

Chapter 12

Splitting Enables Overcoming the Curse of Dimensionality

Jérôme Darbon and Stanley J. Osher

Abstract In this chapter we briefly outline a new and remarkably fast algorithm for solving a large class of high dimensional Hamilton-Jacobi (H-J) initial value problems arising in optimal control and elsewhere [1]. This is done without the use of grids or numerical approximations. Moreover, by using the level set method [8] we can rapidly compute projections of a point in \mathbb{R}^n , n large to a fairly arbitrary compact set [2]. The method seems to generalize widely beyond what will we present here to some nonconvex Hamiltonians, new linear programming algorithms, differential games, and perhaps state dependent Hamiltonians.

1 Introduction

We begin with the Hamilton-Jacobi (HJ) initial value problem

$$\begin{cases} \frac{\partial \varphi}{\partial t}(x, t) + H(\nabla_x \varphi(x, t)) = 0 & \text{in } \mathbb{R}^n \times (0, +\infty) \\ \varphi(x, 0) = J(x) & \forall x \in \mathbb{R}^n. \end{cases} \quad (12.1)$$

We assume $J : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex and one coercive, i.e., $\lim_{\|x\|_2 \rightarrow +\infty} \frac{J(x)}{\|x\|_2} = +\infty$, $H : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex and positively one homogeneous (we sometimes relax all these assumptions).

J. Darbon (✉)
CNRS/CMLA-Ecole Normale Supérieure de Cachan, Cachan, France
e-mail: darbon@cmla.ens-cachan.fr

S.J. Osher
Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, USA
e-mail: sjo@math.ucla.edu

A good example of this is

$$H(v) = \|v\|_2.$$

Here $\|v\|_p = (\sum_{i=1}^n |v_i|^p)^{\frac{1}{p}}$ for $p \geq 1$ and $\langle x, v \rangle = \sum_{i=1}^n x_i v_i$.

Let us consider a convex Lipschitz function J having the property that, for Ω a convex compact set of \mathbb{R}^n

$$\begin{cases} J(x) < 0 & \text{for any } x \in \text{int } \Omega \\ J(x) = 0 & \text{for any } x \in (\Omega \setminus \text{int } \Omega) \\ J(x) > 0 & \text{for any } x \in (\mathbb{R}^n \setminus \Omega). \end{cases}$$

We call this level set initial data. Then the set of points for which $\varphi(x, t) = 0, t > 0$ are exactly those at a distance t , from the boundary of Ω . In fact given $\bar{x} \notin \Omega$, then the closest point x_{opt} from \bar{x} to $(\Omega \setminus \text{int } \Omega)$ is exactly

$$x_{\text{opt}} = \bar{x} - t \frac{\nabla \varphi(\bar{x}, t)}{\|\nabla \varphi(\bar{x}, t)\|_2}. \tag{12.2}$$

To solve (12.1) we use the Hopf formula [5]

$$\varphi(x, t) = (J^* + tH)^*(x) = -\min_{v \in \mathbb{R}^n} \{J^*(v) + tH(v) - \langle x, v \rangle\},$$

where the Fenchel-Legendre transform $f^* : \mathbb{R}^n \rightarrow R \cup \{+\infty\}$ of the convex function f is defined by

$$f^*(v) = \sup_{x \in \mathbb{R}^n} \{\langle v, x \rangle - f(x)\}.$$

Moreover, for free we get that the minimizer satisfies

$$\arg \min_{v \in \mathbb{R}^n} \{J^*(v) + tH(v) - \langle x, v \rangle\} = \nabla_x \varphi(x, t). \tag{12.3}$$

whenever $\varphi(\cdot, t)$ is differentiable at x . Let us note here that our algorithm computes $\varphi(x, t)$ but also $\nabla_x \varphi(x, t)$.

Also, we can use the Hopf-Lax formula [5, 6] to solve (12.1).

$$\varphi(x, t) = \min_{z \in \mathbb{R}^n} \left\{ J(z) + tH^* \left(\frac{x-z}{t} \right) \right\} \tag{12.4}$$

for convex H .

From (12.4) it is easy to show that if we have k different initial value problems $i = 1, \dots, k$

$$\begin{cases} \frac{\partial \varphi_i}{\partial t}(x, t) + H(\nabla_x \varphi_i(x, t)) = 0, & \text{in } \mathbb{R}^n \times (0, +\infty) \\ \varphi_i(x, 0) = J_i(x) & \forall x \in \mathbb{R}^n \end{cases}$$

with the usual hypotheses, then (12.4) implies, for any $x \in \mathbb{R}^n$, $t > 0$

$$\varphi_i(x, t) = \min_{z \in \mathbb{R}^n} \left\{ J_i(z) + tH^* \left(\frac{x-z}{t} \right) \right\}.$$

So

$$\min_{i=1,k} \varphi_i(x, t) = \min_{z \in \mathbb{R}^n} \left\{ \min_{i=1,\dots,k} \left\{ J_i(z) + tH^* \left(\frac{x-z}{t} \right) \right\} \right\}$$

solves the initial value problem

$$\begin{cases} \frac{\partial \varphi}{\partial t}(x, t) + H(\nabla_x \varphi(x, t)) = 0, & \text{in } \mathbb{R}^n \times (0, +\infty) \\ \varphi(x, 0) = \min_{i=1,\dots,k} J_i(x) & \forall x \in \mathbb{R}^n. \end{cases} \quad (12.5)$$

This means that if $\Omega = \cup_{i=1,\dots,k} \Omega_i$, where each Ω_i is compact and convex and may overlap, then we can easily compute the set of all points at distance t from Ω which is exactly the solution to (12.5) where each J_i is a level set function for Ω_i . Moreover, at every point \bar{x} outside of $\bar{\Omega}$ for which there is one i such that $\varphi_i(\bar{x}, t) < \varphi_{i'}(\bar{x}, t)$ for any $i \neq i'$, then the closest point x_{opt} to \bar{x} and Ω is again

$$x_{opt} = \bar{x} - t \frac{\nabla_x \varphi_i(\bar{x}, t)}{|\nabla_x \varphi_i(\bar{x}, t)|}.$$

If there are several i for which $\varphi_i(\bar{x}, t)$ is the minimum among all k of them, then $\nabla_x \varphi$ will be “multivalued”, i.e., it will have jumps, but any of the x_{opt} defined above will be a closest point on Ω to \bar{x} .

2 Split Bregman

We solve the optimization problem (12.3) by using the split Bregman algorithm [4, 3, 9] as follows

$$v^{k+1} = \arg \min_{v \in \mathbb{R}^n} \{ J^*(v) - \langle x, v \rangle + \frac{\lambda}{2} \|d^k - v - b^k\|_2^2 \}, \quad (12.6)$$

$$d^{k+1} = \arg \min_{d \in \mathbb{R}^n} \left\{ tH(d) + \frac{\lambda}{2} \|d - v^{k+1} - b^k\|_2^2 \right\} \quad (12.7)$$

$$b^{k+1} = b^k + v^{k+1} - d^{k+1}. \quad (12.8)$$

Here the sequences $(v^k)_{k \in \mathbb{N}}$, $(d^k)_{k \in \mathbb{N}}$ both converge to $\nabla_x \varphi(x, t)$. Let us emphasize again that our numerical algorithm not only computes the solution $\varphi(x, t)$ but also computes $\nabla_x \varphi(x, t)$ when $\varphi(\cdot, t)$ is differentiable.

Both (12.6) and (12.7), up to change of variables, can be reformulated as finding the unique minimizer of

$$\arg \min_w \left\{ \alpha f(w) + \frac{1}{2} \|w - z\|_2^2 \right\}$$

which is the proximal map of f . Equation (12.6) can be solved if either J^* or J have easily computable proximal maps, which often occurs, especially if one of them is smooth.

Equation (12.7) can be easily solved if $H(d) = \|d\|_2$ via the shrink₂ operator defined by

$$\text{shrink}_2(z, \alpha) = \begin{cases} \frac{z}{\|z\|_2} \max(\|z\|_2 - \alpha, 0) & \text{if } z \neq 0 \\ 0 & \text{if } z = 0 \end{cases}$$

and we have

$$\arg \min_w \left\{ \alpha \|w\|_2 + \frac{1}{2} \|w - z\|_2^2 \right\} = \text{shrink}_2(z, \alpha)$$

If $H(d) = \|d\|_1$ we use shrink₁ operator defined as follows for any $i = 1, \dots, n$

$$(\text{shrink}_1(z, \alpha))_i = \begin{cases} z_i - \alpha & \text{if } z_i > \alpha \\ 0 & \text{if } |z_i| \leq \alpha \\ z_i + \alpha & \text{if } z_i < -\alpha \end{cases}$$

and we have

$$\arg \min_w \left\{ \alpha \|w\|_1 + \frac{1}{2} \|w - z\|_2^2 \right\} = \text{shrink}_1(z, \alpha).$$

To solve (12.7) for more general $H(d)$ convex one homogeneous or to find the proximal map for f of that type we use the fact that H^* is the characteristic function of a closed convex set $C \subset \mathbb{R}^n$

$$H^* = I_C.$$

By using the Moreau identity [7] we realize that the proximal map of H can be obtained by projecting onto C . To do this projection, we merely solve the eikonal equation with level set initial data for C via split Bregman as above in (12.6), (12.7), (12.8) with $H(d) = \|d\|_2$. This is easy using the shrink₂ operator. We then use (12.2) to obtain the projection and repeat the entire iteration.

3 Numerical Experiments

Numerical experiments on an Intel Laptop Core i5-5300U running at 2.3 GHz are now presented. We consider diagonal matrices D defined by $D_{ii} = 1 + \frac{1+i}{n}$ for $i = 1, \dots, n$. We also consider matrices A defined by $A_{ii} = 2$ for $i = 1, \dots, n$ and $A_{ij} = 1$

for $i, j = 1, \dots, n$. Table 12.1 presents the average time (in seconds) to evaluate the solution over 1,000,000 samples (x, t) uniformly drawn in $[-10, 10]^n \times [0, 10]$. The convergence is remarkably rapid: 10^{-6} to 10^{-8} seconds on a standard laptop, per function evaluation. Figure 12.1 depicts 2-dimensional slices at different times for the (H-J) equation with a weighted ℓ_1 Hamiltonian $H = \|D \cdot\|_1$, initial data $J = \frac{1}{2} \|\cdot\|_2^2$ and $n = 8$.

n	$\ y\ _1$	$\ y\ _2$	$\ y\ _\infty$	$\ y\ _D$	$\ y\ _A$
4	6.36e-08	1.20e-07	2.69e-07	7.00e-07	8.83e-07
8	6.98e-08	1.28e-07	4.89e-07	1.07e-06	1.57e-06
12	8.72e-08	1.56e-07	7.09e-07	1.59e-06	2.23e-06
16	9.24e-08	1.50e-07	9.92e-07	2.04e-06	2.95e-06

Table 12.1 Time results in seconds for the average time per call for evaluating the solution of the HJ-PDE with the initial data $J = \frac{1}{2} \|\cdot\|_2^2$, several Hamiltonians and various dimensions n .

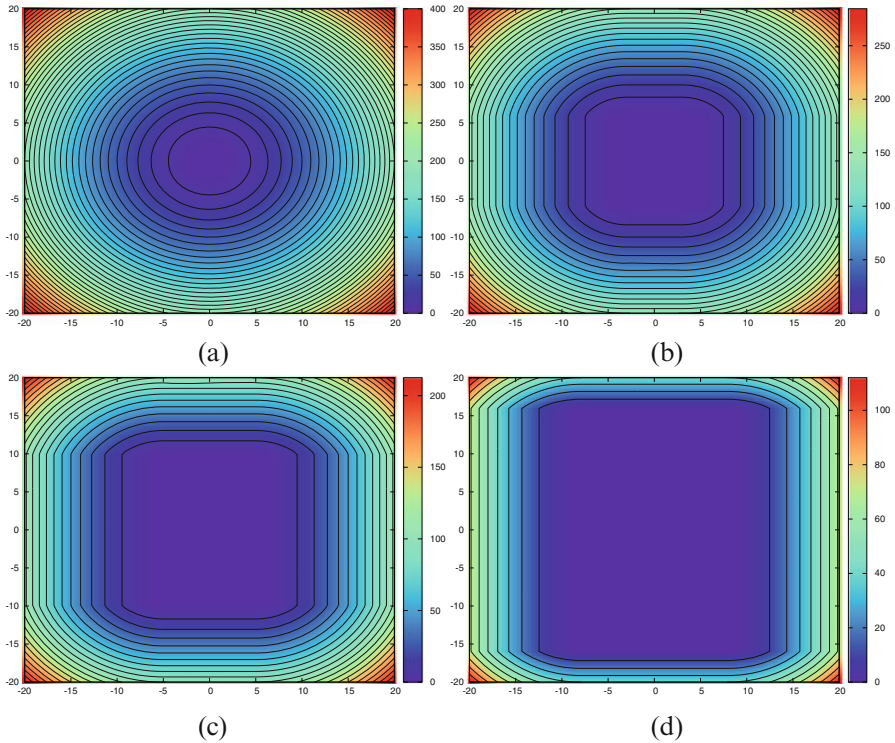


Fig. 12.1 Evaluation of the solution $\phi((x_1, x_2, 0, 0, 0, 0, 0, 0)^\dagger, t)$ of the HJ-PDE with initial data $J = \frac{1}{2} \|\cdot\|_2^2$ and Hamiltonian $H = \|D \cdot\|_1$ for $(x_1, x_2) \in [-20, 20]^2$ for different times t . Plots for $t = 0, 3, 5, 8$ and respectively depicted in (a), (b), (c), and (d). The level lines multiple of 10 are superimposed on the plots.

4 Summary and Future Work

We have derived a very fast and totally parallelizable method to solve a large class of high dimensional, state independent H-J initial value problems. We do this using the Hopf formula and convex optimization via splitting, which overcomes the “curse of dimensionality”. This is also done without the use of grids or numerical approximations, yielding not only the solution, but also its gradient.

We also, as a step in this procedure, very rapidly compute the projections from a point in \mathbb{R}^n , n large, to a fairly arbitrary compact set.

In future work, we expect to extend this set of ideas to nonconvex Hamiltonians, including some that arise in differential games, to new linear programming algorithms, to fast methods for redistancing in level set methods and, hopefully, to a wide class of state dependent Hamiltonians.

Acknowledgements Research supported by ONR grants N000141410683, N000141210838 and DOE grant DE-SC00183838.

References

1. Darbon, J., Osher, S.: Algorithms for overcoming the curse of dimensionality for certain Hamilton-Jacobi equations arising in control theory and elsewhere. *Research in the Mathematical Sciences* (to appear)
2. Darbon, J., Osher, S.: Fast projections onto compact sets in high dimensions using the level set method, Hopf formulas and optimization. (In preparation)
3. Glowinski, R., Marroco, A.: Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité d’une classe de problèmes de Dirichlet non linéaires. *ESAIM: Mathematical Modelling and Numerical Analysis* **9**(R2), 41–76 (1975)
4. Goldstein, T., Osher, S.: The split Bregman method for L1-regularized problems. *SIAM Journal on Imaging Sciences* **2**(2), 323–343 (2009)
5. Hopf, E.: Generalized solutions of non-linear equations of first order (First order nonlinear partial differential equation discussing global locally-Lipschitzian solutions via Jacoby theorem extension). *Journal of Mathematics and Mechanics* **14**, 951–973 (1965)
6. Lax, P.D.: *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*. SIAM, Philadelphia, PA (1990)
7. Moreau, J.J.: Proximité et dualité dans un espace hilbertien. *Bulletin de la Société Mathématique de France* **93**, 273–299 (1965)
8. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics* **79**(1), 12–49 (1988)
9. Yin, W., Osher, S.: Error forgetting of Bregman iteration. *Journal of Scientific Computing* **54**(2–3), 684–695 (2013)