

ON THE VARIATIONAL DUAL FORMULATION OF THE NASH SYSTEM AND AN ADAPTIVE CONVEX GRADIENT-FLOW APPROACH TO NONLINEAR PDES

DMITRY VOROTNIKOV AND AMIT ACHARYA

ABSTRACT. We investigate the influence of base states on the consistency of the dual variational formulation for quadratic systems of PDEs, which are not necessarily conservative (typical examples include the noise-free Nash system with a quadratic Hamiltonian and multiple players). We identify a sufficient condition under which consistency holds over large time intervals. In particular, in the single-player case, there exists a sequence of base states (each exhibiting full consistency) that converges in mean to zero. We also prove existence of variational dual solutions to the noise-free Nash system for arbitrary base states. Furthermore, we propose a scheme based on Hilbertian gradient flows that, starting from an arbitrary base state, generates a sequence of new base states that is expected to converge to a solution of the original PDE.

Keywords: convex duality, generalized optimal transport, gradient flow
MSC [2020] **35D99, 35L40, 37K58, 41A60, 49Q99**

1. INTRODUCTION

Consider the noise-free Nash system with the quadratic Hamiltonian and N players, cf. [21]:

$$-\partial_\tau \psi_i + \frac{1}{2} |\partial_{x_i} \psi_i|^2 + \sum_{j \neq i} \partial_{x_j} \psi_j \partial_{x_j} \psi_i = 0, \quad i = 1, \dots, N. \quad (1.1)$$

Here¹ the unknown function is $\psi = \psi(\tau, x)$, where $\tau \in [0, T]$ and x belongs to the periodic box $\mathbb{T}^N$. The system is complemented with the terminal condition

$$\psi(T, x) = \psi_*(x). \quad (1.2)$$

More generally, if the states of the players are p -dimensional, we have the following system, cf. [17]:

$$-\partial_\tau \psi_i + \sum_{l=1}^p \left[\frac{1}{2} |\partial_{x_{il}} \psi_i|^2 + \sum_{j \neq i} \partial_{x_{jl}} \psi_j \partial_{x_{jl}} \psi_i \right] = 0, \quad i = 1, \dots, N. \quad (1.3)$$

Here $\tau \in [0, T]$ and x is in the periodic box $\mathbb{T}^{N \times p}$.

Then (1.3) — as well as its particular case (1.1) — can be written in the form

$$-\partial_\tau \psi + \mathfrak{U}(\nabla \psi \otimes \nabla \psi) = 0, \quad \psi(T, x) = \psi_*(x), \quad (1.4)$$

¹The Einstein summation convention is not used here but will be adopted in Section 3.

where $\mathfrak{U} : \mathbb{R}_s^{n \times n} \rightarrow \mathbb{R}^N$ is a finite-dimensional linear operator (hereafter $n = pN^2$). More explicitly, the elements of $\mathbb{R}_s^{n \times n}$ can be represented in the form $A = \{A_{ijk,hqr}\}$, where i, j, h and q vary between 1 and N , whereas k and r vary between 1 and p ; then $\mathfrak{U}$ acts according to

$$(\mathfrak{U}A)_i = \frac{1}{2} \sum_{l=1}^d \left[A_{iil,iil} + \sum_{j \neq i} (A_{jjl,ijl} + A_{ijl,jjl}) \right]. \quad (1.5)$$

Setting $t := T - \tau$, and slightly abusing the notation by letting $\psi(t, x) := \psi(\tau, x)$, we rewrite the problem in the form

$$\partial_t \psi + \mathfrak{U}(\nabla \psi \otimes \nabla \psi) = 0, \quad \psi(0, x) = \psi_*(x). \quad (1.6)$$

The particular case $N = 1$ and $\mathfrak{U} = \frac{1}{2} \text{Tr}$ corresponds to the classical quadratic Hamilton-Jacobi (H-J) equation.

To the best of our knowledge, **no solvability results are currently available for the noise-free Nash system** with $N > 1$, cf. [21]. This absence is related to the lack of a comparison principle for H-J systems. In particular, there is no analogue of viscosity solutions for the noise-free Nash system. Weak solutions in the sense of distributions can still be defined (cf. Definition 2.1), though, of course, without any expectation of uniqueness. Furthermore, due to the lack of compactness, no existence result for such solutions is currently known. Finally, no selection criterion has been established to address the non-uniqueness phenomenon, although vanishing viscosity may provide a possible approach.

Beyond the natural applications to game theory of the H-J systems, another primary motivation of our study comes from the modern theory of mechanics of defects in continuum mechanics, with many outstanding classical as well as cutting-edge applications in engineering and science. The following set of PDEs is a ‘simplest’ model for the dissipative dynamics of localized, non-singular dislocation line defects and cracks in elastic solids and geophysical rupture [2], [12], [27], [33], [34]: given a domain $\Omega \subset \mathbb{R}^3$, $0 \leq \varepsilon \ll 1$, and a multi-well, non-convex $f : \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}$, we look for $P : \Omega \times [0, T] \rightarrow \mathbb{R}^{3 \times 3}$, $\theta : \Omega \times [0, T] \rightarrow \mathbb{R}^3$ that solve²

$$\begin{aligned} \partial_t P &= -\text{curl } P \times X \left[\left(-(\nabla \theta - P) + \varepsilon \text{curl curl } P + \partial_P f(P) \right)^T \text{curl } P \right], \\ \partial_{tt} \theta &= \text{div}(\nabla \theta - P), \end{aligned}$$

with appropriate initial and boundary conditions (see [1] and earlier references therein for a model without simplifying assumptions). Particular physically meaningful ansätze often allow assuming P to have N non-zero components, $N \leq 9$, possibly varying in $d \leq 3$ essential directions in space (similar reductions can be implemented for the field θ as well). For $N = 1$, $d = 1, 2$, robust computational techniques have been developed for this set of equations generating useful physical insights into defect dynamics in elastic solids [34], cracks [27], nematic liquid crystals [33], and geophysical rupture dynamics [12], [34]. For $N > 1$ this is a H-J-like system, and, to our knowledge, no theory or numerical schemes

²Notation: $(X[AB])_i = \sum_{j,k,M} \epsilon_{ijk} A_{jM} B_{Mk}$; $(\text{curl } P)_{Mi} = \sum_{j,k} \epsilon_{ijk} \partial_j P_{Mk}$, where ϵ_{ijk} is the Levi-Civita symbol.

are available for it — see [4] for a first analysis in the static and [8], [10] in the dynamic settings. The dual methodology studied in this paper marks the beginning of a rigorous analysis of such dynamical models; for a model and analysis of plasticity from dislocations in the setting of integral currents, see [28], and within a phase-field setting see [11] as a representative of the work of Monneau and co-workers related to dislocations and H-J type equations. In the practical realm, the variational structure and the gradient flow scheme that we will propose already has the significant advantage of suggesting a natural finite dimensional Rayleigh-Ritz discretization for a multidimensional system involving nonlinear transport which does not easily, if at all, lend itself to discretization based on existing numerical algorithms.

In this paper, we are concerned with the variational dual formulation of such nonlinear systems, with an emphasis on the noise-free Nash system. The idea originates with Brenier [14], [15], who proposed, for the incompressible Euler equations and related models, to study solutions that minimize the Lagrangian action, i.e., the time integral of the kinetic energy (for recent applications of the least action principle as a selection criterion for systems of hyperbolic conservation laws, see [19], [20]). Although this minimization problem may fail to have a solution, it leads to a dual problem with more favorable convexity properties. Brenier derived an explicit relation linking smooth solutions of the incompressible Euler equations on short time intervals with the *variational dual solutions*. In [16], this technique was applied to the multi-stream Euler-Poisson system. A numerical implementation for the dual variational formulation of the quadratic porous medium equation and the Burgers equation has recently been done in [26]. In [31], [32], the first author extended Brenier’s approach by identifying structures in nonlinear PDEs that allow variational dual formulations with favorable properties. These dual problems are related to optimal transport, specifically its *ballistic* variant [26], [32]. Notably, the functionals arising in this context are not expected to be self-dual, so the formalism of [18] is not applicable.

A similar variational dual formulation has been proposed by the second author and several collaborators, see [1], [5], [6], [7], [9], [25], [29] and the references therein. A key feature of this approach is that the dual problem is built around a *base state*, which can be seen as an “initial guess” for the solution of a PDE in question; accordingly, the kinetic energy is replaced with an appropriate relative energy with positive-definite Hessian. From this perspective, the base states in [14], [15], [26], [31], [32] are simply zero.

The primary goal of these approaches is to define convex variational principles for conservative and non-conservative (including dissipative) physical models to facilitate their solution, whether exact or approximate (through computation). For instance, they allow, using methods of the calculus of variations, the definition and proofs of existence of variational dual solutions to the Euler-Lagrange PDE systems of physical energy functionals whose global minimizers do not exist [29, Sec. 5], [6]. However, a somewhat unsettling aspect of the duality scheme with the zero base state is that many consistency results are obtained only on small time intervals, cf. [14, Theorems 2.3 and 3.1], [31, Theorem 3.8], [16, Theorem 5.1]. In [32], the first author succeeded in extending the consistency to large time intervals by introducing suitable time-dependent weights; as an application, he

derived a variant of Dafermos principle for various PDE models that admit the variational dual formulation.

In Section 2, we investigate how the deployment of base states (without any time-dependent weights) influences consistency. We identify a broad class of base states that ensure consistency of the dual formulation of the noise-free Nash system over any interval $[0, T]$ (Theorem 2.4). As a consequence, in the single-player setting ($N = 1$), we build a sequence of base states — each exhibiting full consistency — that converges in mean to zero (Corollary 2.6). Moreover, we show that, for $N > 1$, a variational dual solution to the noise-free Nash system exists and belongs to a certain L^2 -space, without any restrictions on base states and initial data (Theorem 2.11).

Subsequently, in Section 3 we propose a novel scheme, based on Hilbertian gradient flows and suitable for a wide class of nonlinear PDEs, that produces, from an initial base state, a sequence of new base states anticipated to converge to a solution of the original PDE. The proposed adaptive gradient flow scheme is a natural outgrowth of the practical experience with computational schemes based on Newton's method for approximating the Euler-Lagrange system of the dual variational principle. From the very first nonlinear problem solved in [23] (Euler's ODE system for dynamics of a rigid body), it was clear that the idea of base states was crucial to the practical success of the scheme, further confirmed in [24] in the context of (inviscid) Burgers (see also [26]) and in [29] for nonconvex elastostatics and dynamics (in space dimension 1). In [25], base state changes with step-size control in a Newton approximation scheme were used, and the possibility of a gradient flow-based approach was alluded to. The gradient flow scheme formalized here in detail constitutes a new and logically natural development of these ideas. We work in an exact setting, i.e., without approximation.

Our gradient flow scheme proceeds through several stages of gradient descent in the Hilbert space of functions of t and x , with respect to a fictitious, time-like variable s along which the evolution unfolds. Transitions between stages correspond to the selection of a new base state, unambiguously determined by the scheme, without disrupting the continuity of the overall trajectory. In Section 4, using a simple toy model, we demonstrate how the switching of base states, the convergence of the scheme, and the eventual generation of a solution occur in practice. Section 5 presents our scheme in the particular case of the noise-free Nash system.

2. THE DUAL VARIATIONAL FORMULATION OF THE NOISE-FREE NASH SYSTEM. CONSISTENCY AND EXISTENCE OF SOLUTIONS.

Basic notation. Let $\Omega := \mathbb{T}^{N \times p}$, and denote for brevity $X = L^2(\Omega)$. We use the notations $\mathbb{R}^{n \times n}$ and $\mathbb{R}_s^{n \times n}$ for the spaces of $n \times n$ matrices and symmetric matrices, resp., with the scalar product generated by the Frobenius norm. Let $X_s^{n \times n}$ be the subspace of $X^{n \times n}$ consisting of symmetric-matrix-valued functions. The parentheses $(\cdot, \cdot)$ will stand for the scalar products in X^n and $X_s^{n \times n}$. For $A, B \in X_s^{n \times n}$, we write $A \geq B$ and $A > B$ when $A - B$ is a positive-semidefinite-matrix-function and a positive-definite-matrix-function, resp. The symbol I stands for the identity matrix of a relevant size. Denote by $\Pi : X^n \rightarrow X^n$ the

averaging operator defined by

$$\Pi v(x) = \int_{\Omega} v(y) dy.$$

2.1. The variational dual formulation of the Nash system. We begin by presenting the variational dual formulation of the noise-free Nash system, following the approach outlined in [14], [31] and employing the base states introduced in [1].

We set $v := \nabla \phi$, $v_0 := \nabla \phi_*$, and define the differential operator L by

$$Lw = -\nabla(\mathfrak{L}w). \quad (2.1)$$

Then a natural (Burgers-like) weak formulation of our problem (1.6) consists in finding a function $v : [0, T] \rightarrow X^n$ such that

$$\int_0^T [(v - v_0, \partial_t a) + (v \otimes v, L^* a)] dt = 0 \quad (2.2)$$

for all sufficiently smooth vector fields $a : [0, T] \rightarrow X^n$, $a(T) = 0$. Here

$$L^* = \mathfrak{L}^* \operatorname{div}$$

is the adjoint operator. **Note that this weak formulation includes the information that v is a gradient and the initial condition.** Indeed, multiplying (1.6) by $\operatorname{div} a$, integrating w.r.t. $(0, T) \times \Omega$ and observing that

$$\int_0^T (\partial_t \psi, \operatorname{div} a) dt = - \int_0^T (\partial_t v, a) dt = \int_0^T (v - v_0, \partial_t a) dt \quad (2.3)$$

we get (2.2). Conversely, if b is any smooth divergence-free (w.r.t. x) field on $[0, T] \times \Omega$, test (2.2) with $a(t) = \int_T^t b(s) ds$ to obtain $\int_0^T (v - v_0, b) dt = 0$, which yields that $v - v_0$ is a gradient, and hence $v = \nabla \psi$ for some $\psi : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ that is determined up to adding a function of t . With this at hand, and a being now chosen arbitrary within its range, we observe that (2.2) implies (up to regularity issues) that

$$(\psi(0) - \psi_*, \operatorname{div} a(0)) + \int_0^T (\partial_t \psi + \mathfrak{L}(\nabla \psi \otimes \nabla \psi), \operatorname{div} a) dt = 0,$$

which means that there exists a function $g(t)$ and a constant c such that $\partial_t \psi + \mathfrak{L}(\nabla \psi \otimes \nabla \psi) = g(t)$ and $\psi(0) = \psi_* + c$. Since ψ has been defined up to adding a function of time, without loss of generality we may assume that $g(t) \equiv 0$ and $c = 0$ (otherwise replace ψ with $\psi - c - \int_0^t g(s) ds$).

Let us now rewrite problem (2.2) in terms of the test functions $B := L^* a$ and $E := \partial_t a$. The link between B and E can alternatively be described by the condition

$$\int_0^T [(B, \partial_t \Psi) + (E, L \Psi)] dt = 0 \quad (2.4)$$

for all smooth vector fields $\Psi : [0, T] \rightarrow X_s^{n \times n}$, $\Psi(0) = 0$, cf. [31]. In particular, this implies that $B(t, x)$ a.e. belongs to the range of the operator L^* . Note that (2.2) becomes

$$\int_0^T [(v - v_0, E) + (v \otimes v, B)] dt = 0. \quad (2.5)$$

Motivated by this discussion, we adopt the following definition.

Definition 2.1 (Weak solutions). Let $v_0 \in X^n$. A function $v \in L^2((0, T) \times \Omega; \mathbb{R}^n)$ is a *weak solution* to (1.6) (more precisely, to its Burgers-like formulation) if it satisfies (2.5) for all pairs

$$(E, B) \in L^2((0, T) \times \Omega; \mathbb{R}^n) \times L^\infty((0, T) \times \Omega; \mathbb{R}^{n \times n}) \quad (2.6)$$

meeting the constraint (2.4).

Following [1], [3], fix a *base state* $\bar{v} \in L^2((0, T) \times \Omega; \mathbb{R}^n)$, and consider the problem of finding a weak solution to (1.6), regardless of the fact that there is no existence theorem for the weak solutions, that minimizes the time integral of the relative energy

$$K_{v, \bar{v}}(t) := \frac{1}{2}(v(t) - \bar{v}(t), v(t) - \bar{v}(t)). \quad (2.7)$$

This can be implemented via the saddle-point problem

$$\mathcal{I}(v_0, \bar{v}, T) = \inf_v \sup_{E, B: (2.4)} \int_0^T [K_{v, \bar{v}}(t) + (v - v_0, E) + (v \otimes v, B)] dt. \quad (2.8)$$

(If a weak solution does not exist, then $\mathcal{I}(v_0, \bar{v}, T) = +\infty$). Equivalently,

$$\mathcal{I}(v_0, \bar{v}, T) = \inf_v \sup_{E, B: (2.4)} \int_0^T \left[(v - v_0, E - \bar{v}) + \frac{1}{2}(v \otimes v, I + 2B) \right] dt + \mathfrak{C}(\bar{v}, v_0), \quad (2.9)$$

where

$$\mathfrak{C}(\bar{v}, v_0) := \int_0^T \left[-(v_0, \bar{v}) + \frac{1}{2}(\bar{v}, \bar{v}) \right] dt. \quad (2.10)$$

The infimum in (2.8) is taken over all $v \in L^2((0, T) \times \Omega; \mathbb{R}^n)$, and the supremum is taken over all pairs (E, B) satisfying (2.6) and the linear constraint (2.4).

The dual problem is

$$\mathcal{J}(v_0, \bar{v}, T) = \sup_{E, B: (2.4)} \inf_v \int_0^T \left[(v - v_0, E - \bar{v}) + \frac{1}{2}(v \otimes v, I + 2B) \right] dt + \mathfrak{C}(\bar{v}, v_0), \quad (2.11)$$

where v, E, B are varying in the same function spaces as above.

It is easy to see that any solution to (2.11) necessarily satisfies

$$I + 2B \geq 0 \text{ a.e. in } (0, T) \times \Omega. \quad (2.12)$$

Following [14], [31], consider the nonlinear functional

$$\mathcal{K} : L^2((0, T) \times \Omega; \mathbb{R}^n) \times L^\infty((0, T) \times \Omega; \mathbb{R}^{n \times n}) \rightarrow \mathbb{R}$$

defined by the formula

$$\mathcal{K}(Q, B) = \inf_{z \otimes z \leq M} \int_0^T \left[(z, Q) + \frac{1}{2}(M, I + 2B) \right] dt, \quad (2.13)$$

where the infimum is taken over all pairs $(z, M) \in L^2((0, T) \times \Omega; \mathbb{R}^n) \times L^1((0, T) \times \Omega; \mathbb{R}_s^{n \times n})$. Note that $\mathcal{K}(Q, B)$ is a negative-semidefinite quadratic form w.r.t. Q for fixed B . Mimicking [31, Remark 3.6], it is possible to see that (2.11) is equivalent to

$$\mathcal{J}(v_0, \bar{v}, T) = \sup_{E, B: (2.4), (2.12)} \int_0^T \left[-(v_0, E) + \frac{1}{2}(\bar{v}, \bar{v}) \right] dt + \mathcal{K}(E - \bar{v}, B), \quad (2.14)$$

where the supremum is taken over all pairs (E, B) belonging to the class (2.6).

Remark 2.2. For the sake of generality, in what follows v_0 and $\bar{v}$ are not necessarily gradients because the dual problem is well-defined even without this restriction.

2.2. Consistency and existence of solutions. In this section, we first show that a broad class of base states ensures consistency of the dual formulation over any interval $[0, T]$. Specifically, any matrix-valued positive semidefinite “density” $G(t, x)$ and the “velocity field” $u(t, x)$ solving the generalized “transport” equation

$$\partial_t G + 2\mathcal{U}^* \operatorname{div}(Gu) = 0, \quad G(T, \cdot) \equiv I \quad (2.15)$$

define a base state with no duality gap and an explicit correspondence between the solutions of the original problem and the variational dual solutions. As a corollary, in the single-player case ($N = 1$), we construct a sequence of base states (each exhibiting full consistency and a one-to-one correspondence between the “primal” and “dual” solutions) that converges in mean to zero. Finally, we establish the existence of a variational dual solution to the noise-free Nash system for any base state $\bar{v}$ and $N > 1$.

Remark 2.3. Although our main focus is on the noise-free Nash system, we are confident that the results below can be extended to other quadratic PDEs, for instance, to those studied in [14], [31].

Theorem 2.4 (Consistency). *Let $v_0 \in X^n$ and v be a weak solution to (1.6). Select any pair $(G, u) \in L^\infty((0, T) \times \Omega; \mathbb{R}_s^{n \times n}) \times L^2((0, T) \times \Omega; \mathbb{R}^n)$ satisfying the constraints*

$$G \geq 0 \text{ a.e. in } (0, T) \times \Omega \quad (2.16)$$

and

$$\int_0^T [(G - I, \partial_t \Psi) - 2(Gu, L\Psi)] dt = 0 \quad (2.17)$$

for all smooth vector fields $\Psi : [0, T] \rightarrow X_s^{n \times n}$, $\Psi(0) = 0$. Assume also that

$$\bar{v} = G(v - u) \text{ a.e. in } (0, T) \times \Omega. \quad (2.18)$$

Then $\mathcal{I}(v_0, \bar{v}, T) = \mathcal{J}(v_0, \bar{v}, T)$, and the pair (E_+, B_+) defined by

$$E_+ := -Gv + \bar{v}, \quad B_+ := \frac{1}{2}(G - I) \quad (2.19)$$

belongs to the class (2.6) and maximizes (2.14). In particular, if $G > 0$ a.e. in an open subset $\mathcal{O} \subset (0, T) \times \Omega$, we can retrieve the weak solution v (inside $\mathcal{O}$) from the variational dual solution (E_+, B_+) by applying the formula

$$v = (I + 2B_+)^{-1}(\bar{v} - E_+) \text{ a.e. in } \mathcal{O}. \quad (2.20)$$

Proof. By construction, $\bar{v} \in L^2((0, T) \times \Omega; \mathbb{R}^n)$. Consequently, the pair (E_+, B_+) belongs to $L^2((0, T) \times \Omega; \mathbb{R}^n) \times L^\infty((0, T) \times \Omega; \mathbb{R}_s^{n \times n})$. Moreover, it follows from (2.17) and (2.18) that the pair (E_+, B_+) verifies (2.4), while (2.16) implies (2.12) for B_+ .

By (2.5) and (2.8),

$$\mathcal{I}(v_0, \bar{v}, T) \leq \int_0^T K_{v, \bar{v}}(t) dt.$$

In view of (2.14), in order to prove that there is no duality gap — and that (E_+, B_+) is a variational dual solution — it suffices to show that

$$\int_0^T -(v_0, E_+) dt + \mathcal{K}(E_+ - \bar{v}, B_+) \geq \int_0^T \left[K_{v, \bar{v}}(t) - \frac{1}{2}(\bar{v}, \bar{v}) \right] dt. \quad (2.21)$$

But since v , in particular, satisfies (2.5) with the test functions (E_+, B_+) , we have

$$\int_0^T [(v - v_0, E_+) + (v \otimes v, B_+)] dt = 0. \quad (2.22)$$

Hence, by the first definition in (2.19),

$$\int_0^T [-(v_0, E_+) + (v \otimes v, B_+)] dt = \int_0^T [(v \otimes v, G) - (v, \bar{v})] dt, \quad (2.23)$$

which in view of the second definition in (2.19) gives

$$-\int_0^T [(v_0, E_+) + (v \otimes v, B_+)] dt = \int_0^T [(v \otimes v, I) - (v, \bar{v})] dt. \quad (2.24)$$

Consequently,

$$\begin{aligned} & \int_0^T -(v_0, E_+) dt + \mathcal{K}(E_+ - \bar{v}, B_+) \\ &= \int_0^T -(v_0, E_+) dt + \inf_{z \otimes z \leq M} \int_0^T \left[(z, -Gv) + \frac{1}{2}(M, I + 2B_+) \right] dt \\ &\geq \int_0^T -(v_0, E_+) dt + \inf_{z \in L^2((0, T) \times \Omega; \mathbb{R}^n)} \int_0^T \left[-(z, (I + 2B_+)v) + \frac{1}{2}(z \otimes z, I + 2B_+) \right] dt \\ &= \int_0^T - \left[(v_0, E_+) + \frac{1}{2}((I + 2B_+)v, v) \right] dt \\ &= \int_0^T \left[(v \otimes v, I) - (v, \bar{v}) - \frac{1}{2}(v, v) \right] dt = \int_0^T \left[K_{v, \bar{v}}(t) - \frac{1}{2}(\bar{v}, \bar{v}) \right] dt. \end{aligned}$$

If $G > 0$ a.e. in an open set $\mathcal{O}$, then (2.19) obviously yields (2.20). $\square$

Remark 2.5. Constraints (2.16) and (2.17) may be interpreted as a weak form of a generalized “transport” equation (2.15) where the matricial positive-semidefinite “density” $G(t, x)$ and the “velocity field” $u(t, x)$ may be freely chosen. The simplest pair satisfying (2.16), (2.17) is $(G, u) \equiv (I, 0)$, which yields $\bar{v} = v$. In this connection, we note that the consistency of the dual formulation of the incompressible Euler equations when the base state coincides with the solution was established in [9, Theorem 3.6].

Corollary 2.6. *Let $N = 1$ (and hence $n = p$), and let v be the gradient of the viscosity solution ψ to the Hamilton-Jacobi equation (1.6). Assume that v_0 and v are continuous in Ω and $[0, T] \times \Omega$, resp. Then there exists a sequence $\{\bar{v}_m\}$ of continuous base states with the following properties:*

- i) $\bar{v}_m \rightarrow 0$ strongly in $L^1((0, T) \times \Omega; \mathbb{R}^n)$;
- ii) for every $m \in \mathbb{N}$, $\mathcal{I}(v_0, \bar{v}_m, T) = \mathcal{J}(v_0, \bar{v}_m, T)$;
- iii) for every $m \in \mathbb{N}$, the dual problem (2.14) with $\bar{v} = \bar{v}_m$ has a smooth solution (E_m, B_m) satisfying $I + 2B_m > 0$ in $(0, T) \times \Omega$;
- iv) the gradient v of the solution to the Hamilton-Jacobi equation can be fully retrieved from the variational dual solutions according to the formula

$$v = (I + 2B_m)^{-1}(\bar{v}_m - E_m). \quad (2.25)$$

Proof. Since v is continuous in $[0, T] \times \Omega$, there exists a sequence $\{u_m\}$ of smooth vector fields on $[0, T] \times \Omega$ that converges to v uniformly. Let $\rho_m : [0, T] \times \Omega \rightarrow \mathbb{R}$ be the smooth solution of the linear transport equation

$$\partial_t \rho_m + \operatorname{div}(\rho_m u_m) \quad (2.26)$$

that satisfies the terminal condition

$$\rho_m(T, x) = 1. \quad (2.27)$$

Since each u_m is smooth and $\frac{d}{dt} \rho_m(t) = \rho_m(t) \operatorname{div} u_m$ along any corresponding characteristic, ρ_m is bounded away from 0 and $+\infty$ (not uniformly in m). Set $G_m := \rho_m I$, and fix a smooth matrix field $\Psi : [0, T] \rightarrow X_s^{n \times n}$, $\Psi(0) = 0$. Multiply (2.26) by $\operatorname{Tr} \Psi$, and integrate by parts over $(0, T) \times \Omega$ to obtain

$$\int_0^T [((\rho_m - 1)I, \partial_t \Psi) + (G_m u_m, \nabla \operatorname{Tr} \Psi)] dt = 0. \quad (2.28)$$

In particular, letting $\Psi(t, x) = tI$ gives

$$\|\rho_m\|_{L^1((0, T) \times \Omega)} = \|1\|_{L^1((0, T) \times \Omega)}.$$

Remembering that for $N = 1$ we have $L = -\nabla \mathfrak{U} = -\frac{1}{2} \nabla \operatorname{Tr}$, we conclude that the pairs (G_m, u_m) satisfy the constraints (2.16) and (2.17) of Theorem 2.4. Let $\bar{v}_m := \rho_m(v - u_m) =$

$G_m(v - u_m)$. Then

$$\begin{aligned} \|\bar{v}_m\|_{L^1((0,T)\times\Omega;\mathbb{R}^n)} &\leq \|\rho_m\|_{L^1((0,T)\times\Omega)} \|u_m - v\|_{L^\infty((0,T)\times\Omega;\mathbb{R}^n)} \\ &= \|1\|_{L^1((0,T)\times\Omega)} \|u_m - v\|_{L^\infty((0,T)\times\Omega;\mathbb{R}^n)} \rightarrow 0. \end{aligned}$$

It remains to apply Theorem 2.4 after having observed that $G_m > 0$ everywhere in $\mathcal{O} := (0, T) \times \Omega$ and that the functions $E_m := -G_m v + \bar{v}_m = -\rho_m u_m$, $B_m := \frac{1}{2}(G_m - I) = \frac{1}{2}(\rho_m - 1)I$ are smooth. $\square$

Remark 2.7. In the pioneering work [14], Brenier considered the variational dual formulation (with zero base state) for the quadratic Burgers equation — or, equivalently, to the one-dimensional quadratic H–J — which corresponds to $n = N = p = 1$ and $\bar{v} = 0$ in our framework. He shows explicitly that, in that setting, the dual solution $(\rho v, \rho)$ is merely a vector-valued Radon measure (see also Remark 2.12 below, which explains the loss of regularity) and that the support of the measure ρ is strictly smaller than $[0, T] \times \Omega$; here ρ can be interpreted as the probability measure transported by v whose marginal at time T is the Lebesgue measure on Ω (cf. the proof of Corollary 2.6). Consequently, v cannot be recovered as the Radon-Nikodym derivative of ρv w.r.t. ρ outside of the support of ρ . Corollary 2.6 shows that a small perturbation of the base state $\bar{v} = 0$ can dramatically improve the regularity of the dual solution and secure one-to-one correspondence between v and the variational dual solution. The same observations apply in the case $N = 1$ with $p > 1$; under these conditions, the existence of a variational dual solution to H–J with $\bar{v} = 0$ — again a vector-valued Radon measure — follows from [31, Theorem 4.6].

Remark 2.8. The assumption that v is continuous in Corollary 2.6 is, of course, restrictive, as it rules out the formation of shocks. Nevertheless, if v is merely essentially bounded, we can take *any* smooth vector field u_m and solve the transport equation (2.26) with the terminal condition (2.27). Then the dual formulation with the (essentially bounded) base state $\bar{v}_m := \rho_m(v - u_m)$ is fully consistent (in particular, there is a one-to-one correspondence between the “primal” and “dual” solutions and no duality gap). Hence, the class of “good” base states for v with shocks is still very large.

We recall the following technical definition, cf. [31], that is needed to handle the solvability of the dual problem:

Definition 2.9. The operator L is said to satisfy the *trace condition* if the following holds: for any $\zeta \in D(L^*)$ such that the eigenvalues of the matrix $-L^*\zeta(x)$ are uniformly bounded from above by a constant k for a.e. $x \in \Omega$, the eigenvalues of the matrix $L^*\zeta(x)$ are also uniformly bounded from above a.e. in Ω by a constant that depends only on k , and not on ζ or on x .

Lemma 2.10 (Validity of the trace condition). *Assume that $N > 1$. Then the operator L defined by (2.1) satisfies the trace condition.*

Proof. It suffices to prove that $\mathfrak{U}$ satisfies the trace condition. Using (1.5), for any element $y \in \mathbb{R}^N$ we explicitly compute

$$\begin{aligned} (\mathfrak{U}^*(y))_{ijl, iil} &= (\mathfrak{U}^*(y))_{iil, ijl} = \frac{1}{2}y_j, \\ (\mathfrak{U}^*(y))_{ijl, hqr} &= 0 \text{ for all other entries.} \end{aligned}$$

Consequently,

$$\text{Tr } \mathfrak{U}^*(y) = \frac{p}{2} \sum_{i=1}^N y_i. \quad (2.29)$$

Let $k \geq 0$ be such that $kI + \mathfrak{U}^*(y) \geq 0$. This entails, for every pair of indices $i \neq j$, that

$$\left(\frac{2k + y_i}{y_j} \middle| \frac{y_j}{2k} \right) = 2 \left(\frac{k + (\mathfrak{U}^*(y))_{iil, iil}}{(\mathfrak{U}^*(y))_{ijl, iil}} \middle| \frac{(\mathfrak{U}^*(y))_{iil, ijl}}{k + (\mathfrak{U}^*(y))_{ijl, ijl}} \right) \geq 0.$$

Hence,

$$(y_j)^2 \leq 4k^2 + 2ky_i, \quad i \neq j. \quad (2.30)$$

It follows by some elementary algebra that the trace (2.29) is bounded from above by a constant that merely depends on k , and not on y . This entails that the eigenvalues of $\mathfrak{U}^*(y)$ are also uniformly controlled from above. $\square$

Theorem 2.11 (Existence of a variational dual solution). *Assume that $N > 1$. Then for any $v_0 \in X^n$ and $\bar{v} \in L^2((0, T) \times \Omega; \mathbb{R}^n)$ there exists a maximizer (E, B) to (2.14) in the class (2.6), and $\mathfrak{C}(\Pi\bar{v}, \Pi v_0) \leq \mathcal{J}(v_0, \bar{v}, T) < +\infty$.*

Proof. It suffices to consider the pairs (E, B) that meet the restrictions (2.4), (2.12). In particular, the pair $(E, B) = (\Pi\bar{v}, 0)$ satisfies (2.4), (2.12) and hence is admissible. (Indeed, for this pair $\int_0^T [(B, \partial_t \Psi) + (E, L\Psi)] dt = \int_0^T (\text{div } \Pi\bar{v}, \mathfrak{U}\Psi) dt = 0$ for all smooth $\Psi : [0, T] \rightarrow X_s^{n \times n}$.) Testing (2.14) with this pair, and remembering that $|\Omega| = 1$, we see that

$$\begin{aligned} \mathcal{J}(v_0, \bar{v}, T) &\geq \int_0^T \left[-(v_0, \Pi\bar{v}) + \frac{1}{2}(\bar{v}, \bar{v}) \right] dt + \mathcal{K}(\Pi\bar{v} - \bar{v}, 0) \\ &= \int_0^T \left[-(\Pi v_0, \Pi\bar{v}) + \frac{1}{2}(\bar{v}, \bar{v}) \right] dt - \frac{1}{2} \int_0^T (\bar{v} - \Pi\bar{v}, \bar{v} - \Pi\bar{v}) dt \\ &= \int_0^T \left[-(\Pi v_0, \Pi\bar{v}) + \frac{1}{2}(\Pi\bar{v}, \Pi\bar{v}) \right] dt = \mathfrak{C}(\Pi\bar{v}, \Pi v_0). \end{aligned}$$

Let (E_m, B_m) be a maximizing sequence. Since $\mathfrak{C}(\Pi\bar{v}, \Pi v_0) \leq \mathcal{J}(v_0, \bar{v}, T)$, without loss of generality we may assume that

$$\mathfrak{C}(\Pi\bar{v}, \Pi v_0) \leq \int_0^T \left[-(v_0, E_m) + \frac{1}{2}(\bar{v}, \bar{v}) \right] dt + \mathcal{K}(E_m - \bar{v}, B_m). \quad (2.31)$$

The eigenvalues of $-B_m$ are uniformly bounded from above because $I + 2B_m \geq 0$. Consequently, Lemma 2.10 and Definition 2.9 yield $I + 2B_m \leq kI$ with some constant $k > 0$ a.e.

in $(0, T) \times \Omega$. By the definition of $\mathcal{K}$ in (2.13), we have

$$\mathcal{K}(E_m - \bar{v}, B_m) \leq \inf_{z \otimes z \leq M} \int_0^T \left[(z, E_m - \bar{v}) + \frac{k}{2}(M, I) \right] dt = -\frac{1}{2k} \int_0^T (E_m - \bar{v}, E_m - \bar{v}) dt. \quad (2.32)$$

We infer that

$$\begin{aligned} \frac{1}{4k} \int_0^T (E_m, E_m) dt &\leq \frac{1}{2k} \int_0^T (E_m - \bar{v}, E_m - \bar{v}) dt + \frac{1}{2k} \int_0^T (\bar{v}, \bar{v}) dt \\ &\leq -\int_0^T (v_0, E_m) dt + \frac{k+1}{2k} \int_0^T (\bar{v}, \bar{v}) dt - \mathfrak{C}(\Pi \bar{v}, \Pi v_0) \\ &\leq \frac{k+1}{2k} \int_0^T [(\bar{v}, \bar{v}) + 2k(v_0, v_0)] dt - \mathfrak{C}(\Pi \bar{v}, \Pi v_0) + \frac{1}{4(k+1)} \int_0^T (E_m, E_m) dt, \end{aligned} \quad (2.33)$$

which gives a uniform $L^2((0, T) \times \Omega; \mathbb{R}^n)$ -bound on E_m . Taking into account (2.32) and (2.33), we infer that the right-hand side of (2.31) is uniformly bounded, whence $\mathcal{J}(v_0, \bar{v}, T) < +\infty$. The functional $\mathcal{K}$ is concave and upper semicontinuous on $L^2((0, T) \times \Omega; \mathbb{R}^n) \times L^\infty((0, T) \times \Omega; \mathbb{R}_s^{n \times n})$ as an infimum of affine continuous functionals, cf. (2.13). The functional $\int_0^T (v_0, \cdot) dt$ is a linear bounded functional on $L^2((0, T) \times \Omega; \mathbb{R}^n)$. Consequently, every weak-* accumulation point of (E_m, B_m) is a maximizer of (2.14). Note that the constraints (2.4), (2.12) are preserved by the limit. $\square$

Remark 2.12. The trace condition is not valid for $N = 1$, cf. [32], which explains why, as already mentioned in Remark 2.7, for $N = 1$ and $\bar{v} = 0$ the variational dual solution to the quadratic H-J is merely a vector Radon measure and does not belong to the class (2.6).

3. THE UTILITY OF BASE STATES IN PRACTICAL CONSIDERATIONS: A FORMAL CONVEX GRADIENT FLOW SCHEME

In this section we describe a formal scheme for obtaining weak solutions to a class of PDEs; the class includes many equations of continuum mechanics. The scheme is quite generally applicable, up to minor adjustments, including the case of the Nash system, see Section 5 below. Comparable adaptive numerical schemes employing variational dual solutions and adjustable base states — but not based on the gradient flow idea developed here — for solving nonlinear differential equations have recently been demonstrated in [23], [24], [25], [29]. These methods were developed in connection with the nonlinear system of ODEs of Euler for the dynamics of a rigid body, nonconvex elastostatic and dynamics of a bar (without regularization), (inviscid) Burgers equation, and the problem of traveling waves of a dispersive, nonlocal, nonlinear semi-discrete Burgers equation.

3.1. Variational setup. Following [3] for the notation and setup of the problem, let lower-case Latin indices belong to the set $\{1, \dots, m\}$ representing rectangular Cartesian spatial coordinates, t is time, and we will employ the Einstein summation convention. Let upper-case Latin indices belong to the set $\{1, 2, 3, \dots, u\}$, indexing the components of the array of primal variables, $U \in \mathbb{R}^u$. This array contains variables whose partial derivatives, w.r.t

(t, x) , of at most first order appear in the governing system of primal equations - thus, a conversion to a first-order system of the governing PDE is employed. Consider the system of equations

$$\mathcal{C}_{\Gamma I} \partial_t U_I + \partial_{x_j} \mathcal{F}_{\Gamma j}(U) + G_{\Gamma}(U) + V_{\Gamma}(t, x) = 0 \text{ in } \Omega \times (0, T), \quad \Gamma = 1, \dots, d \quad (3.1a)$$

$$\mathcal{C}_{\Gamma I} U_I(x, 0) = \mathcal{C}_{\Gamma I} U_I^{(0)}(x) \text{ specified on } \Omega \text{ (initial conditions),} \quad (3.1b)$$

$$(\mathcal{F}_{\Gamma j}(U) n_j)|_{(t,x)} = (B_{\Gamma j} n_j)|_{(t,x)} \text{ specified on } \partial\Omega_{\Gamma} \text{ (boundary conditions),} \quad (3.1c)$$

where Ω is a given bounded domain in $\mathbb{R}^m$ with sufficiently regular boundary $\partial\Omega \supset \bigcup_{\Gamma} \partial\Omega_{\Gamma}$ and unit normal vector n , and the upper-case Greek indices index the number of equations. Here, $\mathcal{C} \in \mathbb{R}^{d \times m}$ is a fixed matrix, and $\mathcal{F}, G, V, B$ and $U^{(0)}$ are, respectively, $\mathbb{R}^{d \times m}$ -, d -, d -, $\mathbb{R}^{d \times m}$ - and u -dimensional given functions of their argument.

Define the pre-dual functional by forming the scalar products of (3.1a) with the dual fields D , $(t, x) \mapsto D(t, x) \in \mathbb{R}^d$ (see (3.1a)), integrating by parts, substituting the prescribed initial and boundary conditions (ignoring, for now, space-time boundary contributions that are not specified) *and subtracting an auxiliary potential H as shown*:

$$\begin{aligned} \widehat{S}_H[U, D] = & \int_{\Omega} \int_0^T (-\mathcal{C}_{\Gamma I} U_I \partial_t D_{\Gamma} - \mathcal{F}_{\Gamma j}|_U \partial_{x_j} D_{\Gamma} + (G_{\Gamma}|_U + V_{\Gamma}|_{(t,x)}) D_{\Gamma} - H(U, t, x)) \, dt dx \\ & - \int_{\Omega} \mathcal{C}_{\Gamma I} U_I^{(0)}(x) D_{\Gamma}(x, 0) \, dx + \sum_{\Gamma} \int_{\partial\Omega_{\Gamma}} \int_0^T B_{\Gamma j} D_{\Gamma} n_j \, dt \, d\mathcal{H}^{m-1}, \end{aligned} \quad (3.2)$$

(where the arguments (t, x) are suppressed except to display the explicit dependence of V, H and in the initial condition).

Define

$$\begin{aligned} (t, x) \mapsto \mathcal{D}(t, x) &:= (\partial_t D(t, x), \nabla D(t, x), D(t, x)) \in \mathbb{R}^l, \quad l := d + md + d \\ \mathcal{L}_H(U, \mathcal{D}, t, x) &:= -\mathcal{C}_{\Gamma I} U_I \partial_t D_{\Gamma} - \mathcal{F}_{\Gamma j}|_U \partial_{x_j} D_{\Gamma} + (G_{\Gamma}|_U + V_{\Gamma}(t, x)) D_{\Gamma} - H(U, t, x). \end{aligned} \quad (3.3)$$

The (t, x) dependence of H , introduced in [1, Sec. 5] with the stated goal of aiding in the recovery of non-unique solutions of the primal problem, includes the deployment of “*base states*” in the design of the dual functional³. These are prescribed fields $(t, x) \mapsto \bar{U}(t, x)$. We now require that the choice of H , given t, x and a base state $\bar{U}$, be such that there exists a non-empty, open, convex subset $\mathcal{O} = \mathcal{O}(t, x) \subset \mathbb{R}^l$ and a function $U^{(H)} : \mathcal{O} \times \mathbb{R}^u \rightarrow \mathbb{R}^u$ satisfying

$$\partial_U \mathcal{L}_H(U^{(H)}(\mathcal{D}, \bar{U}), \mathcal{D}, t, x) = 0 \quad \forall \quad \mathcal{D} \in \mathcal{O}, \quad (3.4a)$$

and

$$\partial_U \mathcal{L}_H(U^*, 0, t, x) = 0 \text{ admits the unique solution } U^* = \bar{U}(t, x). \quad (3.4b)$$

³As an example, in [24, Sec. 5.4] piecewise-in-time solutions of Burgers with small viscosity are used as base states to recover entropy solutions to inviscid Burgers, when written in H-J form.

We refer to such a function $U^{(H)}$ as a *Dual-to-Primal* (DtP) mapping. Assume that a *dual-to-primal* (DtP) “change of variables” mapping $U^{(H)}$ exists. Given D that

- (1) satisfies $\mathcal{D}(t, x) \subset \mathcal{O}(t, x)$,
- (2) has a well-defined trace on the space-time boundary,
- (3) satisfies the boundary constraints on the parts of the space-time boundary *complementary* to those that appear explicitly in (3.1):

$$D_\Gamma = D_\Gamma^* \quad \text{for } t = T \quad \text{or } x \notin \partial\Omega_\Gamma, \quad \Gamma = 1, \dots, d, \quad (3.5)$$

where $D^* : [0, T] \times \bar{\Omega} \rightarrow \mathbb{R}^d$ is a fixed, sufficiently regular, *arbitrarily chosen* function satisfying $\mathcal{D}^*(t, x) \subset \mathcal{O}$,

we define the *dual* functional as

$$\begin{aligned} S_H[D; \bar{U}] &:= \hat{S}_H \left[U^{(H)}(\mathcal{D}, \bar{U}), D \right] \\ &= \int_0^T \int_\Omega \mathcal{L}_H \left(U^{(H)}(\mathcal{D}, \bar{U}), \mathcal{D}, \bar{U} \right) dt dx - \int_\Omega \mathcal{C}_{\Gamma I} U_I^{(0)}(x) D_\Gamma(x, 0) dx \\ &\quad + \sum_\Gamma \int_{\partial\Omega_\Gamma} \int_0^T B_{\Gamma j} D_\Gamma n_j dt d\mathcal{H}^{m-1}. \end{aligned} \quad (3.6)$$

Using the facts

- (1) (3.4a),
- (2) $\mathcal{L}_H$ is necessarily affine in its argument $\mathcal{D}$,
- (3) the variations δD vanish on the parts of the boundary where Dirichlet boundary conditions on the dual fields are specified, cf. (3.5),

it is easily verified that, formally, the Euler-Lagrange equations and side conditions of S_H (3.6) are given by the system (3.1), using the substitution $U \rightarrow U^{(H)}(\mathcal{D}, \bar{U})$, for *any* H that allows for the construction of $U^{(H)}$.

In the following, we make the choice $D^* = 0$ for simplicity.

Remark 3.1. The flexibility afforded by the arbitrary choice of D^* can be practically beneficial, especially when solving a problem with a fixed base state (when using a changing base state scheme, the choice $D^* = 0$ is most natural as discussed below). Computational experience with even linear problems like the heat and the linear transport equation in [23], [30] with $\bar{U} = 0$ and $D^* = 0$ shows the development of strong boundary layers and/or large gradients of dual approximations at domain corners. These arise from the demands on the dual solution placed by meeting the arbitrarily set dual Dirichlet boundary condition and the needs of satisfying the primal problem. An obvious solution for such a situation is to note the values of the dual fields in the interior of the domain corresponding to an approximate solution obtained with a first prescription of say $D^* = 0$, smoothly extrapolate, with small gradients, that ‘interior’ field to the boundaries, and then apply the values obtained on the ‘Dirichlet boundaries’ as the modified D^* boundary condition to obtain the exact solution or better approximation.

Before proceeding further, we first show that choices of $H(U, \bar{U})$ satisfying (3.4) are easily made⁴; for instance, suppose

$$H \text{ is a strictly convex function of } U \text{ attaining its unique minimum at } U = \bar{U}, \quad (3.7)$$

e.g.,

$$H(U, \bar{U}) = a \frac{1}{2} |U - \bar{U}|^2 + b \frac{1}{p} |U - \bar{U}|^p, \quad a, b, p \in \mathbb{R}; \quad a > 0; \quad b \geq 0; \quad p > 2.$$

Then, since $\mathcal{L}_H$ is necessarily affine in $\mathcal{D}$,

$$\mathcal{L}_H(U, 0, \bar{U}) = -H(U, \bar{U})$$

and

$$\partial_U \mathcal{L}_H(U^*, 0, \bar{U}) = -\partial_U H(U^*, \bar{U}) = 0$$

admits the unique solution $U^* = \bar{U}$ and

$$-\partial_{UU} \mathcal{L}_H(\bar{U}, 0, \bar{U}) > 0 \quad (3.8)$$

in the sense of symmetric, positive-definite matrices.

By the implicit function theorem⁵, there exists a (w.l.o.g. convex) neighborhood $\mathcal{O} = \mathcal{O}(t, x) \subset \mathbb{R}^l$ containing $\mathcal{D} = 0$ and a mapping $U^{(H)}(\cdot, \bar{U}) : \mathcal{O} \rightarrow \mathbb{R}^u$ satisfying (3.4a), for each fixed choice of t, x and $\bar{U} \in \mathbb{R}^u$.

We also make the following observation, cf. [14]. Let us define

$$\begin{aligned} \tilde{S}_H[D; \bar{U}] &= \sup_U \hat{S}_H[U, D] \\ &= \sup_U \int_{\Omega} \int_0^T \mathcal{L}_H(U(t, x), \mathcal{D}(t, x), \bar{U}(t, x)) \, dt dx + \text{space-time boundary terms from (3.6)} \\ &= \int_{\Omega} \int_0^T \sup_U \mathcal{L}_H(U, \mathcal{D}(t, x), \bar{U}(t, x)) \, dt dx + \text{space-time boundary terms from (3.6)} \end{aligned}$$

(noting that $\mathcal{L}_H$ depends on the field U only through its pointwise values). Since $\hat{S}$ is affine⁶ in D , the functional $\tilde{S}_H[\cdot; \bar{U}]$ is *convex* in the Hilbert space

$$\mathcal{H} := L^2(\Omega \times (0, T); \mathbb{R}^d).$$

Since the set of positive-definite matrices is open in $\mathbb{R}^{u \times u}$, using (3.8) we infer that for a.e. $(t, x) \in [0, T] \times \Omega$ there exists a $\mathbb{R}^l \times \mathbb{R}^u$ -neighborhood centered in $(0, \bar{U}(t, x))$ in which $\mathcal{L}_H(\cdot, \mathcal{D}(t, x), \bar{U}(t, x))$ is strictly concave in U w.r.t. its first argument. Consequently, for

⁴From now on, we employ an abuse of notation and write $H(U, \bar{U})$, $\mathcal{L}_H(U, \mathcal{D}, \bar{U})$ instead of $H(U, t, x)$, $\mathcal{L}_H(U, \mathcal{D}, t, x)$, etc.

⁵Remember that here we work at a formal level and ignore regularity issues.

⁶Rigorously speaking, $\hat{S}_H[U, \cdot]$ has a certain convex domain in $\mathcal{H}$ where $\mathcal{D}$ and the boundary traces are well-defined, and the Dirichlet boundary condition (3.5) holds; outside this domain, we simply set $\hat{S}_H[U, D] := +\infty$; this is enough to guarantee the convexity of $\tilde{S}_H$.

a given base state $\bar{U} : [0, T] \times \Omega \rightarrow \mathbb{R}^u$, there exists a convex set $\mathcal{O}_{\bar{U}}^* \subset \mathcal{H}$ containing zero such that for every $D \in \mathcal{O}_{\bar{U}}^*$ one has

$$\sup_U \mathcal{L}_H(U, \mathcal{D}(t, x), \bar{U}(t, x)) = \mathcal{L}_H\left(U^{(H)}(\mathcal{D}(t, x), \bar{U}(t, x)), \mathcal{D}(t, x), \bar{U}(t, x)\right), \quad (3.9)$$

for a.e. $(t, x) \in [0, T] \times \Omega$

and

$$S_H[D; \bar{U}] = \tilde{S}_H[D; \bar{U}]. \quad (3.10)$$

In other words, for each fixed base state $\bar{U} : [0, T] \times \Omega \rightarrow \mathbb{R}^u$ there exists a convex set $\mathcal{O}_{\bar{U}}^* \subset \mathcal{H}$ in which the dual functional $S_H[D; \bar{U}]$ is convex and its Euler-Lagrange equations and side conditions are given by (3.1), with the substitution $U(t, x) \rightarrow U^{(H)}(\mathcal{D}(t, x), \bar{U}(t, x))$. The DtP map $(\mathcal{D}, \bar{U}) \mapsto U^{(H)}(\mathcal{D}, \bar{U})$ is well-defined for $D \in \mathcal{O}_{\bar{U}}^*$, with the maximizer in (3.10) being an isolated critical point of $\mathcal{L}_H(\cdot, \mathcal{D}(t, x); \bar{U}(t, x))$.

It is this insight that we use next to propose a convex gradient flow technique to attempt to find weak solutions to (3.1), utilizing minimizers of (3.10) for a suitable H **designed by the scheme**.

Remark 3.2. In [29] the minimizers of $\tilde{S}_H$ were called *variational dual solutions* and the ones among these that allow the recovery of a weak solution of the original PDE via the DtP map were defined as *dual solutions*. A meaningful choice of the auxiliary potential H is essential for the idea of variational dual solutions (as defined in [29]) to be non-vacuous, and they can differ from dual solutions, as happens, in particular, in the presence of a duality gap.

Before proposing the gradient flow scheme for our problem, we make a few further observations:

- (1) The “ L^2 -variation” $\frac{\delta S_H}{\delta D}[D; \bar{U}]$ is **formally** (by the envelope theorem) defined by

$$\begin{aligned} & \int_{\Omega} \int_0^T \frac{\delta S_H}{\delta D}[D(t, x); \bar{U}(t, x)] \delta D(t, x) dt dx \\ &= \int_{\Omega} \int_0^T (-\mathcal{C}_{\Gamma I} U_I \partial_t \delta D_{\Gamma} - \mathcal{F}_{\Gamma j}|_U \partial_{x_j} \delta D_{\Gamma} + (G_{\Gamma}|_U + V_{\Gamma}|_{(t, x)}) \delta D_{\Gamma}) dt dx \\ & \quad - \int_{\Omega} \mathcal{C}_{\Gamma I} U_I^{(0)}(x) \delta D_{\Gamma}(x, 0) dx + \sum_{\Gamma} \int_{\partial \Omega_{\Gamma}} \int_0^T B_{\Gamma j} \delta D_{\Gamma} n_j dt d\mathcal{H}^{m-1}, \end{aligned} \quad (3.11)$$

with the substitution $U(t, x) = U^{(H)}(\mathcal{D}(t, x), \bar{U}(t, x))$.

- (2) We note that

$$\frac{\delta S_H}{\delta D}[D; \bar{U}] \quad \text{depends on (the function) } D \text{ only through } U^{(H)}(\mathcal{D}, \bar{U}). \quad (3.12)$$

- (3) Solving system (3.1) in a weak sense means solving

$$\int_{\Omega} \int_0^T \frac{\delta S_H}{\delta D}[D(t, x); \bar{U}(t, x)] \delta D(t, x) dt dx = 0$$

for D satisfying the Dirichlet BC (3.5) and for all δD satisfying homogeneous Dirichlet BC at the complementary boundary defined in (3.5).

- (4) Suppose $\bar{U}$ is a weak solution to (3.1). Then $D = 0$ is a solution to the dual problem with $D^* = 0$, and defines, through the DtP map, a global in time solution to the primal problem. Thus, the intuition is that if $\bar{U}$ is “close” to a primal solution then, given the convexity properties of the dual problem, it is reasonable to expect to obtain a primal solution (through the DtP map), to which the base state is close, by solving the dual problem.

3.2. The scheme: Gradient flow evolutions in stages.

- (1) We think of a (fake, time-like) variable $\mathbf{s}$ along which a gradient descent in the Hilbert space $\mathcal{H}$ of the functional S_{H_k} is executed in stages indexed by k . Here, H_k is the auxiliary potential used in the stage k . A typical choice is

$$H_k(U) = a \frac{1}{2} |U - \bar{U}^k|^2 + b \frac{1}{p} |U - \bar{U}^k|^p, \quad a, b, p \in \mathbb{R}; \quad a > 0; \quad b \geq 0; \quad p > 2,$$

where the sequence of the base states $\bar{U}^k(t, x)$ will be described below.

- (2) Each stage comprises the interval $[0, \tilde{\mathbf{s}}_k^*)$, where $\tilde{\mathbf{s}}_k^*$ could possibly be $+\infty$, but will be otherwise determined by the scheme as follows.
- (3) We consider dual unknown functions $D^k : [0, T] \times \Omega \times [0, \tilde{\mathbf{s}}_k^*) \rightarrow \mathbb{R}^d$, $(t, x, \mathbf{s}) \mapsto D^k(t, x, \mathbf{s})$, and corresponding variations (test functions) δD to execute the following gradient flow:

for all $\mathbf{s} \in [0, \tilde{\mathbf{s}}_k^*]$ find D^k satisfying the dual boundary conditions by solving, for all δD satisfying their respective homogeneous boundary conditions,

$$\begin{aligned} \int_{[0, T] \times \Omega} \delta D(t, x) \frac{\partial D^k}{\partial \mathbf{s}}(t, x, \mathbf{s}) dt dx \\ = - \int_{[0, T] \times \Omega} \frac{\delta S_{H_k}}{\delta D}[D^k; \bar{U}^k](t, x, \mathbf{s}) \delta D(t, x) dt dx. \end{aligned} \quad (3.13)$$

- (4) Remember that the dissipation of the convex driving energies is non-increasing along such gradient flows (see, e.g., [13]):

$$\frac{d}{d\mathbf{s}} \int_{[0, T] \times \Omega} \left| \frac{\delta S_{H_k}}{\delta D} \right|^2 [D^k; \bar{U}^k](t, x, \mathbf{s}) dt dx \leq 0. \quad (3.14)$$

- (5) We also denote

$$U^{H_k}(t, x, \mathbf{s}) := U^{(H_k)}(\mathcal{D}^k(t, x, \mathbf{s}), \bar{U}^k(t, x))$$

provided the right-hand side is well-defined.

- (6) For $k = 1$, we select the base state $\bar{U}^1(t, x)$ arbitrarily, noting that the closer such a base state is to a solution, the better the chances of the scheme to converge to a solution.
- (7) We now run the first gradient descent (3.13) emanating from $D^1(t, x, 0) := 0$. Here and below, recall that the dual boundary data is defined by $D^* = 0$.

- (8) Then, either the gradient flow equilibrates, possibly at $\tilde{s}_1^* = +\infty$, without violating

$$\begin{aligned} \mathcal{D}^1(t, x, s) &\in \mathcal{O}(t, x), \\ -\partial_{UU} \mathcal{L}_H \left(U^{(H)}(\mathcal{D}^1(t, x, s), \bar{U}^1(t, x, s)), \mathcal{D}^1(t, x, s), \bar{U}^1(t, x) \right) &> 0, \end{aligned} \quad (3.15)$$

for a.e. $(t, x) \in [0, T] \times \Omega$ and for all $s \in [0, \tilde{s}_1^*]$, so that, as a consequence,

$$\begin{aligned} \sup_U \mathcal{L}_H(U, \mathcal{D}^1(t, x, s), \bar{U}^1(t, x)) \\ = \mathcal{L}_H \left(U^{(H)}(\mathcal{D}^1(t, x, s), \bar{U}^1(t, x, s)), \mathcal{D}^1(t, x, s), \bar{U}^1(t, x) \right), \end{aligned}$$

or the flow reaches a fake time $\tilde{s}_1^*$ when (3.15) is violated on a set of positive Lebesgue measure in $[0, T] \times \Omega$.

- (9) If the first alternative holds, then we can apply the DtP map to $D^1(t, x, \tilde{s}_1^*)$ and generate a solution to (3.1) by

$$U(t, x) := U^{H_1}(t, x, \tilde{s}_1^*) = U^{(H_1)}(\mathcal{D}^1(t, x, \tilde{s}_1^*), \bar{U}^1(t, x)). \quad (3.16)$$

Remark 3.3. In cases where the solutions to (3.1) are not unique, we can find only one of them. However, by varying the initial base state $\bar{U}^1$, it is expected to be possible to capture additional solutions as well. In principle, any solution can be retrieved, using observation (3) at the end of Sec. 3.1. In practice, see the approximation of a) a wide variety of traveling waves of a dispersive semi-discrete Burgers equation in [25] (with an algorithm which does not exploit the feature (3.14)), and b) even unstable solutions of the primal problem in a stable manner in [29, Sec. 6.3.4] that deals with the nonconvex elastodynamics of a bar (utilizing only one base state and stage).

- (10) If the second alternative holds and $\tilde{s}_1^*$ is finite, we define $s_1^* := \tilde{s}_1^* - \nu$, for some user-defined tolerance $\nu > 0$. Moreover, if $\tilde{s}_1^* = +\infty$ — which means that the limiting $D(\tilde{s}_1^*)$ lies outside of the “DtP zone” (3.15) although the trajectory $D(s)$ stays inside the “DtP zone” — we define $s_1^* := \mu$, for some user-defined big number $\mu > 0$.

- (11) Then we set

$$\begin{aligned} \bar{U}^2(t, x) &:= U^{H_1}(t, x, s_1^*) = U^{(H_1)}(\mathcal{D}^1(t, x, s_1^*), \bar{U}^1(t, x)) \quad \text{for all } s \in [0, \tilde{s}_2^*], \\ D^2(t, x, 0) &:= 0. \end{aligned} \quad (3.17)$$

- (12) We now execute the second gradient flow to find D^2 . By (3.4b) and (3.17) we always have

$$U^{H_2}(t, x, 0) = \bar{U}^2(t, x) = U^{H_1}(t, x, s_1^*),$$

i.e., the new U -trajectory is a continuation of the old one.

- (13) We continue the iterations until equilibration is achieved at some stage $k_* \in \mathbb{N}$, allowing us to generate a solution

$$U(t, x) := U^{H_{k_*}}(t, x, \tilde{s}_{k_*}^*).$$

(14) Since

$$U^{H_{k+1}}(t, x, 0) = \bar{U}^{k+1}(t, x) = U^{H_k}(t, x, \mathbf{s}_k^*),$$

it follows from (3.12) and (3.14) that the $\mathcal{H}$ -norm of the gradients $\|\frac{\delta S_{H_k}}{\delta D}\|_{\mathcal{H}}$, i.e., the dissipation of the energies does not increase with the course of fake-time, **including the moments of switching between the stages** (with the goal of eventually approaching zero dissipation via the gradient flow).

(15) It remains to prove that one has equilibration after a finite number of steps, i.e., $k_* \in \mathbb{N}$. For analytical purposes $k_* \rightarrow +\infty$ is acceptable, and for computational purposes reaching the threshold $\|\frac{\delta S_{H_k}}{\delta D}\|_{\mathcal{H}} \leq \tau$, for some user-defined tolerance ($0 < \tau \ll 1$), in a finite number of stages is sufficient.

Remark 3.4. It is natural to expect that a starting base state in time-dependent problems can be more effective if the real-time interval $[0, T]$ is divided into smaller disjoint sub-intervals whose closed union covers $[0, T]$, with the dual problem solved by the above gradient flow (in fake time) in each such sub-interval, defined with the real-time initial condition generated by the solution of the previous sub-interval. Thus, in this strategy, base states are switched both in real-time (to provide constant-in-real-time starting base state guesses for each sub-interval, at a minimum), as well as potentially in fake-time for the gradient flow within each sub-interval of real time. In experience with nonlinear transient problems computed without the gradient flow strategy [23, Sec. 5.3]-[24], such switching of base states across sub-intervals in real-time was found to be essential for success (as well as computational efficiency). On the other hand, [25] involves a time-independent problem where switching base states between stages of Newton method based iterations (the analog of the gradient flow stages) was found to be essential.

4. A TOY EXAMPLE OF THE GRADIENT FLOW PERFORMANCE

In this section, we illustrate (using a very simple model) how the switching of the base states, the convergence of the scheme, and the eventual generation of a solution occur in practice.

Consider the following *algebraic* system of equations (so that there is no dependence on t and x and no integration) for the unknown vector $U = (U_1, U_2) \in \mathbb{R}^2$:

$$\mathfrak{Q}_c(U) = 0, \tag{4.1}$$

where

$$\mathfrak{Q}_c(U)_i := (U_1 - c)(U_2 - c) - (U_i - c), \quad i = 1, 2. \tag{4.2}$$

Here c is a scalar parameter (in what follows, we tacitly assume $c \neq -\frac{1}{2}$; the special case $c = -\frac{1}{2}$ is addressed in Remark 4.1). It is obvious that (4.1) has two solutions:

$$(U_1, U_2) = (c, c), \quad (U_1, U_2) = (c + 1, c + 1).$$

Let us see how our gradient flow scheme finds one of them.

For any given base state $\bar{U} \in \mathbb{R}^2$, define the auxiliary potential

$$H(U) = H(U, \bar{U}) := \frac{1}{2}|U - \bar{U}|^2. \tag{4.3}$$

Mimicking (3.2), define the pre-dual functional by forming the scalar product of (4.1) with the dual variable $D = (D_1, D_2) \in \mathbb{R}^2$:

$$\mathcal{L}_H(U, D, \bar{U}) := \widehat{S}_H[U, D] := (U_1 - c)(U_2 - c)(D_1 + D_2) - (U_1 - c)D_1 - (U_2 - c)D_2 - H(U). \quad (4.4)$$

Taking the variation w.r.t. U , we observe that the DtP map $U^{(H)}$ is defined (provided $|D_1 + D_2| \neq 1$) by the formula $U^{(H)}(D) = u$, where $u = (u_1, u_2)$ is the unique solution of the linear system

$$(u_2 - c)(D_1 + D_2) - D_1 - (u_1 - \bar{U}_1) = 0, \quad (u_1 - c)(D_1 + D_2) - D_2 - (u_2 - \bar{U}_2) = 0. \quad (4.5)$$

Moreover,

$$-\partial_{UU}\mathcal{L}_H(U, D, \bar{U}) = \left(\frac{1}{-D_1 - D_2} \middle| \frac{-D_1 - D_2}{1} \right). \quad (4.6)$$

The “DtP zone” where the evolution of the gradient flow scheme should occur is determined by the following condition, cf. (3.15),

$$-\partial_{UU}\mathcal{L}_H(U^{(H)}(D), D, \bar{U}) > 0, \quad (4.7)$$

which in view of (4.6) is simply equivalent to

$$|D_1 + D_2| < 1. \quad (4.8)$$

For such D , we define

$$S_H[D] := \widehat{S}_H[U^{(H)}, D]. \quad (4.9)$$

This is actually a restriction of the convex functional

$$\widetilde{S}_H[D] := \sup_U \widehat{S}_H[U, D], \quad D \in \mathbb{R}^2,$$

to the “DtP zone” (4.8). A tedious calculation involving (4.5) shows that

$$S_H[D] = \frac{1}{2} \left(D - \bar{U} + \left(\frac{c}{c} \right) \right)^\top \left(\frac{1}{-D_1 - D_2} \middle| \frac{-D_1 - D_2}{1} \right)^{-1} \left(D - \bar{U} + \left(\frac{c}{c} \right) \right) - \frac{|c - \bar{U}|^2}{2},$$

whence S_H is strictly convex within its domain of definition. Moreover, it follows from the envelope theorem that

$$\nabla S_H(D) = \mathfrak{Q}_c(U^{(H)}(D)). \quad (4.10)$$

For definiteness (and somewhat echoing [14]), let our gradient flow scheme start from the zero base state

$$\bar{U}^1 = (0, 0) \quad \text{for all } \mathbf{s} \in [0, \tilde{\mathfrak{s}}_1^*],$$

where $\tilde{\mathfrak{s}}_1^*$ is the duration of the first stage (to be specified below). We run the gradient flow

$$\frac{d}{ds} D^1(\mathbf{s}) = -\nabla S_{H_1}[D^1(\mathbf{s})], \quad D^1(0) = 0 \quad (4.11)$$

— in our toy model the ambient Hilbert space is two-dimensional and the trajectory⁷ (D_1^1, D_2^1) of the gradient flow is determined by a pair of ODEs — starting from $(D_1^1, D_2^1)(0) =$

⁷Remember that the upper index indicates the number of the stage of the scheme and the lower indices denote the components of the trajectory.

$(0,0)$ until it either reaches the boundary of the “DtP zone” in (fake) time $\tilde{s}_1^*$ or equilibrates, possibly at $\tilde{s}_1^* = +\infty$. Since $\bar{U}_1^1 = \bar{U}_2^1$, by symmetry and strict convexity of S_{H_1} the straight line

$$D_1 = D_2 \quad (4.12)$$

is invariant for (4.11), and, since the initial condition belongs to it, the whole trajectory should also belong to it.

It follows from (4.5) that the DtP map restricted to the straight line (4.12) acts as

$$\left(U^{(H_1)}(d, d)\right)_i = \frac{c+d}{2d-1} + c, \quad i = 1, 2, \quad |2d| \neq 1. \quad (4.13)$$

Leveraging (4.10), we observe that the restricted gradient flow (4.11) reduces to

$$d'(\mathbf{s}) = \frac{c+d(\mathbf{s})}{2d(\mathbf{s})-1} - \frac{(c+d(\mathbf{s}))^2}{(2d(\mathbf{s})-1)^2}, \quad d(0) = 0, \quad (4.14)$$

and the trajectory should stay in the “DtP zone”

$$|2d(\mathbf{s})| < 1.$$

Letting

$$\tilde{d} := 2d - 1, \quad \tilde{c} := |2c + 1| > 0,$$

we rewrite this ODE in the form

$$\tilde{d}' = \frac{\tilde{d}^2 - \tilde{c}^2}{2\tilde{d}^2}, \quad \tilde{d}(0) = -1 \quad (4.15)$$

together with the constraint

$$-2 < \tilde{d} < 0. \quad (4.16)$$

Integrating (4.15), we obtain

$$2\tilde{d} + \tilde{c} \ln \left| \frac{\tilde{c} - \tilde{d}}{\tilde{c} + \tilde{d}} \right| = \mathbf{s} - 2 + \tilde{c} \ln \left| \frac{\tilde{c} + 1}{\tilde{c} - 1} \right|.$$

If $0 < \tilde{c} < 2$, $\tilde{c} \neq 1$, the solution $\tilde{d}$ satisfies (4.16). Hence,

$$\tilde{c} \ln \left| \frac{\tilde{c} - \tilde{d}}{\tilde{c} + \tilde{d}} \right| = \mathbf{s} - 2 + \tilde{c} \ln \left| \frac{\tilde{c} + 1}{\tilde{c} - 1} \right| - 2\tilde{d} \rightarrow +\infty$$

as $\mathbf{s} \rightarrow +\infty$, which implies that $\tilde{d}(\mathbf{s})$ converges to $\tilde{d}_\infty := -\tilde{c}$ with an explicit exponential rate. The corresponding attractor of (4.14) is

$$d_\infty := \frac{1}{2} - \left| c + \frac{1}{2} \right|.$$

The DtP map (4.13) applied to this limit generates the solution (c, c) if $c < -\frac{1}{2}$ and $(c+1, c+1)$ if $c > -\frac{1}{2}$ (i. e., it selects the one that is closer to the base state).

If $\tilde{c} = 1$ (i.e., $c = 0$ or $c = -1$), then the flow is constant: $\tilde{d}(\mathbf{s}) = -1$. Accordingly, $\tilde{s}_1^* = 0$ and $d(\tilde{s}_1^*) = 0$, and the DtP map generates the solution $(0, 0)$.

If $\tilde{c} \geq 2$, the trajectory reaches the value $\tilde{d} = -2$ that corresponds to $d = -\frac{1}{2}$ and lies on the boundary of the “DtP zone” at time

$$\tilde{s}_1^* := -2 + \tilde{c} \ln \left| \frac{\tilde{c} + 2}{\tilde{c} - 2} \right| - \tilde{c} \ln \left| \frac{\tilde{c} + 1}{\tilde{c} - 1} \right| \quad (4.17)$$

(with the convention $\tilde{s}_1^* = +\infty$ for $\tilde{c} = 2$). After stepping back to the user-defined moment s_1^* the value of $d(s_1^*)$ is still approximately equal to $-\frac{1}{2}$. We change the base state in accordance with (3.17), and set

$$\bar{U}^2 = (v, v) := \left(\frac{c + d(s_1^*)}{2d(s_1^*) - 1} + c, \frac{c + d(s_1^*)}{2d(s_1^*) - 1} + c \right) \approx \left(\frac{2c + 1}{4}, \frac{2c + 1}{4} \right). \quad (4.18)$$

We run the second stage of our scheme and observe that the trajectories of the second gradient flow

$$\frac{d}{ds} D^2(s) = -\nabla S_{H_2}[D^2(s)], \quad D^2(0) = 0, \quad (4.19)$$

are still on the straight line (4.12), for the same reason as before. Hence, the restricted DtP map acts as

$$\left(U^{(H_2)}(d, d) \right)_i = \frac{c + d - v}{2d - 1} + c, \quad i = 1, 2, \quad |2d| \neq 1, \quad (4.20)$$

with v defined by (4.18). Consequently, the new gradient flow is

$$d'(s) = \frac{c - v + d(s)}{2d(s) - 1} - \frac{(c - v + d(s))^2}{(2d(s) - 1)^2}, \quad d(0) = 0, \quad |2d(s)| < 1. \quad (4.21)$$

Letting

$$\tilde{d} := 2d - 1, \quad \tilde{c} := |2(c - v) + 1|,$$

we obtain the ODE (4.15). We will later show that $\tilde{c} > 0$.

By the same reasoning as above, if

$$0 < |2(c - v) + 1| < 2, \quad (4.22)$$

including the trivial case $|2(c - v) + 1| = 1$, $d(s)$ converges to $d_\infty := \frac{1}{2} - |c - v + \frac{1}{2}|$ with an explicit exponential rate (or immediately). The DtP map (4.20) generates the solution (c, c) if $c - v < -\frac{1}{2}$ and $(c + 1, c + 1)$ if $c - v > -\frac{1}{2}$.

If $|2(c - v) + 1| \geq 2$, the trajectory reaches $d = -\frac{1}{2}$ at the boundary of the “DtP zone” at explicit time $\tilde{s}_2^*$ determined by (4.17) with the updated value of $\tilde{c}$.

We repeat the procedure in the same manner for several additional steps until we reach the equilibration zone (4.22) and obtain a solution by the DtP map. In what follows, we show that we only need a **finite number of steps** and that, at every step, the corresponding constants $\tilde{c}$ are strictly positive.

We first observe that the base states are defined by the reciprocal relation

$$\begin{aligned}\bar{U}^{k+1} = (v_{k+1}, v_{k+1}) &:= \left(\frac{c - v_k + d(\mathbf{s}_k^*)}{2d(\mathbf{s}_k^*) - 1} + c, \frac{c - v_k + d(\mathbf{s}_k^*)}{2d(\mathbf{s}_k^*) - 1} + c \right) \\ &\approx \left(\frac{2(c + v_k) + 1}{4}, \frac{2(c + v_k) + 1}{4} \right), \quad v_1 = 0. \quad (4.23)\end{aligned}$$

Thus,

$$v_k \approx \left(c + \frac{1}{2} \right) (1 - 2^{1-k}),$$

which obviously satisfies the second inequality in (4.22) for k large enough. It remains to prove that $|2(c - v_k) + 1| > 0$, i.e.,

$$c + \frac{1}{2} - v_k \neq 0.$$

Assume for definiteness that $c > -\frac{1}{2}$ (the opposite case is handled in a similar way). Let us prove by induction that

$$c + \frac{1}{2} - v_k \geq 2^{1-k} \left(c + \frac{1}{2} \right) > 0. \quad (4.24)$$

The case $k = 1$ is trivial, so we can assume that

$$c + \frac{1}{2} - v_{k-1} \geq 2^{2-k} \left(c + \frac{1}{2} \right).$$

By construction, $d(\mathbf{s}_{k-1}^*) \in (-\frac{1}{2}, 0)$, so

$$v_k = \frac{c + \frac{1}{2} - v_{k-1} + d(\mathbf{s}_{k-1}^*) - \frac{1}{2}}{2d(\mathbf{s}_{k-1}^*) - 1} + c < \frac{2(c + v_{k-1}) + 1}{4},$$

whence

$$c + \frac{1}{2} - v_k > \frac{1}{2} \left(c + \frac{1}{2} - v_{k-1} \right) \geq 2^{1-k} \left(c + \frac{1}{2} \right).$$

Remark 4.1. We deliberately excluded the case $c = -\frac{1}{2}$ because, when started from the zero base state — as we did — the scheme enters the classical “Buridan’s donkey” dilemma and does not evolve. Indeed, the corresponding DtP map (4.13) is identically zero, and the gradient flow (4.14) becomes

$$d'(\mathbf{s}) = \frac{1}{4}, \quad d(0) = 0. \quad (4.25)$$

At time $\tilde{\mathbf{s}}_1^* = 2$ the trajectory reaches $d = \frac{1}{2}$ at the boundary of the “DtP zone”, but at any $\mathbf{s}_1^* < 2$ the base state produced by the DtP map is $\bar{U}^2 = (0, 0)$, cf. (4.18), i.e., the new base state coincides with the old one. However, a minor perturbation of the zero base state creates a bias and resolves the issue, allowing the scheme to converge to one of the

solutions, as it did above⁸. This indicates that, in potential applications of the scheme to PDEs, certain base states should be avoided (e.g., to prevent excessive symmetry).

5. A BRIEF DESCRIPTION OF THE GRADIENT FLOW SCHEME FOR THE NOISE-FREE NASH SYSTEM

The noise-free Nash system in its Burgers-like formulation (2.2) formally reads

$$\partial_t v + \nabla \mathfrak{U}(v \otimes v) = 0, \quad v(0, x) = v_0(x) := \nabla \psi_*(x), \quad (5.1)$$

complemented with spatially-periodic boundary conditions. Thus, it fits into our general framework (3.1) (up to the boundary conditions that are now much easier to handle; note also that here we are denoting the unknown function for the primal problem by v instead of U and that $u = d = n = pN^2$).

We define the auxiliary potential H as in (4.3):

$$H(v) = \frac{1}{2}|v - \bar{v}|^2. \quad (5.2)$$

Then, denoting the dual variable by a , we observe that

$$\widehat{S}_H[v, a] = - \int_0^T [K_{v, \bar{v}}(t) + (v - v_0, E) + (v \otimes v, B)] dt, \quad (5.3)$$

where $B = \mathfrak{U}^* \operatorname{div} a$ and $E = \partial_t a$ as in Section 1. Accordingly, the “DtP zone” is determined by the condition

$$I + 2B > 0 \text{ a.e. in } (0, T) \times \Omega, \quad (5.4)$$

and the DtP map acts as in Theorem 2.4:

$$U^{(H)}(E, B) = (I + 2B)^{-1}(\bar{v} - E). \quad (5.5)$$

(Here it is appropriate to use the pair (E, B) instead of the full vector $\mathcal{D}$). Hence,

$$\begin{aligned} S_H[a; \bar{v}] &= \widehat{S}_H \left[U^{(H)}(E, B, \bar{v}), a \right] = -\mathfrak{C}(\bar{v}, v_0) \\ &- \int_0^T \left[(U^{(H)}(E, B, \bar{v}) - v_0, E - \bar{v}) + \frac{1}{2} \left((U^{(H)}(E, B, \bar{v})) \otimes (U^{(H)}(E, B, \bar{v})), I + 2B \right) \right] dt \\ &= \int_0^T \left[(v_0, E) - \frac{1}{2}(\bar{v}, \bar{v}) + \frac{1}{2} \left((E - \bar{v}) \otimes (E - \bar{v}), (I + 2B)^{-1} \right) \right] dt. \end{aligned}$$

This is a restriction of the convex functional

$$\widetilde{S}_H[a] := \sup_U \widehat{S}_H[v, a] = \int_0^T \left[(v_0, E) - \frac{1}{2}(\bar{v}, \bar{v}) \right] dt - \mathcal{K}(E - \bar{v}, B)$$

to the “DtP zone” (5.4).

The minimizers of $\widetilde{S}_H$ exist by Theorem 2.11 (but may lie outside of the “DtP zone”) and belong to the class (2.6) (in particular, the corresponding $a = \int_T^t E \in \mathcal{H}$).

⁸The same issue was discussed in [22, Sec. 2.2.2] from a different perspective (not involving a gradient flow) for the dual problem in the simple example of the intersection of a circle with a vertical line.

Consequently,

$$\int_0^T \left(\frac{\delta S_H}{\delta a} [a; \bar{v}], \delta a \right) dt = - \int_0^T [(v - v_0, \delta E) + (v \otimes v, \delta B)] dt \quad (5.6)$$

with $v = (I + 2B)^{-1}(\bar{v} - E)$.

For the first stage of the gradient flow scheme, we select an arbitrary first base state $\bar{v}^1(t, x)$. For example, we can take $\bar{v}^1(t, x) = v_0(x)$. We run the first gradient descent in the Hilbert space $\mathcal{H}$ for the unknown vector function $a^1(t, x, s)$:

$$\begin{aligned} & \int_0^T (\partial_s a^1, \delta a) dt \\ &= \int_0^T [(v^1 - v_0, \delta E) + (v^1 \otimes v^1, \delta B)] dt, \quad a^1(T, x, s) = 0, \quad a^1(t, x, 0) = 0, \end{aligned} \quad (5.7)$$

where

$$v^1(t, x, s) := (I + 2\mathfrak{U}^* \operatorname{div} a^1(t, x, s))^{-1}(\bar{v}^1(t, x) - \partial_t a^1(t, x, s)). \quad (5.8)$$

Here the variation $\delta a(t, x)$ should satisfy

$$\delta a(T, x) = 0.$$

If we test (5.7) with compactly supported δa and integrate by parts, we formally get the following nonlinear degenerate parabolic PDE system of n equations:

$$\partial_s a^1 = -\partial_t v^1 - \nabla \mathfrak{U}(v^1 \otimes v^1), \quad (5.9)$$

with v^1 defined by (5.8). Testing now with arbitrary δa , we infer that the boundary (and initial) conditions are

$$v^1(0, x, s) = v_0(x), \quad a^1(T, x, s) = 0, \quad a^1(t, x, 0) = 0. \quad (5.10)$$

The process is repeated until stage k_* at which we reach the equilibration zone or continued for infinite number of stages. In the first case, the corresponding

$$v(t, x) := v^{k_*}(t, x, \tilde{s}_{k_*}^*)$$

would allegedly be a solution to the Nash system (2.2). In the second case, we generate an infinite sequence of base states

$$\bar{v}^k(t, x) := v^{k-1}(t, x, s_{k-1}^*).$$

If this sequence has accumulation points in some reasonable topology (say, in the weak L^2 -topology), those points can be referred to as “generalized” solutions to the Nash system.

Acknowledgement. This work originated from discussions held at Carnegie Mellon University while DV was visiting AA; DV acknowledges the hospitality of CMU. DV’s work was supported by CMUC⁹ under FCT, grants UID/00324/2025 and UID/PRR/00324/2025, and by KAUST Research Funding (KRF), award ORFS-2024-CRG12-6430.3. AA thanks Marshall Slemrod for suggesting the Nash system for study.

⁹<https://doi.org/10.54499/UID/00324/2025>

REFERENCES

- [1] A. Acharya, “A dual variational principle for nonlinear dislocation dynamics,” *Journal of Elasticity*, vol. 154, no. 1, pp. 383–395, 2023.
- [2] A. Acharya, “Coupled dislocations and fracture dynamics at finite deformation: Model derivation, and physical questions,” *Journal of Materials Science: Materials Theory*, vol. 8, no. 1, p. 6, 2024.
- [3] A. Acharya, “A hidden convexity in continuum mechanics, with application to classical, continuous-time, rate-(in)dependent plasticity,” *Mathematics and Mechanics of Solids*, vol. 30(3), pp. 701–719, 2025.
- [4] A. Acharya, I. Fonseca, L. Ganedi, and K. Stinson, “Vector field models for nematic disclinations,” *Journal of Nonlinear Science*, vol. 33, no. 5, p. 85, 2023.
- [5] A. Acharya and J. Ginster, “A convex variational principle for the necessary conditions of classical optimal control,” *arXiv e-prints*, Feb. 2025. arXiv: [2502.15973](#).
- [6] A. Acharya, J. Ginster, and A. N. Sengupta, “Variational Dual Solutions of Chern-Simons Theory,” *arXiv e-prints*, Nov. 2024. arXiv: [2411.17635](#).
- [7] A. Acharya and A. N. Sengupta, “Action principles for dissipative, non-holonomic Newtonian mechanics,” *Proc. of the Royal Society A.*, vol. 480, no. 2293, Paper No. 20240113, 21, 2024.
- [8] A. Acharya and M. Slemrod, “Existence, uniqueness, and long-time behavior of linearized field dislocation dynamics,” *Quarterly of Applied Mathematics*, vol. LXXXI, pp. 247–258, 2023.
- [9] A. Acharya, B. Strohffolini, and A. Zarnescu, “Variational Dual Solutions for Incompressible Fluids,” *arXiv e-prints*, Sep. 2024. arXiv: [2409.04911](#).
- [10] A. Acharya and L. Tartar, “On an equation from the theory of field dislocation mechanics,” *Bollettino dell’Unione Matematica Italiana*, vol. 9, pp. 409–444, 2011.
- [11] O. Alvarez, P. Hoch, Y. L. Bouar, and R. Monneau, “Dislocation dynamics: Short-time existence and uniqueness of the solution,” *Archive for Rational Mechanics and Analysis*, vol. 181, no. 3, pp. 449–504, 2006.
- [12] A. Arora and A. Acharya, “Emergent fault friction and supershear in a continuum model of geophysical rupture,” *Journal of the Mechanics and Physics of Solids*, p. 105 827, 2024.
- [13] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces* (MOS-SIAM Series on Optimization), Second. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2014, vol. 17, pp. xii+793, Applications to PDEs and optimization.
- [14] Y. Brenier, “The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem,” *Comm. Math. Phys.*, vol. 364, no. 2, pp. 579–605, 2018.
- [15] Y. Brenier, “Examples of hidden convexity in nonlinear pdes,” *Preprint*, 2020.
- [16] Y. Brenier and I. Moyano, “Relaxed solutions for incompressible inviscid flows: A variational and gravitational approximation to the initial value problem,” *Philos. Trans. Roy. Soc. A*, vol. 380, no. 2219, Paper No. 20210078, 12, 2022.

- [17] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, *The master equation and the convergence problem in mean field games* (Annals of Mathematics Studies). Princeton University Press, Princeton, NJ, 2019, vol. 201, pp. x+212.
- [18] N. Ghoussoub, *Self-dual partial differential systems and their variational principles*. Springer, 2009.
- [19] H. Gimperlein, M. Grinfeld, R. J. Knops, and M. Slemrod, “On action rate admissibility criteria,” *arXiv e-prints*, 2025. arXiv: [2503.03491](#).
- [20] H. Gimperlein, M. Grinfeld, R. J. Knops, and M. Slemrod, “The least action admissibility principle,” *Archive for Rational Mechanics and Analysis*, vol. 249, no. 2, p. 22, 2025.
- [21] R. Hynd, “The Hamilton-Jacobi equation, then and now,” *Notices Amer. Math. Soc.*, vol. 68, no. 9, pp. 1457–1467, 2021.
- [22] U. Kouskiya, “Computational analysis of a duality based approach for solving ordinary and partial differential equations,” Ph.D. dissertation, Carnegie Mellon University, 2024.
- [23] U. Kouskiya and A. Acharya, “Hidden convexity in the heat, linear transport, and Euler’s rigid body equations: A computational approach,” *Quarterly of Applied Mathematics*, vol. LXXXII, pp. 673–703, 2024.
- [24] U. Kouskiya and A. Acharya, “Inviscid Burgers as a degenerate elliptic problem,” *Quarterly of Applied Mathematics*, vol. LXXXIII, pp. 315–360, 2025.
- [25] U. Kouskiya, R. L. Pego, and A. Acharya, “Traveling wave profiles for a semi-discrete Burgers equation,” *Physica D*, p. 134 961, 2025.
- [26] J.-M. Mirebeau and E. Stampfli, “Discretization and convergence of the ballistic Benamou-Brenier formulation of the porous medium and Burgers equations,” *arXiv e-prints*, Nov. 2025, Art. no. arXiv:2511.02662. arXiv: [2511.02662](#).
- [27] L. Morin and A. Acharya, “Analysis of a model of field crack mechanics for brittle materials,” *Computer Methods in Applied Mechanics and Engineering*, vol. 386, p. 114 061, 2021.
- [28] F. Rindler, “Energetic solutions to rate-independent large-strain elasto-plastic evolutions driven by discrete dislocation flow,” *Journal of the European Mathematical Society*, vol. 27, no. 7, pp. 2795–2864, 2025.
- [29] S. Singh, J. Ginster, and A. Acharya, “A hidden convexity of nonlinear elasticity,” *Journal of Elasticity*, vol. 156, no. 3, pp. 975–1014, 2024.
- [30] N. Sukumar and A. Acharya, “Variational formulation based on duality to solve partial differential equations: Use of B-splines and machine learning approximants,” *Computer Methods in Applied Mechanics and Engineering*, vol. 441, p. 117 909, 2025.
- [31] D. Vorotnikov, “Partial differential equations with quadratic nonlinearities viewed as matrix-valued optimal ballistic transport problems,” *Arch. Ration. Mech. Anal.*, vol. 243, no. 3, pp. 1653–1698, 2022.
- [32] D. Vorotnikov, “Dafermos’ principle and Brenier’s duality scheme for defocusing dispersive equations,” *arXiv e-prints*, Jan. 2025. arXiv: [2501.05389](#).
- [33] C. Zhang et al., “A non-traditional view on the modeling of nematic disclination dynamics,” *Quarterly of Applied Mathematics*, vol. 75, no. 2, pp. 309–357, 2017.

- [34] X. Zhang, A. Acharya, N. J. Walkington, and J. Bielak, “A single theory for some quasi-static, supersonic, atomic, and tectonic scale applications of dislocations,” *Journal of the Mechanics and Physics of Solids*, vol. 84, pp. 145–195, 2015.

(D. Vortnikov) UNIVERSITY OF COIMBRA, CMUC, DEPARTMENT OF MATHEMATICS, 3000-143 COIMBRA, PORTUGAL

Email address: `mitvorot@mat.uc.pt`

(A. Acharya) DEPARTMENT OF CIVIL & ENVIRONMENTAL ENGINEERING AND CENTER FOR NONLINEAR ANALYSIS, CARNEGIE MELLON UNIVERSITY, PITTSBURGH, 15213, PA, USA

Email address: `acharyaamit@cmu.edu`