

# CONTROLLING CONSERVATION LAWS I: ENTROPY–ENTROPY FLUX

WUCHEN LI, SITING LIU, AND STANLEY OSHER

**ABSTRACT.** We study a class of variational problems for regularized conservation laws with Lax’s entropy-entropy flux pairs. We first introduce a modified optimal transport space based on conservation laws with diffusion. Using this space, we demonstrate that conservation laws with diffusion are “flux–gradient flows”. We next construct variational problems for these flows, for which we derive dual PDE systems for regularized conservation laws. Several examples, including traffic flow and Burgers’ equation, are presented. Incorporating both primal-dual algorithms and monotone schemes, we successfully compute the control of conservation laws.

## 1. INTRODUCTION

Regularized conservation laws<sup>1</sup> are essential classes of dynamics in physics, materials sciences, and mathematical modeling, with applications to inverse problems and AI sampling problems [16, 22, 35]. Examples of conservation law equations include traffic flows, Burgers’ equation, and compressible Navier-Stokes equations, etc.

Consider a system of initial value PDEs below.

$$\partial_t u(t, x) + \mathcal{B}(u(t, x)) = \beta \mathcal{C}(u(t, x)), \quad u(0, x) = u_0, \quad (1)$$

where  $x \in \Omega \subset \mathbb{R}^n$ ,  $u: [0, \infty) \times \Omega \rightarrow \mathbb{R}^d$  is a unknown vector function,  $u_0$  is a given initial condition,  $\mathcal{B}: C^\infty(\Omega; \mathbb{R}^d) \rightarrow C^\infty(\Omega; \mathbb{R}^d)$  is a “conservative” differential operator,  $\mathcal{C}: C^\infty(\Omega; \mathbb{R}^d) \rightarrow C^\infty(\Omega; \mathbb{R}^d)$  is a dissipative differential (diffusion) operator and  $\beta \geq 0$  is a diffusion constant. For simplicity of presentation, we assume that  $\Omega$  is a compact convex set, and the PDE (1) has periodic boundary conditions. E.g.,  $\Omega = \mathbb{T}^n$ , where  $\mathbb{T}^n$  is a  $n$ -dimensional torus.

In this paper, we introduce variational problems related to the PDE (1). They generalize mean-field information dynamics [26, 27]; see Figure 1. Here we shall design suitable *modified optimal transport spaces* for PDE (1), namely mean-field information metric spaces. In these metric spaces, we demonstrate that the PDE (1) has a dissipative variational structure. And we name PDE (1) *flux–gradient flows*. We then design a control problem

---

*Key words and phrases.* Conservation laws; Entropy–entropy flux pairs; Optimal transport; Mean-field games; Hamiltonian flows; Monotone schemes; Primal-dual algorithms.

W. Li thanks the start-up funding from the University of South Carolina, and DEPSCoR Research Collaboration white paper “Transport information geometric modeling of complex systems”. W. Li, S. Liu and S. Osher thank the funding from AFOSR MURI FA9550-18-1-0502 and ONR grants: N00014-18-1-2527, N00014-20-1-2093, and N00014-20-1-2787.

<sup>1</sup>For simplicity, we omit the regularized throughout the paper.

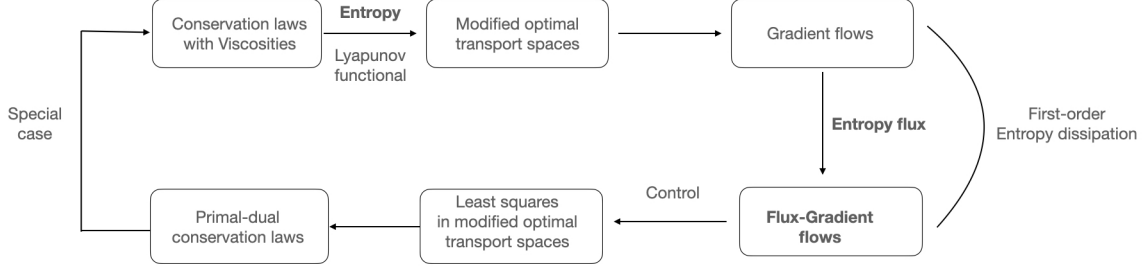


FIGURE 1. We study variational problems of conservation laws in entropy-entropy flux pairs induced metric spaces.

of PDE (1) in metric space. By finding the critical point of control problem, we derive a dual PDE system for equation (1). The primal-dual pair of PDE system satisfies a Hamiltonian flow in the metric space, where the Hamiltonian functional depends on the flux function and “kinetic energy”. Several numerical examples, including traffic flow and Burgers’ equation, are presented.

The main results are sketched below.

**Assumption 1:** Suppose that there exists a function  $G: \mathbb{R}^d \rightarrow \mathbb{R}$ , such that

$$\int_{\Omega} G'(u) \cdot \mathcal{B}(u) dx = 0.$$

**Assumption 2:** Suppose that there exists a “symmetric nonnegative definite” operator  $L_C(u): C^\infty(\Omega) \rightarrow C^\infty(\Omega)$ , such that

$$\int_{\Omega} G'(u) \cdot \mathcal{C}(u) dx = - \int_{\Omega} (G'(u), L_C(u)G'(u)) dx \leq 0.$$

Following [22], under assumption 1 and 2, we demonstrate that

$$\mathcal{G}(u(t, \cdot)) = \int_{\Omega} G(u(t, x)) dx,$$

forms a Lyapunov functional for PDE (1). In other words, along PDE (1), the time derivative of  $\mathcal{G}(u)$  is non-positive:

$$\begin{aligned} \frac{d}{dt} \mathcal{G}(u(t, \cdot)) &= \int_{\Omega} G'(u) \partial_t u dx = \int_{\Omega} G'(u) \cdot (-\mathcal{B}(u) + \beta \mathcal{C}(u)) dx \\ &= \beta \int_{\Omega} G'(u) \cdot \mathcal{C}(u) dx = -\beta \int_{\Omega} (G'(u), L_C(u)G'(u)) dx \leq 0. \end{aligned}$$

Based on the Lyapunov functional  $\mathcal{G}$ , we design optimal control problems of PDE (1). In detail: let  $d = 1$ ,

$$\mathcal{B}(u) = \nabla \cdot f(u), \quad \mathcal{C}(u) = \nabla \cdot (A(u)\nabla u), \quad \mathcal{L}_C(u) := -\nabla \cdot (A(u)G''(u)^{-1}\nabla),$$

where  $f: \mathbb{R} \rightarrow \mathbb{R}^n$  is a flux function and  $A, AG''^{-1}: \Omega \rightarrow \mathbb{R}^{n \times n}$  are both symmetric non-negative definite matrix functions. Given a suitable potential functional  $\mathcal{F}: C^\infty(\Omega; \mathbb{R}) \rightarrow$

$\mathbb{R}$  and a terminal functional  $\mathcal{H}: C^\infty(\Omega; \mathbb{R}) \rightarrow \mathbb{R}$ , consider a variational problem

$$\inf_{u, v, u_1} \int_0^1 \left[ \int_\Omega \frac{1}{2} (v, A(u)G''(u)^{-1}v) dx - \mathcal{F}(u) \right] dt + \mathcal{H}(u_1), \quad (2a)$$

where the infimum is taken among variables  $v: [0, 1] \times \Omega \rightarrow \mathbb{R}^n$ ,  $u: [0, 1] \times \Omega \rightarrow \mathbb{R}$ , and  $u_1: \Omega \rightarrow \mathbb{R}$  satisfying

$$\partial_t u + \nabla \cdot f(u) + \nabla \cdot (A(u)G''(u)^{-1}v) = \beta \nabla \cdot (A(u)\nabla u), \quad u(0, x) = u_0(x). \quad (2b)$$

By Proposition 8, we derive a critical point system of the optimal control problem (2). Define a Lagrange multiplier function  $\Phi: [0, 1] \times \Omega \rightarrow \mathbb{R}$ . Then  $v = \nabla \Phi$ , and  $(u, \Phi)$  satisfies a pair of PDEs:

$$\begin{cases} \partial_t u + \mathcal{B}(u) - L_{\mathcal{C}}(u)(\Phi) - \beta \mathcal{C}(u) = 0, \\ \partial_t \Phi + \frac{\delta}{\delta u} \int_\Omega \left( \frac{1}{2} (\Phi, L_{\mathcal{C}}(u)\Phi) - (\mathcal{B}(u), \Phi) + \beta (\Phi, \mathcal{C}(u)) \right) dx + \frac{\delta}{\delta u} \mathcal{F}(u) = 0, \end{cases} \quad (3a)$$

with both initial and terminal time boundary conditions

$$u(0, x) = u_0(x), \quad \frac{\delta}{\delta u_1(x)} \mathcal{H}(u_1) + \Phi(1, x) = 0. \quad (3b)$$

In above formulations,  $\frac{\delta}{\delta u}$  is the  $L^2$  first variation operator w.r.t. variable  $u$ . Here we aim to study the variational problem (2) and its critical point system (3) in the following two aspects.

- (i) Modeling: We design and control models for conservation law dynamics. A typical example is that we control the kinetic energy of the density under traffic flow.
- (ii) Computation: By choosing some special potential and terminal energies, the minimizer of the mean-field information control problem (2) is consistent with the initial value problem (1). This leads to stable and convergent primal-dual algorithms for initial value conservation laws.

In the literature, there are joint studies towards dynamics in mean-field information metric spaces; see [1, 2, 4, 6, 18, 32, 34, 35, 36] and many references therein. Several types of dynamics in these metric spaces have been studied in recent decades. First, gradient flows have been systematically studied in [1, 9, 14]. Next, Hamiltonian flows with generalizations to differential games have been investigated by [8, 21, 19]. A particular type of Hamiltonian flows, namely Schrödinger bridge systems, and their mean-field generalizations, are widely studied in [3, 11, 15, 12, 24, 23]. Compared to previous works, we focus on conservative-dissipative equations, which are non-gradient flow systems. In particular, we connect the conservation law equations with generalized optimal transport variational structures. Concretely, we study equation (1) associated with a least-square type control problem in the designed metric space. We remark that the control problem in (2) connects with, but is different from, the ones in [5, 6]. In detail, [6] designed variational problems in term of entropy functional  $-\mathcal{F}(u) = \mathcal{G}(u) = \int_\Omega G(u)dx$ , which can solve the initial value conservation laws. And [6] also enforces  $v = 0$  in the control problem (2). In contrast, we control “kinetic energy” in the modified optimal transport space generated by the entropy-entropy flux pairs. Following this approach, the PDE pair (3) can be used to control conservation laws.

The paper is organized below. In section 2, we provide two motivating examples of variational problem (2). In section 3 and 4, we present the main result of this paper. We first present the variational formulation of equation (1) in modified optimal transport space. We next derive a variational problem and a pair of PDEs for conservation laws. In section 5, we design primal-dual algorithms to numerically solve the optimal control problems of conservation laws.

## 2. MOTIVATION

In this section, we provide two examples of conservative-dissipative equations (1). In these examples, we demonstrate variational problems (2) and PDE pairs (3) with their physical and modeling explanations.

**2.1. Viscous Burgers' equation.** Consider a one dimensional viscous Burgers' equation

$$\partial_t u(t, x) + \partial_x \frac{u(t, x)^2}{2} = \beta \partial_{xx} u(t, x).$$

Here  $\partial_x$ ,  $\partial_{xx}$  are the first and the second derivatives w.r.t.  $x$ , and the unknown variable is  $u: \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ . In this case, the conservative and dissipative operators satisfy

$$\mathcal{B}(u) = \partial_x \left( \frac{u^2}{2} \right), \quad \mathcal{C}(u) = \partial_{xx} u.$$

Define a function  $G: \mathbb{R}^1 \rightarrow \mathbb{R}^1$  by

$$G(u) = \frac{u^2}{2}, \quad G'(u) = u.$$

Then assumption 1 is satisfied since

$$\int_{\Omega} G'(u) \cdot \mathcal{B}(u) dx = \int_{\Omega} \partial_x \left( \frac{u^2}{2} \right) \cdot u dx = \int_{\Omega} u^2 \cdot \partial_x u dx = \int_{\Omega} \partial_x \left( \frac{u^3}{3} \right) dx = 0.$$

Denote a ‘‘symmetric nonnegative definite operator’’ by

$$L\mathcal{C}(u) = -\partial_{xx}.$$

Assumption 2 is satisfied since

$$\int_{\Omega} G'(u) \cdot \mathcal{C}(u) dx = \int_{\Omega} u \cdot \partial_{xx} u dx = - \int_{\Omega} |\partial_x u|^2 dx,$$

where the second equality is from the integration by parts formula.

*Formulations:* Variational problem (2) forms an optimal control problem for viscous Burgers' equation. Consider

$$\inf_{u, v, u_1} \int_0^1 \left[ \int_{\Omega} \frac{1}{2} |v(t, x)|^2 dx - \mathcal{F}(u) \right] dt + \mathcal{H}(u_1), \quad (4a)$$

such that

$$\partial_t u(t, x) + \partial_x \left( \frac{u(t, x)^2}{2} \right) + \partial_x v(t, x) = \beta \partial_{xx} u(t, x), \quad u(0, x) = u_0(x). \quad (4b)$$

Here, the minimizer system of variational problem (4) satisfies a pair of PDEs:  $v(t, x) = \nabla \Phi(t, x)$  and

$$\begin{cases} \partial_t u(t, x) + \partial_x \left( \frac{u(t, x)^2}{2} \right) + \partial_{xx} \Phi(t, x) = \beta \partial_{xx} u(t, x), \\ \partial_t \Phi(t, x) + (u(t, x), \partial_x \Phi(t, x)) + \frac{\delta}{\delta u} \mathcal{F}(u)(t, x) = -\beta \partial_{xx} \Phi(t, x). \end{cases}$$

*Velocity control.* The unknown variable  $u$  in Burgers' equation describes the velocity field of the fluid over time. In the variational problem (4), we design a potential field  $\Phi$  to control the evolution of velocity field  $u$ . Here, the background dynamics of  $u$  is the classical initial value viscous Burgers' equation. The designed variational problem is to control a velocity field under suitable running and terminal costs.

**2.2. Traffic flow equation.** Consider a one dimensional traffic flow equation

$$\partial_t u(t, x) + \nabla \cdot (u(t, x)(1 - u(t, x))) = \beta \Delta u(t, x).$$

Here  $\nabla \cdot$ ,  $\Delta$  are divergence, Laplacian operators w.r.t  $x$ , respectively, and the unknown variable is  $u: \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}_+$ . In this case, the conservative and the dissipative operators satisfy

$$\mathcal{B}(u) = \nabla \cdot (u(1 - u)), \quad \mathcal{C}(u) = \Delta u.$$

Define a function  $G: \mathbb{R}_+^1 \rightarrow \mathbb{R}_+^1$  by

$$G(u) = u \log u - u, \quad G'(u) = \log u.$$

Then assumption 1 is satisfied since

$$\int_{\Omega} G'(u) \cdot \mathcal{B}(u) dx = \int_{\Omega} \log u \cdot \nabla \cdot (u(1 - u)) dx = \int_{\Omega} \nabla \cdot \Psi(u) dx = 0,$$

where  $\Psi(u) = \int_0^u (1 - 2z) \log z dz$ . Denote a "symmetric nonnegative definite operator" by

$$L_{\mathcal{C}}(u) = -\nabla \cdot (u \nabla).$$

Assumption 2 is satisfied since

$$\begin{aligned} \int_{\Omega} G'(u) \cdot \mathcal{C}(u) dx &= \int_{\Omega} \log u \cdot \Delta u dx \\ &= \int_{\Omega} (\nabla \log u, \nabla u) dx \\ &= - \int_{\Omega} \|\nabla \log u\|^2 u dx \\ &= - \int_{\Omega} (\log u, -\nabla \cdot (u \nabla \log u)) dx, \end{aligned}$$

where the second equality is from the integration by parts formula and the third equality holds by the fact that  $\frac{\nabla u}{u} = \nabla \log u$ , i.e.,  $\nabla u = u \nabla \log u$ .

*Formulations:* Variational problem (2) forms an optimal control problem for traffic flows. In detail, consider

$$\inf_{u, v, v_1} \int_0^1 \left[ \int_{\Omega} \frac{1}{2} \|v(t, x)\|^2 u(t, x) dx - \mathcal{F}(u) \right] dt + \mathcal{H}(u_1), \quad (5a)$$

such that

$$0 \leq u(t, x) \leq 1, \quad \text{for all } t \in [0, 1],$$

and

$$\partial_t u(t, x) + \nabla \cdot (u(t, x)(1 - u(t, x))) + \nabla \cdot (u(t, x)v(t, x)) = \beta \Delta u(t, x), \quad u(0, x) = u_0(x). \quad (5b)$$

Here, the minimizer system of variational problem (4) satisfies a pair of PDEs. When  $u \in (0, 1)$ , there exists a scalar function  $\Phi$ , such that  $v(t, x) = \nabla \Phi(t, x)$  and

$$\begin{cases} \partial_t u(t, x) + \nabla \cdot (u(t, x)(1 - u(t, x))) + \nabla \cdot (u(t, x)\nabla \Phi(t, x)) = \beta \Delta u(t, x), \\ \partial_t \Phi(t, x) + (1 - 2u(t, x), \nabla \Phi(t, x)) + \frac{1}{2} \|\nabla \Phi(t, x)\|^2 + \frac{\delta}{\delta u(t, x)} \mathcal{F}(u) = -\beta \Delta \Phi(t, x). \end{cases}$$

*Position control.* The unknown variable  $u$  in traffic flows represents the density function of cars (particles) in a given spatial domain. Here, the background dynamics of  $u$  is the classical traffic flow. The control variable is the velocity for enforcing each car's velocity in addition to its background traffic flow dynamics. The goal is to control the “total enforced kinetic energy” of all cars, in which individual cars can determine their velocities through both noises and traffic flow interactions.

We shall demonstrate that variational problems (2) can be formulated in term of general conservation law equations associated with entropy-entropy flux pairs.

### 3. ENTROPY-ENTROPY FLUX AND FLUX-GRADIENT FLOWS

In this section, we first recall Lax's entropy-entropy flux pairs for conservation law equations; see [16, 22]. Here the entropy-entropy flux pair is used to construct a Lyapunov functional for PDE (1). Using this Lyapunov functional with the dissipative operator, we next review and formulate both metric spaces and gradient flows; see [1, 9, 32]. Combining these facts with flux functions, we name PDE (1) *flux-gradient flows* in metric spaces. In later on sections, we shall demonstrate that the flux-gradient flow formulation is useful in designing control problems of conservation laws.

**3.1. Entropy-Entropy flux pairs and Lyapunov functionals.** Consider

$$u: \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}^1, \quad \mathcal{B}(u) = \nabla \cdot f(u), \quad \mathcal{C}(u) = \nabla \cdot (A(u)\nabla u).$$

In this case, equation (1) satisfies

$$\partial_t u(t, x) + \nabla \cdot f(u(t, x)) = \beta \nabla \cdot (A(u)\nabla u), \quad (6)$$

where  $u: \mathbb{R}^n \rightarrow \mathbb{R}$  is a scalar function,  $f = (f_1, \dots, f_n)$  is a flux vector function, and  $A = (A_{ij})_{1 \leq i, j \leq n} \in \mathbb{R}^{n \times n}$  is a semi-positive definite matrix function. If  $\beta = 0$ , equation (6) is a scalar conservation law equation

$$\partial_t u(t, x) + \nabla \cdot f(u(t, x)) = 0. \quad (7)$$

**Definition 1** (Entropy-entropy flux pair condition [22]). *We call  $(G, \Psi)$  an entropy-entropy flux pair for the conservation law (7) if there exists a convex function  $G: \mathbb{R} \rightarrow \mathbb{R}$ , and  $\Psi: \mathbb{R} \rightarrow \mathbb{R}^n$ , such that*

$$\Psi'(u) = f'(u)G'(u).$$

*Remark 1.* We remark that the entropy-entropy flux condition is trivial for scalar conservation laws. Every differentiable convex function  $G$  satisfies this condition. For simplicity of presentation, we only focus on the scalar case, and leave the study of conservation law systems in future work.

In fact, Lax's entropy-entropy flux pair introduces a class of functionals, which can be used as Lyapunov functionals for PDE (1). Denote

$$\mathcal{G}(u) = \int_{\Omega} G(u) dx.$$

For assumption 1,

$$\begin{aligned} \int_{\Omega} G'(u) \cdot \mathcal{B}(u) dx &= \int_{\Omega} G'(u) \nabla \cdot f(u) dx \\ &= \int_{\Omega} G'(u) (f'(u), \nabla u) dx \\ &= \int_{\Omega} (\Psi'(u), \nabla u) dx \\ &= \int_{\Omega} \nabla \cdot \Psi(u) dx = 0, \end{aligned}$$

where we apply the fact that  $f'(u)G'(u) = \Psi'(u)$  and  $\int_{\Omega} \nabla \cdot \Psi(u) dx = 0$  in the last equality.

For assumption 2,

$$\begin{aligned} \int_{\Omega} G'(u) \cdot \mathcal{C}(u) dx &= \int_{\Omega} G'(u) \nabla \cdot (A(u) \nabla u) dx \\ &= - \int_{\Omega} (\nabla G'(u), A(u) \nabla u) dx \\ &= - \int_{\Omega} (\nabla G'(u), A(u) G''(u)^{-1} \nabla G'(u)) dx, \end{aligned}$$

where we apply  $\nabla G'(u) = G''(u) \nabla u$  in the last equality. We observe a fact that  $A(u) G''(u)^{-1} \succeq 0$ , since  $A$  is nonnegative definite and  $G''(u) > 0$ . Hence we know that

$$\begin{aligned} \partial_t \mathcal{G}(u) &= \int_{\Omega} G'(u) \cdot \partial_t u dx \\ &= - \int_{\Omega} G'(u) \cdot \mathcal{B}(u) dx + \beta \int_{\Omega} G'(u) \cdot \mathcal{C}(u) dx \\ &= - \beta \int_{\Omega} (\nabla G'(u), A(u) G''(u)^{-1} \nabla G'(u)) dx \leq 0. \end{aligned}$$

This implies that  $\mathcal{G}(u)$  is a Lyapunov functional for PDE (6).

**3.2. Metric spaces and gradient flows.** We next provide a condition to define a metric space for the unknown variable  $u$ . Here, the metric space connects Lyapunov functionals with dissipative operators through gradient descent flows.

**Definition 2** (Entropy-entropy flux-metric condition). *We call  $(G, \Psi)$  an entropy-entropy flux pair-metric for equation (6) if there exists a convex function  $G: \mathbb{R} \rightarrow \mathbb{R}$ , and  $\Psi: \mathbb{R} \rightarrow \mathbb{R}^n$ , such that*

$$\Psi'(u) = f'(u)G'(u), \quad A(u)G''(u)^{-1} \text{ is a semi positive definite symmetric matrix function.}$$

Under the entropy-entropy flux-metric condition, assumption 2 implies a metric operator below. Define the space of function  $u$  by

$$\mathcal{M} = \left\{ u \in C^\infty(\Omega) : \int_{\Omega} u(x)dx = \text{constant} \right\}.$$

The tangent space of  $\mathcal{M}(u)$  at point  $u$  is defined by

$$T_u\mathcal{M} = \left\{ \sigma \in C^\infty(\Omega) : \int_{\Omega} \sigma(x)dx = 0 \right\}.$$

Denote an elliptic operator  $L_C: C^\infty(\Omega) \rightarrow C^\infty(\mathbb{R})$  by

$$L_C(u) = -\nabla \cdot (A(u)G''(u)^{-1}\nabla).$$

We are ready to define the metric inner product for the dissipation operator.

**Definition 3** (Metric). *The inner product  $\mathbf{g}(u): T_u\mathcal{M} \times T_u\mathcal{M} \rightarrow \mathbb{R}$  is given below.*

$$\begin{aligned} \mathbf{g}(u)(\sigma_1, \sigma_2) &= \int_{\Omega} (\Phi_1, L_C(u)\Phi_2)dx \\ &= - \int_{\Omega} \Phi_1 \nabla \cdot (A(u)G''(u)^{-1}\nabla\Phi_2)dx \\ &= \int_{\Omega} (\nabla\Phi_1, A(u)G''(u)^{-1}\nabla\Phi_2)dx \\ &= \int_{\Omega} \sigma_1\Phi_2dx = \int_{\Omega} \sigma_2\Phi_1dx, \end{aligned}$$

where  $\Phi_i \in C^\infty(\Omega)$  satisfies

$$\sigma_i = -\nabla \cdot (A(u)G''(u)^{-1}\nabla\Phi_i), \quad i = 1, 2.$$

From now on, we call  $(\mathcal{M}, \mathbf{g})$  the metric space. We next review and present the gradient decent flow in metric space  $(\mathcal{M}, \mathbf{g})$ . Here the dissipative part of PDE (1) comes from the gradient flow of the proposed Lyapunov (entropy) functional.

**Proposition 4** (Gradient flow). *Given an energy functional  $\mathcal{E}: \mathcal{M} \rightarrow \mathbb{R}$ , the gradient flow of  $\mathcal{F}$  in  $(\mathcal{M}, \mathbf{g})$  satisfies*

$$\partial_t u = \nabla \cdot (A(u)G''(u)^{-1}\nabla \frac{\delta}{\delta u}\mathcal{E}(u)).$$

If  $\mathcal{E}(u) = \mathcal{G}(u) = \int_{\Omega} G(u)dx$ , then the above gradient flow satisfies

$$\partial_t u = \nabla \cdot (A(u)\nabla u).$$



*Proof.* The derivation of gradient flow follows from the definition. For the completeness of this paper, we still present the proof here. The gradient operator  $\text{grad}\mathcal{E} \in T_u\mathcal{M}$  is defined by

$$\mathbf{g}(u)(\text{grad}\mathcal{E}(u), \sigma) = \int_{\Omega} \left( \frac{\delta}{\delta u} \mathcal{E}(u), \sigma \right) dx,$$

for any  $\sigma \in T_u\mathcal{M}$ . Let  $\sigma = L_{\mathcal{C}}(u)(\Phi) = -\nabla \cdot (A(u)G''(u)^{-1}\nabla\Phi)$ , the above definition forms

$$\begin{aligned} \mathbf{g}(u)(\text{grad}\mathcal{E}(u), \sigma) &= \int_{\Omega} (\text{grad}\mathcal{E}(u), \Phi) dx \\ &= \int_{\Omega} \left( \frac{\delta}{\delta u} \mathcal{E}(u), L_{\mathcal{C}}(u)(\Phi) \right) dx \\ &= \int_{\Omega} (L_{\mathcal{C}}(u) \left( \frac{\delta}{\delta u} \mathcal{E}(u) \right), \Phi) dx. \end{aligned}$$

Since the above equality holds for any smooth function  $\Phi$ , we let

$$\text{grad}\mathcal{E}(u) = L_{\mathcal{C}}(u) \left( \frac{\delta}{\delta u} \mathcal{E}(u) \right) = -\nabla \cdot \left( A(u)G''(u)^{-1}\nabla \frac{\delta}{\delta u} \mathcal{E}(u) \right).$$

Hence the gradient decent flow satisfies

$$\partial_t u = -\text{grad}\mathcal{E}(u) = \nabla \cdot \left( A(u)G''(u)^{-1}\nabla \frac{\delta}{\delta u} \mathcal{E}(u) \right).$$

If  $\mathcal{E}(u) = \mathcal{G}(u)$ , then

$$\begin{aligned} \partial_t u &= \nabla \cdot (A(u)G''(u)^{-1}\nabla \frac{\delta}{\delta u} \mathcal{G}(u)) \\ &= \nabla \cdot (A(u)G''(u)^{-1}\nabla G'(u)) \\ &= \nabla \cdot (A(u)G''(u)^{-1}G''(u)\nabla u) \\ &= \nabla \cdot (A(u)\nabla u). \end{aligned}$$

□

*Remark 2.* An example of metric  $g$  is the Wasserstein-2 metric, which has been widely studied in optimal transport literature [1, 36]. In other words, let  $G(u) = u \log u - u$ ,  $A(u) = \mathbb{I}$ , then  $L_{\mathcal{C}}(u) = -\nabla \cdot (u\nabla)$ . In this case, the heat equation is the gradient flow of entropy functional  $\int_{\Omega} u \log u - u dx$  in Wasserstein-2 space. Recently, generalized optimal transport metrics and nonlinear diffusions have been widely studied in [9, 14, 31].

**3.3. Flux–gradient flows.** In summary, we illustrate the relation among entropy-entropy flux pairs, gradient flows and PDE (6). On the one hand, the entropy-entropy flux pairs introduce a Lyapunov functional, along which the entropy-entropy flux flow is non positive. In addition,  $\frac{dG}{dt} + \nabla \cdot \Psi \leq 0$ . This entropy condition [22] picks out the unique physical weak solution for inviscid conservation law (7). On the other hand, both Lyapunov functional and dissipative operator define a metric space, under which the dissipative operator forms the gradient flow. These facts imply a formulation for PDE (6). We call it *flux–gradient flows*.

In detail, equation (6) can be written below.

$$\partial_t u + \nabla \cdot f(u) = \beta \nabla \cdot (A(u)G''(u)^{-1} \nabla \frac{\delta}{\delta u} \mathcal{G}(u)), \quad (8)$$

where the flux function  $f$  satisfies

$$\int_{\Omega} f(u) \cdot \nabla \frac{\delta}{\delta u} \mathcal{G}(u) dx = 0.$$

The above formulation of the PDE (6) is a gradient flow equation added with the flux function. We can consider its general formulation. We keep the metric operator invariant and replace the Lyapunov functional  $\mathcal{G}(u)$  by a general energy functional  $\mathcal{E}(u)$ .

**Definition 5** (Flux–gradient flow). *Given an energy functional  $\mathcal{E}: \mathcal{M} \rightarrow \mathbb{R}$ , consider a class of PDE*

$$\partial_t u + \nabla \cdot (f_1(x, u)) = \beta \nabla \cdot (A(u)G''(u)^{-1} \nabla \frac{\delta}{\delta u} \mathcal{E}(u)), \quad (9)$$

where  $f_1: \Omega \times \mathbb{R}^1 \rightarrow \mathbb{R}^n$  is a flux function satisfying

$$\int_{\Omega} f_1(x, u) \cdot \nabla \frac{\delta}{\delta u(x)} \mathcal{E}(u) dx = 0. \quad (10)$$

If  $\mathcal{E}(u) = \mathcal{G}(u) = \int_{\Omega} G(u) dx$  and  $f_1(x, u) = f(u)$ , then equation (9) forms PDE (6).

Equations in formulation (9) provide a class of conservative–dissipative dynamics.

**Proposition 6** (Entropy-entropy flux-production). *Energy functional  $\mathcal{E}(u)$  is a Lyapunov functional for PDE (9). In other words, the following dissipation result holds. Suppose  $u(t, x)$  is the solution of equation (9), then*

$$\frac{d}{dt} \mathcal{E}(u(t, \cdot)) = -\beta \mathcal{I}_{\mathcal{E}}(u(t, \cdot)) \leq 0,$$

where the functional  $\mathcal{I}_{\mathcal{E}}: \mathcal{M} \rightarrow \mathbb{R}_+$  is defined by

$$\mathcal{I}_{\mathcal{E}}(u) = \int_{\Omega} \left( \nabla \frac{\delta}{\delta u} \mathcal{E}(u), A(u)G''(u)^{-1} \nabla \frac{\delta}{\delta u} \mathcal{E}(u) \right) dx. \quad (11)$$

*Proof.* The proof follows directly from Definition 5. In other words,

$$\begin{aligned} \frac{d}{dt} \mathcal{E}(u) &= \int_{\Omega} \frac{\delta}{\delta u} \mathcal{E}(u) \cdot \partial_t u dx \\ &= \int_{\Omega} \left[ -\frac{\delta}{\delta u} \mathcal{E}(u) \nabla \cdot f_1(x, u) + \beta \frac{\delta}{\delta u} \mathcal{E}(u) \nabla \cdot (A(u)G''(u)^{-1} \nabla \frac{\delta}{\delta u} \mathcal{E}(u)) \right] dx \\ &= \int_{\Omega} (f_1(x, u), \nabla \frac{\delta}{\delta u} \mathcal{E}(u)) dx - \beta \int_{\Omega} \left( \nabla \frac{\delta}{\delta u} \mathcal{E}(u), A(u)G''(u)^{-1} \nabla \frac{\delta}{\delta u} \mathcal{E}(u) \right) dx \\ &= -\beta \mathcal{I}_{\mathcal{E}}(u), \end{aligned}$$

where the third equality applies the integration by parts formula and the last equality uses condition (10).  $\square$

*Remark 3.* We remark that if  $\mathcal{E}(u) = \mathcal{G}(u) = \int_{\Omega} (u \log u - u) dx$  and  $A(u) = \mathbb{I}$ , then the functional

$$\mathcal{I}_{\mathcal{E}}(u) = \int_{\Omega} \|\nabla \log u\|^2 u dx = \int_{\Omega} \frac{\|\nabla u(x)\|^2}{u(x)} dx,$$

is known as the Fisher information functional. A known fact is that the dissipation of Lyapunov/entropy functional  $\mathcal{G}(u) = \int_{\Omega} u \log u - u dx$  along the heat flow equals to the negative Fisher information functional. This fact is called the de Bruijn equality. Here, the de-Bruijn type equalities hold naturally in conservation laws with entropy-entropy flux pairs. In future works, we shall study the dynamical effect of flux function in metric spaces; see related techniques developed in [26].

*Remark 4.* We remark that equation (9) are generalized variational formulations for flux-gradient flows in metric spaces. They have potential applications in designing Markov-Chain-Monte-Carlo algorithms, and deriving the neural network variational algorithms for conservation laws; see [28]. We leave the detailed studies of these areas in future works.

#### 4. CONTROLLING CONSERVATION LAWS

In this section, we present the main results of this paper. We study the variational problems for conservation laws (6). From now on, we assume that the entropy-entropy flux-metric condition in Definition 2 holds.

We first design an optimal control problem over flux-gradient flows in a metric space.

**Definition 7** (Optimal control of conservation laws). *Given smooth functionals  $\mathcal{F}, \mathcal{H}: \mathcal{M} \rightarrow \mathbb{R}$ , consider a variational problem*

$$\inf_{u, v, u_1} \int_0^1 \left[ \int_{\Omega} \frac{1}{2} (v, A(u) G''(u)^{-1} v) dx - \mathcal{F}(u) \right] dt + \mathcal{H}(u_1), \quad (12a)$$

where the infimum is taken among variables  $v: [0, 1] \times \Omega \rightarrow \mathbb{R}^n$ ,  $u: [0, 1] \times \Omega \rightarrow \mathbb{R}$ , and  $u_1: \Omega \rightarrow \mathbb{R}$  satisfying

$$\partial_t u + \nabla \cdot f(u) + \nabla \cdot (A(u) G''(u)^{-1} v) = \beta \nabla \cdot (A(u) \nabla u), \quad u(0, x) = u_0(x). \quad (12b)$$

Here the equation (12b) is a control dynamic for the conservation law (6). The objective functional is an enforced “kinetic energy” minus a “potential energy” in the metric space  $(\mathcal{M}, \mathbf{g})$ . If the control variable  $v(t, x) = 0$  for all  $t = [0, 1]$ ,  $x \in \Omega$ , then dynamics (12b) becomes the original conservation law equation (6).

*Remark 5.* We notice that variational problem (12) is a generalized dynamical optimal transport problem. In other words, if  $u: [0, 1] \times \Omega \rightarrow \mathbb{R}_+$ ,  $A(u) = \mathbb{I}$ ,  $G(u) = u \log u - u$ ,  $\beta = 0$ ,  $f = 0$ ,  $\mathcal{F}(u) = 0$  and  $u_1$  is a fixed function with  $\int_{\Omega} u_0 dx = \int_{\Omega} u_1 dx$ , then problem (12) forms

$$\inf_{u, v} \left\{ \int_0^1 \int_{\Omega} \frac{1}{2} \|v\|^2 u dx dt : \partial_t u + \nabla \cdot (uv) = 0, \quad u(0, x) = u_0(x), \quad u(1, x) = u_1(x) \right\}.$$

The above minimization is known as Benamou-Brenier’s formula [4] studied in classical optimal transport problems. Here the variational problem (12) generalizes the dynamical optimal transport, which contains the inverse of the Hessian operator of entropy functionals

to model the “kinetic energy”. In particular, we formulate the conservation laws in the constraint set.

We next derive critical point systems of variational problem (12). They are Hamiltonian flows in  $(\mathcal{M}, \mathbf{g})$  associated with conservation laws.

**Proposition 8** (Hamiltonian flows of conservation laws). *The critical point system of variational problem (12) is given below. There exists a function  $\Phi: [0, 1] \times \Omega \rightarrow \mathbb{R}$ , such that*

$$v(t, x) = \nabla \Phi(t, x),$$

and

$$\begin{cases} \partial_t u + \nabla \cdot f(u) + \nabla \cdot (A(u)G''(u)^{-1}\nabla \Phi) = \beta \nabla \cdot (A(u)\nabla u), \\ \partial_t \Phi + (\nabla \Phi, f'(u)) + \frac{1}{2}(\nabla \Phi, (A(u)G''(u)^{-1})'\nabla \Phi) + \frac{\delta}{\delta u} \mathcal{F}(u) = -\beta \nabla \cdot (A(u)\nabla \Phi) + \beta (\nabla \Phi, A'(u)\nabla u). \end{cases} \quad (13)$$

Here  $'$  represents the derivative w.r.t. variable  $u$ . The initial and terminal time conditions satisfy

$$u(0, x) = u_0(x), \quad \frac{\delta}{\delta u_1} \mathcal{H}(u_1) + \Phi(1, x) = 0.$$

*Proof.* We first rewrite the variables  $(u, v)$  in variational formula (12) by  $(u, m)$ , where

$$m(t, x) = A(u(t, x))G''(u(t, x))^{-1}v(t, x) = V(u(t, x))v(t, x).$$

Here we denote  $V(u(t, x)) = A(u(t, x))G''(u(t, x))^{-1}$ . In this case, the variational problem (12) forms

$$\inf_{m, u, u_1} \left\{ \int_0^1 \left[ \int_{\Omega} \frac{1}{2}(m, V(u)^{-1}m) dx - \mathcal{F}(u) \right] dt + \mathcal{H}(u_1) : \right. \\ \left. \partial_t u + \nabla \cdot f(u) + \nabla \cdot m = \beta \nabla \cdot (A(u)\nabla u), \quad \text{fixed } u_0. \right\}. \quad (14)$$

Denote the Lagrange multiplier of problem (14) by  $\Phi: [0, 1] \times \Omega \rightarrow \mathbb{R}$ . Consider the following saddle point problem

$$\inf_{m, u, u_1} \sup_{\Phi} \mathcal{L}(m, u, u_1, \Phi).$$

In the above formula, we have

$$\begin{aligned} \mathcal{L}(m, u, u_1, \Phi) &= \int_0^1 \int_{\Omega} \left[ \frac{1}{2}(m, V(u)^{-1}m) + \Phi \left( \partial_t u + \nabla \cdot f(u) + \nabla \cdot m - \beta \nabla \cdot (A(u)\nabla u) \right) \right] dx dt \\ &\quad - \int_0^1 \mathcal{F}(u) dt + \mathcal{H}(u_1) \\ &= \int_0^1 \int_{\Omega} \left[ \frac{1}{2}(m, V(u)^{-1}m) + \Phi \left( \nabla \cdot f(u) + \nabla \cdot m - \beta \nabla \cdot (A(u)\nabla u) \right) \right] dx dt \\ &\quad + \int_{\Omega} (\Phi(1, x)u_1(x) - \Phi(0, x)u_0(x)) dx - \int_0^1 \int_{\Omega} \partial_t \Phi u dx dt - \int_0^1 \mathcal{F}(u) dt + \mathcal{H}(u_1), \end{aligned}$$

where we use the integration by parts formula w.r.t.  $t$  in the second equality. In other words,

$$\int_0^1 \int_{\Omega} \Phi \partial_t u dx dt = \int_{\Omega} \Phi(1, x) u(1, x) dx - \int_{\Omega} \Phi(0, x) u(0, x) dx - \int_0^1 \int_{\Omega} \partial_t \Phi u dx dt.$$

We next derive the critical point for the above saddle point problem. In other words, consider

$$\begin{cases} \frac{\delta}{\delta m} \mathcal{L} = 0 \\ \frac{\delta}{\delta u} \mathcal{L} = 0 \\ \frac{\delta}{\delta u_1} \mathcal{L} = 0 \\ \frac{\delta}{\delta \Phi} \mathcal{L} = 0 \end{cases} \Rightarrow \begin{cases} V(u)^{-1} m = \nabla \Phi, \\ -\frac{1}{2} (m, V(u)^{-1} V(u)' V(u)^{-1} m) - \frac{\delta}{\delta u} \mathcal{F} - \partial_t \Phi - (\nabla \Phi, f'(u)) \\ \quad - \beta \nabla \cdot (A(u) \nabla \Phi) + \beta (\nabla \Phi, A'(u) \nabla u) = 0, \\ \Phi_1 + \frac{\delta}{\delta u_1} \mathcal{H}(u_1) = 0, \\ \partial_t u + \nabla \cdot f(u) + \nabla \cdot m - \beta \nabla \cdot (A(u) \nabla u) = 0, \end{cases}$$

where  $\frac{\delta}{\delta m}$ ,  $\frac{\delta}{\delta u}$ ,  $\frac{\delta}{\delta u_1}$ ,  $\frac{\delta}{\delta \Phi}$  are  $L^2$  first variational derivatives w.r.t. functions  $m$ ,  $u$ ,  $u_1$ ,  $\Phi$ , respectively. We thus derive the pair of PDEs (13) in  $\mathcal{M}(\Omega)$ .  $\square$

We next present the Hamiltonian formalism for the PDE system (13).

**Proposition 9** (Hamiltonian flows in metric space). *PDE system (13) has the following Hamiltonian flow formulation.*

$$\partial_t u = \frac{\delta}{\delta \Phi} \mathcal{H}_{\mathcal{G}}(u, \Phi), \quad \partial_t \Phi = -\frac{\delta}{\delta u} \mathcal{H}_{\mathcal{G}}(u, \Phi),$$

where we define the Hamiltonian functional  $\mathcal{H}_{\mathcal{G}}: \mathcal{M} \times C^\infty(\Omega) \rightarrow \mathbb{R}$  by

$$\mathcal{H}_{\mathcal{G}}(u, \Phi) = \int_{\Omega} \left[ \frac{1}{2} (\nabla \Phi, A(u) G''(u)^{-1} \nabla \Phi) + (\nabla \Phi, f(u)) - \beta (\nabla \Phi, A(u) \nabla u) \right] dx + \mathcal{F}(u). \quad (15)$$

In other words, the Hamiltonian functional  $\mathcal{H}_{\mathcal{G}}(u, \Phi)$  is conserved along dynamics (13). Suppose  $(u, \Phi)$  are solutions of equation (13), then

$$\frac{d}{dt} \mathcal{H}_{\mathcal{G}}(u, \Phi) = 0.$$

**Proposition 10** (Functional Hamilton-Jacobi equations of conservation laws). *The Hamilton-Jacobi equation in  $(\mathcal{M}, \mathbf{g})$  for equation (13) satisfies*

$$\begin{aligned} \partial_t \mathcal{U}(t, u) + \int_{\Omega} \left[ \frac{1}{2} \left( \nabla \frac{\delta}{\delta u(x)} \mathcal{U}(t, u), A(u) G''(u)^{-1} \nabla \frac{\delta}{\delta u(x)} \mathcal{U}(t, u) \right) + \left( \nabla \frac{\delta}{\delta u(x)} \mathcal{U}(t, u), f(u) \right) \right. \\ \left. - \beta \left( \nabla \frac{\delta}{\delta u(x)} \mathcal{U}(t, u), A(u) \nabla u \right) \right] dx + \mathcal{F}(u) = 0, \end{aligned}$$

where  $\mathcal{U}: [0, 1] \times L^2(\Omega) \rightarrow \mathbb{R}$  is a value functional.

*Proof of Proposition 9 and Proposition 10.* The proof follows from the definition of Hamiltonian dynamics. We can check it directly by using the first variation operators. In other words,

$$\frac{\delta}{\delta \Phi} \mathcal{H}_{\mathcal{G}}(u, \Phi) = -\nabla \cdot (A(u) G''(u)^{-1} \nabla \Phi) - \nabla \cdot f(u) + \beta \nabla \cdot (A(u) \nabla u),$$

and

$$\begin{aligned} \frac{\delta}{\delta u} \mathcal{H}_{\mathcal{G}}(u, \Phi) &= \frac{1}{2} (\nabla \Phi, (A(u)G''(u)^{-1})' \nabla \Phi) + (\nabla \Phi, f'(u)) + \frac{\delta}{\delta u} \mathcal{F}(u) \\ &\quad - \beta (\nabla \Phi, A'(u) \nabla u) + \beta \nabla \cdot (A(u) \nabla \Phi). \end{aligned}$$

Clearly,  $\mathcal{H}_{\mathcal{G}}(u, \Phi)$  is conserved since

$$\begin{aligned} \frac{d}{dt} \mathcal{H}_{\mathcal{G}}(u, \Phi) &= \int_{\Omega} \left( \frac{\delta}{\delta u} \mathcal{H}_{\mathcal{G}} \cdot \partial_t u + \frac{\delta}{\delta \Phi} \mathcal{H}_{\mathcal{G}} \cdot \partial_t \Phi \right) dx \\ &= \int_{\Omega} \left( \frac{\delta}{\delta u} \mathcal{H}_{\mathcal{G}} \cdot \frac{\delta}{\delta \Phi} \mathcal{H}_{\mathcal{G}} - \frac{\delta}{\delta \Phi} \mathcal{H}_{\mathcal{G}} \cdot \frac{\delta}{\delta u} \mathcal{H}_{\mathcal{G}} \right) dx \\ &= 0. \end{aligned}$$

This finishes the proof of Proposition 9.

We next use the Hamiltonian functional to formulate the Hamilton-Jacobi equation in  $(\mathcal{M}, \mathbf{g})$ . Define a functional by  $\mathcal{U}: \mathbb{R}_+ \times \mathcal{M} \rightarrow \mathbb{R}$ . Let  $\Phi(t, x) = \frac{\delta}{\delta u(x)} \mathcal{U}(t, u)$  in Hamiltonian flow (13). We obtain the Hamilton-Jacobi equation in metric space  $(\mathcal{M}, \mathbf{g})$ , where

$$\partial_t \mathcal{U}(t, u) + \mathcal{H}_{\mathcal{G}}(u, \frac{\delta}{\delta u} \mathcal{U}(t, u)) = 0.$$

This finishes the proof of Proposition 10.  $\square$

*Remark 6* (Fisher information regularizations). We remark that the Fisher information functional is also connected with control problems of conservation laws. See similar studies in [21]. We leave the study about information functional regularizations of conservation laws in a sequential work.

**4.1. Examples.** In this subsection, we list several examples of control problems for scalar conservation laws.

**Example 1** (Controlling heat equations). *Consider the heat equation*

$$\partial_t u = \beta \Delta u,$$

where  $u: [0, +\infty) \times \Omega \rightarrow \mathbb{R}_+$  is the probability density function. It satisfies PDE (6), where

$$u \in \mathbb{R}_+^1, \quad f(u) = 0, \quad A(u) = \mathbb{I}.$$

In this case, a function pair  $(G, \Psi)$ , where  $G$  is a convex function and  $\Psi = 0$ , satisfies the entropy-entropy flux condition. In other words,

$$\Psi'(u) = f'(u)G'(u) = 0.$$

In particular, if  $G(u) = u \log u - u$ , then the variational problem (12) satisfies

$$\inf_{u, v, u_1} \int_0^1 \left[ \int_{\Omega} \frac{1}{2} |v|^2 u dx - \mathcal{F}(u) \right] dt + \mathcal{H}(u_1),$$

where the infimum is taken among variables  $u, v, u_1$  satisfying

$$\partial_t u + \nabla \cdot (uv) = \beta \Delta u, \quad u(0, x) = u_0(x).$$

Here the minimizer system is given below. There exists a function  $\Phi$ , such that  $v = \nabla\Phi$  and

$$\begin{cases} \partial_t u + \nabla \cdot (u \nabla \Phi) = \beta \Delta u, \\ \partial_t \Phi + \frac{1}{2} \|\nabla \Phi\|^2 + \frac{\delta}{\delta u} \mathcal{F}(u) = -\beta \Delta \Phi. \end{cases}$$

The above optimal control problem and critical point system has been widely studied in the optimal transport ( $\mathcal{F} = 0$ ,  $\beta = 0$ ), Schrödinger bridge problems ( $\mathcal{F} = 0$ ,  $\beta > 0$ ) and potential mean field games. There is a Hamiltonian formalism for the PDE system  $(u, \Phi)$ . In other words,

$$\partial_t u = \frac{\delta}{\delta \Phi} \mathcal{H}_G(u, \Phi), \quad \partial_t \Phi = -\frac{\delta}{\delta u} \mathcal{H}_G(u, \Phi),$$

where the Hamiltonian functional satisfies

$$\mathcal{H}_G(u, \Phi) = \int_{\Omega} \left[ \frac{1}{2} \|\nabla \Phi\|^2 u - \beta (\nabla \Phi, \nabla u) \right] dx + \mathcal{F}(u).$$

**Example 2** (Controlling scalar conservation laws). Consider

$$\partial_t u + \nabla \cdot f(u) = \beta \Delta u,$$

where  $u: [0, +\infty) \times \Omega \rightarrow \mathbb{R}$  and  $A = \mathbb{I}$ .

(i) Let  $G(u) = u \log u - u$ . Then variational problem (12) satisfies

$$\inf_{u, v, u_1} \int_0^1 \left[ \int_{\Omega} \frac{1}{2} |v|^2 u dx - \mathcal{F}(u) \right] dt + \mathcal{H}(u_1),$$

where the infimum is taken among variables  $u, v, u_1$  satisfying

$$\partial_t u + \nabla \cdot f(u) + \nabla \cdot (uv) = \beta \Delta u, \quad u(0, x) = u_0(x).$$

Here the minimizer system is given below. There exists a function  $\Phi$ , such that  $v = \nabla\Phi$  and

$$\begin{cases} \partial_t u + \nabla \cdot f(u) + \nabla \cdot (u \nabla \Phi) = \beta \Delta u, \\ \partial_t \Phi + (\nabla \Phi, f'(u)) + \frac{1}{2} \|\nabla \Phi\|^2 + \frac{\delta}{\delta u} \mathcal{F}(u) = -\beta \Delta \Phi. \end{cases}$$

I.e.,

$$\partial_t u = \frac{\delta}{\delta \Phi} \mathcal{H}_G(u, \Phi), \quad \partial_t \Phi = -\frac{\delta}{\delta u} \mathcal{H}_G(u, \Phi),$$

where the Hamiltonian functional satisfies

$$\mathcal{H}_G(u, \Phi) = \int_{\Omega} \left[ \frac{1}{2} \|\nabla \Phi\|^2 u + (\nabla \Phi, f(u)) - \beta (\nabla \Phi, \nabla u) \right] dx + \mathcal{F}(u).$$

(ii) Let  $G(u) = \frac{u^2}{2}$ . Then variational problem (12) satisfies

$$\inf_{u, v, u_1} \int_0^1 \left[ \int_{\Omega} \frac{1}{2} |v|^2 dx - \mathcal{F}(u) \right] dt + \mathcal{H}(u_1),$$

where the infimum is taken among variables  $u, v, u_1$  satisfying

$$\partial_t u + \nabla \cdot f(u) + \nabla \cdot v = \beta \Delta u, \quad u(0, x) = u_0(x).$$

Here the minimizer system is given below. There exists a function  $\Phi$ , such that  $v = \nabla\Phi$  and

$$\begin{cases} \partial_t u + \nabla \cdot f(u) + \nabla \cdot (\nabla\Phi) = \beta\Delta u, \\ \partial_t \Phi + (\nabla\Phi, f'(u)) + \frac{\delta}{\delta u} \mathcal{F}(u) = -\beta\Delta\Phi. \end{cases}$$

I.e.,

$$\partial_t u = \frac{\delta}{\delta\Phi} \mathcal{H}_{\mathcal{G}}(u, \Phi), \quad \partial_t \Phi = -\frac{\delta}{\delta u} \mathcal{H}_{\mathcal{G}}(u, \Phi),$$

where the Hamiltonian functional satisfies

$$\mathcal{H}_{\mathcal{G}}(u, \Phi) = \int_{\Omega} \left[ \frac{1}{2} \|\nabla\Phi\|^2 + (\nabla\Phi, f(u)) - \beta(\nabla\Phi, \nabla u) \right] dx + \mathcal{F}(u).$$

## 5. NUMERICAL METHODS AND EXAMPLES

In this section, we first review classical primal-dual hybrid gradient algorithms (PDHG) and their extensions. We next apply this approach to solve the variational problem defined in equation (12) subject to the constraint involving conservation laws. We design a finite difference scheme to discretize conservation laws and solve the variational problem on grids. Several numerical examples, including Burgers' equation and traffic flow, are presented.

**5.1. PDHG algorithm and its extension.** The PDHG algorithm [10] solves the following constrained convex optimization problem

$$\min_z h(Kz) + g(z),$$

where  $\mathcal{Z}$  is a finite or infinite dimensional Hilbert space,  $h$  and  $g$  are convex functions and  $K : \mathcal{Z} \rightarrow \mathcal{H}$  is a linear operator between Hilbert spaces. This problem can be rewritten in the saddle-point problem form

$$\min_z \max_p \langle Kz, p \rangle_{L^2} + g(z) - h^*(p),$$

where  $h^*(p) = \sup_z \langle Kz, p \rangle_{L^2} - h(z)$ , which is the convex conjugate of  $h$ . The algorithm updates  $z, p$  by taking proximal gradient descent, proximal gradient ascent steps alternatively. At the  $n$ -th iteration, the algorithm updates as follows

$$\begin{aligned} z^{n+1} &= \arg \min_z \langle Kz, \bar{p}^n \rangle_{L^2} + g(z) + \frac{1}{2\tau} \|z - z^n\|_{L^2}^2, \\ p^{n+1} &= \arg \max_p \langle Kz^{n+1}, p \rangle_{L^2} - h^*(p) - \frac{1}{2\sigma} \|p - p^n\|_{L^2}^2, \\ \bar{p}^{n+1} &= 2p^{n+1} - p^n. \end{aligned}$$

Here  $\tau, \sigma$  are stepsizes, which have to satisfy  $\sigma\tau\|K^TK\| < 1$  in order to guarantee the convergence of the algorithm. When the operator  $K$  is nonlinear, we use the extension of PDHG algorithm [13]. The idea is to use the linear approximation of  $K$ :

$$K(z) \approx K(\bar{z}) + \nabla K(\bar{z})(z - \bar{z}).$$



Here, the extension of the PDHG scheme is as follows

$$\begin{aligned}
z^{n+1} &= \arg \min_z \langle z, [\nabla K(z^n)]^T \bar{p}^n \rangle_{L^2} + g(x) + \frac{1}{2\tau} \|z - z^n\|_{L^2}^2, \\
p^{n+1} &= \arg \max_p \langle K(z^{n+1}), p \rangle_{L^2} - h^*(p) - \frac{1}{2\sigma} \|p - p^n\|_{L^2}^2, \\
\bar{p}^{n+1} &= 2p^{n+1} - p^n.
\end{aligned} \tag{16}$$

When  $K$  is some unbounded linear operator, for instance  $K = \nabla$ , the operator norm  $\|K\|$  can increase when we refine the grid size. Consequently, the algorithm may converge slowly due to small stepsizes. We apply a generalization of PDHG, namely the General-proximal Primal-Dual Hybrid Gradient (G-prox PDHG) method from [20]. We choose a proper norm ( $L^2, H^1, \dots$ ) for the proximal step to allow larger stepsizes.

We apply the nonlinear G-prox PDHG algorithm to solve the variational problem (12), in particular its equivalent format (14). For illustration purpose, we use the traffic flow variational problem (5) as an example and give details on the algorithm. Set

$$\begin{aligned}
z &= (u, m), \\
p &= \Phi, \\
K((u, m)) &= \partial_t u + \nabla \cdot f(u) + \nabla \cdot m - \beta \Delta u, \\
g((u, m)) &= \int_0^1 \left( \int_{\Omega} \frac{\|m\|^2}{2u} + \mathbf{1}_{[0,1]}(u) dx - \mathcal{F}(u) \right) dt, \\
h(Kz) &= \begin{cases} 0 & \text{if } Kz = 0 \\ +\infty & \text{else} \end{cases}.
\end{aligned}$$

We rewrite the variational problem as follows:

$$\inf_{u, m} \sup_{\Phi} \mathcal{L}(u, m, \Phi), \quad \text{with } u(0, x) = u_0(x), \quad \Phi(1, x) = -\frac{\delta}{\delta u(1, x)} \mathcal{H}(u), \tag{17}$$

where

$$\begin{aligned}
\mathcal{L}(u, m, \Phi) &= \int_0^1 \left( \int_{\Omega} \frac{\|m\|^2}{2u} + \mathbf{1}_{[0,1]}(u) dx - \mathcal{F}(u) \right) dt \\
&\quad + \int_0^1 \int_{\Omega} \Phi (\partial_t u + \nabla \cdot f(u) + \nabla \cdot m - \beta \Delta u) dx dt,
\end{aligned} \tag{18}$$

and  $f(u) = u(1 - u)$  is the traffic flux function. We denote the indicator function by

$$\mathbf{1}_A(x) = \begin{cases} +\infty & x \notin A \\ 0 & x \in A. \end{cases}$$

We also choose  $L^2$  norm for updating  $(u, m)$  and  $H^1$  norm for  $\Phi$ , specifically

$$\|v\|_{L^2}^2 = \int_0^1 \int_{\Omega} v^2 dx dt, \quad \|v\|_{H_1^1}^2 = \|\nabla v\|_{L^2}^2 + \|\partial_t v\|_{L^2}^2.$$

We summarize the algorithm to solve problem (17) as follows.

---

**Algorithm 1:PDHG for the conservation law control system**


---

**While**  $k < \text{Maximal number of iteration}$

$$(u^{(k+1)}, m^{(k+1)}) = \operatorname{argmin}_u \mathcal{L}(u, m, \bar{\Phi}^{(k)}) + \frac{1}{2\tau} \|u - u^{(k)}\|_{L^2}^2 + \frac{1}{2\tau} \|m - m^{(k)}\|_{L^2}^2;$$

$$\Phi^{(k+1)} = \operatorname{argmax}_{\Phi} \mathcal{L}(u^{(k+1)}, m^{(k+1)}, \Phi) - \frac{1}{2\sigma} \|\Phi - \Phi^{(k)}\|_{H^1}^2;$$

$$\bar{\Phi}^{(k+1)} = 2\Phi^{(k+1)} - \Phi^{(k)};$$


---

**5.2. Finite Difference Discretization.** In this subsection, we review basic numerical concepts of conservation law. More details can be found in [25]. Then we design the finite difference discretization for the control of conservation laws.

**5.2.1. Lax–Friedrichs scheme for the conservation law.** Consider a nonlinear scalar conservation law

$$\begin{cases} \partial_t u + \partial_x f(u) = \beta \partial_{xx} u, \\ u(0, x) = u_0(x). \end{cases} \quad (19)$$

We review the Lax–Friedrichs scheme. Denote a discretization  $\Delta t, \Delta x$  in time and space,  $u_j^k = u(k\Delta t, j\Delta x)$ , then the update follows

$$u_j^{k+1} = u_j^k - \frac{\Delta t}{2\Delta x} \left( f(u_{j+1}^k) - f(u_{j-1}^k) \right) + (\beta + c\Delta x) \frac{\Delta t}{(\Delta x)^2} \left( u_{j+1}^k - 2u_j^k + u_{j-1}^k \right). \quad (20)$$

The last term contains diffusion coefficient  $\beta$  from the original equation and  $c\Delta x$  as coefficient for artificial viscosity with  $c > 0$ . Notice that the monotone scheme gives entropy solutions. Here the definition of monotone scheme is given below.

**Definition 11.** For  $p, q \in \mathbb{N}$ , a scheme

$$u_j^{k+1} = G(u_{j-p-1}^k, \dots, u_{j+q}^k)$$

is called a monotone scheme if  $G$  is a monotonically nondecreasing function of each argument.

In order to guarantee that the scheme (20) is monotone, the following inequalities have to be satisfied:

$$\begin{aligned} 1 - 2(\beta + c\Delta x) \frac{\Delta t}{(\Delta x)^2} &\geq 0, \\ -\frac{\Delta t}{2\Delta x} |f'(u)| + (\beta + c\Delta x) \frac{\Delta t}{(\Delta x)^2} &\geq 0. \end{aligned}$$

As we want the scheme works when  $\beta \rightarrow 0$ , the restriction on  $c$  and space–time stepsizes can be simplified as follows:

$$\begin{aligned} c &\geq \frac{1}{2} |f'(u)|, \\ (\Delta x)^2 &\geq 2(\beta + c\Delta x) \Delta t. \end{aligned}$$

The first inequality suggests the artificial viscosity we need to add. The second one impose a strong restriction on the stepsize in time when  $\beta > 0$ .

**5.2.2. Discretization of the control problem.** We consider the control problem of scalar conservation law defined in  $[0, b] \times [0, 1]$ , where  $b$  is a given constant. We apply the periodic boundary condition on the spatial domain. Given  $N_x, N_t > 0$ , we have  $\Delta x = \frac{b}{N_x}$ ,  $\Delta t = \frac{1}{N_t}$ . For  $x_i = i\Delta x, t_l = l\Delta t$ , define

$$\begin{aligned} u_i^l &= u(t_l, x_i) & 1 \leq i \leq N_x, 0 \leq l \leq N_t, \\ m_{1,i}^l &= (m_{x_1}(t_l, x_i))^+ & 1 \leq i \leq N_x, 0 \leq l \leq N_t - 1, \\ m_{2,i}^l &= -(m_{x_1}(t_l, x_i))^- & 1 \leq i \leq N_x, 0 \leq l \leq N_t - 1, \\ \Phi_i^l &= \Phi(t_l, x_i) & 1 \leq i \leq N_x, 0 \leq l \leq N_t, \\ \Phi_i^{N_t} &= -\frac{\delta}{\delta u(1, x_i)} \mathcal{H}(u_i^{N_t}) & 1 \leq i \leq N_x, \end{aligned}$$

where  $u^+ := \max(u, 0)$  and  $u^- = u^+ - u$ . Note here  $m_{1,i}^l \in \mathbb{R}_+, m_{2,i}^l \in \mathbb{R}_-$ . Denote

$$\begin{aligned} (Du)_i &:= \frac{u_{i+1} - u_i}{\Delta x} \\ [Du]_i &:= ((Du)_i, (Du)_{i-1}) \\ [\widehat{Du}]_i &:= ((Du)_i^+, -(Du)_{i-1}^-) \\ Lap(u)_i &= \frac{u_{i+1} - 2u_i + u_{i-1}}{(\Delta x)^2} \end{aligned}$$

The first conservation law equation adapted from the Lax–Friedrichs scheme is as follows:

$$\frac{1}{\Delta t} (u_i^{l+1} - u_i^l) + \frac{1}{2\Delta x} (f(u_{i+1}^{l+1}) - f(u_{i-1}^{l+1})) + (Dm)_{1,i-1}^l + (Dm)_{2,i}^l = (\beta + c\Delta x) Lap(u)_i^{l+1}, \quad (21)$$

where  $1 \leq i \leq N_x, 0 \leq l \leq N_t - 1$ . We choose  $c \geq \frac{1}{2} \max_u |f'(u)|$ . Unlike the Lax–Friedrichs scheme in explicit form, we use an implicit discretization in time to encode the feedback control structure. Another benefit of the implicit scheme is that it allows a larger mesh size in time.

Following the discretization of the conservation law, the discrete saddle point problem has the following form:

$$\min_{u, m} \max_{\Phi} L(u, m, \Phi),$$

where

$$\begin{aligned}
L(u, m, \Phi) = & \Delta x \Delta t \sum_{\substack{1 \leq i \leq N_x \\ 1 \leq l \leq N_t}} \frac{(m_{1,i}^{l-1})^2 + (m_{2,i}^{l-1})^2}{2u_i^l} - \Delta t \sum_{1 \leq l \leq N_t} \mathcal{F}(u^l) \\
& + \Delta x \sum_{1 \leq i \leq N_x} \mathcal{H}(u_i^{N_t}) + \sum_{\substack{1 \leq i \leq N_x \\ 1 \leq l \leq N_t}} \mathbf{1}_{[0,1]}(u_i^l) \\
& + \Delta x \Delta t \sum_{\substack{1 \leq i \leq N_x \\ 0 \leq l \leq N_t - 1}} \Phi_i^l \left( \frac{1}{\Delta t} (u_i^{l+1} - u_i^l) + \frac{1}{2\Delta x} (f(u_{i+1}^{l+1}) - f(u_{i-1}^{l+1})) \right) \\
& \quad + (Dm)_{1,i-1}^l + (Dm)_{2,i}^l - (\beta + c\Delta x) \text{Lap}(u_i^{l+1}).
\end{aligned}$$

By taking the first order derivative of  $u_i^l$ , we automatically get the implicit finite difference scheme for the dual equation of  $\Phi$  that is backward in time. The positive and negative parts of  $(D\phi)_i^l$  are split, which help enhance the monotonicity of the discrete Hamiltonian.

$$\frac{1}{\Delta t} (\Phi_i^{l+1} - \Phi_i^l) + \frac{(\Phi_{i+1}^l - \Phi_{i-1}^l)}{2\Delta x} (f'(u_i^l)) + \frac{1}{2} \|\widehat{D\Phi}_i^l\|^2 + \frac{\delta \mathcal{F}(u_i^l)}{\delta u} = -(\beta + c\Delta x) \text{Lap}(\Phi)_i^l, \tag{22}$$

for  $1 \leq i \leq N_x$ ,  $1 \leq l \leq N_t$ .

We remark that the above discretizations, when  $f(u) = 0$ , reduce to the finite difference scheme for the mean-field game system proposed in [7]. The discrete form of the Hamiltonian functional (15) at  $t = t_l$  takes the form

$$H_G(u, \Phi) = \sum_{1 \leq i \leq N_x} \left( \frac{1}{2} \|\widehat{D\Phi}_i^l\|^2 u_i^l + \frac{(\Phi_{i+1}^l - \Phi_{i-1}^l)}{2\Delta x} (f(u_i^l)) + (\beta + c\Delta x) u_i^l \text{Lap}(\Phi)_i^l \right) - \mathcal{F}(u^l).$$

We shall verify the conservation of Hamiltonian functional with numerical examples.

**5.2.3. Solve conservation laws via primal-dual algorithms.** When  $\mathcal{F} = 0$ ,  $\mathcal{H} = c$  for some constant  $c$ , the variational problem (17) becomes classical conservation laws with initial data. In this case, no control will be enforced on the density function  $u$ . Therefore, the density movement is only determined by the flux term  $(f(u))_x$ . This means that the problem is reduced to initial value conservation laws. In this scenario, our approach proposed in Algorithm 1 provides an alternative way to solve the nonlinear conservation law with implicit discretization in time. In the implementation, we observed that this method allows a larger mesh size in time, thanks to the primal-dual variational structure. The method is also highly parallelizable as the proximal gradient descent step for  $(u, m)$  is point-wise operation for each  $(l, i)$ . We demonstrate this part with two examples in the next section.

**5.3. Numerical examples.** In this subsection, we present numerical examples for control of conservation laws in one dimensional space.

5.4. **Example 1.** We consider the Burgers' equation on  $(t, x) \in [0, 1] \times [0, 4]$ .

$$\partial_t u + \partial_x f(u) = 0, \quad u(0, x) = \begin{cases} 1 & 2 \leq x \leq 3 \\ 0 & \text{else} \end{cases},$$

where  $f(u) = \frac{1}{2}u^2$ ,  $\mathcal{F} = 0$ ,  $\mathcal{H} = 0$ ,  $\beta = 0$ ,  $k = 0.5$ . The entropy solution to this problem at  $t = 1$  satisfies

$$u(1, x) = \begin{cases} 1 & 3 \leq x \leq 3.5 \\ (x - 2) & 2 < x \leq 3 \\ 0 & \text{else} \end{cases}.$$

In the following examples, we use the same spatial time domain with  $N_t = 50$ ,  $N_x = 100$  for the finite difference scheme. We solve the Burgers' equation using two approaches: solve the control problem; use the forward Lax–Friedrichs scheme. Figure 2 shows the solutions to the Burgers' equation. We can see that both numerical solutions are consistent with the exact entropy solution despite some numerical diffusions. From two plots on the right, there are clear formations of rarefaction wave and shock. We have verified that the shock travels at speed  $v = \frac{1}{2}$ .

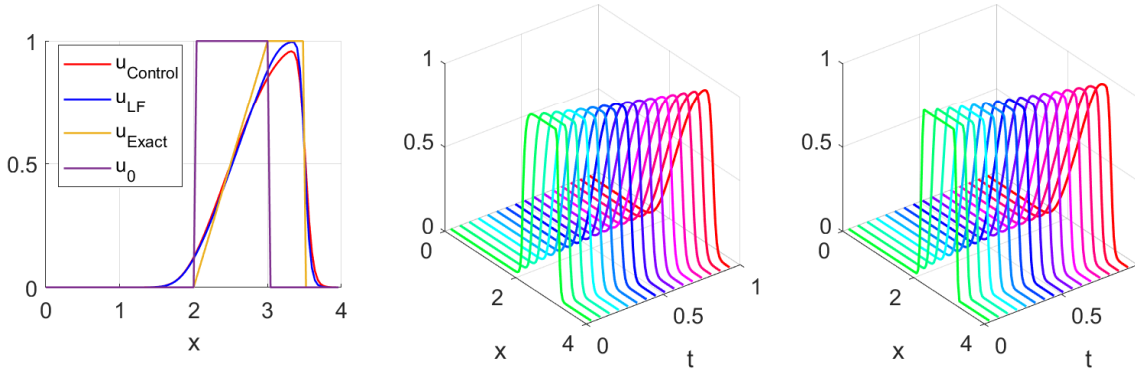


FIGURE 2. Numerical results for the Burgers' equation. Left: a comparison with the exact solution at  $t = 1$ ; middle: the numerical solution via solving the control problem; right: the numerical solution to the conservation law using Lax–Friedrichs scheme.

5.5. **Example 2.** We consider the traffic flow equation

$$\partial_t u + \partial_x f(u) = 0, \quad u(0, x) = \begin{cases} 0.8, & 1 \leq x \leq 2 \\ 0 & \text{else} \end{cases},$$

where  $f(u) = \frac{1}{2}u(1 - u)$ ,  $\mathcal{F} = 0$ ,  $\mathcal{H} = 0$ ,  $\beta = 0$ ,  $k = 0.5$ . The entropy solution to this problem at time  $t = 1$  is

$$u(1, x) = \begin{cases} 1 & 1.2 \leq x \leq 1.4 \\ \frac{1}{2}(3 - x) & 1.4 < x \leq 3 \\ 0 & \text{else} \end{cases}.$$

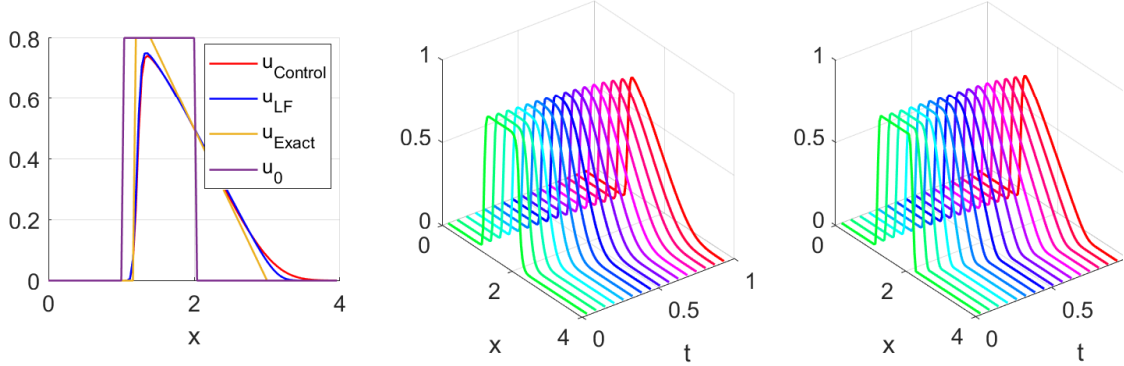


FIGURE 3. Numerical results for the traffic flow equation. Left: a comparison with the exact solution at  $t = 1$ ; middle: the numerical solution via solving the control problem; right: the numerical solution to the conservation law using Lax–Friedrichs scheme.

We can see from Figure 3 (left) that the solution to the control problem gives the same solution to the conservation law, which matches the entropy solution. In the two plots on the right, we observe that both shocks and rarefaction waves are formed. This traffic flow example describes a group of cars  $u_0$  waiting at the traffic light at  $x = 2$ . At the time  $t = 0$ , the red light turns green. But this group of cars doesn't move at a uniform constant speed. Instead, the car, whose originally position at time  $t = 0$  is closer to the red light ( $x = 0$ ), moves faster. The density at position  $x < 0$  doesn't change until time  $t = \frac{1}{2}(3 - x)$ .

**5.6. Example 3.** We again consider the traffic flow equation with  $f(u) = u(1 - u)$ ,  $\beta = 0.1$ ,  $\mathcal{F} = 0$ ,  $c = 0.5$ . The final cost functional  $\mathcal{H}(u(1, \cdot)) = \mu \int_{\Omega} u(1, x) \log\left(\frac{u(1, x)}{u_1}\right) dx$ ,  $\mu > 0$ . We set  $\mu = 1$ . In this case, the density  $u(1, \cdot)$  will tend to form a ‘similar’ distribution as  $u_1$ . When  $\int_{\Omega} u_1 dx = \int_{\Omega} u_0 dx$  and  $\mu \rightarrow +\infty$ , this final cost functional is equivalent to imposes the constraint that  $u(1, \cdot) = u_1$ . We also compare the result from the control of conservation law with a mean-field game problem, i.e.,  $f = 0$ :

$$\begin{cases} \partial_t u + \nabla \cdot (u \nabla \Phi) = \beta \Delta u, \\ \partial_t \Phi + \frac{1}{2} \|\nabla \Phi\|^2 + \frac{\delta}{\delta u} \mathcal{F}(u) = -\beta \Delta \Phi, \\ u(0, x) = u_0(x), \quad \Phi(1, x) = -\frac{\delta}{\delta u_1} \mathcal{H}(u(1, \cdot)). \end{cases} \quad (23)$$

As shown in Figure 4, we set

$$\begin{aligned} u_0 &= 0.001 + 0.9e^{-10(x-2)^2}, \\ u_1 &= 0.001 + 0.45e^{-10(x-1)^2} + 0.45e^{-10(x-3)^2}. \end{aligned}$$

The solutions are presented in Figure 5. We can see that in the mean-field game setup, there is an even split of the density. While for the control of conservation law, more density travels towards the location of the right-side Gaussian distribution. This demonstrates the difference between solutions in mean-field games and the ones in control of conservation laws.

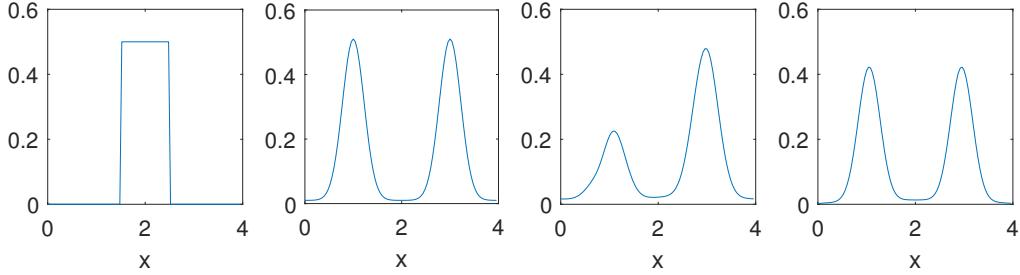


FIGURE 4. From left to right: initial configurations of  $u_0 = 0.001 + 0.9e^{-10(x-2)^2}$ ,  $u_1 = 0.001 + 0.45e^{-10(x-1)^2} + 0.45e^{-10(x-3)^2}$ , solution  $u(1, x)$  for the control of conservation law, solution  $u(1, x)$  for the mean-field game problem.

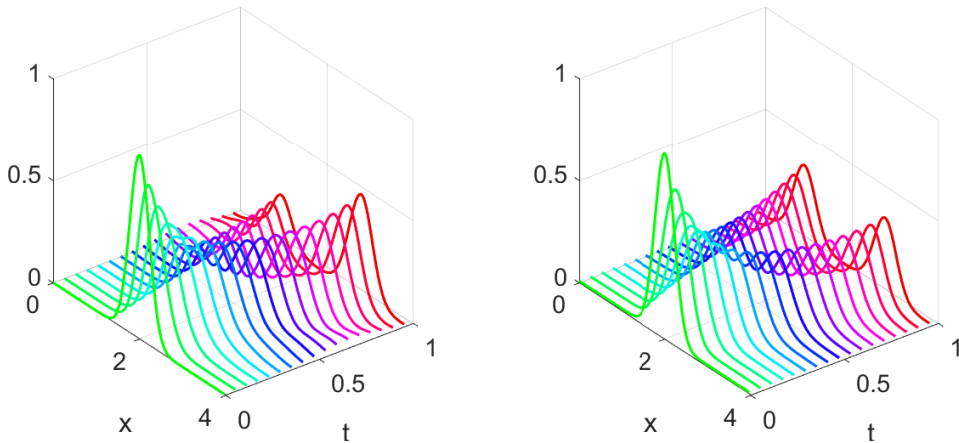


FIGURE 5. Left: solution  $u(t, x)$  for the problem of controlling the conservation law; right: solution  $u(t, x)$  for the mean-field game problem.

**5.7. Example 4.** Consider the traffic flow equation with  $f(u) = u(1 - u)$ ,  $\beta = 10^{-3}$ ,  $\mathcal{F} = -\alpha \int_{\Omega} u \log(u) dx$ ,  $\alpha \geq 0$ . The final cost functional  $\mathcal{H}(u(1, \cdot)) = \int_{\Omega} u(1, x)g(x)dx$ . The initial density and final cost function are as follows

$$u_0(x) = \begin{cases} 0.4 & 0.5 \leq x \leq 1.5 \\ 10^{-3} & \text{else} \end{cases}, \quad g(x) = -0.1 \sin(2\pi x).$$

The term  $\mathcal{F} = -\alpha \int_{\Omega} u \log(u) dx$ ,  $\alpha \geq 0$  penalizes the density for getting too concentrate. Figure 6 shows the effect of the term  $\alpha u \log(u)$ , where the density is more spreading in space for the case  $\alpha = 1$  than the case  $\alpha = 0.5$  and  $\alpha = 0$ . We can also see from the  $u$  profile at the terminal time. In Figure 7 (middle),  $\alpha = 0$  case has  $u(1, \cdot)$  has the most concentrated densities than others. We also numerically verify that the Hamiltonian functional is preserved over time in this example. In Figure 7 (right), the numerical Hamiltonian  $H_G(u, \Phi)$  is preserved with error of order of  $O(\Delta x)$ .

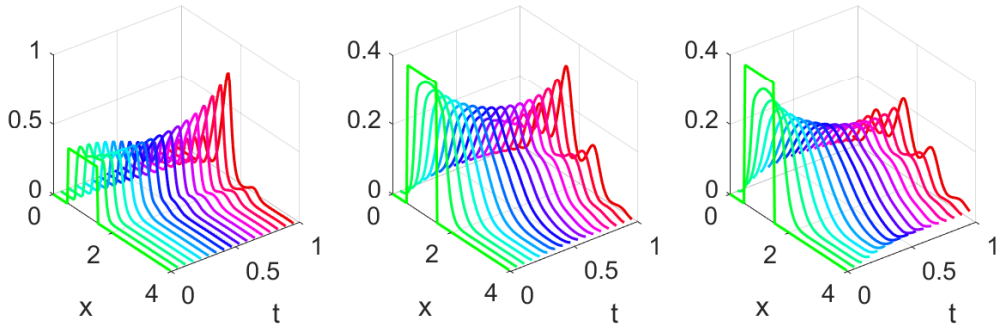


FIGURE 6. Solution  $u(t, x)$  for the problem of controlling the conservation law. From left to right:  $\alpha = 0, 0.5, 1$ .

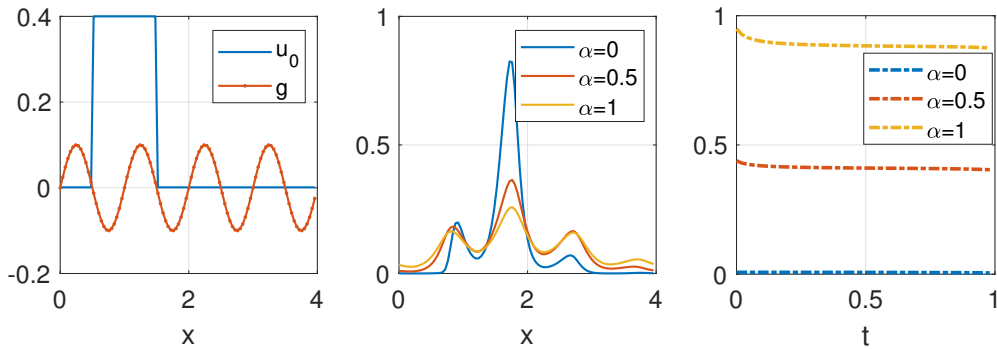


FIGURE 7. Left: boundary conditions for the control problems  $u_0$ . Middle: solution  $u(1, x)$  for the problem of controlling the conservation law. Right: the numerical Hamiltonian  $H_G(u, \Phi)$ .

## REFERENCES

- [1] L. Ambrosio, N. Gigli and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zurich, Birkhauser Verlag, Basel, 2nd ed. 2008.
- [2] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Communications in Partial Differential Equations*, 26 (1), 2000.
- [3] J. Backhoff, G. Conforti, I. Gentil, and C. Leonard. The mean field Schrödinger problem: ergodic behavior, entropy estimates and functional inequalities. *Probability Theory and Related Fields*, 2020.
- [4] J.D. Benamou and Y. Brenier. A Computational Fluid Mechanics Solution to the Monge-Kantorovich Mass Transfer Problem. *Numerische Mathematik*, 84(3):3750–393, 2000.
- [5] Y. Brenier. The Initial Value Problem for the Euler Equations of Incompressible Fluids Viewed as a Concave Maximization Problem. *Commun. Math. Phys.* 364, 579–605, 2018.
- [6] Y. Brenier and I. Moyano. Relaxed solutions for incompressible inviscid flows: A variational and gravitational approximation to the initial value problem. *arXiv:2104.14803*, 2021.
- [7] L. Briceño-Arias, D. Kalise, Z. Kobeissi, M. Laurière, A. M. González, and F. J. Silva. On the implementation of a primal-dual algorithm for second order time-dependent mean field games with local couplings. *ESAIM: Proceedings and Surveys*, 65:330–348, 2019.



- [8] P. Cardaliaguet, F. Delarue, J. Lasry, and P. Lions. The master equation and the convergence problem in mean field games. *arXiv:1509.02505*, 2015.
- [9] J. A. Carrillo, S. Lisini, G. Savare and D. Slepcev. Nonlinear mobility continuity equations and generalized displacement convexity. *Journal of Functional Analysis* 258(4):1273-1309, 2009.
- [10] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision.*, 40(1):120-145, 2011.
- [11] Y. Chen, T. Georgiou, and M. Pavon. On the Relation Between Optimal Transport and Schrodinger Bridges: A Stochastic Control Viewpoint. *J Optim Theory Appl*, 169:671-691, 2016.
- [12] A. Chiarini, G. Conforti and L. Tamanini. Schrodinger Problem for Lattice Gases: A Heuristic Point of View. *Geometric Science of Information*, 2021.
- [13] C. Clason and T. Valkonen. Primal-dual extragradient methods for nonlinear nonsmooth pde-constrained optimization. *SIAM Journal on Optimization*, 27(3):1314–1339, 2017.
- [14] J. Dolbeault, B. Nazaret, and G. Savare A new class of transport distances. *Calculus of Variations and Partial Differential Equations*, (2):193–231, 2009.
- [15] M.D. Donsker, and S.R.S. Varadhan. Large deviations from a hydrodynamic scaling limit. *Commun. Pure Appl. Math*, 42(3), 243–270, 1989.
- [16] L.C. Evans. A Survey of Entropy Methods for Partial Differential Equations. *Bull. Amer. Math. Soc.*, 41, 409-438, 2004.
- [17] W. Gangbo, W. Li, S. Osher, and M. Puthawala. Unnormalized Optimal Transport. *Journal of Computational Physics*, 2019.
- [18] W. Gangbo, T. Nguyen, and A. Tudorascu. Hamilton-Jacobi equations in the Wasserstein space. *Meth. Appl. Anal.* Vol. 15 no 2, 155-184, 2008.
- [19] M. Huang, R.P. Malhame, and P. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221-252, 2006.
- [20] M. Jacobs, F. Léger, W. Li, and S. Osher. Solving Large-Scale Optimization Problems with a Convergence Rate Independent of Grid Size. *SIAM J. Numer. Anal.*, 57(3), 1100–1123, 2019.
- [21] J. M. Lasry, and P. L. Lions, Mean field games. *Japanese Journal of Mathematics*, 2, 229-260, 2007.
- [22] P. Lax. Shock Waves and Entropy. *Contributions to Nonlinear Functional Analysis*, Proceedings of a Symposium Conducted by the Mathematics Research Center, the University of Wisconsin–Madison, April 12–14, 1971.
- [23] F. Leger and W. Li. Hopf–Cole transformation via generalized Schrödinger bridge problem. *Journal of Differential Equations*, Volume 274, 788–827, 2021.
- [24] C. Leonard. A survey of the Schrodinger problem and some of its connections with optimal transport. *DCDS*, 34(4):1533–1574, 2014.
- [25] R. J. LeVeque and R. J. LeVeque. *Numerical methods for conservation laws*, volume 132. Springer, 1992.
- [26] W. Li. Transport information geometry: Riemannian calculus on probability simplex. *Information Geometry*, 2021.
- [27] W. Li. Hessian metric via transport information geometry. *Journal of Mathematical Physics*, 62, 033301, 2021.
- [28] W. Li, and S. Osher. Constrained dynamical optimal transport and its Lagrangian formulation. *arXiv:1807.00937*, 2018.
- [29] W. Li, and L. Ying. Hessian transport gradient flows. *Research in the Mathematical Sciences*, 6, 34, 2019.
- [30] S. Lisini, D. Matthes, and G. Savaré. Cahn–Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *Journal of differential equations*, (253): 814-850, 2012.
- [31] A. Mielke. A gradient structure for reaction diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 4(4):1329, 2011.
- [32] A. Mielke. Free energy, free entropy, and a gradient structure for thermoplasticity. In “K. Weinberg and A. Pandolfi (eds): *Innovative Numerical Approaches for Multi-Field and Multi-Scale Problems*”. Lecture Notes in Appl. Comp. Mechanics Vol. 81, pages 135-160, Springer 2016.

- [33] A. Mielke, D. R. M. Renger, and M. A. Peletier. A generalization of Onsagers reciprocity relations to gradient flows with nonlinear mobility. *J. Non-Equil. Thermodyn.*, 41(2), 141-149, 2016.
- [34] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations*, 26(1-2):101–174, 2001.
- [35] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, no. 950, iv+141, 202, 2009.
- [36] C. Villani. *Optimal Transport: Old and New*. Number 338 in Grundlehren Der Mathematischen Wissenschaften. Springer, Berlin, 2009.

*Email address:* `wuchen@mailbox.sc.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH CAROLINA, COLUMBIA

*Email address:* `siting6@math.ucla.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES

*Email address:* `sjo@math.ucla.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES