization:

First way of discretization:

One can verify that (18) is a Hamiltonian system with the Hamiltonian

$$\mathscr{H}(\rho, S, t) = \sum_{i \in V} \sum_{j \in N(i)} (\exp(S_j - S_i) - 1) m_{ij}^t \rho_i.$$

Furthermore, we can verify the transition rate

$$Q_{ji}(S, \rho, t) = f_{ij}(S_i - S_j, \rho, t) = e^{S_i - S_j} m_{ji}^t \geq 0.$$

We can construct a nonlinear Markov process asscoiated to the solution (18) of Schrödinger Bridge problem (17).

Second way of discretization:

Second way of discretization:

Based on action-minimizing formulation (16), We consider the optimal control problem

$$\min_{\rho,v}\left\{\int_0^1 (\langle v,v\rangle_{\theta^U(\rho)}+\frac{1}{8}\mathcal{I}_G(\rho))dt\right\}, \tag{19}$$

$$\partial\rho+\mathrm{div}_G^{\theta^U}(\rho v)=0,\ \rho(\cdot,0)=\rho_a,\ \rho(\cdot,1)=\rho_b.$$

Recall

$$\langle v,v\rangle_{\theta(\rho)}=\frac{1}{2}\sum_{(j,l)\in E}\theta_{jl}(\rho)v_{jl}^2,\quad (\mathrm{div}_G^{\theta}(\rho v))_j=-\sum_{l\in N(j)}\theta_{jl}(\rho)v_{jl},$$

defined as before. We directly discretize Fisher Information $\mathcal{I}(\rho)$ and define

$$\mathcal{I}_G(\rho)=\frac{1}{2}\sum_{(i,j)\in E}(\log(\rho_i)-\log(\rho_j))^2\widetilde{\theta}_{ij}(\rho),$$

where $\widetilde{\theta}$ is some weight function, not necessarily equal to θ^U before.

Second way of discretization:

Second way of discretization:

Similar to our previous treatments, recall we are using the upwind weight θ^U, we can verify the optimal solution is solved by the following Hamiltonian system

$$
\begin{aligned}
\frac{d\rho_i}{dt} &= \sum_{j\in N(i)} \rho_j(S_j - S_i)^- - \rho_i(S_j - S_i)^+, \\
\frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)} ((S_j - S_i)^+)^2 &= \frac{1}{8}\frac{\partial}{\partial\rho_i}\mathcal{I}_G(\rho).
\end{aligned}
\tag{20}
$$

Second way of discretization:

Similar to our previous treatments, recall we are using the upwind weight θ^U, we can verify the optimal solution is solved by the following Hamiltonian system

$$\begin{aligned} \frac{d\rho_i}{dt} &= \sum_{j\in N(i)} \rho_j(S_j - S_i)^- - \rho_i(S_j - S_i)^+, \\ \frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)} ((S_j - S_i)^+)^2 &= \frac{1}{8}\frac{\partial}{\partial\rho_i}\mathcal{I}_G(\rho). \end{aligned} \tag{20}$$

It is not hard to verify the Hamiltonian of (20) is

$$\mathscr{H}(\rho, S) = \frac{1}{2}\sum_{i\in V}\sum_{j\in N(i)} \rho_i((S_j - S_i)^+)^2 - \frac{1}{8}\mathcal{I}_G(\rho). \tag{21}$$

Second way of discretization:

Similar to our previous treatments, recall we are using the upwind weight θ^U, we can verify the optimal solution is solved by the following Hamiltonian system

$$\begin{aligned} \frac{d\rho_i}{dt} &= \sum_{j\in N(i)} \rho_j(S_j - S_i)^- - \rho_i(S_j - S_i)^+, \\ \frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)}((S_j - S_i)^+)^2 &= \frac{1}{8}\frac{\partial}{\partial\rho_i}\mathcal{I}_G(\rho). \end{aligned} \tag{20}$$

It is not hard to verify the Hamiltonian of (20) is

$$\mathscr{H}(\rho, S) = \frac{1}{2}\sum_{i\in V}\sum_{j\in N(i)} \rho_i((S_j - S_i)^+)^2 - \frac{1}{8}\mathcal{I}_G(\rho). \tag{21}$$

By aforementioned argument regarding upwind θ^U, we can also associated (20) with a nonlinear Markov process as the solution to Schrödinger Bridge problem (19).

Comparison of two SBPs on graph

	Entropy-minimization SBP	Action-minimization SBP
Origin	Derived from (17)	Derived from (16)
Hamiltonian system	$\frac{d}{dt}\rho_t = \rho_t Q(S_t, t)$ $\frac{d}{dt}S_i = -\sum_{j\in N(i)}(e^{S_j - S_i} - 1)m_{ij}^t$	$\frac{d}{dt}\rho_t = \rho_t Q(S_t)$ $\frac{dS_i}{dt} + \frac{1}{2}\sum_{j\in N(i)}((S_j - S_i)^+)^2 = \frac{1}{8}\frac{\partial}{\partial \rho_i}\mathcal{I}_G(\rho)$
$\mathscr{H}$	$\sum_{i\in V}\sum_{j\in N(i)}(\exp(S_j - S_i) - 1)m_{ij}^t\rho_i$	$\frac{1}{2}\sum_{i\in V}\sum_{j\in N(i)}\rho_i((S_j - S_i)^+)^2 - \frac{1}{8}\mathcal{I}_G(\rho)$
$Q_{ji}, j \neq i$	$e^{S_i - S_j} m_{ji}^t \geq 0$ Hamiltonian process exists	$(S_i - S_j)^+$ Hamiltonian process exists
Reference R	stochastic process induced by linear generator $Q = \{m_{ij}^t\}$	stochastic process induced by nonlinear generator related to the Fisher Information $\mathcal{I}_G(\rho)$

- For more discussion on the *periodicity* of Schrödinger Bridge problems, please check our work[1].

[1] J. Cui, S. Liu, H. Zhou, What is a stochastic Hamiltonian process on finite graph? An optimal transport answer, Journal of Differential Equations, 2021

Conclusion & Future direction

In this work, we introduce a novel definition and theoretical framework for Hamiltonian process on graph.

Conclusion & Future direction

In this work, we introduce a novel definition and theoretical framework for Hamiltonian process on graph.

Possible future research directions

Conclusion & Future direction

In this work, we introduce a novel definition and theoretical framework for Hamiltonian process on graph.

Possible future research directions

1. Well posedness & long time existence of the proposed Hamiltonian process;

Conclusion & Future direction

In this work, we introduce a novel definition and theoretical framework for Hamiltonian process on graph.

Possible future research directions

1. Well posedness & long time existence of the proposed Hamiltonian process;
2. Consistency between the proposed Hamiltonian process on graph and Hamiltonian dynamics in continuous space;

Conclusion & Future direction

In this work, we introduce a novel definition and theoretical framework for Hamiltonian process on graph.

Possible future research directions

1. Well posedness & long time existence of the proposed Hamiltonian process;
2. Consistency between the proposed Hamiltonian process on graph and Hamiltonian dynamics in continuous space;
3. Optimal mean-field control on graph.

Thank you!

Welcome to any comments or questions.