Are Neural Networks Collision Resistant?

Marco Benedetti,[1] Andrej Bogdanov,[2] Enrico M. Malatesta,[1] Marc Mézard,[1] Gianmarco Perrupato,[1] Alon Rosen,[1] Nikolaj I. Schwartzbach,[1] and Riccardo Zecchina[1]

[1]*Department of Computing Sciences, Bocconi University, Milano, Italy*
[2]*University of Ottawa, Ottawa, Ontario, Canada*

(Received 12 November 2025; accepted 23 April 2026; published 5 June 2026)

Collision-resistant hash functions are a fundamental cryptographic primitive that rely on the computational hardness of finding two inputs that produce the same output. Motivated by this problem, we study the complexity of finding collisions in a family of neural networks with oscillating activation functions. A neural network trained on a classification task is specified by a set of weights assigning a label to each data point, and a collision is defined as two distinct weight configurations that produce the same labeling. We show that, within this class of neural networks, the space of collisions exhibits an overlap gap property, whereby certain overlap values between distinct solutions are forbidden. This property is a geometric feature of the solution landscape in high-dimensional random constraint satisfaction problems that has recently emerged as a powerful indicator of algorithmic barriers. Our analysis predicts a regime in which efficient algorithms fail to find collisions. This prediction is supported by numerical experiments using approximate message passing algorithms, which cease to return collisions well below the threshold predicted by theory. Neural networks, therefore, provide a class of candidate collision-resistant functions that, for suitable parameter choices, depart from existing constructions based on lattices. Beyond their cryptographic relevance, our results reveal forms of computational hardness in large neural networks that may be of independent interest.

DOI: 10.1103/6shh-9h5m

Subject Areas: Complex Systems, Statistical Physics

I. INTRODUCTION

Neural networks are models of computation at the core of modern machine learning. In feedforward networks used as classifiers, given an input vector of size N, the network computes an output by alternating layers of linear transformations based on "weight" matrices with nonlinear activation functions. Arguably, the simplest neural network is the *binary perceptron* [1,2], consisting of a single layer and binary weights $\mathbf{x} \in \{-1,1\}^N$. Given a matrix of P data points in N-dimensional space, $\mathbf{A} \in \mathbb{R}^{P\times N}$ (also called the *disorder*), the perceptron outputs the labels $f_{\mathbf{A}}(\mathbf{x}) = \varphi(\mathbf{A}\mathbf{x})$, where the activation function $\varphi : \mathbb{R} \to \{-1,1\}$ is applied elementwise.

Here, we consider the perceptron from a dual point of view. Given the disorder $\mathbf{A}$, we use as inputs the weights $\mathbf{x}$ and study the function $f_{\mathbf{A}}(\mathbf{x})$ in a regime of large dimensions where $N, P \to \infty$ with fixed $\alpha = P/N$. Several constraint satisfaction problems (CSPs) arise in this context, listed in order of nonincreasing algorithmic difficulty:

(i) *Inversion*—Given a matrix $\mathbf{A} \in \mathbb{R}^{P\times N}$ and a vector of labels $\boldsymbol{\ell} \in \{-1,1\}^P$, find any set of inputs $\mathbf{x} \in \{-1,1\}^N$ such that $f_{\mathbf{A}}(\mathbf{x}) = \boldsymbol{\ell}$.

(ii) *Second preimage*—Given a matrix $\mathbf{A} \in \mathbb{R}^{P\times N}$ and a "teacher" $\mathbf{x} \in \{-1,1\}^N$ generating labels $\boldsymbol{\ell} = f_{\mathbf{A}}(\mathbf{x}) \in \{-1,1\}^P$, find any "student" $\mathbf{x}' \in \{-1,1\}^N$ with $\mathbf{x}' \neq \mathbf{x}$ such that $f_{\mathbf{A}}(\mathbf{x}') = \boldsymbol{\ell}$.

(iii) *Collision finding*—Given a matrix $\mathbf{A} \in \mathbb{R}^{P\times N}$, find any two inputs $\mathbf{x} \neq \mathbf{x}' \in \{-1,1\}^N$ such that $f_{\mathbf{A}}(\mathbf{x}) = f_{\mathbf{A}}(\mathbf{x}')$.

We are interested in the typical case complexity of CSPs, where the matrix $\mathbf{A}$ has random identically distributed independent entries, and also the labels are identically distributed independent random in the inversion problem. Statistical physics has a long tradition of studying random CSPs, such as inversion and second-preimage variants [3–6], yet the collision-finding problem itself had not been addressed within this framework until now. Notice the difference between second preimage and collision finding: In the latter, there is full freedom to choose both $\mathbf{x}$ and $\mathbf{x}'$ in a way that might depend on $\mathbf{A}$, while, in the former, one of the two is fixed. A function f for which there are no polynomial time algorithms that find collisions [7] is said to be collision resistant.

The collision-finding problem, on the other hand, has been extensively studied in cryptography. A function for

which the collision-finding problem is hard is said to be *collision resistant*. If, in addition, the function is *shrinking*, i.e., the output size P is smaller than the input size N, it is known as a *collision-resistant hash function* (CRH). CRHs are of fundamental importance and form the basis of many cryptographic protocols and security guarantees. In fact, they are an integral part to securing data privacy [8], in applications including commitment schemes [9], encryption [10], and secure computation [11]. In this paper, our main goal is to understand whether the collision-finding problem in neural networks with a suitably chosen activation function is algorithmically hard.

Statistical physics studies on random Boolean satisfiability [4,12–14] have linked typical-case algorithmic hardness to the presence of peculiar geometrical properties in the space of solutions of CSPs. Building upon this geometric approach [15–18], it has been rigorously shown that the presence of a disconnectivity feature known as the overlap gap property (OGP) in the solution space of CSPs leads to the failure of stable algorithms [19,20] and larger classes including online algorithms and low-degree polynomials. In many CSPs, the most efficient solvers identified to date belong to this class, giving rise to the conjecture that OGP is a marker of computational hardness [21].

Examples are the symmetric binary perceptron problem [22], number partitioning [23], maximum independent set on random graphs [24], and the optimization of p-spin glass Hamiltonians [25,26].

Our main result is the identification of a novel candidate for collision-resistant hash function based on a family of neural networks, where the OGP is used as a criterion for collision resistance between pairs of inputs of large Hamming distance. We then argue that, by applying an error-correcting code upstream of a perceptron with an oscillating activation, one obtains a function that remains shrinking and is collision resistant.

These new candidate CRHs are markedly different from existing constructions from, e.g., lattices and hint at a new source of hardness for use in cryptography.

II. SUMMARY OF OUR CONTRIBUTIONS

We investigate the collision resistance of a family of neural networks obtained by composing an error-correcting code (ECC) with a perceptron equipped with a particular choice of activation function. Specifically, we consider the following activation, which we refer to as the square-wave perceptron (SWP) [27]:

$$\varphi_\delta(h) = \operatorname{sgn}\left(\frac{h}{2\delta} \bmod 1\right), \tag{1}$$

where $x \bmod 1$ is the unique real number in $[-\frac{1}{2},\frac{1}{2})$ that differs from x by an integer and $\delta \in \mathbb{R}^+$ is a parameter that controls the oscillation period. The function (1) is a square wave with period 2δ. The choice of activation is ultimately guided by the goal of obtaining a shrinking collision resistant function which, as we show, is crucially dependent on δ.

Some properties of the SWP have been studied in [28] in average-case variants of inversion and second-preimage problems. Here, we focus on the collision problem. In Sec. III, we study the existence of collisions in the large size limit, in the regime where the inputs are at an extensive Hamming distance or, equivalently, when the overlap $q_1 = \mathbf{x}^\top \mathbf{y}/N$ between the two inputs $\mathbf{x}, \mathbf{y} \in \{-1, 1\}^N$ forming a collision is strictly smaller than one, $|q_1| \le 1 - \Omega(1)$. We show that above a certain threshold of α, depending on q_1 and δ, with high probability collisions do not exist. The value of this threshold computed by using the replica method with a replica-symmetric ansatz [29] coincides with a first-moment computation, leading us to conjecture that the replica-symmetric estimate coincides with the actual satisfiable (SAT) or unsatisfiable (UNSAT) transition in the collision problem. This conjecture is supported by numerical simulations based on exhaustive search of collisions on finite-size systems.

In Sec. IV, we move to the main focus of the paper, namely, determining if there is a region of α where collisions exist but are hard to find. Using the first-moment method we show that, conditioning to $|q_1| < 1$, the space of solutions of the collision-finding problem presents an OGP, which implies the failure of stable algorithms. We conjecture that, in this regime, finding collisions with $|q_1| < 1$ is infeasible.

By tuning the parameter δ, we show that the conjectured hard region extends to values of α smaller than one, which is a requirement to construct a hash function. More precisely, we show that there exist α_*, δ_*, and $q_* < 1$ such that for $\delta < \delta_*$ the space of collisions with overlap $q_1 < q_*$ exhibits OGP for $\alpha_\star < \alpha < 1$. Still, a solver could target collisions with overlap $q_* < q < 1$, where our analysis does not imply collision resistance for $\alpha < 1$. To remedy this, in Sec. V, we propose a collision-resistant hash function obtained by composing an error-correcting code with the square-wave perceptron. The error-correcting code removes collisions with overlaps in the region $q_* < q < 1$, leading to a bona fide collision-resistant hash function.

In Sec. VI, we provide some numerical evidence supporting our conjectured CRHs. We study numerically two algorithmic attacks to find collisions, showing that both of them stop working at values of α smaller than the value α_*, where we predict the presence of an OGP.

III. EXISTENCE OF COLLISIONS

Consider the number of collisions with internal overlap q_1:

$$Z(q_1; \boldsymbol{A}) \coloneqq \sum_{\boldsymbol{c} \in \{\pm 1\}^{2N}} \mathbb{X}_{\boldsymbol{A}}(\boldsymbol{c}) \delta(q_i(\boldsymbol{c}) - q_1), \tag{2}$$

where $\mathbb{X}_{\boldsymbol{A}}(\boldsymbol{c})$ is an indicator function, equal to one if the pair of inputs $\boldsymbol{c} = (\mathbf{x}, \mathbf{y})$ forms a collision and zero otherwise:

$$\mathbb{X}_{\mathbf{A}}(\boldsymbol{c}) = \prod_{\mu=1}^{P} \Theta[f_{\mathbf{A}}(\mathbf{x})_\mu f_{\mathbf{A}}(\mathbf{y})_\mu]. \tag{3}$$

In this notation, $\Theta(\cdot)$ is the Heaviside step function and $f_{\mathbf{A}}(\cdot)_\mu$ the μth element of vector $f_{\mathbf{A}}(\cdot)$. We have also defined the "*internal*" overlap of a collision $\boldsymbol{c}$ as the normalized scalar product between the colliding inputs

$$q_i(\boldsymbol{c}) = \frac{1}{N} \mathbf{x}^\top \mathbf{y}. \tag{4}$$

In the limit $N, P \to \infty$ with $\alpha = P/N$ fixed, the random variable $1/N \log Z$ concentrates on its average value, the so-called *quenched* free entropy [30]:

$$\Phi(q_1) = \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\boldsymbol{A}} \log Z(q_1; \boldsymbol{A}). \tag{5}$$

Hence, $Z(q_1;\boldsymbol{A}) = \exp[N\Phi(q_1) + o(N)]$, and if $\Phi < 0$, one has $Z(\boldsymbol{A}) = 0$ with high probability in the large-N limit. Therefore, collisions with internal overlap q_1 exist up to a threshold $\alpha_c(q_1)$ where $\Phi = 0$. We estimated $\alpha_c(q_1)$ using two complementary approaches. First, we derived an upper bound $\alpha_c^a(q_1)$ to $\alpha_c(q_1)$ by using the first-moment method. Because of Markov's inequality, the probability that the random variable $Z(q_1;\boldsymbol{A}) \geq 1$ is bounded by

$$P[Z(q_1;\boldsymbol{A}) \geq 1] \leq \mathbb{E}_{\boldsymbol{A}} Z(q_1;\boldsymbol{A}) = e^{N\Phi^{\mathrm{a}}(q_1)}, \tag{6}$$

where the *annealed* free entropy $\Phi^{\mathrm{a}}(q_1)$ is given by

$$\Phi^{\mathrm{a}}(q_1) = \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}_{\boldsymbol{A}} Z(q_1;\boldsymbol{A}). \tag{7}$$

The value of α for which the annealed free entropy vanishes, therefore, gives an upper bound to $\alpha_c(q_1)$. As we show in Supplemental Material [31], this is given by

$$\alpha_c^{\mathrm{a}}(q_1) = -\frac{\log(2) + H_B(q_1)}{\log \int Dz F_\varphi(\sqrt{q_1}z; \sqrt{1-q_1})}, \tag{8}$$

where

$$H_B(q_1) \equiv -\frac{1-q_1}{2} \log\left(\frac{1-q_1}{2}\right) - \frac{1+q_1}{2} \log\left(\frac{1+q_1}{2}\right), \tag{9}$$

$$F_\varphi(x;\sigma) \equiv 1 - 2I_\varphi(x;\sigma)(1 - I_\varphi(x;\sigma)), \tag{10}$$

$$I_\varphi(x;\sigma) \equiv \int Dh \Theta[\varphi(\sigma h + x)]. \tag{11}$$

Second, we applied the replica method [29] to compute the replica symmetric (RS) upper bound $\Phi^{\mathrm{RS}} \geq \Phi$ to the quenched free entropy (5), and the corresponding upper bound $\alpha_c^{\mathrm{RS}} \geq \alpha_c$. Details can be found in Supplemental Material [31]. Here, we mention that previous applications of the replica method using the RS approximation [32] have already been shown to give the correct result for the storage capacity in inversion problems [33,34] with binary inputs, whereas for continuous inputs one needs replica symmetry breaking [35–38].

Interestingly, we found that the more refined RS computation of the quenched free entropy (5) does not change the result found using the first-moment method $\Phi^{\mathrm{RS}}(q_1) = \Phi^{\mathrm{a}}(q_1)$. This leads also to the same prediction for the value of α up to which collisions exist $\alpha_c^{\mathrm{RS}}(q_1) = \alpha_c^{\mathrm{a}}(q_1)$, suggesting that the annealed bound is tight and, in this case, exact.

Figure 1 displays α_c^{RS} as a function of q_1 for various values of δ. For $\alpha > \alpha_c^{\mathrm{RS}}(q_1)$, the number of collisions is exponentially small. For all δ strictly larger than zero, $\alpha_c(q_1)$ diverges in the limit $q_1 \to 1$ as $\sim(1-q_1)^{-1/2}$. This is to be expected: Trivial collisions with $q_1 = 1$, i.e., $\mathbf{x} = \mathbf{y}$, are always present. The divergence around $q_1 \sim 1$ means that subextensive collisions exist for all $\alpha = O(1)$ and independent of the activation function. In Supplemental Material [31], we show that the computation of the annealed free entropy for $q_1 \approx 1 - 1/N$ consistently predicts $\alpha_c^{\mathrm{a}} \propto N^{1/2}$.

Finally, notice that, for each value of q_1, decreasing δ leads to a smaller value of α_c^{RS}. In particular, in the small δ limit, which we call the *strong hashing limit*, one obtains

$$\alpha_c^{\mathrm{a}}(q_1) = 1 + \frac{H_B(q_1)}{\log 2}, \tag{12}$$

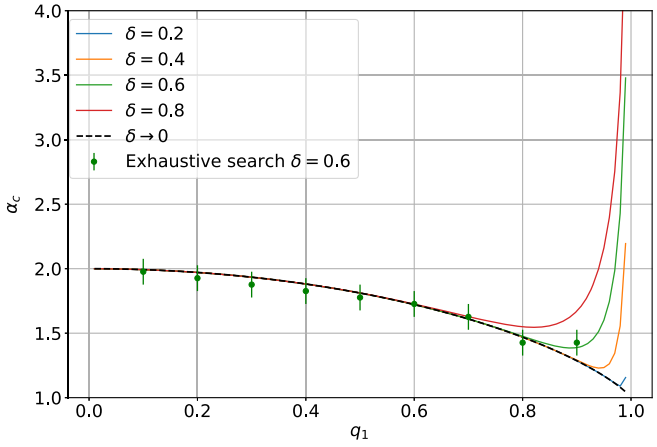

FIG. 1. Full lines: analytic prediction for the transition α_c, above which no collision exists versus the internal overlap q_1, for different values of δ. Dots: numerical estimates from exhaustive search.

which is a monotonically decreasing function of q_1. This curve is shown as black dashed in Fig. 1. Dots in Fig. 1 present a numerical estimate of α_c, obtained by exhaustively searching the space of weights and enumerating how many collisions exist with a given value of q_1 (see Supplemental Material [31] for details). The procedure takes an exponential time in N, so the estimate is based on data with small $N = 14, 16, \ldots, 24$. Still, finite-size effects are small, and the extrapolated values are in excellent agreement with the analytic prediction.

IV. ALGORITHMIC BARRIERS TO COLLISION FINDING

A. Multioverlap gap property

The overlap gap property is a geometric property of the solution space of CSPs, which is proven to imply the failure of stable algorithms. Informally, the stability of an algorithm $\mathcal{A}$, defined as a map from the instance space to the solution space of a CSP, requires that a small perturbation of the input results in a small perturbation of the output. In various CSPs, the best known solvers have been shown to belong to this class, fostering the belief that the presence of OGP is a signature of general algorithmic hardness [19]. Examples of algorithms proven to satisfy this stability property are the Kim-Roche algorithm for the symmetric binary perceptron [22], approximate message passing [25], low-degree polynomials, Langevin dynamics, and low-depth Boolean circuits [26,39].

In this section, we discuss a variant of OGP called multioverlap gap property (m-OGP) [19,20]. In order to define it, we need to introduce a distance between two arbitrary collisions. Take two arbitrary colliding pairs $\boldsymbol{c}_a, \boldsymbol{c}_b \in \{-1, 1\}^{2N}$; we define their "*external*" overlap as

$$q_e(\boldsymbol{c}_a, \boldsymbol{c}_b) \equiv \frac{1}{2N}(\mathbf{x}_a^\top \mathbf{x}_b + \mathbf{y}_a^\top \mathbf{y}_b). \tag{13}$$

The overlaps range from -1 to 1. Let $m \geq 2$ be an integer. We say that the function $f_{\mathbf{A}}(\mathbf{x}) : \mathbb{R}^N \to \mathbb{R}^P$ exhibits m-OGP for the collision-finding problem, with parameters $-1 \leq q_1 \leq 1$ and $-1 \leq \varepsilon_1 < \varepsilon_2 \leq 1$ if, given a random $\mathbf{A}$, it holds with high probability that there is no set of m pairs $\boldsymbol{c}_a = (\mathbf{x}_a, \mathbf{y}_a)$, $a = 1, \ldots, m$, such that the following conditions are met:

(1) $f_{\mathbf{A}}(\mathbf{x}_a) = f_{\mathbf{A}}(\mathbf{y}_a)$ with $\mathbf{x}_a \neq \mathbf{y}_a$, for every $a = 1 \ldots m$;
(2) $q_i(\boldsymbol{c}_a) = q_1$, for every $a = 1, \ldots, m$;
(3) $\varepsilon_1 \leq q_e(\boldsymbol{c}_a, \boldsymbol{c}_b) \leq \varepsilon_2$ for all $a \neq b$.

The first condition imposes that each $\boldsymbol{c}_a$ is a collision. The parameter q_1 in the second condition controls the internal overlap of each collision as defined in Eq. (4); the third condition imposes a gap on the external overlap between each collision pair. We define, respectively, $q_1 = O(1) < 1$ and $q_1 = 1 - o(1)$ as the "extensive" and "subextensive" regimes.

It is important to note that obstruction to stable algorithms does not require a specific value of m [19]. In other words, independent of $m \geq 2$, $\alpha^m_{\mathrm{OGP}}(q_1)$ serves as an upper bound on the values of α where stable algorithms can find collisions with internal overlap q_1. Furthermore, if $m' > m$, then m-OGP implies m'-OGP, meaning that larger values of m can provide stronger bounds on the region of algorithmic hardness.

The m-OGP has a simple geometric interpretation. It requires that, for every m-tuple of collisions with a certain internal overlap, there are at least two collisions that are "close" or "far apart" in the sense of definition (13).

In the following, we study the presence of m-OGP in the collision-finding problem for the function $f_{\mathbf{A}}$, in the limit where both the dimensions of the input N and the output P of the function scale to infinity, with their ratio fixed to $\alpha = P/N = O(1)$. In this high-dimensional limit, we show the existence of a region of α delimited by $\alpha^m_{\mathrm{OGP}}(q_1)$, such that for $\alpha > \alpha^m_{\mathrm{OGP}}(q_1)$ the m-OGP holds for the collision problem with internal overlap q_1. Moreover, $\alpha^m_{\mathrm{OGP}}(q_1)$ is strictly lower than the value $\alpha_c(q_1)$, above which no collisions with internal overlap q_1 can be found. We emphasize that the analytical formulas we derive in Supplemental Material [31] are valid for generic nonlinearity φ. However, in the main text, we focus specifically on the SWP activation function, due to its appealing cryptographic properties.

B. Existence of m-OGP in the collision-finding problem

The quantity of interest in establishing the presence of m-OGP is the number of m-tuples of collisions $\{\boldsymbol{c}_a = (\mathbf{x}_a, \mathbf{y}_a)\}_{a=1}^m$ with reciprocal external overlap p and internal overlap q_1:

$$\mathcal{N}_m(p, q_1; \boldsymbol{A}) \coloneqq \sum_{\{\boldsymbol{c}_a\}_{a=1}^m} \prod_{a=1}^m [\mathbb{X}_{\boldsymbol{A}}(\boldsymbol{c}_a)\delta(q_i(\boldsymbol{c}_a) - q_1)] \times \prod_{a<b} \delta(q_e(\boldsymbol{c}^a, \boldsymbol{c}^b) - p). \tag{14}$$

We say that the problem has an m-OGP when, for a given value of the internal overlap q_1, there exists a range of external overlaps p where $\mathcal{N}_m(p, q_1; \boldsymbol{A}) = 0$ with high probability over the realization of the disorder. Note that this is analogous to condition 3 in the definition of m-OGP (see Sec. IV A). The existence of such an excluded region can be proved by noticing that, since $\mathcal{N}_m(p, q_1; \mathbf{A})$ is a non-negative integer, Markov inequality gives

$$P[\mathcal{N}_m(p, q_1; \mathbf{A}) > 0] \leq \mathbb{E}_{\mathbf{A}} \mathcal{N}_m(p, q_1; \mathbf{A}) = e^{mN\phi^{\mathrm{a}}_m(p, q_1)}, \tag{15}$$

where the annealed entropy ϕ^{a}_m is defined as

$$\phi_m^{\mathrm{a}}(p,q_1)=\lim_{N\to\infty}\frac{1}{mN}\ln\mathbb{E}_{\boldsymbol{A}}\mathcal{N}_m(p,q_1;\boldsymbol{A}). \tag{16}$$

If, given q_1, for some value of α there exists an interval of p where $\phi_m^{\mathrm{a}}(p,q_1)\le 0$, then Eq. (15) proves that with high probability there are no collisions within such an external overlap range. We define $\alpha_{\mathrm{OGP}}^m(q_1)$ as the lowest value of α such that this condition is met. Notice that, because of Markov's inequality, $\alpha_{\mathrm{OGP}}^m(q_1)$ computed using the annealed entropy in Eq. (16) is only an upper bound to the value α_{OGP}, above which an excluded region actually exists. Note that the purpose of our calculation is to identify a regime where OGP is provably present rather than characterizing the exact location of the OGP threshold. In some models, this bound is known to be tight: A matching second-moment argument proves the absence of m-OGP below $\alpha_{\mathrm{ann}}^{\mathrm{OGP}}(q_1)$. This has been rigorously established in certain regimes for the Ising p-spin glass and for random k-SAT [40] as well as for the symmetric binary perceptron [22].

The computation of Eq. (16) is detailed in Supplemental Material [31]; here, we state only the final result. One finds that $\phi_m^a(p,q_1)$ can be computed by maximizing a function $\phi_m(Q)$ over a suitable space of matrices $Q\in\mathcal{F}(p,q_1)$:

$$\phi_m^a(p,q_1)=\max_{Q\in\mathcal{F}(p,q_1)}\phi_m(Q). \tag{17}$$

The detailed expression of $\phi_m(Q)$ is reported in Supplemental Material [31]. Q is a $2m\times 2m$ matrix that represents the covariance matrix of m-clones of colliding inputs. Namely, denoting by α an index that runs over the $2m$ inputs of the m collision clones, Q is defined as

$$Q_{\alpha\beta}=\frac{1}{N}\sum_{k=1}^{N}w_{k\alpha}w_{k\beta}. \tag{18}$$

The maximization over the entries of Q in Eq. (17) is performed over $\mathcal{F}(p,q_1)$, which is the set of covariance matrices Q that fix the internal overlap q_1 of the m collisions and the external overlap among them to p. In other words, the entries of Q should satisfy the following set of constraints:

$$\frac{1}{2}(Q_{2s-1,2t-1}+Q_{2s,2t})=p \quad \forall\ s<t\in[m], \tag{19a}$$

$$Q_{2s-1,2s}=q_1 \quad \forall\ s\in[m]. \tag{19b}$$

Finding the maximum in Eq. (17) over the space of covariance matrices Q can be done numerically; however, the complexity of such maximization increases considerably with m. Therefore, we have made an educated guess on the structure of the covariance matrix Q achieving the global maximum of the function $\phi_m(Q)$. We considered the so-called *symmetric* ansatz, which imposes

$$Q_{2s-1,2t-1}=Q_{2s,2t}=p \quad \forall\ s<t\in[m], \tag{20a}$$

$$Q_{2s-1,2t}=Q_{2s,2t-1}=q_0 \quad \forall\ s<t\in[m]. \tag{20b}$$

This restricts the number of optimization parameters to just one (i.e., the parameter q_0). This ansatz, on which our results rely, corresponds to assuming that the optimal Q is symmetric under permutation between the clone indexes, as the free entropy is, and can be proven to be correct in certain limit cases (see Secs. C1 and C2 in Supplemental Material [31]).

Figure 2 exemplifies the typical behavior of $\phi_m^{\mathrm{a}}(p)$, for various values of α, including the onset value of the OGP at $\alpha\sim 0.7$. Such curves are continuous but not everywhere differentiable (see Supplemental Material [31] for more details). Figure 3 displays α_{OGP}^m as a function of q_1 for $m=5$ and several values of δ. Decreasing δ leads to a decrease in α_{OGP} for all q_1. For any δ strictly larger than zero, the α_{OGP}^m curve is nonmonotonic in q_1 and starts to increase when $q_1\lesssim 1$, as observed for $\alpha_c(q_1)$. We stress that this threshold is strictly below the one corresponding to the existence of collisions. Between $\alpha_{\mathrm{OGP}}^m(q_1)$ and $\alpha_c(q_1)$, despite the presence of exponentially many collisions, we conjecture that finding one of them is algorithmically hard in the average case.

C. Strong hashing limit

The strong hashing limit (SHL) $\delta\to 0$ can be explicitly studied analytically. Moreover, it is a simple case in which the validity of the symmetric ansatz can be theoretically tested. In particular, as we show in Supplemental Material [31] (see Secs. C1 and C2), a detailed study of the cases $m=2$ and $m=3$ for generic covariance matrices in the limit $\delta\to 0$ reveals that the global maximization procedure (17) leads to a symmetric (20) covariance matrix Q.

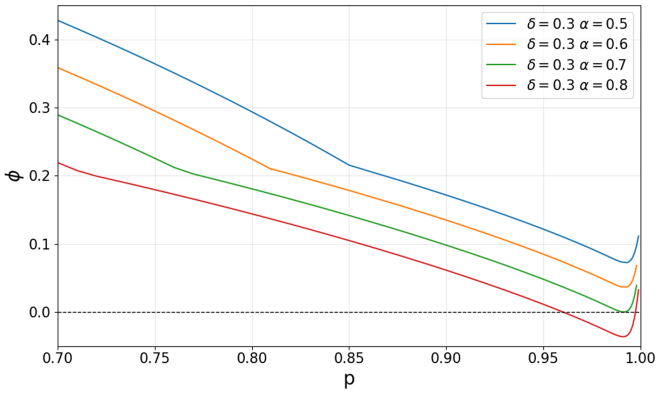

FIG. 2. Typical behavior of $\phi_m^{\mathrm{a}}(p)$, for various values of α, at $q_1=0.6$ and $m=5$. The activation is a square wave, with $\delta=0.3$. The onset value of the OGP is $\alpha=0.7$. For $\alpha>0.7$, an interval of external overlaps p where $\phi<0$ is present.

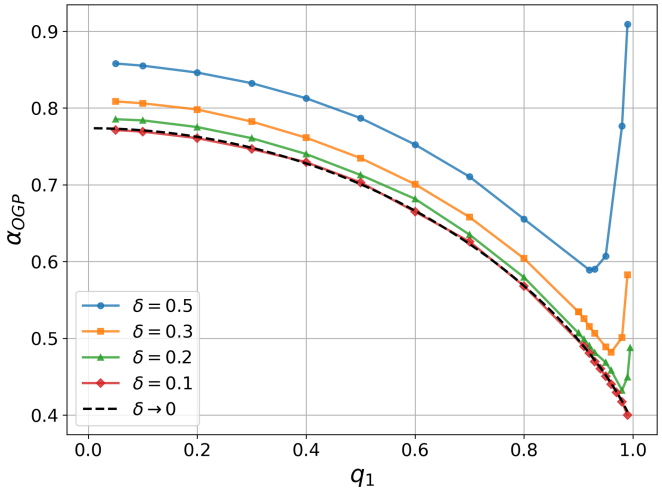

FIG. 3. α^m_{OGP} versus the internal overlap q_1 for the square-wave activation function, for various values of δ, $m = 5$.

In the SHL, the annealed entropy curve becomes a monotonic function of the external overlap p. This allows us to explicitly find the value of $\alpha^m_{\mathrm{OGP}}(q_1)$ by performing the limit $p \to 1$ for every m. One finds

$$\alpha^m_{\mathrm{OGP}}(q_1) = \frac{\log 2 + H_B(q_1)}{\log(1+m)}. \tag{21}$$

Note that this threshold is monotonically decreasing in m and goes to 0 for large m. In Fig. 4, we show the OGP threshold given by Eq. (21) as a function of q_1 and for different values of m.

In Fig. 3, we plot the $\delta \to 0$ OGP transition in black dashed for $m = 5$ together with predictions for finite values of δ. Notice that, for $\delta = 0.1$ already, the curve is practically indistinguishable from the SHL.

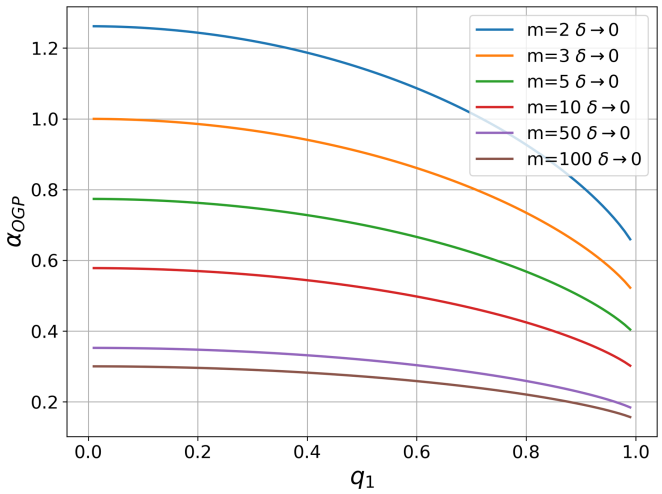

FIG. 4. α^m_{OGP} versus the internal overlap q_1 for the square-wave activation function in the limit $\delta \to 0$ for several values of m.

Since we have established the validity of the symmetric ansatz only for $m = 2$ and $m = 3$ in the SHL, this does not rule out the possibility that the true global maximum is not symmetric, especially for $\delta > 0$. We have, therefore, compared the analytical prediction of the free entropy under the symmetric ansatz, $\phi^{\mathrm{a}}_m(p, q_1)$, with a numerical estimate obtained by an exhaustive search in the space of m-tuples of collisions. The results presented in Supplemental Material [31] are in agreement with the theory.

V. HASH FUNCTIONS FROM COLLISION-RESISTANT NEURAL NETWORKS

A. Collision-resistant function

As already observed, when $\delta > 0$, all m-OGP curves diverge in the limit $q_1 \to 1$. Since an adversary is allowed to look for collisions with arbitrary value of q_1, the region $q_1 \approx 1$ represents a potential flaw in the security of the hash function. To remedy this, recall that an ECC is a function that maps distinct inputs into "code words" separated by large Hamming distance. In order to obtain a collision-resistant function, one can apply an ECC upstream of the network, restricting the space of inputs to $q_1 \le q_{\mathrm{code}}$. The results in Sec. IV B then imply that the composition $f_{\mathbf{A}} \circ \mathrm{ECC}$ is collision resistant as long as α is higher than the OGP threshold for an adversarially chosen $q_1 \le q_{\mathrm{code}}$:

$$\alpha > \alpha_{\mathrm{adv}}(q_{\mathrm{code}}) \coloneqq \max_{q_1 \le q_{\mathrm{code}}} \alpha_{\mathrm{OGP}}(q_1). \tag{22}$$

B. Collision-resistant hash function

We now show how, by composing the neural network with an ECC with appropriate q_{code} value, it is possible to build a CRH. We define the code rate r as the ratio between the input and output bits of the ECC. The compression ratio of the neural network composed with the code is then $\tilde{\alpha} \equiv \alpha / r$. The Gilbert-Varshamov bound [6] shows the existence of a binary code satisfying $r(q_{\mathrm{code}}) = 1 - H_2(q_{\mathrm{code}})$, where

$$H_2(q_{\mathrm{code}}) = -\frac{1-q_{\mathrm{code}}}{2}\log_2\frac{1-q_{\mathrm{code}}}{2} - \frac{q_{\mathrm{code}}+1}{2}\log_2\frac{q_{\mathrm{code}}+1}{2}.$$

The lowest compression rate compatible with collision resistance is then

$$\tilde{\alpha}(q_{\mathrm{code}}) = \frac{\alpha_{\mathrm{adv}}(q_{\mathrm{code}})}{r(q_{\mathrm{code}})} \tag{23}$$

obtained using a code matching the Gilbert-Varshamov bound and selecting the smallest α that guarantees hardness of collision finding for $q_1 \le q_{\mathrm{code}}$. Figure 5 shows $\tilde{\alpha}(q_{\mathrm{code}})$ as a function of q_{code}, for $m = 5$ and various δ values. In the

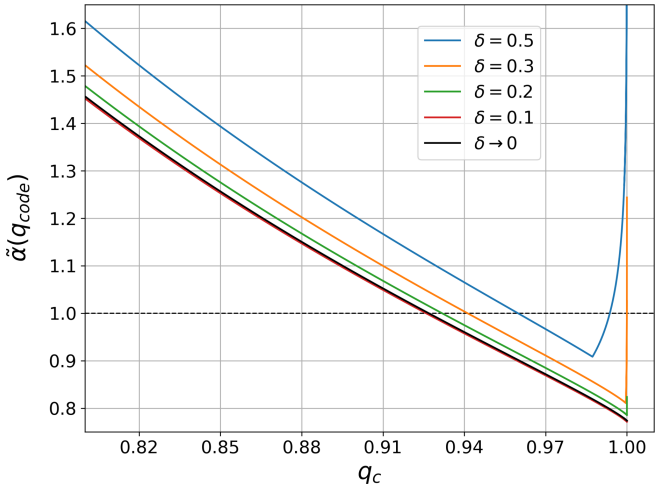

FIG. 5. The lowest compression rate of $NN \circ \mathrm{ECC}$, i.e., $\alpha_{\mathrm{adv}}(q_{\mathrm{code}})/r(q_{\mathrm{code}})$, as a function of q_{code}, for different values of δ. $\alpha_{\mathrm{adv}}(q_{\mathrm{code}})$ is given by an m-OGP computation with $m = 5$.

interval of q_{code} values where $\tilde{\alpha}(q_{\mathrm{code}}) < 1$, we conjecture that $f_A \circ \mathrm{ECC}$ is a CRH.

C. Comparison with Ajtai's function

Postquantum cryptography seeks primitives whose security can be based on problems that are believed to remain intractable even for quantum algorithms. Among the various approaches, lattice-based cryptography has emerged as one of the most promising directions, offering efficient CRH constructions whose security lies on worst-case to average-case reductions.

Ajtai's function [41] stands as a landmark in lattice-based cryptography, as it laid the foundation for subsequent, more efficient lattice-based constructions. Formally, it acts on $\mathbf{x} \in \{0,1\}^N$ as

$$f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} \pmod{q}. \tag{24}$$

Here, $q \in \mathbb{N}$ is a parameter and $\mathbf{A} \in \mathbb{Z}_q^{P\times N}$ is chosen uniformly at random, where $\mathbb{Z}_q$ is a ring with q elements. This function is shrinking when $P \log_2 q < N$. For certain values of $q, P = \mathrm{poly}(N)$ satisfying this bound, Ajtai's function can be shown to be a CRH, assuming it is hard to find short vectors in certain integer lattices [41,42]. Namely, the security relies on the worst-case hardness of the *short integer solution* problem [43] that for a matrix $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$ asks to find a "short" vector $\mathbf{x} \in \mathbb{Z}_q^m$ such that $\mathbf{A}\mathbf{x} = 0$, a problem for which, for certain values of q, M, N, the only known algorithms run in time $2^{\Omega(n)}$ [44,45], even with quantum computers.

Note that our function bears some resemblance to Ajtai's function, as the oscillating nature of our activation is reminiscent of the mod operation. So our parameter δ plays a role similar to q in Ajtai's function. However, there are also important differences. First of all, in Ajtai's case, the worst- to average-case reduction, which is a crucial requirement for its collision resistance, is guaranteed when q scales as $\mathrm{poly}(N)$, whereas in our case we can maintain $\delta = O(1)$. Second, the output spaces differ: Our function maps to $\pm 1^P$, while Ajtai's maps to $\mathbb{F}_q^P$. Moreover, in our case, the entries of $\boldsymbol{A}$ are independent random variables which are drawn from an *arbitrary* distribution. Note also that our security guarantee via the OGP criterion depends on only the first two moments of such distribution. Finally, and most importantly, the periodicity of Ajtai's function is believed to be a central aspect of its hardness, while our results do not rely on the periodicity of the activation. Indeed, we have repeated the OGP analysis for an activation with a finite number $2K$ of oscillations, followed by a constant plateau, i.e., a "truncated" version of the SWP. Specifically, oscillations span a range $[-\gamma, \gamma]$, where γ is a constant with N. As we show in Supplemental Material [31], it is possible to find a region for q_{code} where our function composed with the ECC is a CRH. Preliminary checks indicate that qualitatively similar results also arise in a randomized version of the SWP, where the periodicity is broken by randomizing the points at which the activation function switches sign.

These results suggest that what matters the most for our hardness criterion is not the periodicity of the function but rather the frequency of sign switches of the activation, even if they are located in a limited region near the origin.

Qualitatively, our function can be considered a generalization of Ajtai's function, where one considers a specific bit of its output, when outputs in $\mathbb{Z}_q$ are represented as binary strings, rather than the least significant bit. The corresponding modification of Ajtai's function would not be secure due to the performance of the best-known lattice algorithms [46,47].

VI. ALGORITHMIC ATTACKS

In this section, we present two algorithmic strategies for finding collisions: One involves a local search starting from a random point on the hypercube, where local moves are performed to explore potential collisions, while the other is based on approximate message passing.

A. Local search from a random reference

Perhaps the most straightforward strategy to look for a collision in a binary perceptron is the following. Take a random binary configuration $\boldsymbol{w}$. To construct a collision, one can attempt a sequence of single bit flips on the elements of $\boldsymbol{w}$, stopping as soon as the resulting vector $\boldsymbol{w}'$ collides with $\boldsymbol{w}$. In this section, we argue that this strategy is unfeasible if the number of patterns is proportional to N. In order to do so, let us count the average number $\mathcal{N}_t$ of collisions obtained by flipping $t = O(1)$ bits from the reference $\boldsymbol{w}$. One gets (see Supplemental Material [31])

$$\mathcal{N}_t \approx \binom{N}{t} e^{-a(\delta) t^{1/2} P N^{-1/2}}, \tag{25}$$

where $a(\delta)$ is a positive constant in N, which depends on δ. Equation (25) implies that if $P = O(N)$, in the limit $N \to \infty$ there are no collisions in the finite-size neighborhood of typical vertices of the hypercube, indicating that a local search from a random reference $\boldsymbol{w}$ is an unfeasible strategy.

B. Extensive distances from a random reference

Consider again a random vertex of the hypercube $\boldsymbol{w}$. In this section, instead of looking at a finite-distance neighborhood of $\boldsymbol{w}$, we are interested in the algorithmic problem of finding collisions at an extensive distance. Note that in this case the probability of finding a collision is exponentially small in N, but there is an exponential number of configurations and the two can compensate [see Eq. (25)]. In order to do so, we study the performance of a message-passing algorithm inspired by statistical physics called reinforced approximate message passing (rAMP). Variants of the rAMP algorithms are known to be effective polynomial-time heuristics to solve inversion and second-preimage problems in perceptrons with binary synapses [16,18,48–51]. We write in Supplemental Material [31] the details of the implementation and the analysis of the performance with square-wave activation. The complexity of rAMP is $O(T_{\rm sol} NP)$, where $NP \propto N^2$ comes from the fact that the algorithm involves matrix vector multiplications with the pattern matrix $\boldsymbol{A}$, and $T_{\rm sol}$ is the number of iterations required to find a solution. The quantity $T_{\rm sol}$ displays a power-law behavior $T_{\rm sol} \approx N^{b(\alpha)}$ in the system size, with a critical exponent $b(\alpha) \approx 1/(\alpha_r - \alpha)$. Therefore, the total complexity is $O(N^{2+b(\alpha)})$. The exponent $b(\alpha)$ is usually extrapolated numerically (see Ref. [49]).

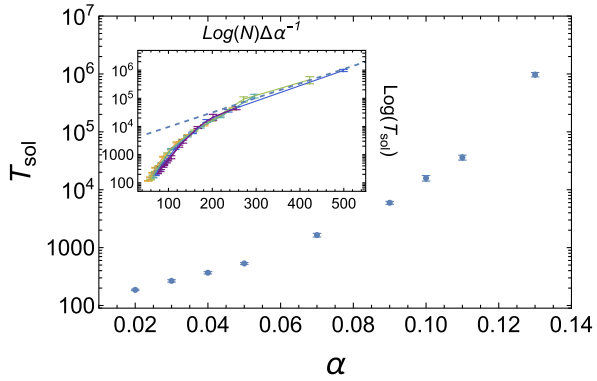

FIG. 6. Number of iterations $T_{\rm sol}$ required by rAMP to find a solution for $N = 8 \times 10^3$ and $\delta = 0.6$, as a function of the constraint density α. Inset: $T_{\rm sol}$ has an exponential growth of the form $T_{\rm sol} \approx N^{b(\alpha)}$, with $b(\alpha) \propto (\alpha_r - \alpha)^{-1} \equiv \Delta\alpha^{-1}$ in the proximity of $\alpha_r \approx 0.15$. To test this behavior, we show $\log T_{\rm sol}$ as a function of $\log N/\Delta\alpha$, for different values of N. The curves collapse for large $\log N/\Delta\alpha$. From bottom left to bottom right, $N = n \times 10^3$, with $n = 1, 2, 4, 8, 16, 32$. The dashed line is a scaling function of the form $\exp(ax + b)$.

In order to test the hardness of the collision-finding problem, we run rAMP on the SWP with $\delta = 0.6$, where the annealed computation of the previous sections predicts the presence of an OGP. We find that rAMP outputs collisions having internal normalized overlap compatible with zero up to $\alpha_{\rm rAMP} \approx 0.15$, where $\alpha_{\rm rAMP}$ is estimated by a numerical characterization of the exponent $b(\alpha)$ (see Fig. 6). We note that $\alpha_{\rm rAMP}$ is strictly smaller than the OGP estimate for $q_1 = 0$, that for $m = 5$ and $\delta = 0.6$ is $\alpha^5_{\rm OGP} \approx 0.9$ (extrapolating from Fig. 3).

VII. CONCLUSIONS

The square-wave perceptron, together with its truncated variant, defines a class of neural-network-based functions that, in a suitable regime, admit collisions while exhibiting features consistent with collision resistance, as suggested by the presence of an overlap gap property. By combining these functions with error-correcting codes, we obtain a novel family of candidate collision-resistant hash functions. To our knowledge, this is the first instance where a geometric criterion rooted in statistical physics is employed to argue the security of cryptographic primitives. A direction for future work is to investigate whether more complex neural network architectures also display this property.

ACKNOWLEDGMENTS

E. M. M. acknowledges the MUR-Prin 2022 funding Prot. No. 20229T9EAT, financed by the European Union (Next Generation EU). This paper is supported by PNRR-PE-AI FAIR project funded by the NextGeneration EU program. N. I. S. was supported by European Research Council (ERC) under the EU's Horizon 2020 research and innovation program (Grant Agreement No. 101019547). A. R. was supported by ERC under the EU's Horizon 2020 research and innovation program (Grant Agreement No. 101019547) and a Cariplo CRYPTONOMEX grant.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

[1] W. S. McCulloch and W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, Bull. Math. Biophys. **5**, 115 (1943).

[2] T. M. Cover, *Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition*, IEEE Trans. Electron. Comput. **EC-14**, 326 (1965).

[3] S. Kirkpatrick and B. Selman, *Critical behavior in the satisfiability of random boolean expressions*, Science **264**, 1297 (1994).

[4] M. Mézard, G. Parisi, and R. Zecchina, *Analytic and algorithmic solution of random satisfiability problems*, Science **297**, 812 (2002).

[5] F. Krzakała, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová, *Gibbs states and the set of solutions of random constraint satisfaction problems*, Proc. Natl. Acad. Sci. U.S.A. **104**, 10318 (2007).

[6] M. Mezard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, New York, 2009).

[7] Formally, collision resistance is defined as follows: For every efficient algorithm $C(\cdot)$ and any constant $c > 0$, the probability that $C(\mathbf{A})$ outputs $\mathbf{x} \neq \mathbf{y}$ with $f_{\mathbf{A}}(\mathbf{x}) = f_{\mathbf{A}}(\mathbf{y})$ is smaller than N^{-c} for any sufficiently large N (where the randomness is taken over the coins used by C).

[8] S. Goldwasser, S. Micali, and C. Rackoff, *The knowledge complexity of interactive proof systems*, SIAM J. Comput. **18**, 186 (2019).

[9] M. Blum, *Coin flipping by telephone a protocol for solving impossible problems*, ACM SIGACT News **15**, 23 (1983).

[10] W. Diffie and M. E. Hellman, *New directions in cryptography*, IEEE Trans. Inf. Theory **22**, 644 (2022).

[11] A. C. Yao, *Protocols for secure computations*, in *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science (SFCS 1982)* (IEEE, New York, 1982), pp. 160–164.

[12] M. Mézard and R. Zecchina, *Random k-satisfiability problem: From an analytic solution to an efficient algorithm*, Phys. Rev. E **66**, 056126 (2002).

[13] M. Mézard, T. Mora, and R. Zecchina, *Clustering of solutions in the random satisfiability problem*, Phys. Rev. Lett. **94**, 197205 (2005).

[14] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, *Determining computational complexity from characteristic 'phase transitions'*, Nature (London) **400**, 133 (1999).

[15] L. Zdeborová and F. Krzakala, *Phase transitions in the coloring of random graphs*, Phys. Rev. E **76**, 031131 (2007).

[16] C. Baldassi, A. Ingrosso, C. Lucibello, L. Saglietti, and R. Zecchina, *Subdominant dense clusters allow for simple learning and high computational performance in neural networks with discrete synapses*, Phys. Rev. Lett. **115**, 128101 (2015).

[17] C. Baldassi, C. Lauditi, E. M. Malatesta, R. Pacelli, G. Perugini, and R. Zecchina, *Learning through atypical phase transitions in overparameterized neural networks*, Phys. Rev. E **106**, 014116 (2022).

[18] C. Baldassi, E. M. Malatesta, G. Perugini, and R. Zecchina, *Typical and atypical solutions in nonconvex neural networks with discrete and continuous weights*, Phys. Rev. E **108**, 024310 (2023).

[19] D. Gamarnik, *The overlap gap property: A topological barrier to optimizing over random structures*, Proc. Natl. Acad. Sci. U.S.A. **118**, e2108492118 (2021).

[20] D. Gamarnik, *Turing in the shadows of Nobel and Abel: An algorithmic story behind two recent prizes*, arXiv:2501.15312.

[21] CSPs with a linear structure are considered somewhat an exception to this hardness conjecture, since, despite the disconnectivity properties of their space of solutions, they can be always solved in polynomial time by Gaussian elimination.

[22] D. Gamarnik, E. C. Kızıldağ, W. Perkins, and C. Xu, *Algorithms and barriers in the symmetric binary perceptron model*, in *Proceedings of the 2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New York, 2022), pp. 576–587.

[23] D. Gamarnik and E. C. Kizildag, *Algorithmic obstructions in the random number partitioning problem*, Ann. Appl. Probab. **33**, 5497 (2023).

[24] D. Gamarnik, E. C. Kizildag, and L. Warnke, *Optimal hardness of online algorithms for large independent sets*, arXiv:2504.11450.

[25] D. Gamarnik and A. Jagannath, *The overlap gap property and approximate message passing algorithms for p-spin models*, Ann. Prob. **49**, 180 (2021).

[26] D. Gamarnik, A. Jagannath, and A. S. Wein, *Hardness of random optimization problems for boolean circuits, low-degree polynomials, and Langevin dynamics*, SIAM J. Comput. **53**, 1 (2024).

[27] As shown in Supplemental Material [31], our analysis holds for a general activation. We concentrate mostly on the family of periodic square-wave activations merely because of their potential relevance in cryptography.

[28] M. Benedetti, A. Bogdanov, E. M. Malatesta, M. Mézard, G. Perrupato, A. Rosen, N. I. Schwartzbach, and R. Zecchina, *Overlap gap and computational thresholds in the square wave perceptron*, J. Stat. Mech. (2025) 123303.

[29] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications* (World Scientific Publishing Company, Singapore, 1987), Vol. 9.

[30] M. Talagrand, *Self averaging and the space of interactions in neural networks*, Random Struct. Algorithms **14**, 199 (1999).

[31] See Supplemental Material at http://link.aps.org/supplemental/10.1103/6shh-9h5m for derivations of the annealed and replica-symmetric free-entropy calculations, details of the multioverlap gap property analysis and numerical simulations.

[32] W. Krauth and M. Mézard, *Storage capacity of memory networks with binary couplings*, J. Phys. **50**, 3057 (1989).

[33] J. Ding and N. Sun, *Capacity lower bound for the Ising perceptron*, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC 2019), Phoenix, AZ, USA* (ACM, New York, NY, 2019), pp. 816–827, 10.1145/3313276.3316383.

[34] B. Huang, *Capacity threshold for the Ising perceptron*, arXiv:2404.18902.

[35] E. Barkai, D. Hansel, and H. Sompolinsky, *Broken symmetries in multilayered perceptrons*, Phys. Rev. A **45**, 4146 (1992).

[36] A. Engel, H. M. Köhler, F. Tschepke, H. Vollmayr, and A. Zippelius, *Storage capacity and learning algorithms for two-layer neural networks*, Phys. Rev. A **45**, 7590 (1992).

[37] C. Baldassi, E. M. Malatesta, and R. Zecchina, *Properties of the geometry of solutions and capacity of multilayer neural*

networks with rectified linear unit activations, Phys. Rev. Lett. **123**, 170602 (2019).

[38] B. L. Annesi, E. M. Malatesta, and F. Zamponi, *Exact full-RSB SAT/UNSAT transition in infinitely wide two-layer neural networks*, SciPost Phys. **18**, 118 (2025).

[39] A. S. Wein, *Optimal low-degree hardness of maximum independent set*, Math. Stat. Learn. **4**, 221 (2022).

[40] E. C. Kızıldağ, *Sharp thresholds for the overlap gap property: Ising p-spin glass and random k-sat*, in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2025), Leibniz International Proceedings in Informatics (LIPIcs)* (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2025), pp. 48:1–48:18.

[41] M. Ajtai, *Generating hard instances of lattice problems*, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing* (ACM, New York, NY, 1996), pp. 99–108, 10.1145/237814.237838.

[42] O. Goldreich, S. Goldwasser, and S. Halevi, *Collision-free hashing from lattice problems*, in *Studies in Complexity and Cryptography: Miscellanea on the Interplay between Randomness and Computation*, edited by O. Goldreich, Lecture Notes in Computer Science Vol. 6650 (Springer, Berlin, Heidelberg, 2011), pp. 30–39, 10.1007/978-3-642-22670-0_5.

[43] O. Regev, *On lattices, learning with errors, random linear codes, and cryptography*, J. Assoc. Comput. Mach. **56**, 1 (2009).

[44] X. Pujol and D. Stehlé, *Solving the shortest lattice vector problem in time* $2^{2.465n}$, Cryptology ePrint Archive (2009).

[45] C.-P. Schnorr and M. Euchner, *Lattice basis reduction: Improved practical algorithms and solving subset sum problems*, Math. Program. **66**, 181 (1994).

[46] A. K. Lenstra, H. W. Lenstra, and L. Lovász, *Factoring polynomials with rational coefficients*, Math. Ann. **261**, 515 (1982).

[47] C.-P. Schnorr, *A more efficient algorithm for lattice basis reduction*, J. Algorithms **9**, 47 (1988).

[48] A. Braunstein and R. Zecchina, *Learning by message passing in networks of discrete synapses*, Phys. Rev. Lett. **96**, 030201 (2006).

[49] C. Baldassi and A. Braunstein, *A max-sum algorithm for training discrete neural networks*, J. Stat. Mech. (2015) P08008.

[50] C. Baldassi, C. Borgs, J. T. Chayes, A. Ingrosso, C. Lucibello, L. Saglietti, and R. Zecchina, *Unreasonable effectiveness of learning neural networks: From accessible states and robust ensembles to basic algorithmic schemes*, Proc. Natl. Acad. Sci. U.S.A. **113**, E7655 (2016).

[51] J. Chavas, C. Furtlehner, M. Mézard, and R. Zecchina, *Survey-propagation decimation through distributed local computations*, J. Stat. Mech. (2005) P11016.