Infinite-dimensional Hamilton-Jacobi equations for statistical inference on sparse graphs

Tomas Dominguez* Jean-Christophe Mourrat†

Abstract

We study the well-posedness of an infinite-dimensional Hamilton-Jacobi equation posed on the set of non-negative measures and with a monotonic non-linearity. Our results will be used in a companion work to propose a conjecture and prove partial results concerning the asymptotic mutual information in the assortative stochastic block model in the sparse regime. The equation we consider is naturally stated in terms of the Gateaux derivative of the solution, unlike previous works in which the derivative is usually of transport type. We introduce an approximating family of finite-dimensional Hamilton-Jacobi equations, and use the monotonicity of the non-linearity to show that no boundary condition needs to be prescribed to establish well-posedness. The solution to the infinite-dimensional Hamilton-Jacobi equation is then defined as the limit of these approximating solutions. In the special setting of a convex non-linearity, we also provide a Hopf-Lax variational representation of the solution.

1 Introduction and main results

A recent approach to describe the asymptotic free energy of a mean-field disordered system is as the solution to a Hamilton-Jacobi equation. Spin-glass models have typically led to infinite-dimensional Hamilton-Jacobi equations of transport type [27, 29, 30] while statistical inference problems have given rise to finite-dimensional Hamilton-Jacobi equations defined on closed convex cones [7, 8, 10, 11, 26, 28]. A general well-posedness theory for the former was established in [12] while one for the latter was developed in [13]. In [19] we will propose to describe the asymptotic mutual information in the sparse stochastic block model in terms of a Hamilton-Jacobi equation posed over a space of probability measures, but featuring derivatives of "affine" rather than transport type. The purpose of this paper is to develop a well-posedness theory for such an infinite-dimensional Hamilton-Jacobi equation; we expect that this type of equation will appear in other mean-field problems with sparse interactions. While our setting is different, the techniques we use here draw heavily upon the arguments introduced in [12] and [13].

Let us describe the class of infinite-dimensional Hamilton-Jacobi equations that we consider. We denote by $\mathscr{M}_s$ the space of signed measures on $[-1,1]$,

$$\mathscr{M}_s = \big\{\mu \mid \mu \text{ is a signed measure on } [-1,1]\big\}, \tag{1.1}$$

and by $\mathscr{M}_+$ the cone of non-negative measures on this interval,

$$\mathscr{M}_+ = \big\{\mu \in \mathscr{M}_s \mid \mu \text{ is a non-negative measure}\big\}. \tag{1.2}$$

*Department of Mathematics, University of Toronto, tomas.dominguezchiozza@mail.utoronto.ca

†Department of Mathematics, ENS Lyon and CNRS, jean-christophe.mourrat@ens-lyon.fr

We follow the convention that a signed measure can only take finite values, and in particular, every $\mu \in \mathscr{M}_+$ must have finite total mass. We fix a continuously differentiable function $g : [-1,1] \to \mathbb{R}$, and for each measure $\mu \in \mathscr{M}_+$, define the function $G_\mu : [-1,1] \to \mathbb{R}$ by

$$G_\mu(x) = \int_{-1}^{1} g(xy)\,\mathrm{d}\mu(y). \tag{1.3}$$

We introduce the cone of functions

$$\mathscr{C}_\infty = \big\{ G_\mu \mid \mu \in \mathscr{M}_+ \big\} \tag{1.4}$$

as well as the non-linearity $\mathsf{C}_\infty : \mathscr{C}_\infty \to \mathbb{R}$ defined on this cone by

$$\mathsf{C}_\infty(G_\mu) = \frac{1}{2}\int_{-1}^{1} G_\mu(x)\,\mathrm{d}\mu(x) = \frac{1}{2}\int_{-1}^{1}\int_{-1}^{1} g(xy)\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x). \tag{1.5}$$

This non-linearity is well-defined by the Fubini-Tonelli theorem. Indeed, if $G_\mu = G_\nu$ for some measures $\mu, \nu \in \mathscr{M}_+$, then

$$\int_{-1}^{1} G_\mu(x)\,\mathrm{d}\mu(x) = \int_{-1}^{1} G_\nu(x)\,\mathrm{d}\mu(x) = \int_{-1}^{1}\int_{-1}^{1} g(xy)\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(x), \tag{1.6}$$

while

$$\int_{-1}^{1} G_\nu(x)\,\mathrm{d}\nu(x) = \int_{-1}^{1} G_\mu(x)\,\mathrm{d}\nu(x) = \int_{-1}^{1}\int_{-1}^{1} g(xy)\,\mathrm{d}\mu(y)\,\mathrm{d}\nu(x), \tag{1.7}$$

and the symmetry of the map $(x,y) \mapsto g(xy)$ implies that these two expressions coincide. Given a function $f : [0,\infty) \times \mathscr{M}_+ \to \mathbb{R}$ and measures $\mu, \nu \in \mathscr{M}_+$, we denote by $D_\mu f(t,\mu;\nu)$ the Gateaux derivative of the function $f(t,\cdot)$ at the measure μ in the direction ν,

$$D_\mu f(t,\mu;\nu) = \lim_{\varepsilon \to 0} \frac{f(t,\mu+\varepsilon\nu) - f(t,\mu)}{\varepsilon}. \tag{1.8}$$

We will say that the Gateaux derivative of $f(t,\cdot)$ at the measure $\mu \in \mathscr{M}_+$ admits a density if there exists a bounded measurable function $x \mapsto D_\mu f(t,\mu,x)$ defined on the interval $[-1,1]$ with

$$D_\mu f(t,\mu;\nu) = \int_{-1}^{1} D_\mu f(t,\mu,x)\,\mathrm{d}\nu(x) \tag{1.9}$$

for every measure $\nu \in \mathscr{M}_+$. We will often abuse notation and identify the density $D_\mu f(t,\mu,\cdot)$ with the Gateaux derivative $D_\mu f(t,\mu)$. The purpose of this paper is to establish the well-posedness of the infinite-dimensional Hamilton-Jacobi equation

$$\begin{cases} \partial_t f(t,\mu) = \mathsf{C}_\infty\big(D_\mu f(t,\mu)\big) & \text{on } \mathbb{R}_{>0} \times \mathscr{M}_+, \\ f(0,\mu) = \psi(\mu) & \text{on } \mathscr{M}_+. \end{cases} \tag{1.10}$$

under appropriate assumptions on the kernel g and the initial condition $\psi : \mathscr{M}_+ \to \mathbb{R}$. In particular, these assumptions will imply that, in a suitably weak sense, the Gateaux derivative of the solution

$D_\mu f(t,\mu)$ belongs to the cone $\mathscr{C}_\infty$ for all $t \geq 0$ and $\mu \in \mathscr{M}_+$. Before stating these assumptions and the results they lead to precisely, let us describe the general strategy we will follow.

To obtain the well-posedness of the Hamilton-Jacobi equation (1.10), we will project it from the infinite-dimensional space of measures $\mathscr{M}_+$ to a family of finite-dimensional spaces of measures $\mathscr{M}_+^{(K)}$ with dimension monotone in some integer parameter $K \geq 1$. The well-posedness of each of these projected equations will be obtained using techniques similar to those in [13], and the limit as K tends to infinity of these projected solutions will be shown to exist using techniques similar to those in [12, 27, 29]; we take the limit thus obtained as the definition of the solution to the infinite-dimensional Hamilton-Jacobi equation (1.10).

In previous works, derivatives of transport type were the primary focus of investigation, and it was thus natural to discretize the space of measures by restricting to measures of the form $K^{-1}\sum_{k=1}^K \delta_{x_k}$, only allowing the x_k's to vary but keeping the weight of each atom fixed. Due to the nature of the derivatives appearing in (1.10), we choose instead here to define our finite-dimensional approximating space as the cone of non-negative measures supported on dyadic rationals in the interval $[-1,1]$. That is, we allow the weights to vary, provided that they remain non-negative, but keep the positions of the atoms fixed. Given an integer $K \geq 1$, we write

$$\mathscr{D}_K = \left\{ k = \frac{i}{2^K} \mid -2^K \leq i < 2^K \right\} \tag{1.11}$$

for the set of dyadic rationals on $[-1,1]$ at scale K. It will be convenient to index vectors using the set of dyadic rationals, writing $x = (x_k)_{k\in\mathscr{D}_K} \in \mathbb{R}^{\mathscr{D}_K}$. We denote the set of discrete measures supported on the dyadic rationals at scale K in the interval $[-1,1]$ by

$$\mathscr{M}_+^{(K)} = \left\{ \mu \in \mathscr{M}_+ \mid \mu = \frac{1}{|\mathscr{D}_K|}\sum_{k\in\mathscr{D}_K} x_k\delta_k \text{ for some } x = (x_k)_{k\in\mathscr{D}_K} \in \mathbb{R}_{\geq 0}^{\mathscr{D}_K} \right\}. \tag{1.12}$$

A natural way to project a general measure $\mu \in \mathscr{M}_+$ onto $\mathscr{M}_+^{(K)}$ is via the mapping

$$x^{(K)}(\mu) = \big(|\mathscr{D}_K|\mu\left[k, k+2^{-K}\right)\big)_{k\in\mathscr{D}_K} \in \mathbb{R}_{\geq 0}^{\mathscr{D}_K}. \tag{1.13}$$

For $\mu \in \mathscr{M}_+^{(K)}$, the image of μ is simply the sequence of weights of the measure μ at each point in $\mathscr{D}_K$, up to multiplication by $|\mathscr{D}_K|$. The inverse of this mapping assigns to each $x \in \mathbb{R}_{\geq 0}^{\mathscr{D}_K}$ the measure

$$\mu_x^{(K)} = \frac{1}{|\mathscr{D}_K|}\sum_{k\in\mathscr{D}_K} x_k\delta_k \in \mathscr{M}_+^{(K)}. \tag{1.14}$$

We can use these projections to devise finite-dimensional approximations to the Hamilton-Jacobi equation (1.10). These will be posed on the cone $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}^{\mathscr{D}_K}$. Indeed, any real-valued function $f : [0,\infty) \times \mathscr{M}_+^{(K)} \to \mathbb{R}$ may be identified with the function

$$f^{(K)}(t,x) = f\big(t, \mu_x^{(K)}\big) \tag{1.15}$$

defined on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}^{\mathscr{D}_K}$. Moreover, the Gateaux derivative at the measure $\mu \in \mathscr{M}_+^{(K)}$ may be identified with the gradient $|\mathscr{D}_K|\nabla f^{(K)}(t, x^{(K)}(\mu))$ by duality. Indeed, for any direction $\nu \in \mathscr{M}_+^{(K)}$,

$$D_\mu f(t,\mu;\nu) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} f^{(K)}(t, x^{(K)}(\mu) + \varepsilon x^{(K)}(\nu)) = \nabla f^{(K)}(t, x^{(K)}(\mu)) \cdot x^{(K)}(\nu). \tag{1.16}$$

The additional factor of $|\mathscr{D}_K|$ appears because $x^{(K)}(\nu)$ has ℓ^1-norm $|\mathscr{D}_K|$ whenever ν is a probability measure. The corresponding initial condition becomes the function $\psi^{(K)} : \mathbb{R}^{\mathscr{D}_K}_{\geq 0} \to \mathbb{R}$ defined by

$$\psi^{(K)}(x) = \psi\big(\mu_x^{(K)}\big). \tag{1.17}$$

The cone (1.4) and the non-linearity (1.5) may be projected in a similar manner. We introduce the symmetric matrix

$$G^{(K)} = \frac{1}{|\mathscr{D}_K|^2}\big(g(kk')\big)_{k,k' \in \mathscr{D}_K} \in \mathbb{R}^{\mathscr{D}_K \times \mathscr{D}_K}, \tag{1.18}$$

and observe that for every $\mu \in \mathscr{M}_+^{(K)}$ and $k \in \mathscr{D}_K$,

$$G_\mu(k) = \sum_{k' \in \mathscr{D}_K} g(kk')\mu(k') = \frac{1}{|\mathscr{D}_K|}\sum_{k' \in \mathscr{D}_K} g(kk')x^{(K)}(\mu)_{k'} = |\mathscr{D}_K|\big(G^{(K)}x^{(K)}(\mu)\big)_k. \tag{1.19}$$

This motivates the definition of the projected cone,

$$\mathscr{C}_K = \Big\{G^{(K)}x^{(K)}(\mu) \in \mathbb{R}^{\mathscr{D}_K} \mid \mu \in \mathscr{M}_+^{(K)}\Big\} = \Big\{G^{(K)}x \in \mathbb{R}^{\mathscr{D}_K} \mid x \in \mathbb{R}^{\mathscr{D}_K}_{\geq 0}\Big\}, \tag{1.20}$$

and the projected non-linearity $\mathsf{C}_K : \mathscr{C}_K \to \mathbb{R}$ defined by

$$\mathsf{C}_K\big(G^{(K)}x\big) = \frac{1}{2}G^{(K)}x \cdot x = \frac{1}{2|\mathscr{D}_K|^2}\sum_{k,k' \in \mathscr{D}_K} g(kk')x_k x_{k'} = \mathsf{C}_\infty\big(G_{\mu_x^{(K)}}\big). \tag{1.21}$$

This projected non-linearity can be shown to be well-defined using the same argument that showed the non-linearity (1.5) was well-defined. With this notation, our finite-dimensional approximation of the Hamilton-Jacobi equation (1.10) reads

$$\partial_t f^{(K)}(t,x) = \mathsf{C}_K\big(\nabla f^{(K)}(t,x)\big) \quad \text{on} \quad \mathbb{R}_{>0} \times \mathbb{R}^{\mathscr{D}_K}_{\geq 0} \tag{1.22}$$

subject to the initial condition $f^{(K)}(0,x) = \psi^{(K)}(x)$ on $\mathbb{R}^{\mathscr{D}_K}_{\geq 0}$. To study this equation we will introduce an appropriate extension $\mathsf{H}_K : \mathbb{R}^{\mathscr{D}_K} \to \mathbb{R}$ of the non-linearity C_K, and instead consider the Hamilton-Jacobi equation

$$\partial_t f^{(K)}(t,x) = \mathsf{H}_K\big(\nabla f^{(K)}(t,x)\big) \quad \text{on} \quad \mathbb{R}_{>0} \times \mathbb{R}^{\mathscr{D}_K}_{>0} \tag{1.23}$$

subject to the initial condition $f^{(K)}(0,x) = \psi^{(K)}(x)$ on $\mathbb{R}^{\mathscr{D}_K}_{\geq 0}$. Notice that we have used the cone $\mathbb{R}^{\mathscr{D}_K}_{>0}$ as opposed to the more intuitive cone $\mathbb{R}^{\mathscr{D}_K}_{\geq 0}$. As detailed below, the monotonicity of the projected non-linearity (1.21) and the ideas regarding boundary conditions of Hamilton-Jacobi equations with suitable non-linearities developed in [13, 18, 31] make these two choices equivalent. In particular, it will not be necessary to endow the projected Hamilton-Jacobi equation with a boundary condition. Remembering that this Hamilton-Jacobi equation appears in the context of statistical inference makes this insight rather reassuring. Indeed, the statistical inference model does not suggest an obvious choice of boundary condition—given that we are ultimately interested in the identification of the value of the solution at a point in $\mathbb{R}_{\geq 0} \times \{0\}$, we would at least not want to use a Dirichlet boundary condition there! In earlier works, the imposition of a Neumann-type boundary condition was observed to be a workable option [27, 28, 29]. In [13], it was shown that this somewhat artificial choice is not necessary, and no boundary condition needs to be specified, because the non-linearity "points in the right direction".

We now state the precise assumptions that will allow us to obtain the well-posedness of the projected Hamilton-Jacobi equations and establish the convergence of their solutions. In the same spirit as [12, 13, 27, 29], we will need the initial conditions $\psi^{(K)}$ and ψ to satisfy a certain number of Lipschitz continuity assumptions. Given an integer $d \geq 1$, we introduce the normalized-ℓ^1 and normalized-$\ell^{1,*}$ norms, defined for every $x, y \in \mathbb{R}^d$ by

$$|||x|||_1 = \frac{1}{d}\sum_{k=1}^{d} |x_k| \quad \text{and} \quad |||y|||_{1,*} = \max_{k \leq d} d|y_k|. \tag{1.24}$$

The underlying dimension $d \geq 1$ will be kept implicit but will always be clear from the context. The normalized-ℓ^1 norm is meant to measure elements of $\mathbb{R}^{\mathscr{D}_k}$ with a scaling that is consistent with our identification of this space with the space of measures $\mathscr{M}_+^{(K)}$. The normalized-$\ell^{1,*}$ norm serves to measure elements of the dual space, and is defined so that the Hölder-type inequality $x \cdot y \leq |||x|||_1 \, |||y|||_{1,*}$ is valid.

The key continuity assumption on the projected initial condition $\psi^{(K)}$ that will make it possible to establish the well-posedness of the projected Hamilton-Jacobi equations will be Lipschitz continuity with respect to the normalized-ℓ^1 norm. Another way to encode this property is to require the initial condition $\psi : \mathscr{M}_+ \to \mathbb{R}$ to be Lipschitz continuous with respect to the total variation distance on $\mathscr{M}_+$,

$$\mathrm{TV}(\mu, \nu) = \sup\big\{|\mu(A) - \nu(A)| \mid A \text{ is a measurable subset of } [-1,1]\big\}. \tag{1.25}$$

The normalized-$\ell^{1,*}$ norm will play its part when discussing the Lipschitz continuity of the projected non-linearity (1.21). To determine the convergence of the projected solutions, it will be important to assume that the initial condition $\psi : \mathscr{M}_+ \to \mathbb{R}$ is Lipschitz continuous with respect to the Wasserstein distance on the set of probability measures $\Pr[-1,1]$,

$$W(\mathbb{P}, \mathbb{Q}) = \sup\left\{ \left| \int_{-1}^{1} h(x)\,\mathrm{d}\mathbb{P}(x) - \int_{-1}^{1} h(x)\,\mathrm{d}\mathbb{Q}(x) \right| \mid \|h\|_{\mathrm{Lip}} \leq 1 \right\}. \tag{1.26}$$

Here $\|\cdot\|_{\mathrm{Lip}}$ denotes the Lipschitz semi-norm

$$\|h\|_{\mathrm{Lip}} = \sup_{x \neq x' \in [-1,1]} \frac{|h(x) - h(x')|}{|x - x'|} \tag{1.27}$$

defined on the space of functions $h : [-1,1] \to \mathbb{R}$. The final assumption on the initial condition will ensure that, in a sense to be made precise, the solution to the projected Hamilton-Jacobi equation has a bounded gradient close to the projected cone $\mathscr{C}_K$ defined in (1.20). It would of course be more convenient to assume that the gradient really belongs to $\mathscr{C}_K$, rather than only being close to it, but unlike in earlier works, this stronger property does not hold in the context of the main application we have in mind in [19]. To impose the boundedness of the gradient, fix $a > 0$, and for each integer $K \geq 1$ introduce the closed convex set

$$\mathscr{K}_{a,K} = \left\{ G^{(K)}x \in \mathbb{R}^{\mathscr{D}_K} \mid x \in \mathbb{R}^{\mathscr{D}_K}_{\geq 0} \text{ and } |||x|||_1 \leq a \right\} \subset \mathscr{C}_K. \tag{1.28}$$

Given a closed convex set $\mathscr{K} \subset \mathbb{R}^d$, write

$$\mathscr{K}' = \mathscr{K} + B_{d^{-1/2}}(0) \tag{1.29}$$

for the neighborhood of radius $d^{-1/2}$ around $\mathscr{K}$ in the normalized-$\ell^{1,*}$ norm. Here

$$B_r(x) = \left\{x' \in \mathbb{R}^d \mid |||x' - x|||_{1,*} \leq r\right\} \tag{1.30}$$

denotes the closed ball of radius $r > 0$ centered around $x \in \mathbb{R}^d$ relative to the normalized-$\ell^{1,*}$ norm. We will say that a Lipschitz continuous function $h : \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ has its gradient in $\mathscr{K}$ if

$$\nabla h \in L^\infty\big(\mathbb{R}^d_{\geq 0}; \mathscr{K}\big). \tag{1.31}$$

Recall that a Lipschitz continuous function is differentiable almost everywhere by Rademacher's theorem (see Theorem 6 in Chapter 5.8 of [20]), so the spatial gradient ∇h is well-defined as an element of L^∞, and the condition (1.31) requires that this object take values in $\mathscr{K}$ almost everywhere. A non-differential criterion for the gradient of a Lipschitz continuous function to lie in a closed convex set is given in Proposition B.2, and will be used frequently throughout the paper. As will be shown below, assuming that the initial condition has its gradient in $\mathscr{K}'_{a,K}$ suffices to ensure that the gradient of the solution remains in this set at all times. Notice that this is insufficient to be able to evaluate the non-linearity C_K at the gradient of the solution; however, under suitable Lipschitz continuity properties of the extension H_K, it ensures that the projected Hamilton-Jacobi equation (1.23) should be an adequate replacement for the Hamilton-Jacobi equation (1.22). In particular, it justifies defining the solution to the infinite-dimensional Hamilton-Jacobi (1.10) as the limit of the solutions to the projected Hamilton-Jacobi equation (1.23). Besides some smoothness, the only constraint we will impose on the kernel $g : [-1,1] \to \mathbb{R}$ is that it be strictly positive. Among other things, this assumption ensures that a non-negative measure $\mu \in \mathscr{M}_+$ cannot have a large total mass unless the function G_μ takes large values. In summary, the assumptions on the kernel $g : [-1,1] \to \mathbb{R}$ and the initial condition $\psi : \mathscr{M}_+ \to \mathbb{R}$ required for the validity of our main results are the following.

H1 The kernel $g : [-1,1] \to \mathbb{R}$ is continuously differentiable and bounded away from zero by some positive constant $m > 0$,

$$g(x) \geq m. \tag{1.32}$$

H2 The initial condition $\psi : \mathscr{M}_+ \to \mathbb{R}$ is Lipschitz continuous with respect to the total variation distance (1.25),

$$|\psi(\mu) - \psi(\nu)| \leq \|\psi\|_{\mathrm{Lip,TV}} \mathrm{TV}(\mu, \nu) \tag{1.33}$$

for all measures $\nu, \mu \in \mathscr{M}_+$.

H3 There exists $a > 0$ such that the initial condition $\psi : \mathscr{M}_+ \to \mathbb{R}$ has the property that each of the projected initial conditions (1.17) has its gradient in the set $\mathscr{K}'_{a,K}$,

$$\nabla \psi^{(K)} \in L^\infty\big(\mathbb{R}^d_{\geq 0}; \mathscr{K}'_{a,K}\big). \tag{1.34}$$

H4 The initial condition $\psi : \Pr[-1,1] \to \mathbb{R}$ is Lipschitz continuous with respect to the Wasserstein distance (1.26),

$$|\psi(\mathbb{P}) - \psi(\mathbb{Q})| \leq \|\psi\|_{\mathrm{Lip},W} W(\mathbb{P},\mathbb{Q}) \tag{1.35}$$

for all probability measures $\mathbb{P},\mathbb{Q} \in \Pr[-1,1]$.

Observe that the hypothesis (**H2**) on the initial condition implies that the projected initial conditions (1.17) are Lipschitz continuous with respect to the normalized-ℓ^1 norm,

$$\big|\psi^{(K)}(x) - \psi^{(K)}(x')\big| \leq \|\psi\|_{\mathrm{Lip,TV}} \mathrm{TV}\big(\mu_x^{(K)}, \mu_{x'}^{(K)}\big) \leq \|\psi\|_{\mathrm{Lip,TV}} |||x - x'|||_1. \tag{1.36}$$

With these assumptions at hand, it is natural to wonder why we cannot simply invoke the main result in [13] to obtain the well-posedness of the projected Hamilton-Jacobi equation (1.23). The setting proposed in [13] is that of a Hamilton-Jacobi equation posed on a cone $\mathscr{C}$ and with a non-linearity that is defined over the cone $\mathscr{C}$ as well; the key assumption to establish well-posedness is that the non-linearity and the initial condition have their gradients in the cone $\mathscr{C}$. In our context, the non-linearity is initially only well-defined on the cone $\mathscr{C}_\infty$, or $\mathscr{C}_K$ for the projected equations, and we must make sure that the gradient of the solution remains in this space. This suggests that we try to use the results in [13] with $\mathscr{C} = \mathscr{C}_\infty$, or $\mathscr{C}_K$ for the projected equations. However, our problem, say for the projected equations, is naturally posed over $\mathbb{R}^{\mathscr{D}_K}_{\geq 0}$ rather than $\mathscr{C}_K$, and moreover, the gradient of the non-linearity that appears in our setting is not in $\mathscr{C}_K$, although it is in $\mathbb{R}^{\mathscr{D}_K}_{\geq 0}$. To make matters more complicated, the gradient of the finite-dimensional initial condition, and therefore also of the solution, does not quite belong to $\mathscr{C}_K$, although it is in the closed convex set $\mathscr{K}'_{a,K}$. Despite all this, we will show that the somewhat richer geometry of our problem can be dealt with using arguments that are similar to those in [13].

We now describe the structure of these arguments in more detail, and state our main results. We will first show that for any $R > 0$, it is possible to define a non-linearity $\mathsf{H}_{K,R} : \mathbb{R}^{\mathscr{D}_K} \to \mathbb{R}$ which agrees with the projected non-linearity C_K on a large enough ball $\mathscr{C}_K \cap B_R(0)$, and is uniformly Lipschitz continuous. We will then obtain the well-posedness of the projected Hamilton-Jacobi equation

$$\partial_t f^{(K)}(t,x) = \mathsf{H}_{K,R}\big(\nabla f^{(K)}(t,x)\big) \quad \text{on} \quad \mathbb{R}_{>0} \times \mathbb{R}^{\mathscr{D}_K}_{>0} \tag{1.37}$$

subject to the initial condition $f^{(K)}(0,x) = \psi^{(K)}(x)$ on $\mathbb{R}^{\mathscr{D}_K}_{\geq 0}$. Finally, we will show that the solutions to these projected Hamilton-Jacobi equations admit a limit as K tends to infinity. We will verify that this limit does not depend on the choice of the extension $\mathsf{H}_{K,R}$, provided that R is chosen sufficiently large, and define it to be the solution to the infinite-dimensional Hamilton-Jacobi equation (1.10).

To state our main well-posedness results, we introduce additional notation. Given functions $h : \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ and $u : [0,\infty) \times \mathbb{R}^d_{\geq 0} \to \mathbb{R}$, we define the semi-norms

$$|||h|||_{\mathrm{Lip},1} = \sup_{x \neq x' \in \mathbb{R}^d_{\geq 0}} \frac{|h(x) - h(x')|}{|||x - x'|||_1} \quad \text{and} \quad [u]_0 = \sup_{\substack{t>0 \\ x \in \mathbb{R}^d_{\geq 0}}} \frac{|u(t,x) - u(0,x)|}{t}, \tag{1.38}$$

and introduce the space of functions with Lipschitz initial condition that grow at most linearly in time,

$$\mathfrak{L} = \big\{u : [0,\infty) \times \mathbb{R}^d_{\geq 0} \to \mathbb{R} \mid u(0,\cdot) \text{ is Lipschitz continuous and } [u]_0 < \infty\big\}, \tag{1.39}$$

as well as its subset of uniformly Lipschitz continuous functions,

$$\mathfrak{L}_{\mathrm{unif}} = \Big\{ u \in \mathfrak{L} \mid \sup_{t \geq 0} |||u(t,\cdot)|||_{\mathrm{Lip},1} < \infty \Big\}. \tag{1.40}$$

The main well-posedness results for the projected Hamilton-Jacobi equation (1.37) and the infinite-dimensional Hamilton-Jacobi equation (1.10) now read as follows.

Theorem 1.1. *Under assumptions* (**H1**)-(**H3**)*, the projected Hamilton-Jacobi equation* (1.37) *with* $R > 0$ *admits a unique viscosity solution* $f_R^{(K)} \in \mathfrak{L}_{\mathrm{unif}}$ *subject to the initial condition* $\psi^{(K)}$*. Moreover,* $f_R^{(K)}$ *has its gradient in the set* $\mathscr{K}'_{a,K}$ *and satisfies the Lipschitz bound*

$$\sup_{t>0} |||f_R^{(K)}(t,\cdot)|||_{\mathrm{Lip},1} = |||\psi^{(K)}|||_{\mathrm{Lip},1} \leq \|\psi\|_{\mathrm{Lip,TV}}. \tag{1.41}$$

Theorem 1.2. *Suppose* (**H1**)-(**H4**)*, and given an integer* $K \geq 1$ *and a real number* $R > \|\psi\|_{\mathrm{Lip,TV}}$*, denote by* $f_R^{(K)} \in \mathfrak{L}_{\mathrm{unif}}$ *the unique viscosity solution to the Hamilton-Jacobi equation* (1.37) *constructed in Theorem 1.1. For every* $t \geq 0$ *and every measure* $\mu \in \mathscr{M}_+$*, the limit*

$$f(t,\mu) = \lim_{K \to \infty} f_R^{(K)}(t, x^{(K)}(\mu)) \tag{1.42}$$

exists, is finite and is independent of R. The value of this limit is defined to be the solution to the infinite-dimensional Hamilton-Jacobi equation (1.10)*.*

Solutions to (1.10) satisfy a comparison principle, since a comparison principle also holds for solutions to the projected Hamilton-Jacobi equation (1.37), by Corollary A.2.

As is apparent, and similarly to [27, 29], we content ourselves here with identifying the solution to (1.10) as the limit of our finite-dimensional approximations. This will suffice for our purposes, and we leave open the question of providing a more intrinsic characterization of the solution to (1.10), as was achieved in [12] in a related context.

In addition to these well-posedness results, we also obtain a Hopf-Lax variational representation for the solution to the infinite-dimensional Hamilton-Jacobi equation (1.10) in the case when the non-linearity C_∞ is convex. Hopf-Lax formulas for related problems have been explored in [9, 10, 12, 13]. In [19], this variational representation will allow us to verify that, in the disassortative regime, our conjectured asymptotic mutual information for the sparse stochastic block model coincides with the value of the asymptotic mutual information established in [14]. The convexity condition on C_∞ boils down to the requirement that the mapping $(x,y) \mapsto g(xy)$ be non-negative definite, and can be phrased as follows.

H5 The kernel $g : [-1,1] \to \mathbb{R}$ satisfies the property

$$\int_{-1}^{1} \int_{-1}^{1} g(xy) \,\mathrm{d}\mu(x) \,\mathrm{d}\mu(y) \geq 0 \tag{1.43}$$

for every signed measure $\mu \in \mathscr{M}_s$.

Theorem 1.3. *If* (**H1**)-(**H5**) *hold, then the unique solution* $f : [0,\infty) \times \mathscr{M}_+ \to \mathbb{R}$ *to the infinite-dimensional Hamilton-Jacobi equation* (1.10) *constructed in Theorem 1.2 admits the Hopf-Lax variational representation*

$$f(t,\mu) = \sup_{\nu \in \mathscr{M}_+} \left\{ \psi(\mu + t\nu) - \frac{t}{2} \int_{-1}^{1} G_\nu(y)\,\mathrm{d}\nu(y) \right\} \tag{1.44}$$

for every $t > 0$ *and* $\mu \in \mathscr{M}_+$*. Moreover, the supremum in* (1.44) *is achieved at some* $\nu^* \in \mathscr{M}_+$*, and whenever the initial condition* ψ *admits a Gateaux derivative at the measure* $\mu + t\nu^*$ *with a density* $x \mapsto D_\mu \psi(\mu + t\nu^*, x)$ *belonging to the cone* $\mathscr{C}_\infty$*, we have*

$$G_{\nu^*} = D_\mu \psi(\mu + t\nu^*, \cdot). \tag{1.45}$$

For the purposes of our companion work [19], it will also be important to identify solutions to equations of the form (1.10) with a kernel g that does not satisfy the positivity assumption (**H1**). The idea will be to introduce a new kernel which satisfies (**H1**) by translating g, and to deduce the well-posedness of the equation with kernel g from the well-posedness of the equation with the translated kernel. For this strategy to work, we will replace the assumption (**H3**) on the initial condition by a stronger assumption which we now describe.

For every $a \in \mathbb{R}$ introduce the set of measures with mass a,

$$\mathscr{M}_{a,+} = \{\mu \in \mathscr{M}_+ \mid \mu[-1,1] = a\}, \tag{1.46}$$

as well as the set of functions

$$\mathscr{C}_{a,\infty} = \{G_\mu \mid \mu \in \mathscr{M}_{a,+}\}. \tag{1.47}$$

The assumption (**H3**) on the initial condition will essentially be replaced by the assumption that its Gateaux derivative lies in the set $\mathscr{C}_{a,\infty}$ for some $a \in \mathbb{R}$. As before, it will be convenient to state this as an assumption on the projected initial conditions (1.17). For every integer $K \geq 1$ introduce the set of projected measures with mass a,

$$\mathscr{M}_{a,+}^{(K)} = \left\{ \mu \in \mathscr{M}_+^{(K)} \mid \mu[1,1] = a \right\}, \tag{1.48}$$

and write

$$\mathscr{K}_{=a,K} = \left\{ G^{(K)} x \in \mathbb{R}^{\mathscr{D}_K} \mid x \in \mathbb{R}^{\mathscr{D}_K}_{\geq 0} \text{ and } |||x|||_1 = a \right\} \tag{1.49}$$

for its associated set of functions. We replace (**H3**) by the following stronger assumption.

H3' There exists $a > 0$ such that the initial condition $\psi : \mathscr{M}_+ \to \mathbb{R}$ has the property that each of the projected initial conditions (1.17) has its gradient in the set $\mathscr{K}'_{=a,K}$,

$$\nabla \psi^{(K)} \in L^\infty\big(\mathbb{R}^d_{\geq 0}; \mathscr{K}'_{=a,K}\big). \tag{1.50}$$

Formal calculations now suggest a way to modify the solution to the infinite-dimensional Hamilton-Jacobi equation (1.10) if the kernel g is not assumed to satisfy (**H1**) but is translated by a large enough constant so that it becomes positive. Given a continuously differentiable kernel $g : [-1,1] \to \mathbb{R}$, fix $b \in \mathbb{R}$ such that the modified kernel

$$\widetilde{g}_b(z) = g(z) + b \tag{1.51}$$

is strictly positive. For every $\mu \in \mathscr{M}_+$ define the modified function $\widetilde{G}_{b,\mu} : [-1,1] \to \mathbb{R}$,

$$\widetilde{G}_{b,\mu}(x) = \int_{-1}^{1} \widetilde{g}_b(xy)\,\mathrm{d}\mu(y), \tag{1.52}$$

the modified cone of functions,

$$\widetilde{\mathscr{C}}_{b,\infty} = \big\{\widetilde{G}_{b,\mu} \mid \mu \in \mathscr{M}_+\big\}, \tag{1.53}$$

and the modified non-linearity $\widetilde{\mathsf{C}}_{b,\infty} : \widetilde{\mathscr{C}}_{b,\infty} \to \mathbb{R}$,

$$\widetilde{\mathsf{C}}_{b,\infty}\big(\widetilde{G}_{b,\mu}\big) = \frac{1}{2}\int_{-1}^{1} \widetilde{G}_{b,\mu}(x)\,\mathrm{d}\mu(x) = \frac{1}{2}\int_{-1}^{1}\int_{-1}^{1} \widetilde{g}_b(xy)\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x). \tag{1.54}$$

Notice that the additional constant b in $\widetilde{g}_b$ induces a shift in the expression above that depends only on the total mass of the measure μ. This suggests that, under assumption (**H3'**), if $\widetilde{f}_b$ is a solution to the infinite-dimensional Hamilton-Jacobi equation

$$\begin{cases} \partial_t \widetilde{f}(t,\mu) = \widetilde{\mathsf{C}}_{b,\infty}\big(D_\mu \widetilde{f}(t,\mu)\big) & \text{on } \mathbb{R}_{>0} \times \mathscr{M}_+, \\ \widetilde{f}(0,\mu) = \widetilde{\psi}_b(\mu) & \text{on } \mathscr{M}_+, \end{cases} \tag{1.55}$$

for the initial condition $\widetilde{\psi}_b : \mathscr{M}_+ \to \mathbb{R}$ defined by

$$\widetilde{\psi}_b(\mu) = \psi(\mu) + ab\int_{-1}^{1} \mathrm{d}\mu, \tag{1.56}$$

then the function

$$f_b(t,\mu) = \widetilde{f}_b(t,\mu) - ab\int_{-1}^{1} \mathrm{d}\mu - \frac{a^2 bt}{2} \tag{1.57}$$

should be a solution to the infinite-dimensional Hamilton-Jacobi equation (1.10). We omit the dependence of $\widetilde{f}_b$, f_b and $\widetilde{\psi}_b$ on a since this constant is given to us and fixed by (**H3'**). The following result renders this construction precise and ensures that it is independent of the choice of b.

Theorem 1.4. *Fix a continuously differentiable kernel $g : [-1,1] \to \mathbb{R}$ and assume that* (**H2**)*,* (**H3'**)*, and* (**H4**) *hold. Let $b \in \mathbb{R}$ be such that the function $\widetilde{g}_b$ defined in* (1.51) *is positive on $[-1,1]$, let $\widetilde{\psi}_b$ be defined by* (1.56)*, and let $\widetilde{f}_b$ be the solution to the infinite-dimensional Hamilton-Jacobi equation* (1.55) *constructed in Theorem 1.2. The function f_b given by* (1.57) *does not depend on the choice of $b \in \mathbb{R}$, and it is defined to be the solution to the infinite-dimensional Hamilton-Jacobi equation* (1.10).

Combining this well-posedness result with the Hopf-Lax representation formula in Theorem 1.3 shows that under the additional assumption (**H5**), the function (1.57) admits a variational representation.

Theorem 1.5. *Fix a continuously differentiable kernel $g : [-1,1] \to \mathbb{R}$ satisfying (**H2**), (**H3'**) and (**H4**). Suppose that there exists $b \in \mathbb{R}$ such that the translated kernel $\widetilde{g}_b$ in (1.51) is strictly positive on $[-1,1]$ and satisfies (**H5**). Suppose moreover that for every $\mu \in \mathscr{M}_+$, the initial condition ψ admits a Gateaux derivative with density $x \mapsto D_\mu \psi(\mu, x)$ belonging to the set $\mathscr{C}_{a,\infty}$. Then, the unique solution $f : [0,\infty) \times \mathscr{M}_+ \to \mathbb{R}$ to the infinite-dimensional Hamilton-Jacobi equation (1.10) constructed in Theorem 1.4 admits the Hopf-Lax variational representation*

$$f(t,\mu) = \sup_{\nu \in \mathscr{M}_{a,+}} \left\{ \psi(\mu + t\nu) - \frac{t}{2} \int_{-1}^{1} G_\nu(y) \,\mathrm{d}\nu(y) \right\} \tag{1.58}$$

for every $t > 0$ and $\mu \in \mathscr{M}_+$. Moreover, the supremum in (1.58) is achieved at some $\nu^* \in \mathscr{M}_{a,+}$ with

$$G_{\nu^*} = D_\mu \psi(\mu + t\nu^*, \cdot). \tag{1.59}$$

We briefly review related works on Hamilton-Jacobi equations in infinite dimensions. The study of equations posed on infinite-dimensional Banach spaces was initiated in [15, 16, 17]. The assumptions imposed on the Banach space preclude the possibility to apply the results presented there to the space of bounded measures. The existence of solutions is obtained via a connection with differential games. An example is also given in which solutions to natural finite-dimensional approximations fail to converge to the solution of the infinite-dimensional equation. We do not expect this phenomenon to occur for the problem we consider in this paper, and in any case, our definition of the solution as the limit of finite-dimensional approximations is the one we make use of in our companion work [19]. Moreover, for the equations of transport type appearing in the context of mean-field spin glasses, it was shown in [12] that finite-dimensional approximations do converge to the intrinsic viscosity solution of the infinite-dimensional equation.

Equations that are posed over a space of probability measures, or more general metric spaces, have been considered in a number of works including [1, 4, 5, 6, 21, 22, 23, 24]. These works revolve around equations involving derivatives of transport type for probability measures over $\mathbb{R}^d$. Since transportation of mass over $\mathbb{R}^d$ can be carried without limit, questions of boundary conditions do not arise there, unlike in the more recent works [12, 27, 29] already cited above in which probability measures over $\mathbb{R}^d_{\geq 0}$ or the space of non-negative definite matrices are considered. We are not aware of previous works considering equations that involve derivatives of "affine" type, as we do here. In this context, the natural "movements" are different from those appearing for the transport geometry, and our additional constraint that we must deal with non-negative measures is the source of the necessity to address boundary issues.

We close this section with a brief outline of the paper. In Section 2, we introduce a non-decreasing and uniformly Lipschitz continuous non-linearity $\mathsf{H}_{K,R}$ which agrees with the projected non-linearity (1.21) on the intersection between the projected cone (1.20) and a large enough ball, and we define the appropriate notion of solution to the projected Hamilton-Jacobi equation (1.37). The definition of the function $\mathsf{H}_{K,R}$ is inspired by Proposition 6.8 in [27] and Lemma 2.5 in [13]. We then leverage the well-posedness results established in Appendix A to prove Theorem 1.1. In Section 3, with the well-posedness of the projected Hamilton-Jacobi equations (1.37) at hand, we modify the arguments in Section 3.2 of [29] and Section 3.3 of [12] to obtain the convergence of solutions as described in Theorem 1.2. In Section 4 we proceed as in Section 6 of [13] to obtain an approximate Hopf-Lax variational representation for the solution to the projected Hamilton-Jacobi equation (1.37). By taking an appropriate limit in this variational formula, Theorem 1.3 is established in Section 5. Section 6 is devoted to the proof of Theorem 1.4 and Theorem 1.5. So as to not disrupt the flow of the paper, the main technical arguments required to establish the well-posedness of the projected Hamilton-Jacobi equation (1.37) have been postponed to Appendix A. All the results in this appendix appear in [13] in some form but have been reproduced here for the reader's convenience. It is also worth pointing out that we treat a slightly different setting to the one in [13]. Indeed, our initial condition has its gradient close to a convex set and not in it, and our non-linearity is monotonic with respect to a different convex set, as discussed below (1.36). Appendix B reviews a fundamental duality theorem from convex analysis, establishes a non-differential criterion for a Lipschitz function to have its gradient in a closed convex set and provides a refresher on the basic properties of semi-continuous functions, which play a role in Appendix A when running Perron's argument for the existence of solutions.

Acknowledgements. *We would like to warmly thank Hong-Bin Chen and Jiaming Xia for sharing a preliminary version of [13] with us. Their work considerably simplified our task, allowing us in particular to discard the much more complicated approach we had originally envisaged.*

2 Construction of finite-dimensional approximations

In this section, we prove Theorem 1.1, and in particular establish the well-posedness of the projected Hamilton-Jacobi equation (1.37). Throughout the paper, we say that a function $h:\mathbb{R}^d\to\mathbb{R}$ is non-decreasing if for every $x,x'\in\mathbb{R}^d$ with $x'-x\in\mathbb{R}^d_{\geq 0}$, we have $h(x')-h(x)\geq 0$. More generally, given a closed convex cone $\mathscr{C}\subset\mathbb{R}^d$, and denoting its dual cone by $\mathscr{C}^*$ (see (B.1) for the definition), we say that a function $h:\mathbb{R}^d\to\mathbb{R}$ is $\mathscr{C}^*$-non-decreasing if for every $x,x'\in\mathbb{R}^d$ with $x'-x\in\mathscr{C}^*$, we have $h(x')-h(x)\geq 0$. In the case when h is Lipschitz continuous, this is equivalent to the requirement that h have its gradient in $\mathscr{C}$ (see Proposition B.2). To alleviate notation and strive for generality, fix an integer dimension $d\geq 1$ and a symmetric matrix $G\in\mathbb{R}^{d\times d}$ for which there exist positive constants $m,M>0$ with

$$\frac{m}{d^2}\leq G_{kk'}\leq\frac{M}{d^2} \tag{2.1}$$

for all $1\leq k,k'\leq d$. Consider the cone

$$\mathscr{C}=\big\{Gx\in\mathbb{R}^d\mid x\in\mathbb{R}^d_{\geq 0}\big\}, \tag{2.2}$$

the closed convex set

$$\mathscr{K}_a=\big\{Gx\in\mathbb{R}^d\mid x\in\mathbb{R}^d_{\geq 0}\text{ and }|||x|||_1\leq a\big\}, \tag{2.3}$$

and the non-linearity $\mathsf{C}:\mathscr{C}\to\mathbb{R}$ defined by

$$\mathsf{C}(Gx)=\frac{1}{2}Gx\cdot x. \tag{2.4}$$

This mapping is well-defined, for the same reason as that explained below (1.5). Recall the definition of the enlarged set (1.29) and of the normalized-ℓ^1 and normalized-$\ell^{1,*}$ norms in (1.24). The first important result of this section will be the definition of a uniformly Lipschitz continuous and non-decreasing non-linearity $\mathsf{H}_R : \mathbb{R}^d \to \mathbb{R}$ which agrees with C on the intersection of the cone $\mathscr{C}$ and a large enough ball. We will then obtain the well-posedness of the Hamilton-Jacobi equation associated with this non-linearity,

$$\partial_t f(t,x) = \mathsf{H}_R\big(\nabla f(t,x)\big) \quad \text{on} \quad \mathbb{R}_{>0} \times \mathbb{R}^d_{>0}, \tag{2.5}$$

subject to a Lipschitz continuous initial condition $\psi : \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ with

$$|\psi(x) - \psi(y)| \leq |||\psi|||_{\mathrm{Lip},1} |||x-y|||_1 \quad \text{and} \quad \nabla\psi \in L^\infty(\mathbb{R}^d_{\geq 0}; \mathscr{K}_a'). \tag{2.6}$$

By (**H3**) and (1.36), the Hamilton-Jacobi equation (2.5) corresponds to the projected Hamilton-Jacobi equation (1.37) for the choices $d = |\mathscr{D}_K|$, $G = G^{(K)}$ and $\psi = \psi^{(K)}$. Theorem 1.1 will therefore be an immediate consequence of the main well-posedness result of this section.

To establish the well-posedness of the Hamilton-Jacobi equation (2.5) using the results in Appendix A, it will be important that the extended non-linearity $\mathsf{H}_R : \mathbb{R}^d \to \mathbb{R}$ be Lipschitz continuous and non-decreasing. Let us start by verifying that these properties are satisfied locally on the cone by the original non-linearity (2.4). It will be convenient to note that

$$|||x|||_1 \leq \frac{1}{m} |||Gx|||_{1,*} \tag{2.7}$$

for all $x \in \mathbb{R}^d_{\geq 0}$.

Lemma 2.1. *The non-linearity* (2.4) *is locally Lipschitz continuous with respect to the normalized-$\ell^{1,*}$ norm,*

$$|\mathsf{C}(y) - \mathsf{C}(y')| \leq \frac{1}{m}\big(|||y|||_{1,*} + |||y'|||_{1,*}\big) |||y-y'|||_{1,*} \tag{2.8}$$

for all $y, y' \in \mathscr{C}$.

Proof. Fix $y, y' \in \mathscr{C}$ with $y = Gx$ and $y' = Gx'$ for some $x, x' \in \mathbb{R}^d_{\geq 0}$. The symmetry of G and the Cauchy-Schwarz inequality imply that

$$|\mathsf{C}(y) - \mathsf{C}(y')| \leq |G(x-x')\cdot x| + |G(x-x')\cdot x'| \leq \big(|||x|||_1 + |||x'|||_1\big) |||y-y'|||_{1,*}.$$

It follows by (2.7) that

$$|\mathsf{C}(y) - \mathsf{C}(y')| \leq \frac{1}{m}\big(|||y|||_{1,*} + |||y'|||_{1,*}\big) |||y-y'|||_{1,*}.$$

This completes the proof. ∎

Lemma 2.2. *The non-linearity* (2.4) *is non-decreasing.*

Proof. Fix $y, y' \in \mathscr{C}$ with $y \leq y'$ (by this we mean that $y' - y \in \mathbb{R}^d_{\geq 0}$), and let $x, x' \in \mathbb{R}^d_{\geq 0}$ be such that $y = Gx$ and $y' = Gx'$. Observe that

$$2\mathsf{C}(y) = Gx\cdot x = y \cdot x \leq y' \cdot x = Gx' \cdot x = Gx \cdot x' = x' \cdot y \leq x' \cdot y' = Gx' \cdot x' = 2\mathsf{C}(y').$$

This completes the proof. ∎

Extending the non-linearity (2.4) to $\mathbb{R}^d$ while preserving these two key properties requires some care. For each $R > 0$, we will define a non-decreasing function $\mathsf{H}_R : \mathbb{R}^d \to \mathbb{R}$ which is uniformly Lipschitz continuous with respect to the normalized-$\ell^{1,*}$ norm and agrees with the non-linearity (2.4) on the intersection of the cone (2.2) and the ball $B_R = B_R(0)$ defined in (1.30). The definition of this extension is inspired by Proposition 6.8 in [27] and Lemma 2.5 in [13].

Proposition 2.3. *For every $R > 0$, there exists a non-decreasing non-linearity $\mathsf{H}_R : \mathbb{R}^d \to \mathbb{R}$ which agrees with C on $\mathscr{C} \cap B_R$ and satisfies the Lipschitz continuity property*

$$|\mathsf{H}_R(y) - \mathsf{H}_R(y')| \leq \frac{8RM}{m^2} |||y - y'|||_{1,*} \tag{2.9}$$

for all $y, y' \in \mathbb{R}^d$.

Proof. The proof proceeds in two steps: first we regularize C by defining a non-decreasing and uniformly Lipschitz continuous function which agrees with C on $\mathscr{C} \cap B_R$, and then we extend this regularization to $\mathbb{R}^d$.

Step 1: regularizing C.

By Lemma 2.1, the non-linearity (2.4) satisfies the Lipschitz bound

$$|\mathsf{C}(y) - \mathsf{C}(y')| \leq \frac{4R}{m} |||y - y'|||_{1,*}$$

for all $y, y' \in \mathscr{C} \cap B_{2R}$. With this in mind, let $L = \frac{4R}{m}$, and define the regularized non-linearity $\widetilde{\mathsf{C}}_R : \mathscr{C} \to \mathbb{R}$ by

$$\widetilde{\mathsf{C}}_R(y) = \begin{cases} \max\Big(\mathsf{C}(y), \mathsf{C}(0) + 2L\big(|||y|||_{1,*} - R\big)\Big) & \text{if } y \in \mathscr{C} \cap B_{2R}, \\ \mathsf{C}(0) + 2L\big(|||y|||_{1,*} - R\big) & \text{if } y \in \mathscr{C} \setminus B_{2R}. \end{cases}$$

To see that $\widetilde{\mathsf{C}}_R$ agrees with C on $\mathscr{C} \cap B_R$, observe that for any $y \in \mathscr{C} \cap B_R$,

$$\mathsf{C}(0) + 2L\big(|||y|||_{1,*} - R\big) \leq \mathsf{C}(0) = 0 \leq \mathsf{C}(y),$$

where the last inequality uses the non-negativity of the components of G. It will also be convenient to note that by Lipschitz continuity of C on $\mathscr{C} \cap B_{2R}$,

$$\mathsf{C}(0) + 2L\big(|||y|||_{1,*} - R\big) = \mathsf{C}(0) + 2LR = \mathsf{C}(0) + L|||y|||_{1,*} \geq \mathsf{C}(y)$$

for any $y \in \mathscr{C} \cap \partial B_{2R}$. This shows that $\widetilde{\mathsf{C}}_R$ is continuous. To establish the non-decreasingness of $\widetilde{\mathsf{C}}_R$, fix $y, y' \in \mathscr{C}$ with $y \leq y'$. If $y, y' \in B_{2R}$, then the non-decreasingness of C in Lemma 2.2 implies that $\mathsf{C}(y) \leq \mathsf{C}(y')$. Combining this with the fact that $|||y|||_{1,*} \leq |||y'|||_{1,*}$ reveals that $\widetilde{\mathsf{C}}_R(y) \leq \widetilde{\mathsf{C}}_R(y')$. On the other hand, if $y \in B_{2R}$ and $y' \in \mathscr{C} \setminus B_{2R}$, then

$$\mathsf{C}(y) \leq \mathsf{C}(0) + L|||y|||_{1,*} \leq \mathsf{C}(0) + L|||y'|||_{1,*} + L\big(|||y'|||_{1,*} - 2R\big) = \widetilde{\mathsf{C}}_R(y')$$

and

$$\mathsf{C}(0) + 2L\big(|||y|||_{1,*} - R\big) \leq \mathsf{C}(0) + 2L\big(|||y'|||_{1,*} - R\big) = \widetilde{\mathsf{C}}_R(y').$$

Once again $\widetilde{\mathsf{C}}_R(y) \le \widetilde{\mathsf{C}}_R(y')$. Finally, if $y \in \mathscr{C} \setminus B_{2R}$, then $2R \le |||y|||_{1,*} \le |||y'|||_{1,*}$ so $y' \in \mathscr{C} \setminus B_{2R}$ and clearly $\widetilde{\mathsf{C}}_R(y) \le \widetilde{\mathsf{C}}_R(y')$. This establishes the non-decreasingness of the regularized non-linearity $\widetilde{\mathsf{C}}_R$. We now show that this non-linearity is uniformly Lipschitz continuous. The reverse triangle inequality implies that the map $y \mapsto \mathsf{C}(0) + 2L\big(|||y|||_{1,*} - R\big)$ is Lipschitz continuous with Lipschitz constant at most $2L$. Recall that the maximum of two Lipschitz continuous maps with Lipschitz constants at most L_1 and L_2, respectively, is Lipschitz continuous with Lipschitz constant at most $\max(L_1, L_2)$. This means that $\widetilde{\mathsf{C}}_R$ is Lipschitz continuous with Lipschitz constant at most $2L$ when it is restricted to $\mathscr{C} \cap B_{2R}$ or $\mathscr{C} \setminus B_{2R}$. For $y, y' \in \mathscr{C}$ with $y \in B_{2R}$ and $y' \in \mathscr{C} \setminus B_{2R}$, we distinguish two cases. On the one hand, if $\widetilde{\mathsf{C}}_R(y) = \mathsf{C}(0) + 2L\big(|||y|||_{1,*} - R\big)$, the reverse triangle inequality shows that

$$\big|\widetilde{\mathsf{C}}_R(y) - \widetilde{\mathsf{C}}_R(y')\big| \le 2L\big||||y|||_{1,*} - |||y'|||_{1,*}\big| \le 2L|||y - y'|||_{1,*}.$$

On the other hand, if $\widetilde{\mathsf{C}}_R(y) = \mathsf{C}(y)$, then the reverse triangle inequality reveals that

$$\begin{aligned}\widetilde{\mathsf{C}}_R(y) - \widetilde{\mathsf{C}}_R(y') &\le \mathsf{C}(0) + L|||y|||_{1,*} - \mathsf{C}(0) - 2L\big(|||y'|||_{1,*} - R\big) \le L|||y - y'|||_{1,*} + L\big(2R - |||y'|||_{1,*}\big) \\ &\le L|||y - y'|||_{1,*}\end{aligned}$$

while the lower bound $\widetilde{\mathsf{C}}_R(y) = \mathsf{C}(y) \ge \mathsf{C}(0) + 2L\big(|||y|||_{1,*} - R\big)$ yields

$$\widetilde{\mathsf{C}}_R(y') - \widetilde{\mathsf{C}}_R(y) = 2L\big(|||y'|||_{1,*} - |||y|||_{1,*}\big) \le 2L|||y - y'|||_{1,*}.$$

This shows that $\widetilde{\mathsf{C}}_R$ is a non-decreasing function which agrees with C on $\mathscr{C} \cap B_R$ and satisfies the Lipschitz continuity property

$$\big|\widetilde{\mathsf{C}}_R(y) - \widetilde{\mathsf{C}}_R(y')\big| \le \frac{8R}{m}|||y - y'|||_{1,*} \tag{2.10}$$

for all $y, y' \in \mathscr{C}$.

Step 2: extending to $\mathbb{R}^d$.

To extend the regularization of the non-linearity (2.4) to $\mathbb{R}^d$, define the function $\mathsf{H}_R : \mathbb{R}^d \to \mathbb{R}$ by

$$\mathsf{H}_R(y) = \inf\Big\{\widetilde{\mathsf{C}}_R(w) \mid w \in \mathscr{C} \text{ with } w \ge y\Big\}. \tag{2.11}$$

Let $\iota = (1, \dots, 1) \in \mathbb{R}^d$ and observe that the vector $v = \frac{G\iota}{m}$ belongs to $\mathscr{C}$ and satisfies the bounds

$$\frac{1}{d} \le v_k \le \frac{M}{dm} \tag{2.12}$$

for $1 \le k \le d$. In particular, the infimum in (2.11) is never taken over the empty set. Moreover, the non-decreasingness of $\widetilde{\mathsf{C}}_R$ and the fact that this function agrees with C on $\mathscr{C} \cap B_R$ imply that H_R also agrees with C on $\mathscr{C} \cap B_R$. To see that H_R is non-decreasing, fix $y, y' \in \mathbb{R}^d$ with $y \ge y'$, and let $w \in \mathscr{C}$ be such that $w \ge y'$. Since $w \ge y$, the definition of H_R gives $\mathsf{H}_R(y) \le \widetilde{\mathsf{C}}_R(w)$, and taking the infimum over all such w shows that $\mathsf{H}_R(y) \le \mathsf{H}_R(y')$. To establish the Lipschitz continuity of H_R, fix $y, y' \in \mathbb{R}^d$ and let $z = |||y - y'|||_{1,*} v \in \mathscr{C}$. Recalling (2.12) reveals that for any $1 \le k \le d$,

$$y_k - y'_k \le \|y - y'\|_\infty = \frac{1}{d}|||y - y'|||_{1,*} \le v_k|||y - y'|||_{1,*} = z_k.$$

This means that $z \geq y - y'$. In particular, if $w \in \mathscr{C}$ is such that $w \geq y'$, then $w + z \in \mathscr{C}$ with $w + z \geq y$. It follows by (2.10), (2.11) and (2.12) that

$$\mathsf{H}_R(y) - \widetilde{\mathsf{C}}_R(w) \leq \widetilde{\mathsf{C}}_R(w+z) - \widetilde{\mathsf{C}}_R(w) \leq \frac{8R}{m} |||z|||_{1,*} \leq \frac{8RM}{m^2} |||y - y'|||_{1,*}.$$

Taking the infimum over all such w and reversing the roles of y and y' completes the proof. ■

With this extended non-linearity at hand, we can now establish the well-posedness of the Hamilton-Jacobi equation (2.5). The appropriate notion of solution for (2.5) will be that of a viscosity solution. With Appendix A in mind, let us fix a non-linearity $\mathsf{H} : \mathbb{R}^d \to \mathbb{R}$ and a domain $\mathscr{D} \subset \mathbb{R}^d_{\geq 0}$, and define the notion of a viscosity solution for the more general Hamilton-Jacobi equation

$$\partial_t f(t,x) = \mathsf{H}\big(\nabla f(t,x)\big) \quad \text{on} \quad \mathbb{R}_{>0} \times \mathscr{D}. \tag{2.13}$$

A brief review of the definition and elementary properties of semi-continuous functions is provided in Appendix B.

Definition 2.4. *An upper semi-continuous function $u : [0,\infty) \times \mathscr{D} \to \mathbb{R}$ is said to be a viscosity subsolution to (2.13) if, given any $\phi \in C^\infty\big((0,\infty) \times \mathscr{D}\big)$ with the property that $u - \phi$ has a local maximum at $(t^*, x^*) \in (0,\infty) \times \mathscr{D}$,*

$$\big(\partial_t \phi - \mathsf{H}(\nabla \phi)\big)(t^*, x^*) \leq 0. \tag{2.14}$$

Definition 2.5. *A lower semi-continuous function $v : [0,\infty) \times \mathscr{D} \to \mathbb{R}$ is said to be a viscosity supersolution to (2.13) if, given any $\phi \in C^\infty\big((0,\infty) \times \mathscr{D}\big)$ with the property that $v - \phi$ has a local minimum at $(t^*, x^*) \in (0,\infty) \times \mathscr{D}$,*

$$\big(\partial_t \phi - \mathsf{H}(\nabla \phi)\big)(t^*, x^*) \geq 0. \tag{2.15}$$

Definition 2.6. *A continuous function $f \in C\big([0,\infty) \times \mathscr{D}\big)$ is said to be a viscosity solution to (2.13) if it is both a viscosity subsolution and a viscosity supersolution to (2.13).*

The existence and uniqueness results for Hamilton-Jacobi equations on positive half-spaces developed in Appendix A now give the well-posedness of the Hamilton-Jacobi equation (2.5). It will be convenient to remember that any closed and convex set $\mathscr{K} \subset \mathbb{R}^d$ may be represented as the intersection of the closed and affine half-spaces which contain it,

$$\mathscr{K} = \big\{x \in \mathbb{R}^d \mid x \cdot v \geq c \text{ for all } (v,c) \in \mathscr{A}\big\}, \tag{2.16}$$

where

$$\mathscr{A} = \big\{(v,c) \in \mathbb{R}^{d+1} \mid x \cdot v \geq c \text{ for all } x \in \mathscr{K} \text{ and } \|v\| = 1\big\} \tag{2.17}$$

for any norm $\|\cdot\|$. A proof of this classical result may be found in Corollary 4.2.4 of [25].

Proposition 2.7. *For every $R > 0$, the Hamilton-Jacobi equation (2.5) admits a unique viscosity solution $f_R \in \mathfrak{L}_{\mathrm{unif}}$ subject to the initial condition ψ. Moreover, f_R has its gradient in the set $\mathscr{K}'_a$ and satisfies the Lipschitz bound*

$$\sup_{t>0} |||f_R(t,\cdot)|||_{\mathrm{Lip},1} = |||\psi|||_{\mathrm{Lip},1}. \tag{2.18}$$

Proof. To alleviate notation, we fix $R > 0$ and omit all dependencies on $R > 0$. We invoke Corollary A.12 to find a viscosity solution $f \in \mathfrak{L}_{\text{unif}}$ to the Hamilton-Jacobi equation (2.5) subject to the initial condition ψ with

$$\sup_{t>0} |||f(t,\cdot)|||_{\text{Lip},1} = |||\psi|||_{\text{Lip},1}. \tag{2.19}$$

We will now show that f has its gradient in the closed convex set $\mathscr{K}'_a$. Denote by $\mathscr{A}$ the set (2.17) associated with $\mathscr{K}'_a$ and the Euclidean norm $\|\cdot\|_2$. For each $(v,c) \in \mathscr{A}$ introduce the closed convex cone

$$\mathscr{H}_v = \{x \in \mathbb{R}^d \mid x \cdot v \geq 0\}$$

as well as the function $g_{v,c} : [0,\infty) \times \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ defined by $g_{v,c}(t,x) = f(t,x) - cx \cdot v$. It is readily verified that $g_{v,c}$ satisfies the Hamilton-Jacobi equation

$$\partial_t g(t,x) = \widetilde{\mathsf{H}}_R\big(\nabla g(t,x)\big) \quad \text{ on } \mathbb{R}_{>0} \times \mathbb{R}^d_{\geq 0}$$

subject to the initial condition $g_{v,c}(0,x) = \psi(x) - cx \cdot v$ for the non-linearity $\widetilde{\mathsf{H}}_R(y) = \mathsf{H}_R(y + cv\iota)$, where $\iota = (1)_{\mathscr{D}_K} \in \mathbb{R}^{\mathscr{D}_K}$. Moreover, this initial condition is $\mathscr{H}_v^*$-non-decreasing (the definition of being $\mathscr{H}_v^*$-non-decreasing was introduced at the beginning of this section). Indeed, we have $\mathscr{H}_v^* = \mathbb{R}v$ by the biduality result in Proposition B.1. Moreover, for any $x, x' \in \mathbb{R}^d_{\geq 0}$ with $x' - x = tv$ for some $t \in \mathbb{R}$, the fact that $(x' - x) \cdot z \geq tc$ for all $z \in \mathscr{K}'_a$ and the characterization of ψ having its gradient in the set $\mathscr{K}'_a$ given in Proposition B.2 imply that

$$g_{v,c}(x') - g_{v,c}(x) = \psi(x') - \psi(x) - c(x' - x) \cdot v \geq tc - tcv \cdot v = 0.$$

It follows by Proposition A.13 that $g_{v,c}$ is $\mathscr{H}_v^*$-non-decreasing. We now fix $x, x' \in \mathbb{R}^d_{\geq 0}$ with the property that for all $z \in \mathscr{K}'_a$, we have $(x' - x) \cdot z \geq c$. If $v = \frac{x'-x}{\|x'-x\|_2}$ and $c' = \frac{c}{\|x'-x\|_2}$, then $(v, c') \in \mathscr{A}$ so the function $g_{v,c'}$ is $\mathscr{H}_v^*$-non-decreasing. This implies that

$$f(t,x') - f(t,x) = g_{v,c'}\big(t, x + \|x' - x\|_2 v\big) - g_{v,c'}(t,x) + c'\|x' - x\|_2 v \cdot v \geq c$$

which means that f has its gradient in $\mathscr{K}'_a$ by Proposition B.2. This completes the proof. ■

Proof of Theorem 1.1. Writing $\mathsf{H}_{K,R}$ for the extension of the non-linearity C_K constructed in Proposition 2.3, the desired result is now an immediate consequence of Proposition 2.7 and the Lipschitz bound (1.36). ■

3 Well-posedness of the infinite-dimensional equation

In this section, we establish the convergence of the solutions to the projected Hamilton-Jacobi equations (1.37) as stated in Theorem 1.2. In the notation of Theorem 1.1, given an integer $K \geq 1$ and some $R > 0$, write $f_R^{(K)} \in \mathfrak{L}_{\text{unif}}$ for the unique solution to the Hamilton-Jacobi equation (2.5) subject to the initial condition $\psi^{(K)}$. Recall that $f_R^{(K)}$ has its gradient in the set $\mathscr{K}'_{a,K}$ and satisfies the Lipschitz bound

$$\sup_{t\geq 0} |||f_R^{(K)}(t,\cdot)|||_{\text{Lip},1} = |||\psi^{(K)}|||_{\text{Lip},1} \leq \|\psi\|_{\text{Lip,TV}}. \tag{3.1}$$

To prove the existence of the limit

$$f_R(t,\mu) = \lim_{K\to\infty} f_R^{(K)}\big(t, x^{(K)}(\mu)\big) \tag{3.2}$$

we will appropriately adapt the arguments in Section 3.2 of [29] and Section 3.3 of [12]. Given two integers $K' > K$, it will be convenient to introduce the projection map $P^{(K,K')} : \mathbb{R}_{\geq 0}^{\mathscr{D}_{K'}} \to \mathbb{R}_{\geq 0}^{\mathscr{D}_K}$ defined by

$$P^{(K,K')}x = x^{(K)}\big(\mu_x^{(K')}\big) \tag{3.3}$$

as well as the lifting map $L^{(K,K')} : \mathbb{R}^{\mathscr{D}_K} \to \mathbb{R}^{\mathscr{D}_{K'}}$ given by

$$L^{(K,K')}x = \big(\widetilde{x}_k\big)_{k\in\mathscr{D}_K}, \tag{3.4}$$

where $\widetilde{x}_k = (x_k,\ldots,x_k) \in \mathbb{R}^{2^{K'-K}}$. A key observation in proving the existence of the limit (3.2) is that

$$P^{(K,K')}L^{(K,K')}x = x. \tag{3.5}$$

The following technical lemmas will also play their part. The first two translate non-differential properties of a non-differentiable function into differential properties of a smooth function at any point where the difference of these functions is locally maximal. The third analyzes the transformation of the pairs (v,c) in the representation (2.16) of a closed convex set by the projection map (3.3), and the fourth shows that these pairs can be used to quantify the distance from a point to the closed convex set they define.

Lemma 3.1. *Fix a Lipschitz function $u \in C\big((0,\infty)\times\mathbb{R}_{\geq 0}^d\big)$ with $L = \sup_{t>0} |||u(t,\cdot)|||_{\mathrm{Lip},1} < \infty$. If $\phi \in C^\infty\big((0,\infty)\times\mathbb{R}_{>0}^d\big)$ is a smooth function with the property that $u-\phi$ has a local maximum at the point $(t^*,x^*) \in (0,\infty)\times\mathbb{R}_{>0}^d$, then $|||\nabla\phi(t^*,x^*)|||_{1,*} \leq L$. An identical statement holds at a local minimum.*

Proof. Since $u-\phi$ has a local maximum at $(t^*,x^*) \in (0,\infty)\times\mathbb{R}_{>0}^d$, for every $\varepsilon > 0$ small enough and $x \in \mathbb{R}_{\geq 0}^d$,

$$\phi(t^*,x^*+\varepsilon x) - \phi(t^*,x^*) \geq u(t^*,x^*+\varepsilon x) - u(t^*,x^*) \geq -\varepsilon L|||x|||_1.$$

Dividing by ε and letting ε tend to zero reveals that

$$\nabla\phi(t^*,x^*)\cdot x \geq -L|||x|||_1.$$

Choosing $x_k = -d\,\mathrm{sgn}\big(\partial_{x_k}\phi(t^*,x^*)\big)e_k$ for each $1 \leq k \leq d$ completes the proof. ∎

Lemma 3.2. *Fix a closed convex set $\mathscr{K}' \subset \mathbb{R}^d$ and a Lipschitz function $u \in C\big((0,\infty)\times\mathbb{R}_{\geq 0}^d\big)$ with $\nabla u \in L^\infty((0,\infty)\times\mathbb{R}_{\geq 0}^d;\mathscr{K}')$. Any smooth function $\phi \in C^\infty\big((0,\infty)\times\mathbb{R}_{>0}^d\big)$ with the property that $u-\phi$ has a local maximum at $(t^*,x^*) \in (0,\infty)\times\mathbb{R}_{>0}^d$ is such that $\nabla\phi(t^*,x^*) \in \mathscr{K}'$. An identical statement holds at a local minimum.*

Proof. Recall the representation (2.16) of the closed convex set $\mathscr{K}'$ as the intersection of the closed and affine half-spaces which contain it, and fix $(v,c) \in \mathscr{A}$. Since $u-\phi$ has a local maximum at $(t^*,x^*) \in (0,\infty)\times\mathbb{R}^d_{>0}$, for every $\varepsilon > 0$ small enough,

$$\phi(t^*,x^*+\varepsilon v)-\phi(t^*,x^*) \geq u(t^*,x^*+\varepsilon v)-u(t^*,x^*) \geq \varepsilon c.$$

The second inequality uses the characterization of $\nabla u \in \mathscr{K}'$ given in Proposition B.2 and the fact that $x\cdot\varepsilon v \geq \varepsilon c$ for all $x \in \mathscr{K}'$. Dividing by ε and letting ε tend to zero reveals that

$$\nabla\phi(t^*,x^*)\cdot v \geq c$$

for all $(v,c) \in \mathscr{A}$. It follows that $\nabla\phi(t^*,x^*) \in \mathscr{K}'$. This completes the proof. ∎

Lemma 3.3. *Fix two integers $K' > K$ large enough and a pair of vectors $(v,c) \in \mathbb{R}^{\mathscr{D}_{K'}}\times\mathbb{R}$ with $|||v|||_1 = 1$ such that, for every $x \in \mathscr{K}'_{a,K'}$, we have $v\cdot x \geq c$. Then, for every $y \in \mathscr{K}'_{a,K}$,*

$$P^{(K,K')}v\cdot y \geq c-\frac{2}{2^{K/2}}. \tag{3.6}$$

Proof. Fix $y \in \mathscr{K}_{a,K'}$, and find vectors $u^{(K)} \in \mathbb{R}^{\mathscr{D}_K}_{\geq 0}$ and $w^{(K)} \in \mathbb{R}^{\mathscr{D}_K}$ with

$$y = G^{(K)}u^{(K)}+w^{(K)}, \quad |||u^{(K)}|||_1 \leq a \quad \text{and} \quad |||w^{(K)}|||_{1,*} \leq \frac{1}{2^{K/2}}.$$

Consider the vector

$$u^{(K')}_{k'} = \frac{|\mathscr{D}_{K'}|}{|\mathscr{D}_K|}u^{(K)}_{k'}\mathbb{1}\{k' \in \mathscr{D}_K\}$$

in $\mathbb{R}^{\mathscr{D}_{K'}}$, and for each $k' \in \mathscr{D}_{K'}$, write $\underline{k}'$ for the unique dyadic $\underline{k}' \in \mathscr{D}_K$ with $k' \in [\underline{k}',\underline{k}'+2^{-K})$. Observe that

$$P^{(K,K')}v\cdot y = P^{(K,K')}v\cdot G^{(K)}u^{(K)}+P^{(K,K')}v\cdot w^{(K)} = v\cdot\big(G^{(K')}u^{(K')}+\alpha^{(K')}\big)+P^{(K,K')}v\cdot w^{(K)}$$

for the vector $\alpha^{(K')} \in \mathbb{R}^{\mathscr{D}_{K'}}$ defined by

$$\alpha^{(K')}_{k'} = \frac{1}{|\mathscr{D}_{K'}|^2}\sum_{k'\in\mathscr{D}_{K'}}\big(g(\underline{k}k')-g(kk')\big)u^{(K')}_{k'}.$$

Since $G^{(K')}u^{(K')} \in \mathscr{K}'_{a,K'}$, the defining property of v, the fact that $|||v|||_1 = 1$ and Hölder's inequality give the lower bound

$$P^{(K,K')}v\cdot y \geq c-|||\alpha^{(K')}|||_{1,*}-|||w^{(K)}|||_{1,*} \geq c-|||\alpha^{(K')}|||_{1,*}-\frac{1}{2^{K/2}}, \tag{3.7}$$

where we have used that $|||P^{(K,K')}v|||_1 \leq |||v|||_1$ and $|||w^{(K)}|||_{1,*} \leq 2^{-K/2}$. The mean value theorem reveals that

$$|||\alpha^{(K')}|||_{1,*} \leq \frac{\|g'\|_\infty}{2^K}|||u^{(K)}|||_1 \leq \frac{a\|g'\|_\infty}{2^K}.$$

Substituting this into (3.7) and taking K large enough completes the proof. ∎

Lemma 3.4. *Fix a closed convex set $\mathscr{K}' \subset \mathbb{R}^d$, and recall its representation* (2.16) *with* $\|\cdot\| = |||\cdot|||_1$ *as the intersection of the closed and affine half-spaces which contain it. If $x \in \mathbb{R}^d$ and $\varepsilon > 0$ are such that $x \cdot v \geq c - \varepsilon$ for all $(v,c) \in \mathscr{A}$, then there exist $y \in \mathscr{K}'$ and $z \in \mathbb{R}^d$ with $x = y + z$ and* $|||z|||_{1,*} \leq \varepsilon$.

Proof. Let $y \in \mathscr{K}'$ denote a projection of $x \in \mathbb{R}^d$ onto the set $\mathscr{K}'$ with respect to the normalized-$\ell^{1,*}$ norm. More precisely, let $y \in \mathscr{K}'$ be any minimizer of the map $y' \mapsto |||y' - x|||_{1,*}$ over points $y' \in \mathscr{K}'$. The existence of such a projection is guaranteed by the fact that $\mathscr{K}'$ is closed. If $y = x$, then the desired conclusion is immediate, so from now on we assume that $y \neq x$. Introduce the set

$$\mathscr{I} = \big\{k \leq d \mid d|x_k - y_k| = |||x - y|||_{1,*}\big\}$$

of indices at which $|||x - y|||_{1,*}$ is achieved, and define the vector $v \in \mathbb{R}^d$ by

$$v_k = \frac{d}{|\mathscr{I}|} \operatorname{sgn}(y_k - x_k) \mathbb{1}\{k \in \mathscr{I}\}.$$

We now show that $(v,c) \in \mathscr{A}$ for $c = v \cdot y$. By construction $|||v|||_1 = 1$, so suppose for the sake of contradiction that there exists $y' \in \mathscr{K}'$ with $(y' - y) \cdot v = y' \cdot v - c < 0$. This means that

$$(y - y') \cdot v = \frac{d}{|\mathscr{I}|} \sum_{k \in \mathscr{I}} (y_k - y'_k) \operatorname{sgn}(y_k - x_k) > 0.$$

In particular, a coordinate $k^* \in \mathscr{I}$ at which the quantity $(y_k - y'_k) \operatorname{sgn}(y_k - x_k)$ is maximized over $k \in \mathscr{I}$ must satisfy

$$(y_{k^*} - y'_{k^*}) \operatorname{sgn}(y_{k^*} - x_{k^*}) > 0.$$

At this point, fix $t \in (0,1)$ small enough so that $\operatorname{sgn}(y_k - x_k + t(y'_k - y_k)) = \operatorname{sgn}(y_k - x_k)$ for every $k \in \mathscr{I}$. For such a value of $t > 0$,

$$|||y - x + t(y' - y)|||_{1,*} = d(y_{k^*} - x_{k^*} + t(y'_{k^*} - y_{k^*})) \operatorname{sgn}(y_{k^*} - x_{k^*}) < |||x - y|||_{1,*}.$$

Since the point $y'' = y + t(y' - y)$ is a convex combination of $y, y' \in \mathscr{K}'$, it must lie in the convex set $\mathscr{K}'$. This contradicts the fact that y minimizes the map $y'' \mapsto |||y'' - x|||_{1,*}$ over points $y'' \in \mathscr{K}'$, and shows that $(v,c) \in \mathscr{A}$. It follows that

$$\varepsilon \geq c - x \cdot v = v \cdot (y - x) = |||x - y|||_{1,*}.$$

Setting $z = x - y$ completes the proof. ■

We are now in a position to prove Theorem 1.2.

Proof of Theorem 1.2. To alleviate notation, until otherwise stated, we fix $R > \|\psi\|_{\mathrm{Lip,TV}}$ and keep all dependencies on R implicit. The existence of the limit (3.2) will be established by showing that the sequence $(f^{(K)}(t, x^{(K)}(\mu)))_K$ is Cauchy. With this in mind, fix $K' > K$ and introduce the function

$$f^{(K,K')}(t,x) = f^{(K)}\big(t, P^{(K,K')}x\big) \tag{3.8}$$

defined on $\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}^{\mathscr{D}_{K'}}$. Since $x^{(K)}(\mu)=P^{(K,K')}x^{(K')}(\mu)$, the Cauchy condition may be expressed in terms of this function as

$$\left|f^{(K')}\big(t,x^{(K')}(\mu)\big)-f^{(K)}\big(t,x^{(K)}(\mu)\big)\right|=\left|f^{(K')}\big(t,x^{(K')}(\mu)\big)-f^{(K,K')}\big(t,x^{(K')}(\mu)\big)\right|. \tag{3.9}$$

To control the right-hand side of this expression, we will first show that $f^{(K,K')}$ is an approximate viscosity solution to the Hamilton-Jacobi equation (2.5) satisfied by $f^{(K')}$, and then we will leverage the comparison principle in Corollary A.12.

<u>*Step 1: $f^{(K,K')}$ is an approximate viscosity solution.*</u>

Consider a function $\phi_{K'}\in C^\infty\big((0,\infty)\times\mathbb{R}_{>0}^{\mathscr{D}_{K'}}\big)$ with the property that $f^{(K,K')}-\phi_{K'}$ achieves a local maximum at $(t^*,x^*)\in(0,\infty)\times\mathbb{R}_{>0}^{\mathscr{D}_{K'}}$. To be more precise, suppose that

$$\sup_{B_{K'}(r)}\big(f^{(K,K')}-\phi_{K'}\big)=\big(f^{(K,K')}-\phi_{K'}\big)(t^*,x^*),$$

where

$$B_{K'}(r)=\Big\{(t,x)\in(0,\infty)\times\mathbb{R}_{\geq 0}^{\mathscr{D}_{K'}}\mid |t-t^*|+|||x-x^*|||_1\leq r\Big\}$$

is the ball of radius $r>0$ centered at (t^*,x^*). Decreasing $r>0$ if necessary, assume without loss of generality that

$$B_{K'}(r)\subset(0,\infty)\times\mathbb{R}_{>0}^{\mathscr{D}_{K'}}.$$

Assume also that $\phi_{K'}\in C^\infty\big((0,\infty)\times\mathbb{R}^{\mathscr{D}_{K'}}\big)$; this can be ensured by replacing $\phi_{K'}$ with $\eta\phi_{K'}$ for some $\eta\in C^\infty\big(\mathbb{R}^{\mathscr{D}_{K'}}\big)$ which is identically one on $B_{K'}(r)$ and vanishes outside $\mathbb{R}_{\geq 0}^{\mathscr{D}_{K'}}$. With these simplifications at hand, introduce the smooth function

$$\phi_K(t,y)=\phi_{K'}\big(t,x^*+L^{(K,K')}y-L^{(K,K')}P^{(K,K')}x^*\big)$$

defined on $(0,\infty)\times\mathbb{R}_{>0}^{\mathscr{D}_K}$. We will now show that the function ϕ_K admits a local maximum at $(t^*,P^{(K,K')}x^*)$. It will be convenient to notice that for any $y\in\mathbb{R}_{\geq 0}^{\mathscr{D}_K}$

$$P^{(K,K')}\big(x^*+L^{(K,K')}y-L^{(K,K')}P^{(K,K')}x^*\big)=P^{(K,K')}x^*+y-P^{(K,K')}x^*=y\in\mathbb{R}_{>0}^{\mathscr{D}_K} \tag{3.10}$$

by (3.5). To simplify notation, let $y^*=P^{(K,K')}x^*\in\mathbb{R}_{\geq 0}^{\mathscr{D}_K}$ and introduce the ball

$$B_K(r)=\big\{(t,y)\in(0,\infty)\times\mathbb{R}_{>0}^{\mathscr{D}_K}\mid |t-t^*|+|||y-y^*|||_1\leq r\big\}\subset(0,\infty)\times\mathbb{R}_{>0}^{\mathscr{D}_K}$$

of radius $r>0$ centered at (t^*,y^*). Given $(t,y)\in B_K(r)$, let $z_y=x^*+L^{(K,K')}y-L^{(K,K')}P^{(K,K')}x^*$ in such a way that by (3.10),

$$f^{(K)}(t,y)-\phi_K(s,y)=f^{(K,K')}(t,z_y)-\phi_{K'}(t,z_y).$$

Observe that

$$|t-t^*|+|||z_y-x^*|||_1=|t-t^*|+|||L^{(K,K')}y-L^{(K,K')}y^*|||_1=|t-t^*|+|||y-y^*|||_1\leq r$$

so $(t,z_y)\in B_{K'}(r)$. It follows that

$$\sup_{B_K(r)}\big(f^{(K)}-\phi_K\big)\leq\sup_{B_{K'}(r)}\big(f^{(K,K')}-\phi_{K'}\big)=\big(f^{(K,K')}-\phi_{K'}\big)(t^*,x^*)=\big(f^{(K)}-\phi_K\big)(t^*,y^*)$$

which means that ϕ_K admits a local maximum at $(t^*, y^*) \in (0,\infty) \times \mathbb{R}^{\mathscr{D}_K}_{>0}$. Since $f^{(K)}$ is a viscosity subsolution to the Hamilton-Jacobi equation (2.5) with $\nabla f^{(K)} \in \mathscr{K}'_{a,K}$, Lemma 3.2 implies that

$$\nabla \phi_K(t^*, y^*) \in \mathscr{K}'_{a,K} \quad \text{and} \quad \big(\partial_t \phi_K - \mathsf{H}_K\big(\nabla \phi_K\big)\big)(t^*, y^*) \le 0.$$

To write this expression in terms of the original test function $\phi_{K'}$, notice that

$$\partial_t \phi_K(t^*, y^*) = \partial_t \phi_{K'}(t^*, x^*) \quad \text{and} \quad \nabla \phi_K(t^*, y^*) = \frac{|\mathscr{D}_{K'}|}{|\mathscr{D}_K|} P^{(K,K')} \nabla \phi_{K'}(t^*, x^*).$$

This means that

$$\frac{|\mathscr{D}_{K'}|}{|\mathscr{D}_K|} P^{(K,K')} \nabla \phi_{K'}(t^*, x^*) \in \mathscr{K}'_{a,K} \quad \text{and} \quad \partial_t \phi_{K'}(t^*, x^*) - \mathsf{H}_K\left(\frac{|\mathscr{D}_{K'}|}{|\mathscr{D}_K|} P^{(K,K')} \nabla \phi_{K'}(t^*, x^*)\right) \le 0.$$

The first of these conditions gives vectors $u^{(K)} \in \mathbb{R}^{\mathscr{D}_K}_{\ge 0}$ and $w^{(K)} \in \mathbb{R}^{\mathscr{D}_K}$ with

$$\frac{|\mathscr{D}_{K'}|}{|\mathscr{D}_K|} P^{(K,K')} \nabla \phi_{K'}(t^*, x^*) = G^{(K)} u^{(K)} + w^{(K)}, \quad |||u^{(K)}|||_1 \le a \quad \text{and} \quad |||w^{(K)}|||_{1,*} \le \frac{1}{2^{K/2}}. \tag{3.11}$$

Observe that

$$|||G^{(K)} u^{(K)}|||_{1,*} \le \frac{|\mathscr{D}_{K'}|}{|\mathscr{D}_K|} |||P^{(K,K')} \nabla \phi_{K'}(t^*, x^*)|||_{1,*} + \frac{1}{2^{K/2}} \le |||\nabla \phi_{K'}(t^*, x^*)|||_{1,*} + \frac{1}{2^{K/2}}.$$

Since $f^{(K,K')} - \phi_{K'}$ achieves a local maximum at (t^*, x^*), Lemma 3.1 and (3.1) imply that

$$|||\nabla \phi_{K'}(t^*, x^*)|||_{1,*} \le \|\psi\|_{\mathrm{Lip,TV}}.$$

Recalling that $R > \|\psi\|_{\mathrm{Lip,TV}}$ and taking K large enough ensures that $G^{(K)} u^{(K)} \in \mathscr{C}_K \cap B_R$. It follows by the Lipschitz continuity of H_K established in Proposition 2.3 that

$$\begin{aligned} \partial_t \phi_{K'}(t^*, x^*) - \mathsf{C}_K\big(G^{(K)} u^{(K)}\big) &\le \partial_t \phi_{K'}(t^*, x^*) - \mathsf{H}_K\left(\frac{|\mathscr{D}_{K'}|}{|\mathscr{D}_K|} P^{(K,K')} \nabla \phi_{K'}(t^*, x^*)\right) + \frac{8RM}{2^{K/2} m^2} \\ &\le \frac{8RM}{2^{K/2} m^2}. \end{aligned}$$

At this point, introduce the vector $u^{(K')} \in \mathbb{R}^{\mathscr{D}_{K'}}_{\ge 0}$ defined by

$$u^{(K')}_{k'} = \frac{|\mathscr{D}_{K'}|}{|\mathscr{D}_K|} u^{(K)}_{k'} \mathbb{1}\{k' \in \mathscr{D}_K\}$$

in such a way that

$$\mathsf{C}_{K'}\big(G^{(K')} u^{(K')}\big) = \frac{1}{|\mathscr{D}_K|^2} \sum_{k,k' \in \mathscr{D}_K} g(kk') u^{(K)}_k u^{(K)}_{k'} = \mathsf{C}_K\big(G^{(K)} u^{(K)}\big),$$

and therefore,

$$\partial_t \phi_{K'}(t^*, x^*) - \mathsf{C}_{K'}\big(G^{(K')} u^{(K')}\big) \le \frac{8RM}{2^{K/2} m^2}. \tag{3.12}$$

We now show that, up to an error vanishing with K, the term $G^{(K')}u^{(K')}$ in this expression may be replaced by $\nabla\phi_{K'}(t^*,x^*)$. This is where Lemma 3.3 will play its part. Recall the representation (2.16) with $\|\cdot\| = |||\cdot|||_1$ of $\mathscr{K}'_{a,K'}$ as the intersection of the closed and affine half-spaces which contain it, and fix $(v,c)\in\mathscr{A}$ with $|||v|||_1 = 1$. The characterization of $\nabla f^{(K)} \in \mathscr{K}'_{a,K}$ given in Proposition B.2 and Lemma 3.3 imply that for every $\varepsilon > 0$ small enough,

$$
\begin{aligned}
\phi_{K'}(t^*,x^*+\varepsilon v) - \phi_{K'}(t^*,x^*) &\geq f^{(K)}\big(t^*,P^{(K,K')}x^* + \varepsilon P^{(K,K')}v\big) - f^{(K)}\big(t^*,P^{(K,K')}x^*\big)\\
&\geq \varepsilon\left(c - \frac{2}{2^{K/2}}\right)
\end{aligned}
$$

Dividing by ε and letting ε tend to zero reveals that $\nabla\phi_{K'}(t^*,x^*)\cdot v \geq c - \frac{2}{2^{K/2}}$. Invoking Lemma 3.4 gives $\alpha^{(K')} \in \mathbb{R}^{\mathscr{D}_{K'}}_{\geq 0}$ and $\beta^{(K')}\in\mathbb{R}^{\mathscr{D}_{K'}}$ with

$$
\nabla\phi_{K'}(t^*,x^*) = G^{(K')}\alpha^{(K')} + \beta^{(K')}, \quad |||\alpha^{(K')}|||_1 \leq a \quad \text{and} \quad |||\beta^{(K')}|||_{1,*} \leq \frac{2}{2^{K/2}}. \tag{3.13}
$$

At this point, fix $k\in\mathscr{D}_K$ and $k'\in[k,k+2^{-K})$. The mean value theorem implies that

$$
\begin{aligned}
|\mathscr{D}_{K'}|\big|\big(G^{(K')}u^{(K')}\big)_{k'} - \partial_{x_{k'}}\phi_{K'}(t^*,x^*)\big| &= \left|\frac{1}{|\mathscr{D}_K|}\sum_{k''\in\mathscr{D}_K} g(k'k'')u^{(K)}_{k''} - |\mathscr{D}_{K'}|\partial_{x_{k'}}\phi_{K'}(t^*,x^*)\right|\\
&\leq \big||\mathscr{D}_K|\big(G^{(K)}u^{(K)}\big)_k - |\mathscr{D}_{K'}|\partial_{x_{k'}}\phi_{K'}(t^*,x^*)\big|\\
&\qquad + \frac{\|g'\|_\infty |||u^{(K)}|||_1}{2^K}.
\end{aligned}
$$

Remembering (3.11) and (3.13), noticing that $|\mathscr{D}_K| = 2^{K+1}$ and using the mean value theorem once again shows that

$$
\begin{aligned}
|\mathscr{D}_K|\big(G^{(K)}u^{(K)}\big)_k &= |\mathscr{D}_K|\sum_{\ell=0}^{2^{K'-K}-1}\partial_{x_{k+\frac{\ell}{2^{K'}}}}\phi_{K'}(t^*,x^*) - |\mathscr{D}_K|w^{(K)}_k\\
&= \frac{|\mathscr{D}_K|}{|\mathscr{D}_{K'}|^2}\sum_{\ell=0}^{2^{K'-K}-1}\sum_{k''\in\mathscr{D}_{K'}} g\Big(k+\frac{\ell}{2^{K'}}\cdot k''\Big)\alpha^{(K')}_{k''} + |\mathscr{D}_K|\sum_{\ell=0}^{2^{K'-K}-1}\beta^{(K')}_{k+\frac{\ell}{2^{K'}}} - |\mathscr{D}_K|w^{(K)}_k\\
&= \frac{|\mathscr{D}_K|}{|\mathscr{D}_{K'}|^2}\sum_{\ell=0}^{2^{K'-K}-1}\sum_{k''\in\mathscr{D}_{K'}} g(k'k'')\alpha^{(K')}_{k''} + \mathscr{O}_1\left(\frac{\|g'\|_\infty|||\alpha^{(K')}|||_1}{2^K} + \frac{3}{2^{K/2}}\right)\\
&= |\mathscr{D}_{K'}|\big(G^{(K')}\alpha^{(K')}\big)_{k'} + \mathscr{O}_1\left(\frac{4}{2^{K/2}}\right) = |\mathscr{D}_{K'}|\partial_{x_{k'}}\phi_{K'}(t^*,x^*) + \mathscr{O}_1\left(\frac{5}{2^{K/2}}\right),
\end{aligned}
$$

where we have written $X = Y + \mathscr{O}_1(Z)$ to mean that $|X-Y|\leq Z$. In the third equality we used that $|||\beta^{(K')}|||_{1,*} + |||w^{(K)}|||_{1,*} \leq 3\cdot 2^{-K/2}$, and in the fourth equality we used that $|||\alpha^{(K')}|||_1 \leq a$ and increased K if necessary. It follows that

$$
|||G^{(K')}u^{(K')} - \nabla\phi_{K'}(t^*,x^*)|||_{1,*} \leq \frac{5}{2^{K/2}} + \frac{a\|g'\|_\infty}{2^K} \leq \frac{6}{2^{K/2}},
$$

where we have used that $|||u^{(K)}|||_1 \leq a$ and increased K if necessary. Combining this bound with the Lipschitz continuity of $\mathsf{H}_{K'}$ established in Proposition 2.3 and (3.12) reveals that

$$\partial_t \phi_{K'}(t^*,x^*) - \mathsf{H}_{K'}\big(\nabla \phi_{K'}(t^*,x^*)\big) \leq \mathscr{E}_K \tag{3.14}$$

for the error term

$$\mathscr{E}_K = \frac{56RM}{2^{K/2}m^2}.$$

In particular, the function $(t,x) \mapsto f^{(K,K')}(t,x) - \mathscr{E}_K t$ is a viscosity subsolution to the Hamilton-Jacobi equation (2.5) satisfied by $f^{(K')}$. An identical argument shows that $(t,x) \mapsto f^{(K,K')}(t,x) + \mathscr{E}_K t$ is a viscosity supersolution to the Hamilton-Jacobi equation (2.5) satisfied by $f^{(K')}$.

Step 2: comparison principle.

Using (3.1) and (1.36), it is readily verified that $f^{(K,K')}$ and $f^{(K')}$ are uniformly Lipschitz continuous in the x variable relative to the normalized-ℓ^1 norm with Lipschitz constant at most $L = \|\psi\|_{\mathrm{Lip,TV}}$. Indeed, for any $t > 0$ and all $x,x' \in \mathbb{R}^{\mathscr{D}_{K'}}_{\geq 0}$,

$$|f^{(K,K')}(t,x) - f^{(K,K')}(t,x')| \leq \|\psi\|_{\mathrm{Lip,TV}} |||P^{(K,K')}x - P^{(K,K')}x'|||_1 \leq \|\psi\|_{\mathrm{Lip,TV}} |||x-x'|||_1.$$

If $V = |||\mathsf{H}_K|||_{\mathrm{Lip},1,*}$, then the comparison principle in Corollary A.12 implies that for any $R' \in \mathbb{R}$, the map

$$(t',x') \mapsto f^{(K,K')}(t',x') - f^{(K')}(t',x') - (2L+1)\big(|||x'|||_1 + Vt' - R'\big)_+ - \mathscr{E}_K t' \tag{3.15}$$

achieves its supremum on $\{0\} \times \mathbb{R}^{\mathscr{D}_{K'}}_{\geq 0}$. We now choose $R' = |||x^{(K')}(\mu)|||_1 + Vt$ and distinguish two cases. On the one hand, if $t' = 0$ and $|||x'|||_1 \geq (2L+1)R'$, then (3.15) is bounded by

$$2L|||x'|||_1 - (2L+1)(|||x'|||_1 - R') = R' - |||x'|||_1 \leq 0, \tag{3.16}$$

where we have used the fact that $f^{(K,K')}(0,0) = f^{(K')}(0,0)$. On the other hand, if $t' = 0$ and $|||x'|||_1 \leq (2L+1)R'$, then (**H4**) implies that (3.15) is bounded by

$$\big|\psi^{(K')}(x') - \psi^{(K)}(P^{(K,K')}x')\big| \leq \|\psi\|_{\mathrm{Lip},W} |||x'|||_1 W\big(\bar{\mu}^{(K')}_{x'}, \bar{\mu}^{(K)}_{P^{(K,K')}x'}\big). \tag{3.17}$$

To estimate this Wasserstein distance, fix a Lipschitz function $h : [-1,1] \to \mathbb{R}$ with $\|h\|_{\mathrm{Lip}} \leq 1$ and observe that

$$\begin{aligned}\Big|\int_{-1}^1 h(y)\,\mathrm{d}\big(\mu^{(K')}_{x'} - \mu^{(K)}_{P^{(K,K')}x'}\big)(y)\Big| &\leq \frac{1}{|\mathscr{D}_{K'}|}\sum_{k\in\mathscr{D}_K}\sum_{\ell=0}^{2^{K'-K-1}} x'_{k+\frac{\ell}{2^{K'}}}\Big|h\Big(k+\frac{\ell}{2^{K'}}\Big) - h(k)\Big| \\ &\leq \frac{1}{|\mathscr{D}_{K'}|}\sum_{k\in\mathscr{D}_K}\sum_{\ell=0}^{2^{K'-K-1}} x'_{k+\frac{\ell}{2^{K'}}}\frac{\ell}{2^{K'}} \leq \frac{|||x'|||_1}{2^K}.\end{aligned}$$

Taking the supremum over all such h and recalling (3.17) shows that (3.15) is bounded by

$$\frac{\|\psi\|_{\mathrm{Lip},W}(2L+1)R'}{2^K} \tag{3.18}$$

whenever $t' = 0$ and $|||x'|||_1 \leq (2L+1)R'$. Combining this with (3.16) reveals that the map (3.15) is uniformly bounded by (3.18). Choosing $t' = t$ and $x' = x^{(K')}(\mu)$, and recalling the choice of R' yields

$$f^{(K,K')}\big(t, x^{(K')}(\mu)\big) - f^{(K')}\big(t, x^{(K')}(\mu)\big) \leq \frac{\|\psi\|_{\mathrm{Lip},W}(2L+1)}{2^K}\big(|||x^{(K')}(\mu)|||_1 + Vt\big) + \mathscr{E}_K t.$$

Together with (3.9) and an identical argument with the roles of $f^{(K,K')}$ and $f^{(K')}$ reversed, this implies that

$$\big|f^{(K')}\big(t, x^{(K')}(\mu)\big) - f^{(K)}\big(t, x^{(K)}(\mu)\big)\big| \leq \frac{\|\psi\|_{\mathrm{Lip},W}(2L+1)}{2^K}\big(\mu[-1,1] + Vt\big) + \mathscr{E}_K t.$$

Since $V = |||\mathsf{H}_K|||_{\mathrm{Lip},1,*}$ is independent of K by Proposition 2.3 and $\mathscr{E}_K$ tends to zero as K tends to infinity, the sequence $(f^{(K)}(t, x^{(K)}(\mu)))_K$ is Cauchy. This establishes the existence of the limit (3.2) for each fixed $R > \|\psi\|_{\mathrm{Lip,TV}}$. All that remains is to show that this limit is independent of R.

Step 3: independence on R.

To show that the limit (3.2) is independent of R, fix $R' > R > \|\psi\|_{\mathrm{Lip,TV}}$ as well as $K \geq 1$ large enough. The idea will be to show that, up to an error vanishing with K, the function $f_R^{(K)}$ satisfies the Hamilton-Jacobi equation defining $f_{R'}^{(K)}$. The equality of the limit (3.2) associated with R and R' will then follow from the comparison principle in Corollary A.12. Consider $\phi \in C^\infty\big((0,\infty) \times \mathbb{R}_{\geq 0}^{\mathscr{D}_K}\big)$ with the property that $f_R^{(K)} - \phi$ achieves a local maximum at the point $(t^*, x^*) \in (0,\infty) \times \mathbb{R}_{>0}^{\mathscr{D}_K}$. Since $f_R^{(K)}$ is a viscosity subsolution to the Hamilton-Jacobi equation (2.5) associated with the non-linearity $\mathsf{H}_{K,R}$,

$$\big(\partial_t \phi - \mathsf{H}_{K,R}(\nabla \phi)\big)(t^*, x^*) \leq 0.$$

The fact that $f_R^{(K)}$ has its gradient in the set $\mathscr{K}'_{a,K}$ together with (3.1), Lemma 3.1 and Lemma 3.2 implies that

$$\nabla \phi(t^*, x^*) \in \mathscr{K}'_{a,K} \quad \text{and} \quad |||\nabla \phi(t^*, x^*)|||_{1,*} \leq \|\psi\|_{\mathrm{Lip,TV}}.$$

It is therefore possible to find $u \in \mathbb{R}_{\geq 0}^{\mathscr{D}_K}$ and $w \in \mathbb{R}^{\mathscr{D}_K}$ with

$$\nabla \phi(t^*, x^*) = G^{(K)} u + w, \quad |||u|||_1 \leq a \quad \text{and} \quad |||w|||_{1,*} \leq \frac{1}{2^{K/2}}.$$

Observe that

$$|||G^{(K)} u|||_{1,*} \leq |||\nabla \phi(t^*, x^*)|||_{1,*} + |||w|||_{1,*} \leq \|\psi\|_{\mathrm{Lip,TV}} + \frac{1}{2^{K/2}},$$

so increasing K if necessary, it is possible to ensure that $G^{(K)} u \in \mathscr{C}_K \cap B_R \subset \mathscr{C}_K \cap B_{R'}$. It follows by the Lipschitz continuity of $\mathsf{H}_{K,R}$ established in Proposition 2.3 that

$$\begin{aligned}\big(\partial_t \phi - \mathsf{H}_{K,R'}(\nabla \phi)\big)(t^*, x^*) &\leq \partial_t \phi(t^*, x^*) - \mathsf{C}_K\big(G^{(K)} u\big) + \frac{8R'M}{2^{K/2} m^2} \\ &\leq \big(\partial_t \phi - \mathsf{H}_{K,R}(\nabla \phi)\big)(t^*, x^*) + \frac{8(R'+R)M}{2^{K/2} m^2} \leq \mathscr{E}_K\end{aligned}$$

for the error term

$$\mathscr{E}_K = \frac{8(R'+R)M}{2^{K/2} m^2}.$$

In particular, the function $(t,x) \mapsto f_R^K(t,x) - \mathscr{E}_K t$ is a viscosity subsolution to the Hamilton-Jacobi equation (2.5) defining $f_{R'}^K$. An identical argument shows that $(t,x) \mapsto f_R^K(t,x) + \mathscr{E}_K t$ is a viscosity supersolution to the Hamilton-Jacobi equation (2.5) defining $f_{R'}^K$. It follows by the comparison principle in Corollary A.12 that for every $\mu \in \mathscr{M}_+$ and $t \geq 0$,

$$\left| f_R^{(K)}\big(t, x^{(K)}(\mu)\big) - f_{R'}^{(K)}\big(t, x^{(K)}(\mu)\big) \right| \leq \mathscr{E}_K t.$$

Letting K tend to infinity completes the proof. ■

4 Approximate Hopf-Lax formula in finite dimensions

In this section, we revisit the Hamilton-Jacobi equation studied in Section 2, and under the additional assumption that the matrix $G \in \mathbb{R}^{d\times d}$ in (2.1) is non-negative definite, we establish an approximate Hopf-Lax formula for the unique solution to the Hamilton-Jacobi equation (2.5). By an approximate Hopf-Lax formula we mean that the error between the Hopf-Lax function we will define and the solution to the Hamilton-Jacobi equation constructed in Proposition 2.7 tends to zero when the dimension d tends to infinity. This will be used in the next section to establish an exact Hopf-Lax formula for the solution to the infinite-dimensional Hamilton-Jacobi equation (1.10). It will be convenient to introduce the bilinear form

$$(x,y)_G = Gx \cdot y \tag{4.1}$$

associated with the non-negative definite matrix G, as well as its induced semi-norm

$$\|x\|_G = \sqrt{(x,x)_G}. \tag{4.2}$$

In this notation, the non-linearity (2.4) may be written as

$$\mathsf{C}(Gx) = \frac{1}{2} Gx \cdot x = \frac{1}{2}\|x\|_G^2. \tag{4.3}$$

In particular, the non-linearity (2.4) is a convex function. This convexity property will allow us to establish an approximate Hopf-Lax formula for the Hamilton-Jacobi equation (2.5). We define the Hopf-Lax function $f_{\mathrm{HL}} : [0,\infty) \times \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ by

$$f_{\mathrm{HL}}(t,x) = \sup_{y \in \mathbb{R}^d_{\geq 0}} \left\{ \psi(x+y) - \frac{\|y\|_G^2}{2t} \right\}. \tag{4.4}$$

The main result of this section is the following.

Proposition 4.1 (Hopf-Lax)**.** *Fix an initial condition $\psi : \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ satisfying (2.6) with its gradient in the set $\mathscr{K}_a'$, let $R > |||\psi|||_{\mathrm{Lip},1}$, and denote by $f : [0,\infty) \times \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ the unique solution to the Hamilton-Jacobi equation (2.5) constructed in Proposition 2.7. If $G \in \mathbb{R}^{d\times d}$ is non-negative definite, then for all $(t,x) \in [0,\infty) \times \mathbb{R}^d_{\geq 0}$,*

$$|f(t,x) - f_{\mathrm{HL}}(t,x)| \leq \frac{t}{\sqrt{d}}\left(R + a + \frac{8RM}{m^2}\right). \tag{4.5}$$

To prove this result, we first verify that the convex dual of the mapping $y \mapsto \|y\|_G^2/2$ is the non-linearity C. We next show that the function (4.4) satisfies the right initial condition, and that the supremum in its definition is attained. We next argue that this function satisfies a semigroup property, and deduce that it belongs to $\mathfrak{L}_{\mathrm{unif}}$. Finally we show that, in a sense to be made precise, it is an approximate solution to the the Hamilton-Jacobi equation (2.5). The estimate (4.5) will then be a consequence of the comparison principle in Corollary A.12. It will be convenient to note that for every $z \in \mathbb{R}^d_{\geq 0}$, we have

$$\|z\|_G^2 \geq \frac{m}{d^2} \sum_{k,k'=1}^{d} z_k z_{k'} = m\,|||z|||_1^2. \tag{4.6}$$

Lemma 4.2. *If $G \in \mathbb{R}^{d\times d}$ is non-negative definite, then for every $z \in \mathscr{C}$,*

$$\mathsf{C}(z) = \sup_{y \in \mathbb{R}^d_{\geq 0}} \left\{ y \cdot z - \frac{\|y\|_G^2}{2} \right\}. \tag{4.7}$$

Moreover, the supremum is attained at any point $x \in \mathbb{R}^d_{\geq 0}$ with $z = Gx$.

Proof. We can represent each $z \in \mathscr{C}$ in the form of Gx for some $x \in \mathbb{R}^d_{\geq 0}$. Using that G is non-negative definite, we can appeal to the Cauchy-Schwarz inequality to assert that

$$y \cdot Gx = (x,y)_G \leq \|x\|_G \, \|y\|_G \leq \frac{1}{2}\|x\|_G^2 + \frac{1}{2}\|y\|_G^2.$$

We thus obtain that

$$\mathsf{C}(Gx) = \frac{1}{2}\|x\|_G^2 \geq \sup_{y \in \mathbb{R}^d_{\geq 0}} \left(y \cdot Gx - \frac{\|y\|_G^2}{2} \right).$$

For the converse inequality, we simply test the supremum with $y = x$. ■

Lemma 4.3. *If $G \in \mathbb{R}^{d\times d}$ is non-negative definite and $\psi : \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ is a Lipschitz continuous initial condition with $\nabla\psi \in L^\infty(\mathbb{R}^d_{\geq 0}; \mathscr{K}_a')$, then for every $x \in \mathbb{R}^d_{\geq 0}$,*

$$f_{\mathrm{HL}}(0,x) = \psi(x). \tag{4.8}$$

Proof. For $t = 0$, we interpret the definition of f as

$$f_{\mathrm{HL}}(0,x) = \sup_{\substack{y \in \mathbb{R}^d_{\geq 0} \\ \|y\|_G = 0}} \psi(x+y). \tag{4.9}$$

Recalling (4.6), we see that the only $y \in \mathbb{R}^d_{\geq 0}$ with $\|y\|_G = 0$ is $y = 0$. Together with (4.9), this completes the proof. ■

It will slightly simplify our arguments below to notice that the supremum in (4.4) is achieved.

Lemma 4.4. *Fix an initial condition $\psi : \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ satisfying (2.6). If $G \in \mathbb{R}^{d\times d}$ is non-negative definite, then for any point $(t,x) \in (0,\infty) \times \mathbb{R}^d_{\geq 0}$, there exists $y \in \mathbb{R}^d_{\geq 0}$ with*

$$f_{\mathrm{HL}}(t,x) = \psi(x+y) - \frac{\|y\|_G^2}{2t}. \tag{4.10}$$

Proof. Combining (4.6) with the Lipschitz continuity of ψ reveals that

$$\psi(x+y) - \frac{\|y\|_G^2}{2t} \le \psi(x) + |||y|||_1 \Big(|||\psi|||_{\mathrm{Lip},1} - \frac{m}{2t} |||y|||_1 \Big).$$

We can thus restrict the supremum in (4.4) to those y's in $\mathbb{R}^d_{\ge 0}$ that satisfy $|||y|||_1 \le \frac{2t}{m} |||\psi|||_{\mathrm{Lip},1}$. Since we are now optimizing a continuous function over a compact set, it is clear that the supremum is achieved. ■

Lemma 4.5 (Semigroup property)**.** *Fix an initial condition $\psi : \mathbb{R}^d_{\ge 0} \to \mathbb{R}$ satisfying (2.6). If $G \in \mathbb{R}^{d\times d}$ is non-negative definite, then for every pair $t > s > 0$ and $x \in \mathbb{R}^d_{\ge 0}$,*

$$f_{\mathrm{HL}}(t,x) = \sup_{y\in\mathbb{R}^d_{\ge 0}} \left\{ f_{\mathrm{HL}}(s,x+y) - \frac{\|y\|_G^2}{2(t-s)} \right\}. \tag{4.11}$$

Proof. Fix $y,z \in \mathbb{R}^d_{\ge 0}$. Since $\|\cdot\|_G^2$ is a convex mapping, we have

$$\left\| \frac{y+z}{t} \right\|_G^2 \le \frac{s}{t} \left\| \frac{y}{s} \right\|_G^2 + \frac{t-s}{t} \left\| \frac{z}{t-s} \right\|_G^2.$$

Substituting this bound into (4.4) yields

$$f_{\mathrm{HL}}(t,x) \ge \psi(x+y+z) - \frac{t}{2} \left\| \frac{y+z}{t} \right\|_G^2 \ge \psi(x+y+z) - \frac{s}{2} \left\| \frac{y}{s} \right\|_G^2 - \frac{t-s}{2} \left\| \frac{z}{t-s} \right\|_G^2.$$

Taking the supremum over all $y \in \mathbb{R}^d_{\ge 0}$ gives

$$f_{\mathrm{HL}}(t,x) \ge f_{\mathrm{HL}}(s,x+z) - \frac{\|z\|_G^2}{2(t-s)},$$

and taking the supremum over all $z \in \mathbb{R}^d_{\ge 0}$ establishes the lower bound

$$f_{\mathrm{HL}}(t,x) \ge \sup_{y\in\mathbb{R}^d_{\ge 0}} \left\{ f_{\mathrm{HL}}(s,x+y) - \frac{\|y\|_G^2}{2(t-s)} \right\}.$$

To obtain the matching upper bound, we invoke Lemma 4.4 to find $y \in \mathbb{R}^d_{\ge 0}$ such that

$$f_{\mathrm{HL}}(t,x) = \psi(x+y) - \frac{\|y\|_G^2}{2t}.$$

Defining $z = \frac{t-s}{t} y \in \mathbb{R}^d_{\ge 0}$, we observe that

$$\frac{z}{t-s} = \frac{y}{t} = \frac{y-z}{s}. \tag{4.12}$$

In particular, testing the supremum in (4.4) with $y - z = \frac{s}{t} y \in \mathbb{R}^d_{\ge 0}$ gives

$$f_{\mathrm{HL}}(s,x+z) \ge \psi(x+z+y-z) - \frac{\|y-z\|_G^2}{2s} = \psi(x+y) - \frac{\|y-z\|_G^2}{2s},$$

and thus, using also (4.12), we obtain

$$\begin{aligned} f_{\mathrm{HL}}(s,x+z) - \frac{\|z\|_G^2}{2(t-s)} &\geq \psi(x+y) - \frac{\|y-z\|_G^2}{2s} - \frac{\|z\|_G^2}{2(t-s)} \\ &= \psi(x+y) - \frac{\|y\|_G^2}{2t}. \end{aligned}$$

Taking the supremum over $z \in \mathbb{R}^d_{\geq 0}$ and then over $y \in \mathbb{R}^d_{\geq 0}$ completes the proof. ■

We next prove some regularity properties of the function f_{HL} in (4.4).

Lemma 4.6. *Fix an initial condition $\psi : \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ satisfying (2.6). If $G \in \mathbb{R}^{d\times d}$ is non-negative definite, then $f_{\mathrm{HL}} \in \mathfrak{L}_{\mathrm{unif}}$ with*

$$\sup_{t>0} |||f_{\mathrm{HL}}(t,\cdot)|||_{\mathrm{Lip},1} \leq |||\psi|||_{\mathrm{Lip},1} \quad \textit{and} \quad [f_{\mathrm{HL}}]_0 \leq \frac{|||\psi|||_{\mathrm{Lip},1}^2}{2m}. \tag{4.13}$$

Proof. Fix $(t,x,x') \in (0,\infty) \times \mathbb{R}^d_{\geq 0} \times \mathbb{R}^d_{\geq 0}$ and invoke Lemma 4.4 to find $y \in \mathbb{R}^d_{\geq 0}$ with

$$f_{\mathrm{HL}}(t,x) = \psi(x+y) - \frac{\|y\|_G^2}{2t}.$$

Taking this $y \in \mathbb{R}^d_{\geq 0}$ in (4.4) gives the lower bound

$$f_{\mathrm{HL}}(t,x') \geq \psi(x'+y) - \frac{\|y\|_G^2}{2t},$$

and thus

$$f_{\mathrm{HL}}(t,x) - f_{\mathrm{HL}}(t,x') \leq \psi(x+y) - \psi(x'+y) \leq |||\psi|||_{\mathrm{Lip},1} |||x-x'|||_1.$$

Reversing the roles of x and x' gives $y' \in \mathbb{R}^d_{\geq 0}$ with

$$f_{\mathrm{HL}}(t,x') - f_{\mathrm{HL}}(t,x) \leq \psi(x'+y') - \psi(x+y') \leq |||\psi|||_{\mathrm{Lip},1} |||x-x'|||_1.$$

Combining these two bounds yields the first inequality in (4.13). To establish Lipschitz continuity in time, fix $x \in \mathbb{R}^d_{\geq 0}$ as well as $t > s \geq 0$. The semigroup property in Lemma 4.5 with $y = 0$ implies that

$$f_{\mathrm{HL}}(t,x) \geq f_{\mathrm{HL}}(s,x). \tag{4.14}$$

Using Lemma 4.5 in combination with the first inequality in (4.13) and (4.6) gives

$$f_{\mathrm{HL}}(t,x) \leq f_{\mathrm{HL}}(s,x) + \sup_{y\in\mathbb{R}^d_{\geq 0}} \left\{ |||\psi|||_{\mathrm{Lip},1} |||y|||_1 - \frac{m|||y|||_1^2}{2(t-s)} \right\} \leq f_{\mathrm{HL}}(s,x) + \frac{|||\psi|||_{\mathrm{Lip},1}^2}{2m}(t-s),$$

where we have used the fact that $r \mapsto r - \frac{1}{2}ar^2$ achieves its maximum at $r = 1/a$. Combining this with (4.14) completes the proof. ■

We are finally in a position to prove Proposition 4.1.

Proof of Proposition 4.1. Denote by $\mathscr{E}_d$ an error term that will be defined in the course of the proof. We will proceed in three steps; first, we will show that the function $f_+(t,x) = f_{\mathrm{HL}}(t,x) + \mathscr{E}_d t$ is a viscosity supersolution to the Hamilton-Jacobi equation (2.5), then we will show that the function $f_-(t,x) = f_{\mathrm{HL}}(t,x) - \mathscr{E}_d t$ is a viscosity subsolution to the Hamilton-Jacobi equation (2.5), and finally we will conclude using the comparison principle in Corollary A.12.

Step 1: f_+ viscosity supersolution.

Consider a smooth function $\phi \in C^\infty\big((0,\infty)\times\mathbb{R}^d_{>0}\big)$ with the property that $f_+ - \phi$ has a local minimum at $(t^*,x^*) \in (0,\infty)\times\mathbb{R}^d_{>0}$. Using Proposition B.2, it is readily verified that f_{HL} has its gradient in $\mathscr{K}_a'$ as ψ does. It follows by Lemma 3.2 that $\nabla\phi(t^*,x^*) \in \mathscr{K}_a'$. It is therefore possible to find $u \in \mathbb{R}^d_{\geq 0}$ and $w \in \mathbb{R}^d$ with

$$\nabla\phi(t^*,x^*) = Gu + w, \quad |||u|||_1 \leq a \quad \text{and} \quad |||w|||_{1,*} \leq \frac{1}{\sqrt{d}}.$$

On the one hand, if $s > 0$ is sufficiently small that $t^* - s > 0$, then

$$f_+(t^*-s, x^*+su) - \phi(t^*-s,x^*+su) \geq f_+(t^*,x^*) - \phi(t^*,x^*).$$

On the other hand, taking $su \in \mathbb{R}^d_{\geq 0}$ in Lemma 4.5 reveals that

$$f_{\mathrm{HL}}(t^*,x^*) \geq f_{\mathrm{HL}}(t^*-s,x^*+su) - s\frac{\|u\|_G^2}{2}.$$

It follows that

$$\phi(t^*,x^*) - \phi(t^*-s,x^*+su) + s\frac{\|u\|_G^2}{2} - \mathscr{E}_d s \geq 0.$$

Dividing by $0 < s < t^*$ and letting $s \to 0$ yields

$$\partial_t\phi(t^*,x^*) - u\cdot\nabla\phi(t^*,x^*) + \frac{\|u\|_G^2}{2} - \mathscr{E}_d \geq 0.$$

Recalling that $\nabla\phi(t^*,x^*) = Gu + w$ and using Lemma 4.2, we obtain

$$\partial_t\phi(t^*,x^*) - \mathsf{C}(Gu) - u\cdot w - \mathscr{E}_d \geq 0.$$

By Lemma 3.1 and Lemma 4.6, we have

$$|||Gu|||_{1,*} \leq |||\nabla\phi(t^*,x^*)|||_{1,*} + \frac{1}{\sqrt{d}} \leq |||\psi|||_{\mathrm{Lip},1} + \frac{1}{\sqrt{d}} \leq R,$$

so the Lipschitz continuity of the non-linearity H_R established in Proposition 2.3 implies that

$$\big(\partial_t\phi - \mathsf{H}_R(\nabla\phi)\big)(t^*,x^*) \geq \mathscr{E}_d - |||u|||_1|||w|||_{1,*} - \frac{8RM}{m^2\sqrt{d}} \geq \mathscr{E}_d - \frac{a}{\sqrt{d}} - \frac{8RM}{m^2\sqrt{d}}.$$

This shows that f_+ is a supersolution to the Hamilton-Jacobi equation (2.5) provided that

$$\mathscr{E}_d \geq \frac{a}{\sqrt{d}} + \frac{8RM}{m^2\sqrt{d}}.$$

Step 2: f_- viscosity subsolution.
Consider a smooth function $\phi \in C^\infty\big((0,\infty)\times\mathbb{R}^d_{>0}\big)$ with the property that $f_- - \phi$ has a local maximum at $(t^*,x^*) \in (0,\infty)\times\mathbb{R}^d_{>0}$. Since f_{HL} has its gradient in $\mathscr{K}'_a$ as ψ does, we have $\nabla\phi(t^*,x^*) \in \mathscr{K}'_a$ by Lemma 3.2. It is therefore possible to find $u \in \mathbb{R}^d_{\geq 0}$ and $w \in \mathbb{R}^d$ with

$$\nabla\phi(t^*,x^*) = Gu + w, \quad |||u|||_1 \leq a \quad \text{and} \quad |||w|||_{1,*} \leq \frac{1}{\sqrt{d}}.$$

Suppose for the sake of contradiction that there exists $\delta > 0$ with

$$\big(\partial_t\phi - \mathsf{H}_R(\nabla\phi)\big)(t^*,x^*) \geq \delta > 0.$$

Arguing as in the previous step, this implies that

$$\partial_t\phi(t^*,x^*) - \mathsf{C}(Gu) \geq \delta - \frac{8RM}{m^2\sqrt{d}}.$$

By Lemma 4.2, this may be recast as the assumption that

$$\partial_t\phi(t^*,x^*) - y\cdot Gu + \frac{\|y\|_G^2}{2} \geq \delta - \frac{8RM}{m^2\sqrt{d}}$$

for all $y \in \mathbb{R}^d_{\geq 0}$. By continuity, of $\partial_t\phi$ and $\nabla\phi$, up to redefining $\delta > 0$, we may in fact assume that

$$\partial_t\phi(t',x') - y\cdot\nabla\phi(t',x') + \frac{\|y\|_G^2}{2} \geq \delta - \frac{8RM}{m^2\sqrt{d}} - |||y|||_1|||w|||_{1,*} \tag{4.15}$$

for all $y \in \mathbb{R}^d_{\geq 0}$ and (t',x') sufficiently close to (t^*,x^*). Recalling Lemma 4.5 and arguing as in the proof of Lemma 4.4, it is possible to find $R > 0$ such that, for every $s > 0$ sufficiently small, there exists $y_s \in \mathbb{R}^d_{\geq 0}$ with $|||y_s|||_1 \leq Rs$ and

$$f_{\mathrm{HL}}(t^*,x^*) = f_{\mathrm{HL}}(t^*-s, x^*+y_s) - \frac{\|y_s\|_G^2}{2s}.$$

It follows by the fundamental theorem of calculus and the absurd assumption (4.15) used with $y = \frac{y_s}{s} \in \mathbb{R}^d_{\geq 0}$ that

$$\begin{aligned}
\phi(t^*,x^*) - \phi(t^*-s,x^*+y_s) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}r}\phi\big(rt^* + (1-r)(t^*-s), rx^* + (1-r)(x^*+y_s)\big)\,\mathrm{d}r \\
&= \int_0^1 \big(s\partial_t\phi - y_s\cdot\nabla\phi\big)(t^*+(r-1)s, x^*+(1-r)y_s)\,\mathrm{d}r \\
&\geq s\delta - \frac{\|y_s\|_G^2}{2s} - s\frac{8RM}{m^2\sqrt{d}} - |||y_s|||_1|||w|||_{1,*} \\
&\geq f_{\mathrm{HL}}(t^*,x^*) - f_{\mathrm{HL}}(t^*-s,x^*+y_s) + s\left(\delta - \frac{8RM}{m^2\sqrt{d}} - \frac{R}{\sqrt{d}}\right).
\end{aligned}$$

Rearranging shows that for s sufficiently small,

$$f_-(t^*-s,x^*+y_s)-\phi(t^*-s,x^*+y_s)\geq s\left(\delta-\frac{8RM}{m^2\sqrt{d}}-\frac{R}{\sqrt{d}}+\mathscr{E}_d\right)+f_-(t^*,x^*)-\phi(t^*,x^*).$$

This contradicts the fact that $f-\phi$ admits a local maximum at (t^*,x^*) provided that

$$\mathscr{E}_d\geq\frac{R}{\sqrt{d}}+\frac{8RM}{m^2\sqrt{d}}.$$

Step 3: comparison principle.
Combining step 1 and step 2 shows that, if we define

$$\mathscr{E}_d=\frac{1}{\sqrt{d}}\left(R+a+\frac{8RM}{m^2}\right),$$

then f_+ is a viscosity supersolution to the Hamilton-Jacobi equation (2.5) while f_- is a viscosity subsolution to this equation. Together with Lemma 4.3, Lemma 4.6 and the comparison principle in Corollary A.12, this implies that for every $(t,x)\in[0,\infty)\times\mathbb{R}^d_{\geq 0}$,

$$|f_{\mathrm{HL}}(t,x)-f(t,x)|\leq\mathscr{E}_d t$$

as required. ■

5 Hopf-Lax formula for the infinite-dimensional equation

In this section, we apply Proposition 4.1 to the projected Hamilton-Jacobi equation (1.37) and let K tend to infinity in the resulting variational formula to establish Theorem 1.3. In addition to the assumptions (**H1**)-(**H4**), we will suppose that the kernel $g:[-1,1]\to\mathbb{R}$ is non-negative definite in the sense that it satisfies (**H5**). This assumption is equivalent to the non-negative definiteness of each of the matrices (1.18), and therefore to the convexity of each of the projected non-linearities (1.21). In particular, Proposition 4.1 implies that the unique solution $f^{(K)}:[0,\infty)\times\mathbb{R}^{\mathscr{D}_K}\to\mathbb{R}$ to the projected Hamilton-Jacobi equation (1.37) in $\mathfrak{L}_{\mathrm{unif}}$ subject to the initial condition $\psi^{(K)}$ satisfies

$$f^{(K)}\big(t,x^{(K)}(\mu)\big)=\sup_{y\in\mathbb{R}^{\mathscr{D}_K}_{\geq 0}}\left\{\psi^{(K)}\big(x^{(K)}(\mu)+y\big)-\frac{\|y\|_G^2}{2t}\right\}+\mathscr{O}\left(t|\mathscr{D}_K|^{-1/2}\right). \tag{5.1}$$

Remembering the definition of the projected initial condition (1.17), the projected non-linearity (1.21) and its relationship (1.24) to the non-linearity C_∞ in (1.5) shows that

$$f^{(K)}\big(t,x^{(K)}(\mu)\big)=\sup_{\nu\in\mathscr{M}_+^{(K)}}\left\{\psi(\mu+\nu)-\frac{\mathsf{C}_\infty(G_\nu)}{t}\right\}+\mathscr{O}\left(t|\mathscr{D}_K|^{-1/2}\right), \tag{5.2}$$

where we have made the substitution $y=x^{(K)}(\nu)$ for $\nu\in\mathscr{M}_+^{(K)}$. Using Theorem 1.2 and a simple continuity argument to let K tend to infinity in this expression gives the variational representation formula

$$f(t,\mu) = \sup_{\nu\in\mathscr{M}_+}\left\{\psi(\mu+\nu) - \frac{1}{2t}\int_{-1}^{1} G_\nu(y)\,\mathrm{d}\nu(y)\right\} \tag{5.3}$$

$$= \sup_{\nu\in\mathscr{M}_+}\left\{\psi(\mu+t\nu) - \frac{t}{2}\int_{-1}^{1} G_\nu(y)\,\mathrm{d}\nu(y)\right\}. \tag{5.4}$$

The second of these expressions follows from the first by setting $\nu' = t\nu$. To establish Theorem 1.3, we need to show that the supremum in (5.4) is achieved at some $\nu^* \in \mathscr{M}_+$, and that whenever the initial condition admits a Gateaux derivative at the measure $\mu + t\nu^*$ with density $x \mapsto D_\mu\psi(\mu + t\nu^*, x)$ in $\mathscr{C}_\infty$,

$$G_{\nu^*} = D_\mu(\mu+t\nu^*,\cdot). \tag{5.5}$$

If we ignore the constraint that the optimizers in (5.4) must be non-negative measures, then this latter property is clear from the first order conditions on a maximizer. To prove this rigorously, we first show that a maximizer exists, and we then establish a Cauchy-Schwarz inequality for the non-negative definite kernel $\tilde{g}(x,y) = g(xy)$.

Lemma 5.1. *For every $t \geq 0$ and $\mu \in \mathscr{M}_+$, there exists $\nu^* \in \mathscr{M}_+$ with*

$$f(t,\mu) = \psi(\mu+t\nu^*) - \frac{t}{2}\int_{-1}^{1} G_{\nu^*}(y)\,\mathrm{d}\nu^*(y). \tag{5.6}$$

Proof. Fix a probability measure $\nu \in \Pr[-1,1]$ and a positive constant $\lambda > 0$. The Lipschitz continuity (**H2**) of the initial condition implies that

$$\psi(\mu+\lambda t\nu) \leq \psi(\mu) + \|\psi\|_{\mathrm{Lip,TV}}\mathrm{TV}(0,\lambda t\nu) \leq \psi(\mu) + 2\lambda t\|\psi\|_{\mathrm{Lip,TV}}.$$

On the other hand,

$$\int_{-1}^{1} G_{\lambda\nu}(y)\,\mathrm{d}(\lambda\nu)(y) = \lambda^2\int_{-1}^{1}\int_{-1}^{1} g(xy)\,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y) \geq \lambda^2 m.$$

Combining these two bounds reveals that

$$\psi(\mu+t\nu) - \frac{t}{2}\int_{-1}^{1} G_\nu(y)\,\mathrm{d}\nu(y) \leq \psi(\mu) + 2\lambda t\|\psi\|_{\mathrm{Lip,TV}} - \frac{\lambda^2 tm}{2}.$$

The supremum in (5.4) can therefore be restricted to measures in $\mathscr{M}_+$ with bounded total mass. The existence of a maximizer is now an immediate consequence of Prokhorov's theorem. Indeed, if $(\nu_n) \subset \mathscr{M}_+$ denotes a maximizing sequence, we may assume without loss of generality that each measure in this sequence has total mass bounded by the same constant. It follows by Prokhorov's theorem that this sequence is pre-compact, and therefore admits a subsequential limit with respect to the weak convergence of measures. By continuity of the functional being maximized in (5.4), this weak limit must be a maximizer. This completes the proof. ■

Lemma 5.2. *If g satisfies (**H5**) and $\mu,\nu \in \mathscr{M}_s$ are signed measures, then*

$$\left(\int_{-1}^{1} G_\nu(x)\,\mathrm{d}\mu(x)\right)^2 \leq \left(\int_{-1}^{1} G_\mu(x)\,\mathrm{d}\mu(x)\right)\left(\int_{-1}^{1} G_\nu(x)\,\mathrm{d}\nu(x)\right). \tag{5.7}$$

Proof. This is the Cauchy-Schwarz inequality for the non-negative definite kernel $\tilde{g}(x,y) = g(xy)$, and can be proved in a standard way. Indeed, for every $t \in \mathbb{R}$, let

$$\begin{aligned} P(t) &= \int_{-1}^{1}\int_{-1}^{1} g(xy)\,\mathrm{d}(\mu+t\nu)(x)\,\mathrm{d}(\mu+t\nu)(y) \\ &= \int_{-1}^{1} G_\mu(x)\,\mathrm{d}\mu(x) + 2t\int_{-1}^{1} G_\nu(x)\,\mathrm{d}\mu(x) + t^2\int_{-1}^{1} G_\nu(x)\,\mathrm{d}\nu(x). \end{aligned}$$

This polynomial is non-negative by (**H5**). In particular, its discriminant cannot be positive. This means that

$$2^2\left(\int_{-1}^{1} G_\nu(x)\,\mathrm{d}\mu(x)\right)^2 - 4\left(\int_{-1}^{1} G_\mu(x)\,\mathrm{d}\mu(x)\right)\left(\int_{-1}^{1} G_\nu(x)\,\mathrm{d}\nu(x)\right) \le 0.$$

Rearranging completes the proof. ■

We are finally in a position to prove Theorem 1.3.

Proof of Theorem 1.3. Fix $t > 0$ and $\mu \in \mathscr{M}_+$. Combining Theorem 1.2 with (5.4) shows that the unique solution to the infinite-dimensional Hamilton-Jacobi equation (1.10) admits the Hopf-Lax variational representation (1.44). Moreover, Lemma 5.1 ensures that the supremum in (5.4) is achieved at some $\nu^* \in \mathscr{M}_+$. To establish the final statement in Theorem 1.3, suppose that the initial condition ψ admits a Gateaux derivative at the measure $\mu + t\nu^*$ with density $x \mapsto D_\mu\psi(\mu+t\nu^*,x)$ in $\mathscr{C}_\infty$. For any measure $\eta \in \mathscr{M}_+$, the Gateaux derivative of the functional

$$\nu \mapsto \psi(\mu+t\nu) - \frac{t}{2}\int_{-1}^{1} G_\nu(y)\,\mathrm{d}\nu(y)$$

at the measure ν^* in the direction of $\eta - \nu^*$ is

$$D_\mu\psi\big(\mu+t\nu^*; t(\eta-\nu^*)\big) - t\int_{-1}^{1}\int_{-1}^{1} g(xy)\,\mathrm{d}(\eta-\nu^*)(x)\,\mathrm{d}\nu^*(y). \tag{5.8}$$

Moreover, for every $\varepsilon \in [0,1]$, we have that $\nu^* + \varepsilon(\eta - \nu^*)$ belongs to $\mathscr{M}_+$, and is thus a valid candidate for the optimization problem in (5.4). As a consequence, the quantity in (5.8) must be non-positive. Using also the definition of the Gateaux derivative density in (1.9), we get that

$$t\int_{-1}^{1}\Big(D_\mu\psi(\mu+t\nu^*,x) - \int_{-1}^{1} g(xy)\,\mathrm{d}\nu^*(y)\Big)\,\mathrm{d}\big(\eta-\nu^*\big)(x) \le 0, \tag{5.9}$$

for every $\eta \in \mathscr{M}_+$. The assumption that the density $x \mapsto D_\mu\psi(\mu+t\nu^*,x)$ belongs to the cone $\mathscr{C}_\infty$ gives a measure $\eta^* \in \mathscr{M}_+$ with $G_{\eta^*}(x) = D_\mu\psi(\mu+t\nu^*,x)$. Applying (5.9) to the measure $\eta = \eta^*$ reveals that

$$\int_{-1}^{1}\int_{-1}^{1} g(xy)\,\mathrm{d}\big(\eta^*-\nu^*\big)(y)\,\mathrm{d}\big(\eta^*-\nu^*\big)(x) \le 0.$$

Together with (1.43), this implies that

$$\int_{-1}^{1}\int_{-1}^{1} g(xy)\,\mathrm{d}\big(\eta^*-\nu^*\big)(y)\,\mathrm{d}\big(\eta^*-\nu^*\big)(x) = 0.$$

Applying Lemma 5.2 to the signed measures $\eta^*-\nu^*$ and δ_x for some $x\in[-1,1]$ shows that

$$\int_{-1}^{1} g(xy)\,\mathrm{d}\big(\eta^*-\nu^*\big)(y)=\int_{-1}^{1} g(yz)\,\mathrm{d}\big(\eta^*-\nu^*\big)(y)\,\mathrm{d}\delta_x(z)=0.$$

Rearranging gives $G_{\nu^*}(x)=G_{\eta^*}(x)=D_\mu(\mu+t\nu^*,x)$. Since $x\in[-1,1]$ is arbitrary, this completes the proof. ■

6 The infinite-dimensional equation with arbitrary kernel

In this section, we extend the main results of this paper to the Hamilton-Jacobi equation (1.10) associated with a kernel g that is not necessarily assumed to be positive. Fix $b\in\mathbb{R}$, and recall the definition of the modified kernel $\widetilde{g}_b$ in (1.51), of the modified Hamilton-Jacobi equation (1.55) and of the modified solution f_b in (1.57). Introduce the symmetric matrix

$$\widetilde{G}_b^{(K)}=\frac{1}{|\mathscr{D}_K|^2}\big(\widetilde{g}_b(kk')\big)_{k,k'\in\mathscr{D}_K}\in\mathbb{R}^{\mathscr{D}_K\times\mathscr{D}_K} \tag{6.1}$$

the projected cone

$$\widetilde{\mathscr{C}}_{b,K}=\Big\{\widetilde{G}_b^{(K)}x^{(K)}(\mu)\in\mathbb{R}^{\mathscr{D}_K}\mid \mu\in\mathscr{M}_+^{(K)}\Big\}=\Big\{\widetilde{G}_b^{(K)}x\in\mathbb{R}^{\mathscr{D}_K}\mid x\in\mathbb{R}^{\mathscr{D}_K}_{\geq 0}\Big\} \tag{6.2}$$

and the projected non-linearity $\widetilde{\mathsf{C}}_{b,K}:\widetilde{\mathscr{C}}_{b,K}\to\mathbb{R}$ defined by

$$\widetilde{\mathsf{C}}_{b,K}\big(\widetilde{G}_b^{(K)}x\big)=\frac{1}{2}\widetilde{G}_b^{(K)}x\cdot x=\frac{1}{2|\mathscr{D}_K|^2}\sum_{k,k'\in\mathscr{D}_K}\widetilde{g}_b(kk')x_kx_{k'}. \tag{6.3}$$

Also define the closed convex set

$$\widetilde{\mathscr{K}}_{=a,b,K}=\Big\{\widetilde{G}_b^{(K)}x\in\mathbb{R}^{\mathscr{D}_K}\mid x\in\mathbb{R}^{\mathscr{D}_K}_{\geq 0}\text{ and }|||x|||_1=a\Big\} \tag{6.4}$$

The first order of business will be to verify that the function f_b is well-defined by ensuring that (**H1**)-(**H4**) are satisfied in the context of the infinite-dimensional Hamilton-Jacobi equation (1.55).

Lemma 6.1. *Under the assumptions of Theorem 1.4, the kernel $\widetilde{g}_b$ in* (1.51) *and the initial condition $\widetilde{\psi}_b$ in* (1.56) *satisfy* (**H1**)-(**H4**). *Moreover, each projected initial condition $\widetilde{\psi}_b^{(K)}$ has its gradient in $\widetilde{\mathscr{K}}_{=a,b,K}$.*

Proof. The kernel $\widetilde{g}_b$ satisfies (**H1**) by the choice of b, while the initial condition $\widetilde{\psi}_b$ satisfies (**H2**) by the triangle inequality and the bound

$$\Big|ab\int_{-1}^{1}\mathrm{d}\mu-ab\int_{-1}^{1}\mathrm{d}\nu\Big|\leq a|b|\big|\mu[-1,1]-\nu[-1,1]\big|\leq a|b|\mathrm{TV}(\mu,\nu).$$

An identical argument shows that the initial condition $\widetilde{\psi}_b$ satisfies (**H4**). To verify (**H3**), introduce the closed convex set

$$\widetilde{\mathscr{K}}_{a,b,K}=\Big\{\widetilde{G}_b^{(K)}x\in\mathbb{R}^{\mathscr{D}_K}\mid x\in\mathbb{R}^{\mathscr{D}_K}_{\geq 0}\text{ and }|||x|||_1\leq a\Big\},$$

and fix $c \in \mathbb{R}$ and $x, x' \in \mathbb{R}^{\mathscr{D}_K}$ such that $(x'-x)\cdot z \geq c$ for every $z \in \widetilde{\mathscr{K}'_{a,b,K}}$. Now fix $y \in \mathscr{K}'_{=a,K}$, and represent it as $y = G^{(K)}u + w$ for some $u \in \mathbb{R}^d_{\geq 0}$ and $w \in \mathbb{R}^d$ with $|||u|||_1 = a$ and $|||w|||_{1,*} \leq 2^{-K/2}$. If $z = \widetilde{G}_b^{(K)}u + w$, then

$$z = G^{(K)}u + b|||u|||_1 \iota_K + w = y + ab\iota_K$$

for the vector $\iota_K = (|\mathscr{D}_K|^{-1})_{k \in \mathscr{D}_K} \in \mathbb{R}^{\mathscr{D}_K}_{\geq 0}$. Since $z \in \widetilde{\mathscr{K}_{=a,b,K}} \subset \widetilde{\mathscr{K}'_{a,b,K}}$,

$$(x'-x)\cdot y = (x'-x)\cdot z - (x'-x)\cdot ab\iota_K \geq c - (x'-x)\cdot ab\iota_K.$$

The assumption (**H3'**) and Proposition B.2 therefore imply that

$$\psi^{(K)}(x') - \psi^{(K)}(x) \geq c - (x'-x)\cdot ab\iota_K.$$

Noticing that $x \cdot \iota_K = |||x|||_1$ and rearranging reveals that

$$\widetilde{\psi}_b^{(K)}(x') - \widetilde{\psi}_b^{(K)}(x) \geq c.$$

Together with Proposition B.2, this establishes (**H3**). Notice that this argument only needed the assumption that $(x'-x)\cdot z \geq c$ for every $z \in \widetilde{\mathscr{K}_{=a,b,K}}$ so it also shows that $\widetilde{\psi}_b^{(K)}$ has its gradient in $\widetilde{\mathscr{K}_{=a,b,K}}$ by Proposition B.2. This completes the proof. ■

Together with Proposition 2.7 this result implies that for every $R > 0$, the Hamilton-Jacobi equation

$$\partial_t \widetilde{f}^{(K)}(t,x) = \widetilde{\mathsf{H}}_{b,K,R}\big(\nabla \widetilde{f}^{(K)}(t,x)\big) \quad \text{on} \quad \mathbb{R}_{>0} \times \mathbb{R}^{\mathscr{D}_K}_{>0} \tag{6.5}$$

admits a unique viscosity solution $\widetilde{f}_{b,R}^{(K)} \in \mathscr{L}_{\mathrm{unif}}$ subject to the initial condition $\widetilde{\psi}_b^{(K)}$ which satisfies the Lipschitz bound

$$\sup_{t \geq 0} |||\widetilde{f}_{b,R}^{(K)}(t,\cdot)|||_{\mathrm{Lip},1} = |||\widetilde{\psi}_b^{(K)}|||_{\mathrm{Lip},1}. \tag{6.6}$$

Here $\widetilde{\mathsf{H}}_{b,K,R} : \mathbb{R}^{\mathscr{D}_K} \to \mathbb{R}$ denotes the extension of the non-linearity $\widetilde{\mathsf{C}}_{b,K}$ provided by Proposition 2.3. Since the projected initial condition $\widetilde{\psi}_b^{(K)}$ has its gradient in the closed convex set $\mathscr{K}'_{=a,b,K}$ by Lemma 6.1, an identical argument to that in Proposition 2.7 shows that the solution $\widetilde{f}_{b,R}^{(K)}$ also has its gradient in $\mathscr{K}'_{=a,b,K}$. Moreover, Theorem 1.2 allows us to define the solution to the infinite-dimensional Hamilton-Jacobi equation (1.55) by

$$\widetilde{f}_b(t,\mu) = \lim_{K \to \infty} \widetilde{f}_{b,R}^{(K)}\big(t, x^{(K)}(\mu)\big), \tag{6.7}$$

and guarantees that this limit is independent of $R > 0$ provided that $R > \|\widetilde{\psi}_b\|_{\mathrm{Lip,TV}}$. Using the comparison principle in Corollary A.12 we now show that the limit defining the function (1.57),

$$f_b(t,\mu) = \lim_{K \to \infty} \left(\widetilde{f}_{b,R}^{(K)}\big(t, x^{(K)}(\mu)\big) - ab|||x^{(K)}(\mu)|||_1 - \frac{a^2 bt}{2} \right), \tag{6.8}$$

is independent of b by establishing Theorem 1.4.

Proof of Theorem 1.4. Let $b, b' \in \mathbb{R}$ be such that the kernels $\widetilde{g}_b$ and $\widetilde{g}_{b'}$ are positive on $[-1,1]$, and fix $R > \|\widetilde{\psi}_b\|_{\mathrm{Lip,TV}} + \|\widetilde{\psi}_{b'}\|_{\mathrm{Lip,TV}}$. The idea will be to show that the function

$$f_{b,b'}^{(K)}(t,x) = \widetilde{f}_b^{(K)}(t,x) - a(b-b')|||x|||_1 - \frac{a^2(b-b')t}{2}$$

satisfies the Hamilton-Jacobi equation defining $\widetilde{f}_{b'}^{(K)}$ up to an error vanishing with K. We have omitted the dependence on R, and will continue to do so throughout this proof, as this constant will remain fixed. The equality of f_b and $f_{b'}$ will then follow from the comparison principle in Corollary A.12. Consider $\phi \in C^\infty\big((0,\infty) \times \mathbb{R}_{>0}^{\mathscr{D}_K}\big)$ with the property that $f_{b,b'}^{(K)} - \phi$ achieves a local maximum at the point $(t^*, x^*) \in (0,\infty) \times \mathbb{R}_{>0}^{\mathscr{D}_K}$. Since $\widetilde{f}_b^{(K)}$ is a viscosity subsolution to the Hamilton-Jacobi equation (6.5),

$$\frac{a^2(b-b')}{2} + \partial_t \phi(t^*, x^*) - \widetilde{\mathsf{H}}_{b,K}\big(a(b-b')\iota_K + \nabla\phi(t^*,x^*)\big) \leq 0$$

for the vector $\iota_K = (|\mathscr{D}_K|^{-1})_{k \in \mathscr{D}_K} \in \mathbb{R}_{\geq 0}^{\mathscr{D}_K}$. The fact that $\widetilde{f}_b^{(K)}$ has its gradient in $\widetilde{\mathscr{K}'}_{=a,b,K}$ together with (3.1), Lemma 3.1 and Lemma 3.2 implies that

$$a(b-b')\iota_K + \nabla\phi(t^*,x^*) \in \widetilde{\mathscr{K}'}_{=a,b,K} \quad \text{and} \quad |||a(b-b')\iota_K + \nabla\phi(t^*,x^*)|||_{1,*} \leq \|\widetilde{\psi}_b\|_{\mathrm{Lip,TV}}.$$

It is therefore possible to find $u \in \mathbb{R}_{\geq 0}^{\mathscr{D}_K}$ and $w \in \mathbb{R}^{\mathscr{D}_K}$ with

$$a(b-b')\iota_K + \nabla\phi(t^*,x^*) = \widetilde{G}_b^{(K)} u + w, \quad |||u|||_1 = a \quad \text{and} \quad |||w|||_{1,*} \leq \frac{1}{2^{K/2}}.$$

Observe that

$$|||\widetilde{G}_b^{(K)} u|||_{1,*} \leq |||a(b-b')\iota_K + \nabla\phi(t^*,x^*)|||_{1,*} + |||w|||_{1,*} \leq \|\widetilde{\psi}_b\|_{\mathrm{Lip,TV}} + \frac{1}{2^{K/2}},$$

so increasing K if necessary, it is possible to ensure that $\widetilde{G}_b^{(K)} u \in \widetilde{\mathscr{C}}_{b,K} \cap B_R$. It follows by the Lipschitz continuity of $\widetilde{\mathsf{H}}_{b,K}$ established in Proposition 2.3 that

$$\frac{a^2(b-b')}{2} + \partial_t\phi(t^*,x^*) - \widetilde{\mathsf{C}}_{b,K}\big(\widetilde{G}_b^{(K)} u\big) \leq \frac{8RM}{2^{K/2} m^2}.$$

Observe that

$$\begin{aligned}
\widetilde{\mathsf{C}}_{b,K}\big(\widetilde{G}_b^{(K)} u\big) &= \frac{1}{2}\widetilde{G}_b^{(K)} u \cdot u = \frac{1}{2} G^{(K)} u \cdot u + \frac{1}{2} b |||u|||_1^2 = \frac{1}{2}\widetilde{G}_{b'}^{(K)} u \cdot u + \frac{1}{2}(b-b')a^2 \\
&= \widetilde{\mathsf{C}}_{b',K}\big(\widetilde{G}_{b'}^{(K)} u\big) + \frac{a^2(b-b')}{2}
\end{aligned}$$

so in fact

$$\partial_t\phi(t^*,x^*) - \widetilde{\mathsf{C}}_{b'}^{(K)}\big(\widetilde{G}_{b'}^{(K)} u\big) \leq \frac{8RM}{2^{K/2} m}.$$

To replace $\widetilde{G}_{b'}^{(K)} u$ by $\nabla\phi(t^*,x^*)$ observe that

$$\begin{aligned}
\widetilde{G}_{b'}^{(K)} u &= \widetilde{G}_b^{(K)} u + (b'-b)\iota_K |||u|||_1 = a(b-b')\iota_K + \nabla\phi(t^*,x^*) - w + a(b'-b)\iota_K \\
&= \nabla\phi(t^*,x^*) - w,
\end{aligned}$$

and leverage the Lipschitz continuity of $\widetilde{\mathsf{H}}_{b',K}$ established in Proposition 2.3 to deduce that

$$\big(\partial_t \phi - \widetilde{\mathsf{H}}_{b',K}(\nabla \phi)\big)(t^*,x^*) \leq \mathscr{E}_K$$

for the error term

$$\mathscr{E}_K = \frac{16RM}{2^{K/2}m}.$$

In particular, the function $(t,x) \mapsto f^K_{b,b'}(t,x) - \mathscr{E}_K t$ is a viscosity subsolution to the Hamilton-Jacobi equation defining $\widetilde{f}^K_{b'}$. An identical argument shows that $(t,x) \mapsto f^K_{b,b'}(t,x) + \mathscr{E}_K t$ is a viscosity supersolution to the Hamilton-Jacobi equation defining $\widetilde{f}^K_{b'}$. It follows by the comparison principle in Corollary A.12 that for every $\mu \in \mathscr{M}_+$,

$$\big|f_b^{(K)}\big(t,x^{(K)}(\mu)\big) - f_{b'}^{(K)}\big(t,x^{(K)}(\mu)\big)\big| = \big|f_{b,b'}^{(K)}\big(t,x^{(K)}(\mu)\big) - \widetilde{f}_{b'}^{(K)}\big(t,x^{(K)}(\mu)\big)\big| \leq \mathscr{E}_K t.$$

Letting K tend to infinity completes the proof. ■

Combining this well-posedness result with the Hopf-Lax representation formula in Theorem 1.3 we now prove the Hopf-Lax variational representation for $f = f_b$ stated in Theorem 1.5.

Proof of Theorem 1.5. Fix $b > 0$ large enough so the kernel $\widetilde{g}_b$ is positive on $[-1,1]$ and satisfies (**H5**). Lemma 6.1 and the Hopf-Lax representation formula in Theorem 1.3 imply that for any $t > 0$ and $\mu \in \mathscr{M}_+$,

$$\widetilde{f}_b(t,\mu) = \sup_{\nu \in \mathscr{M}_+} \left\{ \widetilde{\psi}_b(\mu + t\nu) - \frac{t}{2}\int_{-1}^1 \widetilde{G}_{b,\nu}(y)\,\mathrm{d}\nu(y) \right\}. \tag{6.9}$$

Since the Gateaux derivative density $x \mapsto D_\mu \psi(\mu + t\nu)$ belongs to the set $\mathscr{C}_{a,\infty}$ by assumption, there exists a measure $\eta \in \mathscr{M}_{a,+}$ with $D_\mu \psi(\mu + t\nu, \cdot) = G_\eta$. This means that

$$D_\mu \widetilde{\psi}_b(\mu + t\nu, x) = D_\mu \psi(\mu + t\nu, x) + ab = \int_{-1}^1 g(xy)\,\mathrm{d}\eta(y) + ab = \int_{-1}^1 \widetilde{g}_b(xy)\,\mathrm{d}\eta(y),$$

so another application of Theorem 1.3 implies that the supremum in (6.9) is achieved at some $\nu^* \in \mathscr{M}_+$ with

$$\widetilde{G}_{b,\nu^*} = D_\mu \widetilde{\psi}_b(\mu + t\nu^*, \cdot) = \widetilde{G}_{b,\eta}.$$

Evaluating this equality at $x = 0$ reveals that

$$\widetilde{g}_b(0)\int_{-1}^1 \mathrm{d}\nu^*(y) = \widetilde{g}_b(0)\int_{-1}^1 \mathrm{d}\eta(y) = \widetilde{g}_b(0)a.$$

Since $\widetilde{g}_b(0) > 0$ by the choice of b, this means that $\nu^* \in \mathscr{M}_{a,+}$ and

$$\widetilde{f}_b(t,\mu) = \sup_{\nu \in \mathscr{M}_{a,+}} \left\{ \widetilde{\psi}_b(\mu + t\nu) - \frac{t}{2}\int_{-1}^1 \widetilde{G}_{b,\nu}(y)\,\mathrm{d}\nu(y) \right\}.$$

It follows by (1.57) that

$$
\begin{aligned}
f_b(t,\mu) &= \sup_{\nu\in\mathscr{M}_{a,+}} \left\{ \widetilde{\psi}_b(\mu+t\nu) - \frac{t}{2}\int_{-1}^{1} \widetilde{G}_{b,\nu}(y)\,\mathrm{d}\nu(y) - ab\int_{-1}^{1}\mathrm{d}\mu - \frac{a^2bt}{2} \right\} \\
&= \sup_{\nu\in\mathscr{M}_{a,+}} \left\{ \psi(\mu+t\nu) + abt\int_{-1}^{1}\mathrm{d}\nu - \frac{t}{2}\int_{-1}^{1} G_{\nu}(y)\,\mathrm{d}\nu(y) - \frac{bt}{2}\int_{-1}^{1}\int_{-1}^{1}\mathrm{d}\nu\,\mathrm{d}\nu - \frac{a^2bt}{2} \right\} \\
&= \sup_{\nu\in\mathscr{M}_{a,+}} \left\{ \psi(\mu+t\nu) - \frac{t}{2}\int_{-1}^{1} G_{\nu}(y)\,\mathrm{d}\nu(y) \right\}.
\end{aligned}
$$

This completes the proof. ■

A Hamilton-Jacobi equations on positive half-spaces

In this appendix we fix a non-linearity $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ and an initial condition $\psi:\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ with the properties that

A1 the non-linearity $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ is Lipschitz continuous with respect to the normalized-$\ell^{1,*}$ norm,

$$|\mathsf{H}(y)-\mathsf{H}(y')| \leq |||\mathsf{H}|||_{\mathrm{Lip},1,*}|||y-y'|||_{1,*} \tag{A.1}$$

for all $y,y'\in\mathbb{R}^d$;

A2 the non-linearity is non-decreasing,

$$\mathsf{H}(y)\leq\mathsf{H}(y') \tag{A.2}$$

for all $y,y'\in\mathbb{R}^d$ with $y\leq y'$;

A3 the initial condition $\psi:\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ is Lipschitz continuous with respect to the normalized-ℓ^1 norm,

$$|\psi(x)-\psi(x')| \leq |||\psi|||_{\mathrm{Lip},1}|||x-x'|||_1 \tag{A.3}$$

for all $x,x'\in\mathbb{R}^d_{\geq 0}$;

and we establish the well-posedness of the Hamilton-Jacobi equations

$$\partial_t f(t,x) = \mathsf{H}\big(\nabla f(t,x)\big) \quad \text{on} \quad \mathbb{R}_{>0}\times\mathbb{R}^d_{>0} \tag{A.4}$$

and

$$\partial_t f(t,x) = \mathsf{H}\big(\nabla f(t,x)\big) \quad \text{on} \quad \mathbb{R}_{>0}\times\mathbb{R}^d_{\geq 0} \tag{A.5}$$

subject to the initial condition ψ. The appropriate notion of solution for these equations will be that of a viscosity solution as described in Definition 2.6. To establish the well-posedness of these equations we will closely follow [13]. We found it useful to provide full details when applying Perron's method, although our arguments will certainly be seen as classical by experts, and we hope that the reader will also find these details helpful. To be more specific, first, we will use the assumptions (**A1**) and (**A3**) to prove a comparison principle for the Hamilton-Jacobi equation (A.5) which will ensure the uniqueness of solutions to this equation. This will be the content of Section A.1. In Section A.2, we will combine (**A1**)-(**A3**) with the classical Perron method to obtain the existence of a solution to the Hamilton-Jacobi equation (A.5). Finally, in Section A.3, we will leverage (**A2**) to show that solutions to (A.4) and (A.5) coincide. In the last section of this appendix we will show that the unique solution to the Hamilton-Jacobi equations (A.4) and (A.5) preserves the monotonicity of its initial condition. It will be convenient to remember the definition of the function spaces (1.39) and (1.40).

A.1 Comparison principle and Lipschitz continuity of solutions on $\mathbb{R}^d_{\geq 0}$

In this section, we use the arguments in Proposition 3.2 of [29] to obtain a comparison principle on $\mathfrak{L}_{\text{unif}}$ for the Hamilton-Jacobi equation (A.5). Together with the observation that any solution in $\mathfrak{L}$ to the Hamilton-Jacobi equation (A.5) is uniformly Lipschitz continuous with Lipschitz constant bounded by that of its initial condition, this comparison principle will imply the uniqueness of solutions in $\mathfrak{L}$ to the Hamilton-Jacobi equation (A.5). The Lipschitz continuity of the solutions to the Hamilton-Jacobi equation (A.5) will be obtained using the arguments in Proposition 3.4 of [29]. For every $r \in \mathbb{R}$, we denote the positive part of r by $r_+ = \max(r, 0)$.

Proposition A.1. *Fix a non-linearity $\mathsf{H} : \mathbb{R}^d \to \mathbb{R}$ satisfying (**A1**), and let $u, v \in \mathfrak{L}_{\text{unif}}$ be a subsolution and a supersolution to (A.5). Write $L = \max\left(\sup_{t>0} |||u(t,\cdot)|||_{\text{Lip},1}, \sup_{t>0} |||v(t,\cdot)|||_{\text{Lip},1}\right)$ and $V = |||\mathsf{H}|||_{\text{Lip},1,*}$. For every $Q > 2L$ and all $R \in \mathbb{R}$, the map*

$$(t,x) \mapsto u(t,x) - v(t,x) - Q\big(|||x|||_1 + Vt - R\big)_+ \tag{A.6}$$

achieves its supremum on $\{0\} \times \mathbb{R}^d_{\geq 0}$.

Proof. Suppose for the sake of contradiction that there exists $T > 0$ with

$$\sup_{[0,T]\times\mathbb{R}^d_{\geq 0}} \big(u(t,x) - v(t,x) - \varphi(t,x)\big) > \sup_{\mathbb{R}^d_{\geq 0}} \big(u(0,x) - v(0,x) - \varphi(0,x)\big), \tag{A.7}$$

where $\varphi(t,x) = Q\big(|||x|||_1 + Vt - R\big)_+$. The proof proceeds in three steps: first we smoothen and perturb (A.7), then we use a variable doubling argument to obtain a system of inequalities, and finally we contradict this system of inequalities.

Step 1: smoothing and perturbing.

Given $\varepsilon_0 \in (0,1)$ to be determined, let $\theta \in C^\infty(\mathbb{R})$ be an increasing function with

$$(r - \varepsilon_0)_+ \leq \theta(r) \leq r_+$$

for all $r \in \mathbb{R}$. Consider the smoothed normalized-ℓ^1 norm,

$$|||x|||_{1,\varepsilon_0} = \frac{1}{d} \sum_{k=1}^{d} \left(x_k^2 + \varepsilon_0\right)^{\frac{1}{2}},$$

and introduce the function

$$\Phi(t,x) = Q\theta\left(|||x|||_{1,\varepsilon_0} + Vt - R\right)$$

defined on $\mathbb{R}_{\geq 0} \times \mathbb{R}^d_{\geq 0}$. The choice of θ and the bound $(a+b)_+ \leq a_+ + b_+$ imply that

$$\varphi(t,x) \leq \Phi(t,x) + Q\varepsilon_0 \leq \varphi(t,x) + Q\varepsilon_0^{1/2} + Q\varepsilon_0,$$

where we have used that $(a+b)^{\frac{1}{2}} \leq a^{\frac{1}{2}} + b^{\frac{1}{2}}$ for $a, b > 0$. It follows by (A.7) that

$$\sup_{\mathbb{R}^d_{\geq 0}} \big(u(0,x) - v(0,x) - \Phi(0,x)\big) < \sup_{[0,T]\times\mathbb{R}^d_{\geq 0}} \big(u(t,x) - v(t,x) - \Phi(t,x)\big) + Q\varepsilon_0 + Q\varepsilon_0^{1/2},$$

so choosing $\varepsilon_0 > 0$ small enough guarantees that

$$\sup_{[0,T]\times\mathbb{R}^d_{\geq 0}} \big(u(t,x) - v(t,x) - \Phi(t,x)\big) > \sup_{\mathbb{R}^d_{\geq 0}} \big(u(0,x) - v(0,x) - \Phi(0,x)\big). \tag{A.8}$$

This is a smoothed version of the absurd hypothesis (A.7). For technical reasons, it will be convenient to perturb the function Φ. Given $\varepsilon > 0$ to be determined, introduce the function

$$\chi(t,x) = \Phi(t,x) + \frac{\varepsilon}{T-t},$$

defined on $\mathbb{R}_{\geq 0} \times \mathbb{R}^d_{\geq 0}$. Choosing $\varepsilon > 0$ small enough ensures that

$$\sup_{[0,T]\times\mathbb{R}^d_{\geq 0}} \big(u(t,x) - v(t,x) - \chi(t,x)\big) > \sup_{\mathbb{R}^d_{\geq 0}} \big(u(0,x) - v(0,x) - \chi(0,x)\big). \tag{A.9}$$

This is a smoothed and perturbed version of the absurd hypothesis (A.7).

Step 2: system of inequalities.

For each $\alpha \geq 1$, define the function $\Psi_\alpha : [0,T] \times \mathbb{R}^d_{\geq 0} \times [0,T] \times \mathbb{R}^d_{\geq 0} \to \mathbb{R} \cup \{-\infty\}$ by

$$\Psi_\alpha(t,x,t',x') = u(t,x) - v(t',x') - \frac{\alpha}{2}\big(|t-t'|^2 + |||x-x'|||_{1,\varepsilon_0}\big) - \chi(t,x). \tag{A.10}$$

By doubling the variables and introducing the potential in this way, we ensure that whenever $\alpha > 4(L+1)$, the function Ψ_α achieves its supremum at a point $(t_\alpha, x_\alpha, t'_\alpha, x'_\alpha)$ which remains bounded as α tends infinity. Indeed, if we write $C > 0$ for a constant that depends on T, Q, R, V, $u(0,0)$, $[u]_0$, $v(0,0)$ and $[v]_0$ whose value might not be the same at each occurrence, then for any $x \in \mathbb{R}^d_{\geq 0}$ with $|||x|||_{1,\varepsilon_0} > R+1$, the bound $\Phi(t,x) \geq Q(|||x|||_{1,\varepsilon_0} + Vt - R - 1)_+$ reveals that

$$\begin{aligned}
\Psi_\alpha(t,x,t',x') &\leq u(0,x) - v(0,x') - \frac{\alpha}{2}|||x-x'|||_{1,\varepsilon_0} - \Phi(t,x) + C \\
&\leq L\big(|||x|||_1 + |||x'|||_1\big) - \frac{\alpha}{2}|||x-x'|||_{1,\varepsilon_0} - Q|||x|||_{1,\varepsilon_0} + C \\
&\leq (2L-Q)|||x|||_1 + \Big(L - \frac{\alpha}{2}\Big)|||x-x'|||_{1,\varepsilon_0} + C \\
&\leq (2L-Q)|||x|||_1 + C.
\end{aligned}$$

Remembering that $Q > 2L$ and observing that the supremum of (A.10) is bounded from below by $\Psi_\alpha(0,0,0,0)$ ensures that this upper semi-continuous function achieves its supremum at some point $(t_\alpha, x_\alpha, t'_\alpha, x'_\alpha)$ which remains bounded with respect to the normalized-ℓ^1 norm as α tends to infinity. In particular, the potential

$$\alpha\big(|t_\alpha - t'_\alpha|^2 + |||x_\alpha - x'_\alpha|||_1\big) \leq \alpha\big(|t_\alpha - t'_\alpha|^2 + |||x_\alpha - x'_\alpha|||_{1,\varepsilon_0}\big)$$

must remain bounded as α tends to infinity. It follows that, up to the extraction of a subsequence, there exist $t_0 \in [0,T]$ and $x_0 \in \mathbb{R}^d_{\geq 0}$ such that $t_\alpha \to t_0$, $t'_\alpha \to t_0$, $x_\alpha \to x_0$ and $x'_\alpha \to x_0$ as $\alpha \to \infty$. The term $\frac{\varepsilon}{T-t}$ in the definition of χ guarantees that $t_0 \in [0,T)$. On the other hand, the semi-continuity of u, v and χ together with the bounds

$$\big(u - v - \chi\big)(t,x) \leq \Psi_\alpha(t_\alpha, x_\alpha, t'_\alpha, x'_\alpha) \leq u(t_\alpha, x_\alpha) - v(t'_\alpha, x'_\alpha) - \chi(t_\alpha, x_\alpha)$$

imply that

$$(u - v - \chi)(t_0, x_0) = \sup_{[0,T]\times\mathbb{R}^d_{\geq 0}} (u - v - \chi).$$

By (A.9) it must be the case that $t_0 \in (0,T)$. This means that $t_\alpha, t'_\alpha \in (0,T)$ for all α large enough. We have therefore found a sequence of quadruples $((t_\alpha, x_\alpha, t'_\alpha, x'_\alpha))_\alpha$ with $t_\alpha, t'_\alpha \in (0,T)$ such that Ψ_α achieves its supremum at $(t_\alpha, x_\alpha, t'_\alpha, x'_\alpha)$ for α large enough. With this in mind, fix $\alpha \geq 1$ large enough, and introduce the smooth functions $\phi, \phi' \in C^\infty\big((0,T) \times \mathbb{R}^d_{\geq 0}\big)$ defined by

$$\begin{aligned}
\phi(t,x) &= v(t'_\alpha, x'_\alpha) + \frac{\alpha}{2}\big(|t - t'_\alpha|^2 + |||x - x'_\alpha|||_{1,\varepsilon_0}\big) + \chi(t,x), \\
\phi'(t',x') &= u(t_\alpha, x_\alpha) - \frac{\alpha}{2}\big(|t' - t_\alpha|^2 + |||x' - x_\alpha|||_{1,\varepsilon_0}\big) - \chi(t_\alpha, x_\alpha).
\end{aligned}$$

Since $(t_\alpha, x_\alpha, t'_\alpha, x'_\alpha)$ maximizes Ψ_α, the function $u - \phi$ achieves a local maximum at the point $(t_\alpha, x_\alpha) \in (0,\infty) \times \mathbb{R}^d_{\geq 0}$ while $v - \phi'$ achieves a local minimum at $(t'_\alpha, x'_\alpha) \in (0,\infty) \times \mathbb{R}^d_{\geq 0}$. It follows by definition of a viscosity solution that

$$\big(\partial_t \phi - \mathsf{H}(\nabla \phi)\big)(t_\alpha, x_\alpha) \leq 0 \quad \text{and} \quad \big(\partial_t \phi' - \mathsf{H}(\nabla \phi')\big)(t'_\alpha, x'_\alpha) \geq 0. \tag{A.11}$$

This is the system of inequalities that we now strive to contradict.
Step 3: reaching a contradiction.
A direct computation gives

$$\begin{aligned}
\big(\partial_t \phi - \mathsf{H}(\nabla \phi)\big)(t_\alpha, x_\alpha) &= \alpha(t_\alpha - t'_\alpha) + \partial_t \Phi(t_\alpha, x_\alpha) + \frac{\varepsilon}{(T-t)^2} - \mathsf{H}\big(\nabla \phi(t_\alpha, x_\alpha)\big), \\
\big(\partial_t \phi' - \mathsf{H}(\nabla \phi')\big)(t'_\alpha, x'_\alpha) &= \alpha(t_\alpha - t'_\alpha) - \mathsf{H}\big(\nabla \phi'(t'_\alpha, x'_\alpha)\big),
\end{aligned}$$

where

$$\begin{aligned}
\partial_{x_k} \phi(t_\alpha, x_\alpha) &= \frac{\alpha}{2d} \cdot \frac{\big(x_\alpha - x'_\alpha\big)_k}{((x_\alpha - x'_\alpha)_k^2 + \varepsilon_0)^{\frac{1}{2}}} + \partial_{x_k} \Phi(t_\alpha, x_\alpha), \\
\partial_{x_k} \phi'(t'_\alpha, x'_\alpha) &= \frac{\alpha}{2d} \cdot \frac{\big(x_\alpha - x'_\alpha\big)_k}{((x_\alpha - x'_\alpha)_k^2 + \varepsilon_0)^{\frac{1}{2}}}.
\end{aligned}$$

It follows by the definition of $V = |||\mathsf{H}|||_{\mathrm{Lip},1,*}$ that

$$\begin{aligned}\big(\partial_t\phi' - \mathsf{H}(\nabla\phi')\big)(t'_\alpha,x'_\alpha) &\le \alpha(t_\alpha - t'_\alpha) - \mathsf{H}\big(\nabla\phi(t_\alpha,x_\alpha)\big) + V|||\nabla\phi(t_\alpha,x_\alpha)) - \nabla\phi'(t'_\alpha,x'_\alpha)|||_{1,*}\\ &\le \alpha(t_\alpha - t'_\alpha) - \mathsf{H}\big(\nabla\phi(t_\alpha,x_\alpha)\big) + Vd\max_{1\le k\le d}|\partial_{x_k}\Phi(t_\alpha,x_\alpha)|.\end{aligned}$$

A direct computation shows that $Vd|\partial_{x_k}\Phi(t_\alpha,x_\alpha)| \le \partial_t\Phi(t_\alpha,x_\alpha)$, so this can be bounded further by

$$\begin{aligned}\big(\partial_t\phi' - \mathsf{H}(\nabla\phi')\big)(t'_\alpha,x'_\alpha) &\le \alpha(t_\alpha - t'_\alpha) + \partial_t\Phi(t_\alpha,x_\alpha) - \mathsf{H}\big(\nabla\phi(t_\alpha,x_\alpha)\big)\\ &< \big(\partial_t\phi - \mathsf{H}(\nabla\phi)\big)(t_\alpha,x_\alpha) \le 0,\end{aligned}$$

where the strict inequality is due to the term $\frac{\varepsilon}{(T-t)^2}$ and the final inequality leverages the first inequality in (A.11). This contradicts the second inequality in (A.11) and completes the proof. ■

Corollary A.2. *Fix a non-linearity $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ satisfying (**A1**). If $u,v\in\mathfrak{L}_{\mathrm{unif}}$ are respectively a subsolution and a supersolution to (A.5), then*

$$\sup_{\mathbb{R}_{\ge 0}\times\mathbb{R}^d_{\ge 0}}\big(u(t,x)-v(t,x)\big) = \sup_{\mathbb{R}^d_{\ge 0}}\big(u(0,x)-v(0,x)\big). \tag{A.12}$$

Proof. Suppose for the sake of contradiction that there exists $(t^*,x^*)\in(0,\infty)\times\mathbb{R}^d_{\ge 0}$ with

$$u(t^*,x^*) - v(t^*,x^*) > \sup_{\mathbb{R}^d_{\ge 0}}\big(u(0,x)-v(0,x)\big).$$

Applying the comparison principle in Proposition A.1 with $R = |||x^*|||_1 + Vt^*$ yields a contradiction and completes the proof. ■

Proposition A.3. *Fix a non-linearity $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ satisfying (**A1**). If $f\in\mathfrak{L}$ is a viscosity solution to the Hamilton-Jacobi equation (A.5), then*

$$\sup_{t\ge 0}|||f(t,\cdot)|||_{\mathrm{Lip},1} = |||f(0,\cdot)|||_{\mathrm{Lip},1}. \tag{A.13}$$

Proof. Let $L = |||f(0,\cdot)|||_{\mathrm{Lip},1}$, and suppose for the sake of contradiction that there exists $T>0$ with

$$\begin{aligned}&\sup_{[0,T]\times\mathbb{R}^d_{\ge 0}\times\mathbb{R}^d_{\ge 0}}\big(f(t,x)-f(t,x')-L|||x-x'|||_1\big)\\ &\qquad > 0 \ge \sup_{x,x'\in\mathbb{R}^d_{\ge 0}}\big(f(0,x)-f(0,x')-L|||x-x'|||_1\big).\end{aligned} \tag{A.14}$$

The proof proceeds in three steps: first we perturb (A.14), then we use a variable doubling argument to obtain a system of inequalities, and finally we contradict this system of inequalities.
Step 1: perturbing.
Given $\varepsilon_0\in(0,1)$ to be determined, let $\theta\in C^\infty(\mathbb{R})$ be an increasing function with

$$(r-\varepsilon_0)_+ \le \theta(r) \le r_+$$

for all $r\in\mathbb{R}$, and consider the smoothed normalized-ℓ^1 norm,

$$|||x|||_{1,\varepsilon_0}=\frac{1}{d}\sum_{k=1}^{d}\left(x_k^2+\varepsilon_0\right)^{\frac{1}{2}}.$$

For a constant $R\in\mathbb{R}$ to be chosen, $Q>2L$ and $V=|||\mathsf{H}|||_{\mathrm{Lip},1,*}$, introduce the function

$$\Phi(t,x)=Q\theta\left(|||x|||_{1,\varepsilon_0}+Vt-R\right)$$

defined on $\mathbb{R}_{\geq 0}\times\mathbb{R}^d_{\geq 0}$. Given $\varepsilon>0$ to be determined, consider the functions

$$u_\varepsilon(t,x)=f(t,x)-\Phi(t,x)-\frac{\varepsilon}{T-t}\quad\text{and}\quad v(t,x)=f(t,x)+\Phi(t,x)$$

defined on $[0,T]\times\mathbb{R}^d_{\geq 0}$. Remembering (A.14) and choosing $R>0$ large enough and $\varepsilon,\varepsilon_0>0$ small enough guarantees that

$$\sup_{[0,T]\times\mathbb{R}^d_{\geq 0}\times\mathbb{R}^d_{\geq 0}}\left(u_\varepsilon(t,x)-v(t,x')-L|||x-x'|||_{1,\varepsilon_0}\right)$$
$$>\sup_{\mathbb{R}^d_{\geq 0}\times\mathbb{R}^d_{\geq 0}}\left(u_\varepsilon(0,x)-v(0,x')-L|||x-x'|||_1\right).\qquad\text{(A.15)}$$

This is a perturbed version of the absurd hypothesis (A.14). Before moving onto the variable doubling argument, observe that u_ε is a viscosity subsolution to the Hamilton-Jacobi equation (A.5) while v is a viscosity supersolution to this equation. Indeed, fix $\phi\in C^\infty\left((0,\infty)\times\mathbb{R}^d_{\geq 0}\right)$ with the property that $u_\varepsilon-\phi$ has a local maximum at $(t^*,x^*)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$. This means that the map

$$(t,x)\mapsto f(t,x)-\left(\phi(t,x)+\Phi(t,x)+\frac{\varepsilon}{T-t}\right)$$

has a local maximum at $(t^*,x^*)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$. It follows by the viscosity subsolution criterion for f that

$$\partial_t\phi(t^*,x^*)+\partial_t\Phi(t^*,x^*)+\frac{\varepsilon}{(T-t^*)^2}-\mathsf{H}\big(\nabla(\phi+\Phi)\big)(t^*,x^*)\leq 0.$$

A direct computation shows that $dV|\partial_{x_k}\Phi(t^*,x^*)|\leq\partial_t\Phi(t^*,x^*)$, so the definition of $V=|||\mathsf{H}|||_{\mathrm{Lip},1,*}$ implies that

$$\big(\partial_t\phi-\mathsf{H}(\nabla\phi)\big)(t^*,x^*)\leq\partial_t\phi(t^*,x^*)-\mathsf{H}\big(\nabla(\phi+\Phi)\big)(t^*,x^*)+V|||\nabla\Phi(t^*,x^*)|||_{1,*}\leq 0$$

which means that u_ε is a viscosity subsolution to the Hamilton-Jacobi equation (A.5). An identical argument reveals that v is a viscosity supersolution to this equation.

Step 2: system of inequalities.

Fix $\delta\in(0,1)$, and for each $\alpha\geq 1$ define the function $\Psi_\alpha:\mathbb{R}_{\geq 0}\times\mathbb{R}^d_{\geq 0}\times\mathbb{R}_{\geq 0}\times\mathbb{R}^d_{\geq 0}\to\mathbb{R}\cup\{-\infty\}$ by

$$\Psi_\alpha(t,x,t',x')=u_\varepsilon(t,x)-v(t',x')-\frac{\alpha}{2}|t-t'|^2-(L+\delta t)|||x-x'|||_{1,\varepsilon_0}.\qquad\text{(A.16)}$$

By doubling the variables and introducing the potential in this way, we ensure that the function Ψ_α achieves its supremum at a point $(t_\alpha, x_\alpha, t'_\alpha, x'_\alpha)$ which remains bounded as α tends infinity. Indeed, if we write $C > 0$ for a constant that depends on T, Q, R, V, $f(0,0)$ and $[f]_0$ whose value might not be the same at each occurrence, then for any $x \in \mathbb{R}^d_{\geq 0}$ with $|||x|||_{1,\varepsilon_0} > R+1$,

$$\Psi_\alpha(t,x,t',x') \leq L\big(|||x|||_1 + |||x'|||_1\big) - \Phi(t,x) - L|||x-x'|||_1 + C \leq (2L-Q)|||x|||_1 + C$$

where we have used the bound $\Phi(t,x) \geq Q(|||x|||_{1,\varepsilon_0} + Vt - R - 1)_+$. An analogous bound can be obtained with x replaced by x'. Remembering that $Q > 2L$ and observing that the supremum of (A.16) is bounded from below by $\Psi_\alpha(0,0,0,0)$ ensures that this function achieves its supremum at some point $(t_\alpha, x_\alpha, t'_\alpha, x'_\alpha)$ which remains bounded with respect to the normalized-ℓ^1 norm as α tends to infinity. In particular, the term $\alpha|t-t'|^2$ must remain bounded as α tends to infinity. It follows that, up to the extraction of a subsequence, there exist $t_0 \in [0,T]$ and $x_0, x'_0 \in \mathbb{R}^d_{\geq 0}$ such that $t_\alpha \to t_0$, $t'_\alpha \to t_0$, $x_\alpha \to x_0$ and $x'_\alpha \to x'_0$ as $\alpha \to \infty$. The term $\frac{\varepsilon}{T-t}$ in the definition of u_ε guarantees that $t_0 \in [0,T)$. On the other hand, if $C_1 > T(|||x_\alpha|||_{1,\varepsilon_0} + |||x'_\alpha|||_{1,\varepsilon_0})$ for all $\alpha \geq 1$, then the semi-continuity of u together with the bound

$$\Psi_\alpha(t_\alpha, x_\alpha, t'_\alpha, x'_\alpha) \geq -C_1\delta + \sup_{[0,T]\times\mathbb{R}^d_{\geq 0}\times\mathbb{R}^d_{\geq 0}} \big(u_\varepsilon(t,x) - v(t,x') - L|||x-x'|||_{1,\varepsilon_0}\big)$$

implies that

$$\Psi_\alpha(t_0, x_0, t_0, x'_0) \geq -C_1\delta + \sup_{[0,T]\times\mathbb{R}^d_{\geq 0}\times\mathbb{R}^d_{\geq 0}} \big(u_\varepsilon(t,x) - v(t,x') - L|||x-x'|||_{1,\varepsilon_0}\big). \tag{A.17}$$

To leverage this bound, observe that

$$\begin{aligned}\Psi_\alpha(0,x_0,0,x'_0) &= u_\varepsilon(0,x_0) - v(0,x'_0) - (L+\delta t)|||x_0 - x'_0|||_{1,\varepsilon_0} \\ &\leq \sup_{\mathbb{R}^d_{\geq 0}\times\mathbb{R}^d_{\geq 0}} \big(u_\varepsilon(0,x) - v(0,x') - L|||x-x'|||_1\big)\end{aligned} \tag{A.18}$$

and

$$\Psi_\alpha(t_0,x_0,t_0,x_0) = u_\varepsilon(t_0,x_0) - v(t_0,x_0) \leq \sup_{[0,T]\times\mathbb{R}^d_{\geq 0}} \Big(-2\Phi(t,x) - \frac{\varepsilon}{T-t}\Big).$$

Since θ is non-decreasing this can be bounded further by

$$\Psi_\alpha(0,x_0,0,x'_0) \leq \sup_{\mathbb{R}^d_{\geq 0}} \Big(-2\Phi(0,x) - \frac{\varepsilon}{T}\Big) \leq \sup_{\mathbb{R}^d_{\geq 0}\times\mathbb{R}^d_{\geq 0}} \big(u_\varepsilon(0,x) - v(0,x') - L|||x-x'|||_1\big), \tag{A.19}$$

where we have used the fact that $-2\Phi(0,x) - \frac{\varepsilon}{T} = u_\varepsilon(0,x) - v(0,x)$. Combining (A.17), (A.18) and (A.19) with the absurd assumption (A.15) and choosing δ small enough shows that $t_0 \in (0,T)$ and $x_0 \neq x'_0$. This means that, taking a subsequence if necessary, it is possible to guarantee that $t_\alpha, t'_\alpha \in (0,T)$ and $x_\alpha \neq x'_\alpha$ for all α large enough. We have therefore found a sequence of quadruples $((t_\alpha, x_\alpha, t'_\alpha, x'_\alpha))_\alpha$ with $t_\alpha, t'_\alpha \in (0,T)$ and $x_\alpha \neq x'_\alpha$ such that Ψ_α achieves its supremum at

$(t_\alpha,x_\alpha,t'_\alpha,x'_\alpha)$ for α large enough. With this in mind, fix $\alpha \geq 1$ large enough, and introduce the smooth functions $\phi,\phi' \in C^\infty\big((0,T)\times\mathbb{R}^d_{\geq 0}\big)$ defined by

$$\phi(t,x) = v(t'_\alpha,x'_\alpha)+\frac{\alpha}{2}|t-t'_\alpha|^2+(L+\delta t)|||x-x'_\alpha|||_{1,\varepsilon_0},$$
$$\phi'(t',x') = u_\varepsilon(t_\alpha,x_\alpha)-\frac{\alpha}{2}|t'-t_\alpha|^2-(L+\delta t_\alpha)|||x_\alpha-x'|||_{1,\varepsilon_0}.$$

Since $(t_\alpha,x_\alpha,t'_\alpha,x'_\alpha)$ maximizes Ψ_α, the function $u_\varepsilon-\phi$ achieves a local maximum at the point $(t_\alpha,x_\alpha)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$ while $v-\phi'$ achieves a local minimum at $(t'_\alpha,x'_\alpha)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$. It follows by the observation that u_ε is a viscosity subsolution while v is a viscosity supersolution that

$$\big(\partial_t\phi-\mathsf{H}(\nabla\phi)\big)(t_\alpha,x_\alpha)\leq 0 \quad\text{and}\quad \big(\partial_t\phi'-\mathsf{H}(\nabla\phi')\big)(t'_\alpha,x'_\alpha)\geq 0. \tag{A.20}$$

This is the system of inequalities that we now strive to contradict.
Step 3: reaching a contradiction.
A direct computation gives

$$\big(\partial_t\phi-\mathsf{H}(\nabla\phi)\big)(t_\alpha,x_\alpha) = \alpha(t_\alpha-t'_\alpha)+\delta|||x-x'_\alpha|||_{1,\varepsilon_0}-\mathsf{H}\big(\nabla\phi(t_\alpha,x_\alpha)\big),$$
$$\big(\partial_t\phi'-\mathsf{H}(\nabla\phi')\big)(t'_\alpha,x'_\alpha) = \alpha(t_\alpha-t'_\alpha)-\mathsf{H}\big(\nabla\phi'(t'_\alpha,x'_\alpha)\big),$$

where

$$\partial_{x_k}\phi(t_\alpha,x_\alpha)=\frac{L+\delta t_\alpha}{d}\cdot\frac{\big(x_\alpha-x'_\alpha\big)_k}{((x_\alpha-x'_\alpha)_k^2+\varepsilon_0)^{\frac{1}{2}}}=\partial_{x_k}\phi'(t'_\alpha,x'_\alpha).$$

It follows by the second inequality in (A.20) that

$$\big(\partial_t\phi-\mathsf{H}(\nabla\phi)\big)(t_\alpha,x_\alpha)=\big(\partial_t\phi'-\mathsf{H}(\nabla\phi')\big)(t'_\alpha,x'_\alpha)+\delta|||x-x'_\alpha|||_{1,\varepsilon_0}>0.$$

This contradicts the first inequality in (A.20) and completes the proof. ■

Corollary A.4. *Fix a non-linearity* $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ *and an initial condition* $\psi:\mathbb{R}^d\to\mathbb{R}$ *satisfying* (**A1**) *and* (**A3**). *If* $f_1,f_2\in\mathfrak{L}$ *are viscosity solutions to the Hamilton-Jacobi equation* (A.5) *with initial condition* ψ*, then* $f_1=f_2$.

Proof. This is an immediate consequence of Proposition A.3 and Corollary A.2. ■

A.2 Existence of solutions on $\mathbb{R}^d_{\geq 0}$

In this section, we fix a non-linearity $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ and an initial condition $\psi:\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ satisfying (**A1**)-(**A3**), and we use the classical Perron method to establish the existence of solutions to the Hamilton-Jacobi equation (A.5). We closely follow the arguments in Chapter 5 of [2]. It will be convenient to fix a positive constant

$$K>\sup\big\{|\mathsf{H}(y)|\mid y\in\mathbb{R}^d \text{ with } |||y|||_{1,*}\leq|||\psi|||_{\mathrm{Lip},1}\big\}, \tag{A.21}$$

and to define the continuous functions $\underline{u},\overline{u}:[0,\infty)\times\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ by

$$\underline{u}(t,x)=\psi(x)-Kt \quad\text{and}\quad \overline{u}(t,x)=\psi(x)+Kt. \tag{A.22}$$

The importance of these functions stems from the fact that they are a viscosity subsolution and a viscosity supersolution to the Hamilton-Jacobi equation (A.5), respectively.

Lemma A.5. *Fix a non-linearity $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ and an initial condition $\psi:\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ satisfying* (**A1**)-(**A3**)*. The functions $\underline{u}$ and $\overline{u}$ defined in* (A.22) *are a subsolution and a supersolution to the Hamilton-Jacobi equation* (A.5)*, respectively.*

Proof. Consider a smooth function $\phi\in C^\infty\big((0,\infty)\times\mathbb{R}^d_{\geq 0}\big)$ with the property that $\underline{u}-\phi$ has a local maximum at a point $(t^*,x^*)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$. For any $x\in\mathbb{R}^d_{\geq 0}$ and every $\varepsilon>0$,

$$\phi(t^*,x^*+\varepsilon x)-\phi(t^*,x^*)\geq \underline{u}(t^*,x^*+\varepsilon x)-\underline{u}(t^*,x^*)=\psi(x^*+\varepsilon x)-\psi(x^*).$$

Dividing by ε and letting ε tend to zero shows that $\nabla\phi(t^*,x^*)\cdot x\geq\nabla\psi(x^*)\cdot x$ for all $x\in\mathbb{R}^d_{\geq 0}$. It follows that $\nabla\phi(t^*,x^*)\geq\nabla\psi(x^*)$, so (**A2**) and the fact that $t^*>0$ imply that

$$\big(\partial_t\phi-\mathsf{H}(\nabla\phi)\big)(t^*,x^*)\leq\partial_t\underline{u}(t^*,x^*)-\mathsf{H}\big(\nabla\psi(x^*)\big)=-K-\mathsf{H}\big(\nabla\psi(x^*)\big).\tag{A.23}$$

To bound this further, observe that for every $x\in\mathbb{R}^d_{>0}$, $x'\in\mathbb{R}^d$ and $\varepsilon>0$ small enough,

$$\psi(x+\varepsilon x')-\psi(x)\leq\varepsilon|||\psi|||_{\mathrm{Lip},1}|||x'|||_1.$$

Dividing by ε and letting ε tend to zero gives

$$\nabla\psi(x)\cdot x'\leq|||\psi|||_{\mathrm{Lip},1}|||x'|||_1.$$

Choosing $x'=d\,\mathrm{sgn}(\partial_{x_k}\psi(x))e_k$ shows that $|||\nabla\psi(x)|||_{1,*}\leq|||\psi|||_{\mathrm{Lip},1}$. It follows by continuity that $|||\nabla\psi(x^*)|||_{1,*}\leq|||\psi|||_{\mathrm{Lip},1}$ so (A.23) and the definition of K imply that $\underline{u}$ is a subsolution to the Hamilton-Jacobi equation (A.5). An identical argument shows that $\overline{u}$ is a supersolution to the Hamilton-Jacobi equation (A.5). This completes the proof. ∎

The main result of this section will be that the function $f:[0,\infty)\times\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ defined by

$$f(t,x)=\sup_{u\in\mathscr{S}}u(t,x)\tag{A.24}$$

for the set

$$\mathscr{S}=\big\{u:[0,\infty)\times\mathbb{R}^d_{\geq 0}\to\mathbb{R}\mid\underline{u}\leq u\leq\overline{u}\text{ and }u^\star\text{ is a subsolution to (A.5)}\big\}\tag{A.25}$$

is a viscosity solution to the Hamilton-Jacobi equation (A.5). We refer to Appendix B for the definitions and basic properties of lower and upper semi-continuous envelopes of a function u, which we denote by $u_\star$ and $u^\star$ respectively. The strategy will be to show that $f^\star$ is a viscosity subsolution to the Hamilton-Jacobi equation (A.5) while $f_\star$ is a viscosity supersolution to this equation. The comparison principle in Corollary A.2 will then imply that f is a viscosity solution to the Hamilton-Jacobi equation (A.5). Throughout this subsection, we will write

$$B_r(t^*,x^*)=\big\{(t,x)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}\mid|t-t^*|^2+\|x-x^*\|_2^2\leq r^2\big\}.\tag{A.26}$$

for the Euclidean ball of radius $r>0$ centered at the point $(t^*,x^*)\in\mathbb{R}\times\mathbb{R}^d_{\geq 0}$. It is readily verified that $f^\star$ is a subsolution.

Lemma A.6. *If $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ satisfies (**A1**), then the upper semi-continuous envelope $f^\star$ of the function (A.24) is a viscosity subsolution to the Hamilton-Jacobi equation (A.5). In particular, $f\in\mathscr{S}$.*

Proof. Consider a smooth function $\phi\in C^\infty\big((0,\infty)\times\mathbb{R}^d_{\geq 0}\big)$ with the property that $f^\star-\phi$ has a strict local maximum at the point $(t^*,x^*)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$. To be more precise, suppose that

$$\big(f^\star-\phi\big)(t^*,x^*)>\big(f^\star-\phi\big)(t,x)$$

for all $(t,x)\in B_r(t^*,x^*)\setminus\{(t^*,x^*)\}$. By definition of the upper semi-continuous envelope and continuity of ϕ, it is possible to find points $(t_n,x_n)\in B_r(t^*,x^*)$ converging to (t^*,x^*) with

$$(f-\phi)(t_n,x_n)\geq(f^\star-\phi)(t^*,x^*)-\frac{1}{n}$$

for every integer $n\geq 1$. Similarly, by definition of f, it is possible find a sequence of functions $(u_n)\subset\mathscr{S}$ with

$$f(t_n,x_n)-\frac{1}{n}\leq u_n(t_n,x_n)$$

for all integer $n\geq 1$. If $(t_n',x_n')\in B_r(t^*,x^*)$ denotes the maximum of $u_n^\star-\phi$ on $B_r(t^*,x^*)$, then the fact that $u_n^\star$ is a subsolution to (A.5) implies that

$$\big(\partial_t\phi-\mathsf{H}(\nabla\phi)\big)(t_n',x_n')\leq 0. \tag{A.27}$$

Notice that $u_n^\star-\phi$ achieves its maximum on the compact set $B_r(t^*,x^*)$ as it is an upper semi-continuous function by Proposition B.3. Remembering that $u_n^\star\geq u_n$ reveals that

$$(f^\star-\phi)(t_n',x_n')\geq(u_n^\star-\phi)(t_n',x_n')\geq(u_n-f)(t_n,x_n)+(f-\phi)(t_n,x_n)\geq(f^\star-\phi)(t^*,x^*)-\frac{2}{n},$$

where we have used that $u_n^\star\leq f^\star$ as $u_n\leq f$. If (t_∞',x_∞') denotes any subsequential limit of (t_n',x_n'), then the upper semi-continuity of $f^\star$ established in Proposition B.3 gives

$$(f^\star-\phi)(t_\infty',x_\infty')\geq(f^\star-\phi)(t^*,x^*).$$

Since (t^*,x^*) is a strict local maximum of $f^\star-\phi$ on $B_r(t^*,x^*)$, this implies that $(t_\infty',x_\infty')=(t^*,x^*)$. Letting n tend to infinity in (A.27) shows that $f^\star$ is viscosity subsolution to the Hamilton-Jacobi equation (A.5). It is clear by the definition in (A.24) that $\underline{u}\leq f\leq\overline{u}$ so $f\in\mathscr{S}$. This completes the proof. ∎

Showing that $f_\star$ is a viscosity supersolution requires more work, and relies upon the following modification of Lemma 2.12 in [2].

Lemma A.7. *Fix a non-linearity $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ satisfying (**A1**) and (**A2**). If $u\in\mathscr{S}$ is such that $u_\star$ is not a viscosity supersolution to the Hamilton-Jacobi equation (A.5), then there exists $v\in\mathscr{S}$ with $v(t,x)>u(t,x)$ for some $(t,x)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$.*

Proof. The assumption that $u_\star$ is not a viscosity supersolution to the Hamilton-Jacobi equation (A.5) gives $\phi \in C^\infty\big((0,\infty)\times\mathbb{R}^d_{\geq 0}\big)$ and $(t^*,x^*)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$ with the property that $u_\star-\phi$ has a strict local minimum at (t^*,x^*) but $(\partial_t\phi-\mathsf{H}(\nabla\phi))(t^*,x^*)<0$. To be more precise, we will assume that

$$\big(u_\star-\phi\big)(t,x)>\big(u_\star-\phi\big)(t^*,x^*) \tag{A.28}$$

for all $(t,x)\in B_r(t^*,x^*)\setminus\{(t^*,x^*)\}$ and

$$\big(\partial_t\phi-\mathsf{H}(\nabla\phi)\big)(t,x)<-\varepsilon \tag{A.29}$$

for some $\varepsilon>0$ and all $(t,x)\in B_r(t^*,x^*)$. Notice that we must have $u_\star(t^*,x^*)<\overline{u}(t^*,x^*)$. Indeed, if this were not the case, the assumption that $u\in\mathscr{S}$ would imply that $u_\star(t^*,x^*)=\overline{u}(t^*,x^*)$, and therefore

$$\big(\overline{u}-\phi\big)(t,x)\geq\big(u_\star-\phi\big)(t,x)>(u_\star-\phi)(t^*,x^*)=\big(\overline{u}-\phi\big)(t^*,x^*)$$

for all $(t,x)\in B_r(t^*,x^*)\setminus\{(t^*,x^*)\}$. In other words, the supersolution $\overline{u}$ would be such that $\overline{u}-\phi$ achieves a local maximum at (t^*,x^*). This would contradict (A.29). Decreasing $r>0$ if necessary and using the continuity of ϕ and $\overline{u}$, it is therefore possible to find $\delta>0$ with

$$u_\star(t^*,x^*)+\delta<\overline{u}(t^*,x^*)-\delta\leq\overline{u}(t,x)\quad\text{and}\quad\phi(t,x)\leq\phi(t^*,x^*)+\frac{\delta}{2} \tag{A.30}$$

for all $(t,x)\in B_r(t^*,x^*)$. With this in mind, given $\varepsilon'<\frac{1}{4}\min(r^2,\delta)$, introduce the function

$$w(t,x)=\phi(t,x)+\varepsilon'-\|x-x^*\|_2^2-|t-t^*|^2+(u_\star-\phi)(t^*,x^*),$$

and define $v:[0,\infty)\times\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ by

$$v(t,x)=\begin{cases}\max\big(u(t,x),w(t,x)\big) & \text{if } (t,x)\in B_r(t^*,x^*),\\ u(t,x) & \text{if } (t,x)\notin B_r(t^*,x^*).\end{cases}$$

It is clear from the assumption $u\in\mathscr{S}$ that $v\geq u\geq\underline{u}$. Moreover, for $(t,x)\in B_r(t^*,x^*)$,

$$w(t,x)\leq\phi(t^*,x^*)+\frac{\delta}{2}+\frac{\delta}{2}+(u_\star-\phi)(t^*,x^*)=u_\star(t^*,x^*)+\delta\leq\overline{u}(t,x),$$

where we have used (A.30) and the fact that $\varepsilon'\leq\delta/2$. Together with the assumption $u\in\mathscr{S}$, this shows that $v\leq\overline{u}$, and therefore $\underline{u}\leq v\leq\overline{u}$. Furthermore, the definition of the lower semi-continuous envelope gives points $(t_n,x_n)\in B_r(t^*,x^*)$ with $(t_n,x_n)\to(t^*,x^*)$ and $u(t_n,x_n)\to u_\star(t^*,x^*)$. Since $v\geq w$ on $B_r(t^*,x^*)$, it follows that

$$\liminf_{n\to\infty}v(t_n,x_n)\geq\liminf_{n\to\infty}w(t_n,x_n)=\phi(t^*,x^*)+\varepsilon'+(u_\star-\phi)(t^*,x^*)=u_\star(t^*,x^*)+\varepsilon'.$$

This means that for any n large enough

$$v(t_n,x_n)\geq u(t_n,x_n)+\frac{\varepsilon'}{2}>u(t_n,x_n),$$

so there exists a point $(t,x) \in (0,\infty)\times\mathbb{R}^d_{\geq 0}$ with $v(t,x) > u(t,x)$. All that remains is to verify that $v^\star$ is a subsolution to the Hamilton-Jacobi equation (A.5). Consider $\beta \in C^\infty\big((0,\infty)\times\mathbb{R}^d_{\geq 0}\big)$ and $(t_0,x_0) \in (0,\infty)\times\mathbb{R}^d_{\geq 0}$ with the property that $v^\star - \beta$ has a strict local maximum on $B_{r'}(t_0,x_0)$ at (t_0,x_0). The definition of the upper semi-continuous envelope gives points $(t_n,x_n) \in B_{r'}(t_0,x_0)$ converging to (t_0,x_0) with

$$v(t_n,x_n) \geq v^\star(t_0,x_0) - \frac{1}{n}. \tag{A.31}$$

Since $v(t_n,x_n)$ is either equal to $w(t_n,x_n)$ or $u(t_n,x_n)$, passing to a subsequence, it is possible to assume that $v(t_n,x_n) = w(t_n,x_n)$ for all $n \geq 1$ or that $v(t_n,x_n) = u(t_n,x_n)$ for all $n \geq 1$. We treat these two cases separately.

Case 1: $v(t_n,x_n) = w(t_n,x_n)$ for all $n \geq 1$

In this case, we must have $(t_n,x_n) \in B_r(t^*,x^*)$ for all $n \geq 1$. Indeed, for $(t,x) \notin B_{r/2}(t^*,x^*)$,

$$(w-u)(t,x) \leq (w-u_\star)(t,x) \leq (\phi - u_\star)(t,x) + \varepsilon' - \frac{r^2}{4} + (u_\star - \phi)(t^*,x^*) < \varepsilon' - \frac{r^2}{4} \leq 0$$

where we have used (A.28) and the fact that $\varepsilon' < r^2/4$. If (t_n',x_n') denotes the maximum of $w-\beta$ on $B_r(t^*,x^*)\cap B_{r'}(t_0,x_0)$, arguing as in the proof of Lemma A.5 shows that

$$\begin{aligned}\partial_t\beta(t_n',x_n') &= \partial_t w(t_n',x_n') = \partial_t\phi(t_n',x_n') - 2(t_n'-t^*),\\ \nabla\beta(t_n',x_n') &\geq \nabla w(t_n',x_n') = \nabla\phi(t_n',x_n') - 2(x_n'-x^*).\end{aligned}$$

It follows by (**A2**), (**A1**) and (A.29) that

$$\begin{aligned}\big(\partial_t\beta - \mathsf{H}(\nabla\beta)\big)(t_n',x_n') &\leq \partial_t\phi(t_n',x_n') - \mathsf{H}\big(\nabla\phi(t_n',x_n')\big) + 2|t_n'-t^*| + 2|||\mathsf{H}|||_{\mathrm{Lip},1,*}|||x_n'-x^*|||_{1,*}\\ &\leq -\varepsilon + 2|t_n'-t^*| + 2d|||\mathsf{H}|||_{\mathrm{Lip},1,*}\|x_n'-x^*\|_2.\end{aligned}$$

Decreasing r if necessary, it is therefore possible to ensure that $\big(\partial_t\beta - \mathsf{H}(\nabla\beta)\big)(t_n',x_n') \leq 0$. To leverage this bound, observe that by (A.31), the continuity of β and the fact that (t_n,x_n) converges to (t_0,x_0),

$$(v^\star-\beta)(t_n',x_n') \geq (w-\beta)(t_n',x_n') \geq (w-\beta)(t_n,x_n) = (v-\beta)(t_n,x_n) \geq (v^\star-\beta)(t_0,x_0) - \frac{2}{n}.$$

In particular, any subsequential limit (t_∞',x_∞') of (t_n',x_n') must satisfy

$$(v^\star-\beta)(t_\infty',x_\infty') \geq (v^\star-\beta)(t_0,x_0) \quad\text{and}\quad \big(\partial_t\beta - \mathsf{H}(\nabla\beta)\big)(t_\infty',x_\infty') \leq 0.$$

Since (t_0,x_0) is a strict local maximum of $v^\star-\beta$ on $B_{r'}(t_0,x_0)$, the first of these inequalities shows that $(t_\infty',x_\infty') = (t_0,x_0)$ while the second implies the required subsolution criterion.

Case 2: $v(t_n,x_n) = u(t_n,x_n)$ for all $n \geq 1$

In this case, let (t_n',x_n') denote the maximum of $u^\star - \beta$ on $B_{r'}(t_0,x_0)$. Since $u^\star$ is a viscosity subsolution to the Hamilton-Jacobi equation (A.5),

$$\big(\partial_t\beta - \mathsf{H}(\nabla\beta)\big)(t_n',x_n') \leq 0.$$

On the other hand, the inequality $v \geq u$ and (A.31) reveal that

$$\begin{aligned}(v^\star-\beta)(t_n',x_n') \geq (u^\star-\beta)(t_n',x_n') \geq (u^\star-\beta)(t_n,x_n) \geq (u-\beta)(t_n,x_n) = (v-\beta)(t_n,x_n)\\ \geq (v^\star-\beta)(t_0,x_0)-\frac{1}{n},\end{aligned}$$

so any subsequential limit (t_∞',x_∞') of (t_n',x_n') must satisfy

$$(v^\star-\beta)(t_\infty',x_\infty') \geq (v^\star-\beta)(t_0,x_0) \quad \text{and} \quad \big(\partial_t\beta - \mathsf{H}(\nabla\beta)\big)(t_\infty',x_\infty') \leq 0.$$

Since (t_0,x_0) is a strict local maximum of $v^\star-\beta$ on $B_{r'}(t_0,x_0)$, the first of these inequalities shows that $(t_\infty',x_\infty')=(t_0,x_0)$ while the second implies the required subsolution criterion. This completes the proof. ■

Corollary A.8. *If* $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ *satisfies* (**A1**) *and* (**A2**)*, then the lower semi-continuous envelope* $f_\star$ *of the function* (A.24) *is a viscosity supersolution to the Hamilton-Jacobi equation* (A.5)*.*

Proof. Suppose for the sake of contradiction that $f_\star$ is not a supersolution to the Hamilton-Jacobi equation (A.5). Combining Lemma A.6 and Lemma A.7 gives a function $v\in\mathscr{S}$ and a point $(t,x)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$ with $v(t,x)>f(t,x)$. The contradiction

$$f(t,x)=\sup_{u\in\mathscr{S}} u(t,x)\geq v(t,x)>f(t,x)$$

completes the proof. ■

Together with Lemma A.6 and the comparison principle in Corollary A.2, this result allows us to establish the well-posedness of the Hamilton-Jacobi equation (A.5).

Proposition A.9. *If* $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ *and* $\psi:\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ *satisfy* (**A1**)-(**A3**)*, then the Hamilton-Jacobi equation* (A.5) *admits a unique viscosity solution* $f\in\mathfrak{L}$ *subject to the initial condition* ψ*. Moreover,*

$$\sup_{t>0}|||f(t,\cdot)|||_{\mathrm{Lip},1}=|||\psi|||_{\mathrm{Lip},1}. \tag{A.32}$$

Proof. Denote by $f\in\mathfrak{L}$ the function defined in (A.24). Combining Lemma A.6 and Corollary A.8 shows that $f^\star$ is a viscosity subsolution to the Hamilton-Jacobi equation (A.5) while $f_\star$ is a viscosity supersolution to this equation. By Proposition B.3 and continuity of $\underline{u}$ and $\overline{u}$, it is clear that $\underline{u}\leq f_\star\leq f\leq f^\star\leq\overline{u}$. Moreover, any function $h:[0,\infty)\times\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ with $\underline{u}\leq h\leq\overline{u}$ satisfies the bounds

$$\begin{aligned}\psi(x)=\underline{u}(0,x)\leq h(0,x)\leq\overline{u}(0,x)=\psi(x),\\ h(t,x)-h(0,x)\leq\overline{u}(t,x)-\psi(x)=Kt,\\ h(0,x)-h(t,x)\leq\psi(x)-\underline{u}(t,x)\leq Kt.\end{aligned}$$

for all $t>0$ and every $x\in\mathbb{R}^d_{\geq 0}$. In particular, we have $h\in\mathfrak{L}$, and therefore $f_\star,f,f^\star\in\mathfrak{L}$. It follows by the comparison principle in Corollary A.2 that $f^\star\leq f_\star$. Since $f_\star\leq f\leq f^\star$ by Proposition B.3, we must have $f=f_\star=f^\star$. In particular, the function $f\in\mathfrak{L}$ is a continuous viscosity solution to the Hamilton-Jacobi equation (A.5). The uniqueness of such a viscosity solution is guaranteed by Corollary A.4. Recalling Proposition A.3 gives the Lipschitz bound and completes the proof. ■

A.3 Equivalence of solutions on $\mathbb{R}^d_{\geq 0}$ and $\mathbb{R}^d_{>0}$

In this section, we leverage the monotonicity assumption (**A2**) of the non-linearity to show that viscosity solutions to the Hamilton-Jacobi equations (A.4) and (A.5) coincide. Combining this with Proposition A.9 gives a well-posedness theory for the Hamilton-Jacobi equation (A.4).

To ignore the boundary of the upper half-plane, we proceed as in Proposition 2.1 of [13] which is inspired by [18, 31]. The main difference between [13] and [18, 31] is in the definition of a distance-like function to the boundary of the domain on which the Hamilton-Jacobi equation is defined. Since this distance-like function will reappear in the next section when we show that the solution to the Hamilton-Jacobi equation (A.5) preserves the monotonicity of its initial condition, we will define it for a general closed convex cone $\mathscr{K} \subset \mathbb{R}^d$. Given a closed convex cone $\mathscr{K} \subset \mathbb{R}^d$, denote by

$$\mathscr{K}^* = \left\{x \in \mathbb{R}^d \mid x \cdot y \geq 0 \text{ for all } y \in \mathscr{K}\right\} \tag{A.33}$$

its dual cone, and define the distance-like function $\mathrm{d} : \mathscr{K}^* \to \mathbb{R}_{\geq 0}$ by

$$\mathrm{d}(x) = \inf_{\substack{|||y|||_{1,*}=1 \\ y \in \mathscr{K}}} y \cdot x. \tag{A.34}$$

The notion of a dual cone is reviewed in Appendix B. Before stating the main properties of this distance-like function, recall that the super-differential of a function $h : \mathscr{K}^* \to \mathbb{R}$ at a point $x \in \operatorname{int}(\mathscr{K}^*)$ is the set

$$\partial h(x) = \left\{p \in \mathbb{R}^d \mid h(x') \leq h(x) + p \cdot (x' - x) + o(x' - x) \text{ as } x' \to x \text{ in } \mathscr{K}^*\right\}. \tag{A.35}$$

Lemma A.10. *The function* $\mathrm{d} : \mathscr{K}^* \to \mathbb{R}_{\geq 0}$ *defined in* (A.34) *satisfies the following basic properties.*

1. *The infimum defining* $\mathrm{d}(x)$ *is achieved for every* $x \in \mathscr{K}^*$.

2. $\mathrm{d}(x) = 0$ *if and only if* $x \in \partial \mathscr{K}^*$.

3. d *is Lipschitz continuous with respect to the normalized-ℓ^1 norm. Moreover, it has Lipschitz constant at most one.*

4. d *is concave and* $\mathscr{K}^*$*-non-decreasing.*

5. *If* $x \in \operatorname{int}(\mathscr{K}^*)$*, then* $\partial\, \mathrm{d}(x) \subset \mathscr{K}$ *and* $|||p|||_{1,*} \leq 1$ *for any* $p \in \partial\, \mathrm{d}(x)$.

6. *If* $h : \mathbb{R}^d \to \mathbb{R}$ *is a differentiable function and* $x \mapsto h(x) - \frac{1}{\mathrm{d}(x)}$ *achieves a local maximum at a point* $x_0 \in \operatorname{int}(\mathscr{K}^*)$*, then* $-\mathrm{d}(x_0)^2 \nabla h(x_0) \in \partial\, \mathrm{d}(x_0)$.

Proof. We treat each property separately.

1. Consider a sequence $(y_n) \subset \mathscr{K}$ with $|||y_n|||_{1,*} = 1$ and $y_n \cdot x \to \mathrm{d}(x)$. Since (y_n) is uniformly bounded, it admits a subsequential limit $y \in \mathscr{K}$ with $|||y|||_{1,*} = 1$ and $y \cdot x = \mathrm{d}(x)$. We have used the equivalence and continuity of norms as well as the fact that $\mathscr{K}$ is closed. This shows that the infimum in the definition of $\mathrm{d}(x)$ is attained.

2. If $\mathrm{d}(x) = 0$, then there exists $y \in \mathscr{K}$ with $|||y|||_{1,*} = 1$ and $y \cdot x = 0$. This shows that $x \in \partial\mathscr{K}^*$. On the other hand, if $x \in \partial\mathscr{K}^*$, then there exists a non-zero $z \in \mathscr{K}$ with $z \cdot x = 0$. Taking $y = z/|||z|||_{1,*}$ gives $y \in \mathscr{K}$ with $|||y|||_{1,*} = 1$ and $\mathrm{d}(x) \leq y \cdot x = 0$. This shows that $\mathrm{d}(x) = 0$.

3. Fix $x, y \in \mathscr{K}^*$, and let $z \in \mathscr{K}$ with $|||z|||_{1,*} = 1$ be such that $\mathrm{d}(y) = z \cdot y$. By the Cauchy-Schwarz inequality,

$$\mathrm{d}(x) - \mathrm{d}(y) \leq z \cdot x - z \cdot y = z \cdot (y - x) \leq |||z|||_{1,*} |||y - x|||_1 = |||y - x|||_1.$$

Reversing the roles of x and y shows that d is Lipschitz continuous with Lipschitz constant at most one.

4. Fix $x, y \in \mathscr{K}^*$ as well as $t \in [0,1]$, and let $z \in \mathscr{K}$ achieve the infimum for $\mathrm{d}(tx + (1-t)y)$. It is clear that

$$\mathrm{d}(tx + (1-t)y) = z \cdot \big(tx + (1-t)y\big) = t(z \cdot x) + (1-t)(z \cdot y) \geq t\,\mathrm{d}(x) + (1-t)\,\mathrm{d}(y).$$

This shows that d is concave. To see that d is $\mathscr{K}^*$-non-decreasing, fix $x, x' \in \mathscr{K}^*$ with $x' - x \in \mathscr{K}^*$, and let $y \in \mathscr{K}$ attain the infimum defining $\mathrm{d}(x')$. Since $x' - x \in \mathscr{K}^*$,

$$\mathrm{d}(x') - \mathrm{d}(x) \geq y \cdot x' - y \cdot x = (x' - x) \cdot y \geq 0$$

as required.

5. Fix $x \in \mathrm{int}(\mathscr{K}^*)$, $z \in \mathscr{K}^*$ and $p \in \partial\,\mathrm{d}(x)$. Notice that $\partial\,\mathrm{d}(x) \neq \emptyset$ as d is concave. Since $\varepsilon z \in \mathscr{K}^*$ for every $\varepsilon > 0$, the $\mathscr{K}^*$-non-decreasingness of d and the definition of the super-differential imply that

$$0 \leq \mathrm{d}(x + \varepsilon z) - \mathrm{d}(x) \leq p \cdot \varepsilon z + o(\varepsilon z).$$

Dividing by ε and letting ε tend to zero shows that $p \cdot z \geq 0$ for all $z \in \mathscr{K}^*$. It follows by Proposition B.1 that $p \in \mathscr{K}^{**} = \mathscr{K}$. To see that $|||p|||_{1,*} \leq 1$ for $p \in \partial\,\mathrm{d}(x)$, find $\varepsilon > 0$ small enough so that $x + \varepsilon y \in \mathscr{K}^*$ for all $y \in \mathbb{R}^d$ with $|||y|||_1 = 1$. Fix $p \in \partial\,\mathrm{d}(x)$ and $y \in \mathbb{R}^d$ with $|||y|||_1 = 1$. Since $x - \varepsilon y \in \mathscr{K}^*$, the definition of the super-differential implies that

$$\mathrm{d}(x - \varepsilon y) \leq \mathrm{d}(x) - \varepsilon p \cdot y + o(\varepsilon y).$$

Rearranging and using the 1-Lipschitz continuity of d reveals that

$$\varepsilon p \cdot y \leq \varepsilon |||y|||_1 + o(\varepsilon y) = \varepsilon + o(\varepsilon y).$$

Dividing by ε and letting ε tend to zero shows that $p \cdot y \leq 1$ for every $y \in \mathbb{R}^d$ with $|||y|||_1 = 1$. Choosing $y = d\,\mathrm{sgn}(p_k) e_k$ gives $|||p|||_{1,*} \leq 1$.

6. Fix $z \in \mathscr{K}^*$. Since $x_0 \in \mathrm{int}(\mathscr{K}^*)$ is a local maximum of the map $x \mapsto h(x) - \frac{1}{\mathrm{d}(x)}$, for every $\varepsilon > 0$ small enough,

$$h(x_0) - \frac{1}{\mathrm{d}(x_0)} \geq h(x_0 + \varepsilon z) - \frac{1}{\mathrm{d}(x_0 + \varepsilon z)}.$$

Rearranging and using the 1-Lipschitz continuity of d as well as the differentiability of h reveals that

$$\begin{aligned}\mathrm{d}(x_0+\varepsilon z) &\leq \mathrm{d}(x_0)-\mathrm{d}(x_0)\,\mathrm{d}(x_0+\varepsilon z)\big(h(x_0+\varepsilon z)-h(x_0)\big)\\ &= \mathrm{d}(x_0)-\mathrm{d}(x_0)^2\nabla h(x_0)\cdot\varepsilon z+o(\varepsilon z).\end{aligned}$$

This shows that $-\mathrm{d}(x_0)^2\nabla h(x_0)\in\partial\,\mathrm{d}(x_0)$ and completes the proof. ∎

Proposition A.11. *If $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ is a continuous non-linearity satisfying (**A2**), then a continuous function $u:[0,\infty)\times\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ is a viscosity subsolution to the Hamilton-Jacobi equation (A.4) if and only if it is a viscosity subsolution to the Hamilton-Jacobi equation (A.5). An identical statement holds for viscosity supersolutions.*

Proof. The argument for viscosity subsolutions and viscosity supersolutions being almost identical, we focus exclusively on the case of viscosity subsolutions. To begin with, suppose that u is a viscosity subsolution to the Hamilton-Jacobi equation (A.5), and let $\phi\in C^\infty\big((0,\infty)\times\mathbb{R}^d_{>0}\big)$ be a function with the property that $u-\phi$ has a local maximum at $(t^*,x^*)\in(0,\infty)\times\mathbb{R}^d_{>0}$. After modifying ϕ outside of a neighborhood of (t^*,x^*) so that it becomes a smooth function defined on the larger domain $(0,\infty)\times\mathbb{R}^d_{\geq 0}$, we can apply the subsolution criterion for (A.5) and obtain the result.

Conversely, suppose that u is a continuous viscosity subsolution to the Hamilton-Jacobi equation (A.4), and consider a smooth function $\phi\in C^\infty\big((0,\infty)\times\mathbb{R}^d_{\geq 0}\big)$ with the property that $u-\phi$ has a local maximum at $(t^*,x^*)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$. If $x^*\in\mathbb{R}^d_{>0}$ there is nothing to prove, so assume that $x^*\in\partial\mathbb{R}^d_{\geq 0}$. Perturbing the test function ϕ by a small quadratic function if necessary, suppose further that (t^*,x^*) is a strict local maximum of $u-\phi$. To be more precise, assume that

$$u(t,x)-\phi(t,x)<u(t^*,x^*)-\phi(t^*,x^*) \tag{A.36}$$

for any (t,x) other than (t^*,x^*) in the closure of the open neighborhood

$$\mathscr{O}_r=(t^*-r,t^*+r)\times\big(\mathrm{int}\big(B_r(x^*)\big)\cap\mathbb{R}^d_{>0}\big).$$

Decreasing $r>0$ if necessary, it is possible to ensure that $(t^*-r,t^*+r)\subset(0,\infty)$. To establish the subsolution criterion for (A.5), we proceed in two steps: first we show that there exists an almost maximizer of $u-\phi$ in $\mathscr{O}_r$, and then we use a variable doubling argument to conclude.

Step 1: almost maximizer in $\mathscr{O}_r$.

Introduce the distance-like function $\mathrm{d}:\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ defined by

$$\mathrm{d}(x)=\inf_{\substack{|||y|||_{1,*}=1\\ y\in\mathbb{R}^d_{\geq 0}}} y\cdot x.$$

This corresponds to the function (A.34) for the cone $\mathscr{K}=\mathbb{R}^d_{\geq 0}=(\mathbb{R}^d_{\geq 0})^*$. For each $\varepsilon>0$, define the function $\psi_\varepsilon:(0,\infty)\times\mathbb{R}^d_{\geq 0}\to\mathbb{R}\cup\{-\infty\}$ by

$$\psi_\varepsilon(s,y)=u(s,y)-\phi(s,y)-\frac{\varepsilon}{\mathrm{d}(y)}.$$

Since ψ_ε is upper semi-continuous with values in $\mathbb{R}\cup\{-\infty\}$, there exists $(s_\varepsilon,y_\varepsilon)\in\overline{\mathscr{O}}_r$ with

$$\psi_\varepsilon(s_\varepsilon,y_\varepsilon)=\sup_{(s,y)\in\overline{\mathscr{O}}_r}\psi_\varepsilon(s,y).$$

Decreasing r if necessary and combining the continuity of $u-\phi$ with (A.36) gives $(t,x)\in\mathscr{O}_r$ such that

$$u(s,y)-\phi(s,y)<u(t,x)-\phi(t,x)$$

for all $(s,y)\in\overline{\mathscr{O}}_r$ with $s\in\{t^*-r,t^*+r\}$ or $y\in\partial B_r(x^*)\cap\mathbb{R}^d_{\geq 0}$. This means that for every $\varepsilon>0$ small enough, we must have $\psi_\varepsilon(s,y)<\psi_\varepsilon(t,x)$ for all $(s,y)\in\overline{\mathscr{O}}_r$ with $s\in\{t^*-r,t^*+r\}$ or $y\in\partial B_r(x^*)\cap\mathbb{R}^d_{\geq 0}$. Together with the term $\varepsilon/\mathrm{d}(y)$ in ψ_ε, this ensures that $(s_\varepsilon,y_\varepsilon)\in\mathscr{O}_r$.

Step 2: doubling the variables.

Fix $\varepsilon,\delta>0$, and a smooth function $\zeta_\varepsilon:\mathbb{R}\times\mathbb{R}^d\to[0,1]$ with

$$\operatorname{supp}\zeta_\varepsilon\subset\mathscr{O}_r\quad\text{and}\quad\zeta_\varepsilon(s_\varepsilon,y_\varepsilon)=1.$$

Given $0<\varepsilon_0<1$ to be determined, consider the smoothed normalized-ℓ^1 norm,

$$|||x|||_{1,\varepsilon_0}=\frac{1}{d}\sum_{k=1}^d\left(x_k^2+\varepsilon_0\right)^{\frac{1}{2}}.$$

For each $\theta>0$ introduce the modulus of continuity of u on $\overline{\mathscr{O}}_r$,

$$\omega_u(\theta)=\sup\left\{|u(t,x)-u(s,y)|\;\middle|\;(s,x),(s,y)\in\overline{\mathscr{O}}_r\text{ and }|t-s|^2+|||x-y|||_1\leq\theta^2\right\}.$$

Since u is continuous, and therefore uniformly continuous on $\overline{\mathscr{O}}_r$, it is possible to find $\theta>0$ sufficiently small that $\omega_u(\theta)<\delta$. With this $\theta>0$ at hand, define the function $\Psi_{\varepsilon,\delta,\theta}:\overline{\mathscr{O}}_r\times\overline{\mathscr{O}}_r\to\mathbb{R}$ by

$$\Psi_{\varepsilon,\delta,\theta}(t,x,s,y)=u(t,x)-\phi(s,y)-\frac{\varepsilon}{\mathrm{d}(y)}-\frac{2M_u}{\theta^2}|t-s|^2-\frac{2M_u}{\theta^2}|||x-y|||_{1,\varepsilon_0}+\delta\zeta_\varepsilon(s,y),$$

where $M_u=\sup_{(t,x)\in\overline{\mathscr{O}}_r}|u(t,x)|$. Observe that $\Psi_{\varepsilon,\delta,\theta}(s,y,s,y)=\psi_\varepsilon(s,y)+\delta\zeta_\varepsilon(s,y)$. We now show that the maximizer (t_0,x_0,s_0,y_0) of this function belongs to the open set $\mathscr{O}_r\times\mathscr{O}_r$. Given points $(t,x),(s,y)\in\overline{\mathscr{O}}_r\times\overline{\mathscr{O}}_r$ with $|t-s|^2+|||x-y|||_1\leq\theta^2$, the triangle inequality and the definition of the modulus of continuity reveal that

$$\Psi_{\varepsilon,\delta,\theta}(t,x,s,y)\leq\omega_u(\theta)+u(s,y)-\phi(s,y)-\frac{\varepsilon}{\mathrm{d}(y)}+\delta\zeta_\varepsilon(s,y)=\omega_u(\theta)+\psi_\varepsilon(s,y)+\delta\zeta_\varepsilon(s,y).$$

On the other hand, for $(t,x),(s,y)\in\overline{\mathscr{O}}_r\times\overline{\mathscr{O}}_r$ with $|t-s|^2+|||x-y|||_1>\theta^2$, the triangle inequality, the bound $|||x-y|||_{1,\varepsilon_0}\geq|||x-y|||_1$ and the definition of M_u imply that

$$\Psi_{\varepsilon,\delta,\theta}(t,x,s,y)\leq u(s,y)-\phi(s,y)-\frac{\varepsilon}{\mathrm{d}(y)}+\delta\zeta_\varepsilon(s,y)\leq\omega_u(\theta)+\psi_\varepsilon(s,y)+\delta\zeta_\varepsilon(s,y).$$

It follows that for any $(t,x)\in\overline{\mathscr{O}}_r$ and every $(s,y)\in\overline{\mathscr{O}}_r\setminus\operatorname{supp}\zeta_\varepsilon$,

$$\begin{aligned}\Psi_{\varepsilon,\delta,\theta}(t,x,s,y)&\leq\omega_u(\theta)+\psi_\varepsilon(s_\varepsilon,y_\varepsilon)+\delta\zeta_\varepsilon(s,y)=\omega_u(\theta)+\Psi_{\varepsilon,\delta,\theta}(s_\varepsilon,y_\varepsilon,s_\varepsilon,y_\varepsilon)-\delta\\&<\Psi_{\varepsilon,\delta,\theta}(s_\varepsilon,y_\varepsilon,s_\varepsilon,y_\varepsilon),\end{aligned}$$

where we have used that $\zeta_\varepsilon(s_\varepsilon, y_\varepsilon) = 1$ and $\omega_u(\theta) < \delta$. This means that $(s_0, y_0) \in \operatorname{supp} \zeta_\varepsilon \subset \mathscr{O}_r$. To show that (t_0, x_0) also belongs to this open set, suppose that $|t_0 - s_0|^2 + |||x_0 - y_0|||_1 > \theta^2$. The triangle inequality, the bound $|||x_0 - y_0|||_{1,\varepsilon_0} \geq |||x_0 - y_0|||_1$ and the definition of M_u imply that

$$\Psi_{\varepsilon,\delta,\theta}(t_0, x_0, s_0, y_0) \leq u(t_0, x_0) - \phi(s_0, y_0) - \frac{\varepsilon}{\mathrm{d}(y_0)} - 2M_u + \delta \zeta_\varepsilon(s_0, y_0) \leq \Psi_{\varepsilon,\delta,\theta}(s_0, y_0, s_0, y_0),$$

so, up to replacing (t_0, x_0) with (s_0, y_0), we may assume without loss of generality that

$$|t_0 - s_0|^2 + |||x_0 - y_0|||_1 \leq \theta^2. \tag{A.37}$$

Decreasing θ if necessary and recalling that $\mathscr{O}_r$ is open shows that $(t_0, x_0, s_0, y_0) \in \mathscr{O}_r \times \mathscr{O}_r$. Since the function $(t, x) \mapsto \Psi_{\varepsilon,\delta,\theta}(t, x, s_0, y_0)$ has a local maximum at (t_0, x_0), the subsolution criterion for (A.4) implies that

$$\frac{4M_u}{\theta^2}(t_0 - s_0) - \mathsf{H}\left(\frac{2M_u}{\theta^2} z\right) \leq 0 \tag{A.38}$$

for the vector $z \in \mathbb{R}^d$ defined by

$$z_k = \frac{(x_0 - y_0)_k}{d((x_0 - y_0)_k^2 + \varepsilon_0)^{\frac{1}{2}}}.$$

On the other hand, since the function $(s, y) \mapsto \Psi_{\varepsilon,\delta,\theta}(t_0, x_0, s, y)$ achieves its maximum at an interior point $(s_0, y_0) \in \mathscr{O}_r$, a direct computation together with Lemma A.10 shows that

$$\begin{aligned} &\partial_t \phi(s_0, y_0) - \frac{4M_u}{\theta^2}(t_0 - s_0) - \delta \partial_t \zeta_\varepsilon(s_0, y_0) = 0, \\ &\frac{\mathrm{d}(y_0)^2}{\varepsilon}\Big(\nabla \phi(s_0, y_0) + \frac{2M_u}{\theta^2}\widetilde{z} - \delta \nabla \zeta_\varepsilon(s_0, y_0)\Big) \in \partial\, \mathrm{d}(y_0), \end{aligned} \tag{A.39}$$

for the vector $\widetilde{z} \in \mathbb{R}^d$ defined by

$$\widetilde{z}_k = \frac{(y_0 - x_0)_k}{d((y_0 - x_0)_k^2 + \varepsilon_0)^{\frac{1}{2}}} = -z_k.$$

Remembering that $\partial\, \mathrm{d}(y_0) \subset (\mathbb{R}^d_{\geq 0})^* = \mathbb{R}^d_{\geq 0}$ by Lemma A.10, it is possible to find $p \geq 0$ with

$$\frac{2M_u}{\theta^2} z = \nabla \phi(s_0, y_0) - \delta \nabla \zeta_\varepsilon(s_0, y_0) - p.$$

Substituting this and (A.39) into (A.38) and using the fact that the non-linearity H is non-decreasing reveals that

$$\partial_t \phi(s_0, y_0) - \delta \partial_t \zeta_\varepsilon(s_0, y_0) - \mathsf{H}\big(\nabla \phi(s_0, y_0) - \delta \nabla \zeta_\varepsilon(s_0, y_0)\big) \leq 0. \tag{A.40}$$

Recalling that $(s_0, y_0) \in \operatorname{supp} \zeta_\varepsilon$ depends on ε, δ and θ, and that θ was chosen small enough in terms of δ, we would now like to let $\theta \to 0$ and then $\delta \to 0$ in this inequality. Observe that for any $(t, x) \in \overline{\mathscr{O}}_r$,

$$\begin{aligned} u(t_0, x_0) - \phi(s_0, y_0) - \frac{\varepsilon}{\mathrm{d}(y_0)} + \delta \zeta_\varepsilon(s_0, y_0) &\geq \Psi_{\varepsilon,\delta,\theta}(t_0, x_0, s_0, y_0) \\ &\geq u(t, x) - \phi(t, x) - \frac{\varepsilon}{\mathrm{d}(x)} + \delta \zeta_\varepsilon(t, x) \\ &= \psi_\varepsilon(t, x) + \delta \zeta_\varepsilon(t, x). \end{aligned}$$

If we denote by $(t_1,x_1) \in \operatorname{supp}\zeta_\varepsilon \subset \mathcal{O}_r$ and $(t_1',x_1') \in \operatorname{supp}\zeta_\varepsilon \subset \mathcal{O}_r$ subsequential limits of the sequences (t_0,x_0) and (s_0,y_0) as $\theta \to 0$ and then $\delta \to 0$, we must have $t_1 = t_1'$ and $x_1 = x_1'$ by (A.37). Moreover, the subsequential limit (t_1,x_1) must satisfy the inequality

$$u(t_1,x_1) - \phi(t_1,x_1) \geq u(t_1,x_1) - \phi(t_1,x_1) - \frac{\varepsilon}{\mathsf{d}(x_1)} \geq u(t,x) - \phi(t,x) - \frac{\varepsilon}{\mathsf{d}(x)}$$

for all $(t,x) \in \mathcal{O}_r$. Writing $(t_2,x_2) \in \overline{\mathcal{O}}_r$ for a subsequential limit of the sequence (t_1,x_1) as $\varepsilon \to 0$ we find that

$$u(t_2,x_2) - \phi(t_2,x_2) \geq u(t,x) - \phi(t,x)$$

for all $(t,x) \in \mathcal{O}_r$. By continuity of $u - \phi$, this inequality extends to $\overline{\mathcal{O}}_r$. Since (t^*,x^*) is a strict local maximum of $u - \phi$ on $\overline{\mathcal{O}}_r$, we must have $(t_2,x_2) = (t^*,x^*)$. It follows by letting $\theta \to 0$, then $\delta \to 0$ and finally $\varepsilon \to 0$ in (A.40) that

$$\partial_t \phi(t^*,x^*) - \mathsf{H}\big(\nabla\phi(t^*,x^*)\big) \leq 0.$$

This completes the proof. ∎

Corollary A.12. *If* $\mathsf{H} : \mathbb{R}^d \to \mathbb{R}$ *and* $\psi : \mathbb{R}^d_{\geq 0} \to \mathbb{R}$ *satisfy* (**A1**)-(**A3**)*, then the Hamilton-Jacobi equation* (A.4) *admits a unique viscosity solution* $f \in \mathfrak{L}$ *subject to the initial condition* ψ*. Moreover,*

$$\sup_{t>0} |||f(t,\cdot)|||_{\mathrm{Lip},1} = |||\psi|||_{\mathrm{Lip},1}, \tag{A.41}$$

and if $u,v \in \mathfrak{L}_{\mathrm{unif}}$ *are respectively a continuous subsolution and a continuous supersolution to* (A.4)*, then*

$$\sup_{\mathbb{R}_{\geq 0}\times\mathbb{R}^d_{\geq 0}} \big(u(t,x) - v(t,x)\big) = \sup_{\mathbb{R}^d_{\geq 0}} \big(u(0,x) - v(0,x)\big). \tag{A.42}$$

To be more specific, if $L = \max\big(\sup_{t>0} |||u(t,\cdot)|||_{\mathrm{Lip},1}, \sup_{t>0} |||v(t,\cdot)|||_{\mathrm{Lip},1}\big)$ *and* $V = |||\mathsf{H}|||_{\mathrm{Lip},1,*}$*, then for every* $Q > 2L$ *and* $R \in \mathbb{R}$*, the map*

$$(t,x) \mapsto u(t,x) - v(t,x) - Q\big(|||x|||_1 + Vt - R\big)_+ \tag{A.43}$$

achieves its supremum on $\{0\} \times \mathbb{R}^d_{\geq 0}$*.*

Proof. This is an immediate consequence of Proposition A.9, Proposition A.1, Corollary A.2 and Proposition A.11. ∎

A.4 Monotonicity of solutions on $\mathbb{R}^d_{\geq 0}$

Recall that the notion of being $\mathscr{C}^*$-non-decreasing is introduced at the beginning of Section 2. In this section, we follow the arguments in Section 4 of [13] to show that the solution to the Hamilton-Jacobi equation (A.5) preserves the monotonicity of its initial condition. To be more specific, we assume that

A4 the initial condition $\psi : \mathbb{R}^d \to \mathbb{R}$ is $\mathscr{C}^*$-non-decreasing for some closed convex cone $\mathscr{C} \subset \mathbb{R}^d$, and under a mild assumption on the dual cone $\mathscr{C}^*$, we show that the solution to the Hamilton-Jacobi equation (A.5) constructed in Proposition A.9 is also $\mathscr{C}^*$-non-decreasing. This result will be used in Section 2 when the well-posedness of the projected Hamilton-Jacobi equations (1.37) is established by means of Proposition A.9. Indeed, it will allow us to verify the first condition in (2.14) and (2.15).

Proposition A.13. *Fix a non-linearity* $\mathsf{H}:\mathbb{R}^d\to\mathbb{R}$ *and an initial condition* $\psi:\mathbb{R}^d_{\geq 0}\to\mathbb{R}$ *satisfying* (**A1**)-(**A4**). *If* $\mathbb{R}^d_{\geq 0}\cap\operatorname{int}(\mathscr{C}^*)\neq\emptyset$ *and* $f\in\mathfrak{L}$ *is a viscosity solution to the Hamilton-Jacobi equation* (A.5) *subject to the initial condition* ψ, *then* f *is* $\mathscr{C}^*$*-non-decreasing.*

Proof. Introduce the set $\Omega=\big\{(x,x')\in\mathbb{R}^d_{\geq 0}\times\mathbb{R}^d_{\geq 0}\mid x'-x\in\mathscr{C}^*\big\}$, and suppose for the sake of contradiction that there exists $T>0$ with

$$\sup_{\substack{t\in[0,T]\\(x,x')\in\Omega}}\big(f(t,x)-f(t,x')\big)>0\geq\sup_{(x,x')\in\Omega}\big(f(0,x)-f(0,x')\big). \tag{A.44}$$

The proof proceeds in three steps: first we perturb (A.44), then we use a variable doubling argument to obtain a system of inequalities, and finally we contradict this system of inequalities.

<u>Step 1: perturbing.</u>

Let $V=|||\mathsf{H}|||_{\mathrm{Lip},1,*}$ and fix a constant $L>0$ with

$$L>|||\psi|||_{\mathrm{Lip},1}\quad\text{and}\quad|f(t,x)-\psi(x)|\leq Lt$$

for all $(t,x)\in\mathbb{R}_{>0}\times\mathbb{R}^d_{\geq 0}$. The existence of such a constant follows from the assumption $f\in\mathfrak{L}$. Denote by $\mathrm{d}:\mathscr{C}^*\to\mathbb{R}_{\geq 0}$ the distance-like function (A.34) associated with the cone $\mathscr{C}$,

$$\mathrm{d}(y)=\inf_{\substack{|||y'|||_{1,*}=1\\y'\in\mathscr{C}}}y'\cdot y.$$

Fix $y_0\in\mathbb{R}^d_{\geq 0}\cap\operatorname{int}(\mathscr{C}^*)$ as well as $x_0\in\mathbb{R}^d_{\geq 0}$, and let $\theta\in C^\infty(\mathbb{R})$ be an increasing function with $r_+\leq\theta(r)\leq(r+1)_+$ for all $r\in\mathbb{R}$. Given $0<\varepsilon_0<1$ to be determined, consider the smoothed normalized-ℓ^1 norm,

$$|||x|||_{1,\varepsilon_0}=\frac{1}{d}\sum_{k=1}^d\big(x_k^2+\varepsilon_0\big)^{\frac{1}{2}},$$

and introduce the function

$$\Phi(t,x)=\theta\big(|||x|||_{1,\varepsilon_0}+Vt-R\big)$$

defined on $\mathbb{R}_{\geq 0}\times\mathbb{R}^d$, where $R>0$ is chosen large enough so that $\Phi(0,x_0)=0$. Increasing $R>0$ if necessary, it is possible to perturb the inequality (A.44) to ensure that

$$\sup_{\substack{t\in[0,T]\\(x,x')\in\Omega}}\big(f(t,x)-f(t,x')-\Phi(t,x)\big)>0\geq\sup_{(x,x')\in\Omega}\big(f(0,x)-f(0,x')-\Phi(0,x)\big).$$

Picking $\delta > 0$ small enough, it is also possible to guarantee that

$$\sup_{\substack{t\in[0,T]\\(x,x')\in\Omega}} \Big(f(t,x) - f(t,x') - \delta t - \zeta(t,t) - \Phi(t,x) - \frac{\delta}{\mathrm{d}(x'-x)} - 2\delta|||x-x'|||_{1,\varepsilon_0}^2\Big)$$
$$> \sup_{(x,x')\in\Omega} \Big(f(0,x) - f(0,x') - \zeta(0,0) - \Phi(0,x) - \frac{\delta}{\mathrm{d}(x'-x)} - 2\delta|||x-x'|||_{1,\varepsilon_0}^2\Big) \quad \text{(A.45)}$$

for the perturbation function

$$\zeta(t,t') = \frac{\delta}{T-t} + \frac{\delta}{T-t'}.$$

This is a perturbed version of the absurd hypothesis (A.44).

Step 2: system of inequalities.

For each $\alpha \geq 1$, define the function $\Psi_\alpha : [0,T]\times[0,T]\times\Omega\times\mathscr{C}^* \to \mathbb{R}\cup\{-\infty\}$ by

$$\Psi_\alpha(t,t',x,x',y) = f(t,x) - f(t',x') - \Phi(t,x) - \psi_\alpha(x,x',y)$$
$$- \delta t - \zeta(t,t') - \alpha|t-t'|^2 - \delta|||x-x'|||_{1,\varepsilon_0}^2, \quad \text{(A.46)}$$

where

$$\psi_\alpha(x,x',y) = \alpha|||x'-x-y|||_{1,\varepsilon_0}^2 + \frac{\delta}{\mathrm{d}(y)} + \delta|||y|||_{1,\varepsilon_0}^2.$$

Observe that $\Psi_\alpha(t,t,x,x',x'-x)$ coincides with the function being maximized in (A.45). By doubling the variables in this way, we ensure that the function Ψ_α achieves its supremum at a point $(t_\alpha,t'_\alpha,x_\alpha,x'_\alpha,y_\alpha)$ which remains bounded as α tends to infinity. Indeed, if we temporarily fix $\alpha \geq 1$ and let $(t_{\alpha,n},t'_{\alpha,n},x_{\alpha,n},x'_{\alpha,n},y_{\alpha,n})$ be a maximizing sequence for Ψ_α, then the choice of x_0 implies that when n is large enough,

$$\Psi_\alpha(t_{\alpha,n},t'_{\alpha,n},x_{\alpha,n},x'_{\alpha,n},y_{\alpha,n}) \geq \Psi_\alpha(0,0,x_0,x_0+y_0,y_0) = C_0 \quad \text{(A.47)}$$

for the constant $C_0 = f(0,x_0) - f(0,x_0+y_0) - \zeta(0,0) - \frac{\delta}{\mathrm{d}(y_0)} - 2\delta\|y_0\|_{1,\varepsilon_0}^2$. We have used the fact that $y_0 \in \mathbb{R}^d_{\geq 0}$ and that $\mathbb{R}^d_{\geq 0}$ is a cone. Combining this with the Lipschitz bound

$$|f(t,x) - f(t',x')| \leq L(t+t') + L|||x-x'|||_1 \leq L(t+t') + L|||x-x'|||_{1,\varepsilon_0} \quad \text{(A.48)}$$

and the fact that $\Phi(t,x) \geq |||x|||_{1,\varepsilon_0} - R$ reveals that

$$Lt_{\alpha,n} + Lt'_{\alpha,n} + L|||x_{\alpha,n} - x'_{\alpha,n}|||_{1,\varepsilon_0} + R - |||x_{\alpha,n}|||_{1,\varepsilon_0}$$
$$- \delta|||y_{\alpha,n}|||_{1,\varepsilon_0}^2 - \delta|||x_{\alpha,n} - x'_{\alpha,n}|||_{1,\varepsilon_0}^2 \geq C_0. \quad \text{(A.49)}$$

Noticing that $t_{\alpha,n}, t'_{\alpha,n} < T$ due to the presence of ζ in the function Ψ_α and observing that the quadratic function $r \mapsto Lr - \delta r^2$ is bounded by $\frac{L^2}{4\delta}$ gives the uniform boundedness of $x_{\alpha,n}$ and $y_{\alpha,n}$ in both n and α with respect to the normalized-ℓ^1 norm. Rearranging the lower bound (A.49) also shows that

$$2LT + L|||x_{\alpha,n} - x'_{\alpha,n}|||_{1,\varepsilon_0} + R - C_0 \geq \delta\|x_{\alpha,n} - x'_{\alpha,n}\|_{1,\varepsilon_0}^2$$

which gives the uniform boundedness of $x_{\alpha,n} - x'_{\alpha,n}$ and hence $x'_{\alpha,n}$ in both n and α with respect to the normalized-ℓ^1 norm. It is therefore possible to let n tend to infinity along a subsequence to obtain a maximizer $(t_\alpha, t'_\alpha, x_\alpha, x'_\alpha, y_\alpha)$ of Ψ_α all of whose components are bounded by some constant $C_1 > 0$ that is independent of α with respect to the normalized-ℓ^1 norm. Choosing ε_0 small enough, these components will also be assumed to be bounded by $C_1 > 0$ with respect to the smoothed version of the normalized-ℓ^1 norm. We now obtain some essential bounds on the components of this maximizer. Taking the limit as n tends to infinity in the inequality (A.47) reveals that

$$f(t_\alpha, x_\alpha) - f(t'_\alpha, x'_\alpha) - \frac{\delta}{\mathrm{d}(y_\alpha)} - \delta |||y_\alpha|||_{1,\varepsilon_0}^2 - \alpha |t_\alpha - t'_\alpha|^2 - \delta |||x_\alpha - x'_\alpha|||_{1,\varepsilon_0}^2 \geq C_0.$$

Combining this with (A.48) gives

$$\begin{aligned}\alpha |t_\alpha - t'_\alpha|^2 + \frac{\delta}{\mathrm{d}(y_\alpha)} + \delta |||y_\alpha|||_{1,\varepsilon_0}^2 &\leq L(t_\alpha + t'_\alpha) + L|||x_\alpha - x'_\alpha|||_{1,\varepsilon_0} - \delta |||x_\alpha - x'_\alpha|||_{1,\varepsilon_0}^2 - C_0 \\ &\leq 2LT + \frac{L^2}{4\delta} - C_0,\end{aligned}$$

where we again used the fact that the quadratic function $r \mapsto Lr - \delta r^2$ is bounded by $\frac{L^2}{4\delta}$. If we introduce the constant $C_2 = \max\big(2LT + \frac{L^2}{4\delta} - C_0, 1\big)$, this upper bound implies that

$$|t_\alpha - t'_\alpha| \leq \sqrt{\frac{C_2}{\alpha}}, \qquad \mathrm{d}(y_\alpha) \geq \frac{\delta}{C_2}, \qquad \|y_\alpha\|_{1,\varepsilon_0} \leq \sqrt{\frac{C_2}{\delta}}. \tag{A.50}$$

In particular $y_\alpha \in \mathrm{int}(\mathscr{C}^*)$, so $|||p|||_{1,*} \leq 1$ for every $p \in \partial\, \mathrm{d}(y_\alpha)$ by Lemma A.10. To leverage this observation, notice that $y \mapsto \Psi_\alpha(t_\alpha, t'_\alpha, x_\alpha, x'_\alpha, y)$ achieves a local maximum at y_α, and therefore so does $y \mapsto -\psi_\alpha(x_\alpha, x'_\alpha, y)$. It follows by a direct computation and Lemma A.10 that the vector $p \in \mathbb{R}^d$ defined by

$$p_k = \frac{2\,\mathrm{d}(y_\alpha)^2}{\delta}\bigg(\alpha |||x'_\alpha - x_\alpha - y_\alpha|||_{1,\varepsilon_0} \frac{(x'_\alpha - x_\alpha - y_\alpha)_k}{((x'_\alpha - x_\alpha - y_\alpha)_k^2 + \varepsilon_0)^{\frac{1}{2}}} + \delta |||y_\alpha|||_{1,\varepsilon_0} \frac{(y_\alpha)_k}{((y_\alpha)_k^2 + \varepsilon_0)^{\frac{1}{2}}}\bigg)$$

belongs to the super-differential $\partial\, \mathrm{d}(y_\alpha)$. This means that

$$\alpha |||x'_\alpha - x_\alpha - y_\alpha|||_{1,\varepsilon_0} \max_{1 \leq k \leq d} \bigg| \frac{(x'_\alpha - x_\alpha - y_\alpha)_k}{((x'_\alpha - x_\alpha - y_\alpha)_k^2 + \varepsilon_0)^{\frac{1}{2}}} \bigg| \leq \frac{\delta}{2d\,\mathrm{d}(y_\alpha)^2} + \delta |||y_\alpha|||_{1,\varepsilon_0},$$

where we have used the bounds $|||p|||_{1,*} \leq 1$ and $|(y_\alpha)_k| \leq ((y_\alpha)_k^2 + \varepsilon_0)^{\frac{1}{2}}$. To bound this further, suppose that

$$|||x'_\alpha - x_\alpha - y_\alpha|||_{1,\varepsilon_0} > C \tag{A.51}$$

for some constant C to be determined, and let $1 \leq k^* \leq d$ be such that $((x'_\alpha - x_\alpha - y_\alpha)_{k^*}^2 + \varepsilon_0)^{\frac{1}{2}} > C$. Observe that for any $z \in \mathbb{R}^d$ with $(z_{k^*}^2 + \varepsilon_0)^{\frac{1}{2}} > C$,

$$\frac{|z_{k^*}|}{(z_{k^*}^2 + \varepsilon_0)^{\frac{1}{2}}} = \frac{|z_{k^*}| + \sqrt{\varepsilon_0}}{(z_{k^*}^2 + \varepsilon_0)^{\frac{1}{2}}} - \frac{\sqrt{\varepsilon_0}}{(z_{k^*}^2 + \varepsilon_0)^{\frac{1}{2}}} \geq 1 - \frac{\sqrt{\varepsilon_0}}{C} = \frac{C - \sqrt{\varepsilon_0}}{C},$$

where we have used the fact that $|z_k|+\sqrt{\varepsilon_0}\geq (z_k^2+\varepsilon_0)^{\frac{1}{2}}$. Together with (A.50), this implies that

$$\frac{\alpha(C-\sqrt{\varepsilon_0})}{C}|||x'_\alpha-x_\alpha-y_\alpha|||_{1,\varepsilon_0}\leq \frac{\delta}{2d\,\mathrm{d}(y_\alpha)^2}+\delta|||y_\alpha|||_{1,\varepsilon_0}\leq K$$

for the constant $K=\frac{C_2^2}{2d\delta}+\sqrt{\delta C_2}$. Rearranging, remembering (A.51) and choosing $C=\sqrt{\varepsilon_0}+\frac{K}{\alpha}$ reveals that

$$|||x'_\alpha-x_\alpha-y_\alpha|||_1\leq |||x'_\alpha-x_\alpha-y_\alpha|||_{1,\varepsilon_0}\leq \max\Big(C,\frac{KC}{\alpha(C-\sqrt{\varepsilon_0})}\Big)=\frac{K}{\alpha}+\sqrt{\varepsilon_0}.$$

Letting ε_0 tend to zero in this upper bound yields

$$|||x'_\alpha-x_\alpha-y_\alpha|||_1\leq \frac{K}{\alpha}. \tag{A.52}$$

Combining this with the first bound in (A.50) and the fact that each component in the sequence of maximizers $(t_\alpha,t_\alpha,x_\alpha,x'_\alpha,y_\alpha)$ is uniformly bounded by a constant independent of α gives the existence of a subsequential limit $(t_\infty,t_\infty,x_\infty,x'_\infty,x'_\infty-x_\infty)$ with respect to the normalized-ℓ^1 norm. Observe that for any $t\in[0,T)$ and every $(x,x')\in\Omega$,

$$\begin{aligned}f(t_\alpha,x_\alpha)-f(t'_\alpha,x'_\alpha)-\delta t_\alpha-\zeta(t_\alpha,t'_\alpha)-\Phi(t_\alpha,x_\alpha)-\frac{\delta}{\mathrm{d}(y_\alpha)}-\delta|||y_\alpha|||_{1,\varepsilon_0}^2-\delta|||x_\alpha-x'_\alpha|||_{1,\varepsilon_0}\\ \geq \Psi_\alpha(t_\alpha,t'_\alpha,x_\alpha,x'_\alpha,y_\alpha)\geq \Psi_\alpha(t,t,x,x',x'-x).\end{aligned}$$

Taking the supremum over $(t,x,x')\in[0,T]\times\Omega$, recalling that $\Psi_\alpha(t,t,x,x',x'-x)$ coincides with the function being maximized in (A.45) and letting $\alpha\to\infty$ shows that $t_\infty>0$. At this point, we can use the fact that f is a viscosity solution to the Hamilton-Jacobi equation (A.5) to obtain a system of inequalities. Using the second inequality in (A.50), the bound (A.52) and the observation that $t_\infty>0$, fix $\alpha\geq 1$ large enough so that $x'_\alpha-x_\alpha\in\mathrm{int}(\mathscr{C}^*)$ and $t_\alpha,t'_\alpha>0$. Introduce the smooth functions

$$\phi(t,x)=f(t,x)-\Psi_\alpha(t,t'_\alpha,x,x'_\alpha,y_\alpha)\quad\text{and}\quad \phi'(t',x')=f(t',x')+\Psi_\alpha(t_\alpha,t',x_\alpha,x',y_\alpha)$$

defined on $(0,\infty)\times\mathbb{R}^d_{\geq 0}$. Since $(t_\alpha,t'_\alpha,x_\alpha,x'_\alpha,y_\alpha)$ maximizes Ψ_α, the function $f-\phi$ achieves a local maximum at $(t_\alpha,x_\alpha)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$ while the function $f-\phi'$ achieves a local minimum at $(t'_\alpha,x'_\alpha)\in(0,\infty)\times\mathbb{R}^d_{\geq 0}$. It follows by definition of a viscosity solution that

$$\partial_t\phi(t_\alpha,x_\alpha)-\mathsf{H}\big(\nabla\phi(t_\alpha,x_\alpha)\big)\leq 0\quad\text{and}\quad \partial_t\phi'(t'_\alpha,x'_\alpha)-\mathsf{H}\big(\nabla\phi'(t'_\alpha,x'_\alpha)\big)\geq 0. \tag{A.53}$$

This is the system of inequalities that we now strive to contradict.
Step 3: reaching a contradiction.
The choice $V=|||\mathsf{H}|||_{\mathrm{Lip},1,*}$ and a direct computation reveal that

$$\begin{aligned}\partial_t\phi'(t'_\alpha,x'_\alpha)-\mathsf{H}\big(\nabla\phi'(t'_\alpha,x'_\alpha)\big)<\delta+\frac{\delta}{(T-t_\alpha)^2}+V|||\nabla\Phi(t_\alpha,x_\alpha)|||_{1,*}+2\alpha(t_\alpha-t'_\alpha)\\ -\mathsf{H}\big(\nabla\phi(t_\alpha,x_\alpha)\big).\end{aligned}$$

Another direct computation shows that

$$\partial_t \phi(t_\alpha, x_\alpha) = \delta + \frac{\delta}{(T - t_\alpha)^2} + \partial_t \Phi(t_\alpha, x_\alpha) + 2\alpha(t_\alpha - t'_\alpha)$$

and that $dV|\partial_{x_k}\Phi(t_\alpha, x_\alpha)| \leq \partial_t \Phi(t_\alpha, x_\alpha)$. It follows by the first inequality in (A.53) that

$$\partial_t \phi'(t'_\alpha, x'_\alpha) - \mathsf{H}\big(\nabla\phi'(t'_\alpha, x'_\alpha)\big) < \partial_t \phi(t_\alpha, x_\alpha) - \mathsf{H}\big(\nabla\phi(t_\alpha, x_\alpha)\big) \leq 0$$

which contradicts the second inequality in (A.53) and completes the proof. ■

B Background material

In this appendix, we establish three elementary results in analysis. The first is a classical result in convex analysis regarding the bidual of a closed convex cone. Recall that the dual of a convex cone $\mathscr{K} \subset \mathbb{R}^d$ is the closed convex cone

$$\mathscr{K}^* = \big\{x \in \mathbb{R}^d \mid x \cdot y \geq 0 \text{ for all } y \in \mathscr{K}\big\}. \tag{B.1}$$

It is clear that any convex cone $\mathscr{K}$ is always a subset of its bidual $\mathscr{K}^{**}$. Since $\mathscr{K}^{**}$ is closed, a necessary condition for this containment to be an equality is that $\mathscr{K}$ be closed; it turns out that this is also a sufficient condition. This is often deduced from the Hahn-Banach separation theorem [25] or the Fenchel-Moreau theorem [3]. For the reader's convenience we prove this duality result using the Hahn-Banach separation theorem as stated in Theorem 4.1.1 of [25].

Proposition B.1. *If $\mathscr{K} \subset \mathbb{R}^d$ is a non-empty closed convex cone, then $\mathscr{K} = \mathscr{K}^{**}$.*

Proof. It is clear that $\mathscr{K} \subset \mathscr{K}^{**}$. Suppose for the sake of contradiction that there exists $x \in \mathscr{K}^{**}$ with $x \notin \mathscr{K}$. Since $\mathscr{K}$ is a non-empty closed convex set, the Hahn-Banach separation theorem gives $\alpha \in \mathbb{R}^d$ with

$$\alpha \cdot x > \sup\{\alpha \cdot y \mid y \in \mathscr{K}\}. \tag{B.2}$$

Given $x_0 \in \mathscr{K}$, the assumption that $\mathscr{K}$ is closed implies that $0 = \lim_{n\to\infty} \frac{1}{n} x_0 \in \mathscr{K}$. Together with (B.2), this means that $\alpha \cdot x > 0$. If there were $y_0 \in \mathscr{K}$ with $\alpha \cdot y > 0$, the fact that $\mathscr{K}$ is a cone would imply that $\alpha \cdot x \geq \lambda \alpha \cdot y_0$ for all $\lambda > 0$, and letting λ tend to infinity would give a contradiction. It follows by (B.2) that

$$\alpha \cdot x > 0 = \sup\{\alpha \cdot y \mid y \in \mathscr{K}\},$$

where we have used that $0 \in \mathscr{K}$. The lower bound implies that $-\alpha \in \mathscr{K}^*$ while the upper bound gives $x \cdot (-\alpha) < 0$. This contradicts the assumption that $x \in \mathscr{K}^{**}$ and completes the proof. ■

The second also belongs to the realm of convex analysis, and it gives a non-differential characterization of a Lipschitz function having its gradient in a closed convex set.

Proposition B.2. *If $\mathscr{K} \subset \mathbb{R}^d$ is a closed convex set and $\psi : \mathbb{R}^d \to \mathbb{R}$ is a Lipschitz function, then $\nabla\psi \in \mathscr{K}$ if and only if the following holds. For every $c \in \mathbb{R}$ and $x, x' \in \mathbb{R}^d$ with the property that for every $z \in \mathscr{K}$, $(x' - x) \cdot z \geq c$, we have $\psi(x') - \psi(x) \geq c$.*

Proof. Suppose that $\nabla\psi \in \mathscr{K}$, and fix $c \in \mathbb{R}$ and $x, x' \in \mathbb{R}^d$ with $(x'-x)\cdot z \geq c$ for every $z \in \mathscr{K}$. If we knew that ψ was almost everywhere differentiable along the line joining x and x', we could apply the fundamental theorem of calculus to the one-dimensional Lipschitz function $t \mapsto \psi(x+t(x'-x))$ and conclude that

$$\psi(x') - \psi(x) = \int_0^1 \nabla\psi\big(x+t(x'-x)\big)\cdot(x'-x)\,\mathrm{d}t \geq c.$$

Although ψ could fail to be differentiable almost everywhere on the line joining x and x', we will now fix $\varepsilon > 0$ and show that it must be differentiable almost everywhere on some line joining some point $x_\varepsilon \in B_\varepsilon(x)$ and some point $x'_\varepsilon \in B_\varepsilon(x')$. Denote by

$$\mathscr{H} = \big\{y \in \mathbb{R}^d \mid y\cdot(x'-x) = 0\big\} \cong \mathbb{R}^{d-1}$$

the hyperplane perpendicular to the line segment joining x and x', and write

$$\mathscr{A}_{\varepsilon,x} = B_\varepsilon(x) \cap (x+\mathscr{H})$$

for the cross-section of $B_\varepsilon(x)$ through x and perpendicular to the line segment joining x and x'. Denote by $\mathscr{L}$ the set of line segments between points in $\mathscr{A}_{\varepsilon,x}$ and points in $\mathscr{A}_{\varepsilon,x'}$ which are parallel to the line segment joining x and x'. For each $y \in \mathscr{A}_{\varepsilon,x}$, write $\ell_y \in \mathscr{L}$ for the unique line segment in $\mathscr{A}_{\varepsilon,x}$ through y, and introduce the set

$$\mathscr{D}_y = \big\{z \in \ell_y \mid \psi \text{ is not differentiable at } z\big\}$$

of points on ℓ_y at which ψ is not differentiable. If $\mathscr{D}_y$ were a set of positive one-dimensional Lebesgue measure $m_1(\mathscr{D}_y) > 0$ for every $y \in \mathscr{A}_{\varepsilon,x}$, then the d-dimensional Lebesgue measure of the set of points in $\cup_{y\in\mathscr{A}_{\varepsilon,x}}\ell_y$ at which ψ is not differentiable would have positive measure,

$$\int_{\mathscr{A}_{\varepsilon,x}} m_1(\mathscr{D}_y)\,\mathrm{d}y > 0.$$

This would contradict Rademacher's theorem on the almost everywhere differentiability of Lipschitz functions (see Theorem 6 in Chapter 5.8 of [20]). It is therefore possible to find $x_\varepsilon \in \mathscr{A}_{\varepsilon,x}$ with $m_1(\mathscr{D}_{x_\varepsilon}) = 0$. If we write $x'_\varepsilon \in \mathscr{A}_{\varepsilon,x'}$ for the right endpoint of ℓ_{x_ε}, then the fundamental theorem of calculus implies that

$$\psi(x'_\varepsilon) - \psi(x_\varepsilon) = \int_0^1 \nabla\psi\big(x_\varepsilon+t(x'_\varepsilon-x_\varepsilon)\big)\cdot(x'-x)\,\mathrm{d}t \geq c.$$

Letting ε tend to zero shows that $\psi(x') - \psi(x) \geq c$ as required. Conversely, suppose that for every $c \in \mathbb{R}$ and $x, x' \in \mathbb{R}^d$ with the property that for every $z \in \mathscr{K}$, $(x'-x)\cdot z \geq c$, we have $\psi(x') - \psi(x) \geq c$. Assume for the sake of contradiction that there exists $y \in \mathbb{R}^d$ with $\nabla\psi(y) \notin \mathscr{K}$. The Hahn-Banach separation theorem gives $v \in \mathbb{R}^d$ and $\delta > 0$ with

$$v\cdot\nabla\psi(y) + \delta < \inf\{v\cdot z \mid z \in \mathscr{K}\}.$$

It follows that

$$\psi(y+\varepsilon v) - \psi(y) \geq \varepsilon\big(v\cdot\nabla\psi(y) + \delta\big).$$

Dividing by ε and letting ε tend to zero reveals that $\nabla\psi(y)\cdot v \geq v\cdot\nabla\psi(y) + \delta$. This contradiction completes the proof. ■

The third elementary result in analysis that we will prove regards the basic properties of semi-continuous envelopes. To strive for generality, fix a set $X \subset \mathbb{R}^d$ endowed with a norm $\|\cdot\|$. Recall that a function $u : X \to \mathbb{R}$ is said to be *upper semi-continuous* at a point $x \in X$ if

$$u(x) \geq \limsup_{y \to x} u(y) := \lim_{r \searrow 0} \sup \big\{ u(y) \mid y \in X \text{ with } \|y - x\| \leq r \big\}, \tag{B.3}$$

and it is said to be *lower semi-continuous* at a point $x \in X$ if

$$u(x) \leq \liminf_{y \to x} u(y) := \lim_{r \searrow 0} \inf \big\{ u(y) \mid y \in X \text{ with } \|y - x\| \leq r \big\}. \tag{B.4}$$

Moreover, the *upper semi-continuous envelope* of u is the function $u^\star : X \to \mathbb{R}$ defined by

$$u^\star(x) = \limsup_{y \to x} u(y) = \lim_{r \searrow 0} \sup \big\{ u(y) \mid y \in X \text{ with } \|y - x\| \leq r \big\}, \tag{B.5}$$

while its *lower semi-continuous envelope* is the function $u_\star : X \to \mathbb{R}$ defined by

$$u_\star(x) = \liminf_{y \to x} u(y) = \lim_{r \searrow 0} \inf \big\{ u(y) \mid y \in X \text{ with } \|y - x\| \leq r \big\}. \tag{B.6}$$

The following proposition collects the basic properties of semi-continuous envelopes. This result is used in Appendix A.2 with $X = [0, \infty) \times \mathbb{R}^d_{\geq 0}$ and $\|(t, x)\| = |t| + \||x|\|_1$.

Proposition B.3. *The semi-continuous envelopes of a locally bounded function $u : X \to \mathbb{R}$ satisfy the following basic properties.*

1. *$u_\star(x) \leq u(x) \leq u^\star(x)$ for all $x \in X$.*
2. *$u^\star(x) = \min\{v(x) \mid u \leq v$ and v is upper semi-continuous$\}$ for all $x \in X$. In particular, $u^\star$ is upper semi-continuous.*
3. *$u_\star(x) = \max\{v(x) \mid v \leq u$ and v is lower semi-continuous$\}$ for all $x \in X$. In particular, $u_\star$ is lower semi-continuous.*
4. *u is upper semi-continuous at $x \in X$ if and only if $u(x) = u^\star(x)$.*
5. *u is lower semi-continuous at $x \in X$ if and only if $u(x) = u_\star(x)$.*

Proof. To deduce properties of the lower semi-continuous envelope from the corresponding properties of the upper semi-continuous envelope we will leverage the observation that

$$\begin{aligned} u_\star(x) &= \lim_{r \searrow 0} \inf \big\{ u(y) \mid y \in X \text{ with } \|y - x\| \leq r \big\} \\ &= -\lim_{r \searrow 0} \sup \big\{ -u(y) \mid y \in X \text{ with } \|y - x\| \leq r \big\} = -(-u)^\star(x). \end{aligned} \tag{B.7}$$

1. This is immediate from the definition of the semi-continuous envelopes in (B.5) and (B.6).

2. If v is an upper semi-continuous function with $u \le v$, taking the limsup as y tends to x on both sides of the inequality $u(y) \le v(y)$ and leveraging the upper semi-continuity of v reveals that
$$u^\star(x) = \limsup_{y\to x} u(y) \le \limsup_{y\to x} v(y) \le v(x).$$
This implies that
$$u^\star(x) \le \inf\{v(x) \mid u \le v \text{ and } v \text{ is upper semi-continuous}\}.$$
To show that this infimum is achieved and that this inequality is in fact an equality, it suffices to prove that $u^\star$ is itself upper semi-continuous. Fix $x \in X$ as well as $\varepsilon > 0$, and find $r > 0$ with
$$u^\star(x) + \varepsilon > \sup\big\{u(y) \mid y \in X \text{ with } \|y - x\| \le r\big\}.$$
The triangle inequality reveals that for any $z \in X$ with $\|z - x\| < r$,
$$u^\star(x) + \varepsilon \ge \sup\big\{u(y) \mid y \in X \text{ with } \|y - z\| \le r - \|x - z\|\big\} \ge u^\star(z).$$
It follows that $\limsup_{z\to x} u^\star(z) \le u^\star(x)$ so $u^\star$ is upper semi-continuous at x. Since x is arbitrary, this establishes the claim.

3. Combining the previous part with (B.7) shows that
$$u_\star(x) = -(-u)^\star(x) = \max\big\{-v(x) \mid -u \le v \text{ and } v \text{ is upper semi-continuous}\big\}.$$
Observing that v is upper semi-continuous if and only $-v$ is lower semi-continuous establishes the claim.

4. If u is upper semi-continuous at x, then
$$u^\star(x) = \limsup_{y\to x} u(y) \le u(x).$$
Together with the inequality $u(x) \le u^\star(x)$, this shows that $u(x) = u^\star(x)$. On the other hand, if $u^\star(x) = u(x)$, then
$$\limsup_{y\to x} u(y) = u^\star(x) = u(x) \le u(x)$$
so u is upper semi-continuous at x.

5. Observe that u is lower semi-continuous at $x \in X$ if and only if $-u$ is upper semi-continuous at $x \in X$. The previous part implies that this is the case if and only if $-u(x) = (-u)^\star(x)$. Invoking (B.7) completes the proof. ■

References

[1] L. Ambrosio and J. Feng. On a class of first order Hamilton-Jacobi equations in metric spaces. *Journal of Differential Equations*, 256(7):2194–2245, 2014.

[2] M. Bardi and I. Capuzzo-Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Modern Birkhäuser Classics. Birkhäuser Boston, 2008.

[3] D. Bertsekas. *Convex Optimization Theory*. Athena Scientific optimization and computation series. Athena Scientific, 2009.

[4] P. Cardaliaguet. Notes on mean field games. Technical report, 2010.

[5] P. Cardaliaguet and M. Quincampoix. Deterministic differential games under probability knowledge of initial condition. *International Game Theory Review*, 10(1):1–16, 2008.

[6] P. Cardaliaguet and A. Souquière. A differential game with a blind player. *SIAM Journal on Control and Optimization*, 50(4):2090–2116, 2012.

[7] H.-B. Chen. Hamilton-Jacobi equations for nonsymmetric matrix inference. *arXiv e-prints*, 2020.

[8] H.-B. Chen, J.-C. Mourrat, and J. Xia. Statistical inference of finite-rank tensors. *arXiv e-prints*, 2022.

[9] H.-B. Chen and J. Xia. Fenchel-Moreau identities on convex cones. *arXiv e-prints*, 2020.

[10] H.-B. Chen and J. Xia. Hamilton-Jacobi equations for inference of matrix tensor products. *arXiv e-prints*, 2020.

[11] H.-B. Chen and J. Xia. Limiting free energy of multi-layer generalized linear models. *arXiv e-prints*, 2021.

[12] H.-B. Chen and J. Xia. Hamilton-Jacobi equations from mean-field spin glasses. *arXiv e-prints*, 2022.

[13] H.-B. Chen and J. Xia. Hamilton-Jacobi equations with monotone nonlinearities on convex cones. *arXiv e-prints*, 2022.

[14] A. Coja-Oghlan, F. Krzakala, W. Perkins, and L. Zdeborová. Information-theoretic thresholds from the cavity method. *Advances in Mathematics*, 333:694–795, 2018.

[15] M. G. Crandall and P.-L. Lions. Hamilton-Jacobi equations in infinite dimensions. I. Uniqueness of viscosity solutions. *Journal of Functional Analysis*, 62(3):379–396, 1985.

[16] M. G. Crandall and P.-L. Lions. Hamilton-Jacobi equations in infinite dimensions. II. Existence of viscosity solutions. *Journal of Functional Analysis*, 65(3):368–405, 1986.

[17] M. G. Crandall and P.-L. Lions. Hamilton-Jacobi equations in infinite dimensions. III. *Journal of Functional Analysis*, 68(2):214–247, 1986.

[18] M. G. Crandall and R. Newcomb. Viscosity solutions of Hamilton-Jacobi equations at the boundary. *Proceedings of the American Mathematical Society*, 94(2):283–290, 1985.

[19] T. Dominguez and J.-C. Mourrat. Mutual information for the sparse stochastic block model. *arXiv e-prints*, 2022.

[20] L. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.

[21] J. Feng and M. Katsoulakis. A comparison principle for Hamilton-Jacobi equations related to controlled gradient flows in infinite dimensions. *Archive for Rational Mechanics and Analysis*, 192(2):275–310, 2009.

[22] J. Feng and T. G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.

[23] W. Gangbo, T. Nguyen, and A. Tudorascu. Hamilton-Jacobi equations in the Wasserstein space. *Methods and Applications of Analysis*, 15(2):155–183, 2008.

[24] W. Gangbo and A. Święch. Optimal transport and large number of particles. *Discrete and Continuous Dynamical Systems. Series A*, 34(4):1397–1441, 2014.

[25] J. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of Convex Analysis*. Grundlehren Text Editions. Springer Berlin Heidelberg, 2004.

[26] J.-C. Mourrat. Hamilton–Jacobi equations for finite-rank matrix inference. *The Annals of Applied Probability*, 30(5):2234–2260, 2020.

[27] J.-C. Mourrat. Free energy upper bound for mean-field vector spin glasses. *arXiv e-prints*, 2021.

[28] J.-C. Mourrat. Hamilton–Jacobi equations for mean-field disordered systems. *Annales Henri Lebesgue*, 4:453–484, 2021.

[29] J.-C. Mourrat. Nonconvex interactions in mean-field spin glasses. *Probability and Mathematical Physics*, 2(2):61–119, 2021.

[30] J.-C. Mourrat. The Parisi formula is a Hamilton–Jacobi equation in Wasserstein space. *Canadian Journal of Mathematics*, 74(3):607–629, 2022.

[31] P. E. Souganidis. A remark about viscosity solutions of Hamilton-Jacobi equations at the boundary. *Proceedings of the American Mathematical Society*, 96(2):323–329, 1986.