Proceedings of Machine Learning Research vol 291:1–5, 2025 38th Annual Conference on Learning Theory

Open Problem: Fixed-Parameter Tractability of Zonotope Problems

Vincent Froese VINCENT.FROESE@TU-BERLIN.DE
Technische Universität Berlin, Faculty IV, Institute of Software Engineering and Theoretical Computer Science, Algorithmics and Computational Complexity.

Moritz Grillo GRILLO@MATH.TU-BERLIN.DE
Technische Universität Berlin, Faculty II, Institute of Mathematics,
Combinatorial Optimization and Graph Algorithms.

Christoph Hertrich CHRISTOPH.HERTRICH@UTN.DE
University of Technology Nuremberg, Department Liberal Arts and Social Sciences,
Applied Discrete Mathematics Lab.

Martin Skutella MARTIN.SKUTELLA@TU-BERLIN.DE
Technische Universität Berlin, Faculty II, Institute of Mathematics,
Combinatorial Optimization and Graph Algorithms.

Editors: Nika Haghtalab and Ankur Moitra

Abstract

Neural networks with ReLU activation play a key role in modern machine learning. Understanding the functions represented by ReLU networks is a major topic in current research. Recent results are achieved via connections to tropical geometry based on a duality between convex piecewise linear functions and polytopes. It turns out that several questions about properties of functions computed by ReLU neural networks can be answered by solving certain problems on special polytopes called zonotopes. For example, computing the Lipschitz constant of a ReLU network with one hidden layer corresponds to norm maximization over a zonotope. Moreover, deciding whether the ReLU network attains a positive output is equivalent to zonotope non-containment. These problems are known to be NP-hard in general but polynomial-time solvable if the input dimension is constant. However, it is open whether they are *fixed-parameter tractable* (FPT) with respect to the input dimension d, that is, solvable in $f(d)\cdot n^{O(1)}$ time for some function f solely depending on d. Notably, these zonotope problems also arise in other areas such as robotics and control, reachability analysis, pattern recognition, signal processing or political analysis. Thus, settling their parameterized complexity status is of broad interest.

1. Introduction

A ReLU neural network with two layers (that is, one hidden layer with n ReLU neurons without bias) computes a continuous piecewise linear (CPWL) map

$$f\colon \mathbb{R}^d \to \mathbb{R}, \quad f(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i \max(0, \langle \mathbf{w}_i, \mathbf{x}\rangle),$$

where $\mathbf{w}_i \in \mathbb{R}^d$ are the weights of the hidden neurons and $\alpha_i \in \{1, -1\}$ are the output weights. The linear regions of f are determined by a (central) hyperplane arrangement $\mathcal{H}$ in $\mathbb{R}^d$ defined by the corresponding hyperplanes $H_i := \{\mathbf{x} \in \mathbb{R}^d \mid \langle \mathbf{w}_i, \mathbf{x}\rangle = 0\}$. Within a region R of $\mathcal{H}$, it holds $f(\mathbf{x}) = \sum_{i\in A_R} \alpha_i\langle \mathbf{w}_i, \mathbf{x}\rangle$, where $A_R \subseteq [n]$ is the set of *active* neurons (a neuron is active in R, if $\langle \mathbf{w}_i, \mathbf{x}\rangle > 0$ holds for all $\mathbf{x} \in R$). If all output weights α_i are 1, then f is convex.

There is a well-known duality between convex piecewise linear functions and polytopes originating from tropical geometry. Let $\mathcal{F}_d$ be the set of convex piecewise linear and positively homogeneous functions from $\mathbb{R}^d$ to $\mathbb{R}$ and let $\mathcal{P}_d$ be the set of polytopes in $\mathbb{R}^d$. For every $f \in \mathcal{F}_d$, there are $\{\mathbf{a}_i \in \mathbb{R}^d\}_{i\in I}$ such that $f(\mathbf{x}) = \max_{i\in I}\langle \mathbf{a}_i, \mathbf{x}\rangle$ and there is a bijection $\varphi\colon \mathcal{F}_d \to \mathcal{P}_d$ given by $\varphi\big(\max_{i\in I}\langle \mathbf{a}_i, \mathbf{x}\rangle\big) = \operatorname{conv}\{\mathbf{a}_i \mid i \in I\}$ where the inverse is the *support function* $\varphi^{-1}\colon \mathcal{P}_d \to \mathcal{F}_d$ given by $\varphi^{-1}(P)(\mathbf{x}) = \max\big\{\langle \mathbf{x}, \mathbf{y}\rangle \mid \mathbf{y} \in P\big\}$. Furthermore, φ is a semi-ring isomorphism between the semi-rings $(\mathcal{F}_d, \max, +)$ and $(\mathcal{P}_d, \operatorname{conv}, +)$, where $+$ is either the pointwise addition or the Minkowski sum, respectively (see, e.g., Zhang et al. (2018) for more details on this correspondence).

A *zonotope* is a Minkowski sum of line segments, that is, given a *generator* matrix $\mathbf{G} \in \mathbb{R}^{d\times n}$ with columns vectors $\mathbf{g}_i \in \mathbb{R}^d$, the corresponding zonotope is given by

$$Z(\mathbf{G}) \coloneqq \Big\{\mathbf{x} \in \mathbb{R}^d \mid \mathbf{x} \in \sum_{i\in[n]} \operatorname{conv}\{\mathbf{0}, \mathbf{g}_i\}\Big\},$$

where the Minkowski sum of two sets $P, Q \subseteq \mathbb{R}^d$ is defined as $P + Q \coloneqq \{p + q \mid p \in P, q \in Q\}$. Alternatively, a zonotope is the image of a hypercube under a linear transformation. Another equivalent definition is $Z(\mathbf{G}) = \operatorname{conv}\{\sum_{i\in S}\mathbf{g}_i \mid S \subseteq [n]\}$. Zonotopes are well-studied in polyhedral theory (Ziegler (2012); Fukuda (2020)) and appear in various application domains.

Based on the bijection φ, we can translate functions computed by 2-layer ReLU neural networks to zonotopes. In the following, we describe two applications together with the corresponding open questions in more detail.

2. Norm Maximization over Zonotopes

An important concept in deep learning is the Lipschitz constant. The Lipschitz constant (for some L_p-norm) of a function f is

$$L_p(f) \coloneqq \sup_{\mathbf{x}\neq\mathbf{y}} \frac{\|f(\mathbf{x}) - f(\mathbf{y})\|_p}{\|\mathbf{x} - \mathbf{y}\|_p}.$$

Estimating the Lipschitz constant is a well-studied problem with applications in analyzing the robustness, generalization performance and fairness of ReLU neural networks (see e.g. references in (Virmaux and Scaman (2018); Jordan and Dimakis (2020))). Note that if $f\colon \mathbb{R}^d \to \mathbb{R}$ is computed by a 2-layer ReLU network, then its Lipschitz constant equals the maximum norm of the gradient of any linear region (since f is CPWL) (Virmaux and Scaman (2018)). Computing the Lipschitz constant was shown to be NP-hard for the L_2-norm (Virmaux and Scaman (2018)) and also inapproximable in polynomial time for the L_1- and L_∞-norm (Jordan and Dimakis (2020)).

Now, if all output weights of the network computing f are 1 and hence f is convex and nonnegative, then the corresponding polytope $\varphi(f)$ is the zonotope $Z_f \coloneqq Z(\mathbf{W})$ where $\mathbf{W} \in \mathbb{R}^{d\times n}$ are the hidden weights of the neural network. By a well-known duality between zonotopes and hyperplane arrangements, the vertices of Z_f one-to-one correspond to the regions of the hyperplane arrangement defined by $\mathbf{W}$. That is, each vertex of Z_f is the sum $\sum_{i\in A_R}\mathbf{w}_i$ (that is, the gradient) of some region R. Hence, maximizing the L_p-norm over Z_f yields $L_p(f)$. Note that Bodlaender et al. (1990) already showed that L_p-Maximization is NP-hard for every $p \in \mathbb{N}$ even on parallelotopes (zonotopes in $\mathbb{R}^d$ with d linear independent generators).

For the L_2-norm, the problem is also known as (Unconstrained) Quadratic Binary Maximization in mathematical programming. Formally, the problem is to maximize $x^T Q x$ subject to $x \in \{0,1\}^n$, where Q is an $n \times n$ rational symmetric matrix. This is a classic NP-hard combinatorial optimization problem (e.g. modeling Weighted Max-Cut). If Q is positive semi-definite and has rank d, then this problem corresponds exactly to L_2-maximization over zonotopes in $\mathbb{R}^d$ as shown by Ferrez et al. (2005) and is thus solvable in $O(n^{d-1})$ time (by enumerating all vertices of the zonotope). The problem is also known as the Longest Vector Sum problem in the context of pattern recognition and signal processing (Pyatkin (2010); Shenmaier (2018, 2020)). Notably, NP-hardness (Pyatkin (2010)), inapproximability and also $O(n^{d-1})$ time (Shenmaier (2020)) have been proven here seemingly unaware of the connection to zonotopes. However, the parameterized complexity regarding the parameter input dimension d remained unresolved and was stated as an open question by Shenmaier (2020), who also remarked that fixed-parameter tractability (that is, $f(d)n^{O(1)}$ time) holds for the L_1-norm (since this is a polytopal norm with 2^d facets).

Open Question: Is L_p-Maximization over d-Zonotopes fixed-parameter tractable with respect to d for $p > 1$?

It is known that L_p-Maximization on (halfspace represented) polytopes is W[1]-hard (that is, presumably not FPT) with respect to d for $p > 1$ and fixed-parameter tractable for $p = 1$ (Knauer et al. (2015)). Since zonotopes are special polytopes, fixed-parameter tractability might still be possible. Note, however, that the zonotope is given by its generators, which is a more compact representation that can be exponentially smaller.

3. Zonotope (Non)-Containment

Every function $f\colon \mathbb{R}^d \to \mathbb{R}$ computed by a 2-layer ReLU neural network can be associated with two zonotopes Z^+ and Z^- generated by the weights of the hidden neurons with positive, respectively, negative output weights. A fundamental question is whether there exists an input $\mathbf{x} \in \mathbb{R}^d$ such that $f(\mathbf{x}) > 0$. This problem is called 2-Layer ReLU Positivity and is essentially equivalent to deciding whether f is surjective (Froese et al. (2024)). It is NP-hard since it is equivalent to the question whether $Z^+ \nsubseteq Z^-$, which is NP-hard since Zonotope Containment (that is, deciding whether $Z_1 \subseteq Z_2$ for two zonotopes given via generators) is coNP-hard (Kulmburg and Althoff (2021)). Zonotope Containment is a well-studied problem with various applications in robotics and control (see references in (Kulmburg and Althoff (2021)). It is solvable in $O(n^{d-1})$ time (essentially by enumerating the vertices of Z_1) but unknown to be fixed-parameter tractable with respect to d (Froese et al. (2024)).

Open Question: Is d-Zonotope Containment fixed-parameter tractable with respect to d?

Note that Kulmburg and Althoff (2021) showed that the zonotope containment problem is equivalent to maximizing a certain zonotope norm over a zonotope. Hence, containment is also a special case of norm maximization (alternatively, L_p-maximization is the question of containment in an L_p-ball). We remark that for general polytopes P and Q, deciding whether $P \subseteq Q$ is known to be coNP-hard if P is given by halfspaces and Q by its vertices (all other cases of representations are polynomial-time solvable) (Kaibel and Pfetsch (2003)). To the best of our knowledge, fixed-parameter tractability with respect to d is also open for the coNP-hard case of polytope containment.

4. Conclusion

We discussed several open problems regarding the parameterized complexity of certain zonotope problems which appear as fundamental problems in various areas in computer science. As regards the applications in deep learning, one can naturally generalize the problems of computing the Lipschitz constant and deciding positivity to more than one hidden layer. With multiple hidden layers, we expect it to be easier to exclude fixed-parameter tractability. However, the problems then do not correspond to zonotopes anymore, but to more complex polytopes (see e.g. (Hertrich et al. (2021); Hertrich (2022))) represented via ReLU neural networks, which can be seen as extended formulations (Hertrich and Loho (2024)).

References

Hans L. Bodlaender, Peter Gritzmann, Victor Klee, and Jan van Leeuwen. Computational complexity of norm-maximization. *Combinatorica*, 10(2):203–225, 1990.

J.-A. Ferrez, Komei Fukuda, and Thomas M. Liebling. Solving the fixed rank convex quadratic maximization in binary variables by a parallel zonotope construction algorithm. *European Journal of Operations Research*, 166(1):35–50, 2005.

Vincent Froese, Moritz Grillo, and Martin Skutella. Complexity of deciding injectivity and surjectivity of ReLU neural networks. *CoRR*, abs/2405.19805, 2024. Accepted for publication at COLT 2025.

Komei Fukuda. *Polyhedral Computation*. Department of Mathematics, Institute of Theoretical Computer Science ETH Zurich, 2020.

Christoph Hertrich. *Facets of neural network complexity*. PhD thesis, Technical University of Berlin, Germany, 2022.

Christoph Hertrich and Georg Loho. Neural networks and (virtual) extended formulations. *CoRR*, abs/2411.03006, 2024.

Christoph Hertrich, Amitabh Basu, Marco Di Summa, and Martin Skutella. Towards lower bounds on the depth of ReLU neural networks. In *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021*, pages 3336–3348, 2021.

Matt Jordan and Alexandros G. Dimakis. Exactly computing the local lipschitz constant of ReLU networks. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020*, 2020.

Volker Kaibel and Marc E. Pfetsch. Some algorithmic problems in polytope theory. In *Algebra, Geometry, and Software Systems [outcome of a Dagstuhl seminar]*, pages 23–47. Springer, 2003.

Christian Knauer, Stefan König, and Daniel Werner. Fixed-parameter complexity and approximability of norm maximization. *Discretete & Computational Geometry*, 53(2):276–295, 2015.

Adrian Kulmburg and Matthias Althoff. On the co-NP-completeness of the zonotope containment problem. *European Journal of Control*, 62:84–91, 2021.

A. V. Pyatkin. On complexity of a choice problem of the vector subset with the maximum sum length. *Journal of Applied and Industrial Mathematics*, 4:549–552, 2010.

Vladimir Shenmaier. Complexity and approximation of finding the longest vector sum. *Computational Mathematics and Mathematical Physics*, 58(6):850–857, 2018.

Vladimir Shenmaier. Complexity and algorithms for finding a subset of vectors with the longest sum. *Theoretical Computer Science*, 818:60–73, 2020.

Aladin Virmaux and Kevin Scaman. Lipschitz regularity of deep neural networks: analysis and efficient estimation. In *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018*, pages 3839–3848, 2018.

Liwen Zhang, Gregory Naitzat, and Lek-Heng Lim. Tropical geometry of deep neural networks. In *Proceedings of the 35th International Conference on Machine Learning (ICML'18)*, volume 80 of *PMLR*, pages 5824–5832, 2018.

Günter M. Ziegler. *Lectures on Polytopes*. Springer, 2012.