Critical Jammed Phase of the Linear Perceptron

Silvio Franz,[1,2] Antonio Sclocchi,[1] and Pierfrancesco Urbani[3]
[1]LPTMS, Université Paris-Sud 11, UMR 8626 CNRS, Bâtiment 100, 91405 Orsay Cedex, France
[2]Dipartimento di Fisica Università, La Sapienza, Piazzale Aldo Moro 5, I-00185 Roma, Italy
[3]Institut de Physique Théorique, Université Paris Saclay, CNRS, CEA, F-91191 Gif-sur-Yvette, France

(Received 8 March 2019; published 13 September 2019)

Criticality in statistical physics naturally emerges at isolated points in the phase diagram. Jamming of spheres is not an exception: varying density, it is the critical point that separates the unjammed phase where spheres do not overlap and the jammed phase where they cannot be arranged without overlaps. The same remains true in more general constraint satisfaction problems with continuous variables where jamming coincides with the (protocol dependent) satisfiability transition point. In this work we show that by carefully choosing the cost function to be minimized, the region of criticality extends to occupy a whole region of the jammed phase. As a working example, we consider the spherical perceptron with a linear cost function in the unsatisfiable jammed phase and we perform numerical simulations which show critical power laws emerging in the configurations obtained minimizing the linear cost function. We develop a scaling theory to compute the emerging critical exponents.

DOI: 10.1103/PhysRevLett.123.115702

Introduction.—The jamming transition of spheres is a critical point [1]. At jamming, spheres form an *isostatic* network [2] where the number of contacts between them equals exactly the total number of degrees of freedom. Furthermore, the distributions of forces and gaps (i.e., distances) between particles display power laws [3–6] which play a central role in the mechanical and rheological properties of such systems. In Refs. [7,8] the corresponding critical exponents have been computed from the solution of the hard sphere model in infinite dimension. In this analysis the jamming transition is thought of as the infinite pressure limit of hard sphere glassy states. Upon compression, hard sphere glasses undergo a Gardner transition [9]: the glass basins of configurations split into a "fractal landscape" of just marginally stable metastable states, described by full replica symmetry breaking (RSB) [10], with soft excitations and divergent susceptibilities. Within this mean-field scenario, these excitations are responsible for the anomalous rheological response of amorphous solids [11–15]. In the jamming limit this landscape marginality gives rise to the power laws observed in the gaps and forces distributions and predicts the mechanical properties of amorphous jammed packings.

Subsequently, it has been argued [16–18] that the jamming transition can be thought of as a special case of a satisfiability transition for constraint satisfaction problems [19] with continuous variables (CCSP). In a generic CCSP one seeks configurations of variables that satisfy a set of constraints. From this viewpoint, jamming is the (protocol dependent) point that separates a satisfiable (SAT) unjammed phase, where the spheres can be arranged to satisfy the nonoverlapping constraints, from an unsatisfiable (UNSAT) or jammed phase, where some constraints are violated and spheres overlap. In this way, one can generalize the problem of jamming to other situations. The simplest one is borrowed from machine learning and is a nonconvex twist of the perceptron classifier [20]. In Ref. [16] it has been argued that at the satisfiability transition point (meaning at jamming) this CCSP displays analogous power law distributions of gaps and forces whose critical exponents coincide with the corresponding exponents in spheres. Nontrivial generalizations of the perceptron [21,22] retain the same critical behavior.

However, the criticality of the gap and force distributions both in spheres and in the perceptron is generically attributed to the emergence of jamming and should disappear in the jammed, UNSAT phase. This is supported by analytical computations in the perceptron problem with harmonic cost function and by numerical simulations on harmonic soft spheres.

In this work, we show that this is not always the case: changing the potential or cost function from harmonic to linear, we find jamming criticality in an extensive region of the UNSAT, jammed phase, far away from jamming.

We consider the UNSAT, jammed phase of the simplest CCSP, the spherical perceptron model, and we look at local minima of the linear cost function (instead of the harmonic one). We show that even far from jamming, in an extensive region of the phase diagram, the landscape induced by this cost function is nonconvex and composed by metastable minima which are all jamming critical. Indeed, for these minima the positive and negative gap distributions display power law divergences for small argument: surprisingly, the critical exponents describing these power laws coincide

with the corresponding critical exponents of the jamming transition. Moreover, we find that this behavior is associated with isostaticity: even when the model is in the UNSAT phase, there is an extensive number of marginally satisfied constraints, i.e., constraints that are right at the border of satisfaction (like perfectly touching spheres). The jamming-critical phase appears when the number of marginally satisfied constraints equals the number of degrees of freedom of the system which becomes therefore isostatic.

The spherical perceptron with linear cost function has been studied in Refs. [23–25], where the phase diagram was obtained using the replica method and studied at the so-called replica symmetric level in Ref. [23]. While it was known that RSB is needed in the UNSAT, jammed phase, systematic studies beyond 1-RSB [25–27] were not undertaken. Here we show that the emergence of RSB in the UNSAT phase is related to the emergence of jamminglike criticality.

Model.—The spherical perceptron is an optimization problem defined through an N-dimensional vector $\underline{w}$ on the N-dimensional hypersphere $|\underline{w}|^2 = N$ and by a set of $M = \alpha N$ N-dimensional random vectors $\{\underline{\xi}^\mu\}$ whose components are *independent identically distributed* Gaussian random variables with zero mean and unit variance. For each of these vectors one defines the gap h_μ by

$$h_\mu = \frac{1}{\sqrt{N}} \underline{\xi}^\mu \cdot \underline{w} - \sigma. \tag{1}$$

We say that a gap h_μ is (i) satisfied if $h_\mu > 0$, (ii) marginally satisfied if $h_\mu = 0$, and (iii) unsatisfied if $h_\mu < 0$. The variables σ and α are control parameters. One can define an energy or cost function associated with the unsatisfied gaps as

$$H[\underline{w}] = \frac{1}{p} \sum_{\mu=1}^{\alpha N} |h_\mu|^p \theta(-h_\mu), \tag{2}$$

and study its minima. Equation (2) depends on a parameter p that sets the strength of the cost when gap variables are unsatisfied. The harmonic perceptron corresponds to $p = 2$ and has been studied in Refs. [17,18]. Here we set $p = 1$, which defines the linear cost function. This cost function is not very used in soft matter systems but it is common in the machine learning literature where it is called *hinge loss* and plays an important role in support vector machines [28]. Furthermore, the case $p = 1$ marks the boundary where H passes from a convex ($p > 1$) to a nonconvex ($p < 1$) function of each of the h_μ's. We stress, however, that the convexity of H in the h_μ's does not necessarily imply convexity of H in the variables $\underline{w}$ that live on the hypersphere: the loss of convexity is indeed associated with glassiness.

The phase diagram of the spherical perceptron with linear cost function was obtained for $\sigma > 0$ in Ref. [23] (see also Ref. [25]) and is redrawn in Fig. 1 in terms of the control parameters σ and α. It contains two regions separated by the SAT-UNSAT transition line (red line in Fig. 1). Below this line, the problem is SAT (or unjammed), meaning that with probability one, for $N \to \infty$, there are configurations of $\underline{w}$ for which the cost function is strictly zero; i.e., the gaps h_μ are satisfied for all $\mu = 1, \ldots, \alpha N$. Conversely, above this line, the cost function is positive and an extensive number of gaps are unsatisfied: this is the UNSAT, jammed phase. In Fig. 1 we plot also the de Almeida–Thouless (RSB) line [29] (dashed black line): below this line and in the UNSAT phase, the energy landscape is effectively convex and the linear cost function has a unique global minimum. Above this line, convexity is lost and multiple metastable minima emerge. We are interested in studying the properties of these local minima.

Numerical simulations.—Local minima of the linear cost function turn out to be nonanalytic angular points determined by the intersection of hyperplanes. We smooth out the singularity at $h_\mu = 0$ and define a regularized cost function as

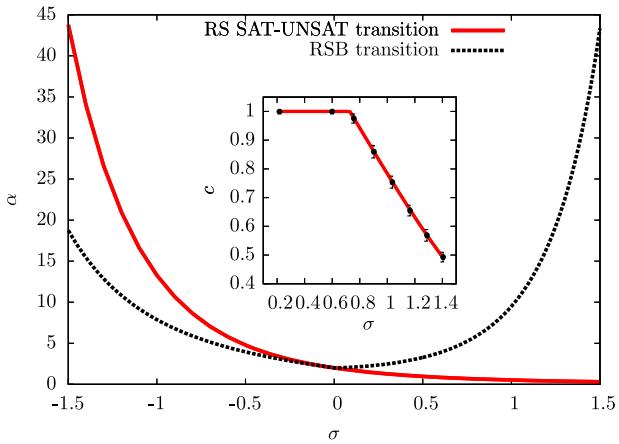

FIG. 1. The phase diagram of the spherical perceptron with linear cost function. The y axis represents the density of constraints $\alpha = M/N$, while the x axis is the control parameter σ that defines the gap variables; see Eq. (1). The red solid line is the SAT-UNSAT (jamming) line (computed under the replica symmetric approximation, exact only for $\sigma > 0$). Below this line the model is SAT, unjammed and one can find zero energy configurations. Above, the problem is UNSAT, jammed and the energy is positive. Above the black dashed line replica symmetry spontaneously breaks down: below this line the problem is convex, while above, the energy landscape is glassy with many local minima with critical properties. The jamming line lies in the RS region for $\sigma > 0$ and in the RSB region for $\sigma < 0$. Inset: The density of contacts c for $\alpha = 5$ as a function of σ. The red line corresponds to the theoretical prediction. We have hypostaticity $c < 1$ in the RS phase and isostaticity, i.e., $c = 1$, in the RSB phase. The dots come from numerical simulations with $N = 1024$ at $\alpha = 5$.

$$H_\varepsilon[\underline{w}] = \sum_{\mu=1}^{\alpha N} \left| h_\mu + \frac{\varepsilon}{2} \right| \theta(-h_\mu - \varepsilon) + \frac{\zeta}{4}(|\underline{w}|^2 - N)^2 + \frac{1}{2\varepsilon} \sum_{\mu=1}^{\alpha N} h_\mu^2 \theta(\varepsilon + h_\mu)\theta(-h_\mu), \tag{3}$$

where $\varepsilon > 0$ and ζ is an arbitrary large constant needed to enforce the spherical constraint $|\underline{w}|^2 = N$. We implemented the numerical minimization of $H_\varepsilon[\underline{w}]$ using the routine BFGS [30] of the SciPy library [31]. For every ε, the algorithm reaches a local minimum. In order to describe the properties of the linear cost function, we need to study the minima when $\varepsilon \to 0^+$. Therefore, we consider a decreasing sequence of values of ε and minimize the cost function at each step (see Supplemental Material for details [32]). We observe that for ε small enough there is an extensive fixed set of gaps in the interval $\mathcal{D} = [-\varepsilon, 0]$. Decreasing ε, these gaps remain in $\mathcal{D}$, indicating that for $\varepsilon \to 0^+$ they become marginally satisfied: we call them contacts in analogy with sphere packings. We define the empirical distribution of gaps as $\rho(h) = \overline{(1/M)\sum_{\mu=1}^{M} \delta(h - h_\mu)}$, where the average is taken over many different realizations of the random vectors $\{\underline{\xi}^\mu\}$. Therefore, for $\varepsilon \to 0^+$, $\rho(h)$ contains a Dirac delta at $h = 0$. Calling $\mathcal{I}_\mathcal{D}$ the total number of contacts, we can define an isostaticity index $c = \mathcal{I}_\mathcal{D}/N$. In the inset of Fig. 1 we plot c as a function of σ for $\alpha = 5$. When replica symmetry is unbroken (RS), $c < 1$, and we say that minima are hypostatic. Conversely, when the minimization is carried in the RSB region, we find $c = 1$, and therefore we say that minima are isostatic.

Once the contacts are identified we construct the statistics of strictly positive and negative gaps. We study what happens to $\rho(h)$ when $h \to 0^\pm$. In the RS-UNSAT phase, $\rho(h \to 0^\pm) \to A_\pm$, with $A_\pm$ being two positive nonuniversal constants that depend on the control parameters. In the RSB region instead we observe that $\rho(h \to 0^\pm) \sim |h|^{-\gamma_\pm}$, where $\gamma_\pm$ are two critical exponents. In Fig. 2, we plot the cumulative distributions of both positive and negative gaps for minima with an average energy of $\langle H \rangle/N = 1.01 \pm 0.02$, therefore far away from the jamming transition. Both distributions display a nontrivial power law for small argument. The critical exponents $\gamma_\pm$ are very close to each other and have numerical value $\gamma_+ \approx \gamma_- \approx 0.41$, which is compatible with the critical exponent γ_J that controls the positive gaps distribution at jamming transition [7]. Moreover, we can obtain the virtual forces $\hat{f}$ associated to contacts [16]. These are defined as the Lagrange multipliers needed to enforce that the corresponding gaps are identically zero [33]. Their empirical distribution is defined as $\rho_f(\hat{f}) = \overline{(1/cN)\sum_{i\in\mathcal{D}} \delta(\hat{f} - \hat{f}_i)}$ and has support in [0, 1]. We find that while in the convex phase $\rho_f(\hat{f})$ is regular at both edges, in the RSB phase it becomes critical, with pseudogaps close to both edges of its support, $\rho_f(\hat{f}) \sim \hat{f}^\theta$ and $\rho_f(\hat{f}) \sim (1 - \hat{f})^{\theta'}$ for $\hat{f} \sim 0^+$ and for $\hat{f} \sim 1^-$, respectively. In Fig. 3, we plot the cumulative distribution of forces as a function of both $\hat{f}$ and $1 - \hat{f}$ as obtained from simulations: we observe two power laws with exponents $\theta \simeq \theta' \simeq 0.42$, again compatible with the critical exponent θ_J that controls the distribution of small forces at jamming.

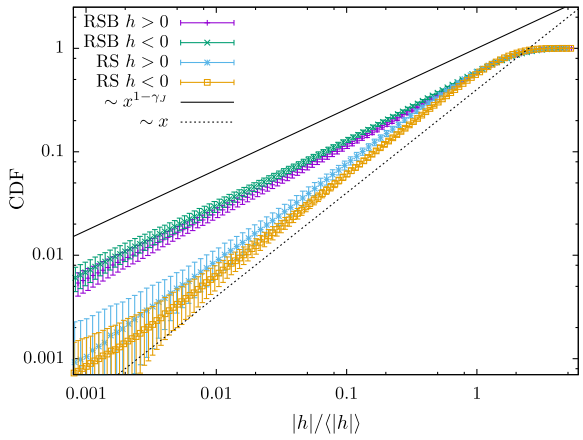

FIG. 2. The cumulative distribution function (CDF) of both strictly positive and strictly negative gaps for $N = 1024$ and $\alpha = 5$. We compare the distributions in the UNSAT-RSB and UNSAT-RS regions. The RSB data refer to minima at an average energy $H/N = 1.01 \pm 0.02$, and the corresponding value of σ is $\sigma = 0.219 \pm 0.004$. Both positive and negative gaps' distributions are compatible with a power law at small argument with exponent $1 - \gamma_J \simeq 0.59$ (black full line). The RS data refer to minima with $H/N = 2.540 \pm 0.013$ and $\sigma = 0.757 \pm 0.010$. One sees there a linear behavior of the CDF, implying a positive probability density function in the origin.

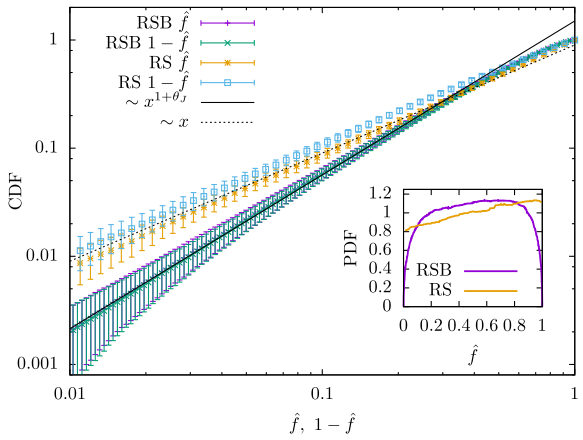

FIG. 3. The CDF of the virtual forces presented as a function of both $\hat{f}$ and $1 - \hat{f}$ for the same parameters as in Fig. 2. In the RSB phase the distribution of forces vanishes as a power law both in $\hat{f} = 0$ and in $\hat{f} = 1$ (purple line in the inset), and both powers are compatible with $\theta_J \simeq 0.42$ corresponding to a CDF with a power 1.42. In the RS region the behavior is linear and the PDF is finite at both edges of its support (yellow line in the inset). The datasets presented in the inset correspond to $\langle H \rangle/N \simeq 2, 3$ for RSB and RS curves respectively.

Therefore, numerical simulations show that when the energy minimization is carried out in the RSB-UNSAT phase, there are three classes of small gaps: an isostatic set of gaps that are identically zero and two sets of positive and negative gaps that accumulate around zero. Furthermore, the marginally satisfied gaps are associated to a critical distribution of virtual forces. Unlike for the harmonic case, here scaling behavior emerges even in the UNSAT phase far from jamming.

Theory.—We analyze the thermodynamic phase diagram of the model using the replica method. Similarly to the case of the jamming transition in spheres and in the nonconvex perceptron, the UNSAT critical phase is associated with scaling behavior that controls the universality of the gaps and forces distributions. However, in the case of the jamming transition, there is a single scaling region that describes small positive gaps and forces. Instead, in the present case we also have two additional power laws describing small negative gaps and virtual forces close to one. This corresponds to the emergence of an additional scaling region. Both regions can be theoretically identified from the replica analysis. Here we sketch the main steps; details are in the Supplemental Material [32]. The phase diagram and the properties of the model can be obtained by studying the zero temperature limit $\beta = 1/T \to \infty$ of its free energy [34]:

$$f = -\frac{1}{\beta N}\overline{\ln \int d\underline{w} e^{-\beta H[\underline{w}]}}, \tag{4}$$

where the overline stands for the average over the random vectors $\underline{\xi}^{\mu}$. The disorder average can be performed using the replica method [18,24]. The RSB phase is described by the probability distribution of the overlap $q = \underline{w}_1 \cdot \underline{w}_2/N$ between different configurations and is captured by the following partial differential equations [10,35], valid for real values of h and for q in an interval $q \in [q_m, q_M] \subset [0,1]$ determined self-consistently:

$$\begin{aligned}\frac{\partial m(q,h)}{\partial q} &= -\frac{1}{2}m''(q,h) - \frac{x(q)}{\lambda(q)}m(q,h)[1+m'(q,h)],\\ \frac{\partial P(q,h)}{\partial q} &= \frac{1}{2}\left[P''(q,h) - 2\frac{x(q)}{\lambda(q)}[P(q,h)m(q,h)]'\right],\end{aligned} \tag{5}$$

where the primes indicate partial derivatives with respect to h, and the boundary conditions are given by

$$\begin{aligned} m(q_M,h) &= (1-q_M)\frac{\partial}{\partial h}\ln \gamma_{1-q_M} \star e^{-\beta|h|\theta(-h)},\\ P(q_m,h) &= \gamma_{q_m}(h+\sigma).\end{aligned} \tag{6}$$

γ_Δ is a Gaussian with zero mean and variance Δ and $\star$ stands for the convolution operation. The function $x(q)$ is directly related to the distribution of the overlap q [10], and we have defined $\lambda(q) = 1 - q_M + \int_q^{q_M} dp x(p)$. At large β, one can get the distribution of virtual forces and gaps from the solution of $P(q,h)$ in the limit $q \to 1$. We analyze Eq. (5) in the $\beta \to \infty$ limit in the RSB-UNSAT phase and show that they admit a scaling solution which accounts for the power laws observed in numerical simulations. In the UNSAT phase, for $\beta \to \infty$ one has $q_M \to 1$. In the limit $0 < 1 - q \ll 1$, two scaling regimes emerge for $m(q,h)$. One concerns the region $h = O(\sqrt{1-q})$, analogue to the one found at jamming [7], and a new one associated to negative gaps for $h = -\hat{\lambda}(q) + O(\sqrt{1-q})$, where $\hat{\lambda}(q) = \lim_{\beta\to\infty}\beta\lambda(q) \simeq (1-q)^{(\kappa-1)/\kappa} \gg \sqrt{1-q}$ (with $\kappa < 2$ a critical exponent). Therefore, we can write

$$m(q,h) = \begin{cases} -\sqrt{1-q}\,\mathcal{M}_+\left(\frac{h}{\sqrt{1-q}}\right) & |h| \sim \sqrt{1-q} \\ -h + \sqrt{1-q}\,\mathcal{M}_-\left(\frac{h+\hat{\lambda}(q)}{\sqrt{1-q}}\right) & h+\hat{\lambda}(q) \sim \sqrt{1-q}. \end{cases} \tag{7}$$

It turns out that the two scaling functions are related by the symmetry relation $\mathcal{M}_-(t) = t + \mathcal{M}_+(-t)$. Moreover, we find that the scaling function $\mathcal{M}_+$ satisfies the same equation that appears at critical jamming transitions [8,18]. At the same time, the function $P(q,h)$ admits the scaling form

$$P(q,h) = \begin{cases} p_+(h) & h \gg \sqrt{1-q} \\ (1-q)^{-a/\kappa} p_0\left(\frac{h}{\sqrt{1-q}}\right) & h \sim \sqrt{1-q} \\ \hat{\lambda}(q)^{-1} p_-(h\hat{\lambda}(q)^{-1}) & -h \sim \hat{\lambda}(q) \\ (1-q)^{-a/\kappa}\tilde{p}_0\left(\frac{h+\hat{\lambda}(q)}{\sqrt{1-q}}\right) & |h+\hat{\lambda}(q)| \sim \sqrt{1-q} \\ \tilde{p}_+(h+\hat{\lambda}(q)) & h+\hat{\lambda}(q) \ll -\sqrt{1-q}. \end{cases} \tag{8}$$

The scaling functions $p_+(t)$ and $\tilde{p}_+(t)$ control, respectively, the distribution of small positive and negative gaps. Furthermore, $\tilde{p}_0(t) = p_0(-t)$ and $p_0(t)$ satisfies again a scaling equation that is exactly the same as the one appearing at critical jamming [8,18]. This analysis implies that the exponents verify $\gamma_+ = \gamma_- = \gamma_J$ and $\theta = \theta' = \theta_J$, with $\gamma_J \simeq 0.41$ and $\theta_J \simeq 0.42$ the critical exponents controlling the gap and force distributions at the jamming point of hard spheres. Finally, the nature of the scaling solution implies that the distribution of gaps has an isostatic delta peak of marginally satisfied gaps, so we get

$$\rho(h) \sim \rho_+ h^{-\gamma}\theta(h) + \rho_-(-h)^{-\gamma}\theta(-h) + \frac{1}{\alpha}\delta(h), \qquad h \to 0, \tag{9}$$

with ρ_+ and ρ_- two positive constants.

Conclusions.—We have analyzed the properties of the UNSAT phase of the spherical perceptron with a linear cost function. In the RS phase the landscape is effectively convex and the global minimum is not critical and hypostatic. When instead the minimization is carried out in the RSB phase, we find that local minima are jamming critical. They are described by an isostatic number of contacts and the distributions of gaps and virtual forces follow power laws whose exponents are the same as the ones characterizing jamming of hard spheres. We have proposed a scaling solution of the RSB equations that agrees with the emerging criticality. There are two clear future directions. First, it will be interesting to understand what happens to the model if we change the cost function to a nonconvex function of the gaps ([.e., $p < 1$ in Eq. (2)]. Furthermore, it will be interesting to study linear cost functions in other CCSPs, and see if this leads to universal critical jammed phases as it happens in the perceptron. We expect that this property is generic within mean field, and our scaling solution extends to high dimensional spheres, multilayer neural nets, etc., [21,22]. More interesting are problems that are not mean field in nature. For finite dimensional spheres the critical exponents of jamming have been shown to be independent of spatial dimension within numerical accuracy [6]. It would be interesting to investigate if the same property holds for the jammed phase of linear soft spheres. We are working in this direction. This may provide a finite dimensional physical system with an extended jamming-critical phase where RSB effects could be tested.

We thank S. Hwang, J. Rocchi, and G. Parisi for discussions. S. F. and A. S. acknowledge a grant from the Simons Foundation (No. 454941, Silvio Franz). P. U. acknowledges the support of "Investissements d'Avenir" LabEx PALM (ANR-10-LABX-0039-PALM) (StatPhysDisSys project). S. F. is a member of the Institut Universitaire de France (IUF). This manuscript was partially prepared during the visit of P. U. to KITP and as such this research was supported in part by the National Science Foundation under Grant No. NSF PHY-1748958.

[1] A. J. Liu and S. R. Nagel, Annu. Rev. Condens. Matter Phys. **1**, 347 (2010).
[2] A. V. Tkachenko and T. A. Witten, Phys. Rev. E **60**, 687 (1999).
[3] M. Wyart, Phys. Rev. Lett. **109**, 125502 (2012).
[4] E. Lerner, G. During, and M. Wyart, Soft Matter **9**, 8252 (2013).
[5] M. Müller and M. Wyart, Annu. Rev. Condens. Matter Phys. **6**, 177 (2015).
[6] P. Charbonneau, E. I. Corwin, G. Parisi, and F. Zamponi, Phys. Rev. Lett. **114**, 125504 (2015).
[7] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, Nat. Commun. **5**, 3725 (2014).
[8] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, J. Stat. Mech. (2014) P10009.
[9] J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, J. Phys. Chem. B **117**, 12979 (2013).
[10] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[11] G. Biroli and P. Urbani, Nat. Phys. **12**, 1130 (2016).
[12] S. Franz and S. Spigler, Phys. Rev. E **95**, 022139 (2017).
[13] Y. Jin, P. Urbani, F. Zamponi, and H. Yoshino, Sci. Adv. **4**, eaat6387 (2018).
[14] Y. Jin and H. Yoshino, Nat. Commun. **8**, 14935 (2017).
[15] C. Rainone, P. Urbani, H. Yoshino, and F. Zamponi, Phys. Rev. Lett. **114**, 015701 (2015).
[16] S. Franz and G. Parisi, J. Phys. A **49**, 145001 (2016).
[17] S. Franz, G. Parisi, P. Urbani, and F. Zamponi, Proc. Natl. Acad. Sci. U.S.A. **112**, 14539 (2015).
[18] S. Franz, G. Parisi, M. Sevelev, P. Urbani, and F. Zamponi, SciPost Phys. **2**, 019 (2017).
[19] C. Moore and S. Mertens, *The Nature of Computation* (Oxford University Press, Oxford, England, 2011).
[20] A. Engel and C. Van den Broeck, *Statistical Mechanics of Learning* (Cambridge University Press, Cambridge, England, 2001).
[21] S. Franz, S. Hwang, and P. Urbani, arXiv:1809.09945.
[22] H. Yoshino, Sci. Post. Phys. **4**, 040 (2018).
[23] M. Griniasty and H. Gutfreund, J. Phys. A **24**, 715 (1991).
[24] G. Györgyi, Phys. Rep. **342**, 263 (2001).
[25] P. Majer, A. Engel, and A. Zippelius, J. Phys. A **26**, 7405 (1993).
[26] G. Györgyi and P. Reimann, Phys. Rev. Lett. **79**, 2746 (1997).
[27] G. Györgyi and P. Reimann, J. Stat. Phys. **101**, 679 (2000).
[28] J. Zhu, S. Rosset, R. Tibshirani, and T. J. Hastie, *Proceedings of the Advances in Neural Information Processing Systems* (2004), pp. 49–56.
[29] J. De Almeida and D. J. Thouless, J. Phys. A **11**, 983 (1978).
[30] R. Fletcher, *Practical Methods of Optimization* (John Wiley & Sons, New York, 2013).
[31] E. Jones, T. Oliphant, P. Peterson *et al.*, SciPy: Open source scientific tools for Python, 2001, http://www.scipy.org/.
[32] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.123.115702 for details on the numerical and theoretical analysis.
[33] The virtual forces can be obtained as the negative gaps in $\mathcal{D}$ rescaled by ε for $\varepsilon \to 0^+$.
[34] E. Gardner, J. Phys. A **21**, 257 (1988).
[35] H.-J. Sommers and W. Dupont, J. Phys. C **17**, 5785 (1984).