Mean field games via probability manifold II

Wuchen Li

IPAM summer school, 2018.

Introduction

Consider a nonlinear Schrödinger equation

$$h\mathrm{i}\frac{\partial}{\partial t}\Psi = -h^2\frac{1}{2}\Delta\Psi + \Psi V(x) + \Psi\int_{\mathbb{R}^d} W(x,y)|\Psi(y)|^2 dy\ .$$

- The unknown $\Psi(t,x)$ is a complex function, $x \in \mathbb{R}^d$, $\mathrm{i} = \sqrt{-1}$, $|\cdot|$ is the modulus of a complex number, h is a positive constant;
- V is a linear potential, W is a mean field interaction potential with $W(x,y) = W(y,x)$.

There are many important properties of the equation, e.g. conservation of total mass, total energy, etc. It is a Hamiltonian system.

History Remark

Brownian motion (Einstein 1905)

$\downarrow$

Schrödinger equation (Schrödinger 1925)

$\downarrow$

Nelson process (Nelson 1966)

$$\underbrace{\hspace{10em}}$$

Optimal transport + Hamiltonian system

Introduction

Optimal transport + Hamiltonian system:

- Related to Schrödinger equations (Nelson, Carlen);
- Related to Mean field games (Larsy, Lions, Gangbo);
- Related to weak KAM theory (Evans);
- Related to 2-Wasserstein metric (Brenier, Villani, Ambrosio);
- Related to Schrödinger bridge problems (Yause, Chen, Georgiou, Pavon, Conforti, Leonard, Flavien).

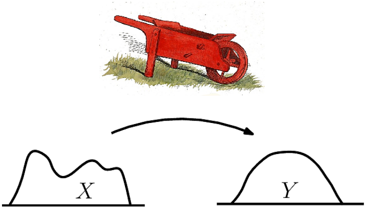
Motivation

Based on optimal transport (OT) and Nelson's idea, we plan to propose a **discrete Schrödinger equation**. Later, we shall show that the derived equation has the following properties:

Method[1]	TSSP	CNFD	ReFD	TSFD	OT+Nelson
Time Reversible	Yes	Yes	Yes	Yes	Yes
Time Transverse Invariant	Yes	No	No	Yes	Yes
Mass Conservation	Yes	Yes	Yes	Yes	Yes
Energy Conservation	No	Yes	Yes	No	Yes
Dispersion Relation	Yes	No	No	Yes	Yes

[1] Antoinea et al (2013), where TSSP: Time Splitting SPectral; CNFD: Crank-Nicolson Finite Difference; ReFD: Relaxation Finite Difference; TSFD: Time splitting Finite Difference.

Optimal transport



The optimal transport problem was first introduced by Monge in 1781, relaxed by Kantorovich by 1940.

It introduces a particular metric on probability set, which can be viewed under various angles:

- Mapping: Monge-Ampére equation ;
- Linear programming ;
- Geometry: Fluid dynamics .

In this talk, we focus on its geometric formulation.

Probability Manifold

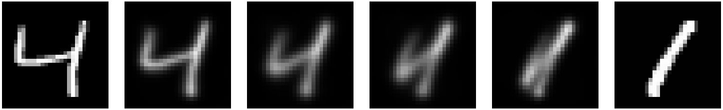
The problem has an important variational formulation (Benamou-Briener 2000):

$$W(\rho^0, \rho^1)^2 := \inf_v \int_0^1 \mathbb{E} v_t^2 \, dt \ ,$$

where $\mathbb{E}$ is the expectation operator and the infimum runs over all vector field v, such that

$$\dot{X}_t = v_t \ , \quad X_0 \sim \rho^0 \ , \quad X_1 \sim \rho^1 \ .$$

Under this metric, the probability set has a Riemannian geometry structure (Lafferty 1988).



Brownian motion and Optimal transport

The gradient flow of (negative) Boltzmann-Shannon entropy

$$\int_{\mathbb{R}^d} \rho(x) \log \rho(x) dx$$

w.r.t. optimal transport distance is:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \big(\rho \nabla \log \rho\big) = \Delta \rho \ .$$

This geometric understanding will be the key for Schrödinger equation.

Nelson's approach

Nelson in 1966 proposed a slightly different problem of optimal transport distance

$$\inf_{b_t}\{\int_0^1 \frac{1}{2}\mathbb{E}\dot{X}_t^2 dt \ : \ \dot{X}_t = b_t + \sqrt{h}\dot{B}_t, \quad X(0) \sim \rho^0, \quad X(1) \sim \rho^1\} \ ,$$

where B_t is a standard Brownian motion in $\mathbb{R}^d$ and h is a small positive constant.

Although Nelson's problem and Schrödinger equation look very different from each other, it can be shown that Schrödinger equation is a critical point of the above variation problem.

Nelson's approach

Rewrite Nelson's problem in terms of densities, i.e. represent X_t by its density ρ:

$$\Pr(X_t \in A) = \int_A \rho(t,x)dx \ .$$

Consider

$$\inf_b \int_0^1 \int \frac{1}{2}[b^2 - hb \cdot \nabla \log \rho]\rho dx \ dt \ ,$$

where the infimum is among all drift vector fields $b(t,x)$, such that

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho b) = \frac{h}{2}\Delta\rho \ , \quad \rho(0) = \rho^0 \ , \quad \rho(1) = \rho^1 \ .$$

Change of variable

The key of Nelson's idea comes from the change of variables

$$v = b - \frac{h}{2} \nabla \log \rho \ .$$

Substituting the v into Nelson's problem, the problem is arrived at

$$\inf_{v} \ \{ \int_0^1 \int_{\mathbb{R}^d} \frac{1}{2} v^2 \rho dx - \frac{h^2}{8} \mathcal{I}(\rho) \ dt \ : \ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0 \} \ ,$$

where the functional

$$\mathcal{I}(\rho) = \int (\nabla \log \rho(x))^2 \rho(x) dx \ ,$$

is called Fisher information. It is worth noting that $\mathcal{I}$ is a key concept in physics and modeling (Frieden 2004).

Critical points

Following the Euler-Lagrange equation in probability set, the critical point of Nelson problem satisfies a pair of equations

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \nabla \cdot (\rho \nabla S) = 0 \\ \dfrac{\partial S}{\partial t} + \dfrac{1}{2}(\nabla S)^2 = -\dfrac{\delta}{\delta \rho(x)}[\dfrac{h^2}{8}\mathcal{I}(\rho)] \end{cases}$$

where $\frac{\delta}{\delta \rho(x)}$ is the L^2 first variation, the first equation is a continuity equation while the second one is a Hamilton-Jacobi equation. Define

$$\Psi(t,x) = \sqrt{\rho(t,x)}e^{\frac{\mathrm{i}S(t,x)}{h}},$$

then Ψ satisfies the linear Schrödinger equation

$$\mathrm{i}h\frac{\partial}{\partial t}\Psi = -\frac{h^2}{2}\Delta\Psi \ .$$

This derivation is also true for other potential energies.

Goals

Following the geometry introduced by optimal transport, we plan to establish a Schrödinger equation on a graph.

Why on graphs?

- Numerics and modeling for nonlinear Schrödinger equations, Mean Field Games;
- Population games;
- Computation of optimal transport metric.

Basic settings

Graph with finite vertices

$$G = (V, E), \quad V = \{1, \cdots, n\}, \quad \text{E is the edge set;}$$

Probability set

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

Linear and interaction potential energies:

$$\mathcal{V}(\rho) = \sum_{i=1}^n V_i \rho_i, \quad \mathcal{W}(\rho) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n W_{ij} \rho_i \rho_j,$$

where V_i, W_{ij} are constants with $W_{ij} = W_{ji}$.

Definition I

We plan to find the discrete analog of Nelson's problem.

First, it is natural to define a *vector field on a graph*

$$v = (v_{ij})_{(i,j)\in E}, \quad \text{satisfying } v_{ij} = -v_{ji}.$$

Next, we define a *divergence operator* of a vector field v on a graph w.r.t a probability measure ρ (Chow, Li, Huang, Zhou 2012):

$$\nabla \cdot (\rho v).$$

Definition II

Let

$$\theta_{ij}(\rho) = \frac{\rho_i + \rho_j}{2} \ .$$

We define an *inner product* of two vector fields v^1, v^2:

$$(v^1, v^2)_\rho := \frac{1}{2} \sum_{(i,j)\in E} v^1_{ij} v^2_{ij} \theta_{ij}(\rho);$$

and a *divergence* of a vector field v at $\rho \in \mathcal{P}(G)$:

$$\mathrm{div}_G(\rho v) := \big(- \sum_{j \in N(i)} v_{ij} \theta_{ij}(\rho) \big)_{i=1}^{n}.$$

Optimal transport distance on a graph

For any ρ^0, $\rho^1 \in \mathcal{P}_o(G)$, consider the optimal transport distance (also named Wassersetin metric) by

$$W(\rho^0, \rho^1)^2 := \inf_v \{ \int_0^1 (v, v)_\rho dt \ : \ \frac{d\rho}{dt} + \mathrm{div}_G(\rho v) = 0, \ \rho(0) = \rho^0, \ \rho(1) = \rho^1 \}.$$

$(\mathcal{P}_o(G), W)$ forms a Riemannian manifold.

Fisher information on a graph

The gradient flow of the Shannon entropy $\mathcal{S}(\rho) = \sum_{i=1}^{n} \rho_i \log \rho_i$ in $(\mathcal{P}(G), \mathcal{W})$ is the diffusion process on a graph:

$$\frac{d\rho}{dt} = -\mathrm{grad}_W \mathcal{S}(\rho) = \mathrm{div}_G(\rho \nabla_G \log \rho) \ .$$

The dissipation of entropy defines the Fisher information on a graph:

$$\mathcal{I}(\rho) = (\mathrm{grad}_W \mathcal{S}(\rho), \mathrm{grad}_W \mathcal{S}(\rho))_\rho = \frac{1}{2} \sum_{(i,j)\in E} (\log \rho_i - \log \rho_j)^2 \theta_{ij}(\rho) \ .$$

Many interesting topics have been extracted from this observation. E.g. entropy dissipation, Log-Sobolev inequalities, Ricci curvature, Yano formula (Annals of Mathematics, 1952, 7 pages).

Discrete Nelson's problem

We introduce Nelson's problem on a graph:

$$\inf_b \int_0^1 \frac{1}{2}(b,b)_\rho - \frac{1}{2}h(\nabla_G \log\rho, b)_\rho - \mathcal{V}(\rho) - \mathcal{W}(\rho)dt,$$

where the infimum is taken among all vector fields b on G, such that

$$\frac{d\rho}{dt} + \mathrm{div}_G(\rho(b - \frac{h}{2}\nabla_G \log\rho)) = 0\ ,\ \rho(0) = \rho^0,\ \rho(1) = \rho^1\ .$$

Derivation

From the change of variables $v = b - \frac{h}{2}\nabla_G \log \rho$, Nelson's problem on a graph can be written as

$$\inf_v \int_0^1 \frac{1}{2}(v, v)_\rho - \frac{h^2}{8}\mathcal{I}(\rho) - \mathcal{V}(\rho) - \mathcal{W}(\rho) dt$$

where the infimum is taken among all discrete vector fields v, such that

$$\frac{d\rho}{dt} + \mathrm{div}_G(\rho v) = 0 \ , \ \rho(0) = \rho^0, \ \rho(1) = \rho^1 \ .$$

Hodge decomposition on graphs

Consider a Hodge decomposition on a graph

$$v = \underset{\text{Gradient}}{\nabla_G S} + \underset{\text{Divergence free}}{u}$$

where the divergence free on a graph means $\mathrm{div}_G(\rho u) = 0$.

Lemma

The discrete Nelson's problem is equivalent to

$$\inf_S \int_0^1 \frac{1}{2}(\nabla_G S, \nabla_G S)_\rho - \frac{h^2}{8}\mathcal{I}(\rho) - \mathcal{V}(\rho) - \mathcal{W}(\rho) dt,$$

where the critical point is taken among all discrete potential vector fields $\nabla_G S$*, such that*

$$\frac{d\rho}{dt} + div_G(\rho \nabla_G S) = 0 \ , \ \rho(0) = \rho^0, \ \rho(1) = \rho^1 \ .$$

Critical points

Applying Euler-Lagrange equation, the solution of Nelson's problem on a graph satisfies an ODE system:

$$\begin{cases} \dfrac{d\rho_i}{dt} + \displaystyle\sum_{j \in N(i)} (S_j - S_i)\theta_{ij}(\rho) = 0 \\ \dfrac{dS_i}{dt} + \dfrac{1}{2} \displaystyle\sum_{j \in N(i)} (S_i - S_j)^2 \dfrac{\partial}{\partial \rho_i}\theta_{ij}(\rho) = -\dfrac{\partial}{\partial \rho_i}[\dfrac{h^2}{8}\mathcal{I}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)] \end{cases}$$

where the first equation is the *continuity equation* on a graph while the second one is the *Hamilton-Jacobi equation* on a graph.

Schrödinger equation on a graph

Denote two real value functions $\rho(t)$, $S(t)$ by

$$\Psi(t) = \sqrt{\rho(t)}e^{\frac{\sqrt{-1}S(t)}{h}} \ .$$

Theorem

Given a graph $G = (V, E)$, a real constant vector $(V_i)_{i=1}^n$ and symmetric matrix $(W_{ij})_{1\le i,j\le n}$. Then every critical point of Nelson problem on the graph satisfies

$$\begin{aligned} h\sqrt{-1}\frac{d\Psi_i}{dt} = \frac{h^2}{2}\Psi_i\{\sum_{j\in N(i)}(\log\Psi_i - \log\Psi_j)\frac{\theta_{ij}}{|\Psi_i|^2} \\ + \sum_{j\in N(i)}|\log\Psi_i - \log\Psi_j|^2\frac{\partial\theta_{ij}}{\partial|\Psi_i|^2}\} \\ + \Psi_i V_i + \Psi_i\sum_{i=1}^n W_{ij}|\Psi_j|^2. \end{aligned}$$

Discrete Laplacian with Hamiltonian structure

We propose a new interpolation of Laplacian operator on a graph

$$\Delta_G \Psi|_i := -\Psi_i \{ \sum_{j \in N(i)} (\log \Psi_i - \log \Psi_j) \frac{\theta_{ij}}{|\Psi_i|^2} + \sum_{j \in N(i)} |\log \Psi_i - \log \Psi_j|^2 \frac{\partial \theta_{ij}}{\partial |\Psi_i|^2} \} .$$

In fact, it is not hard to show that this is consistent with the Laplacian in continuous setting:

$$\Delta \Psi = \Psi(\frac{1}{|\Psi|^2} \nabla \cdot (|\Psi|^2 \nabla \log \Psi) - |\nabla \log \Psi|^2) .$$

What are the **benefits** from this nonlinear interpolation?

Properties

Theorem

Given a graph (V, E) and an initial condition Ψ^0 (complex vector) with positive modulus. There exists a unique solution of Schrödinger equation on the graph for all $t > 0$. Moreover, the solution $\Psi(t)$

(i) *conserves the total mass;*

(ii) *conserves the total energy;*

(iii) *matches the stationary solution (Ground state);*

(iv) *is time reversible;*

(v) *is time transverse invariant.*

Proof of (i) and (ii)

We obtain a Hamiltonian system on the probability space $\mathcal{P}(G)$ w.r.t the discrete optimal transport metric.

$$\frac{d}{dt}\begin{pmatrix}\rho \\ S\end{pmatrix} = \begin{pmatrix}0 & -I \\ I & 0\end{pmatrix}\begin{pmatrix}\frac{\partial}{\partial\rho}\mathcal{H} \\ \frac{\partial}{\partial S}\mathcal{H}\end{pmatrix} \ ,$$

where $I \in \mathbb{R}^{n\times n}$ is the identity matrix and $\mathcal{H}$ is the Hamiltonian:

$$\mathcal{H}(\rho, S) = \frac{1}{2}(\nabla_G S, \nabla_G S)_\rho + \frac{h^2}{8}\mathcal{I}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho) \ .$$

Two points Schrödinger equation

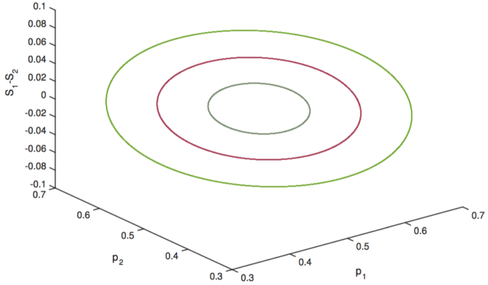


FIGURE 1. The phase portrait of $(\rho_1(t), \rho_2(t), S_1(t) - S_2(t))$ with different initial conditions.

Example: Ground state

Compute the ground state via

$$\min_{\rho \in \mathcal{P}(G)} \quad \frac{h^2}{8}\mathcal{I}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho) \ .$$

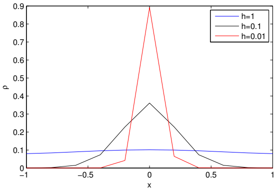


Figure: The plot of ground state's density function. The blue, black, red curves represents $h = 1,\ 0.1,\ 0.01$, respectively.

Discussion

In this talk, we introduce a Schrödinger equation on a graph, which has many dynamical properties. Here the discrete Fisher information plays the main effect. From it, we show that the equation

- exists a unique solution for all time;
- matches the stationary solution.

The discrete Fisher information has been successfully used in Schrödinger equations, computation of optimal transport metric, population games and elsewhere.

Main references

Edward Nelson
Derivation of the Schrödinger Equation from Newtonian Mechanics, 1966.

B. Frieden
Science from Fisher Information: A Unification, 2004.

Shui-Nee Chow, Wuchen Li and Haomin Zhou
Entropy dissipation of Fokker-Planck equations on finite graphs, 2017.

Shui-Nee Chow, Wuchen Li and Haomin Zhou
Schröodinger equation on finite graphs via optimal transport, 2017.

Wuchen Li, Penghang Yin and Stanley Osher.
Computations of optimal transport distance with Fisher information regularization, 2017.