Journal: Physical Review E

Accession code: XW10618E

Article Title: How to count in hierarchical landscapes: A full solution to mean-field complexity

First Author: Jaron Kent-Dobias

AUTHOR QUERIES - TO BE ANSWERED BY THE CORRESPONDING AUTHOR

The following queries have arisen during the typesetting of your manuscript. Please answer these queries by marking the required corrections at the appropriate point in the text.

1	Refs. [9] to [55] out of order we have set sequence please check.
2	Please update all arXiv References.
3	Please update Refs [33,34,53] as able.
4	Please provide all information.
FQ	This funding provider could not be uniquely identified during our search of the FundRef registry (or no Contract or Grant number was detected). Please check information and amend if incomplete or incorrect.
Q	This reference could not be uniquely identified due to incomplete information or improper format. Please check all information and amend if applicable.

Important Notice to Authors

Attached is a PDF proof of your forthcoming article in PRE. Your article has 18 pages and the Accession Code is **XW10618E**. Your paper will be in the following section of the journal: RESEARCH ARTICLES

Please note that as part of the production process, APS converts all articles, regardless of their original source, into standardized XML that in turn is used to create the PDF and online versions of the article as well as to populate third-party systems such as Portico, Crossref, and Web of Science. We share our authors' high expectations for the fidelity of the conversion into XML and for the accuracy and appearance of the final, formatted PDF. This process works exceptionally well for the vast majority of articles; however, please check carefully all key elements of your PDF proof, particularly any equations or tables.

Figures submitted electronically as separate files containing color appear in color in the online journal. However, all figures will appear as grayscale images in the print journal unless the color figure charges have been paid in advance, in accordance with our policy for color in print (https://journals.aps.org/authors/color-figures-print).

No further publication processing will occur until we receive your response to this proof.

ORCIDs

Please follow any ORCID links ( ) after the authors' names and verify that they point to the appropriate record for each author. Requests to add ORCIDs should be sent no later than the first proof revisions. If authors do not subsequently add/authenticate ORCIDs within seven business days, production of the paper will proceed and no further requests to add ORCIDs will be processed. See complete details regarding ORCID requests and ORCID verification at https://journals.aps.org/authors/adding-orcids-during-proof-corrections.

NOTE: If this paper is an Erratum or a Reply, the corresponding author's ORCID may be present if previously provided to APS, but no ORCIDs can be added at proof stage.

Crossref Funder Registry ID

Information about an article's funding sources is now submitted to Crossref to help you comply with current or future funding agency mandates. Crossref's Funder Registry (https://www.crossref.org/services/funder-registry/) is the definitive registry of funding agencies. Please ensure that your acknowledgments include all sources of funding for your article following any requirements of your funding sources. Where possible, please include grant and award ids. Please carefully check the following funder information we have already extracted from your article and ensure its accuracy and completeness:

- Simons Foundation, 454943

Other Items to Check

- Please note that the original manuscript has been converted to XML prior to the creation of the PDF proof, as described above. Please carefully check all key elements of the paper, particularly the equations and tabular data.
- Title: Please check; be mindful that the title may have been changed during the peer-review process.
- Author list: Please make sure all authors are presented, in the appropriate order, and that all names are spelled correctly.
- Please make sure you have inserted a byline footnote containing the email address for the corresponding author, if desired. Please note that this is not inserted automatically by this journal.
- Affiliations: Please check to be sure the institution names are spelled correctly and attributed to the appropriate author(s).
- Receipt date: Please confirm accuracy.
- Acknowledgments: Please be sure to appropriately acknowledge all funding sources.
- Hyphenation: Please note hyphens may have been inserted in word pairs that function as adjectives when they occur before a noun, as in "x-ray diffraction," "4-mm-long gas cell," and "*R*-matrix theory." However, hyphens are deleted from word pairs when they are not used as adjectives before nouns, as in "emission by x rays," "was 4 mm in length," and "the *R* matrix is tested."

  Note also that Physical Review follows U.S. English guidelines in that hyphens are not used after prefixes or before suffixes: superresolution, quasiequilibrium, nanoprecipitates, resonancelike, clockwise.
- Please check that your figures are accurate and sized properly. Make sure all labeling is sufficiently legible. Figure quality in this proof is representative of the quality to be used in the online journal. To achieve manageable file size for online delivery, some compression and downsampling of figures may have occurred. Fine details may have become somewhat fuzzy, especially

in color figures. The print journal uses files of higher resolution and therefore details may be sharper in print. Figures to be published in color online will appear in color on these proofs if viewed on a color monitor or printed on a color printer.

- Please check to ensure that reference titles are given as appropriate.
- Overall, please proofread the entire *formatted* article very carefully. The redlined PDF should be used as a guide to see changes that were made during copyediting. However, note that some changes to math and/or layout may not be indicated.

Ways to Respond

- Web: If you accessed this proof online, follow the instructions on the web page to submit corrections.
- Email: Send corrections to preproofs@aptaracorp.com
  Subject: **XW10618E** proof corrections

PHYSICAL REVIEW E **00**, 004100 (2023)

How to count in hierarchical landscapes: A full solution to mean-field complexity

Jaron Kent-Dobias and Jorge Kurchan

Laboratoire de Physique de l'Ecole Normale Supérieure, Paris, France

(Received 5 September 2022; accepted 24 February 2023; published xxxxxxxxxx)

We derive the general solution for counting the stationary points of mean-field complex landscapes. It incorporates Parisi's solution for the ground state, as it should. Using this solution, we count the stationary points of two models: one with multistep replica symmetry breaking and one with full replica symmetry breaking.

DOI: 10.1103/PhysRevE.00.004100

I. INTRODUCTION

The computation of the number of metastable states of mean-field spin glasses goes back to the beginning of the field. Over 40 years ago, Bray and Moore [1] attempted the first calculation for the Sherrington–Kirkpatrick model, in a paper remarkable for being one of the first applications of a replica symmetry breaking (RSB) scheme. As was clear when the actual ground state of the model was computed by Parisi with a different scheme, the Bray–Moore result was not exact, and the problem has been open ever since [2]. To date, the program of computing the number of stationary points—minima, saddle points, and maxima—of mean-field complex landscapes has been only carried out for a small subset of models, including most notably the (pure) p-spin model ($p > 2$) [3–6] and for similar energy functions inspired by molecular biology, evolution, and machine learning [7–9]. In a parallel development, it has evolved into an active field of probability theory [10–12].

In this paper we present what we argue is the general replica ansatz for the number of stationary points of generic mean-field models, which we expect to include the Sherrington–Kirkpatrick model. It reproduces the Parisi result in the limit of small temperature for the lowest states, as it should.

To understand the importance of this computation, consider the following situation. When one solves the problem of spheres in large dimensions, one finds that there is a transition at a given temperature to a one-step replica symmetry breaking (1RSB) phase at a Kauzmann temperature, and, at a lower temperature, another transition to a full RSB (FRSB) phase (see Refs. [13,14], the so-called "Gardner" phase [15]). Now, this transition involves the lowest equilibrium states. Because they are obviously unreachable at any reasonable timescale, a common question is: what is the signature of the Gardner transition line for higher than equilibrium energy-densities? This is a question whose answers are significant to interpreting the results of myriad experiments and simulations [16–25] (see, for a review [26]). For example, when studying "jamming" at zero temperature, the question is posed as, "On what side of the 1RSB–FRSB transition are high-energy (or low-density) states reachable dynamically?" One approach to answering such questions makes use of "state following," which tracks metastable thermodynamic configurations to their zero temperature limit [27–31]. In the present paper we give a purely geometric approach: We consider the local energy minima at a given energy and study their number and other properties; the solution involves a replica-symmetry breaking scheme that is well-defined and corresponds directly to the topological characteristics of those minima.

Perhaps the most interesting application of this computation is in the context of optimization problems; see, for example, Refs. [32–34]. A question that appears there is how to define a "threshold" level, the lowest energy level that good algorithms can expect to reach. This notion was introduced in the context of the pure p-spin models, as the energy at which level sets of the energy in phase-space percolate, explaining why dynamics never go below that level [35]. The notion of a "threshold" for more complicated landscapes has later been invoked several times, never to our knowledge in a clear and unambiguous way. One of the purposes of this paper is to give a sufficiently detailed characterization of a general landscape so that a meaningful general notion of threshold may be introduced—if this is at all possible.

The format of this paper is as follows. In Sec. II, we introduce the mean-field model of study, the mixed p-spin spherical model. In Sec. III we review details of the equilibrium solution that are relevant to our study of the landscape complexity. In Sec. IV we derive a generic form for the complexity. In Sec. V we make and review the hierarchical replica symmetry breaking ansatz used to solve the complexity. In Sec. VI we write down the solution in a specific and limited regime, which is nonetheless helpful as it gives a foothold for numerically computing the complexity everywhere else. Sec. VII explains aspects of the solution specific to the case of full RSB, and derives the replica symmetric to full FRSB (RS–FRSB) transition line. Sec. VIII details the landscape topology of two example models: a $3+16$ model with a 2RSB ground state and a 1RSB complexity, and a $2+4$ with a FRSB ground state and a FRSB complexity. Finally Sec. IX provides some interpretation of our results.

II. THE MODEL

For definiteness, we consider the mixed p-spin spherical model, whose Hamiltonian

$$H(\mathbf{s}) = -\sum_p \frac{1}{p!}\sum_{i_1\cdots i_p}^N J^{(p)}_{i_1\cdots i_p} s_{i_1}\cdots s_{i_p} \tag{1}$$

is defined for vectors $\mathbf{s} \in \mathbb{R}^N$ confined to the sphere $\|\mathbf{s}\|^2 = N$. The coupling coefficients J are taken at random, with zero mean and variance $\overline{(J^{(p)})^2} = a_p p!/2N^{p-1}$ chosen so that the energy is typically extensive. The overbar will always denote an average over the coefficients J. The factors a_p in the variances are freely chosen constants that define the particular model. For instance, the so-called "pure" models have $a_p = 1$ for some p and all others zero.

The variance of the couplings implies that the covariance of the energy with itself depends on only the dot product (or overlap) between two configurations. In particular, one finds

$$\overline{H(\mathbf{s}_1)H(\mathbf{s}_2)} = Nf\left(\frac{\mathbf{s}_1 \cdot \mathbf{s}_2}{N}\right), \tag{2}$$

where f is defined by the series

$$f(q) = \frac{1}{2}\sum_p a_p q^p. \tag{3}$$

One need not start with a Hamiltonian like Eq. (1), defined as a series: instead, the covariance rule (2) can be specified for arbitrary, nonpolynomial f, as in the "toy model" of Mézard and Parisi [36].

The family of mixed p-spin models may be considered as the most general models of generic Gaussian functions on the sphere. To constrain the model to the sphere, we use a Lagrange multiplier μ, with the total energy being

$$H(\mathbf{s}) + \frac{\mu}{2}(\|\mathbf{s}\|^2 - N). \tag{4}$$

For reasons that will become clear in Sec. IV A 1, we refer to μ as the *stability parameter*. At any stationary point, the gradient and Hessian are given by

$$\nabla H(\mathbf{s}, \mu) = \partial H(\mathbf{s}) + \mu \mathbf{s}, \quad \operatorname{Hess} H(\mathbf{s}, \mu) = \partial\partial H(\mathbf{s}) + \mu I, \tag{5}$$

where $\partial = \frac{\partial}{\partial \mathbf{s}}$ always. An important observation was made by Bray and Dean [37] that gradient and Hessian are independent for Gaussian random functions. The average over disorder breaks into a product of two independent averages, one for any function of the gradient and one for any function of the Hessian. In particular, the number of negative eigenvalues at a stationary point, which sets the index $\mathcal{I}$ of the saddle, is a function of the Hessian alone (see Fyodorov [38] for a detailed discussion).

III. EQUILIBRIUM

Here we review the equilibrium solution, which has been studied in detail [39–42]. For a succinct review, see Ref. [43]. The free energy, averaged over disorder, is

$$\beta F = -\overline{\ln \int d\mathbf{s}\, \delta(\|\mathbf{s}\|^2 - N)\, e^{-\beta H(\mathbf{s})}}. \tag{6}$$

Once n replicas are introduced to treat the logarithm, the fields $\mathbf{s}_a$ can be replaced with the new $n \times n$ matrix field $Q_{ab} \equiv (\mathbf{s}_a \cdot \mathbf{s}_b)/N$. This yields for the free energy

$$\beta F = -1 - \ln 2\pi - \frac{1}{2}\lim_{n\to 0}\frac{1}{n}\left(\beta^2 \sum_{ab}^n f(Q_{ab}) + \ln\det Q\right), \tag{7}$$

which must be evaluated at the Q which maximizes this expression and whose diagonal is one. The solution is generally a hierarchical matrix à la Parisi. The properties of these matrices is reviewed in Sec. A, including how to write down Eq. (7) in terms of their parameters.

The free energy can also be written in a functional form, which is necessary for working with the solution in the limit $k \to \infty$, the so-called full replica symmetry breaking (FRSB). If $P(q)$ is the probability distribution for elements q in a row of the matrix, then define $\chi(q)$ by

$$\chi(q) = \int_q^1 dq' \int_0^{q'} dq''\, P(q''). \tag{8}$$

Since it is the double integral of a probability distribution, χ must be concave, monotonically decreasing, and have $\chi(1) = 0$ and $\chi'(1) = -1$. The function χ turns out to have an interpretation as the spectrum of the hierarchical matrix Q. Using standard arguments, the free energy can be written as a functional over χ as

$$\beta F = -1 - \ln 2\pi - \frac{1}{2}\int_0^1 dq\left(\beta^2 f''(q)\chi(q) + \frac{1}{\chi(q)}\right), \tag{9}$$

which must be maximized with respect to χ given the constraints outlined above.

In our study of the landscape, the free energy will not be directly relevant anywhere except at the ground state, when the temperature is zero or $\beta \to \infty$. Here, the measure will be concentrated in the lowest minima, and the average energy $\langle E\rangle_0 = \lim_{\beta\to\infty}\frac{\partial}{\partial\beta}\beta F$ will correspond to the ground-state energy E_0. The zero temperature limit is most easily obtained by putting $x_i = \tilde{x}_i x_k$ and $x_k = \tilde{\beta}/\beta$, $q_k = 1 - z/\beta$, which ensures the $\tilde{x}_i$, $\tilde{\beta}$, and z have nontrivial limits. Inserting the ansatz and taking the limit, carefully treating the kth term in each sum separately from the rest, one can show after some algebra that

$$\begin{aligned}\tilde{\beta}\langle E\rangle_0 &= \tilde{\beta}\lim_{\beta\to\infty}\frac{\partial(\beta F)}{\partial\beta} \\ &= -\frac{1}{2}z\tilde{\beta}f'(1) - \frac{1}{2}\lim_{n\to 0}\frac{1}{n} \\ &\quad\times\left[\tilde{\beta}^2\sum_{ab}^n f(\tilde{Q}_{ab}) + \ln\det(\tilde{\beta}z^{-1}\tilde{Q} + I)\right],\end{aligned} \tag{10}$$

where $\tilde{Q}$ is a $(k-1)$RSB matrix with entries $\tilde{q}_1 = \lim_{\beta\to\infty} q_1$, ..., $\tilde{q}_{k-1} = \lim_{\beta\to\infty} q_{k-1}$ parameterized by $\tilde{x}_1, \dots, \tilde{x}_{k-1}$. This is a $(k-1)$RSB ansatz whose spectrum in the determinant is scaled by $\tilde{\beta}z^{-1}$ and shifted by 1, with effective temperature $\tilde{\beta}$, and an extra term. In the continuum case, this is

$$\begin{aligned}\tilde{\beta}\langle E\rangle_0 = &-\frac{1}{2}z\tilde{\beta}f'(1) - \frac{1}{2}\int_0^1 dq \\ &\times\left[\tilde{\beta}^2 f''(q)\tilde{\chi}(q) + \frac{1}{\tilde{\chi}(q) + \tilde{\beta}z^{-1}}\right],\end{aligned} \tag{11}$$

where $\tilde{\chi}$ is bound by the same constraints as χ.

The zero temperature limit of the free energy loses one level of replica symmetry breaking. Physically, this is a result of the fact that in kRSB, q_k gives the overlap within a state, i.e., within the basin of a well inside the energy landscape. At zero temperature, the measure is completely localized on the bottom of the well, and therefore the overlap within each state becomes one. We will see that the complexity of low-energy stationary points in Kac–Rice computation is also given by a $(k-1)$RSB anstaz. Heuristically, this is because each stationary point also has no width and therefore overlap one with itself.

IV. LANDSCAPE COMPLEXITY

The stationary points of a function can be counted using the Kac–Rice formula, which integrates over the function's domain a δ function containing the gradient multiplied by the absolute value of the determinant [44,45]. It gives the number of stationary points $\mathcal{N}$ as

$$\mathcal{N} = \int d\mathbf{s}\, d\mu\, \delta\left(\frac{1}{2}(\|\mathbf{s}\|^2 - N)\right) \delta(\nabla H(\mathbf{s},\mu)) \times |\det \operatorname{Hess} H(\mathbf{s},\mu)|. \tag{12}$$

It is more interesting to count stationary points which share certain properties, like energy density E or index density $\mathcal{I}$. These properties can be fixed by inserting additional δ-functions into the integral. Rather than fix the index directly, we fix the trace of the Hessian, which we'll soon show is equivalent to fixing the value μ, and fixing μ fixes the index to within order one. Inserting these δ functions, we arrive at

$$\begin{aligned}\mathcal{N}(E,\mu^*) = \int d\mathbf{s}\, d\mu\, \delta\left(\frac{1}{2}(\|\mathbf{s}\|^2 - N)\right) \delta(\nabla H(\mathbf{s},\mu)) \\ \times |\det \operatorname{Hess} H(\mathbf{s},\mu)|\delta(NE - H(\mathbf{s})) \\ \times \delta(N\mu^* - \operatorname{Tr} \operatorname{Hess} H(\mathbf{s},\mu)).\end{aligned} \tag{13}$$

This number will typically be exponential in N. To find the typical count when disorder is averaged, we want to average its logarithm instead, which is known as the complexity:

$$\Sigma(E,\mu^*) = \lim_{N\to\infty} \frac{1}{N} \overline{\log \mathcal{N}(E,\mu^*)}. \tag{14}$$

If one averages over $\mathcal{N}$ and afterward takes its logarithm, then one arrives at the so-called *annealed* complexity

$$\Sigma_{\mathrm{a}}(E,\mu^*) = \lim_{N\to\infty} \frac{1}{N} \log \overline{\mathcal{N}(E,\mu^*)}. \tag{15}$$

The annealed complexity has been previously computed for the mixed p-spin models [12]. The annealed complexity is known to equal the actual (quenched) complexity in circumstances where there is at most one level of replica symmetry breaking in the model's equilibrium. This is the case for the pure p-spin models, or for mixed models where $1/\sqrt{f''(q)}$ is a convex function. However, it fails dramatically for models with higher replica symmetry breaking. For instance, when $f(q) = \frac{1}{2}(q^2 + \frac{1}{16}q^4)$ (a model we study in detail later), the annealed complexity predicts that minima vanish well before the dominant saddles, a contradiction for any bounded function.

A sometimes more illuminating quantity is the Legendre transform G of the complexity, defined by

$$e^{NG(\hat{\beta},\mu^*)} = \int dE\, e^{-\hat{\beta}E + \Sigma(\hat{\beta},\mu^*)}. \tag{16}$$

There will be a critical value $\hat{\beta}_c$ beyond which the complexity is zero: above this value the measure is split between the lowest $O(1)$ energy states. We shall not study here this regime that interpolates between the dynamically relevant and the equilibrium states, but just mention that it is an interesting object of study.

A. The replicated problem

The replicated Kac–Rice formula was introduced by Ros *et al.* [8], and its effective action for the mixed p-spin model has previously been computed by Folena *et al.* [46]. Here we review the derivation.

To average the complexity over disorder, we must deal with the logarithm. We use the standard replica trick to convert the logarithm into a product, which gives

$$\begin{aligned}\log \mathcal{N}(E,\mu^*) = \lim_{n\to 0} \frac{\partial}{\partial n} \mathcal{N}^n(E,\mu^*) = \lim_{n\to 0} \frac{\partial}{\partial n} \int \prod_a^n d\mathbf{s}_a\, d\mu_a\, \delta\left(\frac{1}{2}(\|\mathbf{s}_a\|^2 - N)\right) \delta(\nabla H(\mathbf{s}_a,\mu_a))\, |\det \operatorname{Hess} H(\mathbf{s}_a,\mu_a)| \\ \times \delta(NE - H(\mathbf{s}_a))\delta(N\mu^* - \operatorname{Tr}\operatorname{Hess} H(\mathbf{s}_a,\mu_a)).\end{aligned} \tag{17}$$

As discussed in Sec. II, it has been shown that to the largest order in N, the Hessian of Gaussian random functions is independent from their gradient, once both are conditioned on certain properties. Here, they are only related by their shared value of μ. Because of this statistical independence, we may write

$$\begin{aligned}\Sigma(E,\mu^*) = \lim_{N\to\infty} \frac{1}{N} \lim_{n\to 0} \frac{\partial}{\partial n} \int \left(\prod_a^n d\mathbf{s}_a\, d\mu_a\right) \overline{\prod_a^n \delta\left(\frac{1}{2}(\|\mathbf{s}_a\|^2 - N)\right) \delta(\nabla H(\mathbf{s}_a,\mu_a))\delta(NE - H(\mathbf{s}_a))} \\ \times \overline{\prod_a^n |\det \operatorname{Hess}(\mathbf{s}_a,\mu_a)|\, \delta(N\mu^* - \operatorname{Tr}\operatorname{Hess} H(\mathbf{s}_a,\mu_a))},\end{aligned} \tag{18}$$

which simplifies matters. The average of the two factors may now be treated separately.

1. The Hessian factors

The spectrum of the matrix $\partial\partial H(\mathbf{s})$ is uncorrelated from the gradient. In the large-N limit, for almost every point and realization of disorder it is a GOE matrix with variance

$$\overline{(\partial_i\partial_j H(\mathbf{s}))^2} = \frac{1}{N}f''(1)\delta_{ij}. \tag{19}$$

Therefore, in that limit its spectrum is given by the Wigner semicircle with radius $\sqrt{4f''(1)}$, or

$$\rho(\lambda) = \begin{cases} \frac{1}{2\pi f''(1)}\sqrt{4f''(1)-\lambda^2} & \lambda^2 \leqslant 4f''(1), \\ 0 & \text{otherwise}. \end{cases} \tag{20}$$

The spectrum of the Hessian $\operatorname{Hess} H(\mathbf{s},\mu)$ is the same semicircle shifted by μ, or $\rho(\lambda+\mu)$. The stability parameter μ thus fixes the center of the spectrum of the Hessian. The semicircle radius $\mu_m = \sqrt{4f''(1)}$ is a kind of threshold. When μ is taken to be within the range $\pm\mu_m$, the critical points have index density

$$\mathcal{I}(\mu) = \int_0^\infty d\lambda\,\rho(\lambda+\mu) = \frac{1}{2} - \frac{1}{\pi}\left[\arctan\left(\frac{\mu}{\sqrt{\mu_m^2-\mu^2}}\right) + \frac{\mu}{\mu_m^2}\sqrt{\mu_m^2-\mu^2}\right]. \tag{21}$$

When $\mu > \mu_m$, the critical points are minima whose sloppiest eigenvalue is $\mu - \mu_m$. When $\mu = \mu_m$, the critical points are marginal minima, with flat directions in their spectrum. This property of μ is why we've named it the stability parameter: it governs the stability of stationary points, and for unstable ones it governs their index.

To largest order in N, the average over the product of determinants factorizes into the product of averages, each of which is given by the same expression depending only on μ [8]. We therefore find

$$\overline{\prod_a^n |\det \operatorname{Hess}(\mathbf{s}_a,\mu_a)|\,\delta(N\mu^* - \operatorname{Tr}\operatorname{Hess} H(\mathbf{s}_a,\mu_a))} \to \prod_a^n e^{N\mathcal{D}(\mu_a)}\delta(N(\mu^*-\mu_a)), \tag{22}$$

where the function $\mathcal{D}$ is defined by

$$\begin{aligned}\mathcal{D}(\mu) &= \frac{1}{N}\overline{\ln|\det\operatorname{Hess} H(s,\mu)|} = \int d\lambda\,\rho(\lambda+\mu)\ln|\lambda| \\ &= \operatorname{Re}\left\{\frac{1}{2}\left[1 + \frac{\mu}{2f''(1)}(\mu - \sqrt{\mu^2-4f''(1)})\right] - \ln\left[\frac{1}{2f''(1)}(\mu - \sqrt{\mu^2-4f''(1)})\right]\right\}.\end{aligned} \tag{23}$$

By fixing the trace of the Hessian, we have effectively fixed the value of the stability μ in all replicas to the value μ^*.

(1) For $\mu^* < \mu_m$, this amounts to fixing the index density. Since the overwhelming majority of saddles have a semicircle distribution, the fluctuations are rarer than exponential.

(2) For the gapped case $\mu^* > \mu_m$, there is an exponentially small probability that $r = 1, 2, \dots$ eigenvalues detach from the semicircle in such a way that the index is in fact $N\mathcal{I} = r$. We shall not discuss these subextensive index fluctuations in this paper, the interested reader may find what is needed in Ref. [11].

2. The gradient factors

The δ functions in the remaining factor are treated by writing them in the Fourier basis. Introducing auxiliary fields $\hat{\mathbf{s}}_a$ and $\hat\beta$ for this purpose, for each replica replica one writes

$$\delta\left(\frac{1}{2}(\|\mathbf{s}_a\|^2 - N)\right)\delta(\nabla H(\mathbf{s}_a,\mu^*))\delta(NE - H(\mathbf{s}_a)) = \int \frac{d\hat\mu}{2\pi}\frac{d\hat\beta}{2\pi}\frac{d\hat{\mathbf{s}}_a}{(2\pi)^N} e^{\frac{1}{2}\hat\mu(\|\mathbf{s}_a\|^2-N)+\hat\beta(NE-H(\mathbf{s}_a))+i\hat{\mathbf{s}}_a\cdot(\partial H(\mathbf{s}_a)+\mu^*\mathbf{s}_a)}. \tag{24}$$

Anticipating a Parisi-style solution, we do not label $\hat\mu$ or $\hat\beta$ with replica indices, since replica vectors will not be broken in the scheme. The average over disorder can now be taken for the pieces which depend explicitly on the Hamiltonian, and since everything is Gaussian this gives

$$\begin{aligned}&\overline{\exp\left[\sum_a^n (i\hat{\mathbf{s}}_a\cdot\partial_a - \hat\beta)H(s_a)\right]} \\ &= \exp\left[\frac{1}{2}\sum_{ab}^n (i\hat{\mathbf{s}}_a\cdot\partial_a - \hat\beta)(i\hat{\mathbf{s}}_b\cdot\partial_b - \hat\beta)\overline{H(\mathbf{s}_a)H(\mathbf{s}_b)}\right] = \exp\left[\frac{N}{2}\sum_{ab}^n (i\hat{\mathbf{s}}_a\cdot\partial_a - \hat\beta)(i\hat{\mathbf{s}}_b\cdot\partial_b - \hat\beta)f\left(\frac{\mathbf{s}_a\cdot\mathbf{s}_b}{N}\right)\right] \\ &= \exp\left\{\frac{N}{2}\sum_{ab}^n\left[\hat\beta^2 f\left(\frac{\mathbf{s}_a\cdot\mathbf{s}_b}{N}\right) - 2i\hat\beta\frac{\hat{\mathbf{s}}_a\cdot\mathbf{s}_b}{N}f'\left(\frac{\mathbf{s}_a\cdot\mathbf{s}_b}{N}\right) - \frac{\hat{\mathbf{s}}_a\cdot\hat{\mathbf{s}}_b}{N}f'\left(\frac{\mathbf{s}_a\cdot\mathbf{s}_b}{N}\right) + \left(i\frac{\hat{\mathbf{s}}_a\cdot\mathbf{s}_b}{N}\right)^2 f''\left(\frac{\mathbf{s}_a\cdot\mathbf{s}_b}{N}\right)\right]\right\}.\end{aligned} \tag{25}$$

We introduce new matrix fields

$$C_{ab} = \frac{1}{N}\mathbf{s}_a \cdot \mathbf{s}_b, \quad R_{ab} = -i\frac{1}{N}\hat{\mathbf{s}}_a \cdot \mathbf{s}_b, \quad D_{ab} = \frac{1}{N}\hat{\mathbf{s}}_a \cdot \hat{\mathbf{s}}_b. \tag{26}$$

Their physical meaning is explained in Sec. IX. By substituting these parameters into the expressions above and then making a change of variables in the integration from $\mathbf{s}_a$ and $\hat{\mathbf{s}}_a$ to these three matrices, we arrive at the form for the complexity

$$\Sigma(E,\mu^*) = \mathcal{D}(\mu^*) + \hat{\beta}E - \frac{1}{2}\hat{\mu} + \lim_{n\to 0}\frac{1}{n} \times \left\{\frac{1}{2}\hat{\mu}\,\mathrm{Tr}\,C - \mu^*\,\mathrm{Tr}\,R + \frac{1}{2}\sum_{ab}[\hat{\beta}^2 f(C_{ab}) + (2\hat{\beta}R_{ab} - D_{ab})f'(C_{ab}) + R_{ab}^2 f''(C_{ab})] + \frac{1}{2}\ln\det\begin{bmatrix} C & iR \\ iR & D\end{bmatrix}\right\}, \tag{27}$$

where $\hat{\mu}$, $\hat{\beta}$, C, R, and D must be evaluated at the extrema of this expression which minimize the complexity. Note that one cannot *minimize* the complexity with respect to these parameters: there is no pure variational problem here. Extremizing with respect to $\hat{\mu}$ is not difficult, and results in setting the diagonal of C to one, fixing the spherical constraint. Maintaining $\hat{\mu}$ in the complexity is useful for writing down the extremal conditions, but when convenient we will drop the dependence.

The same information is contained but better expressed in the Legendre transform

$$G(\hat{\beta},\mu^*) = \mathcal{D}(\mu^*) + \lim_{n\to 0}\frac{1}{n}\left\{-\mu^*\,\mathrm{Tr}\,R + \frac{1}{2}\sum_{ab}[\hat{\beta}^2 f(C_{ab}) + (2\hat{\beta}R_{ab} - D_{ab})f'(C_{ab}) + R_{ab}^2 f''(C_{ab})] + \frac{1}{2}\ln\det\begin{bmatrix} C & iR \\ iR & D\end{bmatrix}\right\}. \tag{28}$$

Denoting $r_d \equiv \frac{1}{n}\mathrm{T}r R$, we can write down the double Legendre transform $K(\hat{\beta}, r_d)$:

$$e^{NK(\hat{\beta},r_d)} = \int dE\, d\mu^* e^{N\{\Sigma(E,\mu^*) - \hat{\beta}E + r_d\mu^* - \mathcal{D}(\mu^*)\}}, \tag{29}$$

given by

$$K(\hat{\beta}, r_d) = \lim_{n\to 0}\frac{1}{n}\left\{\frac{1}{2}\sum_{ab}[\hat{\beta}^2 f(C_{ab}) + (2\hat{\beta}R_{ab} - D_{ab})f'(C_{ab}) + R_{ab}^2 f''(C_{ab})] + \frac{1}{2}\ln\det\begin{bmatrix} C & iR \\ iR & D\end{bmatrix}\right\}, \tag{30}$$

where the diagonal of C is fixed to one and the diagonal of R is fixed to r_d. The variable r_d is conjugate to μ^* and through it to the index density, while $\hat{\beta}$ plays the role of an inverse temperature conjugate to the complexity, that has been used since the beginning of the spin-glass field. In this way $K(\hat{\beta}, r_d)$ contains all the information about saddle densities.

V. REPLICA ANSATZ

Based on previous work on the Sherrington–Kirkpatrick model and the equilibrium solution of the spherical model, we expect C, and R and D to be hierarchical matrices in Parisi's scheme. This assumption immediately simplifies the extremal conditions, since hierarchical matrices commute and are closed under matrix products and Hadamard products. In particular, the determinant of the block matrix can be written as a determinant of a product,

$$\ln\det\begin{bmatrix} C & iR \\ iR & D\end{bmatrix} = \ln\det(CD + R^2). \tag{31}$$

This is straightforward (if strenuous) to write down at kRSB, since the product and sum of the hierarchical matrices is still a hierarchical matrix. The algebra of hierarchical matrices is reviewed in Sec. A. Using the product formula (A3), one can write down the hierarchical matrix $CD + R^2$, and then compute the ln det using the formula (A2).

The extremal conditions are given by differentiating the complexity with respect to its parameters, yielding

$$0 = \frac{\partial\Sigma}{\partial\hat{\mu}} = \frac{1}{2}(c_d - 1), \tag{32}$$

$$0 = \frac{\partial\Sigma}{\partial\hat{\beta}} = E + \lim_{n\to 0}\frac{1}{n}\sum_{ab}[\hat{\beta}f(C_{ab}) + R_{ab}f'(C_{ab})], \tag{33}$$

$$0 = \frac{\partial\Sigma}{\partial C} = \frac{1}{2}[\hat{\mu}I + \hat{\beta}^2 f'(C) + (2\hat{\beta}R - D)\odot f''(C) + R\odot R\odot f'''(C) + (CD + R^2)^{-1}D], \tag{34}$$

$$0 = \frac{\partial\Sigma}{\partial R} = -\mu^* I + \hat{\beta}f'(C) + R\odot f''(C) + (CD + R^2)^{-1}R, \tag{35}$$

$$0 = \frac{\partial\Sigma}{\partial D} = -\frac{1}{2}f'(C) + \frac{1}{2}(CD + R^2)^{-1}C, \tag{36}$$

where $\odot$ denotes the Hadamard product, or the componentwise product. Equation (36) implies that

$$D = f'(C)^{-1} - RC^{-1}R. \tag{37}$$

To these conditions must be added the addition condition that Σ is extremal with respect to $x_1, \dots, x_k$. There is no better way to enforce this condition than to directly differentiate Σ with respect to the xs, and we have

$$0 = \frac{\partial\Sigma}{\partial x_i} 1 \leqslant i \leqslant k. \tag{38}$$

The stationary conditions for the xs are the most numerically taxing.

In addition to these equations, we often want to maximize the complexity as a function of μ^*, to find the most common type of stationary points. These are given by the condition

$$0=\frac{\partial\Sigma}{\partial\mu^*}=\mathcal{D}'(\mu^*)-r_d. \tag{39}$$

Since $\mathcal{D}(\mu^*)$ is effectively a piecewise function, with different forms for μ^* greater or less than μ_m, there are two regimes. When $\mu^*>\mu_m$ and the critical points are minima, Eq. (39) implies

$$\mu^*=\frac{1}{r_d}+r_df''(1). \tag{40}$$

When $\mu^*<\mu_m$ and the critical points are saddles, it implies

$$\mu^*=2f''(1)r_d. \tag{41}$$

It is often useful to have the extremal conditions in a form without matrix inverses, so that the saddle conditions can be expressed using products alone. By simple manipulations, the matrix equations can be written as

$$\begin{aligned}0=[\hat\beta^2f'(C)+(2\hat\beta R-D)\odot f''(C)\\+R\odot R\odot f'''(C)+\hat\mu I]C+f'(C)D,\end{aligned} \tag{42}$$

$$0=[\hat\beta f'(C)+R\odot f''(C)-\mu^*I]C+f'(C)R, \tag{43}$$

$$0=C-f'(C)(CD+R^2). \tag{44}$$

The right-hand side of each of these equations is also a hierarchical matrix, since products, Hadamard products, and sums of hierarchical matrices are such.

VI. SUPERSYMMETRIC SOLUTION

The Kac–Rice problem has an approximate supersymmetry, which is found when the absolute value of the determinant is neglected and the trace of the Hessian is not fixed. This supersymmetry has been studied in great detail in the complexity of the Thouless–Anderson–Palmer (TAP) free energy [47–51]. When the absolute value is dropped, the determinant in (12) can be represented by an integral over Grassmann variables, which yields a complexity depending on "bosons" and "fermions" that share the supersymmetry. The Ward identities associated with the supersymmetry imply that $D=\hat\beta R$ [47]. Under which conditions can this relationship be expected to hold? We find that their applicability is limited to a specific line in the energy and stability plane.

The identity $D=\hat\beta R$ heavily constrains the form that the rest of the solution can take. Assuming the supersymmetry holds, Eq. (34) implies

$$\begin{aligned}0=\hat\mu I+\hat\beta^2f'(C)+\hat\beta R\odot f''(C)\\+R\odot R\odot f'''(C)+\hat\beta(CD+R^2)^{-1}R.\end{aligned} \tag{45}$$

Substituting (35) for the factor $(CD+R^2)^{-1}R$, we find substantial cancellation, and finally

$$0=(\hat\mu+\mu^*)I+R\odot R\odot f'''(C). \tag{46}$$

If C has a nontrivial off-diagonal structure and supersymmetry holds, then the off-diagonal of R must vanish, and therefore $R=r_dI$. Therefore, a supersymmetric ansatz is equivalent to a *diagonal* ansatz for both R and D.

Supersymmetry has further implications. Equations (35) and (36) can be combined to find

$$I=R[\mu^*I-R\odot f''(C)]+(D-\hat\beta R)f'(C). \tag{47}$$

Assuming the supersymmetry holds implies that

$$I=R[\mu^*I-R\odot f''(C)]. \tag{48}$$

Understanding that R is diagonal, we find

$$\mu^*=\frac{1}{r_d}+r_df''(1), \tag{49}$$

which is precisely the condition (40) for dominant minima. Therefore, *the supersymmetric solution counts the most common minima* [49]. When minima are not the most common type of stationary point, the supersymmetric solution correctly counts minima that satisfy (40), but these do not have any other special significance.

Inserting the supersymmetric ansatz $D=\hat\beta R$ and $R=r_dI$, one gets for the complexity

$$\begin{aligned}\Sigma(E,\mu^*)=\mathcal{D}(\mu^*)+\hat\beta E-\mu^*r_d+\frac{1}{2}\hat\beta r_df'(1)\\+\frac{1}{2}r_d^2f''(1)+\frac{1}{2}\ln r_d^2+\frac{1}{2}\lim_{n\to0}\frac{1}{n}\\\times\left[\hat\beta^2\sum_{ab}f(C_{ab})+\ln\det((\hat\beta/r_d)C+I)\right].\end{aligned} \tag{50}$$

From here, it is straightforward to see that the complexity vanishes at the ground-state energy. First, in the ground-state minima will dominate (even if they are marginal), so we may assume Eq. (40). Then, taking $\Sigma(E_0,\mu^*)=0$, gives

$$\begin{aligned}\hat\beta E_0=-\frac{1}{2}r_d\hat\beta f'(1)-\frac{1}{2}\lim_{n\to0}\frac{1}{n}\\\times\left[\hat\beta^2\sum_{ab}^nf(C_{ab})+\ln\det\left(\hat\beta r_d^{-1}C+I\right)\right],\end{aligned} \tag{51}$$

which is precisely the ground-state energy predicted by the equilibrium solution (10) with $r_d=z$, $\hat\beta=\tilde\beta$, and $C=\tilde Q$.

Therefore, a $(k-1)$RSB *ansatz in Kac–Rice will predict the correct ground-state energy for a model whose equilibrium state at small temperatures is* kRSB Moreover, there is an exact correspondence between the saddle parameters of each. If the equilibrium is given by a Parisi matrix with parameters $x_1,\dots,x_k$ and $q_1,\dots,q_k$, then the parameters $\hat\beta$, $r_d,d_d,\tilde x_1,\dots,\tilde x_{k-1}$, and $c_1,\dots,c_{k-1}$ for the complexity in the ground state are

$$\begin{aligned}\hat\beta=\lim_{\beta\to\infty}\beta x_k,\quad\tilde x_i=\lim_{\beta\to\infty}\frac{x_i}{x_k},\quad c_i=\lim_{\beta\to\infty}q_i,\\r_d=\lim_{\beta\to\infty}\beta(1-q_k),\quad d_d=\hat\beta r_d.\end{aligned} \tag{52}$$

Unlike the case for the TAP complexity, this correspondence between landscape complexity and equilibrium solutions only exists at the ground state. We will see in our examples in Sec. VIII that there appears to be little correspondence between these parameters away from the ground state.

The supersymmetric solution produces the correct complexity for the ground state and for a class of minima, including dominant ones. Moreover, it produces the correct parameters for the fields C, R, and D at those points. This is an important foothold in the problem of computing the general complexity. The full saddle point equations at kRSB are not very numerically stable, and a "good" saddle point has a typically small radius of convergence under methods like Newton's algorithm. With the supersymmetric solution in hand, it is possible to take small steps in the parameter space to find nonsupersymmetric numeric solutions, each time ensuring the initial conditions for the solver are sufficiently close to the correct answer. This is the strategy we use in Sec. VIII.

VII. FULL REPLICA SYMMETRY BREAKING

This reasoning applies equally well to FRSB systems. In the end, when the limit of $n \to 0$ is taken, each matrix field can be represented in the canonical way by its diagonal and a continuous function on the domain [0,1] which parameterizes each of its rows, with

$$C \leftrightarrow [c_d, c(x)], \quad R \leftrightarrow [r_d, r(x)], \quad D \leftrightarrow [d_d, d(x)]. \tag{53}$$

The algebra of hierarchical matrices under this continuous parametrization is reviewed in Sec. A. With these substitutions, the complexity becomes

$$\begin{aligned}\Sigma(E,\mu^*) = {}&\mathcal{D}(\mu^*) + \hat\beta E - \mu^* r_d + \frac{1}{2}\big[\hat\beta^2 f(1) \\ &+ (2\hat\beta r_d - d_d)f'(1) + r_d^2 f''(1)\big] \\ &- \frac{1}{2}\int_0^1 dx[\hat\beta^2 f(c(x)) + (2\hat\beta r(x) - d(x))f'(c(x)) \\ &+ r(x)^2 f''(c(x))] + \frac{1}{2}\lim_{n\to 0}\frac{1}{n}\ln\det(CD + R^2).\end{aligned} \tag{54}$$

The formula for the determinant is complicated and can be found by using the product formula (A6) to write CD and R^2, summing them, and finally using the ln det formula (A9). The saddle point equations take the form

$$\begin{aligned}0 = {}&\hat\mu c(x) + [(\hat\beta^2(f'\circ c) + (2\hat\beta r - d)(f''\circ c) \\ &+ r^2(f'''\circ c)) * c](x) + [(f'\circ c) * d](x),\end{aligned} \tag{55}$$

$$\begin{aligned}0 = {}&-\mu^* c(x) + [(\hat\beta(f'\circ c) + r * (f''\circ c)) * c](x) \\ &+ [(f'\circ c) * r](x),\end{aligned} \tag{56}$$

$$0 = c(x) - [(f'\circ c) * (c * d + r * r)](x), \tag{57}$$

where $(ab)(x) = a(x)b(x)$ denotes the hadamard product, $(a * b)(x)$ denotes the functional parametrization of the diagonal of the product of hierarchical matrices AB defined in Eq. (A6), and $(a\circ b)(x) = a[b(x)]$ denotes composition.

A. Supersymmetric complexity

Using standard manipulations, one finds also a continuous version of the supersymmetric complexity

$$\begin{aligned}\Sigma(E,\mu^*) = {}&\mathcal{D}(\mu^*) + \hat\beta E - \mu^* r_d \\ &+ \frac{1}{2}\big[\hat\beta r_d f'(1) + r_d^2 f''(1) + \ln r_d^2\big] \\ &+ \frac{1}{2}\int_0^1 dq\left[\hat\beta^2 f''(q)\chi(q) + \frac{1}{\chi(q) + r_d/\hat\beta}\right],\end{aligned} \tag{58}$$

where $\chi(q) = \int_1^q dq' \int_0^{q'} dq''\, P(q)$ for $P(q)$ the distribution of elements in a row of C, as in the equilibrium case. Like in the equilibrium case, χ must be concave, monotonically decreasing, and have $\chi(1) = 0$, $\chi'(1) = -1$.

First, we use this solution to inspect the ground state of a full RSB system. We know from the equilibrium that in the ground state χ is continuous in the whole range of q. Therefore, the saddle solution found by extremizing

$$0 = \frac{\delta\Sigma}{\delta\chi(q)} = \frac{1}{2}\hat\beta^2 f''(q) - \frac{1}{2}\frac{1}{[\chi(q) + r_d/\hat\beta]^2} \tag{59}$$

over all functions χ. This gives

$$\chi_0(q \mid \hat\beta, r_d) = \frac{1}{\hat\beta}[f''(q)^{-1/2} - r_d]. \tag{60}$$

Satisfying the boundary conditions requires $r_d = f''(1)^{-1/2}$ and $\hat\beta = \frac{1}{2}f'''(1)/f''(1)^{3/2}$. This in turn implies $\mu^* = \frac{1}{r_d} + f''(1)r_d = \sqrt{4f''(1)} = \mu_m$. Therefore, the FRSB ground state is always marginal, as excepted. It is straightforward to check that these conditions are indeed a saddle of the complexity. This has several implications. First, other than the ground state, there are *no* energies at which minima are most numerous; saddles always dominate. As we will see, stable minima are numerous at energies above the ground state, but these vanish at the ground state.

Away from the ground state, this expression still correctly counts a class of nondominant minima. However, like in the equilibrium solution, the function χ which produces an extremal value is not smooth in the entire range [0,1], but adopts a piecewise form

$$\chi(q) = \begin{cases}\chi_0(q \mid \hat\beta, r_d) & q \leqslant q_{\max}, \\ 1 - q & \text{otherwise.}\end{cases} \tag{61}$$

With this ansatz, the complexity must be extremized with respect to r_d and $\hat\beta$, while simultaneously ensuring that $q_{\max}$ is such that $\chi(q)$ is continuous, that is, that $\chi_0(q_{\max} \mid \hat\beta, r_d) = 1 - q_{\max}$. The significance of the minima counted by this method is unclear, but they do represent a nodal line in the off-diagonal parts of R and D. Since, as usual, $\chi(q)$ is related to $c(x)$ by $-\chi'(c(x)) = x$, there is a corresponding $x_{\max}$ given by

$$x_{\max} = -\chi'(q_{\max}) = \frac{1}{2\hat\beta}\frac{f'''(q_{\max})}{f''(q_{\max})^{3/2}}. \tag{62}$$

B. Expansion near the transition

Working with the continuum equations away from the supersymmetric solution is not generally tractable. However, there is another point where they can be treated analytically: near the onset of replica symmetry breaking. Here, the off-diagonal components of C, R, and D are expected to be small. In particular, we expect the functions $c(x)$, $r(x)$, and $d(x)$ to approach zero at the transition, and moreover take the piecewise linear form

$$c(x) = \begin{cases}\bar{c}x & x \leqslant x_{\max} \\ \bar{c}x_{\max} & \text{otherwise}\end{cases}, \quad r(x) = \begin{cases}\bar{r}x & x \leqslant x_{\max} \\ \bar{r}x_{\max} & \text{otherwise}\end{cases},$$
$$d(x) = \begin{cases}\bar{d}x & x \leqslant x_{\max} \\ \bar{d}x_{\max} & \text{otherwise}\end{cases}, \tag{63}$$

with $x_{\max}$ vanishing at the transition, with the slopes $\bar{c}$, $\bar{r}$, and $\bar{d}$ remaining nonzero. This ansatz is informed both by the experience of the equilibrium solution, and by empirical observation within the numerics of Sec. VIII.

Given this ansatz, we take Eqs. (55)–(57), which are true for any x, and integrate them over x. We then expand the result about small $x_{\max}$ to linear order in $x_{\max}$. Equation (56) depends linearly on $\bar{r}$ to all orders, and therefore $\bar{r}$ can be found in terms of $\bar{c}$, yielding

$$\frac{\bar{r}}{\bar{c}} = -\hat{\beta} - \frac{1}{f'(1)+f''(0)}\{r_d[f''(0)+f''(1)] - \mu^*\} + O(x_{\max}). \tag{64}$$

Likewise, Eq. (57) depends linearly on $\bar{d}$ to all orders and can be solved to give

$$\frac{\bar{d}}{\bar{c}} = -2r_d\frac{\bar{r}}{\bar{c}} - \frac{1}{f'(1)}\left\{r_d^2 f''(0) + d_d[f'(1) + f''(0)] - 1\right\} + O(x_{\max}). \tag{65}$$

The equations cannot be used to find the value of $\bar{c}$ without going to higher order in $x_{\max}$, but the transition line can be determined by examining the stability of the replica symmetric complexity. First, we expand the full form for the complexity about small $x_{\max}$ in the same way as we expand the extremal conditions, using Eq. (A9) to treat the determinant. To quadratic order, this gives

$$\Sigma(E,\mu^*) = \mathcal{D}(\mu^*) + \hat{\beta}E - \mu r_d + \frac{1}{2}\left[\hat{\beta}^2 f(1) + (2\hat{\beta}r_d - d_d)f'(1) + r_d^2 f''(1)\right] + \frac{1}{2}\ln\left(d_d + r_d^2\right)$$
$$- \frac{1}{2}\left[\frac{1}{2}\hat{\beta}^2\bar{c}^2 f''(0) + (2\hat{\beta}\bar{r} - \bar{d})\bar{c}f''(0) + \bar{r}^2 f''(0) - \frac{\bar{d}^2 - 2d_d\bar{r}^2 + d_d^2\bar{c}^2 + 4r_d\bar{r}(\bar{d} + d_d\bar{c}) - 2r_d^2(\bar{c}\bar{d} + \bar{r}^2)}{2\left(d_d + r_d^2\right)^2}\right]x_{\max}^2. \tag{66}$$

The spectrum of the Hessian of Σ with evaluated at the RS solution gives its stability with respect to these functional perturbations. When the values of $\bar{r}$ and $\bar{d}$ above are substituted into the Hessian and $\hat{\beta}$, r_d, and d_d are evaluated at their RS values, the eigenvalue of interest takes the form

$$\lambda = -\bar{c}^2\frac{(f'(1) - 2f(1))^2(f'(1) - f''(0))f''(0)}{2(f'(1)+f''(0))(f'(1)^2 - f(1)(f'(1)+f''(1)))^2}(\mu^* - \mu_+^*(E))(\mu^* - \mu_-^*(E)), \tag{67}$$

where

$$\mu_\pm^*(E) = \pm\frac{(f'(1)+f''(0))(f'(1)^2 - f(1)(f'(1)+f''(1)))}{(2f(1)-f'(1))f'(1)f''(0)^{-1/2}} - \frac{f''(1)-f'(1)}{f'(1)-2f(1)}E. \tag{68}$$

This eigenvalue changes sign when μ^* crosses $\mu_\pm^*(E)$. We expect that this is the line of stability for the replica symmetric solution when the transition is RS-FRSB. The numerics in Sec. VIII bear this out.

VIII. GENERAL SOLUTION: EXAMPLES

Though we have only written down an easily computable complexity along a specific (and often uninteresting) line in energy and stability, this computable (supersymmetric) solution gives a numeric foothold for computing the complexity in the rest of that space. First, Eq. (11) is *maximized* with respect to its parameters, since the equilibrium solution is equivalent to a variational problem. Second, the mapping (52) is used to find the corresponding Kac–Rice saddle parameters in the ground state. With these parameters in hand, small steps are then made in energy E or stability μ, after which known these values are used as the initial condition for a saddle-finding problem. In this section, we use this basic numeric idea to map out the complexity for two representative examples: a model with a 2RSB equilibrium ground state and therefore 1RSB complexity in its vicinity, and a model with a FRSB equilibrium ground state, and therefore FRSB complexity as well.

A. 1RSB complexity

It is known that by choosing a covariance f as the sum of polynomials with well-separated powers, one develops 2RSB in equilibrium. This should correspond to 1RSB in Kac–Rice. For this example, we take

$$f(q) = \tfrac{1}{2}\left(q^3 + \tfrac{1}{16}q^{16}\right) \tag{69}$$

established to have a 2RSB ground state [52]. With this covariance, the model sees a replica symmetric to 1RSB transition at $\beta_1 = 1.70615\ldots$ and a 1RSB to 2RSB transition

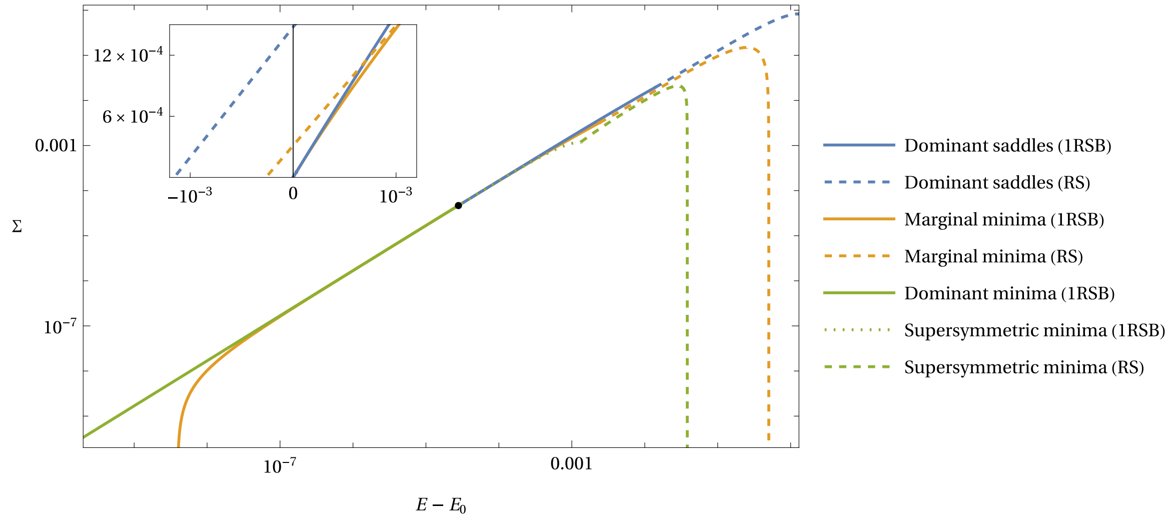

FIG. 1. Complexity of dominant saddles, marginal minima, and dominant minima of the 3 + 16 model. Solid lines show the result of the 1RSB ansatz, while the dashed lines show that of a RS ansatz. The complexity of marginal minima is always below that of dominant critical points except at the black dot, where they are dominant. The inset shows a region around the ground state and the fate of the RS solution.

at $\beta_2 = 6.02198\ldots$. At these transitions, the average energies in equilibrium are $\langle E\rangle_1 = -0.906391\ldots$ and $\langle E\rangle_2 = -1.19553\ldots$, respectively, and the ground-state energy is $E_0 = -1.287\,605\,530\ldots$. Besides these typical equilibrium energies, an energy of special interest for looking at the landscape topology is the *algorithmic threshold*E_{alg}, defined by the lowest energy reached by local algorithms like approximate message passing [53,54]. In the spherical models, this has been proven to be

$$E_{\text{alg}} = -\int_0^1 dq\,\sqrt{f''(q)}. \tag{70}$$

For full RSB systems, $E_{\text{alg}} = E_0$ and the algorithm can reach the ground-state energy. For the pure p-spin models, $E_{\text{alg}} = E_{\text{th}}$, where E_{th} is the energy at which marginal minima are the most common stationary points. Something about the topology of the energy function might be relevant to where this algorithmic threshold lies. For the 3 + 16 model at hand, $E_{\text{alg}} = -1.275\,140\,128\ldots$.

In this model, the RS complexity gives an inconsistent answer for the complexity of the ground state, predicting that the complexity of minima vanishes at a higher energy than the complexity of saddles, with both at a lower energy than the equilibrium ground state. The 1RSB complexity resolves these problems, predicting the same ground state as equilibrium and with a ground-state stability $\mu_0 = 6.480\,764\ldots > \mu_m$. It predicts that the complexity of marginal minima (and therefore all saddles) vanishes at $E_m = -1.287\,605\,527\ldots$, which is very slightly greater than E_0. Saddles become dominant over minima at a higher energy $E_{\text{th}} = -1.287\,575\,114\ldots$. The 1RSB complexity transitions to a RS description for dominant stationary points at an energy $E_1 = -1.273\,886\,852\ldots$. The highest energy for which the 1RSB description exists is $E_{\max} = -0.886\,029\,051\ldots$

The complexity as a function of energy difference from the ground state is plotted in Fig. 1. In that figure, the complexity is plotted for dominant minima and saddles, marginal minima, and supersymmetric minima. A contour plot of the complexity as a function of energy E and stability μ is shown in Fig. 2. That plot also shows the RS–1RSB transition line in the complexity. For minima, the complexity does not inherit a 1RSB description until the energy is with in a close vicinity of the ground state. However, for high-index saddles the complexity becomes described by 1RSB at quite high energies. This suggests that when sampling a landscape at high energies, high index saddles may show a sign of replica symmetry breaking when minima or inherent states do not.

Figure 3 shows a different detail of the complexity in the vicinity of the ground state, now as functions of the energy difference and stability difference from the ground state. Several of the landmark energies described above are plotted, alongside the boundaries between the "phases." Though E_{alg} looks quite close to the energy at which dominant saddles transition from 1RSB to RS, they differ by roughly 10^{-3}, as evidenced by the numbers cited above. Likewise, though $\langle E\rangle_1$ looks very close to $E_{\max}$, where the 1RSB transition line terminates, they too differ. The fact that E_{alg} is very slightly below the place where most saddle transition to 1RSB is suggestive; we speculate that an analysis of the typical minima connected to these saddles by downward trajectories will coincide with the algorithmic limit. An analysis of the typical nearby minima or the typical downward trajectories from these saddles at 1RSB is warranted [8,55]. Also notable is that E_{alg} is at a significantly higher energy than E_{th}; according to the theory, optimal smooth algorithms in this model stall in a place where minima are exponentially subdominant.

Figure 4 shows the saddle parameters for the 3 + 16 system for notable species of stationary points, notably the most common, the marginal ones, those with zero complexity, and those on the transition line. When possible, these are compared with the same expressions in the equilibrium solution at the same average energy. Besides the agreement at the ground-state energy, there seems to be little correlation between the equilibrium and complexity parameters.

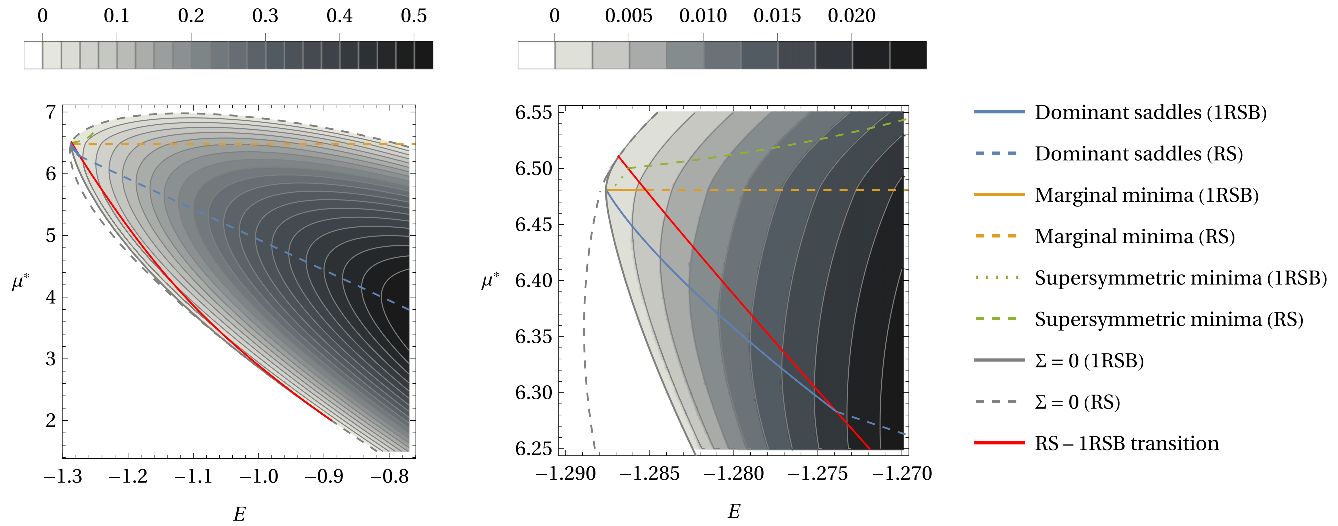

FIG. 2. Complexity of the 3 + 16 model in the energy E and stability μ^* plane. The right shows a detail of the left. Below the horizontal marginal line the complexity counts saddles of increasing index as μ^* decreases. Above the horizontal marginal line the complexity counts minima of increasing stability as μ^* increases.

Of specific note is what happens to d_1 as the 1RSB phase boundary for the complexity meets the zero complexity line. Here, d_1 diverges like

$$d_1 = -\left(\frac{1}{f'(1)} - \left(d_d + r_d^2\right)\right)(1 - x_1)^{-1} + O(1), \tag{71}$$

while x_1 and q_1 both go to one. Note that this is the only place along the phase boundary where q_1 goes to one. The significance of this critical point in the complexity of high-index saddles in worth further study.

B. Full RSB complexity

If the covariance f is chosen to be concave, then one develops FRSB in equilibrium. To this purpose, we choose

$$f(q) = \tfrac{1}{2}\left(q^2 + \tfrac{1}{16}q^4\right), \tag{72}$$

also studied before in equilibrium [41,42]. Because the ground state is FRSB, for this model

$$E_0 = E_{\text{alg}} = E_{\text{th}} = -\int_0^1 dq\,\sqrt{f''(q)} = -1.059\,384\,319\ldots. \tag{73}$$

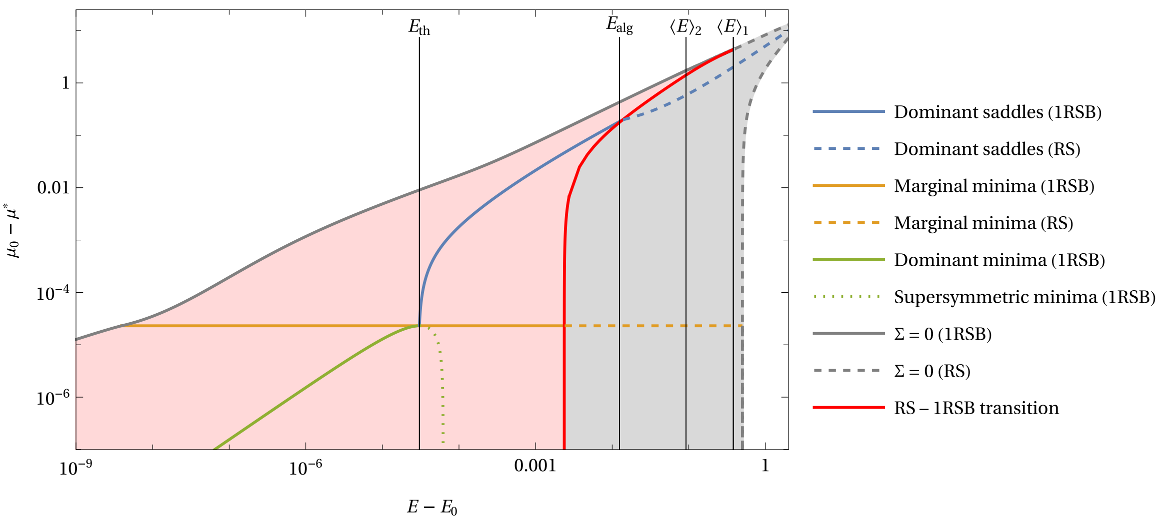

FIG. 3. Detail of the "phases" of the 3 + 16 model complexity as a function of energy and stability. Above the horizontal marginal stability line the complexity counts saddles of fixed index, while below that line it counts minima of fixed stability. The shaded red region to the left of the transition line shows places where the complexity is described by the 1RSB solution, while the shaded gray region to the right of the transition line shows places where the complexity is described by the RS solution. In white regions the complexity is zero. Several interesting energies are marked with vertical black lines: the traditional "threshold" E_{th} where minima become most numerous, the algorithmic threshold E_{alg} that bounds the performance of smooth algorithms, and the average energies at the 2RSB and 1RSB equilibrium transitions $\langle E\rangle_2$ and $\langle E\rangle_1$, respectively. Though the figure is suggestive, E_{alg} lies at slightly lower energy than the termination of the RS–1RSB transition line.

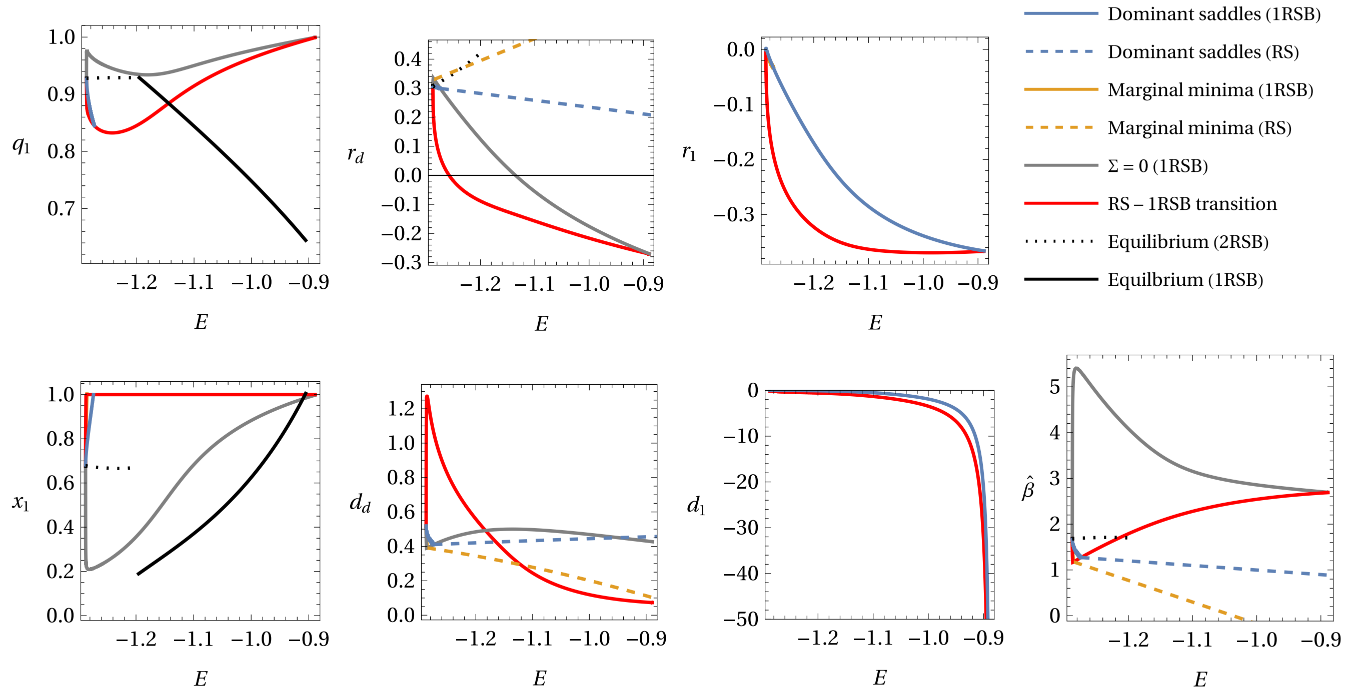

FIG. 4. Comparison of the saddle point parameters for the $3 + 16$ model along different trajectories in the energy and stability space, and with the equilibrium values (when they exist) at the same value of average energy $\langle E\rangle$.

In the equilibrium solution, the transition temperature from RS to FRSB is $\beta_\infty = 1$, with corresponding average energy $\langle E\rangle_\infty = -0.53125\ldots$.

Along the supersymmetric line, the FRSB solution can be found in full, exact functional form. To treat the FRSB away from this line numerically, we resort to finite kRSB approximations. Since we are not trying to find the actual kRSB solution, but approximate the FRSB one, we drop the extremal condition (38) for $x_1, \ldots, x_k$ and instead set

$$x_i = \left(\frac{i}{k+1}\right)x_{\max} \tag{74}$$

and extremize over $x_{\max}$ alone. This dramatically simplifies the equations that must be solved to find solutions. In the results that follow, a 20RSB approximation is used to trace the dominant saddles and marginal minima, while a 5RSB approximation is used to trace the (much longer) boundaries of the complexity.

Figure 5 shows the complexity for this model as a function of energy difference from the ground state for several notable trajectories in the energy and stability plane. Figure 6 shows these trajectories, along with the phase boundaries of the complexity in this plane. Notably, the phase boundary predicted by Eq. (68) correctly predicts where all of the finite kRSB approximations terminate. Like the 1RSB model in the previous subsection, this phase boundary is oriented such that very few, low energy, minima are described by a FRSB solution, while relatively high-energy saddles of high index are also. Again, this suggests that studying the mutual distribution of high-index saddle points might give insight into lower-energy symmetry breaking in more general contexts.

Figure 7 shows the value of $x_{\max}$ along several trajectories of interest. Everywhere along the transition line, $x_{\max}$ continuously goes to zero. Examples of our 20RSB approximations of the continuous functions $c(x)$, $r(x)$, and $d(x)$ are also shown. As expected, these functions approach linear ones as $x_{\max}$ goes to zero with finite slopes.

IX. INTERPRETATION

Let $\langle A\rangle$ be the average of any function A over stationary points with given E and μ^*, i.e.,

$$\langle A\rangle = \frac{1}{\mathcal{N}}\sum_{\mathbf{s}\in\mathcal{S}} A(\mathbf{s}) = \frac{1}{\mathcal{N}}\int d\nu(\mathbf{s})\, A(\mathbf{s}), \tag{75}$$

with

$$d\nu(\mathbf{s}) = d\mathbf{s}\, d\mu\, \delta\left(\frac{1}{2}(\|\mathbf{s}\|^2 - N)\right)\delta(\nabla H(\mathbf{s},\mu))\,|\det \operatorname{Hess} H(\mathbf{s},\mu)|\,\delta(NE - H(\mathbf{s}))\,\delta(N\mu^* - \operatorname{Tr}\operatorname{Hess} H(\mathbf{s},\mu)) \tag{76}$$

the Kac–Rice measure. Note that this definition of the angle brackets, which is in analogy with the typical equilibrium average, is not the same as that used in Sec. VII B for averaging over the off-diagonal elements of a hierarchical matrix. The fields C, R, and D defined in (26) can be related to certain averages of this type.

A. *C*: Distribution of overlaps

First consider C, which has an interpretation nearly identical to that of Parisi's Q matrix of overlaps in the equilibrium case. Its off-diagonal corresponds to the probability distribution $P(q)$ of the overlaps $q = (\mathbf{s}_1 \cdot \mathbf{s}_2)/N$ between stationary points. Let $\mathcal{S}$ be the set of all stationary points with given

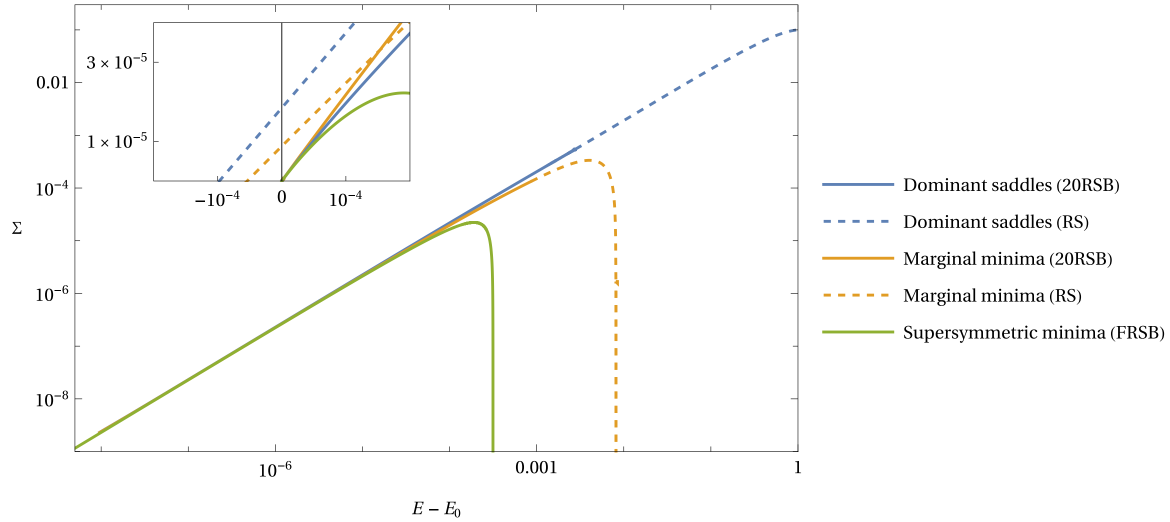

FIG. 5. The complexity Σ of the mixed $2+4$ spin model as a function of distance $\Delta E = E - E_0$ of the ground state. The solid blue line shows the complexity of dominant saddles given by the FRSB ansatz, and the solid yellow line shows the complexity of marginal minima. The dashed lines show the same for the annealed complexity. The inset shows more detail around the ground state.

energy density and index. Then

$$P(q) \equiv \frac{1}{\mathcal{N}^2} \sum_{\mathbf{s}_1 \in \mathcal{S}} \sum_{\mathbf{s}_2 \in \mathcal{S}} \delta\left(\frac{\mathbf{s}_1 \cdot \mathbf{s}_2}{N} - q\right). \tag{77}$$

This is the probability that two stationary points uniformly drawn from the ensemble of all stationary points with fixed E and μ^* happen to be at overlap q. Though these are evaluated for a given energy, index, etc, we shall omit these subindices for simplicity.

The moments of this distribution $q^{(p)}$ are given by

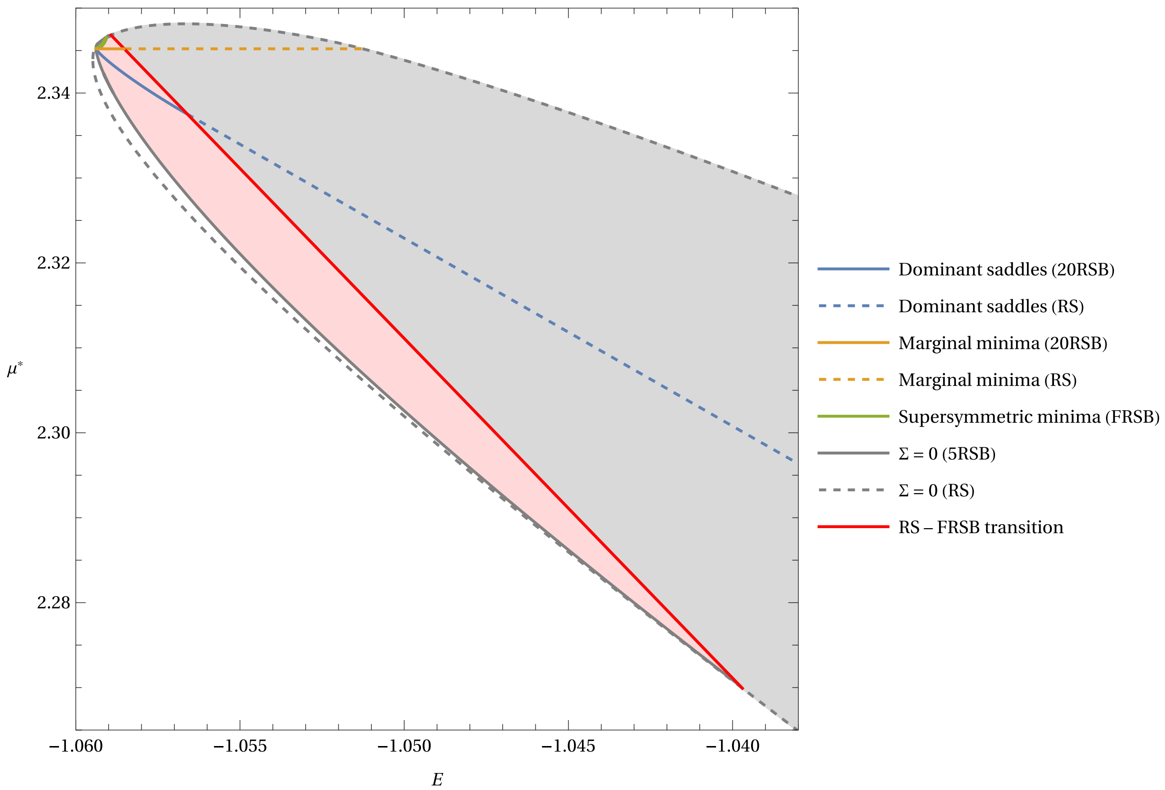

FIG. 6. "Phases" of the complexity for the $2+4$ model in the energy E and stability μ^* plane. The region shaded gray to the right of the transition line shows where the RS solution is correct, while the region shaded red to the left of the transition line shows that where the FRSB solution is correct. The white region shows where the complexity is zero.

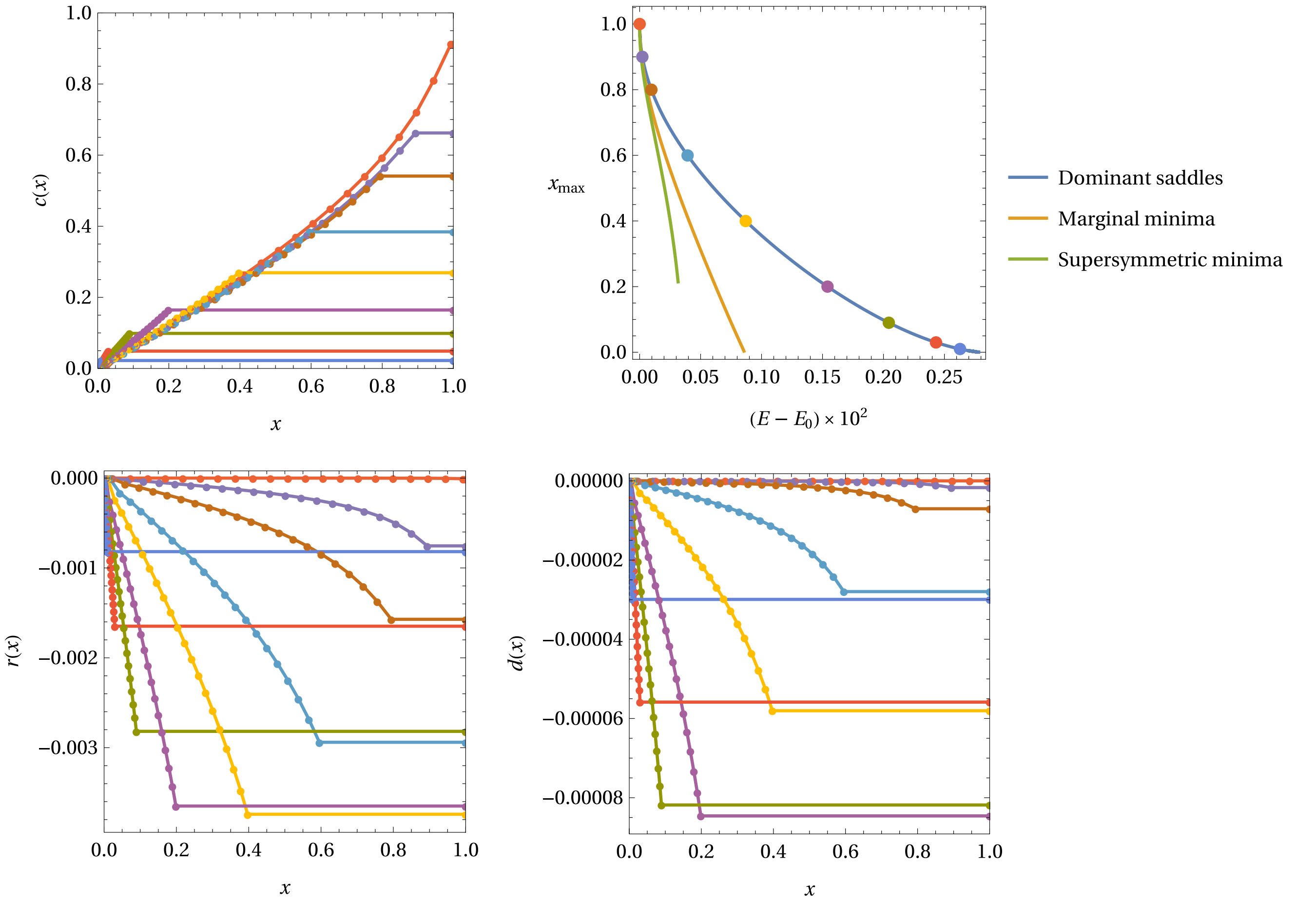

FIG. 7. $x_{\max}$ as a function of E for several trajectories of interest, along with examples of the 20RSB approximations of the functions $c(x)$, $r(x)$, and $d(x)$ along the dominant saddles. Colors of the approximate functions correspond to the points on the $x_{\max}$ plot. The supersymmetric line terminates where the complexity reaches zero, which happens inside the FRSB phase.

$$\begin{aligned} q^{(p)} &\equiv \int_0^1 dq\, q^p P(q) = \frac{1}{N^p}\sum_{i_1\cdots i_p}\langle s_{i_1}\cdots s_{i_p}\rangle\langle s_{i_1}\cdots s_{i_p}\rangle = \frac{1}{N^p}\frac{1}{\mathcal{N}^2}\left\{\sum_{\mathbf{s}_1,\mathbf{s}_2}\sum_{i_1\cdots i_p} s^1_{i_1}\cdots s^1_{i_p} s^2_{i_1}\cdots s^2_{i_p}\right\} \\ &= \frac{1}{\mathcal{N}^2}\left\{\sum_{\mathbf{s}_1,\mathbf{s}_2}\left(\frac{\mathbf{s}_1\cdot\mathbf{s}_2}{N}\right)^p\right\} = \lim_{n\to 0}\left\{\sum_{\mathbf{s}_1,\mathbf{s}_2,\dots,\mathbf{s}_n}\left(\frac{\mathbf{s}_1\cdot\mathbf{s}_2}{N}\right)^p\right\}. \end{aligned} \tag{78}$$

The $(n-2)$ extra replicas provide the normalization, with $\lim_{n\to 0}\mathcal{N}^{n-2} = \mathcal{N}^{-2}$. Replacing the sums over stationary points with integrals over the Kac–Rice measure, the average over disorder (again, for fixed energy and index) gives

$$\begin{aligned} \overline{q^{(p)}} &= \overline{\frac{1}{N^p}\sum_{i_1\cdots i_p}\langle s_{i_1}\cdots s_{i_p}\rangle\langle s_{i_1}\cdots s_{i_p}\rangle} = \lim_{n\to 0}\int \overline{\prod_a^n d\nu(\mathbf{s}_a)}\left(\frac{\mathbf{s}_1\cdot\mathbf{s}_2}{N}\right)^p \\ &= \lim_{n\to 0}\int D[C,R,D]\,(C_{12})^p\, e^{nN\Sigma[C,R,D]} = \lim_{n\to 0}\int D[C,R,D]\,\frac{1}{n(n-1)}\sum_{a\neq b}(C_{ab})^p\, e^{nN\Sigma[C,R,D]}. \end{aligned} \tag{79}$$

In the last line, we have used that there is nothing special about replicas one and two. Using the Parisi ansatz, evaluating by saddle point *summing over all the* $n(n-1)$ *saddles related by permutation* we then have

$$\overline{q^{(p)}} = \int_0^1 dx\, c^p(x) = \int_0^1 dq\, q^p P(q), \quad \text{concluding} \quad P(q) = \frac{dx}{dq} = \left(\frac{dc}{dx}\right)^{-1}\bigg|_{c(x)=q}. \tag{80}$$

FIG. 8. A cartoon visualizing how to interpret replica symmetry breaking solutions in the complexity. The black region show schematically areas where stationary points of a given energy can be found. Left: When the region is connected, pairs of stationary points exist at any overlap, but the vast majority of pairs are orthogonal. Center: When there are exponentially many disconnected regions of similar size, the vast majority of pairs will be found in different, orthogonal regions. Right: When there are a few large disconnected regions, pairs have a comparable probability to be found in different regions or in the same region. This gives rise to two (or more) possible overlaps.

The appeal of Parisi to properties of pure states is unnecessary here, since the stationary points are points.

With this established, we now address what it means for C to have a nontrivial replica-symmetry broken structure. When C is replica symmetric, drawing two stationary points at random will always lead to the same overlap. In the case when there is no linear field and $q_0 = 0$, they will always have overlap zero, because the second point will almost certainly lie on the equator of the sphere with respect to the first. Though other stationary points exist nearby the first one, they are exponentially fewer and so will be picked with vanishing probability in the thermodynamic limit.

When C is replica-symmetry broken, there is a nonzero probability of picking a second stationary point at some other overlap. This can be interpreted by imagining the level sets of the Hamiltonian in this scenario. If the level sets are disconnected but there are exponentially many of them distributed on the sphere, then one will still find zero average overlap. However, if the disconnected level sets are *few*, i.e., less than order N, then it is possible to draw two stationary points from the same set with nonzero probability. Therefore, the picture in this case is of few, large basins each containing exponentially many stationary points. A cartoon of this picture is shown in Fig. 8.

1. A tractable example

One can construct a schematic 2RSB model from two 1RSB models. Consider two independent pure models of size N and with p_1-spin and p_2-spin couplings, respectively, with energies $H_{p_1}(\mathbf{s})$ and $H_{p_2}(\boldsymbol{\sigma})$, and couple them weakly with $\varepsilon\,\boldsymbol{\sigma}\cdot\mathbf{s}$. The landscape of the pure models is much simpler than that of the mixed because, in these models, fixing the stability μ is equivalent to fixing the energy: $\mu = pE$. This implies that at each energy level there is only one type of stationary point. Therefore, for the pure models our formulas for the complexity and its Legendre transforms are functions of one variable only, E, and each instance of μ^* inside must be replaced with pE.

In the joint model, we wish to fix the total energy, not the energies of the individual two models. Therefore, we insert a δ function containing $(E_1 + E_2) - E$ and integrate over E_1 and E_2. This results in a joint complexity (and Legendre transform)

$$e^{N\Sigma(E)} = \int dE_1\,dE_2\,d\lambda\,\exp\{N[\Sigma_1(E_1) + \Sigma_2(E_2) + O(\varepsilon) - \lambda((E_1 + E_2) - E)]\}, \tag{81}$$

$$e^{NG(\hat\beta)} = \int dE\,dE_1\,dE_2\,d\lambda\,\exp\{N[-\hat\beta E + \Sigma_1(E_1) + \Sigma_2(E_2) + O(\varepsilon) - \lambda((E_1 + E_2) - E)]\}. \tag{82}$$

The saddle point is given by $\Sigma_1'(E_1) = \Sigma_2'(E_2) = \hat\beta$, provided that both $\Sigma_1(E_1)$ and $\Sigma_2(E_2)$ are nonzero. In this situation, two systems are "thermalized," and, because many points contribute, the overlap between two global configurations is zero:

$$\frac{1}{2N}\langle(\mathbf{s}^1,\boldsymbol{\sigma}^1)\cdot(\mathbf{s}^2,\boldsymbol{\sigma}^2)\rangle = \frac{1}{2N}[\langle\mathbf{s}^1\cdot\mathbf{s}^2\rangle + \langle\boldsymbol{\sigma}^1\cdot\boldsymbol{\sigma}^2\rangle] = 0. \tag{83}$$

This is the "annealed" phase of a Kac-Rice calculation.

Now start going down in energy, or up in $\hat\beta$: there will be a point E_c or $\hat\beta_c$ at which one of the subsystems (say it is system one) freezes at its lowest energy density, while system two is not yet frozen. At this point, $\Sigma_1(E_1) = 0$ and E_1 is the ground-state energy. At an even higher value $\hat\beta = \hat\beta_f$, both systems will become frozen in their ground states. For $\hat\beta_f > \hat\beta > \hat\beta_c$ one system is unfrozen, while the other is, because of coupling, frozen at inverse temperature $\hat\beta_c$. The overlap between two solutions in this intermediate phase is

$$\begin{aligned}\frac{1}{2N}\langle(\mathbf{s}^1,\boldsymbol{\sigma}^1)\cdot(\mathbf{s}^2,\boldsymbol{\sigma}^2)\rangle &= \frac{1}{2N}[\langle\mathbf{s}^1\cdot\mathbf{s}^2\rangle + \langle\boldsymbol{\sigma}^1\cdot\boldsymbol{\sigma}^2\rangle] \\ &= \frac{1}{2N}\langle\mathbf{s}^1\cdot\mathbf{s}^2\rangle > 0,\end{aligned} \tag{84}$$

which is nonzero because there are only a few low-energy stationary points in system one, and there is a nonvanishing probability of selecting one of them twice. The distribution of this overlap is one-half the overlap distribution of a frozen spin-glass at temperature $\hat\beta$, a 1RSB system like the random energy model. The value of x corresponding to it depends on $\hat\beta$, starting at $x=1$ at $\hat\beta_c$ and decreasing with increasing $\hat\beta$. Globally, the joint complexity of the system is 1RSB, but note that the global overlap between different states is at most $1/2$. At $\hat\beta>\hat\beta_f$ there is a further transition.

This schematic example provides a metaphor for considering what happens in ordinary models when replica symmetry is broken. At some point certain degrees of freedom "freeze" onto a subextensive number of possible states, while the remainder are effectively unconstrained. The overlap measures something in the competition between the number of these unconstrained subregions and their size.

B. *R* and *D*: Response functions

The matrix field R is related to responses of the stationary points to perturbations of the tensors J. One adds to the Hamiltonian a random term $\varepsilon_p\tilde H_p = -\frac{1}{p!}\varepsilon_p\sum_{i_1\cdots i_p}\tilde J_{i_1\cdots i_p}s_{i_1}\cdots s_{i_p}$, where the $\tilde J$ are random Gaussian uncorrelated with the Js and having variance $\overline{\tilde J^2}=p!/2N^{p-1}$. The response to these is

$$\frac{1}{N}\overline{\frac{\partial\langle\tilde H_p\rangle}{\partial\varepsilon_p}} = \lim_{n\to0}\int\left(\prod_a^n d\nu(\mathbf{s}_a)\right)\sum_b^n \times\left[\hat\beta\left(\frac{\mathbf{s}_1\cdot\mathbf{s}_b}{N}\right)^p + p\left(-i\frac{\mathbf{s}_1\cdot\hat{\mathbf{s}}_b}{N}\right)\left(\frac{\mathbf{s}_1\cdot\mathbf{s}_b}{N}\right)^{p-1}\right]. \tag{85}$$

Taking the average of this expression over disorder and averaging over the equivalent replicas in the integral gives, similar to before,

$$\begin{aligned}\frac{1}{N}\overline{\frac{\partial\langle\tilde H_p\rangle}{\partial\varepsilon_p}} &= \lim_{n\to0}\int D[C,R,D]\,\frac{1}{n}\sum_{ab}^n \times\left(\hat\beta C_{ab}^p + pR_{ab}C_{ab}^{p-1}\right)e^{nN\Sigma[C,R,D]}\\ &= \hat\beta + pr_d - \int_0^1 dx\, c^{p-1}(x)[\hat\beta c(x) + pr(x)].\end{aligned} \tag{86}$$

The responses as defined by this average perturbation in the pure p-spin energy can be directly related to responses in the tensor polarization of the stationary points:

$$\frac{1}{N^p}\sum_{i_1\cdots i_p}\overline{\frac{\partial\langle s_{i_1}\cdots s_{i_p}\rangle}{\partial J^{(p)}_{i_1\cdots i_p}}} = \frac{1}{N}\overline{\frac{\partial\langle\tilde H_p\rangle}{\partial\varepsilon_p}}. \tag{87}$$

In particular, when the energy is unconstrained ($\hat\beta=0$) and there is replica symmetry, the above formulas imply that

$$\frac{1}{N}\sum_i\frac{\partial\langle s_i\rangle}{\partial J_i^{(1)}} = r_d, \tag{88}$$

i.e., adding a linear field causes a response in the average stationary point location proportional to r_d. If positive, for instance, then stationary points tend to align with a field. The energy constraint has a significant contribution due to the perturbation causing stationary points to move up or down in energy.

The matrix field D is related to the response of the complexity to perturbations of the variance of the tensors J. This can be found by taking the expression for the complexity and inserting the dependence of f on the coefficients a_p, then differentiating:

$$\frac{\partial\Sigma}{\partial a_p} = \frac{1}{4}\lim_{n\to0}\frac{1}{n}\sum_{ab}^n\left[\hat\beta^2C_{ab}^p + p(2\hat\beta R_{ab} - D_{ab})C_{ab}^{p-1} + p(p-1)R_{ab}^2C_{ab}^{p-2}\right]. \tag{89}$$

In particular, when the energy is unconstrained ($\hat\beta=0$) and there is no replica symmetry breaking,

$$\frac{\partial\Sigma}{\partial a_1} = -\frac{1}{4}\lim_{n\to0}\frac{1}{n}\sum_{ab}D_{ab} = -\frac{1}{4}d_d, \tag{90}$$

i.e., adding a random linear field decreases the complexity of solutions by an amount proportional to d_d in the variance of the field.

When the saddle point of the Kac–Rice problem is supersymmetric,

$$\frac{\partial\Sigma}{\partial a_p} = \frac{\hat\beta}{4}\overline{\frac{1}{N^p}\sum_{i_1\cdots i_p}\frac{\partial\langle s_{i_1}\cdots s_{i_p}\rangle}{\partial J^{(p)}_{i_1\cdots i_p}}} + \lim_{n\to0}\frac{1}{n}\sum_{ab}^n p(p-1)R_{ab}^2C_{ab}^{p-2}, \tag{91}$$

and in particular for $p=1$,

$$\frac{\partial\Sigma}{\partial a_1} = \frac{\hat\beta}{4}\overline{\frac{1}{N}\sum_i\frac{\partial\langle s_i\rangle}{\partial J_i^{(1)}}}, \tag{92}$$

i.e., the change in complexity due to a linear field is directly related to the resulting magnetization of the stationary points for supersymmetric minima.

X. CONCLUSION

We have constructed a replica solution for the general problem of finding saddles of random mean-field landscapes, including systems with many steps of RSB. For systems with full RSB, we find that minima are exponentially subdominant with respect to saddles at all energy densities above the ground state. The solution should be subjected to standard checks, like the examination of its stability with respect to other RSB schemes. The solution contains valuable geometric information that has yet to be extracted in all detail, for example, considering several copies of the system [56], or the extension to complex variables [57,58].

A first and very important application of the method here is to perform the calculation for high dimensional spheres, where it would give us a clear understanding of what happens in realistic low-temperature jamming dynamics [59]. More simply, examining the landscape of a spherical model with a glass to glass transition from 1RSB to RS, like the $2+4$ model when a_4 is larger than we have taken it in our example, might give insight into the cases of interest for Gardner physics [41,42]. In any case, our analysis of typical 1RSB and FRSB landscapes indicates that the highest energy signature of RSB phases is in the overlap structure of the high-index saddle points. Though measuring the statistics of saddle points is difficult to imagine for experiments, this insight could find application in simulations of glass formers, where saddle-finding methods are possible.

A second application is to evaluate in more detail the landscape of these RSB systems. In particular, examining the complexity of stationary points with nonextensive indices (like rank-one saddles), the complexity of pairs of stationary points at fixed overlap, or the complexity of energy barriers [10,60]. These other properties of the landscape might shed light on the relationship between landscape RSB and dynamical features, like the algorithmic energy E_{alg}, or the asymptotic level reached by physical dynamics. For our 1RSB example, because E_{alg} is just below the energy where dominant saddles transition to a RSB complexity, we speculate that E_{alg} may be related to the statistics of minima connected to the saddles at this transition point.

ACKNOWLEDGMENTS

The authors thank Valentina Ros for helpful discussions. J.K.-D. and J.K. are supported by Simons Foundation Grant No. 454943.

APPENDIX: HIERARCHICAL MATRIX DICTIONARY

Each row of a hierarchical matrix is the same up to permutation of their elements. The so-called kRSB ansatz has $k+2$ different values in each row. If A is an $n\times n$ hierarchical matrix, then $n-x_1$ of those entries are a_0, x_1-x_2 of those entries are a_1, and so on until x_k-1 entries of a_k, and one entry of a_d, corresponding to the diagonal. Given such a matrix, there are standard ways of producing the sum and determinant that appear in the free energy. These formulas are, for an arbitrary kRSB matrix A with a_d on its diagonal (recall $q_d=1$),

$$\lim_{n\to0}\frac{1}{n}\sum_{ab}^{n}A_{ab}=a_d-\sum_{i=0}^{k}(x_{i+1}-x_i)a_i, \tag{A1}$$

$$\begin{aligned}\lim_{n\to0}\frac{1}{n}\ln\det A=&\frac{a_0}{a_d-\sum_{i=0}^{k}(x_{i+1}-x_i)a_i}+\frac{1}{x_1}\log\left[a_d-\sum_{i=0}^{k}(x_{i+1}-x_i)a_i\right]\\&-\sum_{j=1}^{k}\left(x_j^{-1}-x_{j+1}^{-1}\right)\log\left[a_d-\sum_{i=j}^{k}(x_{i+1}-x_i)a_i-x_ja_j\right],\end{aligned} \tag{A2}$$

where $x_0=0$ and $x_{k+1}=1$. The sum of two hierarchical matrices results in the sum of each of their elements: $(a+b)_d=a_d+b_d$ and $(a+b)_i=a_i+b_i$. The product AB of two hierarchical matrices A and B is given by

$$(a*b)_d=a_db_d-\sum_{j=0}^{k}(x_{j+1}-x_j)a_jb_j, \tag{A3}$$

$$(a*b)_i=b_da_i+a_db_i-\sum_{j=0}^{i-1}(x_{j+1}-x_j)a_jb_j+(2x_{i+1}-x_i)a_ib_i-\sum_{j=i+1}^{k}(x_{j+1}-x_j)(a_ib_j+a_jb_i). \tag{A4}$$

There is a canonical mapping between the parametrization of a hierarchical matrix described above and a functional parametrization that is particularly convenient in the twin limit $n\to0$ and $k\to\infty$ [61,62]. The distribution of diagonal elements of a matrix A is parameterized by a continuous function $a(x)$ on the interval [0,1], while its diagonal is still called a_d. Define for any function g the average

$$\langle g\rangle=\int_0^1dx\,g(x). \tag{A5}$$

The sum of two hierarchical matrices so parameterized results in the sum of these functions. The product AB of hierarchical matrices A and B gives

$$(a*b)_d=a_db_d-\langle ab\rangle, \tag{A6}$$

$$(a*b)(x)=(b_d-\langle b\rangle)a(x)+(a_d-\langle a\rangle)b(x)-\int_0^xdy\,[a(x)-a(y)][b(x)-b(y)]. \tag{A7}$$

The sum over all elements of a hierarchical matrix A gives

$$\lim_{n\to0}\frac{1}{n}\sum_{ab}A_{ab}=a_d-\langle a\rangle. \tag{A8}$$

810 The ln det = Tr ln becomes

$$\lim_{n\to 0}\frac{1}{n}\ln\det A = \ln(a_d - \langle a\rangle) + \frac{a(0)}{a_d - \langle a\rangle} - \int_0^1 \frac{dx}{x^2}\ln\left(\frac{a_d - \langle a\rangle - xa(x) + \int_0^x dy\, a(y)}{a_d - \langle a\rangle}\right). \tag{A9}$$

[1] A. J. Bray and M. A. Moore, Metastable states in spin glasses, J. Phys. C: Solid State Phys. **13**, L469 (1980).
[2] G. Parisi, Infinite Number of Order Parameters for Spin-Glasses, Phys. Rev. Lett. **43**, 1754 (1979).
[3] H. Rieger, The number of solutions of the Thouless-Anderson-Palmer equations for p-spin-interaction spin glasses, Phys. Rev. B **46**, 14655 (1992).
[4] A. Crisanti and H. J. Sommers, Thouless-Anderson-Palmer approach to the spherical p-spin spin glass model, J. Phys. I **5**, 805 (1995).
Q
[5] A. Cavagna, I. Giardina, and G. Parisi, An investigation of the hidden structure of states in a mean-field spin-glass model, J. Phys. A: Math. Gen. **30**, 7021 (1997).
[6] A. Cavagna, I. Giardina, and G. Parisi, Stationary points of the Thouless-Anderson-Palmer free energy, Phys. Rev. B **57**, 11251 (1998).
[7] A. Maillard, G. Ben Arous, G. Biroli, Landscape complexity for the empirical risk of generalized linear models, in *Proceedings of the 1st Mathematical and Scientific Machine Learning Conference*, edited by J. Lu and R. Ward, Proceedings of Machine Learning Research (July 2020), Vol. 107, pp. 287–327.
[8] V. Ros, G. Ben Arous, G. Biroli, and C. Cammarota, Complex Energy Landscapes in Spiked-Tensor and Simple Glassy Models: Ruggedness, Arrangements of Local Minima, and Phase Transitions, Phys. Rev. X **9**, 011003 (2019).
[9] A. Altieri, F. Roy, C. Cammarota, and G. Biroli, Properties of Equilibria and Glassy Phases of the Random Lotka-Volterra Model with Demographic Noise, Phys. Rev. Lett. **126**, 258301 (2021).
[10] A. Auffinger, G. Ben Arous, and J. Černý, Random matrices and complexity of spin glasses, Commun. Pure Appl. Math. **66**, 165 (2012).
[11] A. Auffinger and G. Ben Arous, Complexity of random smooth functions on the high-dimensional sphere, Ann. Probab. **41**, 4214 (2013).
[12] G. Ben Arous, E. Subag, and O. Zeitouni, Geometry and temperature chaos in mixed spherical spin glasses at low temperature: The perturbative regime, Commun. Pure Appl. Math. **73**, 1732 (2019).
[13] D. J. Gross, I. Kanter, and H. Sompolinsky, Mean-Field Theory of the Potts Glass, Phys. Rev. Lett. **55**, 304 (1985).
[14] E. Gardner, Spin glasses with p-spin interactions, Nucl. Phys. B **257**, 747 (1985).
[15] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, Fractal free-energy landscapes in structural glasses, Nat. Commun. **5**, 3725 (2014).
[16] H. Xiao, A. J. Liu, and D. J. Durian, Probing Gardner Physics in an Active Quasithermal Pressure-Controlled Granular System of Noncircular Particles, Phys. Rev. Lett. **128**, 248001 (2022).
[17] C. L. Hicks, M. J. Wheatley, M. J. Godfrey, and M. A. Moore, Gardner Transition in Physical Dimensions, Phys. Rev. Lett. **120**, 225501 (2018).
[18] Q. Liao and L. Berthier, Hierarchical Landscape of Hard Disk Glasses, Phys. Rev. X **9**, 011049 (2019).
[19] R. C. Dennis and E. I. Corwin, Jamming Energy Landscape is Hierarchical and Ultrametric, Phys. Rev. Lett. **124**, 078002 (2020).
[20] P. Charbonneau, Y. Jin, G. Parisi, C. Rainone, B. Seoane, and F. Zamponi, Numerical detection of the Gardner transition in a mean-field glass former, Phys. Rev. E **92**, 012316 (2015).
[21] H. Li, Y. Jin, Y. Jiang, and J. Z. Y. Chen, Determining the nonequilibrium criticality of a Gardner transition via a hybrid study of molecular simulations and machine learning, Proc. Natl. Acad. Sci. USA **118**, e2017392118 (2021).
[22] A. Seguin and O. Dauchot, Experimental Evidence of the Gardner Phase in a Granular Glass, Phys. Rev. Lett. **117**, 228001 (2016).
[23] K. Geirhos, P. Lunkenheimer, A. Loidl, Johari-Goldstein relaxation far below T_g: Experimental Evidence for the Gardner Transition in Structural Glasses? Phys. Rev. Lett. **120**, 085705 (2018).
[24] A. P. Hammond and E. I. Corwin, Experimental observation of the marginal glass phase in a colloidal glass, Proc. Natl. Acad. Sci. USA **117**, 5714 (2020).
[25] S. Albert, G. Biroli, F. Ladieu, R. Tourbot, and P. Urbani, Searching for the Gardner Transition in Glassy Glycerol, Phys. Rev. Lett. **126**, 028001 (2021).
[26] L. Berthier, G. Biroli, P. Charbonneau, E. I. Corwin, S. Franz, and F. Zamponi, Gardner physics in amorphous solids and beyond, J. Chem. Phys. **151**, 010901 (2019).
[27] C. Rainone, P. Urbani, H. Yoshino, and F. Zamponi, Following the Evolution of Hard Sphere Glasses in Infinite Dimensions Under External Perturbations: Compression and Shear Strain, Phys. Rev. Lett. **114**, 015701 (2015).
[28] G. Biroli and P. Urbani, Breakdown of elasticity in amorphous solids, Nat. Phys. **12**, 1130 (2016).
[29] C. Rainone and P. Urbani, Following the evolution of glassy states under external perturbations: The full replica symmetry breaking solution, J. Stat. Mech. (2016) 053302.
[30] G. Biroli and P. Urbani, Liu-Nagel phase diagrams in infinite dimension, SciPost Phys. **4**, 020 (2018).
[31] P. Urbani and F. Zamponi, Shear Yielding and Shear Jamming of Dense Hard Sphere Glasses, Phys. Rev. Lett. **118**, 038001 (2017).
[32] D. Gamarnik and A. Jagannath, The overlap gap property and approximate message passing algorithms for p-spin models, Ann. Probab. **49**, 180 (2021).

[33] A. El Alaoui, A. Montanari, M. Sellke, Sampling from the Sherrington-Kirkpatrick Gibbs measure via algorithmic
2 stochastic localization (2022), arXiv:2203.05093v1.

[34] B. Huang, M. Sellke, Tight Lipschitz hardness for optimizing
3 mean-field spin glasses (2021), arXiv:2110.07847v1.

[35] L. F. Cugliandolo and J. Kurchan, Analytical Solution of the off-Equilibrium Dynamics of a Long-Range Spin-Glass Model, Phys. Rev. Lett. **71**, 173 (1993).

[36] M. Mézard and G. Parisi, Manifolds in random media: Two
Q extreme cases, J. Phys. I **2**, 2231 (1992).

[37] A. J. Bray, D. S. Dean, Statistics of Critical Points of Gaussian Fields on Large-Dimensional Spaces, Phys. Rev. Lett. **98**, 150201 (2007).

[38] Y. V. Fyodorov and I. Williams, Replica symmetry breaking condition exposed by random matrix calculation of landscape complexity, J. Stat. Phys. **129**, 1081 (2007).

[39] A. Crisanti, H. J. Sommers, The spherical p-spin interaction spin glass model: The statics, Z. Phys. B **87**, 341 (1992).

[40] A. Crisanti, H. Horner, H. J. Sommers, The spherical p-spin interaction spin-glass model, Z. Phys. B **92**, 257 (1993).

[41] A. Crisanti, L. Leuzzi, Spherical $2 + p$ Spin-Glass Model: An Exactly Solvable Model for Glass to Spin-Glass Transition, Phys. Rev. Lett. **93**, 217203 (2004).

[42] A. Crisanti, L. Leuzzi, Spherical $2 + p$ spin-glass model: An analytically solvable model with a glass-to-glass transition, Phys. Rev. B **73**, 014412 (2006).

[43] G. Folena, The mixed p-spin model: Selecting, following and losing states, theses, Université Paris-Saclay & Universitá degli
4 studi La Sapienza, Rome (2020).

[44] S. O. Rice, The distribution of the maxima of a random curve, Am. J. Math. **61**, 409 (1939).

[45] M. Kac, On the average number of real roots of a random algebraic equation, Bull. Amer. Math. Soc. **49**, 314 (1943).

[46] G. Folena, S. Franz, F. Ricci-Tersenghi, Rethinking Mean-Field Glassy Dynamics and its Relation with the Energy Landscape: The Surprising Case of the Spherical Mixed p-spin model, Phys. Rev. X **10**, 031045 (2020).

[47] A. Annibale, A. Cavagna, I. Giardina, G. Parisi, and E. Trevigne, The role of the Becchi–Rouet–Stora–Tyutin supersymmetry in the calculation of the complexity for the Sherrington–Kirkpatrick model, J. Phys. A: Math. Gen. **36**, 10937 (2003).

[48] A. Annibale, A. Cavagna, I. Giardina, and G. Parisi, Supersymmetric complexity in the Sherrington-Kirkpatrick model, Phys. Rev. E **68**, 061103 (2003).

[49] A. Annibale, G. Gualdi, and A. Cavagna, Coexistence of supersymmetric and supersymmetry-breaking states in spherical spin-glasses, J. Phys. A: Math. Gen. **37**, 11311 (2004).

[50] A. Cavagna, I. Giardina, and G. Parisi, Cavity method for supersymmetry-breaking spin glasses, Phys. Rev. B **71**, 024422 (2005).

[51] I. Giardina, A. Cavagna, G. Parisi, Supersymmetry and metastability in disordered systems, in *Proceedings of the 31st Workshop of the International School of Solid State Physics*, edited by C. Beck, G. Benedek, A. Rapisarda, and C. Tsallis (2005), pp. 204–209.

[52] A. Crisanti, L. Leuzzi, and M. Paoluzzi, Statistical mechanical approach to secondary processes and structural relaxation in glasses and glass formers, Eur. Phys. J. E **34**, 98 (2011).

[53] A. El Alaoui, A. Montanari, Algorithmic thresholds in mean-field spin glasses (2020), arXiv:2009.11481v1.

[54] A. El Alaoui, A. Montanari, and M. Sellke, Optimization of mean-field spin glasses, Ann. Probab. **49**, 2922 (2021).

[55] V. Ros, G. Biroli, and C. Cammarota, Dynamical instantons and activated processes in mean-field glass models, SciPost Phys. **10**, 002 (2021).

[56] A. Cavagna, I. Giardina, and G. Parisi, Structure of metastable states in spin glasses by means of a three replica potential, J. Phys. A: Math. Gen. **30**, 4449 (1997).

[57] J. Kent-Dobias and J. Kurchan, Complex complex landscapes, Phys. Rev. Res. **3**, 023064 (2021).

[58] J. Kent-Dobias and J. Kurchan, Analytic continuation over complex landscapes, J. Phys. A: Math. Theor. **55**, 434006 (2022).

[59] T. Maimbourg, J. Kurchan, and F. Zamponi, Solution of the Dynamics of Liquids in the Large-Dimensional Limit, Phys. Rev. Lett. **116**, 015902 (2016).

[60] V. Ros, G. Biroli, and C. Cammarota, Complexity of energy barriers in mean-field glassy systems, Europhys. Lett. **126**, 20003 (2019).

[61] G. Parisi, Magnetic properties of spin glasses in a new mean-field theory, J. Phys. A: Math. Gen. **13**, 1887 (1980).

[62] M. Mézard and G. Parisi, Replica field theory for random man-
ifolds, J. Phys. I **1**, 809 (1991). Q