

A direct verification argument for the Hamilton-Jacobi equation continuum limit of nondominated sorting

Jeff Calder

*Department of Mathematics
University of California, Berkeley
Email: jcalder@berkeley.edu*

Abstract

Nondominated sorting is a combinatorial algorithm that sorts points in Euclidean space into layers according to a partial order. It was recently shown that nondominated sorting of random points has a Hamilton-Jacobi equation continuum limit. The original proof, given in [1], relies on a continuum variational problem. In this paper, we give a new proof using a direct verification argument that completely avoids the variational interpretation. We believe this may be generalized to apply to other stochastic homogenization problems for which there is no obvious underlying variational principle.

Keywords: Continuum limit, Hamilton-Jacobi equation, viscosity solution, stochastic homogenization, nondominated sorting, longest chain problem

2000 MSC: 35D40,

2000 MSC: 60F15,

2000 MSC: 60C05

1. Introduction

Many problems in science and engineering require the sorting, or ordering, of large amounts of multivariate data. Since there is no canonical linear criterion for sorting data in dimensions greater than one, many different methods for sorting have been proposed to address various problems (see, e.g.,

The research described in this paper was partially supported by NSF grant DMS-1500829.

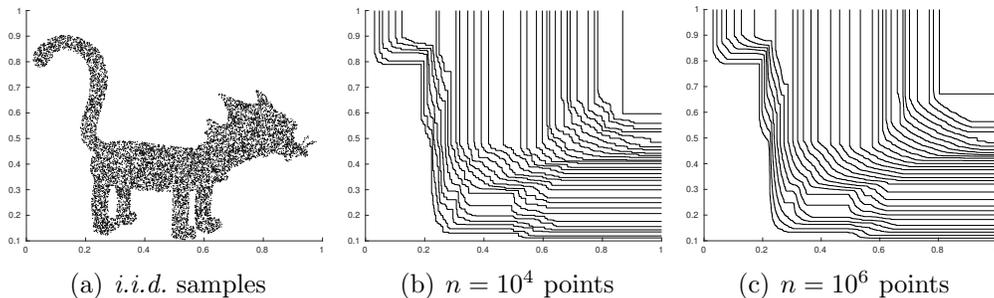


Figure 1: Examples of Pareto fronts corresponding to *i.i.d.* random variables X_1, \dots, X_n drawn from the distribution depicted in (a). In (b) and (c), we show 30 evenly spaced Pareto fronts for $n = 10^4$ and $n = 10^6$, respectively.

[2, 3, 4, 5]). Many of these algorithms abandon the idea of a linear ordering, and instead sort the data into layers according to some set of criteria.

We consider here *nondominated sorting*, which arranges a set of points in Euclidean space into layers by repeatedly removing the set of minimal elements. Let \leq denote the coordinatewise partial order on \mathbb{R}^d defined by

$$x \leq y \iff x_i \leq y_i \text{ for all } i.$$

The *first nondominated layer*, also called the *first Pareto front* and denoted \mathcal{F}_1 , is exactly the set of minimal elements of S with respect to \leq , and the deeper fronts are defined recursively as follows:

$$\mathcal{F}_k = \text{Minimal elements of } S \setminus \bigcup_{i < k} \mathcal{F}_i.$$

This peeling process eventually exhausts the entire set S , and the result is a partition of S based on Pareto front index, which is often called *Pareto depth* or *rank*. Figure 1 gives an illustration of nondominated sorting of a random set S .

Nondominated sorting is widely used in multi-objective optimization, where it is the basis of the popular and effective genetic and evolutionary algorithms [5, 6, 7, 8, 9]. Of course, multi-objective optimization is ubiquitous in engineering and scientific contexts, such as control theory and path planning [10, 11, 12], gene selection and ranking [13, 14, 15, 16, 17, 18, 19], data clustering [20], database systems [21, 22, 23], image processing and computer vision [24, 25], and some recent machine learning problems [23, 26, 27].

Nondominated sorting is also equivalent to the longest chain problem, which has a long history in probability and combinatorics [28, 29, 30, 31]. A *chain* in \mathbb{R}^d is a finite sequence of points that is totally ordered with respect to \leq . Let X_1, \dots, X_n be n distinct points in \mathbb{R}^d and define

$$U_n(x) = \ell(\{X_1, \dots, X_n\} \cap [0, x]), \quad (1.1)$$

where $\ell(\mathcal{O})$ denotes the length of a longest chain in the finite set $\mathcal{O} \subseteq \mathbb{R}^d$. The notation $[0, x]$ is a special case of the more general interval notation

$$[x, z] = \{y \in \mathbb{R}^d : x \leq y \leq z\} = \prod_{i=1}^d [x_i, z_i]$$

that we shall use throughout the paper. Let $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \dots$ denote the Pareto fronts obtained by applying nondominated sorting to $S := \{X_1, \dots, X_n\}$. Then $x \in \mathcal{F}_1$ if and only if there are no other points $y \in S$ with $y \leq x$, i.e., $U_n(x) = 1$. A point $x \in S$ is on the second Pareto front \mathcal{F}_2 if and only if all points $y \in S$ satisfying $y \leq x$ are on the first front, and one such point exists. For any such $y \in \mathcal{F}_1$, $\{y, x\}$ is a chain of length $\ell = 2$ in $S \cap [0, x]$, and we see that $x \in \mathcal{F}_2 \iff U_n(x) = 2$. Peeling off successive Pareto fronts and repeating this argument yields

$$x \in \mathcal{F}_k \iff U_n(x) = k.$$

Hence, the Pareto fronts $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \dots$ are embedded into the level sets (or jump sets) of the longest chain function U_n , as depicted in Figure 1.

In [1], we proved the following continuum limit for nondominated sorting.

Theorem 1. *Let X_1, X_2, X_3, \dots be a sequence of i.i.d. random variables in \mathbb{R}^d with density f , and define U_n by (1.1). Suppose there exists an open and bounded set $\mathcal{O} \subseteq \mathbb{R}_+^d$ with Lipschitz boundary such that f is uniformly continuous on \mathcal{O} , and $f(x) = 0$ for $x \notin \mathcal{O}$. Then $n^{-\frac{1}{d}}U_n \rightarrow u$ uniformly on $[0, \infty)^d$ with probability one, where u is the unique nondecreasing¹ viscosity solution of the Hamilton-Jacobi equation*

$$(P) \quad \begin{cases} u_{x_1} \cdots u_{x_d} = (c_d/d)^d f & \text{in } \mathbb{R}_+^d, \\ u = 0 & \text{on } \partial\mathbb{R}_+^d. \end{cases}$$

and $c_d > 0$ are the universal constants given in [30].

¹We say u is nondecreasing if $x_i \mapsto u(x)$ is nondecreasing for all i .

This result shows that nondominated sorting of massive datasets, which are common in big data applications [32, 33, 34], can be approximated by solving a partial differential equation (PDE). In [35, 36], we exploited this idea to propose a fast approximate nondominated sorting algorithm based on estimating f from the data X_1, \dots, X_n and solving (P) numerically.

The proof of Theorem 1, given in [1], is based on the following continuum variational problem:

$$u(x) = c_d \sup \int_0^1 f(\gamma(t))^{\frac{1}{d}} (\gamma'_1(t) \cdots \gamma'_d(t))^{\frac{1}{d}} dt, \quad (1.2)$$

where the supremum is over all C^1 curves $\gamma : [0, 1] \rightarrow [0, \infty)^d$ that are monotone nondecreasing in all coordinates and satisfy $\gamma(1) = x$. This can be viewed as a type of stochastic homogenization of the longest chain problem, where the curves γ are continuum versions of chains. The variational problem (1.2) first appeared in two dimensions in [31]. The PDE (P) arises as the Hamilton-Jacobi-Bellman equation associated with (1.2) [37]. We used the same ideas to prove a similar Hamilton-Jacobi equation continuum limit for the directed last passage percolation model in statistical physics, which also has an obvious discrete variational formulation [38].

It is possible to view Theorem 1 in the context of stochastic homogenization of Hamilton-Jacobi equations. Indeed, U_n can be interpreted as a discontinuous viscosity solution of the Hamilton-Jacobi equation

$$\left. \begin{aligned} U_{n,x_1} \cdots U_{n,x_d} &= \sum_{i=1}^n \delta_{X_i} && \text{in } \mathbb{R}_+^d, \\ U_n &= 0 && \text{on } \partial\mathbb{R}_+^d. \end{aligned} \right\} \quad (1.3)$$

Since the right hand side of (1.3) is highly singular, it is worth discussing in what sense we interpret U_n to be a solution of (1.3). The obvious approach is to mollify the right hand side to obtain a sequence U_n^ε of approximate solutions. It is possible so show that as $\varepsilon \rightarrow 0$, the sequence U_n^ε converges pointwise to CU_n , where the constant $C > 0$ depends on the mollification kernel used. Intriguing as this observation is, it is unclear whether (1.3) can be used directly to prove Theorem 1.

There is a growing literature on stochastic homogenization of Hamilton-Jacobi equations. Many of the proofs are based on applying subadditive ergodic theory to a representation formula for the solution, like the Hopf-Lax formula, or more generally, a control theoretic interpretation [39, 40, 41,

42, 43, 44, 45, 46, 47]. These techniques are similar in spirit to our original proof of Theorem 1 that was based on the variational representation of u given in (1.2). Recently, there has been significant interest in stochastic homogenization techniques that do not require a variational interpretation [48, 49, 50, 51, 52].

In this paper we give an alternative proof of Theorem 1 that uses a direct verification argument, and completely avoids using the variational interpretation of (P). The argument is based on a heuristic derivation of (P) that originally appeared in [1]. We review this argument, which is reminiscent of dimensional analysis, in Section 2.1. We believe this proof is new in the stochastic homogenization literature, and it seems to be more robust than arguments based on variational principles. We present the proof in the simplest setting in this paper, but we believe it can be substantially generalized. In future work, we plan to apply this proof technique to other stochastic growth models that do not stem from underlying variational principles.

2. Main result

To simplify the presentation, we model the data here using a Poisson point process. Given a nonnegative function $f \in L^1_{loc}(\mathbb{R}^d)$, we denote by Π_f a Poisson point process with intensity function f . This means that Π_f is a random at most countable subset of \mathbb{R}^d , and for every bounded Borel set $A \subseteq \mathbb{R}^d$, the cardinality of $\Pi_f \cap A$, denoted $N(A)$, is a Poisson random variable with mean $\int_A f dx$. Furthermore, for disjoint A and B , the random variables $N(A)$ and $N(B)$ are independent. It is worthwhile to mention that in the special case where $\int_{\mathbb{R}^d} f dx = 1$, one way to construct Π_{nf} for any $n > 0$ is by setting

$$\Pi_{nf} := \{X_1, \dots, X_N\}, \quad (2.1)$$

where X_1, X_2, X_3, \dots is an *i.i.d.* sequence of random variables with density f and N is a Poisson random variable with mean n . For more details on Poisson point processes, we refer the reader to Kingman's book [53].

For nonnegative $f \in L^1_{loc}(\mathbb{R}^d)$ we define

$$U_n(x) = \ell(\Pi_{nf} \cap [0, x]). \quad (2.2)$$

Let us recall Ishii's notion of viscosity solution [54] for the Hamilton-Jacobi equation

$$H(Du) = f \quad \text{in } \mathcal{O}, \quad (2.3)$$

where $f : \mathcal{O} \rightarrow \mathbb{R}$ may be discontinuous. Here, $\mathcal{O} \subseteq \mathbb{R}^d$ is open and H is assumed continuous. We recall the upper semicontinuous envelope of $f : \mathcal{O} \rightarrow \mathbb{R}$ is defined by

$$f^*(x) := \inf_{r>0} \sup_{y \in B(x,r)} f(y),$$

and the lower semicontinuous envelope is given by

$$f_*(x) := \sup_{r>0} \inf_{y \in B(x,r)} f(y).$$

The function f^* is the smallest upper semicontinuous function that is pointwise greater than f , and f_* is the greatest lower semicontinuous function that is less than f .

Definition 1 (Viscosity solution). *We say a continuous function $u : \mathcal{O} \rightarrow \mathbb{R}$ is a viscosity subsolution of (2.3) if for every $x \in \mathcal{O}$ and every $\varphi \in C^\infty(\mathbb{R}^d)$ such that $u - \varphi$ has a local maximum at x*

$$H(D\varphi(x)) \leq f^*(x).$$

Similarly, we say a continuous function $u : \mathcal{O} \rightarrow \mathbb{R}$ is a viscosity supersolution of (2.3) if for every $x \in \mathcal{O}$ and every $\varphi \in C^\infty(\mathbb{R}^d)$ such that $u - \varphi$ has a local minimum at x

$$H(D\varphi(x)) \geq f_*(x).$$

A continuous function $u : \mathcal{O} \rightarrow \mathbb{R}$ is a viscosity solution of (2.3) if u is both a viscosity sub- and supersolution.

Ishii's notion of viscosity solution reduces to the usual notion when f is continuous. In the context of this paper, $f \in L^1_{loc}(\mathbb{R}^d)$ is the intensity function of a Poisson point process, and Ishii's notion of viscosity solution depends on the particular choice of representation of $f \in L^1_{loc}(\mathbb{R}^d)$. It seems natural to seek a refinement of Ishii's notion that is independent of this representation.

Suppose $f : \mathcal{O} \rightarrow \mathbb{R}$ is measurable. We define the *essential upper semicontinuous envelope* of f by

$$f^e(x) := \inf_{r>0} \text{ess sup}_{y \in B(x,r)} f(y),$$

and the *essential lower semicontinuous envelope* by

$$f_e(x) := \sup_{r>0} \text{ess inf}_{y \in B(x,r)} f(y).$$

The function f^e is the smallest upper semicontinuous function that is almost everywhere greater than f , and f_e is the greatest lower semicontinuous function that is almost everywhere less than f .

We propose the following refinement of Ishii's notion of viscosity solution.

Definition 2 (Essential viscosity solution). *We say a continuous function $u : \mathcal{O} \rightarrow \mathbb{R}$ is an essential viscosity subsolution of (2.3) if for every $x \in \mathcal{O}$ and every $\varphi \in C^\infty(\mathbb{R}^d)$ such that $u - \varphi$ has a local maximum at x*

$$H(D\varphi(x)) \leq f^e(x).$$

Similarly, we say a continuous function $u : \mathcal{O} \rightarrow \mathbb{R}$ is an essential viscosity supersolution of (2.3) if for every $x \in \mathcal{O}$ and every $\varphi \in C^\infty(\mathbb{R}^d)$ such that $u - \varphi$ has a local minimum at x

$$H(D\varphi(x)) \geq f_e(x).$$

A continuous function $u : \mathcal{O} \rightarrow \mathbb{R}$ is an essential viscosity solution of (2.3) if u is both an essential viscosity sub- and supersolution.

The notion of essential viscosity solution is clearly independent of the choice of representation of the intensity function f . Since $f^e \leq f^*$ and $f_e \geq f_*$, an essential viscosity sub- (resp. super-) solution is also an Ishii viscosity sub- (resp. super) solution for any choice of representation of f .

In this paper, we give a direct proof of the following result.

Theorem 2. *Let $f \in L_{loc}^\infty(\mathbb{R}^d)$ be nonnegative such that (P) has at most one nondecreasing essential viscosity solution. Then there exists a unique nondecreasing essential viscosity solution $u \in C([0, \infty)^d)$ of (P) and with probability one*

$$n^{-\frac{1}{d}}U_n \longrightarrow u \text{ locally uniformly on } [0, \infty)^d \text{ as } n \rightarrow \infty.$$

Figure 2 gives a visual illustration of Theorem 2.

The proof of Theorem 2 also shows the existence of an essential viscosity solution of (P) for every nonnegative $f \in L_{loc}^\infty(\mathbb{R}^d)$. In [1, 38], we showed that nondecreasing Ishii viscosity solutions of (P) are unique provided either (1) f is uniformly continuous on an open and bounded set $\mathcal{O} \subseteq \mathbb{R}_+^d$ with Lipschitz boundary, and $f(x) = 0$ for $x \notin \mathcal{O}$, or (2) f is continuous on $[0, \infty)^d$. The first case allows f to be discontinuous across the Lipschitz

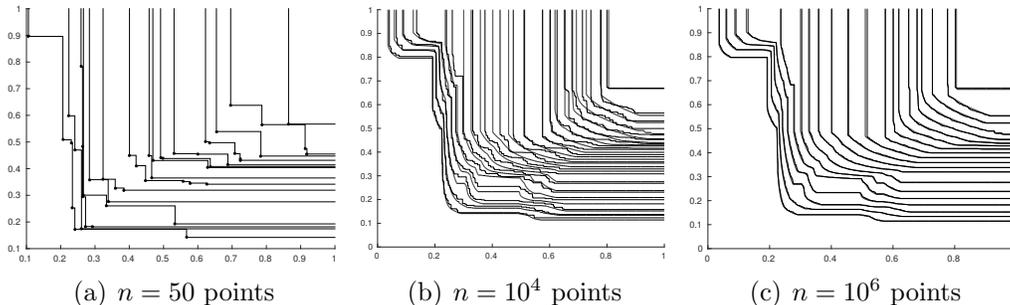


Figure 2: Visualization of the continuum limit given by Theorem 2 applied to the distribution in Figure 1. In (b) and (c), we compare the Pareto fronts to the level sets of the viscosity solution of (P).

boundary of an open set. In either case, Ishii’s notion of viscosity solution agrees with our proposed notion of essential viscosity solution, provided we choose representations of f for which $f^e = f^*$ and $f_e = f_*$. We leave the question of uniqueness of essential viscosity solutions in more general settings to future work.

In the proof of Theorem 2, we assume the family of Poisson point processes $\{\Pi_{nf}\}_{n \in \mathbb{N}}$ is defined on a common probability space and we apply deterministic PDE arguments to the individual realizations

$$U_n^\omega(x) = \ell(\Pi_{nf}^\omega \cap [0, x]).$$

In the literature on the longest chain problem, it is common to take a construction for which $n \mapsto U_n^\omega$ is monotone nondecreasing for all realizations ω . Indeed, this was necessary to obtain almost sure convergence in some early works (see [30, 55]). Due to the sharp concentration of measure results established by Frieze [56], Bollobás and Brightwell [57], and more recently by Talagrand [58], the monotonicity property is no longer required for almost sure convergence. Thus, our proof of Theorem 2 does not assume any particular construction of $\{\Pi_{nf}\}_{n \in \mathbb{N}}$ (see Theorem 3).

We should mention, however, that if we do construct a family of Poisson point processes $\{\Pi_{tf}\}_{t \in \mathbb{R}_+}$ on a common probability space in such a way that

$$\Pi_{sf} \subseteq \Pi_{tf} \quad \text{whenever } s \leq t, \tag{2.4}$$

then the function $U_t(x) = \ell(\Pi_{tf} \cap [0, x])$ is monotone nondecreasing as a function of $t \in \mathbb{R}_+$ for *every* realization. It then follows directly from Theorem 2

that

$$t^{-\frac{1}{d}}U_t \longrightarrow u \text{ locally uniformly on } [0, \infty)^d$$

with probability one as $t \rightarrow \infty$. One way to ensure that (2.4) is satisfied is to construct Π_{tf} via a generalization of (2.1) [53].

2.1. Informal derivation

As motivation, let us give an informal derivation of the Hamilton-Jacobi PDE continuum limit (P). This derivation was originally given in a slightly different form in [1], and our proof of Theorem 2 is based on these heuristics. Suppose $f : \mathbb{R}^d \rightarrow [0, \infty)$ is continuous and that $n^{-\alpha}U_n \longrightarrow u \in C^1(\mathbb{R}^d)$ uniformly with probability one for some $\alpha \in (0, 1]$. The argument below not only derives (P), but also suggests the order of growth $\alpha = \frac{1}{d}$.

Let $x \in (0, \infty)^d$. Since U_n is nondecreasing, the uniform limit u must be nondecreasing and so $u_{x_i}(x) \geq 0$ for all i . Let us assume that $u_{x_i}(x) > 0$ for all i . Fix $v \in \mathbb{R}^d$ with $\langle Du(x), v \rangle > 0$ and consider the quantity $n^\alpha(u(x+v) - u(x))$. This is approximately the number of Pareto fronts passing between x and $x+v$ when n is large. When counting these fronts, we may restrict ourselves to Poisson points falling in the set

$$A = \{y \in \mathbb{R}^d : y \leq x+v \text{ and } u(y) \geq u(x)\}. \quad (2.5)$$

This is because any points in $\{y \in \mathbb{R}^d : u(y) < u(x)\}$ will be on previous Pareto fronts and only points that are coordinatewise less than $x+v$ can influence the Pareto rank of $x+v$. See Figure 3 for a depiction of this region and some quantities from the derivation. Since $u_{x_i}(x) > 0$ for all i , and u is C^1 , A is well approximated for small $|v|$ by a simplex with orthogonal corner, and the points within A are approximately uniformly distributed. Let $N = N(A)$ denote the number of Poisson points falling in A . Since we can scale the simplex into a standard simplex while preserving the Pareto fronts within A , it is reasonable to conjecture that the number of Pareto fronts within A (or the length of a longest chain in A) is asymptotic to $c_1 N^\alpha$ for some constant $c_1 > 0$, independent of x .

By the law of large numbers, we have $N \approx n \int_A f(y) dy$. Hence when $|v| > 0$ is small

$$n^\alpha(u(x+v) - u(x)) \approx c \left(n \int_A f(y) dy \right)^\alpha \approx c_1 n^\alpha |A|^\alpha f(x)^\alpha, \quad (2.6)$$

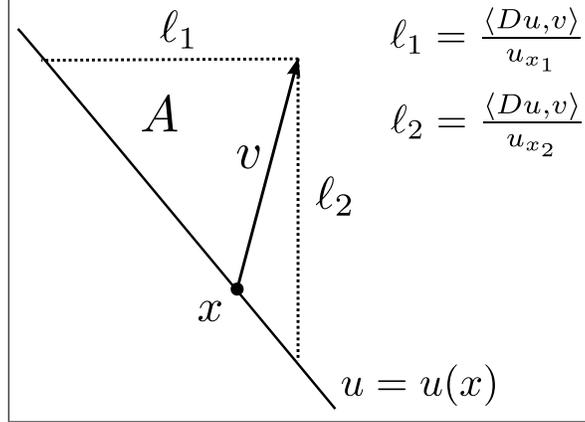


Figure 3: Some quantities from the informal derivation of the Hamilton-Jacobi PDE (P).

where $|A|$ denotes the Lebesgue measure of A . Let ℓ_1, \dots, ℓ_d denote the side lengths of the simplex A . Then $|A| \approx c_2 \ell_1 \cdots \ell_d$ for a constant c_2 . Since $x + v - \ell_i e_i$ lies approximately on the tangent plane to the level set $\{y : u(y) = u(x)\}$, we see that

$$\langle Du(x), v - \ell_i e_i \rangle \approx 0.$$

This gives $\ell_i \approx u_{x_i}(x)^{-1} \langle Du(x), v \rangle$, and hence

$$|A| \approx c_2 (u_{x_1}(x)^{-1} \cdots u_{x_d}(x)^{-1}) \langle Du(x), v \rangle^d. \quad (2.7)$$

For small $|v|$, we can combine (2.6) and (2.7) to obtain

$$\langle Du(x), v \rangle \approx u(x+v) - u(x) \approx c_1 c_2^\alpha f(x)^\alpha (u_{x_1}(x)^{-\alpha} \cdots u_{x_d}(x)^{-\alpha}) \langle Du(x), v \rangle^{\alpha d}.$$

Therefore

$$u_{x_1}(x)^\alpha \cdots u_{x_d}(x)^\alpha \approx C f(x)^\alpha \langle Du(x), v \rangle^{\alpha d - 1},$$

where $C = c_1 c_2^\alpha$. Since the left hand side is independent of v , this suggests that $\alpha = \frac{1}{d}$ and

$$u_{x_1} \cdots u_{x_d} = C^d f \quad \text{on } \mathbb{R}^d.$$

Although this derivation is informal, it is straightforward and conveys the essence of the result. Inspecting the argument, we see that the main ingredients are an invariance with respect to scaling along coordinate axes, and some basic properties of the coordinatewise partial order. It is intriguing

that we can derive the order of growth $n^{\frac{1}{d}}$ and the continuum limit (P) from such basic information. This seems analogous to dimensional analysis arguments in applied mathematics, which use ideas like self-similarity and scaling to derive fundamental laws for natural phenomena [59]. A fundamental feature of dimensional analysis is that, like our argument above, it usually gives no information about the constants of proportionality. Contrary to dimensional analysis, which is typically nonrigorous, we show in this paper that the argument above can be made rigorous using the framework of viscosity solutions.

2.2. Outline of proof

Our new proof of Theorem 2 is a rigorous version of the heuristic dimensional analysis argument given in Section 2.1. The main ideas used in our proof draw some inspiration from the Barles-Souganidis framework for convergence of numerical schemes [60]. The Barles-Souganidis framework establishes convergence of any numerical scheme that is monotone, consistent, and stable, provided the PDE admits a strong comparison principle. In our proof of Theorem 2, we view U_n as a numerical approximation of the viscosity solution of (P), and use ideas that parallel monotonicity, stability and consistency to prove convergence to the viscosity solution of (P).

The longest chain problem is monotone in the following sense: For any finite sets $A, B \subseteq \mathbb{R}^d$,

$$A \subseteq B \implies \ell(A) \leq \ell(B). \tag{2.8}$$

When proving convergence of numerical schemes to viscosity solutions, monotonicity of the scheme is used to replace the numerical solution by a smooth test function for which consistency is trivial. The monotonicity property (2.8) is used for precisely the same purpose in the proof of Theorem 2, though consistency is non-trivial in this setting.

Stability of numerical schemes for PDE usually refers to a uniform bound on the numerical solutions that is independent of the grid resolution. Here, we use a type of asymptotic equicontinuity. Namely, in Theorem 4 we show that a bound of the form

$$|U_n(x) - U_n(y)| \lesssim C|x - y|^{\frac{1}{d}n^{\frac{1}{d}}} \tag{2.9}$$

holds for all x, y with probability one. The stability estimate (2.9) and the monotonicity of U_n can be combined with the Arzelà-Ascoli theorem to show

that for every realization ω in a probability one event Ω , there exists a subsequence $n_k^{-\frac{1}{d}} U_{n_k}^\omega$ converging locally uniformly on $[0, \infty)^d$ to a locally Hölder-continuous function U^ω .

The final step in the proof is to use consistency and monotonicity to identify U^ω as the unique nondecreasing essential viscosity solution of (P). Our consistency result, Theorem 5, is a statement on the asymptotic length of a longest chain in sets resembling the approximate simplex (2.5) used in the heuristic derivation. Namely, for a smooth test function φ with $\varphi_{x_i} > 0$ for all i , we prove asymptotics of the form

$$\ell(\Pi_{nf} \cap A_\varepsilon) \sim \left(\frac{c_d}{d} \left(\frac{f(x)}{\varphi_{x_1}(x) \cdots \varphi_{x_d}(x)} \right)^{\frac{1}{d}} \varepsilon + O(\varepsilon^2) \right) n^{\frac{1}{d}},$$

as $n \rightarrow \infty$, where

$$A_\varepsilon(x) = \left\{ y \in B(x, \sqrt{\varepsilon}) : y \leq x \text{ and } \varphi(y) \geq \varphi(x) - \varepsilon \right\}.$$

Notice the set $A_\varepsilon(x)$ is similar to the set A used in the heuristic derivation and defined in (2.5). The main difference is that $A_\varepsilon(x)$ is defined using a smooth test function, and more importantly, it is defined by looking “backward” from the point x , instead of looking “forward” as in (2.5). This respects the natural domain of dependence of the nondominated sorting problem, and allows us to use the monotonicity of ℓ in the proof of Theorem 2. This is analogous to the idea of an upwind numerical scheme for Hamilton-Jacobi equations.

3. Analysis of longest chain problem

Let us give a brief history of the longest chain problem. Let X_1, \dots, X_n be *i.i.d.* random variables uniformly distributed on the unit hypercube $[0, 1]^d$, and let $\ell_n = \ell(\{X_1, \dots, X_n\})$ be the length of a longest chain. Hammersley [29] was the first to study the asymptotics of ℓ_n , motivated by a connection to Ulam’s famous problem of finding the longest monotone subsequence in a random sequence of numbers [28]. Hammersley showed that in dimension $d = 2$, $n^{-\frac{1}{2}} \ell_n \rightarrow c > 0$ in probability as $n \rightarrow \infty$, and he conjectured that $c = 2$. The value $c = 2$ was later established in [61, 62]. Hammersley’s proof is now a classic application of subadditive ergodic theory.

Bollobás and Winkler [30] extended Hammersley's result to dimensions $d \geq 3$ showing that $n^{-\frac{1}{d}}\ell_n \rightarrow c_d > 0$ in probability. They also established the bounds

$$\frac{d^2}{d!^{\frac{1}{d}}\Gamma\left(\frac{1}{d}\right)} \leq c_d < e,$$

which by Stirling's formula shows that $\lim_{d \rightarrow \infty} c_d = e$. Aside from $c_2 = 2$, the exact values of c_d are still unknown. When the random variables $\{\ell_n\}_{n \in \mathbb{N}}$ are defined on a common probability space in such a way that $n \mapsto \ell_n$ is monotone nondecreasing along sample paths, the stronger almost sure convergence can be obtained via an observation of Kesten [55].

Here, we consider the longest chain among Poisson points. For $t > 0$ let

$$L_t = \ell(\Pi_t \cap [0, 1]^d).$$

Then L_t has the same distribution as

$$\bar{L}_t = \ell(\Pi_1 \cap [0, t^{\frac{1}{d}}]^d).$$

Notice that the family $\{\bar{L}_t\}_{t > 0}$ is defined on a common probability space and $t \mapsto \bar{L}_t$ is monotone nondecreasing along sample paths. Subadditive ergodic theory [29, 63] can be employed to show that

$$t^{-\frac{1}{d}}\bar{L}_t \rightarrow c_d \text{ almost surely as } t \rightarrow \infty.$$

Since L_t and \bar{L}_t merely have the same distribution, this argument only gives convergence in probability for L_t .

Since we have made no assumptions about how the Poisson point process Π_{n_f} is to be constructed, the proof of our main result (Theorem 2) relies on establishing almost sure convergence for sequences of random variables that have the same distribution as L_t , but may not necessarily satisfy the monotonicity condition. Using the concentration of measure results provided by Talagrand's isoperimetric theory [58] we can establish the following result.

Theorem 3. *For $t > 0$, let $L_t = \ell(\Pi_t \cap [0, 1]^d)$. Let $\{t_n\}_{n \in \mathbb{N}}$ be an increasing sequence of positive real numbers such that*

$$\sum_{n=1}^{\infty} \exp\left(-\varepsilon t_n^{\frac{1}{d}}\right) < \infty \text{ for all } \varepsilon > 0. \quad (3.1)$$

Then

$$t_n^{-\frac{1}{d}}L_{t_n} \longrightarrow c_d \text{ completely as } n \rightarrow \infty.$$

Recall that a sequence of random variables $X_1, X_2, \dots, X_n, \dots$ *converges completely* to a random variable X if

$$\sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty \text{ for all } \varepsilon > 0.$$

By the Borel-Cantelli lemma, complete convergence implies almost sure convergence. The advantage of complete convergence is that it depends only on the distributions of X_n and X , and not on any particular construction of the random variables. We believe Theorem 3 is likely well-known in the probability literature, but to the best of our knowledge it has not been published. For the sake of completeness, we have sketched the proof of Theorem 3 in the appendix.

3.1. Stability

In this section, we prove in Theorem 4 a stability result for the longest chain problem. We first have a preliminary lemma.

Lemma 1. *Let $f \in L_{loc}^{\infty}(\mathbb{R}^d)$ be nonnegative, and define U_n by (2.2). Then for any $x, y \in [0, \infty)^d$*

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}}(U_n(x) - U_n(y)) \leq c_d M^{\frac{d-1}{d}} \|f\|_{L^{\infty}([0, M]^d)}^{\frac{1}{d}} |x - y|^{\frac{1}{d}} \quad (3.2)$$

with probability one, where $M = \max\{x_1, y_1, x_2, y_2, \dots, x_d, y_d\}$.

Proof. We first show that for each $h > 0$ and $i \in \{1, \dots, d\}$

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}}(U_n(x + he_i) - U_n(x)) \leq c_d M^{\frac{d-1}{d}} \|f\|_{L^{\infty}([x_i e_i, x + he_i])}^{\frac{1}{d}} h^{\frac{1}{d}}, \quad (3.3)$$

holds with probability one. Without loss of generality we may assume that $i = 1$. If $x_j = 0$ for some $j \geq 2$, then $x, x + he_1 \in \partial\mathbb{R}_+^d$, and hence $U_n(x) = U_n(x + he_1) = 0$ with probability one. Thus, we may assume that $x_j > 0$ for all $j \geq 2$.

Let $L = U_n(x + he_1)$ and let $C = \{X_1, \dots, X_L\}$ be a chain in $\Pi_{n_f} \cap [0, x + he_1]$ of length L . We can split this chain into two chains, C_1 and C_2 , such that C_1 lies in the rectangle $[0, x]$, and C_2 belongs exclusively to $R := [x_1 e_1, x + he_1]$. See Figure 4 for a depiction of the setup. By the definition of U_n

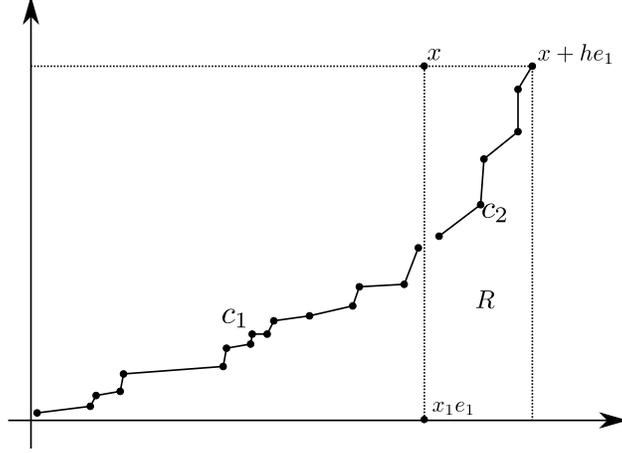


Figure 4: A depiction of the chains C_1 and C_2 from the proof of Lemma 1.

$$U_n(x + he_1) = \ell(C_1) + \ell(C_2) \leq U_n(x) + \ell(\Pi_{nf} \cap R). \quad (3.4)$$

Let $K = \text{ess sup}_R f$ and set

$$g(x) = \begin{cases} K - f(x), & \text{if } x \in R \\ 0, & \text{otherwise.} \end{cases}$$

Let Π_{ng} be a Poisson point process with intensity ng and let $\Pi = \Pi_{nf} \cup \Pi_{ng}$. Then Π is a Poisson point process with intensity λ , where

$$\lambda(x) = nf(x) + ng(x) = \begin{cases} nK, & \text{if } x \in R \\ nf(x), & \text{otherwise.} \end{cases}$$

Let $N = hx_2x_3 \cdots x_dKn$. By scaling, $\ell(\Pi \cap R)$ has the same distribution as $\ell(\Pi_N \cap [0, 1]^d)$. By Theorem 3 and the inclusion $\Pi_{nf} \subseteq \Pi$ we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{nf} \cap R) &\leq \lim_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi \cap R) \\ &= c_d(x_2 \cdots x_d)^{\frac{1}{d}} K^{\frac{1}{d}} h^{\frac{1}{d}} \\ &\leq c_d M^{\frac{d-1}{d}} K^{\frac{1}{d}} h^{\frac{1}{d}} \end{aligned}$$

with probability one. Combining this with (3.4) establishes (3.3).

We can apply (3.3) in each coordinate and note that

$$\sum |x_i - y_i|^{\frac{1}{d}} \leq d^{1-\frac{1}{2d}} |x - y|^{\frac{1}{d}}$$

to establish the desired result (3.2). \square

Theorem 4 (Stability). *Let $f \in L_{loc}^\infty(\mathbb{R}^d)$ be nonnegative, and define U_n by (2.2). Then with probability one*

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}}(U_n(x) - U_n(y)) \leq c_d d M^{\frac{d-1}{d}} \|f^e\|_{L^\infty([0, M]^d)}^{\frac{1}{d}} |x - y|^{\frac{1}{d}} \quad (3.5)$$

for all $x, y \in [0, \infty)^d$, where $M = \max\{x_1, y_1, x_2, y_2, \dots, x_d, y_d\}$.

It is important to emphasize in Theorem 4 that (3.5) holds with probability one *simultaneously* for all $x, y \in [0, \infty)^d$. This is a far stronger statement than Lemma 1, where it was shown that for fixed $x, y \in [0, \infty)^d$, (3.5) holds with probability one. This stronger result gives us a form of compactness that is used in the proof of Theorem 2.

Proof. For $x, y \in [0, \infty)^d$ let $\Omega_{x,y}$ denote the event that (3.2) holds for x, y . By Lemma 1, $\mathbb{P}(\Omega_{x,y}) = 1$. Let

$$\Omega = \bigcap \left\{ \Omega_{x,y} : x, y \in \mathbb{Q}^d \cap [0, \infty)^d \right\}.$$

Being the countable intersection of probability one events, Ω has probability one.

Let $x, y \in [0, \infty)^d$. Let $\hat{x}, \hat{y} \in \mathbb{Q}^d \cap [0, \infty)^d$ such that $x \leq \hat{x}$ and $\hat{y} \leq y$, and set

$$\widehat{M} = \max\{\hat{x}_1, \hat{y}_1, \hat{x}_2, \hat{y}_2, \dots, \hat{x}_d, \hat{y}_d\}.$$

Then since U_n is nondecreasing

$$U_n(x) - U_n(y) \leq U_n(\hat{x}) - U_n(\hat{y}),$$

and therefore

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}}(U_n^\omega(x) - U_n^\omega(y)) \leq c_d d \widehat{M}^{\frac{d-1}{d}} \|f\|_{L^\infty([0, \widehat{M}]^d)}^{\frac{1}{d}} |\hat{x} - \hat{y}|^{\frac{1}{d}} \quad \text{for all } \omega \in \Omega.$$

Since we can choose \hat{x} and \hat{y} arbitrarily close to x and y , respectively, we obtain

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}}(U_n^\omega(x) - U_n^\omega(y)) \leq c_d d M^{\frac{d-1}{d}} \|f^e\|_{L^\infty([0, M]^d)}^{\frac{1}{d}} |x - y|^{\frac{1}{d}} \quad \text{for all } \omega \in \Omega,$$

which completes the proof. \square

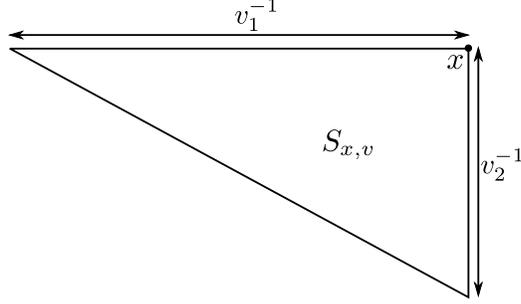


Figure 5: Illustration of the simplex $S_{x,v}$.

3.2. Consistency

We now establish our main consistency result. For $x, v \in (0, \infty)^d$, we define

$$S_{x,v} = \left\{ z \in \mathbb{R}^d : z \leq x \text{ and } \langle x - z, v \rangle \leq 1 \right\}. \quad (3.6)$$

The set $S_{x,v}$ is a simplex with orthogonal corner at x and side lengths $1/v_1, \dots, 1/v_d$. See Figure 5 for an illustration of $S_{x,v}$. We also define $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^d$.

Lemma 2. *Let $x, v \in (0, \infty)^d$ and let $\{t_n\}_{n \in \mathbb{N}}$ be any increasing sequence of positive real numbers satisfying (3.1). Then*

$$\lim_{n \rightarrow \infty} t_n^{-\frac{1}{d}} \ell(\Pi_{t_n} \cap S_{x,v}) = \frac{c_d}{d(v_1 \cdots v_d)^{\frac{1}{d}}} \quad (3.7)$$

in the sense of complete convergence.

Proof. We first show that

$$\lim_{n \rightarrow \infty} t_n^{-\frac{1}{d}} \ell(\Pi_{t_n} \cap S_{\mathbf{1},\mathbf{1}}) = \frac{c_d}{d} \text{ completely.} \quad (3.8)$$

Note that we can write

$$S_{\mathbf{1},\mathbf{1}} = \left\{ x \in [0, 1]^d : x_1 + \cdots + x_d \geq d - 1 \right\}.$$

Let K be a positive integer and define

$$I = \left\{ x \in [0, 1]^d : Kx \in \mathbb{Z}^d \text{ and } x_1 + \cdots + x_d = d - 1 - \frac{d}{K} \right\}.$$

It follows from this construction that *every* chain in $S_{\mathbf{1},\mathbf{1}}$ is contained entirely in $[x, \mathbf{1}]$ for some $x \in I$. Since I is a finite set

$$\ell(\Pi_t \cap S_{\mathbf{1},\mathbf{1}}) \leq \max \left\{ \ell(\Pi_t \cap [x, \mathbf{1}]) : x \in I \right\}. \quad (3.9)$$

Fix $x \in I$ and let $N_n = t_n \prod_{i=1}^d (1 - x_i)$. Since $\ell(\Pi_{t_n} \cap [x, \mathbf{1}])$ has the same distribution as $\ell(\Pi_{N_n} \cap [0, 1]^d)$ we have by Theorem 3 that

$$\lim_{n \rightarrow \infty} t_n^{-\frac{1}{d}} \ell(\Pi_{t_n} \cap [x, \mathbf{1}]) = c_d \prod_{i=1}^d (1 - x_i)^{\frac{1}{d}} \leq \frac{c_d}{d} \sum_{i=1}^d (1 - x_i) = c_d \left(\frac{1}{d} + \frac{1}{K} \right),$$

in the sense of complete convergence. Combining this with (3.9) and noting that K was arbitrary yields

$$\limsup_{n \rightarrow \infty} t_n^{-\frac{1}{d}} \ell(\Pi_{t_n} \cap S_{\mathbf{1},\mathbf{1}}) \leq \frac{c_d}{d} \quad (3.10)$$

completely.

For the other direction, let $z = \mathbf{1}(d - 1)/d$. By Theorem 3

$$\liminf_{n \rightarrow \infty} t_n^{-\frac{1}{d}} \ell(\Pi_{t_n} \cap S_{\mathbf{1},\mathbf{1}}) \geq \lim_{n \rightarrow \infty} t_n^{-\frac{1}{d}} \ell(\Pi_{t_n} \cap [z, \mathbf{1}]) = c_d \prod_{i=1}^d (1 - z_i)^{\frac{1}{d}} = \frac{c_d}{d}.$$

in the sense of complete convergence. This establishes (3.8).

We now establish (3.7) with a simple scaling argument. Let $x, v \in (0, \infty)^d$ and define $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ by

$$\Phi(y) = ((y_1 - x_1)v_1 + 1, (y_2 - x_2)v_2 + 1, \dots, (y_d - x_d)v_d + 1).$$

Since $v_i > 0$ for all i , the mapping Φ preserves the partial order \leq and therefore

$$\ell(\Pi_{t_n} \cap S_{x,v}) = \ell(\Phi(\Pi_{t_n} \cap S_{x,v})). \quad (3.11)$$

Let Π be the Poisson point process induced by the mapping Φ , i.e., $\Pi = \Phi(\Pi_n)$. Then Π is a Poisson point process on \mathbb{R}^d with constant intensity $N_n = t_n/(v_1 \cdots v_d)$. Since $\Phi(S_{x,v}) = S_{\mathbf{1},\mathbf{1}}$, we have by (3.11) that $\ell(\Pi_{t_n} \cap S_{x,v})$ and $\ell(\Pi_{N_n} \cap S_{\mathbf{1},\mathbf{1}})$ have the same distribution. The result now follows directly from (3.8). \square

Lemma 3. *Let $f \in L_{loc}^\infty(\mathbb{R}^d)$ be nonnegative. Then for any $x, v \in (0, \infty)^d$*

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{nf} \cap S_{x,v}) \leq \frac{c_d}{d} \left(\frac{\text{ess sup}_{S_{x,v}} f}{v_1 \cdots v_d} \right)^{\frac{1}{d}}, \quad (3.12)$$

and

$$\liminf_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{nf} \cap S_{x,v}) \geq \frac{c_d}{d} \left(\frac{\text{ess inf}_{S_{x,v}} f}{v_1 \cdots v_d} \right)^{\frac{1}{d}} \quad (3.13)$$

with probability one.

Proof. Let $x, v \in (0, \infty)^d$, $K = \text{ess sup}_{S_{x,v}} f$, and define

$$g(x) = \begin{cases} K - f(x), & \text{if } x \in S_{x,v}, \\ 0, & \text{otherwise.} \end{cases}$$

Let Π_{ng} be a Poisson point process with intensity ng and let $\Pi = \Pi_{nf} \cup \Pi_{ng}$. Then Π is a Poisson point process with intensity λ , where

$$\lambda(x) = nf(x) + ng(x) = \begin{cases} nK, & \text{if } x \in S_{x,v} \\ nf(x), & \text{otherwise.} \end{cases}$$

By Lemma 2

$$\lim_{n \rightarrow \infty} (Kn)^{-\frac{1}{d}} \ell(\Pi \cap S_{x,v}) = \frac{c_d}{d(v_1 \cdots v_d)^{\frac{1}{d}}}$$

with probability one, and hence

$$\lim_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi \cap S_{x,v}) = \frac{c_d}{d} \left(\frac{\text{ess sup}_{S_{x,v}} f}{v_1 \cdots v_d} \right)^{\frac{1}{d}}$$

with probability one. Since $\Pi_{nf} \subseteq \Pi$, $\ell(\Pi_{nf} \cap S_{x,v}) \leq \ell(\Pi \cap S_{x,v})$, which establishes (3.12).

To prove (3.13), note first that if $\text{ess inf}_{S_{x,v}} f = 0$, then the result is trivial. Hence, we may take $m := \text{ess inf}_{S_{x,v}} f > 0$. For each point X from Π_{nf} that falls in $S_{x,v}$, color X red with probability $m/f(X)$. The red points form another Poisson point process on $S_{x,v}$ with intensity mn [53]. Let us denote this Poisson point process by Π^r . By Lemma 2, we have that

$$\lim_{n \rightarrow \infty} (mn)^{-\frac{1}{d}} \ell(\Pi^r \cap S_{x,v}) = \frac{c_d}{d(v_1 \cdots v_d)^{\frac{1}{d}}} \quad \text{with probability one.} \quad (3.14)$$

Since $\Pi^r \subseteq \Pi_{nf}$, $\ell(\Pi^r \cap S_{x,v}) \leq \ell(\Pi_{nf} \cap S_{x,v})$. Combining this with (3.14) completes the proof. \square

Lemma 4. *Let $f \in L_{loc}^\infty(\mathbb{R}^d)$ be nonnegative. Then with probability one*

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{nf} \cap S_{x,v}) \leq \frac{c_d}{d} \left(\frac{\sup_{S_{x,v}} f^e}{v_1 \cdots v_d} \right)^{\frac{1}{d}}, \quad (3.15)$$

and

$$\liminf_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{nf} \cap S_{x,v}) \geq \frac{c_d}{d} \left(\frac{\inf_{S_{x,v}} f_e}{v_1 \cdots v_d} \right)^{\frac{1}{d}}, \quad (3.16)$$

for all $x, v \in (0, \infty)^d$.

Proof. Let us prove (3.15); the proof of (3.16) is similar. For each $x, v \in (0, \infty)^d$ let $\Omega_{x,v}$ denote the event that (3.12) holds for x, v . By Lemma 3, $\mathbb{P}(\Omega_{x,v}) = 1$. Let us set

$$\Omega = \bigcap \left\{ \Omega_{x,v} : x, v \in \mathbb{Q}^d \cap (0, \infty)^d \right\}.$$

Being the countable intersection of probability one events, Ω has probability one.

Let $x, v \in (0, \infty)^d$ and let $\omega \in \Omega$. Let $q \in (0, \infty)^d \cap \mathbb{Q}^d$ such that $q_i < v_i$ for all i , and let $y \in (0, \infty)^d \cap \mathbb{Q}^d$ such that $y \geq x$ and

$$\frac{1}{q_j} \geq \frac{1 + \langle y - x, v \rangle}{v_j}. \quad (3.17)$$

We claim that $S_{x,v} \subseteq S_{y,q}$. To see this, let $z \in S_{x,v}$ and write

$$\langle y - z, q \rangle \stackrel{(3.17)}{\leq} \frac{\langle y - z, v \rangle}{1 + \langle y - x, v \rangle} = \frac{\langle x - z, v \rangle + \langle y - x, v \rangle}{1 + \langle y - x, v \rangle} \leq 1,$$

where we used in the last step that $\langle x - z, v \rangle \leq 1$. Therefore $z \in S_{y,q}$. It follows that $\ell(\Pi_{nf} \cap S_{x,v}) \leq \ell(\Pi_{nf} \cap S_{y,q})$, and from (3.12) we deduce

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{nf}^\omega \cap S_{x,v}) \leq \frac{c_d}{d} \left(\frac{\text{ess sup}_{S_{y,q}} f}{q_1 \cdots q_d} \right)^{\frac{1}{d}}. \quad (3.18)$$

Sending $y \rightarrow x$ in such a way that $y \geq x$ and $y \in \mathbb{Q}^d \cap (0, \infty)^d$ yields

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{nf}^\omega \cap S_{x,v}) \leq \frac{c_d}{d} \left(\frac{\sup_{S_{x,q}} f^e}{q_1 \cdots q_d} \right)^{\frac{1}{d}}.$$

The result follows by noting that the limit above holds for all $q \in (0, \infty)^d \cap \mathbb{Q}^d$ with $q_i < v_i$ for all i . \square

We denote by $B(x, r)$ the open ball of radius $r > 0$ centered at $x \in \mathbb{R}^d$. We now prove our main consistency result.

Theorem 5 (Consistency). *Let $f \in L_{loc}^\infty(\mathbb{R}^d)$ be nonnegative. Then with probability one*

$$\limsup_{\varepsilon \rightarrow 0^+} \limsup_{n \rightarrow \infty} \varepsilon^{-1} n^{-\frac{1}{d}} \ell(\Pi_{nf} \cap A_\varepsilon(x_0)) \leq \frac{c_d}{d} \left(\frac{f^e(x_0)}{\varphi_{x_1}(x_0) \cdots \varphi_{x_d}(x_0)} \right)^{\frac{1}{d}}, \quad (3.19)$$

and

$$\liminf_{\varepsilon \rightarrow 0^+} \liminf_{n \rightarrow \infty} \varepsilon^{-1} n^{-\frac{1}{d}} \ell(\Pi_{nf} \cap A_\varepsilon(x_0)) \geq \frac{c_d}{d} \left(\frac{f_e(x_0)}{\varphi_{x_1}(x_0) \cdots \varphi_{x_d}(x_0)} \right)^{\frac{1}{d}}, \quad (3.20)$$

for all $x_0 \in (0, \infty)^d$ and all $\varphi \in C^2(\mathbb{R}^d)$ such that $\varphi_{x_i}(x_0) > 0$ for all i , where

$$A_\varepsilon(x_0) := \left\{ x \in B(x_0, \sqrt{\varepsilon}) : x \leq x_0 \text{ and } \varphi(x) \geq \varphi(x_0) - \varepsilon \right\}. \quad (3.21)$$

Proof. Let Ω be the event that (3.15) and (3.16) hold for all $x, v \in (0, \infty)^d$. By Lemma 4, Ω has probability one. Let $x_0 \in (0, \infty)^d$, $\varepsilon > 0$ and let $\varphi \in C^2(\Omega)$ such that $\varphi_{x_i}(x_0) > 0$ for all i . Fix $\omega \in \Omega$.

We will give the proof of (3.19); the proof of (3.20) is very similar. Since $\varphi_{x_i}(x_0) > 0$, there exists $m > 0$ such that $\varphi_{x_i}(x) \geq m$ for all $x \in A_\varepsilon$, $i \in \{1, \dots, d\}$, and all $\varepsilon > 0$ sufficiently small. It follows that for any $x \in A_\varepsilon$

$$\begin{aligned} \varphi(x_0) - \varphi(x) &= \int_0^1 \langle D\varphi(x + t(x_0 - x)), x_0 - x \rangle dt \geq m \sum_{i=1}^d (x_{0,i} - x_i) \\ &\geq \frac{m|x_0 - x|}{\sqrt{d}}. \end{aligned} \quad (3.22)$$

We therefore deduce

$$|x_0 - x| \leq \frac{\sqrt{d}\varepsilon}{m} \text{ for all } x \in A_\varepsilon. \quad (3.23)$$

Since $\varphi \in C^2$, there exists a constant $C > 0$ such that

$$\varphi(x) - \varphi(x_0) \leq \langle D\varphi(x_0), x - x_0 \rangle + C\varepsilon^2 \quad (3.24)$$

for all $x \in A_\varepsilon$ and $\varepsilon > 0$ sufficiently small. It follows that

$$A_\varepsilon(x_0) \subseteq \left\{ x \in \mathbb{R}^d : x \leq x_0 \text{ and } \langle D\varphi(x_0), x_0 - x \rangle \leq \varepsilon + C\varepsilon^2 \right\}.$$

Setting $q_\varepsilon = D\varphi(x_0)/(\varepsilon + C\varepsilon^2)$ we have $A_\varepsilon(x_0) \subseteq S_{x_0, q_\varepsilon}$, and hence

$$\ell(\Pi_{n_f}^\omega \cap A_\varepsilon(x_0)) \leq \ell(\Pi_{n_f}^\omega \cap S_{x_0, q_\varepsilon}).$$

By (3.15)

$$\limsup_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{n_f}^\omega \cap A_\varepsilon(x_0)) \leq \frac{c_d}{d} \left(\frac{\sup_{S_{x_0, q_\varepsilon}} f^e}{\varphi_{x_1}(x_0) \cdots \varphi_{x_d}(x_0)} \right)^{\frac{1}{d}} (\varepsilon + C\varepsilon^2).$$

Sending $\varepsilon \rightarrow 0^+$ completes the proof. \square

4. Proof of main result

We now have the proof of Theorem 2.

Proof. Let Ω denote the event that the conclusions of Theorem 4 and Theorem 5 hold, and $U_n \equiv 0$ on $\partial\mathbb{R}_+^d$ for all n . Then $\mathbb{P}(\Omega) = 1$. Let us fix a realization $\omega \in \Omega$. The remainder of the proof is split into several steps.

1. We first use stability (Theorem 4) to obtain a compactness result. Since $U_n^\omega(0) = 0$ for all n , it follows from Theorem 4 that the sequence $\{n^{-\frac{1}{d}}U_n^\omega(x)\}_{n \in \mathbb{N}}$ is bounded for all $x \in [0, \infty)^d$. By a diagonal argument, there exists a subsequence $\{U_{n_k}^\omega\}_{k \in \mathbb{N}}$ such that for all $x \in [0, \infty)^d \cap \mathbb{Q}^d$, $\{n_k^{-\frac{1}{d}}U_{n_k}^\omega(x)\}_{k \in \mathbb{N}}$ is a convergent sequence, whose limit we denote by $U^\omega(x)$. By Theorem 4

$$\begin{aligned} |U^\omega(x) - U^\omega(y)| &= \lim_{n_k \rightarrow \infty} n_k^{-\frac{1}{d}} |U_{n_k}^\omega(x) - U_{n_k}^\omega(y)| \\ &\leq c_d d M^{\frac{d-1}{d}} \|f^e\|_{L^\infty([0, M]^d)}^{\frac{1}{d}} |x - y|^{\frac{1}{d}} \end{aligned} \quad (4.1)$$

for all $x, y \in \mathbb{Q}^d \cap [0, \infty)^d$, where $M = \max\{x_1, y_1, x_2, y_2, \dots, x_d, y_d\}$. Hence, we can extend U^ω uniquely to a function $U^\omega \in C([0, \infty)^d)$ such that for every $M > 0$, $U^\omega \in C^{0, \frac{1}{d}}([0, M]^d)$ and

$$[U^\omega]_{\frac{1}{d}; [0, M]^d} \leq c_d d M^{\frac{d-1}{d}} \|f^e\|_{L^\infty([0, M]^d)}^{\frac{1}{d}}. \quad (4.2)$$

Furthermore, U^ω is nondecreasing and $U^\omega \equiv 0$ on $\partial\mathbb{R}_+^d$.

We claim that $n_k^{-\frac{1}{d}}U_{n_k}^\omega \rightarrow U^\omega$ locally uniformly on $[0, \infty)^d$. The proof of this is similar to [1, Theorem 1]. We include it here for completeness. Fix $M > 0$ and let $\varepsilon > 0$. Let $T \in \mathbb{N}$ and for any multi-index $\alpha \in \mathbb{Z}^d$, let $x_\alpha = \alpha/T$. Let I be the set of multi-indices $\alpha \in \mathbb{Z}^d$ for which $x_\alpha \in [0, M]^d$. Since U^ω is continuous on $[0, M]^d$, we can choose T large enough so that

$$U^\omega(x_{\alpha+1}) - U^\omega(x_\alpha) < \varepsilon \quad \text{for all } \alpha \in I. \quad (4.3)$$

Since I is a finite set and $x_\alpha \in \mathbb{Q}^d$ for all $\alpha \in I$, we deduce

$$\lim_{k \rightarrow \infty} \max_{\alpha \in I} \left| n_k^{-\frac{1}{d}}U_{n_k}^\omega(x_\alpha) - U^\omega(x_\alpha) \right| = 0. \quad (4.4)$$

Let $y \in [0, M]^d$ and let $\alpha \in I$ such that $x_\alpha \leq y \leq x_{\alpha+1}$. Since U_n^ω and U^ω are nondecreasing we have

$$n^{-\frac{1}{d}}U_n^\omega(y) - U^\omega(y) \leq n^{-\frac{1}{d}}U_n^\omega(x_\beta) - U^\omega(x_{\beta-1}) \stackrel{(4.3)}{<} n^{-\frac{1}{d}}U_n^\omega(x_\beta) - U^\omega(x_\beta) + \varepsilon,$$

where $\beta = \alpha + \mathbf{1}$. Similarly, we deduce

$$n^{-\frac{1}{d}}U_n^\omega(y) - U^\omega(y) \geq n^{-\frac{1}{d}}U_n^\omega(x_\alpha) - U^\omega(x_{\alpha+1}) \stackrel{(4.3)}{>} n^{-\frac{1}{d}}U_n^\omega(x_\alpha) - U^\omega(x_\alpha) - \varepsilon.$$

Combining these inequalities yields

$$\|n^{-\frac{1}{d}}U_n^\omega - U^\omega\|_{L^\infty([0, M]^d)} < \max_{\alpha \in I} \left| n^{-\frac{1}{d}}U_n^\omega(x_\alpha) - U^\omega(x_\alpha) \right| + \varepsilon. \quad (4.5)$$

Invoking (4.4) we see that

$$\limsup_{k \rightarrow \infty} \|n_k^{-\frac{1}{d}}U_{n_k}^\omega - U^\omega\|_{L^\infty([0, M]^d)} < \varepsilon.$$

Sending $\varepsilon \rightarrow 0$ establishes the claim.

2. We now show that U^ω is an essential viscosity subsolution of (P). For simplicity, let us set

$$V_k = n_k^{-\frac{1}{d}}U_{n_k}^\omega \quad \text{and} \quad V = U^\omega.$$

Fix $M > 0$ and let $x_0 \in (0, M)^d$. Let $\varphi \in C^2(\mathbb{R}^d)$ such that $V - \varphi$ has a local maximum at x_0 . Since V is nondecreasing, $\varphi_{x_i}(x_0) \geq 0$ for all i . If

$\varphi_{x_i}(x_0) = 0$ for some i , then the subsolution property is trivially satisfied. Hence, we may assume that $\varphi_{x_i}(x_0) > 0$ for all i . Without loss of generality, we may also assume that $V - \varphi$ has a strict maximum at x_0 , relative to the set $[0, M]^d$. Then there exists a sequence $\{x_k\}_{k \in \mathbb{N}}$ in $[0, M]^d$ converging to x_0 such that $V_k - \varphi$ has a maximum at x_k , relative to $[0, M]^d$.

Let $r, m > 0$ such that $B(x_0, r) \subseteq (0, M)^d$ and $\varphi_{x_i}(x) > m$ for all $x \in B(x_0, r)$ and all i . Let $\varepsilon \in (0, 1)$ such that

$$\varepsilon < \min \left\{ \frac{m^2}{4d}, r^2 \right\}. \quad (4.6)$$

Since $V_k \rightarrow V$ uniformly on $[0, M]^d$ and $x_k \rightarrow x_0$, there exists $K > 0$ such that for all $k > K$

$$\varphi(x_0) - \varphi(x_k) + V_k(x_k) - V_k(x_0) < \varepsilon^2. \quad (4.7)$$

Let

$$\mathcal{A}_{k,\varepsilon} = \{z \in [0, x_0] : V_k(z) \geq V_k(x_0) - \varepsilon\}. \quad (4.8)$$

We claim that $\mathcal{A}_{k,\varepsilon} \subseteq B(x_0, \sqrt{\varepsilon})$ for all $k > K$. To see this, let $z \in B(x_0, r)$ such that $z \leq x_0$ and note that

$$\begin{aligned} \varphi(x_0) - \varphi(z) &= \int_0^1 \langle D\varphi(z + t(x_0 - z)), x_0 - z \rangle dt \geq m \sum_{i=1}^d (x_{0,i} - z_i) \\ &\geq \frac{m|x_0 - z|}{\sqrt{d}}. \end{aligned} \quad (4.9)$$

Since $V_k - \varphi$ has a maximum at x_k ,

$$V_k(z) - \varphi(z) \leq V_k(x_k) - \varphi(x_k) \quad \text{for } z \in [0, M]^d. \quad (4.10)$$

Combining this with (4.6), (4.7) and (4.9) we deduce

$$\begin{aligned} V_k(z) &\leq V_k(x_k) + \varphi(z) - \varphi(x_k) \\ &= V_k(x_0) + \varphi(z) - \varphi(x_0) + \varphi(x_0) - \varphi(x_k) + V_k(x_k) - V_k(x_0) \\ &< V_k(x_0) - \frac{m}{\sqrt{d}}\sqrt{\varepsilon} + \varepsilon^2 \\ &< V_k(x_0) - \varepsilon \end{aligned}$$

for all $z \in \partial B(x_0, \sqrt{\varepsilon})$ with $z \leq x_0$, and $k > K$. Since V_k is nondecreasing, $\mathcal{A}_{k,\varepsilon} \subseteq B(x_0, \sqrt{\varepsilon})$, establishing the claim.

We now claim that

$$\varepsilon n_k^{\frac{1}{d}} \leq \ell(\Pi_{n_k f}^\omega \cap \mathcal{A}_{k,\varepsilon}). \quad (4.11)$$

To see this, since U_n is integer valued, we can write $\mathcal{A}_{k,\varepsilon}$ as

$$\mathcal{A}_{k,\varepsilon} = \{z \in [0, x_0] : U_{n_k}^\omega(z) \geq U_{n_k}^\omega(x_0) - \lfloor \varepsilon n_k^{\frac{1}{d}} \rfloor\}, \quad (4.12)$$

where $\lfloor t \rfloor$ is the largest integer less than or equal to $t \in \mathbb{R}$. Since $\mathcal{A}_{k,\varepsilon} \subseteq B(x_0, r) \subseteq (0, \infty)^d$ and $U_{n_k}^\omega(y) = 0$ for $y \in \partial \mathbb{R}_+^d$, we must have that $U_{n_k}^\omega(x_0) - \lfloor \varepsilon n_k^{\frac{1}{d}} \rfloor \geq 1$. Let $L = U_{n_k}^\omega(x_0)$ and let $X_1^\omega \leq \dots \leq X_L^\omega$ be a chain in $\Pi_{n_k f}^\omega \cap [0, x_0]$ of length L . For any $q \geq L - \lfloor \varepsilon n_k^{\frac{1}{d}} \rfloor$, $U_{n_k}^\omega(X_q^\omega) \geq q \geq L - \lfloor \varepsilon n_k^{\frac{1}{d}} \rfloor$, and hence $X_q^\omega \in \mathcal{A}_{k,\varepsilon}$. Hence the chain $X_q^\omega, \dots, X_L^\omega$ belongs to $\mathcal{A}_{k,\varepsilon}$ for $q = L - \lfloor \varepsilon n_k^{\frac{1}{d}} \rfloor$ and

$$\varepsilon n_k^{\frac{1}{d}} \leq \lfloor \varepsilon n_k^{\frac{1}{d}} \rfloor + 1 = L - q + 1 \leq \ell(\Pi_{n_k f}^\omega \cap \mathcal{A}_{k,\varepsilon}),$$

which establishes the claim.

By (4.7) and (4.10)

$$V_k(z) - V_k(x_0) \leq \varphi(z) - \varphi(x_0) + \varepsilon^2$$

for all $k > K$ and $z \in [0, M]^d$. Since $\mathcal{A}_{k,\varepsilon} \subseteq B(x_0, \sqrt{\varepsilon})$ it follows that

$$\mathcal{A}_{k,\varepsilon} \subseteq \{z \in B(x_0, \sqrt{\varepsilon}) : x \leq x_0 \text{ and } \varphi(z) \geq \varphi(x_0) - \varepsilon - \varepsilon^2\} =: A_{\varepsilon+\varepsilon^2}(x_0) \quad (4.13)$$

for $k > K$, where $A_{\varepsilon+\varepsilon^2}(x_0)$ is defined as in Theorem 5. Invoking (4.11) and the monotonicity of ℓ we have

$$\varepsilon \leq \limsup_{k \rightarrow \infty} n_k^{-\frac{1}{d}} \ell(\Pi_{n_k f}^\omega \cap \mathcal{A}_{k,\varepsilon}) \leq \limsup_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{n f}^\omega \cap A_{\varepsilon+\varepsilon^2}(x_0)).$$

Since Theorem 5 holds for all $\omega \in \Omega$ and $\varphi_{x_i}(x_0) > 0$, we deduce

$$\begin{aligned} 1 &\leq \limsup_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} (\varepsilon + \varepsilon^2)^{-1} n^{-\frac{1}{d}} \ell(\Pi_{n f}^\omega \cap A_{\varepsilon+\varepsilon^2}(x_0)) \\ &\leq \frac{c_d}{d} \left(\frac{f^e(x_0)}{\varphi_{x_1}(x_0) \cdots \varphi_{x_d}(x_0)} \right)^{\frac{1}{d}}, \end{aligned}$$

and hence

$$\varphi_{x_1}(x_0) \cdots \varphi_{x_d}(x_0) \leq \frac{C_d^d}{d^d} f^e(x_0).$$

3. We now show that V is an essential viscosity supersolution of (P). Let $\varphi \in C^2(\mathbb{R}^d)$ such that $V - \varphi$ has a local minimum at x_0 . Without loss of generality, we may assume that x_0 is a minimum of $V - \varphi$ relative to the set $[0, M]^d$. Since V is nondecreasing, $\varphi_{x_i}(x_0) \geq 0$ for all i . For $\lambda > 0$, set

$$\varphi^\lambda(x) = \varphi(x) + \lambda(x_1 + \cdots + x_d).$$

Then $V - \varphi^\lambda$ has a strict minimum at x_0 , relative to the set $[0, x_0]$, and $\varphi_{x_i}^\lambda(x_0) > 0$ for all i . Therefore, there exists a sequence $\{x_k\}_{k \in \mathbb{N}}$ in $[0, x_0]$ converging to x_0 such that $V_k - \varphi^\lambda$ has a minimum at x_k , relative to $[0, x_0]$.

Let $r, m > 0$ such that $B(x_0, r) \subseteq (0, M)^d$ and $\varphi_{x_i}^\lambda(x) > m$ for all $x \in B(x_0, r)$ and all i . Let $\varepsilon > 0$ and let $X_1^\omega, \dots, X_j^\omega$ be a chain in $\Pi_{n_k f}^\omega \cap \mathcal{A}_{k, \varepsilon}$. Then by the definition of $\mathcal{A}_{k, \varepsilon}$ (4.12)

$$U_{n_k}^\omega(X_1^\omega) \geq U_{n_k}^\omega(x_0) - \lfloor \varepsilon n_k^{\frac{1}{d}} \rfloor,$$

and therefore

$$U_{n_k}^\omega(x_0) \geq U_{n_k}^\omega(X_j^\omega) \geq U_{n_k}^\omega(X_1^\omega) + j - 1 \geq U_{n_k}^\omega(x_0) - \lfloor \varepsilon n_k^{\frac{1}{d}} \rfloor + j - 1.$$

Hence $j \leq \varepsilon n_k^{\frac{1}{d}} + 1$ and therefore

$$\ell(\Pi_{n_k f}^\omega(\mathcal{A}_{k, \varepsilon})) \leq \varepsilon n_k^{\frac{1}{d}} + 1. \quad (4.14)$$

Since $V_k - \varphi^\lambda$ has a minimum at x_k

$$V_k(z) - V_k(x_0) \geq \varphi^\lambda(z) - \varphi^\lambda(x_0) + \varphi^\lambda(x_0) - \varphi^\lambda(x_k) + V_k(x_k) - V_k(x_0),$$

for $z \in [0, x_0]$. Since $x_k \rightarrow x_0$ and $V_k \rightarrow V$ locally uniformly, we can choose k larger, if necessary, so that

$$V_k(z) - V_k(x_0) \geq \varphi^\lambda(z) - \varphi^\lambda(x_0) - \varepsilon^2 \quad \text{for all } z \in [0, x_0].$$

It follows that

$$\mathcal{A}_{k, \varepsilon} \supseteq \{z \in B(x_0, \sqrt{\varepsilon}) : z \leq x_0 \text{ and } \varphi^\lambda(z) \geq \varphi^\lambda(x_0) + \varepsilon^2 - \varepsilon\} =: A_{\varepsilon - \varepsilon^2}(x_0),$$

where $A_{\varepsilon-\varepsilon^2}(x_0)$ is as defined in Theorem 5. Invoking (4.14) and the monotonicity of ℓ

$$\varepsilon n_k^{\frac{1}{d}} + 1 \geq \ell(\Pi_{n_k f}^\omega \cap \mathcal{A}_{k,\varepsilon}) \geq \ell(\Pi_{n_k f}^\omega \cap A_{\varepsilon-\varepsilon^2}(x_0)),$$

and therefore

$$\varepsilon \geq \liminf_{n \rightarrow \infty} n^{-\frac{1}{d}} \ell(\Pi_{n f}^\omega \cap A_{\varepsilon-\varepsilon^2}(x_0)).$$

Since Theorem 5 holds for all $\omega \in \Omega$ and $\varphi_{x_i}^\lambda(x_0) > 0$

$$\begin{aligned} 1 &\geq \liminf_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} (\varepsilon - \varepsilon^2)^{-1} n^{-\frac{1}{d}} \ell(\Pi_{n f}^\omega \cap A_{\varepsilon-\varepsilon^2}(x_0)) \\ &\geq \frac{c_d}{d} \left(\frac{f_\varepsilon(x_0)}{\varphi_{x_1}^\lambda(x_0) \cdots \varphi_{x_d}^\lambda(x_0)} \right)^{\frac{1}{d}}. \end{aligned}$$

Thus, we arrive at

$$(\varphi_{x_1}(x_0) + \lambda) \cdots (\varphi_{x_d}(x_0) + \lambda) \geq \frac{c_d^d}{d^d} f_\varepsilon(x_0).$$

Since $\lambda > 0$ was arbitrary, V is an essential viscosity supersolution of (P).

4. By the assumed uniqueness of nondecreasing essential viscosity solutions of (P), we have $U^\omega = u$, where u is the unique nondecreasing essential viscosity solution of (P). Since we can apply the same argument to any subsequence of $\{n^{-\frac{1}{d}} U_n^\omega\}_{n \in \mathbb{N}}$, and extract a further subsequence converging locally uniformly to u , we find that $n^{-\frac{1}{d}} U_n^\omega \rightarrow u$ locally uniformly on $[0, \infty)^d$ for all $\omega \in \Omega$, where Ω has probability one. \square

Acknowledgments

The author is grateful to Lawrence C. Evans and Fraydoun Rezakhanlou for valuable discussions about this work, and to the anonymous referees whose insightful comments improved aspects of the paper.

Appendix A. Complete convergence for longest chain problem

We sketch the proof of Theorem 3.

Proof. Let $\varepsilon > 0$. Let X_1, X_2, X_3, \dots be a sequence of independent and uniformly distributed random variables on $[0, 1]^d$ and let $\ell_n = \ell(\{X_1, \dots, X_n\})$ be the length of a longest chain. Let N denote the cardinality of $\Pi_t \cap [0, 1]^d$. Conditioned on $N = n$, L_t and ℓ_n have the same distribution, so by the usual tail bounds on Poisson random variables we deduce

$$P(L_t \geq \lambda) \leq P\left(\ell_{\lfloor t+t^{\frac{3}{4}} \rfloor} \geq \lambda\right) + \exp\left(\frac{-t^{\frac{1}{2}}}{2}\right). \quad (\text{A.1})$$

By [30] there exist constants $0 < C_1 < C_2$ such that

$$C_1 n^{\frac{1}{d}} \leq \mathbb{E}[\ell_n] \leq C_2 n^{\frac{1}{d}} \quad (\text{A.2})$$

for all $n \geq 1$, and

$$\lim_{n \rightarrow \infty} n^{-\frac{1}{d}} \mathbb{E}[\ell_n] = c_d > 0. \quad (\text{A.3})$$

Since the longest chain function ℓ is a *configuration function* (see [58, Definition 7.1.7]), we have

$$P(\ell_n \leq \mathbb{M}[\ell_n] - \lambda), P(\ell_n \geq \mathbb{M}[\ell_n] + \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{4(\mathbb{M}[\ell_n] + \lambda)}\right), \quad (\text{A.4})$$

for any $\lambda > 0$, where $\mathbb{M}[\ell_n]$ denotes any median of ℓ_n . A short computation involving (A.2) and (A.4) yields

$$|\mathbb{E}[\ell_n] - \mathbb{M}[\ell_n]| \leq C \log(n) n^{\frac{1}{2d}}. \quad (\text{A.5})$$

and we find that $n^{-\frac{1}{d}} \mathbb{M}[\ell_n] \rightarrow c_d$ as $n \rightarrow \infty$. Fix T large enough so that

$$|\mathbb{M}[\ell_n] - c_d n^{\frac{1}{d}}| \leq \frac{\varepsilon}{2} n^{\frac{1}{d}} \quad \text{for all } n > T. \quad (\text{A.6})$$

It is well known (see, e.g., [64]) that

$$|\mathbb{E}[\ell_n] - \mathbb{E}[\ell_k]| \leq C|n - k|^{\frac{1}{d}}.$$

Combining this with (A.5) we deduce

$$|\mathbb{M}[\ell_n] - \mathbb{M}[\ell_k]| \leq C \left(\log(k) k^{\frac{1}{2d}} + \log(n) n^{\frac{1}{2d}} + |n - k|^{\frac{1}{d}} \right). \quad (\text{A.7})$$

Recalling (A.1), (A.4) and (A.6) we have

$$P(L_t \geq c_d t^{\frac{1}{d}} + \varepsilon t^{\frac{1}{d}}) \leq 2 \exp\left(-C\varepsilon^2 t^{\frac{1}{d}}\right) + \exp\left(\frac{-t^{\frac{1}{2}}}{2}\right). \quad (\text{A.8})$$

for all $t > T$ sufficiently large. The other inequality is similar, and the result follows from (3.1). \square

References

- [1] J. Calder, S. Esedoğlu, A. O. Hero III, A Hamilton-Jacobi equation for the continuum limit of non-dominated sorting, *SIAM Journal on Mathematical Analysis* 46 (1) (2014) 603–638.
- [2] V. Barnett, The ordering of multivariate data, *Journal of the Royal Statistical Society. Series A (General)* (1976) 318–355.
- [3] C. G. Small, A survey of multidimensional medians, *International Statistical Review* (1990) 263–277.
- [4] R. Y. Liu, J. M. Parelius, K. Singh, Multivariate analysis by data depth: descriptive statistics, graphics and inference, *The Annals of Statistics* 27 (3) (1999) 783–858.
- [5] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation* 6 (2) (2002) 182–197.
- [6] C. Fonseca, P. Fleming, Genetic algorithms for multiobjective optimization: formulation, discussion and generalization, *Proceedings of the Fifth International Conference on Genetic Algorithms* 1 (1993) 416–423.
- [7] C. Fonseca, P. Fleming, An overview of evolutionary algorithms in multiobjective optimization, *Evolutionary Computation* 3 (1) (1995) 1–16.
- [8] K. Deb, *Multi-objective optimization using evolutionary algorithms*, Wiley, Chichester, UK, 2001.
- [9] N. Srinivas, K. Deb, Multiobjective optimization using nondominated sorting in genetic algorithms, *Evolutionary Computation* 2 (3) (1994) 221–248.
- [10] I. Mitchell, S. Sastry, Continuous path planning with multiple constraints, in: *IEEE Conference on Decision and Control*, Vol. 5, 2003, pp. 5502–5507.
- [11] A. Kumar, A. Vladimirovsky, An efficient method for multiobjective optimal control and optimal control subject to integral constraints, *Journal of Computational Mathematics* 28 (4) (2010) 517–551.

- [12] N. Madavan, Multiobjective optimization using a Pareto differential evolution approach, in: Proceedings of the Congress on Evolutionary Computation (CEC), 2002.
- [13] T. Speed, Statistical analysis of gene expression microarray data, Vol. 11, Chapman & Hall/CRC, 2003.
- [14] A. O. Hero III, G. Fleury, Posterior Pareto front analysis for gene filtering, in: Proceedings of the Workshop on Genomic Signal Processing and Statistics (GENSIPS), 2002.
- [15] A. O. Hero III, Gene selection and ranking with microarray data, in: IEEE International Symposium on Signal Processing and its Applications, Vol. 1, 2003, pp. 457–464.
- [16] A. O. Hero III, G. Fleury, Pareto-optimal methods for gene ranking., VLSI Signal Processing 38 (3) (2004) 259–275.
- [17] G. Fleury, A. O. Hero III, S. Yoshida, T. Carter, C. Barlow, A. Swaroop, Pareto analysis for gene filtering in microarray experiments, in: European Signal Processing Conference (EUSIPCO), 2002.
- [18] G. Fleury, A. O. Hero III, S. Zarepari, A. Swaroop, Gene discovery using Pareto depth sampling distributions, Journal of the Franklin Institute 341 (1–2) (2004) 55–75.
- [19] G. Fleury, A. O. Hero III, S. Zarepari, A. Swaroop, Pareto depth sampling distributions for gene ranking, in: IEEE International Symposium on Biomedical Imaging: Nano to Macro , Vol. 1, 2004, pp. 65–68.
- [20] J. Handl, J. Knowles, Exploiting the trade-off – the benefits of multiple objectives in data clustering, in: Evolutionary Multi-Criterion Optimization, Springer, 2005, pp. 547–560.
- [21] D. Kossmann, F. Ramsak, S. Rost, et al., Shooting stars in the sky: an online algorithm for skyline queries, in: Proceedings of the 28th International Conference on Very Large Data Bases, 2002, pp. 275–286.
- [22] D. Papadias, Y. Tao, G. Fu, B. Seeger, Progressive skyline computation in database systems, ACM Transactions on Database Systems 30 (1) (2005) 41–82.

- [23] K.-J. Hsiao, J. Calder, A. O. Hero III, Pareto-depth for multiple-query image retrieval, *IEEE Transactions on Image Processing* 24 (2) (2015) 583–594.
- [24] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, *Communications on Pure and Applied Mathematics* 42 (5) (1989) 577–685.
- [25] T. Chan, L. Vese, Active contours without edges, *IEEE Transactions on Image Processing* 10 (2) (2001) 266–277.
- [26] K.-J. Hsiao, K. Xu, J. Calder, A. O. Hero III, Multi-criteria anomaly detection using Pareto Depth Analysis, in: *Advances in Neural Information Processing Systems* 25, 2012, pp. 854–862.
- [27] K.-J. Hsiao, K. Xu, J. Calder, A. O. Hero III, Multi-criteria similarity-based anomaly detection using Pareto Depth Analysis, *IEEE Transactions on Neural Networks and Learning Systems* To appear.
- [28] S. Ulam, Monte Carlo calculations in problems of mathematical physics, *Modern Mathematics for the Engineers* (1961) 261–281.
- [29] J. Hammersley, A few seedlings of research, in: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, 1972, pp. 345–394.
- [30] B. Bollobás, P. Winkler, The longest chain among random points in Euclidean space, *Proceedings of the American Mathematical Society* 103 (2) (1988) 347–353.
- [31] J.-D. Deuschel, O. Zeitouni, Limiting curves for i.i.d. records, *The Annals of Probability* 23 (2) (1995) 852–878.
- [32] F.-A. Fortin, S. Grenier, M. Parizeau, Generalizing the improved runtime complexity algorithm for non-dominated sorting, in: *Proceeding of the Fifteenth Annual Conference on Genetic and Evolutionary Computation*, ACM, 2013, pp. 615–622.
- [33] M. T. Jensen, Reducing the run-time complexity of multiobjective EAs: The NSGA-II and other algorithms, *IEEE Transactions on Evolutionary Computation* 7 (5) (2003) 503–515.

- [34] A. Gilbert, M. Strauss, Analysis of data streams: computational and algorithmic challenges, *Technometrics* 49 (3) (2007) 346–356.
- [35] J. Calder, S. Esedoğlu, A. O. Hero III, A PDE-based approach to non-dominated sorting, *SIAM Journal on Numerical Analysis* 53 (1) (2015) 82–104.
- [36] J. Calder, Numerical schemes and rates of convergence for the Hamilton-Jacobi equation continuum limit of nondominated sorting, arXiv preprint:1508.01557.
- [37] M. Bardi, I. Capuzzo-Dolcetta, Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations, Springer, 1997.
- [38] J. Calder, Directed last passage percolation with discontinuous weights, *Journal of Statistical Physics* 158 (45) (2015) 903–949.
- [39] T. Seppäläinen, A growth model in multiple dimensions and the height of a random partial order, *Lecture Notes-Monograph Series* (2007) 204–233.
- [40] P.-L. Lions, P. E. Souganidis, Homogenization of “viscous” Hamilton-Jacobi equations in stationary ergodic media, *Communications in Partial Difference Equations* 30 (3) (2005) 335–375.
- [41] E. Kosygina, F. Rezakhanlou, S. R. Varadhan, Stochastic homogenization of Hamilton-Jacobi-Bellman equations, *Communications on Pure and Applied Mathematics* 59 (10) (2006) 1489–1521.
- [42] E. Kosygina, S. R. Varadhan, Homogenization of Hamilton-Jacobi-Bellman equations with respect to time-space shifts in a stationary ergodic medium, *Communications on Pure and Applied Mathematics* 61 (6) (2008) 816–847.
- [43] P. E. Souganidis, Stochastic homogenization of Hamilton-Jacobi equations and some applications, *Asymptotic Analysis* 20 (1) (1999) 1–11.
- [44] P.-L. Lions, P. E. Souganidis, Correctors for the homogenization of Hamilton-Jacobi equations in the stationary ergodic setting, *Communications on Pure and Applied Mathematics* 56 (10) (2003) 1501–1524.

- [45] F. Rezakhanlou, Continuum limit for some growth models, *Stochastic Processes and Their Applications* 101 (1) (2002) 1–41.
- [46] T. Seppäläinen, Hydrodynamic scaling, convex duality and asymptotic shapes of growth models, *Markov Processes and Related Fields* 4 (1) (1998) 1–26.
- [47] R. Schwab, Stochastic homogenization of Hamilton-Jacobi equations in stationary ergodic spatio-temporal media, *Indiana Univ. Math. J* 58 (2) (2009) 537–581.
- [48] P.-L. Lions, P. E. Souganidis, et al., Stochastic homogenization of Hamilton-Jacobi and “viscous” Hamilton–Jacobi equations with convex nonlinearities–revisited, *Communications in Mathematical Sciences* 8 (2) (2010) 627–637.
- [49] S. N. Armstrong, P. E. Souganidis, Stochastic homogenization of level-set convex Hamilton–Jacobi equations, *International Mathematics Research Notices* 2013 (15) (2013) 3420–3449.
- [50] S. N. Armstrong, P. E. Souganidis, Stochastic homogenization of Hamilton–Jacobi and degenerate Bellman equations in unbounded environments, *Journal de Mathématiques Pures et Appliquées* 97 (5) (2012) 460–504.
- [51] S. N. Armstrong, H. V. Tran, Y. Yu, Stochastic homogenization of a nonconvex Hamilton–Jacobi equation, *Calculus of Variations and Partial Differential Equations* (2015) 1–18.
- [52] F. Rezakhanlou, Continuum limit for some growth models II, *Annals of Probability* (2001) 1329–1372.
- [53] J. F. C. Kingman, *Poisson processes*, Vol. 3, Oxford university press, 1992.
- [54] H. Ishii, Hamilton-jacobi equations with discontinuous hamiltonians on arbitrary open sets, *Bull. Fac. Sci. Eng. Chuo Univ* 28 (28) (1985) 1985.
- [55] H. Kesten, Comment to “Subadditive ergodic theory” by J. F. C. Kingman, *Annals of Probability* 1 (1973) 903.

- [56] A. Frieze, On the length of the longest monotone subsequence in a random permutation, *The Annals of Applied Probability* (1991) 301–305.
- [57] B. Bollobás, G. Brightwell, The height of a random partial order: concentration of measure, *The Annals of Applied Probability* (1992) 1009–1018.
- [58] M. Talagrand, Concentration of measure and isoperimetric inequalities in product spaces, *Publications Mathématiques de l’Institut des Hautes Etudes Scientifiques* 81 (1) (1995) 73–205.
- [59] G. I. Barenblatt, *Scaling*, no. 34 in *Cambridge Texts in Applied Mathematics*, Cambridge University Press, 2003.
- [60] G. Barles, P. E. Souganidis, Convergence of approximation schemes for fully nonlinear second order equations, *Asymptotic Analysis* 4 (3) (1991) 271–283.
- [61] B. F. Logan, L. A. Shepp, A variational problem for random Young tableaux, *Advances in Mathematics* 26 (2) (1977) 206–222.
- [62] A. Vershik, S. Kerov, Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables, *Soviet Doklady Mathematics* 18 (527-531) (1977) 38.
- [63] J. F. C. Kingman, Subadditive ergodic theory, *The Annals of Probability* (1973) 883–899.
- [64] J. M. Steele, *Probability theory and combinatorial optimization*, Vol. 69 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, SIAM, 1997.