SIAM J. APPL. MATH. © 2023 Society for Industrial and Applied Mathematics
Vol. 83, No. 2, pp. 484–509

WASSERSTEIN HAMILTONIAN FLOW WITH COMMON NOISE ON GRAPH*

JIANBO CUI†, SHU LIU‡, AND HAOMIN ZHOU‡

Abstract. We study the Wasserstein Hamiltonian flow with a common noise on the density manifold of a finite graph. Under the framework of the stochastic variational principle, we first develop the formulation of stochastic Wasserstein Hamiltonian flow and show the local existence of a unique solution. We also establish a sufficient condition for the global existence of the solution. Consequently, we obtain the global well-posedness for the nonlinear Schrödinger equations with common noise on a graph. In addition, using Wong–Zakai approximation of common noise, we prove the existence of the minimizer for an optimal control problem with common noise. We show that its minimizer satisfies the stochastic Wasserstein Hamiltonian flow on a graph as well.

Key words. stochastic Hamiltonian flow on graph, density manifold, Wong–Zakai approximation, optimal transport

MSC codes. 58B20, 58J65, 35Q41, 49Q20

DOI. 10.1137/22M1490697

1. Introduction.

The Wasserstein Hamiltonian flow defined on the cotangent bundle of a probability density manifold, also known as the Wasserstein manifold in the literature, has been studied in the past few years (see, e.g., [27, 3, 23, 13]). Its relationship with the Hamiltonian ordinary differential equations (ODEs) has also been well demonstrated via optimal transport theory (see, e.g., [39, 11, 12]). Furthermore, it has been used in the theoretical or numerical analysis of the nonlinear Schrödinger equation (see, e.g., [34, 35, 36, 11, 16]), mass optimal transport (see, e.g., [5, 24, 16, 14]), and the Schrödinger bridge problem (see, e.g., [30, 29, 9, 18]). Extending Wasserstein Hamiltonian flow to account for random perturbations is challenging because not all types of noise can be used to perturb the dynamics on a density manifold in which the nonnegativity of the probability density function and mass conservation must be preserved. Recently, using the concept of common noise, also referred as environment or system noise [7, 6], a stochastic version of Wasserstein Hamiltonian flow is introduced to understand the collective dynamical behavior on a density manifold of the stochastic Hamiltonian ODE defined on continuous phase space [17]. However, little is known if the underlying space becomes discrete, such as a finite graph or a spatial discretization of a continuous space, due to several significant challenges that arise in the discrete space.

Unlike the continuous space, where stochastic Hamiltonian ODEs can be identified and interpreted as the particle dynamics corresponding to the stochastic Wasserstein

*Received by the editors April 18, 2022; accepted for publication (in revised form) December 1, 2022; published electronically April 20, 2023.
https://doi.org/10.1137/22M1490697
Funding: This research is partially supported by Georgia Tech Mathematics Application Portal (GT-MAP) and by research grants NSF DMS-1830225, and ONR N00014-21-1-2891. The research of the first author is partially supported by start-up funds (P0039016, P0041274) from Hong Kong Polytechnic University, the Hong Kong Research Grant Council ECS grant 25302822, and the CAS AMSS-PolyU Joint Laboratory of Applied Mathematics.

†Corresponding author. Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (jianbo.cui@polyu.edu.hk).
‡School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA (sliu459@gatech.edu, hmzhou@gatech.edu).

Hamiltonian flow, such a particle correspondence has not been established in the discrete space, which prevents adopting many well-developed techniques to the discrete case. For example, the particle version of stochastic Hamiltonian ODEs has been used as a push-forward map, a crucial tool in the analysis, to study the dynamical properties on the density manifold [12]. This tool is hard to generalize to a general graph partially because not all of the graph can be embedded into a continuous space [18]. Due to the loss of particle formulation, it is still unclear what kind of noise or random perturbation on the finite graph can be used as a functional replacement of the white noise in the continuous space. In addition, low regularity of noise and the discrete structure of a graph make it harder to analyze the dynamical properties of a Hamiltonian system on a graph.

In this paper, we propose two different strategies to establish the Wasserstein Hamiltonian flow with common noise on the finite graph and investigate their mathematical properties. The first approach is based on the discrete version of the generalized stochastic variational principle, which provides a formulation to construct a stochastic Wasserstein Hamiltonian flow with given initial values. We use the stopping time technique to show its local well-posedness. Using the Poisson bracket, we provide a sufficient condition on the energy terms in the variational principle to ensure the global well-posedness for the resulting system. We further demonstrate that both the nonlinear Schrödinger equation and the logarithmic Schrödinger equations with common noise on a graph satisfy this sufficient condition. Thus they possess global solutions uniquely. In this consideration, it is observed that the Fisher information plays a fundamental role in obtaining the global existence result. This study on stochastic Wasserstein Hamiltonian flows on a graph will be useful for designing novel structure-preserving numerical schemes for stochastic Hamiltonian partial differential equation, such as the stochastic nonlinear Schrödinger equation which emerged from nonlinear optics (see, e.g., [4, 22, 26, 15]), and their related stochastic optimal control problems [17, 19].

The second approach to derive the boundary value formulation of Wasserstein Hamiltonian flow with common noise on a graph is proposed in the framework of stochastic optimal control. Using the Wong–Zakai approximation [41, 40] of common noise and von Neumann's minimax theorem [37], we prove the existence of a minimizer for the stochastic optimal control problems. Under suitable assumptions, we show that their critical point satisfies Wasserstein Hamiltonian flow with common noise on a graph. In addition, the system obtained by the stochastic optimal control approach exhibits highly consistent formulation as do those constructed by using the stochastic variational principle. Yet, they have interesting differences, especially when the local well-posedness for the latter one is no longer valid. In our investigation, these two strategies are complementary to each other in exploring the properties of stochastic Wasserstein Hamiltonian flow on a graph. In particular, we present two stochastic OT(optimal transport) formulations to show the influence of common noise. When the diffusion coefficient in the constraint does not depend on the control (the velocity v), the limit of the stochastic OT problem with Wong–Zakai approximation could be characterized by the stochastic Wasserstein Hamiltonian flow, but its θ-connected component (see [31, 24]) may be different from the deterministic case. In contrast, when the diffusion coefficient in the constraint depends on the control, then the limit of the stochastic OT problem with Wong–Zakai approximation is still unclear and its θ-connected component is the same as the deterministic case.

The organization of this paper is as follows. In section 2, we discuss what the common noise is and why it is used in our study. In section 3, we review the basic notations of the deterministic Wasserstein Hamiltonian flow on a finite graph and the discrete optimal transport theory. In section 4, we present the discrete generalized stochastic variational principle to derive the stochastic Wasserstein Hamiltonian flow on a graph and study several properties of the stochastic Wasserstein Hamiltonian flow. In section 5, we give an alternative way based on stochastic optimal control to derive the stochastic Wasserstein Hamiltonian flow on a graph. Meanwhile, we show the existence of the minimizer and derive its equation using the Wong–Zakai approximation.

2. Common noise. In this section, we borrow some examples to explain what common noise is and why it is a good choice for us to consider here.

The first example is a mean-field game model (see, e.g., [7]). Consider an N-player differential game; the state of each player $X_i(t)$ is a stochastic process described by a stochastic differential equation

$$dX_i(t) = b(t, X_i(t), \mu(t), \alpha_i)dt + \sigma(t, X_i(t), \mu(t))dB_i(t) + \sigma_0(t, X_i(t), \mu(t))dW(t),$$

where b, σ, σ_0 are given functions, α_i is a control variable, $\mu(t) = \frac{1}{N}\sum_{j=1}^N \delta_{X_i(t)}$, and $B_i (i = 1, \ldots, N)$ and W are a one-dimensional independent Brownian motion defined on a completed probability space $(\Omega, \mathbb{P}, \mathcal{F})$. In this model, the Brownian motion B_i is called the idiosyncratic noise, which is introduced to model random perturbations to each individual, while W is a stochastic perturbation independent of individuals (i.e., the same W for all $X_i(t)$) and it is used to model the common disturbance to all players, hence it is called common noise. When $N \to \infty$, μ tends to a random measure reflecting the aggregate behavior of all players. μ is independent of B_i while depending on the common noise W, because the effect from B_i is averaged out but not that for W. In this sense, $\mu(t)$ is a random measure flow perturbed by the common noise W. Conditioned on W, the model recovers the standard mean-field game formulation (see the pioneering works [28, 25]).

The second example is the stochastic nonlinear Schrödinger equation which emerged from nonlinear optics, hydrodynamics, and plasma physics. For instance, in the molecular monolayers arranged in Scheibe aggregates [4, 22], the thermal fluctuations of the phonons are included, which results in a stochastic nonlinear dynamical model given by

$$du = \mathbf{i}\Delta u dt + \lambda \mathbf{i}|u|^2 u dt + \mathbf{i}u \circ dW_t.$$

Here $\lambda \in \mathbb{R}$ is a constant, W is a Wiener process on an infinite-dimensional space, and $\circ$ means that the stochastic integral is taken in the Stratonovich sense. The numerical simulations based on this stochastic model coincide with experimental results reported in [33] when temperatures are lower than $3K$.

Another model, called the nonlinear Schrödinger equation with random dispersion,

$$du = \mathbf{i}\Delta u \circ dW_t + \mathbf{i}\lambda|u|^2 u,$$

is proposed to describe the propagation of signal (see, e.g., [1]), in which W is a standard one-dimensional Brownian motion. In a recent study [17], by using the Madelung transformation $u = \sqrt{\rho}e^{\mathbf{i}S}$ and the stochastic variational principle on the density manifold, it is found that the mathematically equivalent systems in terms of ρ, S, and W for the above two stochastic nonlinear Schrödinger equations can be established. Under this viewpoint, W is a random noise acting on the density function ρ.

Therefore, it is a common noise, because it perturbs the entire density, not an individual particle. In the mean-field game model and nonlinear Schrödinger equations, both μ and ρ remain probability density functions, despite the perturbations by the common noise W. In other words, nonnegativity as well as mass can be preserved under common noise perturbations. Inspired by those examples, we select common noise to establish the stochastic Wasserstein Hamiltonian flow on graphs.

3. Discrete optimal transport and discrete Wasserstein Hamiltonian flow. In this section, we introduce the notations and some known results for the discrete optimal transport problem and Wasserstein Hamiltonian flow [11, 16].

Consider a graph $G=(V,E,\omega)$ with a node set $V=\{a_i\}_{i=1}^N$, an edge set E, and ω_{jl} are the weights of the edges: $\omega_{jl}=\omega_{lj}>0$ if there is an edge between a_j and a_l, and 0 otherwise. Below, we will write $(i,j)\in E$ to denote the edge in E between the vertices a_i and a_j. Throughout the paper, we assume that G is an un-directed, connected graph with no self-loops or multiple edges.

Let us denote the set of discrete probabilities on a graph by $\mathcal{P}(G)$,

$$\mathcal{P}(G)=\left\{(\rho)_{j=1}^N\ :\sum_{j=1}^N \rho_j=1,\rho_j\geq 0 \text{ for } a_j\in V\right\},$$

and let $\mathcal{P}_o(G)$ be its interior (i.e., all $\rho_j>0$ for $a_j\in V$). Let $\mathbb{V}_j$ be a linear potential on each node a_j, and $\mathbb{W}_{jl}=\mathbb{W}_{lj}$ an interactive potential between nodes a_j,a_l. The total linear potential $\mathcal{V}$ and interaction potential $\mathcal{W}$ are given by

$$\mathcal{V}(\rho)=\sum_{i=1}^N \mathbb{V}_i\rho_i,\ \mathcal{W}(\rho)=\frac{1}{2}\sum_{i,j=1}^N \mathbb{W}_{ij}\rho_i\rho_j.$$

We let $N(i)=\{a_j\in V:(i,j)\in E\}$ be the adjacency set of node a_i and $\theta_{ij}(\rho)$ be the density dependent weight on the edge $(i,j)\in E$. Consider the probability weight θ which is defined by $\theta_{ij}(\rho)=\Theta(\rho_i,\rho_j)$ with a continuous differentiable, symmetric, and concave function $\Theta:[0,\infty)\times[0,\infty)\to[0,\infty)$ satisfying $\min(s,t)\leq\Theta(s,t)\leq\max(s,t), s,t\geq 0$. Two typical examples considered in this paper are the average mean $\theta_{ij}^A(\rho_i,\rho_j)=\frac{\rho_i+\rho_j}{2}$ and the logarithmic mean $\theta_{ij}^L(\rho_i,\rho_j)=\frac{\rho_i-\rho_j}{\log(\rho_i)-\log(\rho_j)}$. For more choices of the probability weight functions, we refer to [8, 9, 16, 31].

Define the discrete Lagrange functional on a graph by

$$\mathcal{L}(\rho,v)=\int_0^1\left[\frac{1}{2}\langle v,v\rangle_{\theta(\rho)}-\mathcal{V}(\rho)-\mathcal{W}(\rho)+\alpha L(\rho)-\beta I(\rho)\right]dt, \tag{3.1}$$

where $\rho(\cdot)\in\mathcal{P}_o(G)$ and the vector field v is a skew-symmetric matrix on E. The inner product of two vector fields u,v is defined by

$$\langle u,v\rangle_{\theta(\rho)}:=\frac{1}{2}\sum_{(j,l)\in E} u_{jl}v_{jl}\theta_{jl}(\rho)\omega_{ij}.$$

The parameter $\beta\geq 0$, the *discrete Fisher information* [11, 17], is defined by

$$I(\rho)=\frac{1}{2}\sum_{i=1}^N\sum_{j\in N(i)}\widetilde{\omega}_{ij}|\log(\rho_i)-\log(\rho_j)|^2\widetilde{\theta}_{ij}(\rho), \tag{3.2}$$

and the *discrete entropy* is $L(\rho)=\sum_{i=1}^N(\log(\rho_i)\rho_i-\rho_i)$, where $\alpha\in\mathbb{R}$, $(\widetilde{\omega},\widetilde{\theta})$ can be another pair of weight and density dependent weight on G.

The overall goal of the discrete variational problem is to find the minimizer of $\mathcal{L}(\rho, v)$ subject to the discrete continuity equation on a graph

$$\frac{d\rho_i}{dt} + div_G^{\theta}(\rho v) = 0,$$

where the discrete divergence of the flux function ρv is defined as

$$div_G^{\theta}(\rho v) := -\left(\sum_{l\in N(j)} \sqrt{\omega_{jl}} v_{jl}\theta_{jl}\right).$$

As shown in [16], the critical point (ρ, v) of $\mathcal{L}$ satisfies $v = \nabla_G S := \sqrt{\omega_{jl}}(S_j - S_l)_{(j,l)\in E}$ for some function S defined on G. As a consequence, the minimization problem leads to the discrete Wasserstein–Hamiltonian Hamiltonian system on G whose Hamiltonian is $\mathcal{H}(\rho, S) = \frac{1}{2}\mathcal{K}(S,\rho) + \mathcal{F}(\rho)$ with $\mathcal{K}(S,\rho) := \frac{1}{2}\langle \nabla_G S, \nabla_G S\rangle_{\theta(\rho)}$ and $\mathcal{F}(\rho) := \beta I(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)$. In particular, if $\beta = 0$, $\mathcal{V} = 0$, and $\mathcal{W} = 0$, the infimum of $2\mathcal{L}(\rho, v)$ induces the Wasserstein metric on a finite graph, which is a discrete version of the Benamou–Brenier formula [10]

$$W(\rho^0, \rho^1) := \inf_v \left\{ \sqrt{\int_0^1 \langle v, v\rangle_{\theta(\rho)} dt} \ : \ \frac{d\rho}{dt} + div_G^{\theta}(\rho v) = 0,\ \rho(0) = \rho^0,\ \rho(1) = \rho^1 \right\}.$$

4. Wasserstein Hamiltonian flow with common noise on graph. In this section, we first use the discrete version of the generalized stochastic variational principle in [17] to derive the discrete Wasserstein Hamiltonian flow with common noise. Then we study both the local and global existence of the unique solution for the stochastic Wasserstein Hamiltonian flow on a graph.

Let us briefly introduce the generalized stochastic variational principle or Hamiltonian principle as follows. Define W_δ the linear Wong–Zakai approximation [41] of a standard Wiener process W, i.e., $W_\delta(t) = W(t_k) + \frac{t-t_k}{\delta}(W(t_{k+1}) - W(t_k))$ for $t \in [t_k, t_{k+1})$ with $t_k = k\delta$, on a complete filtered probability space $(\Omega, \mathbb{P}, (\mathcal{F})_{t\geq 0}, \mathcal{F})$. Define the dominated energy and perturbed energy as

$$\begin{aligned} \mathcal{H}_0(\rho, S) &= \mathcal{K}(S,\rho) + \mathcal{F}(\rho) - \alpha L(\rho), \\ \mathcal{H}_1(\rho, S) &= \eta_1 \mathcal{K}(S,\rho) + \eta_2 I(\rho) + \eta_3 \mathcal{V}(\rho) + \eta_4 \mathcal{W}(\rho) - \eta_5 L(\rho) \end{aligned}$$

with different noise intensities $\eta_i \in \mathbb{R}, i = 1, \dots, 5$. We would like to remark that by taking different values for the noise intensities, the above general form covers many well-known problems, such as the stochastic optimal transport on graph, the stochastic Schrödinger equation, and the Schrödinger equation with white noise on a graph. Consider the following stochastic variational principle with the Wong–Zakai approximation W_δ,

$$\mathcal{I}(\rho^0, \rho^T) = \inf\{\mathcal{S}(\rho_t, \Phi_t) | \Delta_{\rho_t}\Phi_t \in \mathcal{T}_{\rho_t}\mathcal{P}_o(G), \rho(0) = \rho^0, \rho(T) = \rho^T\}, \tag{4.1}$$

whose action functional is given by the dual coordinates

$$\begin{aligned} \mathcal{S}(\rho_t, \Phi_t) = \langle \rho(0), \Phi(0)\rangle - \langle \rho(T), \Phi(T)\rangle + \int_0^T \langle \partial_t \Phi(t), \rho_t\rangle + \mathcal{H}_0(\rho_t, \Phi_t) dt \\ + \int_0^T \mathcal{H}_1(\rho_t, \Phi_t)\dot{W}_\delta dt. \end{aligned}$$

Here $\langle\cdot,\cdot\rangle$ is the standard inner product in $\mathbb{R}^N$. Denote $(\Delta_\rho)^\dagger$ as the pseudoinverse of $\Delta_\rho(\cdot) := -div_G^\theta(\rho\nabla(\cdot))$, and $\mathcal{T}_\rho\mathcal{P}_o(G)$ as the tangent space at $\rho \in \mathcal{P}_o(G)$. In particular, when $\eta_1 = 0$, the above generalized Hamiltonian principle becomes the classical variational problem with random potential in Lagrangian formalism.

By using the Lagrange multiplier method, one may verify that the critical point of (4.1) satisfies the following discrete stochastic Wasserstein Hamiltonian flow:

$$\begin{aligned}\frac{d\rho}{dt} &= \frac{\partial}{\partial S}\mathcal{H}_0(\rho,S) + \frac{\partial}{\partial S}\mathcal{H}_1(\rho,S)dW_\delta(t),\\ \frac{dS}{dt} &= -\frac{\partial}{\partial \rho}\mathcal{H}_0(\rho,S) - \frac{\partial}{\partial \rho}\mathcal{H}_1(\rho,S)dW_\delta(t).\end{aligned}$$

Moreover, if ρ^0, ρ^T, are $\mathcal{F}_0$ and $\mathcal{F}_T$ measurable functions and $\mathcal{H}_0, \mathcal{H}_1$ satisfies some growth conditions such as those given in [17], the limit of the above Wasserstein Hamiltonian flow with a Wong–Zakai approximation converges to the stochastic Hamiltonian flow in the Stratonovich sense:

$$\begin{aligned}\frac{d\rho}{dt} &= \frac{\partial}{\partial S}\mathcal{H}_0(\rho,S) + \frac{\partial}{\partial S}\mathcal{H}_1(\rho,S)\circ dW(t),\\ \frac{dS}{dt} &= -\frac{\partial}{\partial \rho}\mathcal{H}_0(\rho,S) - \frac{\partial}{\partial \rho}\mathcal{H}_1(\rho,S)\circ dW(t).\end{aligned} \tag{4.2}$$

However, we would like to remark that it is difficult to rigorously show that (4.2) is the critical point of (4.1) when $\delta \to 0$.

4.1. Properties of Wasserstein Hamiltonian flow with common noise. In this part, we consider the initial value problem of (4.2) with $\rho(0) \in \mathcal{P}_o(G)$ which is $\mathcal{F}_0$–measurable. Let us first consider the local well-posedness of (4.2). For simplicity, we may take $\theta_{ij}(\rho) = \frac{\rho_i+\rho_j}{2}, \widetilde{\theta}_{ij}(\rho) = \frac{\rho_i-\rho_j}{\log(\rho_i)-\log(\rho_j)}$ since the proof for the general case is analogous. In the rest of this paper, we also assume that $\mathbb{V}_i, \mathbb{W}_{ij}, i,j = 1,\dots N$, are deterministic finite numbers, i.e., the linear potential and interaction potential are finite.

Proposition 4.1. *Let $\rho(0) \in \mathcal{P}_o(G)$ and $S(0) \in \mathbb{R}^N$ be $\mathcal{F}_0$-measurable. Then there exists a stopping time $\tau^*(\rho(0), S(0)) > 0$ such that either*

$$\tau^*(\rho(0),S(0)) = +\infty \text{ or } \lim_{t\to\tau^*}\min_{i=1}^N \rho_i(t) = 0 \text{ or } \lim_{t\to\tau^*} S(t) = \infty, \text{ a.s.}$$

Proof. Let $c > 1$. Denote smooth truncation functions θ^1, θ^2 such that

$$\begin{aligned}&\theta_c^1(x) := 1,\ x\in[0,c],\ \theta_c^1(x) = 0, x\in[2c,\infty),\\ &\theta_c^2(x) := 1,\ x\in[1/c,1],\ \theta_c^2(x) = 0, x\in[0,1/2c].\end{aligned}$$

The support of θ_c^1 is chosen as $[0,2c]$ and that of θ_c^2 is $[\frac{1}{2c},1]$. Define $\phi_c^1(S,t), \phi_c^2(\rho,t)$ by

$$\phi_c^1(S,t) = \theta_c^1(\|S\|_{\mathcal{C}([0,t];\mathbb{R}^N)}),\ \ \phi_c^2(\rho,t) = \theta_c^2\left(\min_{i=1}^N \min_{s\in[0,t]}\rho_i(s)\right).$$

Due to the relationship between the Itô integral and Stratonovich integral, we consider the following truncated equation with $c > 0$ large enough:

$$
\begin{aligned}
\frac{d\rho^c}{dt} &= \phi_c^1(S^c,t)\phi_c^2(\rho^c,t)\frac{\partial}{\partial S}\mathcal{H}_0(\rho^c,S^c)dt + \phi_c^1(S^c,t)\phi_c^2(\rho^c,t)\frac{\partial}{\partial S}\mathcal{H}_1(\rho^c,S^c)dW_t \\
&\quad - \frac{1}{2}\phi_c^1(S^c,t)\phi_c^2(\rho^c,t)\frac{\partial^2}{\partial S^2}\mathcal{H}_1(\rho^c,S^c)\frac{\partial}{\partial \rho}\mathcal{H}_1(\rho^c,S^c)dt \\
&\quad + \frac{1}{2}\phi_c^1(S^c,t)\phi_c^2(\rho^c,t)\frac{\partial^2}{\partial \rho\partial S}\mathcal{H}_1(\rho^c,S^c)\frac{\partial}{\partial S}\mathcal{H}_1(\rho^c,S^c)dt, \\
\frac{dS^c}{dt} &= -\phi_c^1(S^c,t)\phi_c^2(\rho^c,t)\frac{\partial}{\partial \rho}\mathcal{H}_0(\rho^c,S^c) - \phi_c^1(S^c,t)\phi_c^2(\rho,t)\frac{\partial}{\partial \rho}\mathcal{H}_1(\rho^c,S^c)dW_t \\
&\quad + \frac{1}{2}\phi_c^1(S^c,t)\phi_c^2(\rho^c,t)\frac{\partial^2}{\partial S\partial \rho}\mathcal{H}_1(\rho^c,S^c)\frac{\partial}{\partial \rho}\mathcal{H}_1(\rho^c,S^c)dt \\
&\quad - \frac{1}{2}\phi_c^1(S^c,t)\phi_c^2(\rho^c,t)\frac{\partial^2}{\partial \rho^2}\mathcal{H}_1(\rho^c,S^c)\frac{\partial}{\partial S}\mathcal{H}_1(\rho^c,S^c)dt.
\end{aligned}
\tag{4.3}
$$

The local Lipschitz continuity of $\mathcal{H}_0(\rho^c,S^c)$ and $\mathcal{H}_1(\rho^c,S^c)$ implies the existence and uniqueness of the global mild solution for the truncated equation by the standard arguments in [38] since $\mathbb{V}_i$ and $\mathbb{W}_{ij}$ are finite for $i,j=1,\dots N$. Thus, for any $T>0$, there always exists a global mild solution $(\rho^c,S^c)\in\mathcal{C}([0,T];\mathbb{R}^N)\times\mathcal{C}([0,T];\mathbb{R}^N)$. Now we define the local solution of (4.3) as follows. For $n\in\mathbb{N}^+$, define the stopping time τ_n by

$$
\tau_n := \inf\{t\in[0,T]: \|S^n\|_{\mathcal{C}([0,t];\mathbb{R}^N)}\ge n\}\wedge \inf\left\{t\in[0,T]: \min_{i=1}^{N}\min_{s\in[0,t]}\rho_i^n(s)\le\frac{1}{n}\right\},
$$

and $\tau_\infty := \sup_{n\in\mathbb{N}}\tau_n$. This is guaranteed by the fact that

$$
Z^n(t) := \|S^n\|_{\mathcal{C}([0,t];\mathbb{R}^N)} + \frac{1}{\min_{i=1}^{N}\min_{s\in[0,t]}\rho_i^n(s)} < \infty
$$

defines an increasing, continuous, and $\mathcal{F}_t$-adapted process with $Z^n(0) = \|S(0)\| + \frac{1}{\min_{i=1}^N \rho_i(0)}$.

For $n\le k$, set $\tau_{k,n} := \inf\{t\in[0,T]: Z^k(t)\ge n\}$. Then we have $\tau_{k,n}\le\tau_k$ and thus $\phi_n^1(S^k,t)=\phi_k^1(S^k,t)=1$, $\phi_n^2(\rho^k,t)=\phi_k^2(\rho^k,t)$ on $\{t\le\tau_{k,n}\}$. This leads to $(\rho^k,S^k)=(\rho^n,S^n)$ and $Z^k=Z^n$ a.s. on $\{t\le\tau_{k,n}\}$. We conclude that $\tau_{k,n}=\tau_n$, a.s., and define the local solution (ρ,S) up to the stopping time τ_∞ by $(\rho,S)=(\rho^n,S^n)$, on $\{t\le\tau_n\}$. ∎

We would like to mention that in [11, 16], the global solution in deterministic case ($\eta_1=\dots=\eta_5=0$) is obtained by using the energy conservation law if $\mathcal{F}(\rho)$ contains the Fisher information $\beta I(\rho)$ with $\beta>0$. In the stochastic case, the existence of a global solution becomes more complicated and depends on the relationship between the deterministic energy $\mathcal{H}_0$ and he perturbed energy $\mathcal{H}_1$.

To see this fact, applying Itô's formula to $\mathcal{H}_0$, we obtain that before τ^n, it holds that

$$
\begin{aligned}
\mathcal{H}_0(\rho^n(t),S^n(t)) &= \mathcal{H}_0(\rho^n(0),S^n(0)) + \int_0^t \frac{\partial\mathcal{H}_0}{\partial\rho}^\top\frac{\partial\mathcal{H}_1}{\partial S}dW_s - \int_0^t\frac{\partial\mathcal{H}_0}{\partial S}^\top\frac{\partial\mathcal{H}_1}{\partial\rho}dW_s \\
&\quad - \frac{1}{2}\int_0^t\left(\frac{\partial\mathcal{H}_0}{\partial\rho}^\top\frac{\partial^2\mathcal{H}_1}{\partial S^2}\frac{\partial\mathcal{H}_1}{\partial\rho} - \frac{\partial\mathcal{H}_0}{\partial\rho}^\top\frac{\partial^2\mathcal{H}_1}{\partial S\partial\rho}\frac{\partial\mathcal{H}_1}{\partial S}\right)ds \\
&\quad + \frac{1}{2}\int_0^t\left(\frac{\partial\mathcal{H}_0}{\partial S}^\top\frac{\partial^2\mathcal{H}_1}{\partial S\partial\rho}\frac{\partial\mathcal{H}_1}{\partial\rho} - \frac{\partial\mathcal{H}_0}{\partial S}^\top\frac{\partial^2\mathcal{H}_1}{\partial\rho^2}\frac{\partial\mathcal{H}_1}{\partial S}\right)ds
\end{aligned}
$$

$$
\begin{aligned}
&+\int_0^t \frac{1}{2}\left(\frac{\partial^2\mathcal{H}_0}{\partial\rho^2}\right)\cdot\left(\frac{\partial\mathcal{H}_1}{\partial S},\frac{\partial\mathcal{H}_1}{\partial S}\right)ds\\
&-\int_0^t \left(\frac{\partial^2\mathcal{H}_0}{\partial\rho\partial S}\right)\cdot\left(\frac{\partial\mathcal{H}_1}{\partial S},\frac{\partial\mathcal{H}_1}{\partial \rho}\right)ds\\
&+\int_0^t \frac{1}{2}\left(\frac{\partial^2\mathcal{H}_0}{\partial S^2}\right)\cdot\left(\frac{\partial\mathcal{H}_1}{\partial \rho},\frac{\partial\mathcal{H}_1}{\partial \rho}\right)ds.
\end{aligned}\tag{4.4}
$$

To get the global existence of the solution, it suffices to show

$$\sup_n \mathbb{E}\left[\sup_{s\in[0,\tau^n]}\mathcal{H}_0(\rho(s),S(s))\right]<\infty.$$

Therefore, a sufficient condition to ensure the global existence of the solution is that

$$|\{\mathcal{H}_0,\mathcal{H}_1\}|+|\{\mathcal{H}_1,\{\mathcal{H}_0,\mathcal{H}_1\}\}|\le c_1\mathcal{H}_0+C_1 \text{ for some } c_1,C_1>0,$$

where $\{\cdot,\cdot\}$ is the Poisson bracket. In particular, when $\{\mathcal{H}_0,\mathcal{H}_1\}=0$, $\mathcal{H}_0$ is an invariant of the stochastic Wasserstein Hamiltonian flow. A typical example is $\mathcal{H}_0$ being a multiple of $\mathcal{H}_1$.

THEOREM 4.1. *Let $\beta>0$, $\alpha\in\mathbb{R}$, $T>0$, $\rho(0)\in\mathcal{P}_o(G)$, and $S(0)\in\mathbb{R}^d$ be $\mathcal{F}_0$-measurable and have the finite second moment. Assume that there exists $c_1,C_1>0$ such that*

$$|\{\mathcal{H}_0,\mathcal{H}_1\}|+|\{\mathcal{H}_1,\{\mathcal{H}_0,\mathcal{H}_1\}\}|\le c_1\mathcal{H}_0+C_1.$$

Then there exists a unique global solution of (4.2) *satisfying $\rho(t)\in\mathcal{P}_o(G), t\in[0,T]$.*

Proof. It suffices to prove that $\rho(t)\in\mathcal{P}_o(G), t\in[0,T]$. Let $\alpha\ge 0$. By applying (4.4), taking expectation, and employing the Burkerholder inequality, we achieve that

$$
\begin{aligned}
\mathbb{E}\left[\sup_{s\in[0,t]}\mathcal{H}_0(\rho(t),S(t))\right]\le \mathbb{E}[\mathcal{H}_0(\rho(0),S(0))]+C\int_0^t\mathbb{E}[\mathcal{H}_0(\rho(s),S(s))+C_1']ds\\
+C\mathbb{E}\left[\left(\int_0^t(\mathcal{H}_0(\rho(s),S(s))+C_1')^2ds\right)^{\frac{1}{2}}\right].
\end{aligned}
$$

Gronwall's inequality leads to

$$\mathbb{E}\left[\sup_{t\in[0,T]}\mathcal{H}_0(\rho(t),S(t))\right]\le C(T,\rho(0),S(0)).$$

It follows that $\sup_{t\in[0,T]}\mathcal{H}_0(\rho(t),S(t))<\infty, a.s.$ Due to the fact that

$$\mathcal{H}_0(\rho,S)=\frac{1}{2}\mathcal{K}(\rho,S)+\beta I(\rho)+\mathcal{V}(\rho)+\mathcal{W}(\rho)-\alpha L(\rho),$$

there always exists a constant $C>0$ such that $\mathcal{H}_0(\rho,S)+C>0$. Therefore, if $\beta>0$, $\alpha\ge 0$, we obtain that

$$\sup_{t\in[0,T]}\mathcal{K}(\rho(t),S(t))<\infty,\text{a.s., and }\sup_{t\in[0,T]}I(\rho(t))<\infty,\text{a.s.}$$

Let us define $C_{Kin} := \sup_{t\in[0,T]} \mathcal{K}(\rho(t), S(t))$. The fact that I is positive infinity on the boundary and the definition of $\mathcal{K}$ yield that

$$\min_{t\in[0,T]} \min_{i=1}^{N} \rho_i(t) > 0, \ \max_{ij\in E} |S_i(t) - S_j(t)| \le \sqrt{\frac{C_{Kin}}{\min_{t\in[0,T]} \min_{i=1}^{N} \rho_i(t)}} < \infty, a.s. \qquad \square$$

We conclude that $\rho_i(t) \in \mathcal{P}_o(G), t \in [0,T]$, a.s. The case that $\alpha < 0$ can be proven by similar steps and the fact that $x\log(x)$ is uniformly bounded in $[0,1]$.

Remark 4.1. Theorem 4.1 still holds for a more general Hamiltonian system which does not contain the Fisher information but any other potential $\mathbb{Z}(\rho)$ provided that $\mathbb{Z}(\rho)$ is smooth, bounded from below in $\mathcal{P}_o(G)$, and it approaches infinity at the boundary of $\mathcal{P}(G)$. We also would like to remark that other choices of $\theta, \widetilde{\theta}$ are available (see, e.g., [16]).

As shown in Theorem 4.1, the lower bound of the density ρ is a positive random variable, a.s. We end this section with three examples of stochastic nonlinear Schrödinger equations on G. When G is a lattice graph, the following examples can be viewed as the spatial approximations of the stochastic nonlinear Schrödinger equation on a continuous space [15].

Example 4.1 (nonlinear Schrödinger equation on graph with common noise [4, 22, 26, 15]). When nonlinear Schrödinger equation on a graph is perturbed by the common noise, it reads

$$\mathbf{i}\frac{du_j}{dt} = -\frac{1}{2}(\Delta_G u)_j + u_j\mathbb{V}_j + u_j\sum_{l=1}^{N}\mathbb{W}_{jl}|u_l|^2 + \sigma_j u_j \circ dW_t. \tag{4.5}$$

Here σ is a potential on G; Δ_G, the nonlinear Laplacian operator on G [11, 16], is defined by

$$\begin{aligned}(\Delta_G u)_j = -u_j\Bigg(\frac{1}{|u_j|^2}\Bigg[&\sum_{l\in N(j)} \omega_{jl}\theta_{jl}(\Im(\log(u_j)) - \Im(\log(u_l)))\\
&+ \sum_{l\in N(j)} \widetilde{\omega}_{jl}\widetilde{\theta}_{jl}(\Re(\log(u_j)) - \Re(\log(u_l))\Bigg]\\
&+ \sum_{l\in N(j)} \omega_{jl}\frac{\partial\theta_{jl}}{\partial\rho_j}|\Im(\log(u_j) - \log(u_l))|^2\\
&+ \sum_{l\in N(j)} \widetilde{\omega}_{jl}\frac{\partial\widetilde{\theta}_{jl}}{\partial\rho_j}|\Re(\log(u_j) - \log(u_l))|^2\Bigg),\end{aligned}$$

where $\Im$ and $\Re$ are the imaginary and real parts of a complex number, respectively.

Denoting the complex form $u_j = \sqrt{\rho_j}e^{\mathbf{i}S_j(t)}, j = 1,\dots,N$, the Madelung system on the graph becomes

$$\begin{aligned}&d\rho_i = \sum_{j\in N(i)} \omega_{ij}(S_i - S_j)\theta_{ij}(\rho)dt,\\
&dS_i + \Bigg(\sum_{j\in N(i)} \frac{1}{2}\omega_{ij}(S_i - S_j)^2\frac{\partial\theta_{ij}}{\partial\rho_i} + \frac{1}{8}\frac{\partial}{\partial\rho_i}I(\rho) + \mathbb{V}_i + \sum_{j\in N(i)} \mathbb{W}_{ij}\rho_j\Bigg)dt + \sigma_i dW_t = 0.\end{aligned} \tag{4.6}$$

We verify that there exists a global unique solution to (4.5) in Example 4.1 as follows. Notice that $\mathcal{H}_0(\rho,S)=\frac{1}{2}\mathcal{K}(\rho,S)+\frac{1}{8}I(\rho)+\mathcal{V}(\rho)+\mathcal{W}(\rho)$, $\mathcal{H}_1(\rho,S)=\sum_{i=1}^N\sigma_i\rho_i$. By calculating the derivatives of $\mathcal{H}_0$ and $\mathcal{H}_1$, we find that

$$\begin{aligned}\frac{\partial\mathcal{H}_0}{\partial\rho_i}&=\frac{1}{2}\sum_{j\in N(i)}\frac{\partial\theta_{ij}(\rho)}{\partial\rho_i}(S_i-S_j)^2\omega_{ij}+\frac{1}{8}\sum_{j\in N(i)}\frac{\partial\widetilde{\theta}_{ij}(\rho)}{\partial\rho_i}|\log(\rho_i)-\log(\rho_j)|^2\widetilde{\omega}_{ij}\\&\quad+\frac{1}{4}\sum_{j\in N(i)}\widetilde{\theta}_{ij}(\rho)(\log(\rho_i)-\log(\rho_j))\frac{1}{\rho_i}\widetilde{\omega}_{ij}+\mathbb{V}_i+\sum_{j=1}^N\mathbb{W}_{ij}\rho_j,\\\frac{\partial\mathcal{H}_0}{\partial S_i}&=\sum_{j\in N(i)}\omega_{ij}\theta_{ij}(\rho)(S_i-S_j),\ \frac{\partial\mathcal{H}_1}{\partial\rho_i}=\sigma_i,\ \frac{\partial\mathcal{H}_1}{\partial S_i}=\frac{\partial^2\mathcal{H}_1}{\partial\rho_i\partial S_j}=\frac{\partial^2\mathcal{H}_1}{\partial S_i\partial S_j}=0,\\\frac{\partial^2\mathcal{H}_0}{\partial S_i\partial S_j}&=-\omega_{ij}\theta_{ij}(\rho),\ \frac{\partial^2\mathcal{H}_0}{\partial^2 S_i}=\sum_{j\in N(i)}\omega_{ij}\theta_{ij}(\rho).\end{aligned}$$

Then it follows that

$$\begin{aligned}\{\mathcal{H}_0,\mathcal{H}_1\}&=-\sum_{i=1}^N\frac{\partial\mathcal{H}_0}{\partial S_i}\frac{\partial\mathcal{H}_1}{\partial\rho_i}=\sum_{i=1}^N\sum_{j\in N(i)}\omega_{ij}\sigma_i(S_j-S_i)\theta_{ij}(\rho),\\\{\mathcal{H}_1,\{\mathcal{H}_0,\mathcal{H}_1\}\}&=-\sum_{i,j=1}^N\frac{\partial\mathcal{H}_1}{\partial\rho_i}\frac{\partial^2\mathcal{H}_0}{\partial S_i\partial S_j}\frac{\partial\mathcal{H}_1}{\partial\rho_j}\\&=-\sum_{i=1}^N\sum_{j\in N(i)}\omega_{ij}\theta_{ij}(\rho)\sigma_i\sigma_j+\sum_{i=1}^N\sum_{j\in N(i)}\omega_{ij}\theta_{ij}(\rho)\sigma_i^2.\end{aligned}$$

By Hölder's and Young's inequalities, using the fact that I is nonnegative, $\rho_i\in[0,1]$, and that $\sigma_i,\mathbb{V}_i,\omega_{ij},\mathbb{W}_{ij}$ are finite numbers, we have

$$\begin{aligned}|\{\mathcal{H}_0,\mathcal{H}_1\}|+|\{\mathcal{H}_1,\{\mathcal{H}_0,\mathcal{H}_1\}\}|&\le\frac{1}{2}\sum_{i=1}^N\sum_{j\in N(i)}\omega_{ij}|\sigma_i|(S_j-S_i)^2\theta_{ij}(\rho)\\&\quad+\frac{1}{2}\sum_{i=1}^N\sum_{j\in N(i)}\omega_{ij}|\sigma_i|\theta_{ij}(\rho)+C(N,\sigma,\omega)\\&\le c_1\mathcal{H}_0+C(N,\sigma,\omega,\mathbb{V},\mathbb{W}),\end{aligned}$$

which implies that the condition of Theorem 4.1 holds, which completes the proof for the unique global solution.

Example 4.2 (logarithmic Schrödinger equation with common noise on a graph [20]). The logarithmic Schrödinger equation on a graph perturbed by the common noise is

$$\mathbf{i}\frac{du_j}{dt}=-\frac{1}{2}(\Delta_G u)_j+u_j\mathbb{V}_j+u_j\sum_{l=1}^N\mathbb{W}_{jl}|u_l|^2-u_j\log(|u_j|^2)+\sigma_ju_j\circ dW_t.\tag{4.7}$$

Here σ is a potential on G. Denoting the complex form $u_j=\sqrt{\rho_j}e^{\mathbf{i}S_j(t)}, j=1,\dots,N$, the Madelung system on a graph follows:

$$
\begin{aligned}
d\rho_i &= \sum_{j\in N(i)} \omega_{ij}(S_i - S_j)\theta_{ij}(\rho)dt,\\
dS_i &+ \left(\sum_{j\in N(i)} \frac{1}{2}\omega_{ij}(S_i - S_j)^2 \frac{\partial \theta_{ij}}{\partial \rho_i} + \frac{1}{8}\frac{\partial}{\partial \rho_i} I(\rho) + \mathbb{V}_i + \sum_{j\in N(i)} \mathbb{W}_{ij}\rho_j - \log(\rho_i) \right) dt\\
&+ \sigma_i dW_t = 0.
\end{aligned} \tag{4.8}
$$

Next, we show that (4.7) in Example 4.2 also satisfies the condition of Theorem 4.1. In this case, $\mathcal{H}_0(\rho, S) = \frac{1}{2}\mathcal{K}(\rho, S) - L(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho) + \frac{1}{8}I(\rho)$ and $\mathcal{H}_1(\rho, S) = \sum_{i=1}^N \rho_i\sigma_i$. The verification is similar to that of (4.5). The main difference is that

$$
\begin{aligned}
\frac{\partial \mathcal{H}_0}{\partial \rho_i} &= \frac{1}{2}\sum_{j\in N(i)} \frac{\partial \theta_{ij}(\rho)}{\partial \rho_i}(S_i - S_j)^2\omega_{ij} + \frac{1}{8}\sum_{j\in N(i)} \frac{\partial \widetilde{\theta}_{ij}(\rho)}{\partial \rho_i} |\log(\rho_i) - \log(\rho_j)|^2 \widetilde{\omega}_{ij}\\
&\quad + \frac{1}{4}\sum_{j\in N(i)} \widetilde{\theta}_{ij}(\rho)(\log(\rho_i) - \log(\rho_j))\frac{1}{\rho_i}\widetilde{\omega}_{ij} - \log(\rho_i) + \mathbb{V}_i + \sum_{j=1}^N \mathbb{W}_{ij}\rho_j.
\end{aligned}
$$

By repeating the same steps in the case of (4.5), one can achieve that

$$
|\{\mathcal{H}_0, \mathcal{H}_1\}| + |\{\mathcal{H}_1, \{\mathcal{H}_0, \mathcal{H}_1\}\}| \le c_1\mathcal{H}_0 + C_1
$$

for some $c_1 > 0$ and $C_1 > 0$.

Example 4.3 (white noise dispersion nonlinear Schrödinger equation on a graph [1, 2]). The stochastic dispersive Schrödinger equation reads

$$
\begin{aligned}
d\rho_i &= \sum_{j\in N(i)} \omega_{ij}(S_i - S_j)\theta_{ij}(\rho)\circ dW_t,\\
dS_i &+ \left(\frac{1}{2}\sum_{j\in N(i)} \omega_{ij}(S_i - S_j)^2 \frac{\partial \theta_{ij}}{\partial \rho_i} + \frac{1}{8}\frac{\partial}{\partial \rho_i} I(\rho) \right) \circ dW_t + \left(\mathbb{V}_i + \sum_{j\in N(i)} \mathbb{W}_{ij}\rho_j \right) dt = 0,
\end{aligned} \tag{4.9}
$$

which is equivalent to

$$
\mathbf{i}\frac{du_j}{dt} = -\frac{1}{2}(\Delta_G u)_j \circ dW_t + \left(u_j\mathbb{V}_j + u_j\sum_{l=1}^N \mathbb{W}_{jl}|u_l|^2 \right) dt.
$$

The existence of a unique local solution of (4.9) is done by Proposition 4.1. It suffices to present an a priori bound of the solution to achieve the global existence. Since $I(\rho)$ appears in the stochastic integral term, we denote $\widetilde{\mathcal{H}_0}(\rho, S) = \frac{1}{2}\mathcal{K}(\rho, S) + \frac{1}{8}I(\rho)$ and $\mathcal{H}_1(\rho, S) = \mathcal{V}(\rho) + \mathcal{W}(\rho)$. Similarly to Theorem 4.1, applying Itô's formula to $\widetilde{\mathcal{H}_0}(\rho(t), S(t))$, one can obtain a sufficient condition for the uniform boundedness of $\widetilde{\mathcal{H}_0}(\rho, S)$, that is,

$$
|\{\widetilde{\mathcal{H}_0}, \mathcal{H}_1\}| + |\{\widetilde{\mathcal{H}_0}, \{\widetilde{\mathcal{H}_0}, \mathcal{H}_1\}\}| \le c_1\widetilde{\mathcal{H}_0} + C_1. \tag{4.10}
$$

Below we will verify the above condition. Calculating the derivatives yields that

$$\frac{\partial \widetilde{\mathcal{H}_0}}{\partial \rho_i} = \frac{1}{2}\sum_{j\in N(i)} \frac{\partial \theta_{ij}(\rho)}{\partial \rho_i}(S_i - S_j)^2\omega_{ij} + \frac{1}{8}\sum_{j\in N(i)} \frac{\partial \widetilde{\theta}_{ij}(\rho)}{\partial \rho_i}|\log(\rho_i) - \log(\rho_j)|^2\widetilde{\omega}_{ij}$$
$$+ \frac{1}{4}\sum_{j\in N(i)} \widetilde{\theta}_{ij}(\rho)(\log(\rho_i) - \log(\rho_j))\frac{1}{\rho_i}\widetilde{\omega}_{ij},$$
$$\frac{\partial \widetilde{\mathcal{H}_0}}{\partial S_i} = \sum_{j\in N(i)} \omega_{ij}\theta_{ij}(\rho)(S_i - S_j),\ \frac{\partial \mathcal{H}_1}{\partial \rho_i} = \mathbb{V}_i + \sum_{j=1}^{N} \mathbb{W}_{ij}\rho_j,$$
$$\frac{\partial \mathcal{H}_1}{\partial S_i} = \frac{\partial^2 \mathcal{H}_1}{\partial \rho_i \partial S_j} = \frac{\partial^2 \mathcal{H}_1}{\partial S_i \partial S_j} = 0,\ \frac{\partial^2 \mathcal{H}_1}{\partial \rho_i \partial \rho_j} = \mathbb{W}_{ij},\ \frac{\partial^2 \widetilde{\mathcal{H}_0}}{\partial S_i \partial \rho_j} = \omega_{ij}\frac{\partial \theta_{ij}(\rho)}{\partial \rho_j}(S_i - S_j),$$
$$\frac{\partial^2 \widetilde{\mathcal{H}_0}}{\partial S_i \partial \rho_i} = \sum_{j\in N(i)} \omega_{ij}\frac{\partial \theta_{ij}(\rho)}{\partial \rho_i}(S_i - S_j),\ \frac{\partial^2 \widetilde{\mathcal{H}_0}}{\partial S_i \partial S_j} = -\omega_{ij}\theta_{ij}(\rho),\ \frac{\partial^2 \widetilde{\mathcal{H}_0}}{\partial^2 S_i} = \sum_{j\in N(i)} \omega_{ij}\theta_{ij}(\rho).$$

Then it holds that

$$\{\widetilde{\mathcal{H}_0}, \mathcal{H}_1\} = -\sum_{i=1}^{N} \frac{\partial \widetilde{\mathcal{H}_0}}{\partial S_i}\frac{\partial \mathcal{H}_1}{\partial \rho_i} = \sum_{i=1}^{N}\sum_{j\in N(i)} \omega_{ij}\theta_{ij}(\rho)(S_j - S_i)\left(\mathbb{V}_i + \sum_{k=1}^{N} \mathbb{W}_{ij}\rho_j\right),$$
$$\{\widetilde{\mathcal{H}_0}, \{\widetilde{\mathcal{H}_0}, \mathcal{H}_1\}\} = -\sum_{i,j=1}^{N} \frac{\partial \widetilde{\mathcal{H}_0}}{\partial \rho_j}\frac{\partial^2 \widetilde{\mathcal{H}_0}}{\partial S_i \partial S_j}\frac{\partial \mathcal{H}_1}{\partial \rho_i} + \frac{\partial \widetilde{\mathcal{H}_0}}{\partial S_j}\frac{\partial^2 \widetilde{\mathcal{H}_0}}{\partial S_i \partial \rho_j}\frac{\partial \mathcal{H}_1}{\partial \rho_i} + \frac{\partial \widetilde{\mathcal{H}_0}}{\partial S_j}\frac{\partial \widetilde{\mathcal{H}_0}}{\partial S_i}\frac{\partial^2 \mathcal{H}_1}{\partial \rho_i \rho_j},$$

where

$$\sum_{i,j=1}^{N} \frac{\partial \widetilde{\mathcal{H}_0}}{\partial \rho_j}\frac{\partial^2 \widetilde{\mathcal{H}_0}}{\partial S_i \partial S_j}\frac{\partial \mathcal{H}_1}{\partial \rho_i}$$
$$= \sum_{i=1}^{N}\sum_{j\in N(i)} \frac{1}{2}\frac{\partial \theta_{ij}(\rho)}{\partial \rho_i}(S_i - S_j)^2\omega_{ij}(-\omega_{ij}\theta_{ij})\left(\mathbb{V}_j - \mathbb{V}_i + \sum_{k=1}^{N} \mathbb{W}_{jk}\rho_k - \sum_{k=1}^{N} \mathbb{W}_{ik}\rho_k\right)$$
$$+ \sum_{i=1}^{N}\sum_{j\in N(i)} \frac{1}{8}\frac{\partial \widetilde{\theta}_{ij}(\rho)}{\partial \rho_i}|\log(\rho_i) - \log(\rho_j)|^2\widetilde{\omega}_{ij}(-\omega_{ij}\theta_{ij})$$
$$\times \left(\mathbb{V}_j - \mathbb{V}_i + \sum_{k=1}^{N} \mathbb{W}_{jk}\rho_k - \sum_{k=1}^{N} \mathbb{W}_{ik}\rho_k\right)$$
$$+ \sum_{i=1}^{N}\sum_{j\in N(i)} \frac{1}{4}\widetilde{\theta}_{ij}(\rho)(\log(\rho_i) - \log(\rho_j))\frac{1}{\rho_i}\widetilde{\omega}_{ij}(-\omega_{ij}\theta_{ij})$$
$$\times \left(\mathbb{V}_j - \mathbb{V}_i + \sum_{k=1}^{N} \mathbb{W}_{jk}\rho_k - \sum_{k=1}^{N} \mathbb{W}_{ik}\rho_k\right)$$

and

$$\sum_{i,j=1}^{N} \frac{\partial \widetilde{\mathcal{H}_0}}{\partial S_j}\frac{\partial^2 \widetilde{\mathcal{H}_0}}{\partial S_i \partial \rho_j}\frac{\partial \mathcal{H}_1}{\partial \rho_i} + \frac{\partial \widetilde{\mathcal{H}_0}}{\partial S_j}\frac{\partial \widetilde{\mathcal{H}_0}}{\partial S_i}\frac{\partial^2 \mathcal{H}_1}{\partial \rho_i \rho_j}$$
$$= \sum_{i=1}^{N}\sum_{j\in N(i)} \omega_{ij}\theta_{ij}(\rho)(S_i - S_j)^2\omega_{ij}\frac{\partial \theta_{ij}}{\partial \rho_i}\left(\mathbb{V}_j - \mathbb{V}_i + \sum_{k=1}^{N} \mathbb{W}_{jk}\rho_k - \sum_{k=1}^{N} \mathbb{W}_{ik}\rho_k\right)$$
$$+ \sum_{i=1}^{N} \mathbb{W}_{ij}\sum_{l\in N(i)} \theta_{il}(\rho)(S_i - S_l)\sum_{k\in N(j)} \theta_{jk}(\rho)(S_j - S_k).$$

By Hölder's and Young's inequalities, the nonnegativity of I and the fact that $\theta_{ij}\le 1$, letting $\sup_{(i,j)\in E}|\frac{\partial\theta_{ij}}{\partial\rho_i}|<\infty$, we have that

$$|\{\widetilde{\mathcal{H}_0},\mathcal{H}_1\}|+\left|\sum_{i,j=1}^N \frac{\partial\widetilde{\mathcal{H}_0}}{\partial S_j}\frac{\partial^2\widetilde{\mathcal{H}_0}}{\partial S_i\partial\rho_j}\frac{\partial\mathcal{H}_1}{\partial\rho_i}+\frac{\partial\widetilde{\mathcal{H}_0}}{\partial S_j}\frac{\partial\widetilde{\mathcal{H}_0}}{\partial S_i}\frac{\partial^2\mathcal{H}_1}{\partial\rho_i\rho_j}\right|\le c_1\widetilde{\mathcal{H}_0}+C_1.$$

Next we estimate $|\sum_{i,j=1}^N \frac{\partial\widetilde{\mathcal{H}_0}}{\partial\rho_j}\frac{\partial^2\widetilde{\mathcal{H}_0}}{\partial S_i\partial S_j}\frac{\partial\mathcal{H}_1}{\partial\rho_i}|$ as follows. Let $\sup_{(i,j)\in E}|\frac{\partial\widetilde{\theta}_{ij(\rho)}}{\partial\rho_i}|<\infty$ and $\sup_{(i,j)\in E}|\frac{\partial\theta_{ij(\rho)}}{\partial\rho_i}|<\infty$. Then it holds that

$$\left|\sum_{i,j=1}^N \frac{\partial\widetilde{\mathcal{H}_0}}{\partial\rho_j}\frac{\partial^2\widetilde{\mathcal{H}_0}}{\partial S_i\partial S_j}\frac{\partial\mathcal{H}_1}{\partial\rho_i}\right|\le c_2\widetilde{\mathcal{H}_0}+C_2+|\mathcal{B}|,$$

where

$$\begin{aligned}\mathcal{B}:=&\sum_{i=1}^N\sum_{j\in N(i)}\frac{1}{4}\widetilde{\theta}_{ij}(\rho)(\log(\rho_i)-\log(\rho_j))\frac{1}{\rho_i}\widetilde{\omega}_{ij}(-\omega_{ij}\theta_{ij})\\&\times\left(\mathbb{V}_j-\mathbb{V}_i+\sum_{k=1}^N\mathbb{W}_{jk}\rho_k-\sum_{k=1}^N\mathbb{W}_{ik}\rho_k\right)\\=&\frac{1}{4}\sum_{i=1}^N\sum_{j\in N(i)}\left(\frac{1}{\rho_i}+\frac{1}{\rho_j}\right)(\log(\rho_j)-\log(\rho_i))\theta_{ij}\widetilde{\theta}_{ij}\omega_{ij}\widetilde{\omega}_{ij}\left(\mathbb{V}_j+\sum_{k=1}^N\mathbb{W}_{jk}\rho_k\right).\end{aligned}$$

Thus, to let the condition (4.10) hold, we can impose the assumption

$$\sup_{(i,j)\in E}\left|\frac{\partial\theta_{ij}}{\partial\rho_i}\right|+\sup_{(i,j)\in E}\left|\frac{\partial\widetilde{\theta}_{ij}}{\partial\rho_i}\right|<\infty. \tag{4.11}$$

Besides, one of the following additional conditions is required, that is, either

$$\left(\frac{1}{\rho_i}+\frac{1}{\rho_j}\right)|\log(\rho_j)-\log(\rho_i)|\theta_{ij}(\rho)\widetilde{\theta}_{ij}(\rho)\le c_3|\log(\rho_j)-\log(\rho_i)|^2\widetilde{\theta}_{ij}(\rho)+C_3 \tag{4.12}$$

with some $c_3,C_3>0$, or

$$\mathbb{V}_j=\mathbb{V}_i,\ \mathbb{W}_{ik}=\mathbb{W}_{jk},\ \text{for all}\, i,j,k\le N. \tag{4.13}$$

To satisfy (4.12), one may take the harmonic average $\theta_{ij}(\rho)=\frac{2}{\frac{1}{\rho_i}+\frac{1}{\rho_j}}$.

We collect the above results for Example 4.3 in the following proposition.

PROPOSITION 4.2. *Let $\beta>0$, $\alpha\in\mathbb{R}$, $T>0$, $\rho(0)\in\mathcal{P}_o(G)$, and $S(0)\in\mathbb{R}^d$ be $\mathcal{F}_0$-measurable and have a finite second moment. Under assumption* (4.11), *in addition, suppose that either* (4.12) *or* (4.13) *holds. Then there exists a unique global solution to* (4.9).

We note that these examples are constructed by the critical points of stochastic variational principles and have not been considered before. To the best of our knowledge, the existing results on the existence of solutions for these three examples are only obtained in continuous spaces.

We end this section by summarizing that for any initial values, the Wasserstein Hamiltonian flow with common noise established here has local well-posedness up to a positive stopping time. Even though we conjecture that the system is the critical point of a stochastic variational principle, it is however unclear how to prove this rigorously. In fact, it is technically challenging to directly analyze the existence and uniqueness of the minimizer, even if a stochastic variational principle can be identified. To address the shortcomings of this approach, we propose another approach based on optimal control formulation to construct the boundary value formulation of Wasserstein Hamiltonian flow with common noise in the next section.

5. Optimal control problem with common noise. Here we consider the following variational principle in the framework of optimal control,

$$\begin{aligned} &\inf_{\rho,v}\left[\int_0^1 \frac{1}{2}\langle v_t, v_t\rangle_{\theta(\rho_t)} dt\right] \\ &\text{subject to } d\rho(t) + div_G^{\theta}(\rho(t)v(t)) + div_G^{\theta}(\rho(t)\nabla_G\Sigma)\circ dW_t = 0 \\ &\text{and } \rho(0,\omega)=\rho_a,\ \rho(1,\omega)=\rho_b, \end{aligned} \tag{5.1}$$

where Σ is a given vector field on G, ρ_a and ρ_b are given $\mathcal{F}_0$-measurable and $\mathcal{F}_1$-measurable densities in $\mathcal{P}_o(G)$. Let $\omega_{ij}=1$ if $(i,j)\in E$. Via a discrete Hopf–Cole transform (see, e.g., [18]), one can show that the critical point of (5.1) formally coincides with that of the discretization of the stochastic Schrödinger bridge problem in [17], i.e.,

$$\inf\{S(\rho_t,\Phi_t) : \Delta_{\rho_t}\Phi_t \in \mathcal{T}_{\rho_t}\mathcal{P}_o(G), \rho(0)=\rho^a, \rho(1)=\rho^b\}.$$

Recall that $\Delta_{\rho_t} = -div_G^{\theta}(\rho\nabla_G(\cdot))$, $\mathcal{T}_{\rho_t}\mathcal{P}_o(G)$ is the tangent space of $\mathcal{P}_o(G)$ at ρ_t, and

$$\begin{aligned} \mathcal{S}(\rho_t,\Phi_t) &= \langle\rho(0),\Phi(0)\rangle - \langle\rho(1),\Phi(1)\rangle + \int_0^1 \langle\partial_t\Phi(t),\rho_t\rangle + \mathcal{H}_0(\rho_t,\Phi_t)dt \\ &\quad + \int_0^1 \mathcal{H}_1(\rho_t,\Phi_t)\circ dW(t) \end{aligned}$$

with $\mathcal{H}_0(\rho,S)=\frac{1}{4}\sum_{ij\in E}(S_i-S_j)^2\theta_{ij}(\rho), \mathcal{H}_1(\rho,S)=\frac{1}{2}\sum_{ij\in E}(\Sigma_i-\Sigma_j)(S_i-S_j)\theta_{ij}(\rho)$. By the Lagrangian multiplier method, the critical point of (5.1), if it exists, is expected to satisfy

$$\begin{aligned} &d\rho_i(t) + \sum_{j\in N(i)}\theta_{ij}(\rho)(S_j-S_i)dt + \sum_{j\in N(i)}\theta_{ij}(\rho)(\Sigma_j-\Sigma_i)\circ dW_t = 0, \\ &dS_i(t) + \sum_{j\in N(i)}\frac{1}{2}(S_j-S_i)^2\frac{\partial\theta}{\partial\rho_i}(\rho_i,\rho_j)dt \\ &\qquad + \sum_{j\in N(i)}(S_i-S_j)(\Sigma_i-\Sigma_j)\frac{\partial\theta}{\partial\rho_i}(\rho_i,\rho_j)\circ dW_t = 0. \end{aligned}$$

However, due to the low regularity of W, it seems difficult to directly show the existence of the minimizer of (5.1). To overcome the challenges, we consider an optimal control problem perturbed by Wong–Zakai approximations of the Wiener process.

5.1. Optimal control perturbed by Wong–Zakai approximations. In this part, we prove the existence of the minimizer of the optimal control problem with Wong–Zakai approximations, which is formulated as

$$\inf_{\rho,v} \left[\int_0^1 \frac{1}{2}\langle v_t, v_t\rangle_{\theta(\rho_t)} dt\right]$$
$$\text{subject to } d\rho(t) + div_G^\theta(\rho(t)v(t)) + div_G^\theta(\rho(t)\nabla_G\Sigma)dW_t^\delta = 0$$
$$\text{and } \rho(0) = \rho_a,\ \rho(1) = \rho_b. \tag{5.2}$$

It should be mentioned that the critical points of (5.2) and (4.1), if they exist, share the same equation. However, it is not clear how to obtain the existence of the minimizer of (4.1), which motivates us to investigate the minimizer of (5.2). To this end, we first illustrate that the value of (5.2) is finite.

Given $\rho^a, \rho^b \in \mathcal{P}(G)$, we define the feasible set $C_F(\rho^a, \rho^b)$ of pairs (ρ, m) :

$$C_F(\rho^a,\rho^b) = \Bigg\{\rho \in H^1([0,1];\mathcal{P}(G)), m \in L^2([0,1];\mathcal{S}^{N\times N}) \Bigg| (\rho(0),\rho(1)) = (\rho^a,\rho^b),$$
$$d\rho_i(t) + \sum_{j\in N(i)} m_{ij}dt + \sum_{j\in N(i)} (\Sigma_j - \Sigma_i)\theta_{ij}(\rho)dW^\delta(t) = 0.\Bigg\}$$

Here $\mathcal{S}^{N\times N}$ denotes the skew-symmetric matrix, N is the node number. We consider an equivalent form of (5.2), i.e., $\inf_{\rho,m}\mathcal{A}(\rho,m)$ over the set $C_F(\rho^a,\rho^b)$, where

$$\mathcal{A}(\rho,m) := \int_0^1 \frac{1}{4}\sum_{(i,j)\in E} L(\theta_{ij}(\rho), m_{ij})dt. \tag{5.3}$$

$L(x,y) = \frac{y^2}{x}$ if $x > 0$, $L(x,y) = 0$ if $x = y = 0$, and $L(x,y) = \infty$ otherwise. The equivalence between (5.2) and $\inf_{\rho,m}(5.3)$ is based on the following reasons.

(5.2) $\geq \inf_{\rho,m}$ (5.3): this part is straightly forward by defining $m_{ij} = \theta_{ij}(S_i - S_j)$. When $\theta_{ij} = 0$, define $m_{ij} = 0$.

(5.2) $\leq \inf_{\rho,m}$ (5.3): For any fixed ρ, denote $v_{ij} = \frac{m_{ij}}{\theta_{ij}}$, and $\mathbb{H}_\rho = \{[v] \subset \mathcal{S}^{N\times N} | w \in [v] \text{ if and only if } v_{ij} = w_{ij} \text{ for } \theta_{ij}(\rho) \neq 0\}$. Under the graph inner product $\langle\cdot,\cdot\rangle_{\theta(\rho)}$, $\mathbb{H}_\rho$ forms a finite-dimensional subspace. Thus ∇_G defines a linear map from the potential functional space (consider $L^2(G)$ such that it is also a Hilbert space) to $\mathbb{H}_\rho$, and div_G defines a map from the matrix space to $L^2(G)$. Denote P_ρ the orthogonal projection in $\mathbb{H}_\rho$ onto the range of ∇_G. Then for any feasible path $(\rho_t, m_t), t \in [0,1]$, in (5.3), one can always find a potential functional S_t such that $P_{\rho_t}v_t = \nabla_G S_t$. Thanks to the fact that $\mathbb{H}_\rho = Ran(\nabla_G) \otimes Ker(div_G)$, we have that $(I - P_{\rho_t})v_t \in Ker(div_G)$ and thus $div(\rho_t v_t) = div_G(\rho_t\nabla_G S_t)$. As a consequence, (ρ_t, S_t) also belongs to the feasible set of (5.2).

Proposition 5.1. *For any $\rho^a, \rho^b \in \mathcal{P}(G)$, there is a path $(\rho,m) \in C_F(\rho^a,\rho^b)$ such that $\mathcal{A}(\rho,m) < \infty$.*

Proof. We use an induction argument on the number of nodes in G. First, consider the case that the cardinality of $V = \{1,2\}$ is 2, the edge $E = \{(1,2),(2,1)\}$ and $\rho^a \neq \rho^b$. Define $\rho_1(t) = \rho_1^a, t \in [0, 1-\delta], \rho_1(t) = \rho_1^a + (\rho_1^b - \rho_1^a)\frac{t-1+\delta}{\delta}, t \in [1-\delta, 1]$. Then it follows that

$$\rho_1(t) - \rho_1(0) = \int_0^t m_{21}(s)ds + \int_0^t \frac{1}{2}(\Sigma_1 - \Sigma_2)dW^\delta(s).$$

Therefore, we get

$$m_{21}(t) = \frac{1}{2}(\Sigma_2 - \Sigma_1)\dot{W}^\delta(t),\ t \in [0, 1-\delta],$$
$$m_{21}(t) = (\rho_1^b - \rho_1^a)\frac{1}{\delta} + \frac{1}{2}(\Sigma_2 - \Sigma_1)\dot{W}^\delta(t),\ t \in [1-\delta, 1],$$

where $\dot{W}^\delta(t) = \frac{W(t_{k+1})-W(t_k)}{\delta}, t_k = k\delta, k \le K-1, K\delta = 1, t \in [t_k, t_{k+1}]$. Notice that

$$\begin{aligned}\int_0^1 m_{21}^2(s)ds &= \frac{1}{4}\int_0^{1-\delta}(\Sigma_2-\Sigma_1)^2(\dot{W}^\delta(t))^2 ds \\ &\quad + \int_{1-\delta}^1 \left[(\rho_1^b-\rho_1^a)\frac{1}{\delta} + \frac{1}{2}(\Sigma_2-\Sigma_1)\dot{W}^\delta(t)\right]^2 ds \\ &\le \frac{1}{4}(\Sigma_2-\Sigma_1)^2 \sum_{k=0}^{K-1}\frac{(W_{t_{k+1}}-W_{t_k})^2}{\delta} + \frac{1}{4}(\Sigma_2-\Sigma_1)^2\frac{(W_{t_K}-W_{t_{K-1}})^2}{\delta} \\ &\quad + \frac{(\rho_1^b-\rho_1^a)^2}{\delta} + (\rho_1^b-\rho_1^a)(\Sigma_2-\Sigma_1)\frac{W_{t_K}-W_{t_{K-1}}}{\delta} \le C(\delta) < \infty, \text{ a.s.}\end{aligned}$$

This covers the case $n = 2$. When $n > 2$, we use the concatenation arguments to show the finiteness of (5.3). Namely, we need show that if there exists $\rho \in \mathcal{P}(G)$ such that $C_F(\rho^a, \rho)$ and $C_F(\rho, \rho^b)$ have feasible paths then $C_F(\rho^a, \rho^b)$ also has a feasible path. Because for any ρ_a, ρ_b, we can set an intermediate state $(0, \dots, 0, 1)$ and show that there are feasible paths connecting ρ_a and $(0, \dots, 0, 1)$, $(0, \dots, 0, 1)$, and ρ_b, respectively. By integrating these two paths continuously, we could construct a feasible path from ρ_a to ρ_b.

Without loss of generality, we may assume that $\rho^b = (0, \dots, 0, 1)$. If the support of ρ^a is the same as ρ^b, then it follows that $\rho^a = \rho^b, (\rho, m) \in C_F(\rho^a, \rho^b)$ as long as

$$\sum_{j\in N(i)} m_{ij}(t) = \sum_{j \in N(i)} (\Sigma_i - \Sigma_j)\theta(\rho_i, \rho_j)\dot{W}^\delta(t).$$

Supposing that the support of ρ^a has an intersection with the first $N-1$ nodes, we iteratively construct a sequence $\widetilde{\rho}^0, \dots, \widetilde{\rho}^{l_0}$ satisfying $\widetilde{\rho}^0 = \rho^a$, $\widetilde{\rho}^{l_0} = \rho^b$, the cardinality of the support of $\widetilde{\rho}^l$ is strictly smaller than that of $\widetilde{\rho}^{l-1}$, and that there is a feasible path connecting $\widetilde{\rho}^{l-1}$ with $\widetilde{\rho}^l$ in the interval $[t_{l-1}, t_l]$, where $t_l = \frac{l}{l_0}$. ∎

Introducing the corresponding saddle scheme formally,

$$\inf_\rho \sup_\lambda \left[\mathcal{A}(\rho, m) - \int_0^1 \langle \lambda, \dot{\rho}(t) + div_G^\theta(\rho(t)v(t)) + div_G^\theta(\rho(t)\nabla_G \Sigma)\dot{W}_t^\delta\rangle dt\right]$$

with $\rho(0) = \rho^a$ and $\rho(1) = \rho^b$, it can be seen that there exists $\lambda \in BV_{loc}([0,1]; \mathbb{R}^N)$ such that the critical point (ρ, v) of (5.2) satisfies

$$\begin{aligned}&\theta_{ij}(\rho)[v_{ij} - (\lambda_i - \lambda_j)] = 0 \ \forall (i,j) \in E, \\ &\langle \dot{\lambda}, \rho\rangle - \frac{1}{4}\sum_{ij} v_{ij}^2\theta_{ij}(\rho) + \frac{1}{2}\sum_{ij}(\Sigma_i - \Sigma_j)(\lambda_i - \lambda_j)\theta_{ij}(\rho)dW^\delta(t) = 0, \ \mathcal{L}^1, \text{ a.e.}\end{aligned}$$

Denote $S_i = -\lambda_i$. When the optimal path does not intersect the boundary of $\mathcal{P}(G)$, the above equations become the stochastic Wasserstein Hamiltonian flow (see, e.g., [17]),

$$\begin{aligned}&\dot{\rho} = \nabla_S \mathcal{H}_0(\rho, S) + \nabla_S \mathcal{H}_1(\rho, S)\dot{W}^\delta, \\ &\dot{S} = -\nabla_\rho \mathcal{H}_0(\rho, S) - \nabla_\rho \mathcal{H}_1(\rho, S)\dot{W}^\delta,\end{aligned} \tag{5.4}$$

where $\mathcal{H}_0(\rho, S) = \frac{1}{4}\sum_{ij\in E}(S_i - S_j)^2\theta_{ij}(\rho), \mathcal{H}_1(\rho, S) = \frac{1}{2}\sum_{ij\in E}(\Sigma_i - \Sigma_j)(S_i - S_j)\theta_{ij}(\rho)$. Indeed, we have the following result.

PROPOSITION 5.2. *Let $\rho^a, \rho^b \in \mathcal{P}(G)$. Assume that $(\rho, m) \in C_F(\rho^a, \rho^b)$ and that $S \in H^1([0,1]; \mathbb{R}^N)$ satisfies*

$$\langle \dot{S}, \rho \rangle + \frac{1}{4} \sum_{ij} (S_i - S_j)^2 \theta_{ij}(\rho) + \sum_{ij} (\Sigma_i - \Sigma_j)(S_i - S_j)\theta_{ij}(\rho) dW^\delta(t) \leq 0, \ \mathcal{L}^1 \text{ a.e.}$$

Then

(i) *it holds that*

$$\langle S(1), \rho^b \rangle - \langle S(0), \rho^a \rangle \leq \mathcal{A}(\rho, m).$$

(ii) *Equality holds in* (i) *if and only if*

$$m_{ij} = \theta_{ij}(\rho)(\nabla_G S)_{ij} \ \forall (i,j) \in E,$$
$$\langle \dot{S}, \rho \rangle + \frac{1}{4} \sum_{ij} (S_i - S_j)\theta_{ij}(\rho) + \sum_{ij} (\Sigma_i - \Sigma_j)(S_i - S_j)\theta_{ij}(\rho) dW^\delta(t) = 0.$$

(iii) *If $\rho \in \mathcal{P}_o(G)$, a.e., then (ρ, S) satisfies* (5.4)*, a.e.*

Proof. By using integration by parts and Hölder's inequality, we get

$$\begin{aligned}
&\langle S(1), \rho^b \rangle - \langle S(0), \rho^a \rangle \\
&\quad = \int_0^1 (\langle m, \nabla_G S \rangle + \langle \rho, \dot{S} \rangle) dt + \int_0^1 \langle \nabla_G \Sigma, \nabla_G S \rangle_{\theta(\rho)} dW^\delta(t) \\
&\quad \leq \mathcal{A}(\rho, m) + \int_0^1 \left(\langle \rho, \dot{S} \rangle + \frac{1}{2} \|\nabla_G S\|^2_{\theta(\rho)} + \langle \nabla_G \Sigma, \nabla_G S \rangle_{\theta(\rho)} \dot{W}^\delta(t) \right) dt \\
&\quad \leq \mathcal{A}(\rho, m).
\end{aligned}$$

From the above estimate, the equality holds if and only if the conditions in (ii) hold. If $\rho \in \mathcal{P}_o(G)$, a.e., we obtain

$$\begin{aligned}
0 &= \langle \rho, \dot{S} + \nabla_\rho \mathcal{H}_0(\rho, S) + \nabla_\rho \mathcal{H}_1(\rho, S) \dot{W}^\delta \rangle \\
&= \langle \rho, \dot{S} \rangle + \mathcal{H}_0(\rho, S) + \mathcal{H}_1(\rho, S) \dot{W}^\delta \leq 0,
\end{aligned}$$

which completes the proof. ∎

Now we focus on the existence of the minimizer of (5.2).

THEOREM 5.1. *Let $\rho^a, \rho^b \in \mathcal{P}(G)$. There exists (ρ^*, v^*, m^*) such that (ρ^*, v^*) minimizes* (5.2) *and (ρ^*, m^*) minimizes $\inf_{\rho,m} \mathcal{A}(\rho, m)$.*

Proof. By Proposition 5.1, there exists a path $(\rho, m) \in C_F(\rho^a, \rho^b)$ such that $\mathcal{A}(\rho, m) \leq C < \infty$ for some constant $C > 0$, which implies that $\|m\|_{L^2([0,T];\mathcal{S}^{n\times n})} \leq 2C$. Then the equation of $\dot{\rho}$, together with the Poincaré–Wirtinger inequality, implies that $\rho \in H^1([0,1]; \mathbb{R}^N)$. The intersection of $C_F(\rho^a, \rho^b)$ with any sublevel set, i.e., $\{(\rho, S) | \mathcal{A}(\rho, \theta(\rho)\nabla_G S) \leq c\}$ for some $c \geq 0$, of $\mathcal{A}$ is a precompact set in the weak topology of $H^1([0,1]; \mathbb{R}^N) \times L^2([0,1]; \mathcal{S}^{N\times N})$. Notice that $\mathcal{A}$ is nonnegative and weakly lower semicontinuous on $H^1([0,1]; \mathbb{R}^N) \times L^2([0,1]; \mathcal{S}^{N\times N})$ (see, e.g., [24]). Thus it achieves its minimum at some path $(\rho^*, m^*) \in C_F(\rho^a, \rho^b)$.

Next we define a measurable vector field v^* as $v^*_{ij}(t) = \frac{m^*_{ij}(t)}{\theta_{ij}(\rho)}$ if $\theta_{ij}(\rho) > 0$, and $v^*_{ij}(t) = 0$ otherwise. As a consequence, we have that $\frac{1}{2}\int_0^1 \|v^*\|^2_{\theta(\rho)} dt = \mathcal{A}(\rho^*, m^*) < \infty$. Then we show that (ρ^*, v^*) is also a minimizer of (5.2). Let (ρ, v) be a feasible set of

(5.2) and set $m_{ij} = \theta_{ij}(\rho) v_{ij}$. It holds that $\mathcal{A}(\rho, m) = \frac{1}{2}\int_0^1 \|v\|^2_{\theta(\rho)} dt < \infty$ and $(\rho, m) \in C_F(\rho^a, \rho^b)$. From the property of (ρ^*, m^*), we have $\int_0^1 \|v^*\|^2_{\theta(\rho^*)} dt \le \int_0^1 \|v\|^2_{\theta(\rho)} dt$. ◻

Now we are in a position to show the following duality property:

$$\min_{(\rho,m)\in C_F(\rho^a,\rho^b)} \mathcal{A}(\rho,m) = \sup_S \Bigg\{ \langle S(1), \rho^b\rangle - \langle S(0), \rho^a\rangle : \sup_\rho \Bigg\{ \langle \dot S, \rho\rangle + \frac{1}{4}\sum_{ij} v_{ij}^2 \theta_{ij}(\rho) + \frac{1}{2}\sum_{ij} (\Sigma_i - \Sigma_j)(S_i - S_j)\theta_{ij}(\rho) dW^\delta(t) \Bigg\} = 0 \Bigg\}. \tag{5.5}$$

The key is using the minimax identity of the following Lagrange multiplier:

$$\mathcal{L}(\rho, m, S) := \langle S(1), \rho^b\rangle - \langle S(0), \rho^a\rangle + \mathcal{A}(\rho, m) - \int_0^1 (\langle \dot S, \rho\rangle + \langle m, \nabla_G S\rangle + \langle \nabla_G \Sigma, \nabla_G S\rangle \dot W^\delta(t)) dt.$$

To ensure the boundedness of S, we consider a subset H^1_R of $H^1([0,1];\mathbb{R}^n)$ which is defined by $H^1_R := \{S \in H^1([0,1];\mathbb{R}^n) : \|S\|_{H^1([0,1];\mathbb{R}^n)} \le R\}, R > 0$. We claim that the following property holds,

$$\inf_{(\rho,m)} \sup_{S\in H^1_R} \mathcal{L}(\rho, m, S) = \sup_{S\in H^1_R} \inf_{(\rho,m)} \mathcal{L}(\rho, m, S), \tag{5.6}$$

by applying the standard minimax theorem in [32, Theorem I.1.1.]. It suffices to prove that H^1_R is convex and compact in the weak topology, $\mathcal{A}$ is convex in the weak topology, $\{S \in H^1_R : \mathcal{L}(\rho, m, S) \ge C\}$ is closed convex set in H^1_R, and $\{(\rho, m) \in \mathcal{C}_F(\rho^a, \rho^b) : \mathcal{L}(\rho, m, S) \le C\}$ is a convex set for any $C \in \mathbb{R}$. All these conditions can be verified since $\mathcal{L}$ is convex in (ρ, m) and linear in λ, and that $H^1([0,1];\mathbb{R}^N)$ is compact in $L^2([0,1];\mathbb{R}^N)$. Furthermore, we also have that

$$\sup_{S\in H^1_R} \mathcal{L}(\rho, m, S) = \mathcal{A}(\rho, m) + R\mathcal{E}(\rho, m), \tag{5.7}$$

where the nonnegative functional $\mathcal{E}$ is defined by

$$\mathcal{E}(\rho, m) := \sup_{S\in H^1_1} \{\langle S(1), \rho^b\rangle - \langle S(0), \rho^a\rangle - \int_0^1 (\langle \dot S, \rho\rangle + \langle m, \nabla_G S\rangle + \langle \nabla_G \Sigma, \nabla_G S\rangle_{\theta(\rho)} \dot W^\delta(t)) dt\}.$$

It can been seen that $\mathcal{E} = 0$ only if $(\rho, m) \in C_F(\rho^a, \rho^b)$ and larger than 0 otherwise. By making use of the lower continuity and convexity of $\mathcal{L}$ and $\mathcal{E}$, it can be seen that for any $R > 0$, there exists $(\rho^{*,R}, m^{*,R})$ such that it minimizes $\mathcal{A} + R\mathcal{E}$. Furthermore, the set $\{(\rho^{*,R}, m^{*,R})\}_{R>0}$ is precompact, which complete the proof by taking $R \to \infty$.

LEMMA 5.1. *The commutative property holds:*

$$\inf_{(\rho,m)} \sup_{S\in H^1} \mathcal{L}(\rho, m, S) = \sup_{S\in H^1} \inf_{(\rho,m)} \mathcal{L}(\rho, m, S). \tag{5.8}$$

Proof. Since $(\rho^*, m^*) \in C_F(\rho^0, \rho^1)$, we have that for any $R > 0$,

$$\mathcal{A}(\rho^*, m^*) = \sup_{S\in H^1_R} \mathcal{L}(\rho^*, m^*, S) \ge \inf_{(\rho,m)} \sup_{S\in H^1_R} \mathcal{L}(\rho, m, S).$$

By (5.6), we get $\mathcal{A}(\rho^*, m^*) \geq \sup_{S\in H^1_R} \inf_{(\rho,m)} \mathcal{L}(\rho, m, S)$. Recall that (ρ^*, m^*) is the minimizer of the optimal control problem with common noise. (5.7) and $(\rho^*, m^*) \in C_F(\rho^0, \rho^1)$ implies that

$$\begin{aligned}\mathcal{A}(\rho^*, m^*) &= \sup_{S\in\mathbb{H}^1_R} \mathcal{L}(\rho^*, m^*, S) \geq \inf_{\rho,m} \sup_{S\in\mathbb{H}^1_R} \mathcal{L}(\rho, m, S)\\ &= \mathcal{A}(\rho^{*,R}, m^{*,R}) + R\mathcal{E}(\rho^{*,R}, m^{*,R}) \geq \mathcal{A}(\rho^{*,R}, m^{*,R}).\end{aligned}$$

Denote the accumulation point of $\{\rho^{*,R}, m^{*,R}\}$ by $(\rho^{*,\infty}, m^{*,\infty})$. It follows that $(\rho^{*,\infty}, m^{*,\infty}) \in C_F(\rho^a, \rho^b)$ and therefore that

$$\mathcal{A}(\rho^*, m^*) \leq \mathcal{A}(\rho^{*,\infty}, m^{*,\infty}).$$

We conclude that

$$\mathcal{A}(\rho^*, m^*) = \mathcal{A}(\rho^{*,\infty}, m^{*,\infty}), \quad \limsup_{R\to+\infty} R\mathcal{E}(\rho^{*,R}, m^{*,R}) = 0.$$

It suffices to prove

$$\inf_{(\rho,m)\in C_F(\rho^0,\rho^1)} \sup_{S\in H^1} \mathcal{L}(\rho, m, S) \leq \sup_{S\in H^1} \inf_{(\rho,m)\in C_F(\rho^0,\rho^1)} \mathcal{L}(\rho, m, S).$$

By using (5.6), we obtain that

$$\mathcal{A}(\rho^{*,\infty}, S^{*,\infty}) \leq \lim_{R\to\infty} \sup_{S\in H^1_R} \inf_{(\rho,m)} \mathcal{L}(\rho, m, S) \leq \sup_{S\in H^1} \inf_{(\rho,m)} \mathcal{L}(\rho, m, S),$$

and that

$$\mathcal{A}(\rho^{*,\infty}, S^{*,\infty}) = \sup_{S\in H^1} \mathcal{L}(\rho^{*,\infty}, m^{*,\infty}, S) \geq \inf_{(\rho,m)} \sup_{S\in H^1} \mathcal{L}(\rho, m, S),$$

which completes the proof. □

THEOREM 5.2. *The dual property* (5.5) *holds.*

Proof. For any $(\rho, m) \in H^1([0,1], \mathbb{R}^N) \times L^2([0,1], \mathcal{S}^{N\times N})$, by (5.7) we have

$$\sup_{S\in H^1} \mathcal{L}(\rho, m, S) = \mathcal{A}(\rho, m) + \mathbb{I}_{C_F(\rho^a,\rho^b)}(\rho, m),$$

where $\mathbb{I}_{C_F(\rho^a,\rho^b)}(\rho, m) = 0$ if $(\rho, m) \in C_F(\rho^a, \rho^b)$, otherwise $\mathbb{I}_{C_F(\rho^a,\rho^b)}(\rho, m) = \infty$. By using (5.8), we achieve that

$$\begin{aligned}\inf_{(\rho,m)} \sup_{S\in H^1} \mathcal{L}(\rho, m, S) &= \inf_{(\rho,m)} \{\mathcal{A}(\rho, m) + \mathbb{I}_{C_F(\rho^a,\rho^b)}(\rho, m)\}\\ &= \inf_{(\rho,m)\in C_F(\rho^0,\rho^1)} \{\mathcal{A}(\rho, m)\}.\end{aligned}$$

Notice that for a fixed $S \in H^1$, using the Hölder inequality, we get

$$\inf_{(\rho,m)} \mathcal{L}(\rho, m, S) = \langle S(1), \rho^b\rangle - \langle S(0), \rho^a\rangle - \int_0^1 \max(H(\dot{S}, \nabla_G S), 0)dt,$$

where $H(\dot{S}, \nabla_G S) := \sup_\rho\{\langle \dot{S}, \rho\rangle + \frac{1}{2}\|\nabla_G S\|^2_{\theta(\rho)} + \langle \nabla_G \Sigma, \nabla_G S\rangle_{\theta(\rho)} \dot{W}^\delta(t))\}$. Thus it follows that

$$\inf_{(\rho,m)\in C_F(\rho^0,\rho^1)} \{\mathcal{A}(\rho, m)\} = \sup_{S\in H^1} \inf_{(\rho,m)} \mathcal{L}(\rho, m, S)$$

if $H(\dot S, \nabla_G S) \le 0$, $\mathcal{L}^1$ a.e. It only needs to show the existence of $\bar S$ such that $H(\dot{\bar S}, \nabla_G \bar S) = 0$ and that $\langle \bar S(1), \rho^b\rangle - \langle \bar S(0), \rho^a\rangle \ge \langle S(1), \rho^b\rangle - \langle S(0), \rho^a\rangle$. To this end, let $\mathcal{O} := \{H(\dot S, \nabla_G S) < 0\}$ and assume that $\mathcal{L}^1(\mathcal{O}) > 0$. Define $\bar S_i = S_i + \alpha$ with $\alpha(t) = -\int_0^t \chi_{\mathcal{O}} H(\dot S, \nabla_G S) ds$. Thus, we get

$$\begin{aligned}\langle \bar S(1), \rho^b\rangle - \langle \bar S(0), \rho^a\rangle - \int_{\mathcal{O}} H(\dot{\bar S}, \nabla_G \bar S) dt &= \langle S(1), \rho^b\rangle - \langle S(0), \rho^a\rangle - \int_{\mathcal{O}} H(\dot S, \nabla_G S) dt \\ &\ge \langle S(1), \rho^b\rangle - \langle S(0), \rho^a\rangle,\end{aligned}$$

which completes the proof. □

Now, we are able to describe the Hamiltonian structure of the minimizer. Following the idea of [24], define $\dot S = \dot S^{sing} + \dot S^{abs}$, where $\dot S^{abs}$ is the absolutely continuous part (w.r.t. $\mathcal{L}_1$) of $\dot S$ and $\dot S^{sing}$ is the singular part (w.r.t. $\mathcal{L}_1$) of $\dot S$, then we have that

$$\begin{aligned}&d\rho(t) + div_G^\theta(\rho(t)\nabla_G S(t)) + div_G^\theta(\rho(t)\nabla_G \Sigma) dW_t^\delta = 0,\\ &\langle \dot S^{abs}, \rho\rangle + \frac{1}{4}\sum_{ij}(S_i - S_j)^2\theta_{ij}(\rho) + \sum_{ij}(\Sigma_i - \Sigma_j)(S_i - S_j)\theta_{ij}(\rho) dW_t^\delta = 0,\ \mathcal{L}^1 \text{ a.e.},\\ &\left\langle \frac{d\dot S^{sing}}{d\mu}, \rho\right\rangle = 0\ \ \forall\, \mu \text{ a.e.},\ \mu \perp \mathcal{L}_1.\end{aligned}$$

The singular means the singular part of $\dot S$ w.r.t. $\mathcal{L}_1$. When the optimal path does not intersect the boundary of $\mathcal{P}(G)$, we recover (5.4). We would like to remark that if the minimizer (ρ, S) is also predictable (see, e.g., [21]), then (5.4) converges to a stochastic Wasserstein Hamiltonian flow driven by the standard Brownian motion when $\delta \to 0$ [17].

In the deterministic case, the θ-connected components have been introduced in [31, 24] to study whether the optimal transfer achieves the boundary of the density manifold in optimal transport on a graph (see, e.g., [31, section 1], [24, section 3]). In this part, we demonstrate that this approach may fail in the stochastic case, such as (5.2). Let $\rho \in \mathcal{P}(G)$. The nodes $i, j \in V$ are called θ-connected, if there exist integers $i_1, \dots, i_k \in V$ such that $i_1 = i, i_k = j, (i_l, i_{l+1}) \in E, l \le k-1$, and $\theta_{i_1 i_2}(\rho)\cdots\theta_{i_{k-1}i_k}(\rho) > 0$. The largest θ-connected set containing i is called the θ-connected component of i. All the θ-components of ρ form a partition of V.

We use the following example to illustrate that θ-connected component may not characterize the optimal path.

Remark 5.1. Let $V = \{1, 2, 3\}, E = \{(1, 2), (2, 3)\}$. Let $\rho^a = (0, 0, 1)$ and $\rho^b = (0, \frac{1}{2}, \frac{1}{2})$. We cannot obtain that ρ connecting ρ^a and ρ^b lies on the boundary as in the deterministic case. To see this fact, assume that $(\rho, m) \in C(\rho^a, \rho^b)$ with $\rho_1 \not\equiv 0$. We have that

$$\begin{aligned}\dot\rho_1 + m_{12} &= (\Sigma_1 - \Sigma_2)\theta_{12}(\rho)\dot W^\delta,\\ \dot\rho_2 + m_{21} + m_{23} &= (\Sigma_2 - \Sigma_1)\theta_{21}(\rho)\dot W^\delta + (\Sigma_2 - \Sigma_3)\theta_{23}(\rho)\dot W^\delta,\\ \dot\rho_3 + m_{32} &= (\Sigma_3 - \Sigma_2)\theta_{32}(\rho)\dot W^\delta.\end{aligned}$$

Then one may define $(\widetilde\rho_1, \widetilde\rho_2, \widetilde\rho_3) = (0, \rho_1 + \rho_2, \rho_3)$ and

$$\widetilde m_{12} = (\Sigma_1 - \Sigma_2)\theta_{12}(\rho)\dot W^\delta,\ \widetilde m_{23} - (\Sigma_2 - \Sigma_3)\dot W^\delta = m_{23} - (\Sigma_2 - \Sigma_3)\theta_{23}(\rho)\dot W^\delta.$$

Then it holds that $\widetilde{\rho}(0)=\rho^a, \widetilde{\rho}(1)=\rho^b$ and $\dot{\widetilde{\rho}}_1=0$. By the definition of $\widetilde{\rho}$, it could be shown that

$$\dot{\widetilde{\rho}}_2+\widetilde{m}_{23}=(\Sigma_2-\Sigma_3)\dot{W}^\delta,$$
$$\dot{\widetilde{\rho}}_3+\widetilde{m}_{32}=(\Sigma_3-\Sigma_2)\dot{W}^\delta.$$

Therefore, we have

$$\mathcal{A}(\rho,m)=\frac{1}{2}\int_0^1\left(\frac{m_{12}^2}{\theta_{12}(\rho)}+\frac{m_{23}^2}{\theta_{23}(\rho)}\right)dt$$

and

$$\mathcal{A}(\widetilde{\rho},\widetilde{m})=\frac{1}{2}\int_0^1\frac{((\Sigma_1-\Sigma_2)\theta_{12}(\widetilde{\rho})\dot{W}^\delta)^2}{\theta_{12}(\widetilde{\rho})}+(m_{23}+\frac{1}{2}\rho_1(\Sigma_2-\Sigma_3)\dot{W}^\delta)^2dt.$$

However, we may not have $\mathcal{A}(\widetilde{\rho},\widetilde{m})\leq\mathcal{A}(\rho,m)$.

In the next subsection, we consider an optimal control problem with a special stochastic perturbation, where the θ-connect method still works in the stochastic case.

5.2. Optimal control problem with a special stochastic perturbation. Now we consider a special perturbation of the optimal control problem, that is,

$$\begin{aligned}&\inf_{\rho,v}\left[\int_0^1\frac{1}{2}\langle v_t,v_t\rangle_{\theta(\rho_t)}dt\right]\\&\text{subject to } d\rho(t)+div_G^\theta(\rho(t)v(t))dt+div_G^\theta(\rho(t)v(t))dW_t^\delta=0\\&\text{and } \rho(0)=\rho_a,\ \rho(1)=\rho_b.\end{aligned}\tag{5.9}$$

Note that (5.9) is different from (5.2) since the diffusion term in the constraint involves $v(t)$.

Given $\rho^a,\rho^b\in\mathcal{P}(G)$, we define the feasible set $C_F(\rho^a,\rho^b)$ of pairs (ρ,m) such that

$$\rho\in H^1([0,1];\mathcal{P}(G)), m\in L^2([0,1];\mathcal{S}^{n\times n}),\ (\rho(0),\rho(1))=(\rho^a,\rho^b)$$

and

$$d\rho_i(t)+\sum_{j\in N(i)}m_{ij}dt+\sum_{j\in N(i)}m_{ij}dW^\delta(t)=0.$$

We consider the equivalent form of (5.9), $\inf_{\rho,m}\mathcal{A}(\rho,m)$ over the set $C_F(\rho^a,\rho^b)$, where $\mathcal{A}$ is defined in (5.3).

PROPOSITION 5.3. *For any $\rho^a,\rho^b\in\mathcal{P}(G)$, there is a path $(\rho,m)\in C_F(\rho^a,\rho^b)$ such that $\mathcal{A}(\rho,m)<\infty$.*

Proof. The proof is similar to that of Proposition 5.1. We use an introduction argument on the nodes number of G. First, consider the case that the cardinality of $V=\{1,2\}$ is 2, the edge $E=\{(1,2),(2,1)\}$, and $\rho^a\neq\rho^b$. Define $\rho_1(t)=\rho_1^a, t\in[0,1-\delta], \rho_1(t)=\rho_1^a+(\rho_1^b-\rho_1^a)\frac{t-1+\delta}{\delta}, t\in[1-\delta,1]$. Then it follows that

$$\rho_1(t)-\rho_1(0)=\int_0^t m_{21}(s)(1+\dot{W}^\delta(s))ds.$$

Therefore, we get

$$m_{21}(t) = 0,\ t \in [0, 1-\delta],$$
$$m_{21}(t)(1 + \dot{W}^{\delta}(t)) = (\rho_1^b - \rho_1^a)\frac{1}{\delta},\ t \in [1-\delta, 1],\ \mathcal{L}^1 \text{ a.e.},$$

where $\dot{W}^{\delta}(t) = \frac{W(t_{k+1}) - W(t_k)}{\delta}, t_k = k\delta, k \le K-1, K\delta = 1, t \in [t_k, t_{k+1}]$. Notice that

$$\int_0^1 m_{21}^2(s)ds = \int_{1-\delta}^1 \frac{1}{\delta^2}\frac{(\rho_1^b - \rho_1^a)^2}{(1+\dot{W}^{\delta}(s))^2}ds \le \int_{1-\delta}^1 \frac{(\rho_1^b - \rho_1^a)^2}{(\delta + W_{t_K} - W_{t_{K-1}})^2}ds < \infty, \text{ a.s.}$$

This covers the case $n = 2$. When $n > 2$, we use the concatenation arguments to show the finiteness of $\mathcal{A}$ as in the proof of Proposition 5.9. ∎

Applying the Lagrange multiplier method, when the optimal path does not intersect the boundary of $\mathcal{P}(G)$, the critical point of (5.9) becomes the stochastic Wasserstein Hamiltonian flow (see [17]),

$$\begin{aligned} \dot{\rho} &= \nabla_S \mathcal{H}_0(\rho, S)(1 + \dot{W}^{\delta})^2, \\ \dot{S} &= -\nabla_\rho \mathcal{H}_0(\rho, S)(1 + \dot{W}^{\delta})^2, \end{aligned} \tag{5.10}$$

where $\mathcal{H}_0(\rho, S) = \frac{1}{4}\sum_{ij\in E}(S_i - S_j)^2\theta_{ij}(\rho)$. Following the arguments in section (5.1), one can obtain similar results in Theorems 5.1–5.2.

We would like to point out that in this particular case, we can use the θ-connected components to study whether the optimal transfer achieves the boundary of the density manifold in optimal transport on a graph. We use the following example to illustrate the reason.

Remark 5.2. Let $V = \{1, 2, 3\}, E = \{(1, 2), (2, 3)\}$. Let $\rho^a = (0, 0, 1)$ and $\rho^b = (0, \frac{1}{2}, \frac{1}{2})$. We claim that ρ connecting ρ^a and ρ^b lies on the boundary as in the deterministic case. Assume that $(\rho, m) \in C_F(\rho^a, \rho^b)$ with $\rho_1 \not\equiv 0$. Then we have

$$\begin{aligned} \dot{\rho}_1 + m_{12}(1 + \dot{W}^{\delta}) &= 0, \\ \dot{\rho}_2 + (m_{21} + m_{23})(1 + \dot{W}^{\delta}) &= 0, \\ \dot{\rho}_3 + m_{32}(1 + \dot{W}^{\delta}) &= 0. \end{aligned}$$

Then one may define $(\widetilde{\rho}_1, \widetilde{\rho}_2, \widetilde{\rho}_3) = (0, \rho_1 + \rho_2, \rho_3)$ and $\widetilde{m}_{12} = 0,\ \widetilde{m}_{23} = m_{23}$. Then it holds that $\widetilde{\rho}(0) = \rho^a, \widetilde{\rho}(1) = \rho^b$, and $\dot{\widetilde{\rho}}_1 = 0$. By the definition of $\widetilde{\rho}$, it could be shown that

$$\begin{aligned} \dot{\widetilde{\rho}}_2 + \widetilde{m}_{23}(1 + \dot{W}^{\delta}) &= 0, \\ \dot{\widetilde{\rho}}_3 + \widetilde{m}_{32}(1 + \dot{W}^{\delta}) &= 0. \end{aligned}$$

Therefore, we have

$$\mathcal{A}(\rho, m) = \frac{1}{2}\int_0^1 \left(\frac{m_{12}^2}{\theta_{12}(\rho)} + \frac{m_{23}^2}{\theta_{23}(\rho)}\right)dt$$

and

$$\mathcal{A}(\widetilde{\rho}, \widetilde{m}) = \frac{1}{2}\int_0^1 \frac{m_{23}^2}{\theta_{23}(\widetilde{\rho})}dt = \frac{1}{2}\int_0^1 m_{23}^2 dt.$$

We have $\mathcal{A}(\widetilde{\rho}, \widetilde{m}) < \mathcal{A}(\rho, m)$, which leads to a contradiction.

Next we show the relationship between (5.9) with a small perturbation $\epsilon \dot{W}^\delta$ and the classical optimal transport problem. By defining $\hat{v} = v(1+\epsilon \dot{W}^\delta)$, (5.9) can then be rewritten as

$$\inf_{\rho,\widehat{v}} \left[\int_0^1 \frac{1}{2}\frac{1}{(1+\epsilon \dot{W}^\delta)^2}\langle \widehat{v}_t, \widehat{v}_t\rangle_{\theta(\rho_t)} dt\right]$$
$$\text{subject to } d\rho(t) + div_G^\theta(\rho(t)\widehat{v}(t)) = 0 \text{ and } \rho(0)=\rho_a,\ \rho(1)=\rho_b. \tag{5.11}$$

We show the Γ-convergence of

$$\mathcal{A}^{\epsilon_n}(\rho,m) := \int_0^1 \frac{1}{(1+\epsilon_n \dot{W}^\delta)^2}\sum_{ij}\frac{m_{ij}^2}{\theta^{ij}(\rho)}ds,\ \epsilon_n \to 0.$$

For a given $(\rho,m)\in C_F(\rho^a,\rho^b)$ and for a sequence $(\rho^{\epsilon_n}, m^{\epsilon_n})\in C_F(\rho^a,\rho^b)$ converging to (ρ,m), we have that

$$\liminf_{n\to\infty} \mathcal{A}^{\epsilon_n}(\rho^n,m^n) \ge \liminf_{n\to\infty}\int_0^1 \frac{1}{(1+\epsilon_n|\dot{W}^\delta|)^2}\sum_{ij}\frac{m_{ij}^2}{\theta^{ij}(\rho)}ds \ge \mathcal{A}(\rho,m).$$

By the dominated convergence theorem, it follows that

$$\lim_{\epsilon\to 0}\inf_{\rho,\widehat{v}}\left[\int_0^1 \frac{1}{2}\frac{1}{(1+\epsilon\dot{W}^\delta)^2}\langle\widehat{v}_t,\widehat{v}_t\rangle_{\theta(\rho_t)}dt\right] \le \inf_{\rho,\widehat{v}}\limsup_{\epsilon\to 0}\left[\int_0^1 \frac{1}{2}\frac{1}{(1+\epsilon\dot{W}^\delta)^2}\langle\widehat{v}_t,\widehat{v}_t\rangle_{\theta(\rho_t)}dt\right]$$
$$= \inf_{\rho,v}\left[\int_0^1 \frac{1}{2}\langle v_t,v_t\rangle_{\theta(\rho_t)}dt\right].$$

Combining the above estimates, we have that the limit of optimal control with common noise (5.11) is the classical optimal control a.s.

We would like to make a comparison between the variational problem (5.2) (its critical point (5.4)), and (5.9) (its critical point (5.10)). First, formally speaking, the limit of (5.2) on the continuous space is expected to be

$$\inf_v \int_0^1 \frac{1}{2}\mathbb{E}[|v(t,X_t)|^2]dt$$
$$\text{subject to } dX(t) = v(t,X(t))dt + \nabla\Sigma(X(t))dW_t^\delta$$
$$\text{and } X(0)\sim\rho(0)=\rho_a,\ X(1)\sim\rho(1)=\rho_b.$$

while that of (5.9) is expected to be

$$\inf_v \int_0^1 \frac{1}{2}\mathbb{E}[|v(t,X_t)|^2]dt$$
$$\text{subject to } dX(t) = v(t,X(t))dt + v(t,X(t))dW_t^\delta$$
$$\text{and } X(0)\sim\rho(0)=\rho_a,\ X(1)\sim\rho(1)=\rho_b.$$

Here the expectation $\mathbb{E}$ is conditionally on W.

Second, the limit of the critical point of (5.4) on the density space is

$$d\rho = \frac{\delta\mathcal{H}_0}{\delta S}(\rho,S)dt + \frac{\delta\mathcal{H}_1}{\delta S}(\rho,S)dW_t^\delta,$$
$$dS = -\frac{\delta\mathcal{H}_0}{\delta\rho}(\rho,S) - \frac{\delta\mathcal{H}_1}{\delta\rho}(\rho,S)dW_t^\delta. \tag{5.12}$$

Here $\mathcal{H}_0(\rho,S)=\int_{\mathbb{R}^d}\frac{1}{2}|\nabla S(x)|^2\rho(x)dx, \mathcal{H}_1(\rho,S)=\int_{\mathbb{R}^d}\nabla S(x)\nabla\Sigma(x)\rho(x)dx$. Then under suitable conditions, one can expect that when $\delta\to 0$, the limit of (5.12) becomes the stochastic Wasserstein Hamiltonian flow

$$d\rho=\frac{\delta\mathcal{H}_0}{\delta S}(\rho,S)dt+\frac{\delta\mathcal{H}_1}{\delta S}(\rho,S)\circ dW_t,$$
$$dS=-\frac{\delta\mathcal{H}_0}{\delta\rho}(\rho,S)-\frac{\delta\mathcal{H}_1}{\delta\rho}(\rho,S)\circ dW_t, \quad \text{a.s.}$$

In contrast, the limit of (5.4) on the continuous space is

$$d\rho=\frac{\delta\mathcal{H}_0}{\delta S}(\rho,S)(1+\dot{W}_t^\delta)^2dt,$$
$$dS=-\frac{\delta\mathcal{H}_0}{\delta\rho}(\rho,S)(1+\dot{W}_t^\delta)^2dt. \tag{5.13}$$

By taking $\delta\to 0$, it is still unclear how to define a suitable limit of (5.13) due to the term $\lim_{\delta\to 0}(\dot{W}^\delta)^2$. If the limit exists, what is the difference compared to the original stochastic Wasserstein Hamiltonian flow? Those are interesting questions that can be further investigated in the future. Understanding them can help to design better numerical schemes by combining the stochastic Wasserstein Hamiltonian flow on a graph and some structure-preserving temporal integration.

Third, for a fixed $\delta>0$, one can use the θ-connected components [31, 24] to study whether the optimal transfer achieves the boundary of the density manifold for (5.10) on the discrete graph. However, this method may fail for (5.4).

6. Conclusions. In this paper, using the notion of common noise, we establish the initial value and two-point boundary value problems of stochastic Wasserstein Hamiltonian flows on the finite graph. We show the local well-posedness of the initial value problem always holds, up to a positive time, for stochastic Wasserstein Hamiltonian flow and provide a sufficient condition of its global well-posedness. For the boundary value problem, by exploiting the Wong–Zakai approximation, we obtain the existence of the minimizer of the optimal control problem perturbed by common noise and derive its dual formula. However, many questions remain to be answered. For example, how to show the existence of the minimizer of the optimal control problem driven by the other Wiener process (not common noise)? Does the minimizer exist for the general variational principle with common noise? When considering the lattice graphs, can we get some characterizations of the minimizer for the continuous problem if the mesh size is reduced to zero? These questions are very important for numerical computations of the stochastic Wasserstein Hamiltonian flow and its related control problem. Although our focus is on using common noise in this paper, we hope the results may shed light on the investigation of Wasserstein Hamiltonian flow with other types of noise too.

Acknowledgement. The authors would like to thank the anonymous referees for their useful suggestions which helped improve the quality of the article.

REFERENCES

[1] G. P. Agrawal, *Applications of Nonlinear Fiber Optics*, Academic Press, San Diego, 2001.
[2] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed., Academic Press, San Diego, 2001.
[3] L. Ambrosio and W. Gangbo, *Hamiltonian ODEs in the Wasserstein space of probability measures*, Comm. Pure Appl. Math., 61 (2008), pp. 18–53, https://doi.org/10.1002/cpa.20188.

[4] O. Bang, P. L. Christiansen, F. If, K. O. Rasmussen, and Y. B. Gaididei, *Temperature effects in a nonlinear model of monolayer Scheibe aggregates*, Phys. Rev. E (3), 49 (1994), pp. 4627–4636, https://doi.org/10.1103/PhysRevE.49.4627.

[5] J. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp. 375–393, https://doi.org/10.1007/s002110050002.

[6] R. Carmona and F. Delarue, *Mean field games with common noise and master equations*, in Probabilistic Theory of Mean Field Games with Applications. II, Probab. Theory Stoch. Model. 84, Springer, Cham, Switzerland, 2018.

[7] R. Carmona, F. Delarue, and D. Lacker, *Mean field games with common noise*, Ann. Probab., 44 (2016), pp. 3740–3803, https://doi.org/10.1214/15-AOP1060.

[8] S. Chow, W. Huang, Y. Li, and H. Zhou, *Fokker-Planck equations for a free energy functional or Markov process on a graph*, Arch. Ration. Mech. Anal., 203 (2012), pp. 969–1008, https://doi.org/10.1007/s00205-011-0471-6.

[9] S. Chow, W. Li, C. Mou, and H. Zhou, *Dynamical Schrödinger bridge problems on graphs*, J. Dynam. Differential Equations, 34 (2022), pp. 2511–2530, https://doi.org/10.1007/s10884-021-09977-1.

[10] S. Chow, W. Li, and H. Zhou, *Entropy dissipation of Fokker-Planck equations on graphs*, Discrete Contin. Dyn. Syst., 38 (2018), pp. 4929–4950, https://doi.org/10.3934/dcds.2018215.

[11] S. Chow, W. Li, and H. Zhou, *A discrete Schrödinger equation via optimal transport on graphs*, J. Funct. Anal., 276 (2019), pp. 2440–2469, https://doi.org/10.1016/j.jfa.2019.02.005.

[12] S. Chow, W. Li, and H. Zhou, *Wasserstein Hamiltonian flows*, J. Differential Equations, 268 (2020), pp. 1205–1219, https://doi.org/10.1016/j.jde.2019.08.046.

[13] C. Conforti and M. Pavon, *Extremal flows in Wasserstein space*, J. Math. Phys., 59 (2018), 063502.

[14] J. Cui, L. Dieci, and H. Zhou, *A continuation multiple shooting method for Wasserstein geodesic equation*, SIAM J. Sci. Comput., 44 (2022), pp. A2918–A2943, https://doi.org/10.1137/21M142160X.

[15] J. Cui, J. Hong, and Z. Liu, *Strong convergence rate of finite difference approximations for stochastic cubic Schrödinger equations*, J. Differential Equations, 263 (2017), pp. 3687–3713.

[16] J. Cui, J. Hong, and D. Sheng, *Density function of numerical solution of splitting AVF scheme for stochastic Langevin equation*, Math. Comp., 91 (2022), pp. 2283–2333, https://doi.org/10.1090/mcom/3752.

[17] J. Cui, S. Liu, and H. Zhou, *Stochastic Wasserstein Hamiltonian flows*, arXiv:2111.15163, 2021.

[18] J. Cui, S. Liu, and H. Zhou, *What is a stochastic Hamiltonian process on finite graph? An optimal transport answer*, J. Differential Equations, 305 (2021), pp. 428–457, https://doi.org/10.1016/j.jde.2021.10.009.

[19] J. Cui, S. Liu, and H. Zhou, *Optimal control for stochastic nonlinear Schrödinger equation on graph*, arXiv:2209.05346, 2022, To appear in SIAM J. Control Optim.

[20] J. Cui and L. Sun, *Stochastic logarithmic Schrödinger equations: Energy regularized approach*, arXiv:2102.12607, 2021, To appear in SIAM J. Math. Anal.

[21] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia Math. Appl. 152, 2nd ed., Cambridge University Press, Cambridge, 2014.

[22] G. E. Falkovich, I. Kolokolov, V. Lebedev, and S. K. Turitsyn, *Statistics of soliton-bearing systems with additive noise*, Phys. Rev. E (3), 63 (2001), 025601, https://doi.org/10.1103/PhysRevE.63.025601.

[23] W. Gangbo, H. Kim, and T. Pacini, *Differential Forms on Wasserstein Space and Infinite-Dimensional Hamiltonian Systems*, Mem. Amer. Math. Soc. 211, American Mathematical Society, Providence, RI, 2011, https://doi.org/10.1090/S0065-9266-2010-00610-0.

[24] W. Gangbo, W. Li, and C. Mou, *Geodesics of minimal length in the set of probability measures on graphs*, ESAIM Control Optim. Calc. Var., 25 (2019), 78, https://doi.org/10.1051/cocv/2018052.

[25] M. Huang, R. P. Malhamé, and P. E. Caines, *Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst., 6 (2006), pp. 221–251, http://projecteuclid.org/euclid.cis/1183728987.

[26] V. V. Konotop and L. Vázquez, *Nonlinear Random Waves*, World Scientific, River Edge, NJ, 1994, https://doi.org/10.1142/9789814350259.

[27] J. D. Lafferty, *The density manifold and configuration space quantization*, Trans. Amer. Math. Soc., 305 (1988), pp. 699–741, https://doi.org/10.2307/2000885.

[28] J.-M. Lasry and P.-L. Lions, *Mean field games*, Jpn J. Math., 2 (2007), pp. 229–260, https://doi.org/10.1007/s11537-007-0657-8.

[29] F. Léger and W. Li, *Hopf-Cole transformation via generalized Schrödinger bridge problem*, J. Differential Equations, 274 (2021), pp. 788–827, https://doi.org/10.1016/j.jde.2020.10.029.

[30] C. Léonard, *A survey of the Schrödinger problem and some of its connections with optimal transport*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1533–1574, https://doi.org/10.3934/dcds.2014.34.1533.

[31] J. Maas, *Gradient flows of the entropy for finite Markov chains*, J. Funct. Anal., 261 (2011), pp. 2250–2292, https://doi.org/10.1016/j.jfa.2011.06.009.

[32] J. F. Mertens, S. Sorin, and S. Zamir, *Repeated Games*, Econom. Soc. Monogr. 55, Cambridge University Press, New York, 2015, https://doi.org/10.1017/CBO9781139343275.

[33] D. Möbius and H. Kugn, *Monolayer assemblies of dyes to study the role of thermal collisions in energy transfer*, Israel J. Chem., 18 (1979), pp. 375–384.

[34] E. Nelson, *Derivation of the Schrödinger equation from Newtonian mechanics*, Phys. Rev. (2), 150 (1966), pp. 1079–1085, https://doi.org/10.1103/PhysRev.150.1079.

[35] E. Nelson, *Quantum Fluctuations*, Princeton Ser. Phys., Princeton University Press, Princeton, NJ, 1985.

[36] E. Nelson, *Field theory and the future of stochastic mechanics*, in Stochastic Processes in Classical and Quantum Systems (Ascona, 1985), Lecture Notes in Phys. 262, Springer, Berlin, 1986, pp. 438–469, https://doi.org/10.1007/3-540-17166-5.

[37] J. v. Neumann, *Ein system algebraisch unabhängiger Zahlen*, Math. Ann., 99 (1928), pp. 134–141, https://doi.org/10.1007/BF01459089.

[38] B. K. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, Springer, Berlin, 2003.

[39] C. Villani, *Optimal Transport*, Grundlehren Math. Wiss. 338, Springer, Berlin, 2009, https://doi.org/10.1007/978-3-540-71050-9.

[40] X. Wang, K. Lu, and B. Wang, *Wong-Zakai approximations and attractors for stochastic reaction-diffusion equations on unbounded domains*, J. Differential Equations, 264 (2018), pp. 378–424, https://doi.org/10.1016/j.jde.2017.09.006.

[41] E. Wong and M. Zakai, *On the convergence of ordinary integrals to stochastic integrals*, Ann. Math. Statist., 36 (1965), pp. 1560–1564, https://doi.org/10.1214/aoms/1177699916.