Representing exact viscosity solutions to high dimensional Hamilton–Jacobi PDEs using neural network architectures

Tingwei Meng (joint work with J. Darbon and G. P. Langlois)

Division of Applied Mathematics, Brown University.

Context and Motivations

Our goal is to solve high dimensional Hamilton–Jacobi (HJ) PDEs in the following form (the Hamiltonian H is state independent)

$$\begin{cases} \frac{\partial S}{\partial t}(x,t) + H(\nabla_x S(x,t)) = 0 & x \in \mathbb{R}^n, t \in (0,+\infty), \\ S(x,0) = J(x) & x \in \mathbb{R}^n. \end{cases} \quad (1)$$

We use the approach of neural networks (NNs), because its ability of high performance computation:

- many dedicated hardware implementations for neural networks (e.g., Xilinx AI, Intel AI and many startup companies)
- high performance (high throughput/low latency)
- low energy

We want to answer the following questions

- Can we leverage these computational resources for solving high dimensional HJ PDEs?
- Which class of HJ PDEs can be solved exactly using neural networks?
- Possible interpretations for certain neural networks using HJ PDEs?

The first neural network architecture [1]

Define a neural network (depicted in Fig. 1) with parameters $p_i \in \mathbb{R}^n, \theta_i \in \mathbb{R}, \gamma_i \in \mathbb{R}$ and max pooling activation function as follows

$$f(x,t) = \max_{i\in\{1,\dots,m\}} \{\langle p_i, x\rangle - t\theta_i - \gamma_i\} \quad x \in \mathbb{R}^n, t \geq 0. \quad (2)$$

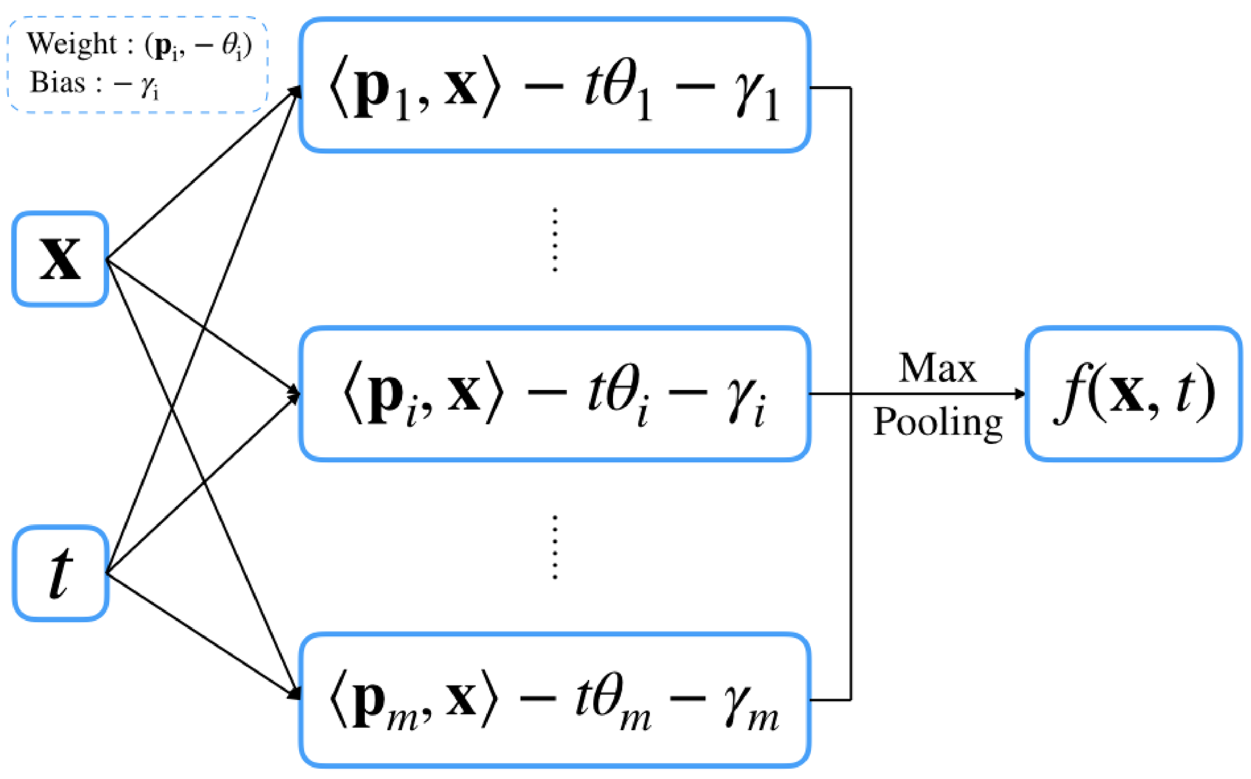

Figure 1: A neural network architecture (2) for solving certain HJ PDEs.

The following result is proved in [1].

- Assumption: Certain assumptions are imposed on $\{(p_i,\theta_i,\gamma_i)\}$. The most restricted one assumes that $\{(p_i,\gamma_i)\}$ satisfies

(A) there exists a convex function $g\colon \mathbb{R}^n \to \mathbb{R}$ such that $g(p_i) = \gamma_i$.

- Definition: the functions $J, H\colon \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ are defined using the parameters as follows

$$J(x) := \max_{i\in\{1,\dots,m\}} \{\langle p_i, x\rangle - \gamma_i\},$$
$$H(p) := \begin{cases} \inf_{\alpha\in\mathcal{A}(p)} \left\{\sum_{i=1}^m \alpha_i\theta_i\right\}, & \text{if } p \in \operatorname{dom} J^*, \\ +\infty, & \text{otherwise,} \end{cases} \quad (3)$$

where the constraint set $\mathcal{A}(p)$ is defined as the set containing all α in the simplex set in $\mathbb{R}^m$ satisfying $\sum_{i=1}^m \alpha_i(p_i,\gamma_i) = (p, J^*(p))$.

- Conclusion: The function f represented by the neural network (2) is the unique uniform continuous viscosity solution to the HJ PDE (1) with Hamiltonian $\tilde{H}$ and initial data J if and only if $\tilde{H}$ satisfies

$$\begin{cases} \tilde{H}(p_i) = H(p_i) \text{ for each } i \in \{1,\dots,m\}, \\ \tilde{H}(p) \geqslant H(p) \text{ for every } p \in \operatorname{dom} J^*. \end{cases} \quad (4)$$

Remarks:

- f solves the HJ PDE (1) with H and J defined in (3).
- We characterize all the HJ PDEs in the form of (1) which can be solved exactly using the neural network architecture in Fig. 1.

The second neural network architecture [2]

Define a neural network (depicted in Fig. 2) with parameters $u_i \in \mathbb{R}^n$ and $a_i \in \mathbb{R}$, and activation function $L\colon \mathbb{R}^n \to \mathbb{R}$ by

$$f_1(x,t) = \min_{i\in\{1,\dots,m\}} \left\{ tL\left(\frac{x-u_i}{t}\right) + a_i \right\} \quad x \in \mathbb{R}^n, t > 0. \quad (5)$$

Assuming L is convex and Lipschitz, then f_1 is a viscosity solution to the HJ PDE (1) with initial data J and Hamiltonian H defined by

$$J(x) := \min_{i\in\{1,\dots,m\}} \{L'_\infty(x-u_i) + a_i\}, \quad H := L^*, \quad (6)$$

where L'_∞ denotes the asymptotic function of L.

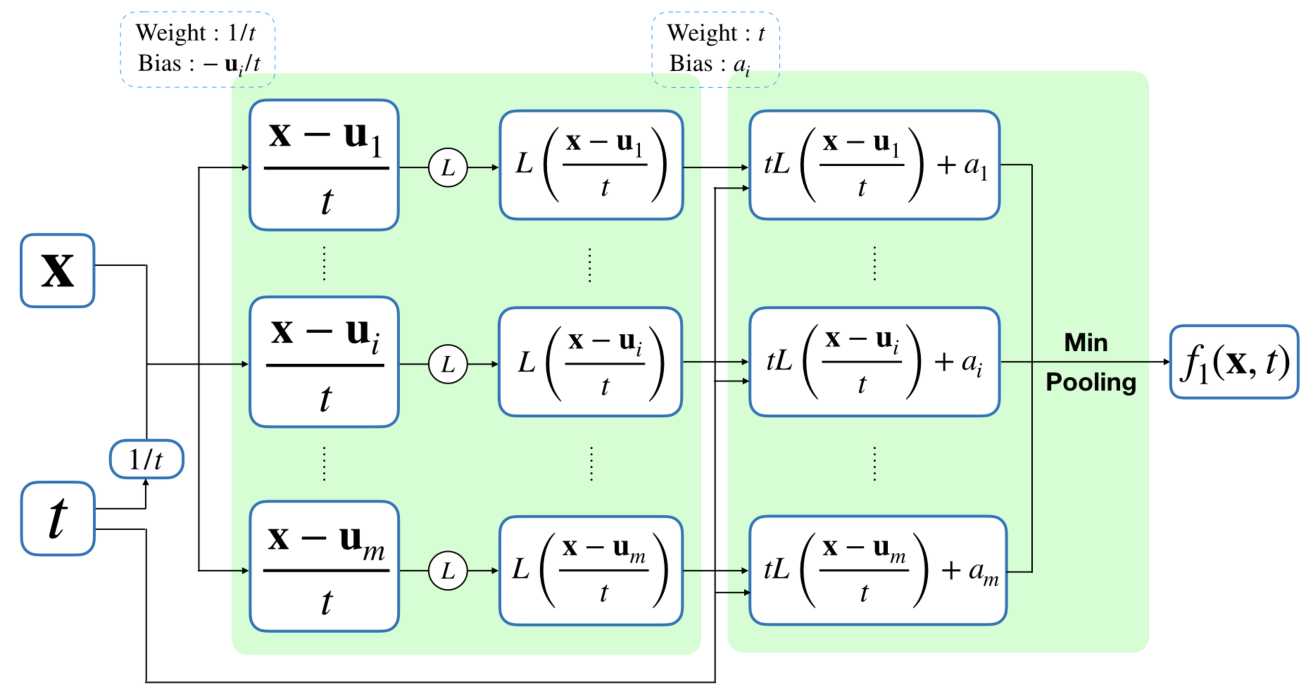

Figure 2: A neural network architecture (5) for solving certain HJ PDEs.

The third neural network architecture [2]

Define a neural network (depicted in Fig. 3) with parameters $v_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$ and activation function $\tilde{J}\colon \mathbb{R}^n \to \mathbb{R}$ as follows

$$f_2(x,t) = \min_{i\in\{1,\dots,m\}} \left\{ \tilde{J}(x - tv_i) + tb_i \right\} \quad x \in \mathbb{R}^n, t \geq 0. \quad (7)$$

Assume $\tilde{J}\colon \mathbb{R}^n \to \mathbb{R}$ is concave, and $\{(v_i,b_i)\}$ satisfies (A). Then f_2 is a viscosity solution to the HJ PDE (1) with J and H defined by

$$J = \tilde{J}, \quad H(p) = \max_{i\in\{1,\dots,m\}} \{\langle p, v_i\rangle - b_i\}. \quad (8)$$

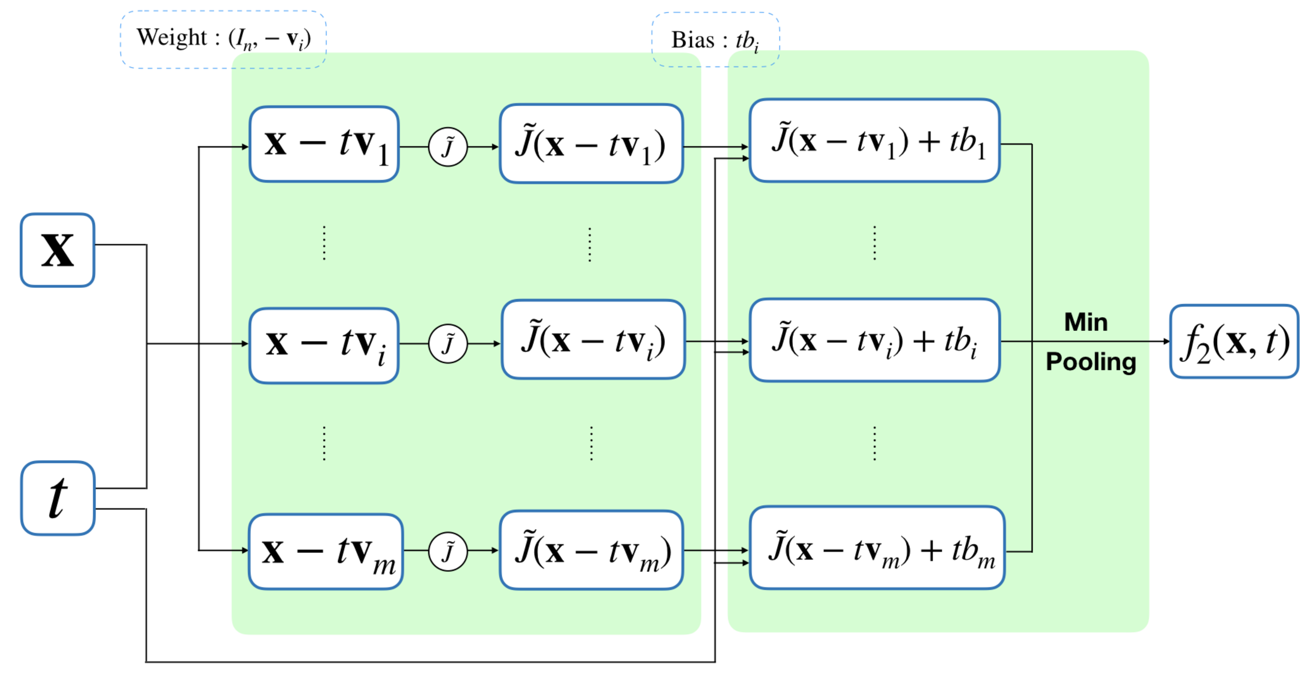

Figure 3: A neural network architecture (7) for solving certain HJ PDEs.

Conclusion

In the following table, we summarize the type of HJ PDEs which can be solved using our proposed NNs.

	NN in (2) (Fig. 1)	NN in (5) (Fig. 2)	NN in (7) (Fig. 3)
J	convex and piecewise affine	specific form in (6)	concave
H	satisfies (4) (e.g., concave or piecewise affine in $\operatorname{dom} J^*$)	convex with bounded domain	convex and piecewise affine

References

[1] Darbon, J., Langlois, G. P., & Meng, T. (2019). Overcoming the curse of dimensionality for some Hamilton–Jacobi partial differential equations via neural network architectures. https://arxiv.org/abs/1910.09045

[2] Darbon, J., & Meng, T. (2020). On some neural network architectures that can represent viscosity solutions of certain high dimensional Hamilton–Jacobi partial differential equations. https://arxiv.org/abs/2002.09750

 Work supported by NSF DMS 1820821. April 20, 2020