Hamilton-Jacobi Equations for the Continuum Limits of Sorting and Percolation Problems

Jeff Calder $^1$  Selim Esedoḡlu $^1$  Alfred Hero $^2$

$^1$Department of Mathematics, University of Michigan

$^2$Dept. of Electrical Engineering and Computer Science, University of Michigan

SIAM Annual Meeting, Chicago
July 8, 2014
Outline

Background
  Motivating example
  Non-dominated sorting

Continuum limit of non-dominated sorting
  Main Result
  Non-rigorous derivation
  Basic ingredients of proof

Continuum limit for directed last passage percolation (DLPP)
  Intro to DLPP
  Main result

Numerical scheme for PDE
  Definition of scheme
  Convergence
  Applications

References
Motivating example: Google Goggles

Query image

Retrieved images
Facial recognition

**Problem:** Retrieve images containing faces from a large database $S$. 

- Look for features that are characteristic of faces. Examples include eyes, nose, mouth, ears, etc.
- For each of these features, define an objective function $f_i(I) = 1 - \text{Probability that image } I \text{ has feature } i$.
- Solve the multi-objective optimization problem: $\arg\min_{I \in S} (f_1(I), \ldots, f_d(I))$. 
Facial recognition

Problem: Retrieve images containing faces from a large database $S$.

One approach:

1. Look for features that are characteristic of faces.
   - Eyes, nose, mouth, ears, etc.
Facial recognition

Problem: Retrieve images containing faces from a large database $S$.

One approach:

1. Look for features that are characteristic of faces.
   ▶ Eyes, nose, mouth, ears, etc.

2. For each of $d$ features, define an objective function

   $$f_i(I) = 1 - \text{Probability that image } I \text{ has feature } i,$$
Facial recognition

**Problem:** Retrieve images containing faces from a large database $S$.

**One approach:**

1. Look for features that are characteristic of faces.
   - Eyes, nose, mouth, ears, etc.

2. For each of $d$ features, define an objective function

$$f_i(I) = 1 - \text{Probability that image } I \text{ has feature } i,$$

3. Solve the multi-objective optimization problem:

$$\arg \min_{I \in S} (f_1(I), \ldots, f_d(I)).$$
Multi-objective optimization

How do we solve the multi-objective optimization problem

$$\arg \min_{I \in S} (f_1(I), \ldots, f_d(I))$$

Basic approach:

1. Choose some weights $\alpha_i \in [0, 1]$ with $\sum_{i=1}^d \alpha_i = 1$ and define $f_{\alpha}(I) = \alpha_1 f_1(I) + \alpha_2 f_2(I) + \cdots + \alpha_d f_d(I)$.

2. Solve the scalarized optimization problem $\arg \min_{I \in S} f_{\alpha}(I)$.

Problems:

1. Difficult to choose weights
2. Ignores relevant solutions
Multi-objective optimization

How do we solve the multi-objective optimization problem

$$\underset{I \in S}{\arg\min} (f_1(I), \ldots, f_d(I)) ?$$

Basic approach:

1. Choose some weights $\alpha_i \in [0, 1]$ with $\sum_{i=1}^{d} \alpha_i = 1$ and define

$$f_{\alpha}(I) = \alpha_1 f_1(I) + \alpha_2 f_2(I) + \cdots + \alpha_d f_d(I).$$
Multi-objective optimization

How do we solve the multi-objective optimization problem

$$\arg\min_{I \in S} (f_1(I), \ldots, f_d(I))$$

Basic approach:

1. Choose some weights $\alpha_i \in [0, 1]$ with $\sum_{i=1}^d \alpha_i = 1$ and define

$$f_{\alpha}(I) = \alpha_1 f_1(I) + \alpha_2 f_2(I) + \cdots + \alpha_d f_d(I).$$

2. Solve the scalarized optimization problem

$$\arg\min_{I \in S} f_{\alpha}(I).$$
Multi-objective optimization

How do we solve the multi-objective optimization problem

$$\arg \min_{I \in S} (f_1(I), \ldots, f_d(I))?$$

Basic approach:

1. Choose some weights $\alpha_i \in [0, 1]$ with $\sum_{i=1}^{d} \alpha_i = 1$ and define

   $$f_\alpha(I) = \alpha_1 f_1(I) + \alpha_2 f_2(I) + \cdots + \alpha_d f_d(I).$$

2. Solve the scalarized optimization problem

   $$\arg \min_{I \in S} f_\alpha(I).$$

Problems:

1. Difficult to choose weights
2. Ignores relevant solutions
Basic approach
Basic approach
Basic approach
Basic approach
Non-dominated solutions
Non-dominated solutions
Non-dominated solutions
Non-dominated solutions
Non-dominated solutions
Non-dominated solutions
Non-dominated sorting

Let $X_1, \ldots, X_n$ be points in $\mathbb{R}^d$ and set $S = \{X_1, \ldots, X_n\}$.

Define the partial order

$$x \leq y \iff x_i \leq y_i \text{ for all } i \in \{1, \ldots, d\}.$$  

We say $x \in S$ is minimal if there are no elements $y \in S$ with $y \neq x$ and $y \leq x$. 
Non-dominated sorting

Let \( X_1, \ldots, X_n \) be points in \( \mathbb{R}^d \) and set \( S = \{ X_1, \ldots, X_n \} \).

Define the partial order

\[
x \leq y \iff x_i \leq y_i \text{ for all } i \in \{1, \ldots, d\}.
\]

We say \( x \in S \) is minimal if there are no elements \( y \in S \) with \( y \neq x \) and \( y \leq x \).

Definition

Non-dominated sorting is the process of arranging \( S \) into layers \( \mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \ldots \), by repeated removal of the set of minimal elements:

\[
\mathcal{F}_1 = \text{minimal elements of } S,
\]

\[
\mathcal{F}_k = \text{minimal elements of } S \setminus \bigcup_{j \leq k-1} \mathcal{F}_j.
\]
Applications

**Multi-criteria optimization**
- Genetic algorithms [Deb et al., 2002]
- Gene selection and ranking [Hero, 2003]
- Database systems [Papadias et al., 2005]
Applications

Multi-criteria optimization
- Genetic algorithms [ Deb et al., 2002]
- Gene selection and ranking [ Hero, 2003]
- Database systems [ Papadias et al., 2005]

Combinatorics and probability
- Longest chain in Euclidean space [ Hammersley, 1972]
- Patience sorting [ Aldous and Diaconis, 1999]
- Young Tableaux [ Viennot, 1984]
- Graph theory [ Lou and Sarrafzadeh, 1993]
- Molecular biology [ Pevzner, 2000]
- Integrated circuit design [ Adhar, 2007]

Materials science
- Polynuclear growth [ Pr̈ahofer and Spohn, 2000]
Applications

**Multi-criteria optimization**
- Genetic algorithms [Deb et al., 2002]
- Gene selection and ranking [Hero, 2003]
- Database systems [Papadias et al., 2005]

**Combinatorics and probability**
- Longest chain in Euclidean space [Hammersley, 1972]
- Patience sorting [Aldous and Diaconis, 1999]
- Young Tableaux [Viennot, 1984]
- Graph theory [Lou and Sarrafzadeh, 1993]
- Molecular biology [Pevzner, 2000]
- Integrated circuit design [Adhar, 2007]

**Materials science**
- Polynuclear growth [Prähofer and Spohn, 2000]
Demo: 50 Random samples
Demo: Uniform distribution
Demo: Uniform distribution
Demo: Uniform distribution
Demo: Uniform distribution
Demo: Uniform distribution
Demo: Uniform distribution on $[0, 1]^2 \setminus [0, 0.5]^2$
Demo: Uniform distribution on $[0, 1]^2 \setminus [0, 0.5]^2$
Demo: Uniform distribution on $[0, 1]^2 \setminus [0, 0.5]^2$
Demo: Uniform distribution on $[0, 1]^2 \setminus [0, 0.5]^2$
Demo: Uniform distribution on $[0, 1]^2 \setminus [0, 0.5]^2$
Question

Can we characterize the asymptotic shapes of the Pareto fronts?
Outline

Background
  Motivating example
  Non-dominated sorting

**Continuum limit of non-dominated sorting**
  Main Result
  Non-rigorous derivation
  Basic ingredients of proof

**Continuum limit for directed last passage percolation (DLPP)**
  Intro to DLPP
  Main result

**Numerical scheme for PDE**
  Definition of scheme
  Convergence
  Applications

References
Assumptions

Let \( X_1, \ldots, X_n \) be \( i.i.d. \) random variables with density \( f : \mathbb{R}^d \to \mathbb{R} \).

\((H)\) There exists an open and bounded set \( \Omega \subset \mathbb{R}^d_+ \) with Lipschitz boundary such that \( f|_{\overline{\Omega}} \) is continuous and \( \text{supp}(f) \subset \overline{\Omega} \).
Let $u_n : \mathbb{R}^d \rightarrow \mathbb{N}_0$ be the function that counts the layers $\mathcal{F}_1, \mathcal{F}_2, \ldots$.
Theorem (Calder, Esedoğlu, Hero, 2013)

Suppose (H) holds and let $X_1, \ldots, X_n$ be i.i.d. with density $f$. Then there exists $c_d > 0$ such that

$$n^{-\frac{1}{d}} u_n \longrightarrow c_d U \quad \text{in} \quad L^\infty(\mathbb{R}_+^d) \quad \text{almost surely},$$

where $U \in C^{0, \frac{1}{d}}([0, \infty)^d)$ is the unique Pareto-monotone \footnote{$x \leq y \implies U(x) \leq U(y)$} viscosity solution of

$$(P) \begin{cases} U_{x_1} \cdots U_{x_d} = f & \text{in} \quad \mathbb{R}_+^d \\ U = 0 & \text{on} \quad \partial \mathbb{R}_+^d. \end{cases}$$
Demo: $f = 1 - \chi[0,0.5]^2$
Demo: $f = 1 - \chi_{B_{0.5}(0,0)}$
Non-rigorous derivation of (P)

Suppose we have $X_1, \ldots, X_n$ i.i.d. with density $f \in C(\mathbb{R}^d)$.
Non-rigorous derivation of (P)

Suppose we have \(X_1, \ldots, X_n\) i.i.d. with density \(f \in C(\mathbb{R}^d)\).

Basic geometric considerations:

Surface order growth: \(\#\mathcal{F}_i \sim n^{\frac{d-1}{d}}\).

and

Number of fronts \(\sim n^{\frac{1}{d}}\).
Non-rigorous derivation of (P)

Suppose we have $X_1, \ldots, X_n$ i.i.d. with density $f \in C(\mathbb{R}^d)$.

Basic geometric considerations:

Surface order growth: $\#F_i \sim n^{d-1\over d}$.

and

Number of fronts $\sim n^{1\over d}$.

Let’s suppose that $n^{-1\over d} u_n \rightarrow U \in C^1$. Then we should have

$$F_i \approx \{ x : U(x) = in^{1\over d} \},$$

for $n$ large.
Non-rigorous derivation of (P)

\( \ell_1 = \frac{\langle DU, v \rangle}{U_x} \)

\( \ell_2 = \frac{\langle DU, v \rangle}{U_x} \)

\( \{U = U(x)\} \)
Non-rigorous derivation of (P)

For small $|v|

\langle DU, v \rangle \approx U(x + v) - U(x)$
Non-rigorous derivation of (P)

For small $|v|$

\[
\langle DU, v \rangle \approx U(x + v) - U(x) \\
\approx (\# \text{ fronts in } A)n^{-\frac{1}{d}}
\]
Non-rigorous derivation of (P)

For small $|v|$

$$\langle DU, v \rangle \approx U(x + v) - U(x)$$

$$\approx (\# \text{ fronts in } A)n^{-\frac{1}{d}}$$

$$\approx (\# \text{ samples in } A)^{\frac{1}{d}}n^{-\frac{1}{d}}$$
Non-rigorous derivation of (P)

For small $|v|$

\[
\langle DU, v \rangle \approx U(x + v) - U(x)
\]
\[
\approx (\# \text{ fronts in } A)n^{-\frac{1}{d}}
\]
\[
\approx (\# \text{ samples in } A)\frac{1}{d} n^{-\frac{1}{d}}
\]
\[
\approx (n|A|f(x))^{\frac{1}{d}} n^{-\frac{1}{d}}.
\]
Non-rigorous derivation of (P)

For small $|v|

\langle DU, v \rangle \approx U(x + v) - U(x)

\approx (\# \text{ fronts in } A)n^{-\frac{1}{d}}

\approx (\# \text{ samples in } A)^{\frac{1}{d}} n^{-\frac{1}{d}}

\approx (n |A| f(x))^{\frac{1}{d}} n^{-\frac{1}{d}}.

Using $|A| \approx \frac{(DU, v)^d}{U_{x_1} \cdots U_{x_d}}$ we have

\langle DU, v \rangle \approx \left( \frac{f(x)}{U_{x_1} \cdots U_{x_d}} \right)^{\frac{1}{d}} \langle DU, v \rangle$
Non-rigorous derivation of (P)

For small $|v|$

$$\langle DU, v \rangle \approx U(x + v) - U(x)$$

$$\approx (\# \text{ fronts in } A)n^{-\frac{1}{d}}$$

$$\approx (\# \text{ samples in } A)\frac{1}{d} n^{-\frac{1}{d}}$$

Using $|A| \approx \frac{\langle DU, v \rangle^d}{U_{x_1} \ldots U_{x_d}}$ we have

$$\langle DU, v \rangle \approx \left( \frac{f(x)}{U_{x_1} \ldots U_{x_d}} \right)^{\frac{1}{d}} \langle DU, v \rangle$$

$$U_{x_1} \ldots U_{x_d} = f$$
Basic ingredients of the proof
Basic ingredients of the proof
Basic ingredients of the proof
Basic ingredients of the proof
Basic ingredients of the proof
Basic ingredients of the proof
A chain in \( S \) is a sequence \( x_1, \ldots, x_\ell \) such that

\[
x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_\ell
\]

We can alternatively define \( u_n : \mathbb{R}^d \to \mathbb{N}_0 \) by

\[
u_n(x) := \text{Length of a longest chain in } \{X_i : X_i \leq x\}.
\]
Basic ingredients of the proof

1. For $X_1, \ldots, X_n$ i.i.d. uniform on $[0, 1]^d$ [Hammersley, 1972]

   Length of a longest chain in $\{X_1, \ldots, X_n\} \sim c_d n^{\frac{1}{d}}$ almost surely.
Basic ingredients of the proof

1. For $X_1, \ldots, X_n$ \textit{i.i.d.} uniform on $[0, 1]^d$ [Hammersley, 1972]

   Length of a longest chain in $\{X_1, \ldots, X_n\} \sim c_d n^{\frac{1}{d}}$ almost surely.

2. This leads formally to a variational problem for $U$

   \[
   U(x) = \sup_{\gamma' \in \mathbb{R}_+^d, \gamma(1) = x} \int_0^1 f(\gamma(t)) \frac{1}{d} (\gamma_1'(t) \cdots \gamma_d'(t))^{\frac{1}{d}} dt. \quad (1)
   \]

   ▶ Generalization of [Deuschel and Zeitouni, 1995].
Basic ingredients of the proof

1. For $X_1, \ldots, X_n$ i.i.d. uniform on $[0, 1]^d$ [Hammersley, 1972]

   Length of a longest chain in $\{X_1, \ldots, X_n\} \sim c_d n^{1/d}$ almost surely.

2. This leads formally to a variational problem for $U$

$$U(x) = \sup_{\gamma' \in \mathbb{R}_+^d, \gamma(1) = x} \int_0^1 f(\gamma(t)) \frac{1}{d} (\gamma_1'(t) \cdots \gamma_d'(t))^{\frac{1}{d}} \, dt.$$  \hfill (1)

   ▶ Generalization of [Deuschel and Zeitouni, 1995].

3. Prove $U$ is unique solution of (P) under hypotheses (H).
Basic ingredients of the proof

1. For $X_1, \ldots, X_n$ i.i.d. uniform on $[0, 1]^d$ [Hammersley, 1972]

   Length of a longest chain in $\{X_1, \ldots, X_n\} \sim c_d n^{\frac{1}{d}}$ almost surely.

2. This leads formally to a variational problem for $U$

   $U(x) = \sup_{\gamma' \in \mathbb{R}_+^d, \gamma(1) = x} \int_0^1 f(\gamma(t)) \frac{1}{d} (\gamma'_1(t) \cdots \gamma'_d(t))^{\frac{1}{d}} dt. \quad (1)$

   ▶ Generalization of [Deuschel and Zeitouni, 1995].

3. Prove $U$ is unique solution of (P) under hypotheses (H).

4. Prove $n^{-\frac{1}{d}} u_n \rightarrow U$ given by (1).

   ▶ Establish convergence for $f$ piecewise constant on a fixed grid.

   ▶ Relax to $f$ satisfying (H) by approximation argument. Relies on the following Hölder estimate for solutions $U$ of (P)

   $[U]_{\frac{1}{d}} \leq d \|f\|_{L_\infty}^{\frac{1}{d}}$. 
Details of proof

Full details of proof in

Outline

Background
  Motivating example
  Non-dominated sorting

Continuum limit of non-dominated sorting
  Main Result
  Non-rigorous derivation
  Basic ingredients of proof

Continuum limit for directed last passage percolation (DLPP)
  Intro to DLPP
  Main result

Numerical scheme for PDE
  Definition of scheme
  Convergence
  Applications

References
Directed last passage percolation (DLPP)

Given independent random variables $X(i, j)$ for $(i, j) \in \mathbb{N}_0^2$ we define

$$L(M, N) = \max_{p \in \Pi(M, N)} \sum_{(i, j) \in p} X(i, j).$$

Of interest: Asymptotics as $M, N \to \infty$

Applications in

- Zero-temperature directed polymer growth
  [Comets et al., 2004]

- Interacting particle systems (TASEP)
  [Ferrari and Spohn, 2006]

- Randomly growing Young diagrams [Seppäläinen, 1996]
Existing results: \(i.i.d.\) passage times

When \(X(i,j)\) are \(i.i.d.\) and either all geometrically [Johansson, 2000], or all exponentially [Rost, 1981] distributed, we have

\[
\lim_{N \to \infty} \frac{1}{N} L([Nx]) = \mu(x_1 + x_2) + 2\sigma \sqrt{x_1 x_2},
\]

with probability one, where \(\mu = \mathbb{E}(X(i,j))\) and \(\sigma^2 = \text{Var}(X(i,j))\).
Main result: Non-i.i.d. passage times

**Theorem**

Let $\mu : [0, \infty)^2 \to [0, \infty)$. Suppose the passage times $X(i, j)$ are independent geometric (resp. exponential) random variables satisfying

$$\mathbb{E}(X(i, j)) = \mu(iN^{-1}, jN^{-1}).$$
Main result: Non-i.i.d. passage times

Theorem

Let $\mu : [0, \infty)^2 \to [0, \infty)$. Suppose the passage times $X(i, j)$ are independent geometric (resp. exponential) random variables satisfying

$$\mathbb{E}(X(i, j)) = \mu(iN^{-1}, jN^{-1}).$$

Under certain regularity assumptions on $\mu$ we have

$$\frac{1}{N} L([N \cdot]) \longrightarrow U \quad \text{locally uniformly on } [0, \infty)^2,$$

with probability one, where $U$ is the unique viscosity solution of the Hamilton-Jacobi equation

$$\begin{cases} (U_{x_1} - \mu)(U_{x_2} - \mu) = \mu(1 + \mu) \ (\text{resp. } \mu^2) & \text{on } \mathbb{R}_+^2, \\ \min(U_{x_1}, U_{x_2}) \geq \mu & \text{on } \mathbb{R}_+^2, \\ U = \varphi & \text{on } \partial \mathbb{R}_+^2, \end{cases}$$

and $\varphi(x) = (x_1 + x_2) \int_0^1 \mu(tx) \, dt$. 
Proof

Proof is similar to the continuum limit for non-dominated sorting and is based on the following variational interpretation of $U$:

\[
U(x) = \sup_{\gamma' \in \mathbb{R}_+^d : \gamma(1)=x} \int_0^1 \mu(\gamma(t))(\gamma'_1(t) + \gamma'_2(t)) + 2\sigma(\gamma(t))\sqrt{\gamma'_1(t)\gamma'_2(t)} \, dt.
\]

A version of this variational problem appeared in [Rolla and Teixeira, 2008].

Outline

Background
  Motivating example
  Non-dominated sorting

Continuum limit of non-dominated sorting
  Main Result
  Non-rigorous derivation
  Basic ingredients of proof

Continuum limit for directed last passage percolation (DLPP)
  Intro to DLPP
  Main result

Numerical scheme for PDE
  Definition of scheme
  Convergence
  Applications

References
The general form of the PDE we seek to solve numerically is

\[
\begin{cases}
(U_{x_1} - g(x)) \cdots (U_{x_d} - g(x)) = f(x) & \text{on } \mathbb{R}^d_+,
\\
\min(U_{x_1}, \ldots, U_{x_d}) \geq g(x) & \text{on } \mathbb{R}^d_+,
\\
U = \varphi & \text{on } \partial \mathbb{R}^d_+.
\end{cases}
\]
General form of PDE

The general form of the PDE we seek to solve numerically is

\[
\begin{cases}
(U_{x_1} - g(x)) \cdots (U_{x_d} - g(x)) = f(x) & \text{on } \mathbb{R}^d_+ , \\
\min(U_{x_1}, \ldots, U_{x_d}) \geq g(x) & \text{on } \mathbb{R}^d_+ , \\
U = \varphi & \text{on } \partial \mathbb{R}^d_.
\end{cases}
\]

For simplicity we will take \( g, \varphi \equiv 0 \) in the following discussion. This corresponds to the continuum limit for non-dominated sorting:

\[
\begin{cases}
U_{x_1} \cdots U_{x_d} = f(x) & \text{on } \mathbb{R}^d_+ , \\
U = 0 & \text{on } \partial \mathbb{R}^d_+ ,
\end{cases}
\]

with the additional constraint that \( U \) is Pareto-monotone.
Consider a grid with spacing $h > 0$. Natural to use backward difference quotients to define numerical solution $U_h : h\mathbb{N}^d_0 \to \mathbb{R}_+$, i.e.,

$$
\prod_{i=1}^{d} \left( U_h(x) - U_h(x - he_i) \right) = h^d f(x) \quad \text{for all } x \in h\mathbb{N}^d.
$$

(3)
Consider a grid with spacing $h > 0$. Natural to use backward difference quotients to define numerical solution $U_h : h \mathbb{N}_0^d \rightarrow \mathbb{R}_+$, i.e.,

$$
\prod_{i=1}^{d} \left( U_h(x) - U_h(x - he_i) \right) = h^d f(x) \quad \text{for all } x \in h \mathbb{N}_0^d.
$$

Given $U_h(x - he_1), \ldots, U_h(x - he_d)$ and $f(x) \geq 0$, $\exists! \ U_h(x)$ solving (3) with

$$
U_h(x) \geq \max \left( U_h(x - he_1), \ldots, U_h(x - he_d) \right).
$$
Consider a grid with spacing $h > 0$. Natural to use backward difference quotients to define numerical solution $U_h : h\mathbb{N}_0^d \to \mathbb{R}_+$, i.e.,

$$
\prod_{i=1}^{d} \left( U_h(x) - U_h(x - he_i) \right) = h^df(x) \quad \text{for all } x \in h\mathbb{N}^d. \quad (3)
$$

Given $U_h(x - he_1), \ldots, U_h(x - he_d)$ and $f(x) \geq 0$, $\exists! \ U_h(x)$ solving (3) with

$$
U_h(x) \geq \max \left( U_h(x - he_1), \ldots, U_h(x - he_d) \right).
$$

For $d = 2$ we can solve (3) explicitly

$$
U_h(x) = \frac{1}{2} \left( U_h(x-he_1) + U_h(x-he_2) \right) + \frac{1}{2} \sqrt{ \left( U_h(x - he_1) - U_h(x - he_2) \right)^2 + 4h^2f(x)}.
$$
Extend $U_h$ to a function on $[0, \infty)^d$ by setting $U_h(x) = U_h(\lfloor x \rfloor h)$.

Then $U_h$ satisfies

$$(S) \left\{ \begin{array}{ll}
S(h, x, U_h) = f(\lfloor x \rfloor h) & \text{if } x \in \mathbb{R}_+^d \setminus \Gamma_h \\
U_h(x) = 0 & \text{if } x \in \Gamma_h.
\end{array} \right. $$

where

$$S(h, x, u) = \prod_{i=1}^{d} \frac{u(x) - u(x - he_i)}{h}, \quad (4)$$

and

$$\Gamma_h = \left\{ x \in \mathbb{R}_+^d : x_i \leq h \text{ for some } i \right\}.$$
Convergence of numerical scheme

(H) There exists an open and bounded set $\Omega \subset \mathbb{R}^d_+$ with Lipschitz boundary such that $f|_{\overline{\Omega}}$ is continuous and $\text{supp}(f) \subset \overline{\Omega}$.

Theorem (Calder, Esedoḡlu, Hero, 2013)
Assume (H). Then the numerical solutions $U_h$ of (S) converge uniformly as $h \to 0$ to the unique Pareto-monotone viscosity solution $U$ of

$$(P) \begin{cases} U_{x_1} \cdots U_{x_d} = f & \text{in } \mathbb{R}^d_+ \\ U = 0 & \text{on } \partial \mathbb{R}^d_+. \end{cases}$$
Proof sketch

Proof roughly follows well-known framework of [Barles and Souganidis, 1991]

- Strong uniqueness, monotonicity, consistency, stability $\implies$ convergence.
- **Recall:** Strong uniqueness refers to comparison for semicontinuous sub and supersolutions.
Proof sketch

Proof roughly follows well-known framework of [Barles and Souganidis, 1991]

- Strong uniqueness, monotonicity, consistency, stability $\implies$ convergence.
- Recall: Strong uniqueness refers to comparison for semicontinuous sub and supersolutions.

Problem: We can only prove comparison for continuous solutions when $f$ is discontinuous.
Proof sketch

1. Prove a coarse Hölder estimate for $U_h$

$$|U_h(x) - U_h(y)| \leq 2d\|f\|_{L_\infty(\mathbb{R}^d)} \left(|x - y|^\frac{1}{d} + h^\frac{1}{d}\right).$$ (5)
Proof sketch

1. Prove a coarse Hölder estimate for $U_h$

$$|U_h(x) - U_h(y)| \leq 2d\|f\|_{L^\infty(\mathbb{R}^d_+)}^\frac{1}{d}(|x - y|^{\frac{1}{d}} + h^{\frac{1}{d}}). \quad (5)$$

2. Stability of (S) and (5) are sufficient to invoke the Arzelà-Ascoli Theorem and extract a uniformly convergent subsequence

$$U_{h_k} \longrightarrow u \in C^{0, \frac{1}{d}}(\overline{\mathbb{R}^d_+}).$$
Proof sketch

1. Prove a coarse Hölder estimate for $U_h$

$$|U_h(x) - U_h(y)| \leq 2d\|f\|_{L^\infty(\mathbb{R}_+^d)} \left( |x - y|^{\frac{1}{d}} + h^{\frac{1}{d}} \right). \quad (5)$$

2. Stability of (S) and (5) are sufficient to invoke the Arzelà-Ascoli Theorem and extract a uniformly convergent subsequence

$$U_{h_k} \to u \in C^{0,\frac{1}{d}}(\overline{\mathbb{R}_+^d}).$$

3. Consistency and monotonicity of (S) $\implies$ $u$ is a viscosity solution of (P).
Proof sketch

1. Prove a coarse Hölder estimate for $U_h$

$$|U_h(x) - U_h(y)| \leq 2d \|f\|_{L^\infty(\mathbb{R}^d_+)} (|x - y|^{\frac{1}{d}} + h^{\frac{1}{d}}).$$  \hspace{1cm} (5)

2. Stability of (S) and (5) are sufficient to invoke the Arzelà-Ascoli Theorem and extract a uniformly convergent subsequence

$$U_{h_k} \longrightarrow u \in C^{0,\frac{1}{d}}(\overline{\mathbb{R}^d_+}).$$

3. Consistency and monotonicity of (S) $\implies$ $u$ is a viscosity solution of (P).

4. Uniqueness of (P) $\implies$ $u = U$ and hence $U_h \to U$ uniformly.
Numerical scheme: $d = 2$
Numerical scheme: $d = 3$
Fast approximate sorting

Algorithm (PDE-based Ranking)

1. Select \(k\) points from \(X_1, \ldots, X_n\) at random. Call them \(Y_1, \ldots, Y_k\).
2. Select a grid spacing \(h\) for solving the PDE and estimate \(f\) with a histogram aligned to the grid \(h\mathbb{N}_0^d\), i.e.,

\[
\hat{f}(x) = \frac{1}{kh^d} \cdot \# \left\{ Y_i : x \leq Y_i \leq x + h(1, \ldots, 1) \right\} \text{ for } x \in h\mathbb{N}_0^d.
\]

3. Compute the numerical solution \(\hat{U}_h\) on \(h\mathbb{N}_0^d \cap [0, 1]^d\) via \((S)\).
4. Evaluate \(\hat{U}_h(X_i)\) for \(i = 1, \ldots, n\) via interpolation.

Notes:

- Total complexity is \(O(kh^{-d} + n)\).
- If we fix \(k\) and \(h\), independent of \(n\), then Steps 1-3 have \(O(1)\) complexity.
CPU Time

- # Subsamples = \( k = 10^7 \), Grid for solving PDE = \( 250 \times 250 \).
- \( O(n \log n) \) non-dominated sorting of [Felsner and Wernisch, 1999].
Application in anomaly detection

(a) Example trajectories

(b) 50000 Pareto points

Figure: Accuracy scores for PDE-based ranking and subset ranking for sorting $10^9$ Pareto points from the pedestrian anomaly detection problem versus the number of subsamples $k$.

Application: Finding optimal DLPP paths

\[ L(M, N) = \max_{p \in \Pi(M, N)} \sum_{(i,j) \in p} X(i, j). \]

\[ U(x) = \sup_{\gamma' \in \mathbb{R}^d_+ : \gamma(1) = x} J(\gamma) := \int_0^1 \mu(\gamma(t))(\gamma_1'(t) + \gamma_2'(t)) + 2\sigma(\gamma(t))\sqrt{\gamma_1'(t)\gamma_2'(t)} \, dt. \]

**Algorithm 1: Find \( \varepsilon \)-optimal curve**

Given a step size \( \varepsilon > 0 \) and \( x_0 \in \mathbb{R}^2_+ \), we generate \( x_1, \ldots, x_k, \ldots \) as follows:

\[
\begin{align*}
\text{while } x_k & \in \mathbb{R}^2_+ \text{ do} \\
&s_k^* = \arg\max_{s \in [0,1]} \left\{ U(x_k - (1 - s, s)\varepsilon) + 2\sigma(x_k)\varepsilon \sqrt{s(1-s)} \right\}; \\
&x_{k+1} = (x_k - (1 - s_k^*, s_k^*)\varepsilon)_+; \\
\text{end}
\end{align*}
\]

\( x_{k+1} = 0; \)
Applying: Finding optimal DLPP paths

\[ L(M, N) = \max_{p \in \Pi(M, N)} \sum_{(i, j) \in p} X(i, j). \]

\[ U(x) = \sup_{\gamma' \in \mathbb{R}_+^d : \gamma(1) = x} J(\gamma) := \int_0^1 \mu(\gamma(t))(\gamma_1'(t) + \gamma_2'(t)) + 2\sigma(\gamma(t)) \sqrt{\gamma_1'(t)\gamma_2'(t)} \, dt. \]

**Algorithm 2: Find \( \varepsilon \)-optimal curve**

Given a step size \( \varepsilon > 0 \) and \( x_0 \in \mathbb{R}^2_+ \), we generate \( x_1, \ldots, x_k, \ldots \) as follows:

\[
\text{while } x_k \in \mathbb{R}^2_+ \text{ do} \\
\quad s_k^* = \arg\max_{s \in [0, 1]} \left\{ U(x_k - (1 - s, s)\varepsilon) + 2\sigma(x_k)\varepsilon\sqrt{s(1 - s)} \right\}; \\
\quad x_{k+1} = (x_k - (1 - s_k^*, s_k^*)\varepsilon)_+; \\
\text{end} \\
\]

\( x_{k+1} = 0; \)

**Theorem (Calder 2014)**

Let \( \gamma \) be the polygonal curve connecting the points \( x_1, \ldots, x_k, \ldots \) generated by the algorithm. Then

\[ U(x_0) \leq J(\gamma) + C\varepsilon. \]
Application: Finding optimal DLPP paths

\[ L(M, N) = \max_{\rho \in \Pi(M, N)} \sum_{(i, j) \in \rho} X(i, j). \]

\[ U(x) = \sup_{\gamma' \in \mathbb{R}^d_+: \gamma(1) = x} J(\gamma) := \int_0^1 \mu(\gamma(t))(\gamma_1'(t) + \gamma_2'(t)) + 2\sigma(\gamma(t)) \sqrt{\gamma_1'(t)\gamma_2'(t)} \, dt. \]

\[ \mu(x) = 1 - \chi_{[0, 0.5]}^2 \]

\[ \mu(x) = \exp(-10|x-a|^2) + \exp(-10|x-b|^2), \]

where \( a = (1/4, 3/4) \) and \( b = (3/4, 1/4) \).
Future/ongoing work

1. How do we solve (P) numerically in high dimensions?
Future/ongoing work

1. How do we solve (P) numerically in high dimensions?

2. Is there a higher order accuracy numerical scheme for (P) in low dimensions?
   - Current scheme is $O(h^{1/d})$.
   - Preliminary work suggests a scheme that is $O(h)$ for strictly positive Lipschitz densities $f : [0, 1]^d \to \mathbb{R}_+$. 

3. Directed polymers with positive temperature.
Outline

Background
  Motivating example
  Non-dominated sorting

Continuum limit of non-dominated sorting
  Main Result
  Non-rigorous derivation
  Basic ingredients of proof

Continuum limit for directed last passage percolation (DLPP)
  Intro to DLPP
  Main result

Numerical scheme for PDE
  Definition of scheme
  Convergence
  Applications

References


*Communications in Mathematical Physics*, 265(1):1–44.


*Communications in Mathematical Physics*, 209(2):437–476.


