A discrete Schrödinger equation via optimal transport on graphs ☆

Shui-Nee Chow [a], Wuchen Li [b,*], Haomin Zhou [a]

[a] *School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA*
[b] *Department of Mathematics, UCLA, Los Angeles, CA 90095, USA*

ARTICLE INFO

ABSTRACT

In 1966, Edward Nelson presented an interesting derivation of the Schrödinger equation using Brownian motion. Recently, this derivation is linked to the theory of optimal transport, which shows that the Schrödinger equation is a Hamiltonian system on the probability density manifold equipped with the Wasserstein metric. In this paper, we consider similar matters on a finite graph. By using discrete optimal transport and its corresponding Nelson's approach, we derive a discrete Schrödinger equation on a finite graph. The proposed system is quite different from the commonly referred discretized Schrödinger equations. It is a system of nonlinear ordinary differential equations (ODEs) with many desirable properties. Several numerical examples are presented to illustrate the properties.

1. Introduction

The nonlinear Schrödinger equation (NLS) given in the form of

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

plays vital roles in many areas in physical sciences [2,27]. The unknown $\Psi(t,x)$ is a complex wave function for $x \in \mathbb{R}^d$, $h > 0$ is the Planck constant, and $\mathbb{V}(x)$ and $\mathbb{W}(x,y)$ are real valued functions, referred as linear and interaction potentials respectively. Although the NLS provides accurate predictions to various physical phenomena, its formulation is very different from the classical mechanics, and it cannot be easily interpreted by the Newton's law.

To bridge the difference, Edward Nelson provided an compelling approach in 1966 [23]. He derived the NLS by means of classical mechanics in conjunction with variational principles and stochastic diffusion processes [3,16]. To better understand his idea, we recall that the NLS (1) has a fluid dynamics formulation, named Madelung system [19]. By introducing a change of variables, $\Psi(t,x) = \sqrt{\rho(t,x)}e^{iS(t,x)/h}$, one can rewrite (1) as

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \nabla \cdot (\rho \nabla S) = 0 \; ; \\ \dfrac{\partial S}{\partial t} + \dfrac{1}{2}(\nabla S)^2 + \dfrac{h^2}{8}\dfrac{\delta}{\delta\rho(x)}\mathcal{I}(\rho) + \mathbb{V}(x) + \displaystyle\int_{\mathbb{R}^d}\mathbb{W}(x,y)\rho(t,y)dy = 0 \; , \end{cases} \quad (2)$$

where ρ, S are unknown real valued functions, $\frac{\delta}{\delta\rho(x)}$ is the L^2 first variation operator, and $\mathcal{I}(\rho) = \int_{\mathbb{R}^d}(\nabla \log \rho(x))^2\rho(x)dx$ is the Fisher information [14]. Nelson constructed a Lagrangian in the space of probability density functions, and then used calculus of variation to derive (2). We shall give a brief review on his approach in Section 3. Readers can find more details in [24]. Recently, Nelson's approach is linked to the framework of optimal transport theory [6,28], which has been developed in the past few decades [1,4,15,28]. The theory shows that the probability density space equipped with the optimal transport distance, also known as the Wasserstein metric, becomes a Riemannian manifold, and the NLS is a Hamiltonian system on this density manifold [16,26].

In this paper, we consider similar matters in discrete spaces, such as finite graphs. There are two reasons motivating us to conduct this investigation. On one hand, Nelson's derivation is based on a variational principle, which makes his approach more attractive. However, although different formulations of the NLS on graphs have been introduced in physics and mathematics [5,9,12,13,25], not much is known through Nelson's approach, mainly because the theory of discrete optimal transport has not been seriously explored until the past few years [10,18,20]. On the other hand, most of the discrete formulations for the NLS, especially those defined on lattices, are obtained by discretizations of the continuous NLS. Some important properties, such as conservation of energy, or dispersion

relation, or time transverse (gauge) invariant, can be lost due to the discretizations. As reported in a recent survey on numerical methods for the NLS [2], none of the commonly used schemes has all those features simultaneously. We would like to examine whether Nelson's approach can provide a systematic strategy for constructing the discrete NLS on general graphs in order to retain those desirable properties. Our investigation confirms this assertion.

We follow the settings given in [11,17] to derive the discrete NLS, which turns out to be a system of ordinary differential equations. The derivation utilizes the optimal transport distance and the Fokker–Planck equation on a graph. The main results are sketched here.

Consider a weighted finite graph $G = (V, E, \omega)$, where V is the vertex set, E is the edge set, and $\omega_{jl} \in \omega$ the weight of edge $(j, l) \in E$ satisfies $\omega_{lj} = \omega_{jl} > 0$. We assume that G is undirected, contains no self loops or multiple edges. Given a linear potential $\mathbb{V}_j$ on each note j and an interactive potential $\mathbb{W}_{jl}$, with $\mathbb{W}_{lj} = \mathbb{W}_{jl}$, for any two nodes $(j, l) \in E$. Nelson's approach leads to the following ODEs:

$$\begin{cases} \dfrac{d\rho_j}{dt} + \displaystyle\sum_{l \in N(j)} \omega_{jl}(S_j - S_l)\theta_{jl}(\rho) = 0 \; ; \\[4mm] \dfrac{dS_j}{dt} + \dfrac{1}{2} \displaystyle\sum_{l \in N(j)} \omega_{jl}(S_j - S_l)^2 \dfrac{\partial \theta_{jl}}{\partial \rho_j} + \dfrac{h^2}{8} \dfrac{\partial}{\partial \rho_j} \mathcal{I}(\rho) + \mathbb{V}_j + \displaystyle\sum_{l=1}^{n} \mathbb{W}_{jl}\rho_l = 0 \; , \end{cases} \tag{3}$$

where $\rho_j(t)$ and $S_j(t)$ are the probability density and potential function at time t on node j respectively, $N(j) = \{l \in V : (j, l) \in E\}$ is the adjacency set of node j, $\theta_{jl}(\rho) = \frac{1}{2}(\frac{1}{d_j}\rho_j + \frac{1}{d_l}\rho_l)$ represents the weight of probability density on the edge $(j, l) \in E$, $d_j = \sum_{l \in N(j)} \omega_{jl}$ is the volume form at node j and

$$\mathcal{I}(\rho) := \frac{1}{2} \sum_{(j,l) \in E} \omega_{jl}(\log \rho_j - \log \rho_l)^2 \theta_{jl}(\rho)$$

denotes the discrete Fisher information. Using $\rho_j(t)$ and $S_j(t)$, we reconstruct a complex wave equation, for the discrete NLS, on the graph. The derived discrete NLS is very different from the commonly seen ones, with the most notable distinction being a nonlinear graph Laplacian, which has not been reported before.

We shall prove that the initial value ODE (3) is well defined with several favorable properties. For example, it is a Hamiltonian system that conserves the total mass and total energy. It is time reversible and gauge invariant. Its stationary solution is related to the discrete ground state, whose formulation has many desirable properties similar to those in the continuous space. In addition, a Hamiltonian matrix is introduced to study the stability of ground states. This Hamiltonian matrix is a symplectic decomposition of two nonlinear graph Laplacian matrices. One is from discrete optimal transport geometry, and the other is induced by the Hessian matrix of discrete Fisher information.

It is worth noting that there are other approaches to study quantum Markov semi-group equations [8,7,21,22]. Their methods generalize the discrete optimal transport metric into matrix form and establish the Lindblad equations by the associated gradient flows. Compare to them, our approach fixes the discrete optimal transport metric and discusses the related Hamiltonian structures. From numerics and modeling angles, our goals are to establish new discrete schemes for computing Schrödinger equations and for exploring dynamical properties of discrete Hamiltonian flows in the light of optimal transport.

Our paper is arranged as follows. In Section 2, we discuss the necessity of using non-linear Laplacian on graphs. In Section 3, we briefly review Nelson's approach, and then followed by the derivation of (3) in Section 4. We show several interesting dynamical properties of (3) in Sections 5 and 6. Several numerical examples are provided in Section 7.

2. Why nonlinear Laplacian on graphs?

To answer the question, we consider the classical linear Schrödinger equation without potentials,

$$i\frac{\partial}{\partial t}\Psi(t,x) = -\frac{1}{2}\Delta\Psi(t,x) , \quad x \in \mathbb{R}^d. \tag{4}$$

It is well known that (4) admits plane wave solutions given in the form of

$$\Psi(x,t) = Ae^{i(k\cdot x - \mu t)} ,$$

as long as the time frequency μ and the spatial wave number k satisfy the following *dispersion relation*:

$$\mu = \frac{|k|^2}{2} .$$

Such a simple property may become problematic on graphs. To illustrate the challenges we face, let us consider the regular lattice in $\mathbb{R}^d$ for its simplicity, or even periodic lattices if one wishes to avoid dealing with the boundaries. In the lattice, we assume that every node has the same number of adjacent nodes. The weight on each edge is uniformly given by Δx, and the coordinate value for node j is $x_j = j\Delta x$.

On the lattice, any linear spatial discretization of (4) can be expressed as

$$i\frac{d\Psi_j}{dt} = -\frac{1}{2}\sum_{l\in N(j)} C_{jl}\Psi_l , \tag{5}$$

where $\{C_{jl}\}'s$ are selected, not all zeros, constants used to approximate the Laplace operator in (4). Assume that the discrete plane wave $\Psi_j(t) = Ae^{i(k\cdot x_j - \mu t)}$ satisfies (5), we must have

$$\mu e^{i(k \cdot j\Delta x - \mu t)} = -\frac{1}{2} \sum_{l \in N(j)} C_{jl} e^{i(k \cdot l\Delta x - \mu t)} \,,$$

which is equivalent to

$$\mu = -\frac{1}{2} \sum_{l \in N(j)} C_{jl} e^{ik \cdot (l-j)\Delta x} \,.$$

If μ and k satisfy the dispersion relation $\mu = |k|^2/2$, one gets

$$|k|^2/2 = -\frac{1}{2} \sum_{l \in N(j)} C_{jl} e^{ik \cdot (l-j)\Delta x} \,.$$

The left hand side is a quadratic function on the wave number k, while the right hand side is a periodic function consisting of finite terms of trignometric polynomials. This implies that there are at most a finite number of values for k satisfying this relation. Therefore only a finite number of pairs (μ, k) can form the plane wave solutions for (5). In contrast, any dispersion relation satisfying pair (μ, k) gives a plane wave solution for (4), and there are infinitely many of them. We summarize this observation in the following theorem.

Theorem 1. *For any linear spatial discretization of* (4)*, there are at most a finite number of pairs* (μ, k) *satisfying* $\mu = |k|^2/2$ *that can form its discrete plane wave solutions.*

An implication of this theorem is that one cannot expect every pair (μ, k), $\mu = |k|^2/2$, to give a plane wave solution for a given linear spatial discretization. In fact, only a finite number of pairs, which is a measure zero set, can do that. It also suggests that a nonlinear Laplacian on graphs must be used if one wants to construct a spatial discretization scheme that allows any pair (μ, k) to form a plane wave solution. On the other hand, this theorem does not imply that for any given pair (μ, k), one cannot find a linear Laplacian on the graph and use the pair to construct a plane wave solution for (5). However, if a different pair is given, one may have to switch to a different linear Laplacian on the graph.

We also note that nonlinear Laplace operators have been used, and proved to be necessary in the study of Fokker–Planck equations on graphs [10]. These observations motivated us to consider Nelson's approach to systematically construct schemes for (1).

3. Review of Nelson's approach

In this section, we briefly review Nelson's approach [23,24], and explain its connection with optimal transport [26,28]. To simplify the presentation, we do not consider the interactive potential $\mathbb{W}(x, y)$ in this section, even though such a consideration can be obtained in a straightforward manner.

Consider the following stochastic variational problem:

$$\inf_{b}\{\int_0^1 \mathbb{E}[\frac{1}{2}\dot{X}_t^2 - \mathbb{V}(X_t)]dt \; : \; \dot{X}_t = b(t, X_t) + \sqrt{h}\dot{B}_t \; , \; X(0) \sim \rho^0 \; , \; X(1) \sim \rho^1\} \; , \quad (6)$$

where $b(t, x) \in \mathbb{R}^d$ can be any smooth vector field, X_t is a stochastic process with prescribed probability densities ρ^0 and ρ^1 at time 0 and 1 respectively, $h > 0$ represents the noise level, B_t is a standard Brownian motion in $\mathbb{R}^d$ and $\mathbb{E}$ the expectation operator. Under suitable conditions given in [23], Nelson showed that (6) is equivalent to

$$\inf_{b}\{\int_0^1 \mathbb{E}[\frac{1}{2}(b(t, X_t)^2 + h\nabla \cdot b(t, X_t)) - \mathbb{V}(X_t)]dt \; : \; \dot{X}_t = b(t, X_t) + \sqrt{h}\dot{B}_t \; ,$$

$$X(0) \sim \rho^0 \; , \; X(1) \sim \rho^1\} \; . \quad (7)$$

By using the probability density function $\rho(t, x)$ defined as

$$\int_A \rho(t, x)dx = \Pr(X_t \in A) \; , \quad \text{for any measurable set } A \; ,$$

problem (7) is transferred into a deterministic variational problem,

$$\inf_{b} \int_0^1 \int_{\mathbb{R}^d} [\frac{1}{2}(b^2\rho + h\nabla \cdot b)\rho - \mathbb{V}(x)\rho]dxdt \; , \quad (8)$$

in which the density function $\rho(t, x)$ evolves according to the Fokker–Planck equation

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

Noticing the facts $\nabla\rho = \rho\nabla\log\rho$, $\int_{\mathbb{R}^d} \nabla \cdot b\rho dx = -\int_{\mathbb{R}^d} b\nabla\rho dx = \int_{\mathbb{R}^d} b \cdot \rho\nabla\log\rho dx$, and $\Delta\rho = \nabla \cdot (\nabla\rho) = \nabla \cdot (\rho\nabla\log\rho)$, then (8) can be rewritten into the following optimal control problem

$$\inf_{b} \int_0^1 \int_{\mathbb{R}^d} [\frac{1}{2}(b^2 - hb \cdot \nabla\log\rho)\rho - \mathbb{V}(x)\rho]dxdt \; , \quad (9)$$

where

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

The key of Nelson's derivation is based on a *change of variable* given by,

$$v(t, x) := b(t, x) - \frac{h}{2} \nabla \log \rho(t, x) . \tag{10}$$

Substituting v into (9), the problem becomes

$$\inf_v \int_0^1 \{ \int_{\mathbb{R}^d} \frac{1}{2} v^2 \rho dx - \frac{h^2}{8} \mathcal{I}(\rho) - \mathcal{V}(\rho) \} dt , \tag{11}$$

such that

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

Here $\mathcal{V}(\rho) := \int_{\mathbb{R}^d} V(x)\rho(x)dx$ is the linear potential energy and $\mathcal{I}(\rho) := \int_{\mathbb{R}^d} (\nabla \log \rho)^2 \rho dx$ is the Fisher information. The integrant in (11) is the Lagrangian for the optimal control problem.

The critical point (in the sense of Guerra–Morato [23]) of (11) satisfies the Madelung equations

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \nabla \cdot (\rho \nabla S) = 0 ; \\ \dfrac{\partial S}{\partial t} + \dfrac{1}{2} (\nabla S)^2 + \dfrac{\delta}{\delta \rho(x)} \{ \dfrac{h^2}{8} \mathcal{I}(\rho) + \mathcal{V}(\rho) \} = 0 , \end{cases} \tag{12}$$

where $v(t, x) = \nabla S(t, x)$ is the optimal velocity field. The first equation in (12) is known as the continuity equation, while the second one is called Hamilton–Jacobi equation in the literature.

Introducing a complex wave function

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

then $\Psi(t, x)$ satisfies the linear Schrödinger equation

$$hi \frac{\partial}{\partial t} \Psi = -\frac{h^2}{2} \Delta \Psi + \Psi \mathbb{V}(x) .$$

Compared to the optimal transport distance, which can be defined by the well known Benamou–Brenier formula [4],

$$\inf_v \{ \int_0^1 \int_{\mathbb{R}^d} v^2 \rho dx dt \ : \ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0 , \quad \rho(0, \cdot) = \rho^0(\cdot) , \quad \rho(1, \cdot) = \rho^1(\cdot) \}.$$

Nelson's approach can be viewed as a modified optimal transport problem in which the negative of a potential energy and the Fisher information are amended to create the Lagrangian in formulation (11).

4. A Schrödinger equation on a finite graph

Following Nelson's approach, we shall derive the discrete NLS (3) on a graph via discrete optimal transport.

To do so, we first review some basics of the discrete optimal transport theory developed in recent years. Consider a finite graph $G = (V, E, \omega)$. The probability set (simplex) supported on all vertices of G is defined by

$$\mathcal{P}(G) = \{(\rho_j)_{j=1}^n \mid \sum_{j=1}^n \rho_j = 1, \quad \rho_j \geq 0, \quad \text{for any } j \in V\},$$

where ρ_j is the discrete probability function at node j. The interior of $\mathcal{P}(G)$ is denoted by $\mathcal{P}_o(G)$.

Following [11], we introduce some notations and operators on G and $\mathcal{P}(G)$. A *vector field* v on G refers to a *skew-symmetric matrix* on the edges set E:

$$v := (v_{jl})_{(j,l) \in E}, \quad \text{with} \quad v_{lj} = -v_{jl}.$$

Given a function $S = (S_j)_{j=1}^n$ on V, it induces a *potential vector field* $\nabla_G S$ on G as

$$\nabla_G S := (\sqrt{\omega_{jl}}(S_j - S_l))_{(j,l) \in E}.$$

For a probability function $\rho \in \mathcal{P}(G)$ and a vector field v, define the product ρv, called *flux function* on G, by

$$\rho v := (v_{jl}\theta_{jl}(\rho))_{(j,l) \in E},$$

where $\theta_{jl}(\rho)$ is a chosen function on the edge

$$\theta_{jl}(\rho) = \frac{\rho_j + \rho_l}{2}, \quad \text{for any } (j,l) \in E.$$

We remark that θ_{jl} may have other choices, such as the logarithmic mean used in [18]. The *divergence* of flux function ρv on G is defined by

$$\text{div}_G(\rho v) := -\left(\sum_{l \in N(j)} \sqrt{\omega_{jl}} v_{jl} \theta_{jl}(\rho)\right)_{j=1}^n.$$

Given two vector fields $v = (v_{jl})_{(j,l) \in E}$, $u = (u_{jl})_{(j,l) \in E}$ on a graph and $\rho \in \mathcal{P}(G)$, the discrete *inner product* is defined by,

$$(v, u)_\rho := \frac{1}{2} \sum_{(j,l) \in E} v_{jl} u_{jl} \theta_{jl}(\rho) ,$$

where the coefficient $1/2$ accounts for the fact that every edge in G is counted twice, i.e. $(j, l), (l, j) \in E$.

Using the notations, the Wasserstein metric on the graph can be defined by the discrete Benamou–Brenier formula [4],

Definition 2. For any ρ^0, $\rho^1 \in \mathcal{P}_o(G)$, define a metric

$$W(\rho^0, \rho^1) := \inf_v \left\{ \left(\int\limits_0^1 (v, v)_\rho dt \right)^{\frac{1}{2}} : \frac{d\rho}{dt} + \mathrm{div}_G(\rho v) = 0 , \quad \rho(0) = \rho^0 , \quad \rho(1) = \rho^1 \right\} ,$$

where the infimum is taken over all vector fields v on a graph, and ρ is a continuous differentiable curve $\rho : [0, 1] \to \mathcal{P}_o(G)$.

$\mathcal{P}_o(G)$ equipped with the metric W is a Riemannian manifold [10,11].

4.1. Nelson's approach on a finite graph

Now, we are ready to derive the NLS on graph (3) via discrete optimal transport. In the discrete case, the linear and interaction potentials refer to

$$\mathcal{V}(\rho) = \sum_{j=1}^n \mathbb{V}_j \rho_j , \quad \mathcal{W}(\rho) = \frac{1}{2} \sum_{l=1}^n \sum_{j=1}^n \mathbb{W}_{lj} \rho_l \rho_j ,$$

respectively. We start with a discrete analog of Nelson's problem presented in (8),

$$\inf_b \int\limits_0^1 \frac{1}{2}[(b, b)_\rho - h(b, \nabla_G \log \rho)_\rho] - \mathcal{V}(\rho) - \mathcal{W}(\rho)dt , \tag{13}$$

where the infimum is taken over all discrete vector fields b, $\rho(t)$ satisfies the discrete Fokker–Planck equation [10,11]:

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

and $\rho(0) = \rho^0$, $\rho(1) = \rho^1$ are given in $\mathcal{P}_o(G)$. Similar to Nelson's change of variable, we define a new vector field $v = (v_{lj})_{(l,j) \in E}$ on the graph

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

Substituting v into (13), we can write the objective functional as

$$J(v) := \int_0^1 \frac{1}{2}(v,v)_\rho - \frac{h^2}{8}\mathcal{I}(\rho) - \mathcal{V}(\rho) - \mathcal{W}(\rho)dt \ , \tag{14}$$

where v is a vector field on the graph, such that $\rho(t) \in \mathcal{P}_o(G)$ satisfies

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

We note that any given feasible path $\rho(t)$ of (14) is determined by $v(t)$, so we denote functional J only in term of v, and call it discrete Nelson's functional.

We also call $(v(t), \rho(t))$ a **critical point** of (14) if

$$\mathcal{J}(v + \epsilon \delta v) - \mathcal{J}(v) = o(\epsilon) \ , \quad \forall \delta v(t) \in \mathcal{D}, \tag{15}$$

where

$$\mathcal{D} = \{\delta v \in C^\infty[0,1] \ : \ \bar{\rho}(t) \in \mathcal{P}_o(G) \text{ is continuously differentiable, and }, $$

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

In the following theorem, we show that the critical point of the discrete Nelson's functional satisfies (3).

Theorem 3 (*Critical point of discrete Nelson's functional*). *Assume there exists a critical point $(v(t), \rho(t))$ of (14) in the sense of (15), whose components are smooth functions with respect to the time variable. Then $v(t)$ and $\rho(t)$ satisfy the following conditions:*

(a) *$v(t)$ is a potential vector field on the graph, i.e. there exists a function $S(t) = (S_j(t))_{j=1}^n$ defined on the nodes, such that*

$$v_{jl}(t) = \sqrt{\omega_{jl}}(S_j(t) - S_l(t)) \ , \quad \text{for all } t \in [0,1] \text{ and } (j,l) \in E \ .$$

(b) *For every $S(t)$ that induces $v(t)$, there exists a scalar function $C(t) \in \mathbb{R}$, independent of the nodes, such that $\rho(t)$ and $\bar{S}(t)$, defined by $\bar{S}(t) = (S_j(t) - C(t))_{j=1}^n$, satisfy (3):*

$$\begin{cases} \dfrac{d\rho_j}{dt} + \displaystyle\sum_{l \in N(j)} \omega_{jl}(\bar{S}_j - \bar{S}_l)\theta_{jl}(\rho) = 0 \ ; \\[4mm] \dfrac{d\bar{S}_j}{dt} + \dfrac{1}{2}\displaystyle\sum_{l \in N(j)} \omega_{jl}(\bar{S}_j - \bar{S}_l)^2\dfrac{\partial\theta_{jl}}{\partial\rho_j} + \dfrac{h^2}{8}\dfrac{\partial}{\partial\rho_j}\mathcal{I}(\rho) + \mathbb{V}_j + \displaystyle\sum_{l=1}^n \mathbb{W}_{jl}\rho_l = 0 \ , \end{cases}$$

Remark 1. We can rewrite (3) as

$$\frac{d}{dt}\begin{pmatrix} \rho \\ S \end{pmatrix} = \mathbb{J}\begin{pmatrix} \frac{\partial}{\partial \rho}\mathcal{H} \\ \frac{\partial}{\partial S}\mathcal{H} \end{pmatrix} , \tag{16}$$

where $\mathcal{H}$ is the discrete total energy

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

and

$$\mathbb{J} = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}$$

is a symplectic matrix, with $\mathbb{I} \in \mathbb{R}^{n \times n}$ being the identity matrix. The symplectic form (16), identical to (3), is the discrete analog of (12). Following the convention, we call the first equation the discrete continuity equation, and the second one the discrete Hamilton–Jacobi equation.

The proof of Theorem 3 requires the following lemma, which can be viewed as the Hodge decomposition on a graph.

Lemma 4. *Given a vector field* $v = (v_{jl})_{(j,l) \in E}$ *on a graph and a probability density function* $\rho \in \mathcal{P}_o(G)$, *there exists a unique* $\nabla_G S$, *such that*

$$v = \nabla_G S + u , \quad with \quad div_G(\rho u) = 0 . \tag{17}$$

Proof. The detailed proof can be found in [18]. For the completeness of this paper, more importantly to introduce some notations that will be used later in the paper, we sketch the proof here. We define a weighted graph Laplacian matrix $L(\rho) \in \mathbb{R}^{n \times n}$:

$$L(\rho) = -D^T \Theta(\rho)D ,$$

where $D \in \mathbb{R}^{|E| \times |V|}$ is the discrete gradient matrix

$$D_{(j,l) \in E, k \in V} = \begin{cases} \sqrt{\omega_{jl}} & \text{if } j = k ; \\ -\sqrt{\omega_{jl}} & \text{if } l = k ; \\ 0 & \text{otherwise;} \end{cases}$$

and $\Theta \in \mathbb{R}^{|E| \times |E|}$ is the diagonal weighted matrix

$$\Theta_{(j,l) \in E, (j',l') \in E} = \begin{cases} \theta_{jl}(\rho) & \text{if } (j, l) = (j', l') \in E ; \\ 0 & \text{otherwise .} \end{cases}$$

We want to emphasize that $L(\rho)$ depends on ρ. This is different from the commonly seen graph Laplace operators, which are often independent of ρ.

We only need to show that there exists a unique gradient vector field $\nabla_G S$, such that

$$\text{div}_G(\rho \nabla_G S) = L(\rho)S = \text{div}_G(\rho v) .$$

Since $\rho \in \mathcal{P}_o(G)$ and the graph is connected, then

$$S^T L(\rho)S = \frac{1}{2} \sum_{(j,l)\in E} \omega_{jl}(S_j - S_l)^2 \theta_{jl}(\rho) = 0 ,$$

this implies that value 0 must be a simple eigenvalue of the weighted graph Laplacian matrix $L(\rho)$ with eigenvector $\{1, \cdots, 1\}$. Thus there exists a unique solution of S up to constant shrift. Therefore $\nabla_G S$ is unique. $\square$

Furthermore, we can express

$$L(\rho) = T \begin{pmatrix} 0 & & & \\ & \lambda_{sec}(L(\rho)) & & \\ & & \ddots & \\ & & & \lambda_{\max}(L(\rho)) \end{pmatrix} T^{-1} ,$$

where $0 < \lambda_{sec}(L(\rho)) \leq \cdots \leq \lambda_{\max}(L(\rho))$ are eigenvalues of $L(\rho)$ arranged in the ascending order, and T is its corresponding eigenvector matrix. The pseudo-inverse of $L(\rho)$ is defined by

$$L(\rho)^{-1} = T \begin{pmatrix} 0 & & & \\ & \frac{1}{\lambda_{sec}(L(\rho))} & & \\ & & \ddots & \\ & & & \frac{1}{\lambda_{max}(L(\rho))} \end{pmatrix} T^{-1} .$$

Thus $S = L(\rho)^{-1}\text{div}_G(\rho v)$.

Proof of Theorem 3. (a) We prove that $v(t) = \nabla_G S(t)$. From Lemma 4, we have,

$$v(t) = \nabla_G S(t) + u(t) , \quad \text{with} \quad \text{div}_G(\rho(t)u(t)) = 0 \quad \text{for all } t \in [0,1] .$$

We only need to prove $u(t) = 0$ for $t \in [0,1]$ when $v(t)$ is a critical point. Consider a function $w(t) = (w_{jl}(t))_{(j,l)\in E}$ satisfying

$$\text{div}_G(\rho w(t)) = 0 , \quad \text{for } t \in [0,1] .$$

It is clear that $\epsilon w \in \mathcal{D}$ for any $\epsilon > 0$ because $\frac{d\rho}{dt} + \text{div}_G(\rho(v + \epsilon w)) = 0$.

Since (v, ρ) is a critical solution of (14), we must have

$$\lim_{\epsilon \to 0} \frac{\mathcal{J}(v + \epsilon w) - \mathcal{J}(v)}{\epsilon} = 0 . \tag{18}$$

Because $\rho(t)$ keeps the same for the vector field $v + \epsilon w$, it implies

$$2\frac{\mathcal{J}(v + \epsilon w) - \mathcal{J}(v)}{\epsilon} = \int_0^1 \frac{(v + \epsilon w, v + \epsilon w)_\rho - (v, v)_\rho}{\epsilon} dt$$

$$= \int_0^1 \frac{(v, v)_\rho + 2\epsilon(v, w)_\rho + \epsilon^2(w, w)_\rho - (v, v)_\rho}{\epsilon} dt$$

$$= 2\int_0^1 (v, w)_\rho dt + O(\epsilon) = 2\int_0^1 (\nabla_G S + u, w)_\rho dt + O(\epsilon)$$

$$= 2\int_0^1 (\nabla_G S, w)_\rho + (u, w)_\rho dt + O(\epsilon)$$

$$= 2\int_0^1 -\sum_{j=1}^n \operatorname{div}_G(\rho w)|_j S_j + (u, w)_\rho dt + O(\epsilon)$$

$$= \int_0^1 \sum_{(j,l) \in E} u_{jl}(t) w_{jl}(t) \theta_{jl}(\rho(t)) dt + O(\epsilon) ,$$

where the last equality uses the fact $\operatorname{div}_G(\rho w) = 0$. From (18), we get

$$\int_0^1 \sum_{(j,l) \in E} u_{jl}(t) w_{jl}(t) \theta_{jl}(\rho(t)) dt = 0 .$$

In particular, by taking $w(t) = u(t)$, we obtain

$$\int_0^1 \sum_{(j,l) \in E} u_{jl}(t)^2 \theta_{jl}(\rho(t)) dt = 0 .$$

Since $\rho(t) \in \mathcal{P}_o(G)$, $\theta_{jl}(\rho(t)) > 0$ and $u(t) \in \mathcal{D}$, this implies $u(t) = 0$ for $t \in [0, 1]$, which proves (a).

(b) Since $\rho(t) \in \mathcal{P}_o(G)$ is continuous in $[0, 1]$, then $\min_{i \in V, \, t \in [0,1]} \rho_i(t) \geq c_0 > 0$. We consider a perturb function $\rho^\epsilon(t)$ defined by:

$$\rho^\epsilon(t) = \rho(t) + \epsilon \delta \rho(t) ,$$

where $\delta\rho(t) = (\delta\rho_j(t))_{j=1}^n$, $\delta\rho_j(t) \in C^\infty[0,1]$ with $\sum_{j=1}^n \delta\rho_j = 0$ and $\delta\rho(0) = \delta\rho(1) = 0$. Let $\epsilon \sup_{0 \le t \le 1} |\delta\rho(t)| < \frac{1}{2}c_0$, then $\rho^\epsilon(t) \in \mathcal{P}_o(G)$. Thus $L(\rho^\epsilon(t))^{-1}$ is well defined for $t \in [0,1]$, whose entries are smooth. From $S^\epsilon(t) = L(\rho^\epsilon(t))^{-1}\frac{d\rho^\epsilon}{dt}$, then $S^\epsilon(t)$ is smooth with respect to t and ϵ. Since

$$\frac{d\rho^\epsilon}{dt} + \nabla_G(\rho^\epsilon(\nabla_G S^\epsilon - \nabla_G S + \nabla_G S)) = 0,$$

we have $\nabla_G S^\epsilon(t) - \nabla_G S(t) \in D$.

For the simplicity of presentation, we denote

$$\mathcal{F}(\rho) := \frac{h^2}{8}\mathcal{I}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho) .$$

By direct calculations, we have

$$\frac{\mathcal{J}(\nabla_G S^\epsilon) - \mathcal{J}(\nabla_G S)}{\epsilon}$$

$$= \int_0^1 \frac{1}{2\epsilon}[(\nabla_G S^\epsilon, \nabla_G S^\epsilon)_{\rho^\epsilon} - (\nabla_G S, \nabla_G S)_\rho]dt - \int_0^1 \frac{1}{\epsilon}[\mathcal{F}(\rho^\epsilon) - \mathcal{F}(\rho)]dt .$$

$$= \int_0^1 \frac{1}{2\epsilon}[(\nabla_G S^\epsilon, \nabla_G S^\epsilon)_{\rho^\epsilon} - (\nabla_G S, \nabla_G S)_\rho]dt \qquad (\star) \tag{19}$$

$$- \int_0^1 \sum_{j=1}^n \delta\rho_j \frac{\partial}{\partial\rho_j}\mathcal{F}(\rho)dt + O(\epsilon) .$$

We need to estimate $(\star)$. Using the Taylor expansion of $S^\epsilon(t)$ with respect to ϵ, $S^\epsilon(t) = S(t) + \epsilon\delta S(t) + o(\epsilon)$, where $\delta S(t) = \frac{d}{d\epsilon}S^\epsilon(t)|_{\epsilon=0}$, we obtain

$$(\star) = \int_0^1 \frac{1}{2\epsilon}[(\nabla_G S + \epsilon\nabla_G \delta S, \nabla_G S + \epsilon\nabla_G \delta S)_{\rho^\epsilon} - (\nabla_G S, \nabla_G S)_\rho]dt + O(\epsilon)$$

$$= \int_0^1 \frac{1}{2\epsilon}[(\nabla_G S, \nabla_G S)_{\rho^\epsilon} - (\nabla_G S, \nabla_G S)_\rho + 2\epsilon(\nabla_G S, \nabla_G \delta S)_{\rho^\epsilon}]dt + O(\epsilon)$$

$$= \int_0^1 \frac{1}{4}\sum_{j=1}^n \sum_{l \in N(j)} \omega_{jl}(S_j - S_l)^2[\theta_{jl}(\rho + \epsilon\delta\rho) - \theta_{jl}(\rho)]dt + \int_0^1 (\nabla_G S, \nabla_G \delta S)_\rho dt + O(\epsilon) \tag{20}$$

$$= \int_0^1 \frac{1}{4}\sum_{(j,l) \in E} \omega_{jl}(S_j - S_l)^2[\frac{\partial\theta_{jl}}{\partial\rho_j}\delta\rho_j + \frac{\partial\theta_{jl}}{\partial\rho_l}\delta\rho_l]dt + \int_0^1 (\nabla_G S, \nabla_G \delta S)_\rho dt + O(\epsilon)$$

$$= \int_0^1 \frac{1}{2} \sum_{j=1}^n [\delta\rho_j \sum_{l\in N(j)} \omega_{jl}(S_j - S_l)^2 \frac{\partial\theta_{jl}}{\partial\rho_j}]dt + \int_0^1 (\nabla_G S, \nabla_G \delta S)_\rho dt + O(\epsilon) \ .$$

Next we find a connection between $\delta\rho(t)$ and $\delta S(t)$. Notice that

$$\frac{d\rho}{dt} + \epsilon\frac{d}{dt}\delta\rho = \mathrm{div}_G((\rho + \epsilon\delta\rho)\nabla_G(S + \epsilon\delta S + o(\epsilon))) \ .$$

By comparing the order ϵ term, we have

$$\frac{d}{dt}\delta\rho_j = \sum_{l\in N(j)} \omega_{jl}(S_l - S_j)[\frac{\partial\theta_{jl}}{\partial\rho_j}\delta\rho_j + \frac{\partial\theta_{jl}}{\partial\rho_l}\delta\rho_l] + \mathrm{div}_G(\rho\nabla_G\delta S) \ .$$

Then

$$\int_0^1 \sum_{j=1}^n S_j \frac{d}{dt}\delta\rho_j dt$$

$$= \int_0^1 \sum_{j=1}^n \sum_{l\in N(j)} \omega_{lj}S_j(S_l - S_j)[\frac{\partial\theta_{lj}}{\partial\rho_j}\delta\rho_j + \frac{\partial\theta_{jl}}{\partial\rho_l}\delta\rho_l] + \sum_{j=1}^n S_j\mathrm{div}_G(\rho\nabla_G\delta S)|_j dt$$

$$= \int_0^1 (\frac{1}{2}+\frac{1}{2}) \sum_{j=1}^n \delta\rho_j \sum_{l\in N(j)} \omega_{lj}(S_l - S_j)^2 \frac{\partial\theta_{lj}}{\partial\rho_j} + (\nabla_G S, \nabla_G\delta S)_\rho dt \tag{21}$$

$$= (T_1) + \int_0^1 \frac{1}{2} \sum_{j=1}^n \delta\rho_j \sum_{l\in N(j)} \omega_{jl}(S_l - S_j)^2 \frac{\partial\theta_{lj}}{\partial\rho_j} dt + O(\epsilon) \ ,$$

where the last equality is from (20). Substituting (21) into (19), we have

$$0 = \lim_{\epsilon\to 0} \frac{J(\nabla_G S^\epsilon) - J(\nabla_G S)}{\epsilon} = \lim_{\epsilon\to 0} \{T - \int_0^1 \sum_{j=1}^n \delta\rho_j \frac{\partial}{\partial\rho_j}\mathcal{F}(\rho)dt + O(\epsilon)\}$$

$$= \int_0^1 \sum_{j=1}^n S_j \frac{d}{dt}\delta\rho_j dt - \frac{1}{2} \sum_{j=1}^n \delta\rho_j \sum_{l\in N(j)} \omega_{lj}(S_l - S_j)^2 \frac{\partial\theta_{lj}}{\partial\rho_j}dt - \sum_{j=1}^n \delta\rho_j \frac{\partial}{\partial\rho_j}\mathcal{F}(\rho)dt$$

$$= -\int_0^1 \sum_{j=1}^n \delta\rho_j\{\frac{d}{dt}S_j + \frac{1}{2} \sum_{l\in N(j)} \omega_{lj}(S_l - S_j)^2 \frac{\partial\theta_{lj}}{\partial\rho_j} + \frac{\partial}{\partial\rho_j}\mathcal{F}(\rho)\}dt \ ,$$

where the last equality is from integration by parts and $\delta\rho(0) = \delta\rho(1) = 0$. Since $\delta\rho_j(t)$ with $\sum_{j=1}^n \delta\rho_j(t) = 0$ can be any smooth function, we obtain

$$\frac{d}{dt}S_j + \frac{1}{2}\sum_{l\in N(j)}\omega_{jl}(S_j - S_l)^2\frac{\partial\theta_{jl}}{\partial\rho_j} + \frac{\partial}{\partial\rho_j}\mathcal{F}(\rho) = c(t) \, ,$$

for any smooth function $c(t)\in\mathbb{R}$. We denote $C(t) = \int_0^t c(s)ds$, then $\bar{S}_j(t) = S_j(t) - C(t)$, and together with $\rho(t)$, satisfy (3). $\square$

Corollary 5. *Let $(v(t), \rho(t))$ be a critical point of (14), and $S(t)$ a function on G that induces $v(t)$, then $S(t)$ and $\rho(t)$ satisfy (3) if and only if*

$$\sum_{j=1}^n S_j(t)\rho_j(t) = \sum_{j=1}^n S_j(0)\rho_j(0) + \int_0^t \{\frac{1}{2}(\nabla_G S, \nabla_G S)_\rho - \frac{h^2}{8}\mathcal{I}(\rho) - \mathcal{V}(\rho) - 2\mathcal{W}(\rho)\}ds \, . \quad (22)$$

Proof. From the proof of Theorem 3, we know

$$\frac{dS_j}{dt} + \frac{1}{2}\sum_{l\in N(j)}\omega_{jl}(S_j - S_l)^2\frac{\partial\theta_{jl}}{\partial\rho_j} + \frac{\partial}{\partial\rho_j}\mathcal{F}(\rho) = c(t) \, .$$

Then by direct calculations, we obtain

$$\frac{d}{dt}(\sum_{j=1}^n S_j(t)\rho_j(t))$$

$$= \sum_{j=1}^n [\frac{dS_j}{dt}\rho_j(t) + \frac{d\rho_j}{dt}S_j(t)]$$

$$= \sum_{j=1}^n [-\frac{1}{2}\sum_{l\in N(j)}\omega_{jl}(S_j - S_l)^2\frac{\partial\theta_{jl}}{\partial\rho_j} - \frac{\partial}{\partial\rho_j}\mathcal{F}(\rho) + c(t)]\rho_j(t) + (\nabla_G S, \nabla_G S)_\rho$$

$$= \frac{1}{2}(\nabla_G S, \nabla_G S)_\rho - \sum_{j=1}^n \frac{\partial}{\partial\rho_j}\mathcal{F}(\rho)\rho_j(t) + c(t) \, .$$

We note

$$\sum_{j=1}^n \frac{\partial}{\partial\rho_j}\mathcal{F}(\rho)\rho_j = \sum_{j=1}^n \frac{\partial}{\partial\rho_j}(\frac{h^2}{8}\mathcal{I}(\rho) + \mathbf{V}^T\rho + \frac{1}{2}\rho^T\mathbf{W}\rho)\rho_j$$

$$= \sum_{j=1}^n \frac{h^2}{8}\frac{\partial}{\partial\rho_j}\mathcal{I}(\rho)\rho_j + \mathbf{V}^T\rho + \frac{1}{2}\rho^T\mathbf{W}\rho + \frac{1}{2}\rho^T\mathbf{W}\rho \, ,$$

and

$$\sum_{j=1}^{n} \frac{\partial}{\partial \rho_j} \mathcal{I}(\rho) \cdot \rho_j$$

$$= \sum_{j=1}^{n} \sum_{l \in N(j)} \omega_{jl} (\log \rho_l - \log \rho_j)^2 \frac{\partial \theta_{jl}}{\partial \rho_j} \rho_j + 2 \sum_{j=1}^{n} \sum_{l \in N(j)} \omega_{jl} \frac{1}{\rho_j} (\log \rho_j - \log \rho_l) \rho_j$$

$$= \frac{1}{2} \sum_{l=1}^{n} \sum_{l \in N(j)} \omega_{jl} (\log \rho_j - \log \rho_l)^2 \theta_{jl}(\rho) + 2 \sum_{(j,l) \in E} \omega_{jl} (\log \rho_j - \log \rho_l)$$

$$= \mathcal{I}(\rho) .$$

Combining the calculations together, we have

$$\frac{d}{dt} (\sum_{j=1}^{n} S_j(t) \rho_j(t)) = \frac{1}{2} (\nabla_G S, \nabla_G S)_\rho - \mathcal{F}(\rho) - \mathcal{W}(\rho) + c(t) .$$

Therefore $c(t) = 0$ if and only if (22) holds. □

In fact, the construction of S suggests $S = L(\rho)^{-1} \dot{\rho}$. This implies

$$(\nabla_G S, \nabla_G S)_\rho = S^T L(\rho) S = \dot{\rho}^T L(\rho)^{-1} \cdot L(\rho) \cdot L(\rho)^{-1} \dot{\rho} = \dot{\rho}^T L(\rho)^{-1} \dot{\rho} .$$

Thus the discrete Nelson's problem can be rewritten as a geometric variational problem on the probability density manifold $\mathcal{P}_o(G)$

$$\inf_{\rho} \{ \int_0^1 \dot{\rho}^T L(\rho)^{-1} \dot{\rho} - \frac{h^2}{8} \mathcal{I}(\rho) - \mathcal{V}(\rho) - \mathcal{W}(\rho) dt \; : \; \rho(0) = \rho^0 \;, \; \rho(1) = \rho^1 \;, \; \rho(t) \in \mathcal{C} \}, \quad (23)$$

where $\mathcal{C}$ is the set of continuous differentiable curve in $\mathcal{P}_o(G)$, and $L(\rho)^{-1}$ introduces a location dependent inner product on the tangent space at every point in $(\mathcal{P}_o(G), W)$; see [11]. A solution of (16) is a critical point of (23).

4.2. Complex formulations

In this sequel, we reformulate (3) into a complex wave equation. Let us define

$$\Psi(t) = (\Psi_j(t))_{j=1}^{n} = (\sqrt{\rho_j(t)} e^{i \frac{S_j(t)}{h}})_{j=1}^{n} ,$$

where $(\rho(t), S(t))$ are solutions of (3), then $\Psi(t)$ satisfies the following complex value ODE system.

$$hi \frac{d\Psi_j}{dt} = -\frac{h^2}{2} \Delta_G \Psi|_j + \Psi_j V_j + \Psi_j \sum_{l=1}^{n} \mathbb{W}_{jl} |\Psi_l|^2 , \quad (24)$$

in which the Laplacian on graph is defined by

$$\Delta_G \Psi|_j := -\Psi_j \Big(\frac{1}{|\Psi_j|^2} \sum_{l \in N(j)} \omega_{jl} (\log \Psi_j - \log \Psi_l)\theta_{jl} + \sum_{l \in N(j)} \omega_{jl} |\log \Psi_j - \log \Psi_l|^2 \frac{\partial \theta_{jl}}{\partial \rho_j}\Big) .$$

One may wonder when seeing the Laplace operator in such a nonlinear way. However, a closer examination demonstrates that this graph Laplacian is consistent with the one in the continuous case. In fact, we can show the following relationship in the continuous space. Let $\Psi(t, x)$ be a complex function defined in $\mathbb{R}^d$, then

$$\Delta \Psi = \Psi\{\frac{1}{|\Psi|^2} \nabla \cdot (|\Psi|^2 \nabla \log \Psi) - |\nabla \log \Psi|^2\} . \tag{25}$$

Proof of (25). Denote $\Psi(t, x) = \sqrt{\rho(t,x)} e^{i \frac{S(t,x)}{h}} = e^{\frac{1}{2} \log \rho(t,x) + i \frac{S(t,x)}{h}}$, we have

$$\Delta \Psi = \nabla \cdot (\nabla \Psi) = \nabla \cdot [\Psi(\frac{1}{2} \nabla \log \rho + i \frac{\nabla S}{h})]$$

$$= \Psi[(\frac{1}{2} \nabla \log \rho + i \frac{\nabla S}{h})^2 + (\frac{1}{2} \Delta \log \rho + i \Delta \frac{S}{h})]$$

$$= \Psi[\frac{1}{2}(\nabla \log \rho)^2 + \frac{1}{2} \Delta \log \rho + i \nabla \log \rho \cdot \nabla \frac{S}{h} + i \Delta \frac{S}{h} - \frac{1}{4}(\nabla \log \rho)^2 - (\nabla \frac{S}{h})^2]$$

$$= \Psi[\frac{1}{\rho} \nabla \cdot (\rho \nabla(\frac{1}{2} \log \rho + i \frac{S}{h})) - (\frac{1}{2} \nabla \log \rho)^2 - (\nabla \frac{S}{h})^2]$$

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

where the first equality uses $\frac{1}{\rho} \nabla \rho = \nabla \log \rho$, while the second to the last equality uses the fact

$$\frac{1}{\rho} \nabla \cdot (\rho \nabla(\frac{1}{2} \log \rho + i \frac{S}{h}))$$

$$= \frac{1}{\rho}[\frac{1}{2} \nabla \rho \cdot \nabla \log \rho + \rho \Delta \log \rho + i \nabla \rho \cdot \frac{S}{h} + \rho \Delta S]$$

$$= \frac{1}{2}(\nabla \log \rho)^2 + \frac{1}{2} \Delta \log \rho + i \nabla \log \rho \cdot \nabla \frac{S}{h} + i \Delta \frac{S}{h} . \quad \square$$

The nonlinearity in the Laplace operator allows the discrete NLS possessing many desirable dynamical properties, which will be shown in the next two sections.

5. Some properties

For the convenience of presentation, we do not distinguish (3) and its complex wave version (24) in the discussion. The results here are always proposed for formulation (24) while all proofs are based on (3).

Our first task is examining the dispersion relation in the absence of potentials.

Proposition. *For a uniform toroidal graph G, i.e. a graph that every node has the same number of adjacent nodes and the weight on each edge is uniformly given, the plane wave function $\Psi(t) = A(e^{i(k \cdot j\Delta x - \mu t)})_{j=1}^n$, with any $\mu = \frac{1}{2}|k|^2$ and $A \geq 0$, satisfies*

$$i\frac{d}{dt}\Psi = -\frac{1}{2}\Delta_G \Psi .$$

The proposition can be verified by directly substituting the plane wave function in (3).

In what follows, we show that (3) is a well defined ODE system having several desirable properties such as total mass and energy conservation, time reversibility, and gauge invariant. In addition, its interior stationary solution shares the same property as that for the counterpart in the continuous case.

Theorem 6. *Given a simple weighted graph $G = (V, E, \omega)$, a vector $(\mathbb{V}_l)_{l=1}^n$, a symmetric matrix $(\mathbb{W}_{jl})_{1 \leq j,l \leq n}$, and an initial condition $\Psi^0 = (\Psi_j^0)_{j=1}^n$ (complex vector) satisfying*

$$\sum_{j=1}^n |\Psi_j^0|^2 = 1 , \quad |\Psi_j^0|^2 > 0 , \quad \text{for any } j \in V .$$

Then equation (3) has a unique solution $\Psi(t)$ for all $t \in [0, \infty)$. Moreover, $\Psi(t)$ satisfies following properties:

(i) *It conserves the total mass*

$$\sum_{j=1}^n |\Psi_j(t)|^2 = 1 ;$$

(ii) *It conserves the total energy*

$$\mathcal{E}(\Psi(t)) = \mathcal{E}(\Psi^0) ,$$

where $\mathcal{E}$ is a combination of the discrete Kinetic energy $\mathcal{E}_{kin}$, linear potential energy $\mathcal{E}_{pot}$ and interaction potential energy $\mathcal{E}_{int}$, i.e.

$$\mathcal{E}(\Psi) = h^2 \mathcal{E}_{kin}(\Psi) + \mathcal{E}_{pot}(\Psi) + \mathcal{E}_{int}(\Psi) . \tag{26}$$

They are given by the following definitions:

$$\mathcal{E}_{kin}(\Psi) = \frac{1}{4} \sum_{(j,l) \in E} \{[Re(\log \Psi_j - \log \Psi_l)]^2 + [Im(\log \Psi_j - \log \Psi_l)]^2\} \theta_{jl}(|\Psi|^2) \ ,$$

$$\mathcal{E}_{pot}(\Psi) = \sum_{j=1}^{n} \mathbb{V}_j |\Psi_j|^2, \quad \mathcal{E}_{int}(\Psi) = \frac{1}{2} \sum_{j=1}^{n} \sum_{l=1}^{n} \mathbb{W}_{jl} |\Psi_j|^2 |\Psi_l|^2 \ ;$$

(iii) *It is time reversible:*

$$\Psi(t) = \bar{\Psi}(-t) \ ;$$

(iv) *It is time transverse (gauge) invariant: Denote $\Psi^\alpha(t)$ as the solution of* (3) *with $\mathbb{V}^\alpha = (\mathbb{V}_j + \alpha)_{j=1}^n$, where α is a given real constant, then*

$$\Psi^\alpha(t) = \Psi(t) e^{i\frac{\alpha t}{\hbar}} \ ;$$

(v) *A time invariant $\rho^* = (\rho_j^*)_{j=1}^n \in \mathcal{P}_o(G)$ and $\mathbf{S}(t) = \nu t$ form an interior stationary solution of* (3) *if and only if ρ^* is the critical point of $\{\mathcal{E}(\sqrt{\rho}): \rho \in \mathcal{P}(G)\}$, and*

$$\nu = \mathcal{E}(\sqrt{\rho^*}) + \mathcal{E}_{int}(\sqrt{\rho^*}) \ .$$

Remark 2. $\mathcal{E}_{kin}$ in (26) is an analog of the Kinetic energy in continuous case:

$$\int_{\mathbb{R}^d} |\nabla \Psi|^2 dx = \int_{\mathbb{R}^d} ([Re(\nabla \log \Psi)]^2 + [Im(\nabla \log \Psi)]^2) |\Psi|^2 dx \ .$$

Remark 3. Equations (3) are always well defined in the interior of probability set $\mathcal{P}(G)$. In fact, we shall show that the boundary of probability set $\mathcal{P}(G)$ is a repeller for (3).

Remark 4. The property of the interior stationary solution of (3) given in (v) of Theorem 6 mimics a similar property for the ground state of Schrödinger equation in the continuous case. Details will be given in Section 5.1.

Proof. We show that for any given initial condition $\rho^0 \in \mathcal{P}_o(G)$, there exists a unique solution $(\rho(t), S(t))$ for all $t > 0$. Since the right hand side of (3) is locally Lipchitz continuous and $\rho^0 \in \mathcal{P}_o(G)$, from Picard's existence theorem, there exists a unique solution $(\rho(t), S(t))$ in time interval $[0, T(\rho^0))$, where $T(\rho^0)$ is the maximal time that the solution exists. We will prove $T(\rho^0) = +\infty$ by the following claim.

Claim. *For any given $\rho^0 \in \mathcal{P}_o(G)$, there exists a compact set $B \subset \mathcal{P}_o(G)$, such that $T(\rho^0) = \infty$ and $\rho(t) \in B$.*

The proof of Claim is based on two facts. On one hand, the ODE system (3) is a Hamiltonian system on probability set, which conserves the total mass and total energy;

On the other hand, the total energy contains the Fisher information $\mathcal{I}(\rho)$. On the boundary of $\mathcal{P}_o(G)$, $\mathcal{I}(\rho)$ is positive infinity, so is the total energy. From the conservation of total energy, it is not hard to see that the boundary of $\mathcal{P}_o(G)$ is a repeller for $\rho(t)$.

Proof of Claim. We construct a set $B \subset \mathcal{P}(G)$:

$$B = \{\rho \in \mathcal{P}(G) \ : \ \frac{h^2}{8}\mathcal{I}(\rho) \leq \mathcal{E}(\Psi^0) - \min_{\rho \in \mathcal{P}(G)}[\mathcal{V}(\rho) + \mathcal{W}(\rho)]\ \} \ ,$$

where $\mathcal{E}(\Psi^0) = \mathcal{H}(\rho^0, S^0) = \frac{1}{2}(\nabla_G S^0, \nabla_G S^0)_{\rho^0} + \frac{h^2}{8}\mathcal{I}(\rho^0) + \mathcal{V}(\rho^0) + \mathcal{W}(\rho^0) < \infty$. Obviously, B is not empty.

We will prove that B is a compact set and $\rho(t) \subset B$ for all $t > 0$ by following three steps.

Step 1, we prove (i) and (ii) for $t \in [0, T(\rho^0))$. Since

$$\sum_{j=1}^{n} \frac{d\rho_j}{dt} = -\sum_{j=1}^{n} \mathrm{div}_G(\rho \nabla_G S)|_j = 0 \ ,$$

(i) is concluded. For (ii), we need to show

$$\frac{d}{dt}\mathcal{E}(\Psi(t)) = 0 \ ,$$

where $\mathcal{E}(\Psi) = \mathcal{H}(\rho, S)$. Notice (3) has the following symplectic form

$$\frac{d}{dt}\begin{pmatrix} \rho \\ S \end{pmatrix} = \mathbb{J}\begin{pmatrix} \frac{\partial}{\partial \rho}\mathcal{H} \\ \frac{\partial}{\partial S}\mathcal{H} \end{pmatrix} \ ,$$

then

$$\frac{d}{dt}\mathcal{E}(\Psi(t)) = \frac{d}{dt}\mathcal{H}(\rho(t), S(t)) = \sum_{j=1}^{n}\{\frac{\partial}{\partial \rho_j}\mathcal{H}\frac{d}{dt}\rho_j + \frac{\partial}{\partial S_j}\mathcal{H}\frac{d}{dt}S_j\}$$

$$= \sum_{j=1}^{n}\{\frac{\partial}{\partial \rho_j}\mathcal{H}\frac{\partial}{\partial S_j}\mathcal{H} - \frac{\partial}{\partial S_j}\mathcal{H}\frac{\partial}{\partial \rho_j}\mathcal{H}\} = 0 \ .$$

Step 2, we show that $\mathcal{I}(\rho)$ is positive infinity on the boundary, i.e.

$$\lim_{\min_{j \in V}\rho_j \to 0}\mathcal{I}(\rho) = +\infty \ .$$

Assume the above is not true, there exists a constant $M > 0$, such that if $\min_{i \in V}\rho_j = 0$, then

$$M \geq \mathcal{I}(\rho) = \frac{1}{2} \sum_{(j,l) \in E} \omega_{jl} (\log \rho_j - \log \rho_l)^2 \frac{\frac{1}{d_j}\rho_j + \frac{1}{d_l}\rho_l}{2}$$

$$\geq \frac{1}{4} \sum_{(j,l) \in E} \omega_{jl} (\log \rho_j - \log \rho_l)^2 \max\{\rho_j, \rho_l\} \min_k \frac{1}{d_k} \ .$$

Hence for any $(j,l) \in E$, we have

$$\omega_{jl} (\log \rho_j - \log \rho_l)^2 \max\{\rho_j, \rho_l\} \leq 2M < +\infty \ .$$

Since there exists a $j^* \in V$, such that $\rho_{j*} = 0$, the above formula implies that for any $l \in N(j^*)$, $\rho_l = 0$. Since G is connected and V is a finite set, by iterating through the nodes, we get $\rho_1 = \cdots = \rho_n = 0$, which contradicts the fact that $\sum_{j=1}^n \rho_j = 1$.

Step 3, we claim that B is a compact set. This can be easily verified because $\mathcal{I}$ is a lower semi continuous function, and $\mathcal{I}(\rho) = +\infty$ when $\rho \in \mathcal{P}(G) \setminus \mathcal{P}_o(G)$. Hence B is a compact set in $\mathbb{R}^n$.

Let us combine above three steps. Since (3) is a Hamiltonian system in $\mathcal{P}_o(G)$,

$$\mathcal{E}(\Psi(t)) = \mathcal{E}(\Psi^0) = \frac{1}{2}(\nabla_G S(t), \nabla_G S(t))_{\rho(t)} + \frac{h^2}{8}\mathcal{I}(\rho(t)) + \mathcal{V}(\rho(t)) + \mathcal{W}(\rho(t)) \ ,$$

then

$$\frac{h^2}{8}\mathcal{I}(\rho(t)) = \mathcal{E}(\Psi^0) - \frac{1}{2}(\nabla_G S(t), \nabla_G S(t))_{\rho(t)} - (\mathcal{V}(\rho(t)) + \mathcal{W}(\rho(t)))$$

$$\leq \mathcal{E}(\Psi^0) - \min_{\rho \in \mathcal{P}(G)} [\mathcal{V}(\rho) + \mathcal{W}(\rho)] \ .$$

Thus $\rho(t) \in B \subset \mathcal{P}_o(G)$ for all $t > 0$. $\quad\square$

Next, we prove (iii) and (iv). For (iii), since $\Psi_j = \sqrt{\rho_j} e^{i\frac{S_j}{h}}$, its conjugate $\bar{\Psi}$ satisfies

$$\bar{\Psi}_j = \sqrt{\rho_j} e^{i\frac{\bar{S}_j}{h}} \quad \text{with } \bar{S}_j = -S_j \ .$$

Let us look at (3) by changing t to $-t$.

$$\begin{cases} -\dfrac{d\rho_j}{dt} + \displaystyle\sum_{l \in N(j)} \omega_{lj}(S_l - S_j)\theta_{lj}(\rho) = 0 \ ; \\ -\dfrac{dS_j}{dt} + \dfrac{1}{2} \displaystyle\sum_{j \in N(i)} \omega_{jl}(S_j - S_l)^2 \dfrac{\partial \theta_{jl}}{\partial \rho_j} + \dfrac{\partial}{\partial \rho_j}\{\dfrac{h^2}{8}\mathcal{I}(\rho) + \mathcal{W}(\rho) + \mathcal{V}(\rho)\} = 0 \ . \end{cases}$$

Denote $\bar{S} = -S$, then $(\rho(t), S(t))$ and $(\rho(-t), \bar{S}(-t))$ satisfies (3).

For (iv), if $V \to V^\alpha = V + \alpha$, we substitute $\Psi^\alpha(t) = \Psi(t)e^{i\frac{\alpha t}{\hbar}}$ into (3) to get:

$$
\begin{cases}
\dfrac{d\rho_j}{dt} + \displaystyle\sum_{l \in N(j)} \omega_{jl}[(S_l + \alpha t) - (S_j + \alpha t)]\theta_{jl}(\rho) = 0 \; ; \\[3mm]
\dfrac{d}{dt}(S_j + \alpha t) + \dfrac{1}{2}\displaystyle\sum_{l \in N(j)} \omega_{jl}[(S_l + \alpha t) - (S_j + \alpha t)]^2 \dfrac{\partial \theta_{jl}}{\partial \rho_j} + \mathbb{V}_j + \dfrac{\partial}{\partial \rho_j}\{\dfrac{h^2}{8}\mathcal{I}(\rho) + \mathcal{W}(\rho)\} \\[3mm]
\qquad = 0 \; .
\end{cases}
$$

This means that if (ρ, S) are solutions of (3) with V, then $S^\alpha(t) = S + \alpha t$, $\rho^\alpha(t) = \rho(t)$ are solutions of (3) with V^α, i.e.

$$
\Psi^\alpha = \sqrt{\rho^\alpha}e^{i\frac{S^\alpha}{\hbar}} = \sqrt{\rho}e^{i\frac{S}{\hbar}}e^{i\frac{\alpha t}{\hbar}} = \Psi e^{i\frac{\alpha t}{\hbar}} \; .
$$

(v). On one hand, suppose Ψ^* satisfies (3), then

$$
\nu = \frac{\partial}{\partial \rho_j}\left(\frac{h^2}{8}\mathcal{I}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)\right)\big|_{\rho = \rho^*} \; , \quad \text{for any } i \in V \; .
$$

Notice that $\mathcal{E}(\sqrt{\rho}) = \frac{h^2}{8}\mathcal{I}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)$. It is simple to check that ρ^* satisfies the Karush–Kuhn–Tucker conditions of minimization

$$
\min_{\rho}\{\mathcal{E}(\sqrt{\rho}) : \sum_{j=1}^{n} \rho_j = 1, \quad \rho_j > 0\} \; ,
$$

with ν being the Lagrange multiplier. Next, we show

$$
\begin{aligned}
\nu &= \sum_{j=1}^{n} \nu \rho_j^* = \sum_{j=1}^{n} \frac{\partial}{\partial \rho_j}\{\frac{h^2}{8}\mathcal{I} + \mathcal{V} + \mathcal{W}\}\big|_{\rho^*} \cdot \rho_j^* \\[2mm]
&= \frac{h^2}{8}\sum_{j=1}^{n} \frac{\partial}{\partial \rho_j}\mathcal{I}\big|_{\rho^*}\rho_j^* + \sum_{j=1}^{n}[\mathbb{V}_j + \sum_{l=1}^{n}\mathbb{W}_{jl}\rho_l^*]\rho_j^* \\[2mm]
&= \frac{h^2}{8}\sum_{j=1}^{n} \frac{\partial}{\partial \rho_j}\mathcal{I}\big|_{\rho^*}\rho_j^* - \frac{h^2}{8}\mathcal{I}(\rho^*) + \frac{h^2}{8}\mathcal{I}(\rho^*) + \mathcal{V}(\rho^*) + \mathcal{W}(\rho^*) + \mathcal{W}(\rho^*) \\[2mm]
&= \frac{h^2}{8}\left(\sum_{j=1}^{n} \frac{\partial}{\partial \rho_j}\mathcal{I}\big|_{\rho^*}\rho_j^* - \mathcal{I}(\rho^*)\right) + \mathcal{E}(\sqrt{\rho^*}) + \mathcal{W}(\rho^*) \\[2mm]
&= \mathcal{E}(\sqrt{\rho^*}) + \mathcal{E}_{int}(\sqrt{\rho^*}) \; ,
\end{aligned}
\tag{27}
$$

where the last equality is from the fact: $\mathcal{E}_{int}(\sqrt{\rho}) = \mathcal{W}(\rho)$ and $\mathcal{I}(\rho) = \sum_{j=1}^{n} \frac{\partial}{\partial \rho_j}\mathcal{I}(\rho)\rho_j$. On the other hand, suppose ρ^* is a critical point of $\mathcal{E}(\sqrt{\rho})$, then there exists $\lambda \in \mathbb{R}$, such that $\lambda = \frac{\partial}{\partial \rho_j}\{\frac{h^2}{8}\mathcal{I} + \mathcal{V} + \mathcal{W}\}\big|_{\rho^*}$. Since $\nu = \mathcal{E}(\sqrt{\rho^*}) + \mathcal{E}_{int}(\sqrt{\rho^*})$, we have $\lambda =$

$\sum_{j=1}^{n} \frac{\partial}{\partial \rho_j} \{\frac{h^2}{8} \mathcal{I} + \mathcal{V} + \mathcal{W}\}|_{\rho^*} \cdot \rho_j^* = \mathcal{E}(\sqrt{\rho^*}) + \mathcal{E}_{int}(\sqrt{\rho^*}) = \nu$. Thus $\Psi^*(t) = \sqrt{\rho^*} e^{-i\nu t}$ is an interior stationary solution. $\square$

5.1. Ground states

Similar to Nelson's idea in [23], we show that the stationary solution of (3) in Theorem 6 (v) is related to the discrete ground state of the NLS.

Corollary 7. *If* $\mathbb{W}$ *is a semi positive definite matrix, then the stationary state is a ground state* $\Psi^g = \sqrt{\rho^g} e^{-i\nu^g t}$ *if and only if*

$$\Psi^g = \arg\min_{\Psi} \{ \ \mathcal{E}(\Psi) \ : \ \sum_{j=1}^{n} |\Psi_j|^2 = 1 \} \ , \tag{28}$$

with

$$\rho^g = \arg\min_{\rho \in \mathcal{P}(G)} \mathcal{E}(\sqrt{\rho}) \quad and \quad \nu^g = \mathcal{E}(\sqrt{\rho^g}) + \mathcal{E}_{int}(\sqrt{\rho^g}) \ .$$

Proof. From Theorem 6 (v), ρ^g is a critical point of $\min_{\rho \in \mathcal{P}_o(G)} \mathcal{E}(\sqrt{\rho})$ and ν^g is defined as above. We only need to prove Ψ^g is the minimizer of problem (28). In fact,

$$\min_{\Psi} \{ \ \mathcal{E}(\Psi) \ : \ \sum_{j=1}^{n} |\Psi_j|^2 = 1 \} \geq \min_{\rho \in \mathcal{P}(G)} \mathcal{E}(\sqrt{\rho}) \ ,$$

because

$$\mathcal{E}(\Psi) = \frac{1}{4} \sum_{(j,l) \in E} \omega_{jl}(S_j - S_l)^2 \theta_{jl}(\rho) + \frac{h^2}{8} \mathcal{I}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)$$

$$\geq \frac{h^2}{8} \mathcal{I}(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho) = \mathcal{E}(\sqrt{\rho}) \ ,$$

the equality holds if and only if $S_j = S_l$, for any $(j, l) \in E$. Since G is a connected graph, then a ground state $\Psi^g = \sqrt{\rho^g} e^{i \frac{S^g}{h}}$ has the following structure:

$$\rho^g = \arg\min_{\rho \in \mathcal{P}(G)} \mathcal{E}(\sqrt{\rho}) \quad and \quad S_1^g = S_2^g = \cdots = S_n^g \ .$$

Next, we show that the function $\mathcal{E}(\sqrt{\rho}) = \frac{h^2}{8} \mathcal{I}(\rho) + \mathcal{W}(\rho) + \mathcal{V}(\rho)$ is strictly convex. If this is true, we can conclude that ρ^g is a unique minimizer, which is the ground state.

Notice that $\mathcal{W}(\rho) = \frac{1}{2} \sum_{j=1}^{n} \sum_{l=1}^{n} \mathbb{W}_{jl} \rho_j \rho_l$, $\mathcal{V}(\rho) = \sum_{j=1}^{n} \mathbb{V}_j \rho_j$ are convex functionals. So we only need to prove

$$\mathcal{I}(\rho) \ \textit{is a strict convex functional in } \mathcal{P}_o(G).$$

We show this result by proving

$$\min_{\sigma \in T_\rho \mathcal{P}_o(G)} \{\sigma^T \mathrm{Hess}_{\mathbb{R}^n} \mathcal{I}(\rho)\sigma \; : \; \sigma^T \sigma = 1\} > 0 \,. \tag{29}$$

Since the Hessian matrix of $\mathcal{I}$ is

$$\frac{\partial^2}{\partial \rho_l \partial \rho_j} \mathcal{I}(\rho) = \begin{cases} -\frac{1}{\rho_l \rho_j} \omega_{lj} t_{lj} & \text{if } l \in N(j) \,; \\ \frac{1}{\rho_j^2} \sum_{l \in N(j)} \omega_{lj} t_{lj} & \text{if } l = j \,; \\ 0 & \text{otherwise} \,, \end{cases}$$

where

$$t_{lj} = (\frac{1}{d_l}\rho_l - \frac{1}{d_j}\rho_j)(\log \rho_l - \log \rho_j) + (\frac{1}{d_l}\rho_l + \frac{1}{d_j}\rho_j) > 0 \,, \tag{30}$$

hence

$$\sigma^T \mathrm{Hess}_{\mathbb{R}^n} \mathcal{I}(\rho)\sigma = \frac{1}{2} \sum_{(l,j) \in E} t_{lj} \{(\frac{\sigma_j}{\rho_j})^2 + (\frac{\sigma_l}{\rho_l})^2 - 2\frac{\sigma_l}{\rho_l}\frac{\sigma_j}{\rho_j}\}$$

$$= \frac{1}{2} \sum_{(l,j) \in E} t_{lj} (\frac{\sigma_j}{\rho_j} - \frac{\sigma_l}{\rho_l})^2 \geq 0 \,.$$

So $\mathrm{Hess}_{\mathbb{R}^n} \mathcal{I}$ is a semi-positive definite matrix.

Suppose (29) is not true, there exists a unit vector $\sigma^* \in T_\rho \mathcal{P}_o(G)$, such that

$$\sigma^{*T} \mathrm{Hess}_{\mathbb{R}^n} \mathcal{I}(\rho)\sigma^* = \frac{1}{2} \sum_{(l,j) \in E} t_{lj} (\frac{\sigma_l^*}{\rho_l} - \frac{\sigma_j^*}{\rho_j})^2 = 0 \,.$$

Then $\frac{\sigma_1^*}{\rho_1} = \frac{\sigma_2^*}{\rho_2} = \cdots \frac{\sigma_n^*}{\rho_n} = 0$. Combining with $\sum_{j=1}^n \sigma_j^* = 0$, we have $\sigma_1^* = \sigma_2^* = \cdots = \sigma_n^* = 0$, which contradicts that σ^* is a unit vector. $\square$

It is worth mentioning that we have the following eigenvalue problem at the ground state:

$$\nu \Psi_j = -\frac{h^2}{2}\Delta_G \Psi_j + \mathbb{V}_j \Psi_j + \Psi_j \sum_{l=1}^n \mathbb{W}_{jl}|\Psi_l|^2 \,, \quad (\Psi_j)_{j=1}^n \in \mathbb{R}^n \,. \tag{31}$$

The solution of (31) is the ground state configuration, where $|\Psi| = \sqrt{\rho^g}$ and $\nu = \mathcal{E}(\sqrt{\rho^g}) + \mathcal{E}_{int}(\sqrt{\rho^g})$ is the associated energy level.

6. Linearized problems

In this section, we study the linearized problem near the ground state. Consider the Hamiltonian system

$$\frac{d}{dt}\begin{pmatrix} \rho \\ S \end{pmatrix} = \mathbb{J}\begin{pmatrix} \frac{\partial}{\partial\rho}\mathcal{H} \\ \frac{\partial}{\partial S}\mathcal{H} \end{pmatrix} .$$

The ground state $(\rho^g, S^g(t))$ can be viewed as its equilibrium solution, which is clearly the critical point of Hamiltonian

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

Consider the linearized problem of (3)

$$\frac{d}{dt}z = H^{(2)}z ,$$

where $z \in \mathbb{R}^{2n}$, $H^{(2)} \in \mathbb{R}^{2n \times 2n}$ is the *Hamiltonian matrix* at the equilibrium (ρ^g, S^g). Because $(S_j^g)_{j=1}^n$ is a constant vector, we obtain a simple structure for $H^{(2)}$:

$$H^{(2)} := \mathbb{J} \cdot \mathrm{Hess}_{\mathbb{R}^{2n}}\mathcal{H}(\rho, S)|_{(\rho^g, S^g)} = \begin{pmatrix} 0 & L_2(\rho^g) \\ -\mathbb{W} - \frac{h^2}{8}\mathrm{Hess}_{\mathbb{R}^n}\mathcal{I}(\rho^g) & 0 \end{pmatrix} . \quad (32)$$

We estimate the eigenvalue of (32) in a particular case:

$$hi\frac{d\Psi_j}{dt} = -\frac{h^2}{2}\Delta_G\Psi|_j + \alpha\Psi_j|\Psi_j|^2 , \quad (33)$$

where G is a d-dimensional Torus graph. This equation is obtained from (24) by taking $\mathbb{V} = 0$ and $\mathbb{W} = \alpha\mathbb{I}$. It can be viewed as a discrete version of Gross–Pitaevskii equation (GPE), which has been proposed to model the Bose–Einstein condensate.

Proposition. *For discrete GPE* (33), *$H^{(2)}$ has eigenvalues*

$$\alpha_k^+ := +i\sqrt{\frac{1}{4}\lambda_k^2 h^2 + \frac{\alpha\lambda_k}{n}} , \quad \alpha_k^- := -i\sqrt{\frac{1}{4}\lambda_k^2 h^2 + \frac{\alpha\lambda_k}{n}} ,$$

with associated eigenvectors

$$w_k^+ := \begin{pmatrix} v_k \\ i\sqrt{\frac{n^2h^2}{4} + \frac{\alpha n}{\lambda_k}}v_k \end{pmatrix} , \quad w_k^- := \begin{pmatrix} iv_k \\ \sqrt{\frac{n^2h^2}{4} + \frac{\alpha n}{\lambda_k}}v_k \end{pmatrix} \in \mathbb{C}^{2n} ,$$

i.e.

$$H^{(2)}w_k^+ = \alpha_k^+ w_k^+ , \quad H^{(2)}w_k^- = \alpha_k^- w_k^- .$$

Here $Lv_k = \lambda_k v_k$. $\lambda_k \geq 0$, $v_k \in \mathbb{R}^n$ are k-th eigenvalue and eigenvector of graph Laplacian matrix $L = (L(j,l))_{1 \leq j,l \leq n} \in \mathbb{R}^{n \times n}$, where $L(j,l) = \begin{cases} -w_{jl} & \text{if } l \in N(j); \\ \sum_{l \in N(j)} w_{jl} & \text{if } l = j; \\ 0 & \text{otherwise.} \end{cases}$

Proof. Denote $\mathbf{1} = (\frac{1}{n})_{j=1}^n$. From Karush–Kuhn–Tucker conditions, one can easily find $\rho^g = \mathbf{1}$ is the critical point of $\{\frac{h^2}{8}\mathcal{I}(\rho) + \frac{\alpha}{2}\sum_{j=1}^n \rho_j^2 : \rho \in \mathcal{P}(G)\}$. In this case, $L_1(\mathbf{1}) = 2nL$, $L_2(\mathbf{1}) = \frac{1}{n}L$, where $L = \frac{1}{d}D^T D$. So the matrix (32) becomes

$$H^{(2)} = \mathbb{J} \cdot \text{Hess}_{\mathbb{R}^{2n}} \mathcal{H}(\rho, S)|_{(\mathbf{1},S^g)} = \begin{pmatrix} 0 & \frac{1}{n}L \\ -\alpha\mathbb{I} - \frac{nh^2}{4}L & 0 \end{pmatrix} . \tag{34}$$

In fact, we can find all eigenvalues and eigenvectors of (32). Notice that $L \in \mathbb{R}^{n \times n}$ is a semi-positive matrix, and denote $\lambda_k \in \mathbb{R}$, $v_k \in \mathbb{R}^n$, as the k-th eigenvalue and eigenvector of L. Therefore one can check that

$$H^{(2)}w_k^+ = \begin{pmatrix} 0 & \frac{1}{n}L \\ -\alpha\mathbb{I} - \frac{nh^2}{4}L & 0 \end{pmatrix} \begin{pmatrix} v_k \\ i\sqrt{\frac{n^2h^2}{4} + \frac{\alpha n}{\lambda_k}} v_k \end{pmatrix}$$

$$= \begin{pmatrix} i\frac{1}{n}\sqrt{\frac{n^2h^2}{4} + \frac{\alpha n}{\lambda_k}} \cdot \lambda_k v_k \\ -\alpha v_k - \frac{nh^2}{4}\lambda_k v_k \end{pmatrix}$$

$$= +i\sqrt{\frac{\lambda_k^2 h^2}{4} + \frac{\alpha\lambda_k}{n}} \begin{pmatrix} v_k \\ i\sqrt{\frac{n^2h^2}{4} + \frac{\alpha n}{\lambda_k}} v_k \end{pmatrix}$$

$$= \alpha_k^+ w_k^+$$

Similarly, $H^{(2)}w_k^- = \alpha_k^- w_k^-$. $\square$

From these eigenvalues, when $\alpha > -\frac{n}{4}\lambda_k h^2$, the solution $\rho = \mathbf{1}$, $\nabla_G S = 0$ is stable for (33). When $\alpha = -\frac{n}{4}\lambda_k h^2$, bifurcations may happen.

7. Examples

Finally, we demonstrate (3) and (31) by two numerical examples.

Example 1 *(NLS on a two points graph)*. Consider a Hamiltonian:

$$\mathcal{H}(\rho, S) = \frac{1}{2}(S_1 - S_2)^2 \theta_{12}(\rho) + \frac{h^2}{8}(\log \rho_1 - \log \rho_2)^2 \theta_{12}(\rho) + \mathbb{V}_1 \rho_1 + \mathbb{V}_2 \rho_2 ,$$

where $\mathbb{V}_1 = \mathbb{V}_2 = c$. In this case, the solution of (3), $(\rho_1(t), \rho_2(t), S_1(t) - S_2(t)) \in \mathbb{R}^3$, can be plotted using a phase portrait. In Fig. 1, each circle represents a trajectory of (3).
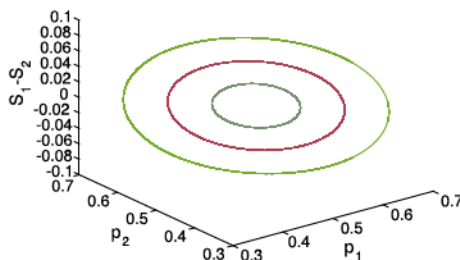
Fig. 1. The phase portrait of $(\rho_1(t), \rho_2(t), S_1(t) - S_2(t))$ with different initial conditions.
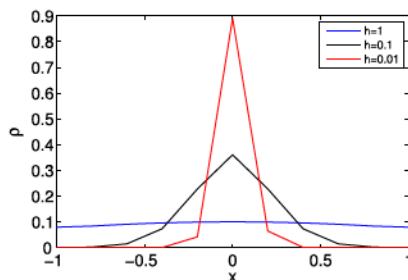

Fig. 2. The plot of ground state ρ^g. The blue, black, red curves represent $h = 1, 0.1, 0.01$, respectively. The color version of the figure is on the web version of this paper.

The ground state $(\frac{1}{2}, \frac{1}{2}, 0)$ is in the center of all these circles, so it is spectrally stable.

Example 2 *(Ground state)*. We demonstrate the ground states on a 1-D lattice graph. Set $\mathbb{W} = 0$. Consider the following minimization problem

$$\rho^g = \arg \min_{\rho \in \mathcal{P}(G)} \sum_{j=1}^{n} \mathbb{V}_j \rho_j + \frac{h^2}{8} \mathcal{I}(\rho) ,$$

where $\mathbb{V}_j = \frac{x_j^2}{2}$. We compute the above minimizer numerically [17] in the interval $[-5, 5]$ with $n = 20$. From Fig. 2, we observe that the ground state approaches to the delta measure supported at 0 when $h \to 0$. This captures the exactly same effect in continuous states, in which the ground state is a Gaussian distribution with variance h^2 [2].

8. Conclusions

In this paper we have introduced a new NLS on finite graphs (3). Compared to the existing work, (3) has distinct features: First, the discrete NLS is introduced via discrete optimal transport. This formulation provides a way to study the discrete NLS from geometric viewpoint. The inverse of the weighted Laplacian $L(\rho)^{-1}$ induces the metric on the probability manifold. Second, the discrete Fisher information $\mathcal{I}(\rho)$ is applied to construct Hamiltonian system. As Nelson deduced in his work, introducing the Fisher information to the Lagrangian is equivalent to adding noise perturbations to the paths,

yet the Hamiltonian structure is still retained. This provides a valuable avenue to study random perturbations to Hamiltonian systems, which is one of the motivations for us to carry out the present investigation. Compared to other existing ways of adding random perturbations, which often destroy the Hamiltonian structure, this strategy has preferred features allowing (3) to mimic many interesting properties of the Schrödinger equation in the continuous case, such as conservation mass and energy. Last but not the least, it introduces the ground state on graph given in (31). Studying the stability problem around the discrete ground state introduces a Hamiltonian matrix, which is a symplectic composition of two modified graph Laplacian matrices. These give insights of the system that can be explored in the future.

Acknowledgment

References

[1] Luigi Ambrosio, Wilfrid Gangbo, Hamiltonian ODEs in the Wasserstein space of probability measures, Comm. Pure Appl. Math. 61 (1) (2008) 18–53.

[2] Xavier Antoine, Weizhu Bao, Christophe Bessee, Computational methods for the dynamics of the nonlinear Schrödinger/Gross–Pitaevskii equations, Comput. Phys. Commun. 1 (23) (2013).

[3] Guido Bacciagaluppi, A conceptual introduction to Nelson's mechanics, in: Endophysics, Time, Quantum and the Subjective, 2005, pp. 367–388.

[4] Jean-David Benamou, Yann Brenier, A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem, Numer. Math. 84 (3) (2000) 375–393.

[5] Jean Bourgain, Abel Klein, Bounds on the density of states for Schrödinger operators, Invent. Math. 194 (1) (2013) 41–72.

[6] Eric Carlen, Stochastic mechanics: a look back and a look ahead, in: Diffusion, Quantum Theory and Radically Elementary Mathematics, vol. 47, 2014, pp. 117–139.

[7] Eric Carlen, Jan Maas, An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker–Planck equation is gradient flow for the entropy, Comm. Math. Phys. 331 (3) (2014) 887–926.

[8] Eric Carlen, Jan Maas, Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance, J. Funct. Anal. 273 (5) (2017).

[9] Sourav Chatterjee, Kay Kirkpatrick, Probabilistic methods for discrete nonlinear Schrödinger equations, Comm. Pure Appl. Math. 65 (5) (2010) 727–757.

[10] Shui-Nee Chow, Wen Huang, Yao Li, Haomin Zhou, Fokker–Planck equations for a free energy functional or Markov process on a graph, Arch. Ration. Mech. Anal. 203 (3) (2012) 969–1008.

[11] Shui-Nee Chow, Wuchen Li, Haomin Zhou, Entropy dissipation of Fokker–Planck equations on graphs, Discrete Contin. Dyn. Syst., Ser. A 38 (10) (2018) 4929–4950.

[12] Pierre Degond, Samy Gallego, Florian Méhats, An asymptotic preserving scheme for the Schrödinger equation in the semiclassical limit, C. R. Math. 345 (9) (2007) 531–536.

[13] J. Chris Eilbeck, Magnus Johansson, The Discrete Nonlinear Schrödinger Equation–20 Years on Proceedings of the Third Conference: Localization & Energy Transfer in Nonlinear Systems, San Lorenzo de El Escorial, Madrid, 17–21 June, 2002.

[14] B. Roy Frieden, Science from Fisher Information: A Unification, Cambridge University Press, 2004.

[15] Wilfrid Gangbo, Truyen Nguyen, Adrian Tudorascu, Hamilton–Jacobi equations in the Wasserstein space, Methods Appl. Anal. 15 (2) (2008) 155–184.

[16] John D. Lafferty, The density manifold and configuration space quantization, Trans. Amer. Math. Soc. 305 (2) (1988) 699–741.

[17] Wuchen Li, A Study of Stochastic Differential Equations and Fokker–Planck Equations with Applications, PhD thesis, Georgia Institute of Technology, 2016.
[18] Jan Maas, Gradient flows of the entropy for finite Markov chains, J. Funct. Anal. 261 (8) (2011) 2250–2292.
[19] E. Madelung, Quanten theorie in Hydrodynamischer Form, Z. Phys. 40 (1927) 322.
[20] Alexander Mielke, A gradient structure for reaction–diffusion systems and for energy-drift-diffusion, Nonlinearity 24 (4) (2011) 13–29.
[21] A. Mielke, M.A. Peletier, D.R.M. Renger, On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion, Potential Anal. 41 (2014) 1293–1327.
[22] Markus Mittnenzweig, Alexander Mielke, An entropic gradient structure for Lindblad equations and couplings of quantum systems to macroscopic models, J. Stat. Phys. 167 (2) (2017).
[23] Edward Nelson, Derivation of the Schrödinger equation from Newtonian mechanics, Phys. Rev. 150 (4) (1966) 1079.
[24] Edward Nelson, Quantum Fluctuations, Princeton University Press, 1985.
[25] Diego Noja, Nonlinear Schrödinger equation on graphs: recent results and open problems, Philos. Trans. R. Soc. (2014).
[26] von Max-K Renesse, An optimal transport view on Schroedinger's equation, Canad. Math. Bull. 55 (4) (2008).
[27] E. Schrödinger, Quantisierung als Eigenwertproblem (zweite Mitteilung), Ann. Phys. 79 (1926) 489–527.
[28] Cédric Villani, Optimal Transport: Old and New, vol. 338, Springer Science & Business Media, 2008.