

Letter

Curvature-driven pathways interpolating between stationary points: the case of the pure spherical 3-spin model

Alessandro Pacco^{1,*} , Giulio Biroli²  and Valentina Ros¹ 

¹ Université Paris-Saclay, CNRS, LPTMS, 91405 Orsay, France

² Laboratoire de Physique de l'Ecole Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, F-75005 Paris, France

E-mail: alessandro.pacco@universite-paris-saclay.fr

Received 1 December 2023

Accepted for publication 18 January 2024

Published 31 January 2024



CrossMark

Abstract

This paper focuses on characterizing the energy profile along pathways connecting different regions of configuration space in the context of a prototypical glass model, the pure spherical p -spin model with $p = 3$. The study investigates pairs of stationary points (local minima or rank-1 saddles), analyzing the energy profile along geodesic paths and comparing them with ‘perturbed’ pathways correlated to the landscape curvature. The goal is to assess the extent to which information from the local Hessian matrices around stationary points can identify paths with lower energy barriers. Surprisingly, unlike findings in other systems, the direction of softest local curvature is not a reliable predictor of low-energy paths, except in the case in which the direction of softest curvature corresponds to an isolated mode of the Hessian. However, other information encoded in the local Hessian does allow the identification of pathways associated with lower energy barriers. We conclude commenting on implications for the system’s activated dynamics.

Keywords: spin glass, random matrix theory, complex random landscapes, disordered systems

Understanding the statistical properties of random energy or cost landscapes in high dimensions is a rather ubiquitous problem. In several cases one is interested in characterizing the profile of the landscape along pathways connecting different regions of configuration space; these regions may correspond to different local minima of an energy function [1, 2] (equivalently,

* Author to whom any correspondence should be addressed.

to fitness local maxima [3, 4]), or to different configurations reached by algorithms optimizing a cost function [5–7]. In many situations, in particular in several glassy systems [8–10], the height of the typical effective barriers crossed during the dynamics increases at low temperatures, thus leading to super-Arrhenius behaviors [11]. Characterizing the paths crossing these barriers and connecting low-energy configurations is a major challenge to develop the theory of glassy dynamics. Note that there are also other kinds of complex high-dimensional landscapes: some with a preferred minimum, as in the problem of protein folding [12], or with many flat minima separated by low barriers, as in machine learning [5, 7, 13, 14]. Here, we focus on the first class of energy landscapes and address the challenge outlined above.

We consider a prototypical glass model, the spherical p -spin model with $p = 3$, given by a random Gaussian function defined on a high-dimensional sphere [15, 16]. This landscape exhibits a multiplicity of stationary points (local minima, but also saddles) that are isolated in configuration space, and separated by energy barriers. We consider pairs of such stationary points and characterize the energy profile along simple paths interpolating between them. In particular, we compute the typical energy profile along the geodesic path, determining how the resulting energy barrier depends on the energy of the stationary points and on their distance in configuration space. We then compare the geodesic energy profile with the profile along ‘perturbed’ pathways, which follow directions in configuration space that are correlated to the landscape curvature around one of the two stationary points.

Our goal is to investigate to what extent one can identify good pathways (i.e. paths associated with lower energy barriers) using the information on the landscape around the two stationary points, encoded in their local Hessian. This question is motivated by studies of finite-dimensional systems of jammed and mildly supercooled particles. For the former, the softest Hessian modes at a given configuration provide directions toward low energy barriers for particle rearrangements [17]. For the latter, the spatial amplitudes of the softest Hessian modes correlate with the regions where irreversible rearrangements of particles take place [18]. Contrary to these findings, we show that in the case of the purely random landscape we consider in this work, the direction of softest local curvature is not a predictor of paths with lower energy barriers, except in the case in which the local Hessian has an isolated mode. However, we argue that having access to the whole local Hessian in general allows to identify pathways associated to lower energy barriers. In the conclusion we also discuss possible physical reasons for the discrepancy between our findings and the ones cited above.

The structure of the energy landscape. The energy function of the pure p -spin spherical model reads:

$$\mathcal{E}(\mathbf{s}) = \sqrt{\frac{p!}{2N^{p-1}}} \sum_{i_1 < \dots < i_p} a_{i_1 \dots i_p} s_{i_1} \dots s_{i_p}, \quad (1)$$

where $\mathbf{s} = (s_1, \dots, s_N) \in \mathbb{R}^N$ satisfies $\mathbf{s}^2 = \sum_{i=1}^N s_i^2 = N$. The couplings $a_{i_1 \dots i_p}$ are i.i.d Gaussian variables with zero mean and unit variance: we denote with $\mathbb{E}[\cdot]$ the average with respect to them. The energy (1) is a Gaussian function with covariance

$$\mathbb{E}[\mathcal{E}(\mathbf{s}^0) \mathcal{E}(\mathbf{s}^1)] = \frac{N}{2} [q(\mathbf{s}^0, \mathbf{s}^1)]^p, \quad (2)$$

where $q(\mathbf{s}^0, \mathbf{s}^1) = N^{-1} (\mathbf{s}^0 \cdot \mathbf{s}^1)$ is the overlap between the configurations, which measures how close they are in configuration space. We denote the energy density of a configuration by $\epsilon(\mathbf{s}) = \lim_{N \rightarrow \infty} N^{-1} \mathcal{E}(\mathbf{s})$. There are two peculiar values of energy density for this model, ϵ_{gs} and ϵ_{th} [15, 19–24]. For $\epsilon < \epsilon_{\text{gs}}$ no configuration is present for typical realizations of the randomness: ϵ_{gs} is the typical energy of the deepest minima, i.e. the ground states. For $\epsilon > \epsilon_{\text{gs}}$

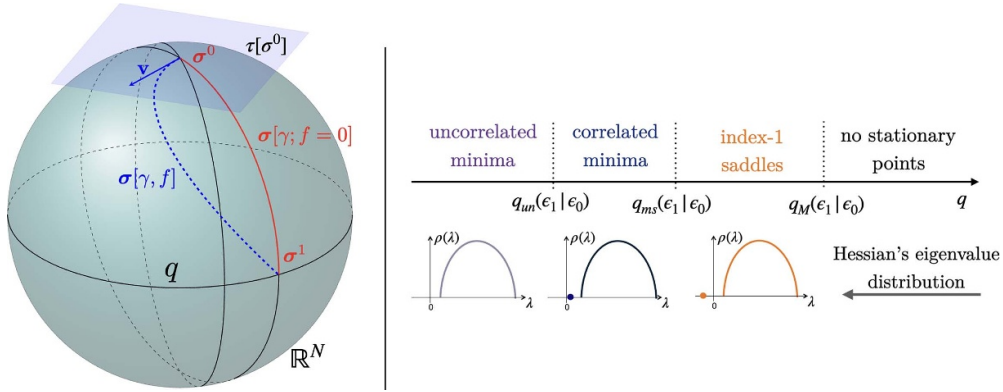


Figure 1. *Left.* Sketch of the configuration space (sphere), with a geodesic path (red thick line) and a perturbed path (blue dashed line) interpolating between two configurations $\sigma^a = s^a / \sqrt{N}$ with $a = 0, 1$. The tangent plane $\tau[\sigma^0]$ to which \mathbf{v} belongs is also sketched. *Right.* Eigenvalue distribution of the Hessian at σ^1 as a function of the overlap q between the σ^a , for fixed ϵ_a .

the landscape presents a number of stationary points (minima, saddles, maxima) that scales exponentially with N . Their stability changes at the threshold $\epsilon_{th} = -\sqrt{2(p-1)/p}$: below ϵ_{th} the exponential majority of the stationary points are local minima; the number of saddles with intensive index $k = o(N)$ [23, 25] (the index counting the number of directions in configuration space along which the curvature of the landscape is negative) is also exponentially large below ϵ_{th} , even though smaller than that of minima. Above ϵ_{th} the stationary points are overwhelmingly saddles with $k = O(N)$. The arrangement of these points in configuration space has also been investigated [19, 20, 24, 26, 27], by determining the typical number $\mathcal{N}_{s^0}(\epsilon, q | \epsilon_0)$ of stationary points \mathbf{s}^1 that are at fixed energy density $\epsilon_1 \in [\epsilon_{gs}, \epsilon_{th}]$ and overlap $\mathbf{s}^0 \cdot \mathbf{s}^1 = Nq$ with a reference minimum \mathbf{s}^0 of the landscape at arbitrary energy density $\epsilon_0 \in [\epsilon_{gs}, \epsilon_{th}]$. For large N , the scaling of the typical value of $\mathcal{N}_{s^0}(\epsilon, q | \epsilon_0)$ is controlled by the constrained complexity [26, 27]:

$$\Sigma(\epsilon_1, q | \epsilon_0) = \lim_{N \rightarrow \infty} \frac{\mathbb{E}[\log \mathcal{N}_{s^0}(\epsilon_1, q | \epsilon_0)]}{N}, \quad (3)$$

where the average is over both the population of local minima at energy density ϵ_0 at fixed randomness, and over the randomness. From now on, we focus on parameters for which $\Sigma(\epsilon_1, q | \epsilon_0) \geq 0$: negative values of this function indicate the region of parameters where stationary points are atypical, i.e. they are found with a probability that decays exponentially fast with N . By monitoring the stability of the stationary points at fixed ϵ_a (with $a = 0, 1$), one finds several transitions as a function of q , represented schematically in figure 1. For q smaller than a threshold $q_{un}(\epsilon_1 | \epsilon_0)$ they are typically local minima with a Hessian spectrum that shows no correlations with \mathbf{s}^0 . For intermediate $q_{un}(\epsilon_1 | \epsilon_0) \leq q \leq q_{ms}(\epsilon_1 | \epsilon_0)$ they are minima that are correlated to \mathbf{s}^0 , in the sense that the eigenvector associated to the minimal eigenvalue of their Hessian (which identifies the direction of minimal curvature of the landscape around the minimum) is oriented towards \mathbf{s}^0 . For large q , the stationary points are typically saddles of index $k = 1$: the landscape has a single direction of negative curvature, which is oriented towards the reference minimum \mathbf{s}^0 . For q larger than a threshold $q_M(\epsilon_1 | \epsilon_0)$, no stationary point is present. See appendix B for a concise account of these results. In this energy landscape thus

there are exponentially many (in N) saddles with energy below ϵ_{th} which surround the reference minimum, from which \mathbf{s}^0 is reachable by following the direction of negative curvature of the saddle [28]. These saddles are possible escape states for the system when it is dynamically trapped in \mathbf{s}^0 , and they are likely to be the first energy barriers that the system crosses in its dynamics, even though one expects frequent returns to \mathbf{s}^0 after these barriers crossings [28]. Given two local minima at arbitrary overlap q with each others, however, it is unknown how high in energy the system has to climb in order to transition between them. In particular, it is unknown if this energy barrier is always much higher than ϵ_{th} , or whether one can find pairs of distant minima that are connected by a sequence of saddles having all energies below ϵ_{th} . To get a proxy of these barriers, we compute the energy profile along simple paths interpolating between pairs of stationary points at arbitrary overlap q .

Interpolating paths and energy profiles. We introduce the vectors $\boldsymbol{\sigma} = \mathbf{s}/\sqrt{N}$ belonging to the N -dimensional sphere of unit radius, $\mathcal{S}_N(1)$, and the scaled energy function $h(\boldsymbol{\sigma}) := \mathcal{E}(\boldsymbol{\sigma}\sqrt{N})\sqrt{2/N}$. We denote with $\mathbf{g}(\boldsymbol{\sigma})$ and $\mathcal{H}(\boldsymbol{\sigma})$ the gradient and the Hessian matrix of the function $h(\boldsymbol{\sigma})$ restricted to $\mathcal{S}_N(1)$. For a fixed realization of the landscape, we consider two configurations $\boldsymbol{\sigma}^0$ and $\boldsymbol{\sigma}^1$ drawn at random from the population of stationary points of $h(\boldsymbol{\sigma})$ such that $\mathbf{g}(\boldsymbol{\sigma}^a) = 0$ for $a = 0, 1$. We extract the stationary points in such a way that each $\boldsymbol{\sigma}^a$ has energy density $\epsilon_a \in [\epsilon_{\text{gs}}, \epsilon_{\text{th}}]$ for $a = 0, 1$, and their overlap $\boldsymbol{\sigma}^0 \cdot \boldsymbol{\sigma}^1$ equals to some $q \in [0, 1]$. We are interested in the energy density profile along paths lying on the surface of $\mathcal{S}_N(1)$, which interpolate between the two stationary points. We parametrize the paths as follows:

$$\boldsymbol{\sigma}[\gamma; f] = \gamma \boldsymbol{\sigma}^1 + \beta[\gamma; f] \boldsymbol{\sigma}^0 + f(\gamma) \mathbf{v}, \quad \gamma \in [0, 1], \quad (4)$$

where f is a continuous function such that $f(0) = f(1) = 0$, and \mathbf{v} is a norm-1 vector orthogonal to both $\boldsymbol{\sigma}^0$ and $\boldsymbol{\sigma}^1$. The condition that $\boldsymbol{\sigma}[\gamma; f]$ lies on $\mathcal{S}_N(1)$ enforces

$$\beta[\gamma; f] = -\gamma q + \sqrt{1 - \gamma^2(1 - q^2) - [f(\gamma)]^2}. \quad (5)$$

Notice that this quantity has to be real, and this imposes some constraints on $f(\gamma)$. When $f \equiv 0$, equation (4) gives the geodesic path connecting the two stationary points. The function f acts as a perturbation of the geodesic, along the direction identified by the vector \mathbf{v} , see figure 1. Therefore, we are restricting to interpolating paths belonging to the low-dimensional subspace spanned by the vectors \mathbf{v} and $\boldsymbol{\sigma}^a$, intersected with the surface of the sphere. We aim at computing the typical energy profile

$$\epsilon_{\mathbf{v}}[\gamma; f] := \lim_{N \rightarrow \infty} \mathbb{E} \left[\frac{h(\boldsymbol{\sigma}[\gamma; f])}{\sqrt{2N}} \right], \quad (6)$$

where the average is over the distribution of stationary points $\boldsymbol{\sigma}^0, \boldsymbol{\sigma}^1$ with energy densities ϵ_0, ϵ_1 and overlap q , and over the realizations of the landscape. The average over $\boldsymbol{\sigma}^0$ and $\boldsymbol{\sigma}^1$ requires some care, since it can be performed within two different protocols (referred to as *quenched* and *annealed* in the terminology of glasses [29]); in appendix D we discuss both averaging schemes, arguing that for the problem at hand they give identical results.

To specify our choice of \mathbf{v} in (4), it is convenient to introduce the vectors

$$\mathbf{e}_{N-1}(\boldsymbol{\sigma}^0) = \frac{q \boldsymbol{\sigma}^0 - \boldsymbol{\sigma}^1}{\sqrt{1 - q^2}}, \quad \mathbf{e}_{N-1}(\boldsymbol{\sigma}^1) = \frac{q \boldsymbol{\sigma}^1 - \boldsymbol{\sigma}^0}{\sqrt{1 - q^2}}. \quad (7)$$

Each $\mathbf{e}_{N-1}(\boldsymbol{\sigma}^a)$ belongs to the tangent plane $\tau[\boldsymbol{\sigma}^a]$ to $\mathcal{S}_N(1)$ at the point $\boldsymbol{\sigma}^a$ (meaning that $\mathbf{e}_{N-1}(\boldsymbol{\sigma}^a) \perp \boldsymbol{\sigma}^a$), and identifies the direction in the tangent plane pointing towards the other configuration $\boldsymbol{\sigma}^{b \neq a}$. Let also \mathbf{x}_i with $i = 1, \dots, N-2$ be an arbitrary basis of the subspace orthogonal to the $\boldsymbol{\sigma}^a$. Finally, we denote with λ_{\min}^a the minimal eigenvalue of $\mathcal{H}(\boldsymbol{\sigma}^a)$ and with $\mathbf{e}_{\min}(\boldsymbol{\sigma}^a)$ the associated eigenvector, and let:

$$u^a = (\mathbf{e}_{\min}(\boldsymbol{\sigma}^a) \cdot \mathbf{e}_{N-1}(\boldsymbol{\sigma}^a))^2. \quad (8)$$

We consider two possible choices for \mathbf{v} . The first corresponds to

$$\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^a = \frac{\mathbf{e}_{\min}(\boldsymbol{\sigma}^a) - \sqrt{u^a} \mathbf{e}_{N-1}(\boldsymbol{\sigma}^a)}{\sqrt{1 - u^a}}, \quad (9)$$

meaning that the path is deformed in the directions of softest curvature of the energy landscape at $\boldsymbol{\sigma}^a$. The second one corresponds to

$$\mathbf{v} \rightarrow \mathbf{v}_{\text{Hess}} = Z \sum_{i=1}^{N-2} \left[\mathbf{x}_i \cdot \tilde{\mathcal{H}}(\boldsymbol{\sigma}^0) \cdot \mathbf{e}_{N-1}(\boldsymbol{\sigma}^0) \right] \mathbf{x}_i \quad (10)$$

where $\tilde{\mathcal{H}}(\boldsymbol{\sigma}^a) = \mathcal{H}(\boldsymbol{\sigma}^a) + p\sqrt{2N}\epsilon_a$ denote the Hessian matrices at the stationary points shifted by a constant, and Z a normalization constant. Consider now $\mathbf{v} \rightarrow \mathbf{v}_{\text{Hess}}$. This choice is motivated by the study of the gradient vector \mathbf{g} at each configuration $\boldsymbol{\sigma}(\gamma)$ along the geodesic path. As we show in appendix I, at each point the gradient has a tangent component \mathbf{g}^{\parallel} to the path, and an orthogonal component \mathbf{g}^{\perp} that is proportional to (10). While \mathbf{g}^{\parallel} obviously vanishes at the value of γ that corresponds to the local maximum of the geodesic energy profile, \mathbf{g}^{\perp} does not, meaning that the maximum of the geodesic profile is not a stationary point of $h(\boldsymbol{\sigma})$. This suggests that interpolating paths associated to lower barriers can be found by deforming the geodesic path in the direction of \mathbf{g}^{\perp} , since this is likely the direction that the path would follow if it was allowed to relax in configuration space by gradient descent [2, 30]. In fact, [6] developed an iterative algorithm to find low energy paths for deep neural networks based on this idea.

The case $p=3$ and the underlying random matrix problem. From now on we focus specifically on $p=3$. This restriction is motivated by the fact that in this case the energy profile (6) can be expressed as a function of the local properties of the landscape at $\boldsymbol{\sigma}^a$ only, i.e. of the local gradients $\mathbf{g}(\boldsymbol{\sigma}^a)$ and Hessian matrices $\mathcal{H}(\boldsymbol{\sigma}^a)$. By implementing the constraints $h(\boldsymbol{\sigma}^a) = \sqrt{2N}\epsilon_a$, $\mathbf{g}(\boldsymbol{\sigma}^a) = 0$ and by using the fact that typically $h(\mathbf{v}) = 0$, one obtains:

$$\begin{aligned} \epsilon_{\mathbf{v}}[\gamma; f] &= (\gamma^3 + 3\gamma^2\beta q) \epsilon_1 + (\beta^3 + 3\beta^2\gamma q) \epsilon_0 - \sqrt{1 - q^2}\gamma\beta f \mathbb{E} \left[\mathbf{v} \cdot \frac{\tilde{\mathcal{H}}(\boldsymbol{\sigma}^0)}{\sqrt{2N}} \cdot \mathbf{e}_{N-1}(\boldsymbol{\sigma}^0) \right] \\ &+ \frac{f^2}{2} \mathbb{E} \left[\gamma \mathbf{v} \cdot \frac{\tilde{\mathcal{H}}(\boldsymbol{\sigma}^1)}{\sqrt{2N}} \cdot \mathbf{v} + \beta \mathbf{v} \cdot \frac{\tilde{\mathcal{H}}(\boldsymbol{\sigma}^0)}{\sqrt{2N}} \cdot \mathbf{v} \right], \end{aligned} \quad (11)$$

where we omitted the dependence of f on γ and of β on γ and f . It follows that for $p=3$ and with our choices of \mathbf{v} , the energy profile depends only on correlations between the entries of the Hessian matrices, whose statistics has to be conditioned to the properties of $\boldsymbol{\sigma}^a$. This statistics is described in detail in appendix E. With the choice (9), in particular, the profile depends on

the matrix elements of the Hessian $\tilde{\mathcal{H}}(\boldsymbol{\sigma}^1)$ on the minimal eigenvector of the Hessian $\mathcal{H}(\boldsymbol{\sigma}^0)$ (equivalently, $\tilde{\mathcal{H}}(\boldsymbol{\sigma}^0)$), or vice-versa. For any $q > 0$, the two matrices are correlated (due to the fact that the landscape at the two points is correlated as (2)) and thus the matrix element is non-trivial. As we argue in appendix G, to determine its typical behavior for large N one needs to compute the typical value of the overlap between arbitrary eigenvectors of the two correlated matrices. This random matrix problem is discussed in recent literature for standard random matrix ensembles [31, 32]; for the particular type of random matrix ensembles describing the statistics of the Hessians $\mathcal{H}(\boldsymbol{\sigma}^a)$, this overlap function has been computed explicitly in [33], and the results are recalled in appendix G. Below, we discuss the properties of the energy profiles that are obtained making use of these results.

Results. For $\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^0$, the interpolating path is as in figure 1. When $\epsilon_0 < \epsilon_{\text{th}}$, the reference stationary point $\boldsymbol{\sigma}^0$ is typically a local minimum: the shifted Hessian $N^{-\frac{1}{2}}\tilde{\mathcal{H}}(\boldsymbol{\sigma}^0)$ has the statistics of matrices extracted from a Gaussian Orthogonal Ensemble (GOE) with variance $6/N$. The eigenvalues distribution of the shifted Hessian follows a semicircular law, and the minimal eigenvalue equals to $-2\sqrt{6}$. The associated eigenvector will show no condensation in the direction of $\mathbf{e}_{N-1}(\boldsymbol{\sigma}^0)$, meaning that typically u^0 takes the value $u_{\text{typ}}^0 = 0$. The path along (4) then reads (see appendix H for a derivation)

$$\epsilon_{\mathbf{v}^0_{\text{soft}}}[\gamma;f] = (\gamma^3 + 3\gamma^2\beta[\gamma;f]q)\epsilon_1 + (\beta^3[\gamma;f] + 3\beta^2[\gamma;f]\gamma q)\epsilon_0 - \sqrt{3}f^2(\gamma)(\beta[\gamma;f] + \gamma q). \quad (12)$$

One finds that $\delta\epsilon_{\mathbf{v}^0_{\text{soft}}}/\delta f = 0$ is satisfied by $f \equiv 0$, and that for $\epsilon_a < \epsilon_{\text{th}}$ it holds $\delta^2\epsilon_{\mathbf{v}^0_{\text{soft}}}/\delta f^2 > 0$ at $f \equiv 0$, meaning that the geodesic path is a minimum of the functional (12). Arbitrary deformations of the interpolating path in the direction of softest curvature at $\boldsymbol{\sigma}^0$ go through regions of the landscape of higher energy density, on average. This observation is confirmed by figure 2 *bottom*, which shows two density plots of equation (12) as a function of γ and f , where f is allowed to take any value within its range of validity that keeps β in (5) well defined. Arbitrary paths are obtained drawing curves connecting the points $f(0) = 0$ and $f(1) = 0$ in a continuous and injective fashion. One finds that the energy profile along these curves is non-monotonic, with a local maximum whose energy we refer to as the energy barrier. In figure 2 the parameters ϵ_0 and ϵ_1 are fixed, while q is tuned in such a way that $\boldsymbol{\sigma}^1$ is either a minimum with an Hessian with a single isolated eigenvalue (*left*) or an index-1 saddle (*right*). The white dotted lines represent the level curves of value ϵ_{th} . The plot confirms that the lowest energy barrier is obtained for the geodesic path $f \equiv 0$, and shows that such barrier is well above ϵ_{th} . The same results are obtained for different values of ϵ_a and q . One finds that the barrier associated to the geodesic path increases when the overlap q decreases and when ϵ_1 increases towards ϵ_{th} . We remark that the profile (12) is not the same that one would obtain choosing \mathbf{v} in (11) as a purely random Gaussian vector \mathbf{v}_{rand} , uncorrelated to the local Hessian: with that choice, all the terms depending on \mathbf{v} in equation (11) vanish on average, and one obtains $\epsilon_{\mathbf{v}_{\text{rand}}}[\gamma;f] = (\gamma^3 + 3\gamma^2\beta q)\epsilon_1 + (\beta^3 + 3\beta^2\gamma q)\epsilon_0$. For fixed $f(\gamma)$ this energy profile is systematically higher than (12). However, the functional is again minimized by $f \equiv 0$.

The case $\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^1$ is richer. In this case, depending on the values of ϵ_a and q , the stationary point $\boldsymbol{\sigma}^1$ is either a rank-1 saddle ($q_{\text{ms}} < q \leq q_{\text{M}}$), a minimum with one isolated mode in the Hessian ($q_{\text{un}} < q \leq q_{\text{ms}}$), or an uncorrelated minimum ($q < q_{\text{un}}$), see figure 1. Whenever the Hessian at $\boldsymbol{\sigma}^1$ has an isolated eigenvalue ($q > q_{\text{un}}$), it is the smallest eigenvalue and its typical value and the typical value u_{typ}^1 are given by

$$\lambda_{\text{typ}}^1 = \frac{3\delta\epsilon_q}{\sqrt{2}q} \left[\frac{1+3q^2}{1-q^2} - \sqrt{1 - \frac{2(1-q^2)^2}{3(1+q^2)\delta\epsilon_q^2}} \right], \quad (13)$$

$$u_{\text{typ}}^1 = \frac{1+q^2}{1+3q^2 - \frac{\sqrt{2}(q-q^3)\lambda_{\text{typ}}^1}{3\delta\epsilon_q}} \left[1 - \frac{(1-q^2)g(\lambda_{\text{typ}}^1)}{12\sqrt{2}(q+q^3)\delta\epsilon_q} \right]$$

where $\delta\epsilon_q = \epsilon_0 - q\epsilon_1$ and $g(\lambda) = 1 + 3q^2\lambda - \text{sign}(\lambda)(1 - q^2)\sqrt{\lambda^2 - 24}$, see appendix E. In this case the energy profile reads:

$$\begin{aligned} \epsilon_{\mathbf{v}_{\text{soft}}}^1[\gamma;f] &= (\gamma^3 + 3\gamma^2\beta q)\epsilon_1 + (\beta^3 + 3\gamma\beta^2 q)\epsilon_0 \\ &+ \gamma\beta f \sqrt{\frac{u_{\text{typ}}^1(1-q^2)}{2(1-u_{\text{typ}}^1)}} \left(\frac{6\sqrt{2}q(\epsilon_0 - q\epsilon_1)}{1-q^2} - \lambda_{\text{typ}}^1 \right) \\ &+ \frac{f^2(\gamma + 2\beta)u_{\text{typ}}^1}{2\sqrt{2}(1-u_{\text{typ}}^1)} \left(\frac{6\sqrt{2}q(\epsilon_0 - q\epsilon_1)}{1-q^2} - \lambda_{\text{typ}}^1 \right) \\ &+ \frac{f^2\beta u_{\text{typ}}^1}{2\sqrt{2}(1-u_{\text{typ}}^1)} \frac{6\sqrt{2}q(q\epsilon_0 - \epsilon_1)}{1-q^2} + \frac{f^2\gamma}{2\sqrt{2}}\lambda_{\text{typ}}^1 \\ &+ \frac{f^2\beta}{2\sqrt{2}(1-u_{\text{typ}}^1)} \int_{-2\sqrt{6}}^{2\sqrt{6}} d\lambda \frac{\sqrt{36-\lambda^2}}{18\pi} \Phi(\lambda_{\text{typ}}^1, \lambda) \lambda. \end{aligned} \quad (14)$$

The function Φ in the last term of this expression gives the typical value of the overlap between the eigenvector associated to λ_{min}^1 and any arbitrary eigenvector of $\mathcal{H}(\sigma^0)$ with eigenvalue λ . Its expression is rather involved, and we report it in appendix G (see equation (G13)).

In figure 2 (top) we show two density plots associated to (14): clearly in this case the geodesic path is no longer optimal, and thus the energy barrier is lowered by deforming the path in the direction of softest curvature at σ^1 . The optimal path is obtained numerically, by selecting the lowest energy point for each γ and by verifying that this leads to a continuous path. In figure 3 we compare the energy profile (14) evaluated along the geodesic and optimal paths, for σ^1 being a correlated minimum ($q_{\text{un}} < q \leq q_{\text{ms}}$) or a saddle ($q_{\text{ms}} < q < q_M$). In the latter case, the optimal path lies entirely below ϵ_{th} (dashed horizontal line). When σ^1 is an uncorrelated minimum, since $u_{\text{typ}}^1 = 0$ the behavior is analogous as in equation (12) with γ and β exchanged in the last term of the expression; the optimal path is again the geodesic one.

Similarly as above, plugging \mathbf{v}_{Hess} into (11) and computing the averages one finds (see appendix I):

$$\epsilon_{\mathbf{v}_{\text{Hess}}}^1[\gamma;f] = (\gamma^3 + 3\gamma^2\beta[\gamma;f]q)\epsilon_1 + (\beta^3[\gamma;f] + 3\beta^2[\gamma;f]\gamma q)\epsilon_0 - \gamma\beta[\gamma;f]f(\gamma) \sqrt{\frac{3(1-q^2)^2}{1+q^2}}. \quad (15)$$

In this case the barrier along the perturbed path is lower than the geodesic one even for $q < q_{\text{un}}$, even though it is still above the threshold for all choices of $\epsilon_a < \epsilon_{\text{th}}$. In figure 4 we plot a comparison between the energy barrier of the geodesic path and that of the optimal paths, for given ϵ_a and varying q . The barriers decrease with q ; when $\mathbf{v} = \mathbf{v}_{\text{Hess}}$ the deformed path is always associated to a lower energy barrier, while for $\mathbf{v} = \mathbf{v}_{\text{soft}}^1$ this is true only for $q > q_{\text{un}}$. For

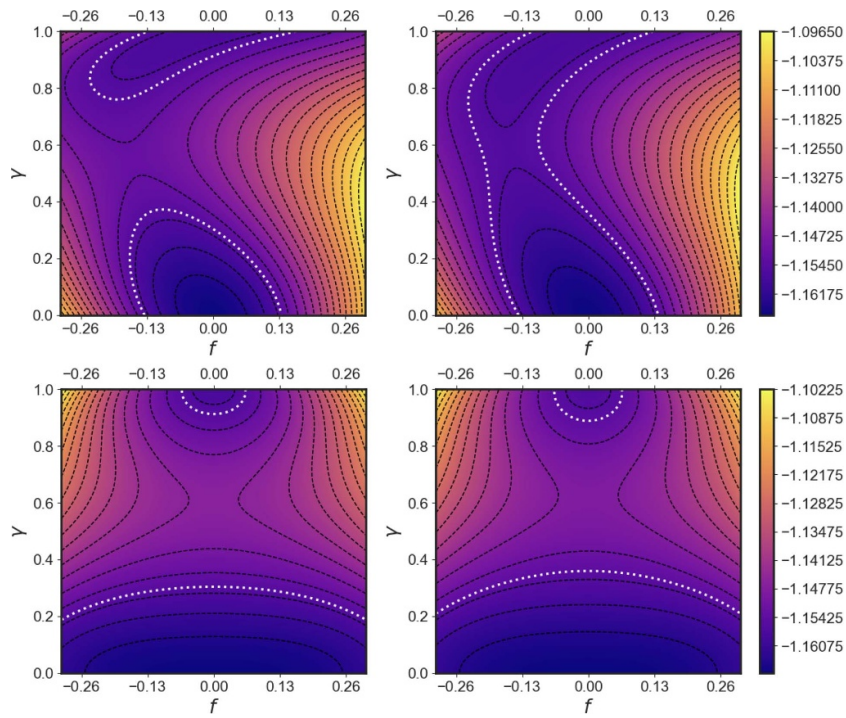


Figure 2. Color plot of $\epsilon_{\mathbf{v}_{\text{soft}}}[\gamma; f]$ as a function of γ and f for $\epsilon_0 = -1.167$, $\epsilon_1 = -1.157$ and $q = 0.66 \in [q_{\text{un}}, q_{\text{ms}}]$ (left) and $q = 0.7 > q_{\text{ms}}$ (right). The plots correspond to $\mathbf{v}_{\text{soft}} \rightarrow \mathbf{v}_{\text{soft}}^1$ (top), and $\mathbf{v}_{\text{soft}} \rightarrow \mathbf{v}_{\text{soft}}^0$ (bottom).

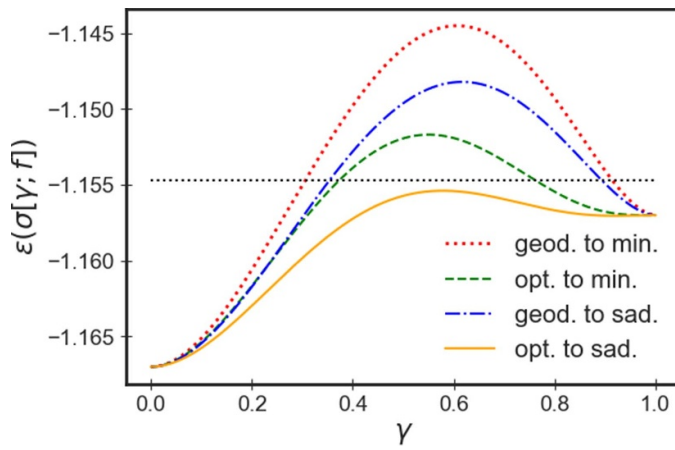


Figure 3. Comparison between the energy profile along the geodesic (geod.) and optimal (opt.) paths, for the same parameters as in figure 2(top) which correspond to σ^1 being a minimum (min.) and index-1 saddle (sad.).

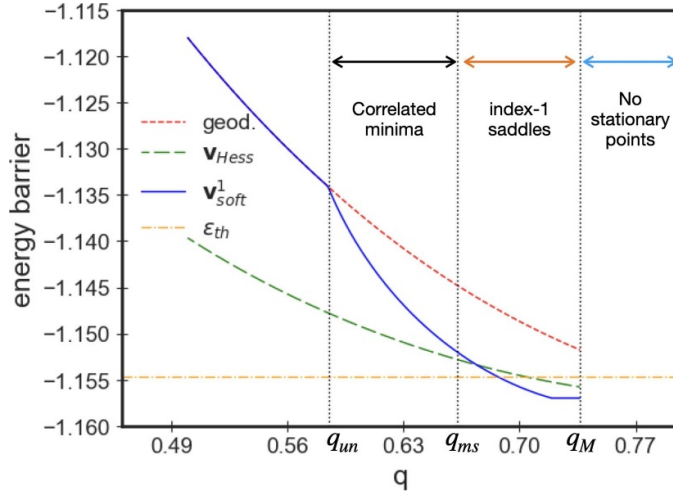


Figure 4. Energy barrier along the geodesic and optimal (perturbed) paths as a function of q , for $\epsilon_0 = -1.167$, $\epsilon_1 = -1.157$ and \mathbf{v} being equal to either $\mathbf{v}_{\text{soft}}^1$ or \mathbf{v}_{Hess} .

large values of q the barrier along the perturbed paths lies below ϵ_{th} , but this is true only within the range $q_{ms} < q < q_M$, when the arrival point is a rank-1 saddle. We find that this remains true for arbitrary values of $\epsilon_a < \epsilon_{\text{th}}$. For the largest $q \lesssim q_M$, the curve associated to $\mathbf{v}_{\text{soft}}^1$ is flat, indicating that the energy profile becomes monotonically increasing in the interval $\gamma \in [0, 1]$ with a maximum at $\gamma = 1$, equal to ϵ_1 .

In view of this comparison, one also understands why the geodesic path is no longer optimal when $\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^1$ and $u_{\text{typ}}^1 \neq 0$: in this case, indeed, $\mathbf{v}_{\text{soft}}^1$ has an $\mathcal{O}(1)$ projection on the vector \mathbf{v}_{Hess} , see appendix I, and thus allowing the path to deviate in the direction of $\mathbf{v}_{\text{soft}}^1$ leads to a lower energy barrier.

Discussion. The motivation for our work is characterizing paths connecting minima in rough energy landscapes. By focusing on a prototypical mean-field glass model, we have found that the energy profile along the geodesic path connecting two local minima always reaches values of energies that are above the threshold one; moreover, interpolating paths that are allowed to deviate from the geodesic along the direction associated to the softest curvature of the landscape at σ^0 are on average associated to higher energy barriers. Therefore, for the 3-spin model the direction of softest local curvature at the departing minimum σ^0 is not a predictor of lower energy barriers, and it is likely to be uncorrelated to good transition paths. This is different from what observed in numerical simulations of finite dimensional jammed and mildly super-cooled particles [17, 18, 34]. A possible explanation is that those cases were focusing on portions of the energy landscapes in which indeed the Hessian plays a crucial role [35–38]: in the former case the energy configurations were drawn close to the jamming transition, whereas in the latter close to the mode-coupling cross-over. We focus instead in a regime which would correspond to a deep super-cooled state. Our results therefore suggest that in this case the Hessian plays a different role—a result that would be interesting to test by studying energy paths and barriers in small systems [39, 40].

On the other hand, the Hessian matrix at σ^0 can still be used to construct directions in configuration space along which to deform the geodesic path, in order to lower the energy barrier: this special direction is obtained acting with the Hessian matrix on the vector that is tangent to the geodesic path at its initial point, and projecting the resulting vector in the subspace orthogonal to the tangent.

When the stationary points are closer and more strongly correlated, the direction of softest curvature at σ^1 corresponds to an isolated mode of the Hessian; this case of nearby, correlated minima is likely to be the one more closely related to the results found in finite dimensional systems [38]: in fact, we find that in this setting deformation of the geodesic path along the direction of the isolated mode *does* lead to lower energy barriers. The eigenvector associated to the isolated mode is localized in the direction connecting the two local minima (even though one can not talk about quasi-localized modes in this framework of fully-connected systems). The barriers of the perturbed paths lie in some cases below the threshold energy: we find that this happens only when σ^1 is a rank-1 saddle. For smaller q , paths lying below the threshold connecting the two minima might exist, but the curvature at σ^a is not enough to identify them.

Our results shed some light on the challenging problem of activated glassy dynamics [41], and in particular the identification of the barriers crossed by typical dynamical paths. The energy barriers found along the interpolating paths are in fact a possible proxy for such barriers, and they give some indications on how the dynamical transition rates between minima depend on the parameters q, ϵ_a . At very low temperature, one expects that the first energy barriers crossed by the system in its dynamics starting from a metastable state σ^0 correspond to the rank-1 saddles in the landscape that are closer to σ^0 [23–27]. These saddles connect the reference minimum to other local minima that are quite close to the original one³, are correlated to it (i.e. which belong to the region $q \in [q_{un}, q_{ms}]$) and lie at much higher energy [28]. The analysis of the interpolating paths complements these findings, as it provides some information on the energy barriers between distant minima at smaller energy difference, that are likely to be connected to the first one by a non-trivial sequence of saddles, i.e. of activated jumps. It is clear that further information can be obtained by identifying the distribution of rank-1 saddles that connect two *given* local minima σ^a by means of a direct counting (i.e. by performing a doubly constrained complexity calculation); we defer this analysis to future work. It could also be interesting to extend this study to mixed models obtained summing terms of the form (1) with different values of p (such as $p = 3$ and $p = 4$), in particular looking at paths interpolating between the different families of marginally stable minima that are exhibited by those models [42, 43]: this is also left to future work.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

³ For any of the rank-1 saddles that lie at $q > q_{ms}$, one can determine the properties of the local minima that are connected to the reference one σ^0 by this saddle [28]. These local minima typically are at overlap $q \in [q_{un}, q_{ms}]$ with σ^0 . We see that the interpolating path between these minima and the reference one have a local maximum that is at values of energies that do not correspond to that of the saddles, but are higher. Therefore, to reach the rank-1 saddle connecting the pair of minima one has to modify the interpolating path by including other directions in configuration space.

Acknowledgments

We thank Stephane d'Ascoli for useful discussions at the earlier stage of this work. VR acknowledges funding by the 'Investissements d'Avenir' LabEx PALM (ANR-10-LABX-0039-PALM). G B acknowledges funding from the Simons Foundation (#454935 Giulio Biroli)

Appendix A. Tangent planes, Riemannian gradients and Hessians

In this appendix, we introduce the main formalism and notation used in the rest of the paper. Consider the energy field $h(\boldsymbol{\sigma})$, defined on the N -dimensional hypersphere of unit radius, $\boldsymbol{\sigma} \in \mathcal{S}_N(1)$. We introduce a symmetric random tensor with components $a_{i_1 \dots i_p}$ (which are mean zero and unit variance random variables) so that

$$h(\boldsymbol{\sigma}) = \sqrt{p!} \sum_{i_1 < \dots < i_p} a_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p} = \frac{1}{\sqrt{p!}} \sum_{i_1, \dots, i_p} a_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p} = \sqrt{\frac{2}{N}} \mathcal{E}(\sqrt{N} \boldsymbol{\sigma}). \quad (\text{A1})$$

We define with $\nabla h(\boldsymbol{\sigma})$ the N -dimensional gradient of the field $h(\boldsymbol{\sigma})$ extended to the space \mathbb{R}^N . Such vector has components:

$$(\nabla h(\boldsymbol{\sigma}))_i = \mathbf{x}_i \cdot \nabla h(\boldsymbol{\sigma}) = \frac{\partial h(\boldsymbol{\sigma})}{\partial \sigma_i}, \quad (\text{A2})$$

where \mathbf{x}_i with $i = 1, \dots, N$ denotes some orthonormal basis of \mathbb{R}^N with components $[\mathbf{x}_i]_j = \delta_{ij}$. Similarly, we define with $\nabla^2 h(\boldsymbol{\sigma})$ the $N \times N$ -dimensional Hessian matrix with components:

$$(\nabla^2 h(\boldsymbol{\sigma}))_{ij} = \mathbf{x}_i \cdot \nabla^2 h(\boldsymbol{\sigma}) \cdot \mathbf{x}_j = \frac{\partial^2 h(\boldsymbol{\sigma})}{\partial \sigma_i \partial \sigma_j}. \quad (\text{A3})$$

Notice that, due to the homogeneity of the field $h(\boldsymbol{\sigma})$, one can show that:

$$\nabla h(\boldsymbol{\sigma}) \cdot \boldsymbol{\sigma} = p h(\boldsymbol{\sigma}), \quad (\text{A4})$$

and similarly

$$\nabla^2 h(\boldsymbol{\sigma}) \cdot \boldsymbol{\sigma} = (p-1) \nabla h(\boldsymbol{\sigma}). \quad (\text{A5})$$

The Riemannian gradient $\mathbf{g}(\boldsymbol{\sigma})$ and Hessian $\mathcal{H}(\boldsymbol{\sigma})$ of $h(\boldsymbol{\sigma})$ take into account the spherical constraint imposed on $\boldsymbol{\sigma}$. We denote with $\tau[\boldsymbol{\sigma}]$ the $(N-1)$ -dimensional tangent plane for each point $\boldsymbol{\sigma}$ of the hypersphere, which is the plane in \mathbb{R}^N spanned by vectors that are orthogonal to the vector $\boldsymbol{\sigma}$ (see figure 1). Let $\mathbf{e}_{i < N}(\boldsymbol{\sigma})$ be an orthonormal basis of such tangent plane,

$$\tau[\boldsymbol{\sigma}] = \text{Span} \{ \mathbf{e}_i(\boldsymbol{\sigma}) \}_{i=1}^{N-1}.$$

This local basis of $\tau[\boldsymbol{\sigma}]$ can be extended to a basis of \mathbb{R}^N by adding $\mathbf{e}_N(\boldsymbol{\sigma}) := \boldsymbol{\sigma}$. The $(N-1)$ -dimensional Riemannian gradient $\mathbf{g}(\boldsymbol{\sigma})$ is defined, component-wise, as the projection of the unconstrained gradient $\nabla h(\boldsymbol{\sigma})$ on the basis elements of the tangent plane $\tau[\boldsymbol{\sigma}]$:

$$\mathbf{g}_\alpha(\boldsymbol{\sigma}) = \nabla h(\boldsymbol{\sigma}) \cdot \mathbf{e}_\alpha(\boldsymbol{\sigma}) \quad \text{for } \alpha < N. \quad (\text{A6})$$

With a slight abuse of notation, in the following we sometimes use the same notation \mathbf{g} for the N -dimensional extended vector with an additional last component $\mathbf{g}_N(\boldsymbol{\sigma}) = 0$. In this way, the unconstrained and Riemannian gradients are related by:

$$\nabla h(\boldsymbol{\sigma}) = \mathbf{g}(\boldsymbol{\sigma}) + ph(\boldsymbol{\sigma})\boldsymbol{\sigma}. \tag{A7}$$

In order to obtain the Riemannian Hessian one can make use of a Lagrange multiplier to enforce the constraint $\boldsymbol{\sigma}^2 = 1$: we define $h_\lambda(\boldsymbol{\sigma}) := h(\boldsymbol{\sigma}) - \frac{\lambda}{2}(\boldsymbol{\sigma}^2 - 1)$. Then $\nabla h_\lambda(\boldsymbol{\sigma}) = \nabla h(\boldsymbol{\sigma}) - \lambda\boldsymbol{\sigma} \stackrel{!}{=} 0 \Rightarrow \lambda = \boldsymbol{\sigma} \cdot \nabla h(\boldsymbol{\sigma})$. From this it follows that

$$\begin{aligned} \mathcal{H}_{\alpha\beta}(\boldsymbol{\sigma}) &:= \mathbf{e}_\alpha(\boldsymbol{\sigma}) \cdot \nabla^2 h_\lambda(\boldsymbol{\sigma}) \cdot \mathbf{e}_\beta(\boldsymbol{\sigma}) \\ &= \mathbf{e}_\alpha(\boldsymbol{\sigma}) \cdot \nabla^2 h(\boldsymbol{\sigma}) \cdot \mathbf{e}_\beta(\boldsymbol{\sigma}) - (\nabla h(\boldsymbol{\sigma}) \cdot \boldsymbol{\sigma}) \delta_{\alpha\beta}, \quad \alpha, \beta \leq N-1. \end{aligned} \tag{A8}$$

The identity (A4) implies that the Riemannian Hessian is obtained from the unconstrained one by shifting by a diagonal matrix proportional to $ph(\boldsymbol{\sigma})$, and projecting on the local tangent plane. Therefore, working with the unconstrained or Riemannian Hessian is essentially the same, provided that one remembers about the shift.

The tangent plane $\tau[\boldsymbol{\sigma}]$, and therefore the basis vectors $\mathbf{e}_\alpha(\boldsymbol{\sigma})$, depend on the particular point in configuration space that one is looking at. For two different configurations, $\boldsymbol{\sigma}^0$ and $\boldsymbol{\sigma}^1$ at overlap $q = \boldsymbol{\sigma}^0 \cdot \boldsymbol{\sigma}^1$, it is convenient to choose the bases on the tangent planes $\tau[\boldsymbol{\sigma}^a]$, with $a = 0, 1$, as follows: first, one can always choose the basis \mathbf{x}_i in the N -dimensional space \mathbb{R}^N in such a way that the first $N - 2$ vectors $\mathbf{x}_1, \dots, \mathbf{x}_{N-2}$ are orthogonal to both $\boldsymbol{\sigma}^0$ and $\boldsymbol{\sigma}^1$. These vectors belong to both tangent planes (because they are orthogonal to both the configuration vectors), and therefore one can choose $\mathbf{e}_i(\boldsymbol{\sigma}^a) = \mathbf{x}_i$ for all $i = 1, \dots, N - 2$. Concretely, in the \mathbf{x}_i basis one can set $\boldsymbol{\sigma}^0 = (0, 0, \dots, 0, 1)$ and $\boldsymbol{\sigma}^1 = (0, 0, \dots, -\sqrt{1 - q^2}, q)$. Then, in the tangent plane $\tau[\boldsymbol{\sigma}^0]$ there remains a single basis vector to be chosen, which will have a non-zero projection with $\boldsymbol{\sigma}^1$; this vector has already been defined in equation (7),

$$\mathbf{e}_{N-1}(\boldsymbol{\sigma}^0) = \frac{q\boldsymbol{\sigma}^0 - \boldsymbol{\sigma}^1}{\sqrt{1 - q^2}}.$$

It has unit norm and it is orthogonal to all the others, since $\mathbf{x}_i \perp \boldsymbol{\sigma}^a$ for any i and $a = 0, 1$. The vector $\boldsymbol{\sigma}^0$ then completes the local basis. With the choice above, $\mathbf{e}_{N-1}(\boldsymbol{\sigma}^0) = (0, 0, \dots, 1, 0)$. Similarly, $\tau[\boldsymbol{\sigma}^1]$ is spanned by $\mathbf{x}_1, \dots, \mathbf{x}_{N-2}$ plus the vector $\mathbf{e}_{N-1}(\boldsymbol{\sigma}^1)$ defined similarly as

$$\mathbf{e}_{N-1}(\boldsymbol{\sigma}^1) = \frac{q\boldsymbol{\sigma}^1 - \boldsymbol{\sigma}^0}{\sqrt{1 - q^2}}.$$

Again, with the choice above $\mathbf{e}_{N-1}(\boldsymbol{\sigma}^1) = (0, 0, \dots, -q, -\sqrt{1 - q^2})$. In summary, we can choose $\mathbf{e}_\alpha(\boldsymbol{\sigma}^0) = \mathbf{e}_\alpha(\boldsymbol{\sigma}^1)$ with $\alpha = 1, \dots, N - 2$ as a basis of the subspace $\text{span}\{\mathbf{x}_\alpha\}$, while the remaining basis vector of the tangent planes $\mathbf{e}_{N-1}(\boldsymbol{\sigma}^a)$ changes depending on which tangent plane we consider (in appendix E.2, we specify the choice of some of the vectors $\mathbf{e}_\alpha(\boldsymbol{\sigma}^0) = \mathbf{e}_\alpha(\boldsymbol{\sigma}^1)$ as well). We denote the sets of these basis vectors as $\mathcal{B}^a = \{\mathbf{e}_\alpha(\boldsymbol{\sigma}^a)\}_{\alpha=1}^N$. All along this work, we assume that this choice of basis is made. Moreover, we will often use the notation $\mathbf{e}_\alpha(\boldsymbol{\sigma}^a) = \mathbf{e}_\alpha^a$, as well as $h(\boldsymbol{\sigma}^a) = h^a$, $\mathbf{g}_\alpha(\boldsymbol{\sigma}^a) = \mathbf{g}_\alpha^a$ and $\mathcal{H}(\boldsymbol{\sigma}^a) = \mathcal{H}^a$ for simplicity.

Appendix B. The complexity calculation: a recap

For completeness, we provide a summary of the main results of [26], upon which part of the present work is built. In [26] the authors compute the quenched complexity of stationary points σ^1 of $h(\sigma)$ at energy density ϵ_1 (meaning that $\mathbf{g}(\sigma^1) = 0$ and $h(\sigma^1) = \sqrt{2N}\epsilon_1$), that are conditioned to be at fixed overlap $\sigma^0 \cdot \sigma^1 = q$ with a given *reference* minimum σ^0 at energy density ϵ_0 . We refer to σ^1 as the *secondary* configuration. The complexity reads

$$\Sigma(\epsilon_1, q|\epsilon_0) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[\log \mathcal{N}_{\sigma^0}(\epsilon_1, q|\epsilon_0)],$$

see equation (3). The expectation value indicates both an average over the disorder, as well as a flat average over all local minima of energy ϵ_0 ; the random variable $\mathcal{N}_{\sigma^0}(\epsilon_1, q|\epsilon_0)$ is the number of such stationary points σ^1 , and it is defined below in equation (B4). To compute the expected value of the logarithm by means of the replica trick, one needs to replicate the secondary configuration σ^1 :

$$\Sigma(\epsilon_1, q|\epsilon_0) = \lim_{N \rightarrow \infty} \lim_{n \rightarrow 0} \frac{M_n(\epsilon_1, q|\epsilon_0) - 1}{Nn} \quad (\text{B1})$$

with

$$M_n(\epsilon_1, q|\epsilon_0) := \mathbb{E} \left[\frac{1}{\mathcal{N}(\epsilon_0)} \int_{\mathcal{S}_N(1)} d\sigma^0 \omega(\sigma^0) \int_{\mathcal{S}_N(1)} \prod_{a=1}^n d\sigma^a \omega_{\epsilon_1, q}(\sigma^a | \sigma^0) \right] \quad (\text{B2})$$

and

$$\begin{aligned} \omega_{\epsilon_0}(\sigma^0) &= |\det \mathcal{H}(\sigma^0)| \delta(h(\sigma^0) - \sqrt{2N}\epsilon_0) \delta(\mathbf{g}(\sigma^0)) \\ \omega_{\epsilon_1, q}(\sigma^a | \sigma^0) &= |\det \mathcal{H}(\sigma^a)| \delta(h(\sigma^a) - \sqrt{2N}\epsilon_1) \delta(\mathbf{g}(\sigma^a)) \delta(\sigma^0 \cdot \sigma^a - q) \end{aligned} \quad (\text{B3})$$

and where

$$\begin{aligned} \mathcal{N}(\epsilon_0) &= \int_{\mathcal{S}_N(1)} d\sigma^0 \omega_{\epsilon_0}(\sigma^0), \\ \mathcal{N}_{\sigma^0}(\epsilon_1, q|\epsilon_0) &= \int_{\mathcal{S}_N(1)} d\sigma^1 \omega_{\epsilon_1, q}(\sigma^1 | \sigma^0) \end{aligned} \quad (\text{B4})$$

represent, respectively, the number of stationary points at energy density ϵ_0 and the number of stationary points at energy density ϵ_1 and overlap q with σ^0 . In order to carry out such computation, in principle one has to replicate the reference configuration σ^0 as well by raising the denominator to the numerator, thus writing:

$$M_n(\epsilon_1, q|\epsilon_0) := \lim_{k \rightarrow 0} \mathbb{E} \left[\int_{\mathcal{S}_N(1)} \prod_{\beta=1}^k d\sigma^{0, \beta} \omega(\sigma^{0, \beta}) \int_{\mathcal{S}_N(1)} \prod_{a=1}^n d\sigma^a \omega_{\epsilon_1, q}(\sigma^a | \sigma^0), \right] \quad (\text{B5})$$

where $\sigma^{0, \beta=1} = \sigma^0$. Due to the isotropy of the correlation function of the random field $h(\sigma)$, it turns out that this expectation value depends on the configurations $\sigma^a, \sigma^{0, \beta}$ only through the overlaps $q_{\alpha\beta}^0 = \sigma^{0, \alpha} \cdot \sigma^{0, \beta}$, $q_{a\beta} = \sigma^a \cdot \sigma^{0, \beta}$ and $q_{ab}^1 = \sigma^a \cdot \sigma^b$, where by construction $q_{\alpha\alpha}^0 = q_{aa}^1 = 1$ and $q_{a1} = \sigma^a \cdot \sigma^{0, \beta=1} = q$. This implies that the integral over the configurations can

be replaced by an integral over these order parameters, which can be computed with the saddle point method. One sees that the saddle point equations for the parameters $q_{\alpha\beta}^0$ enforce $q_{\alpha\beta}^0 = 0$ for all $\alpha \neq \beta = 1, \dots, k$; this reflects the fact that the overlap between metastable states in the pure spherical p -spin model is vanishing [15], i.e. the states are typically orthogonal to each others. This also implies that $q_{a\beta} = 0$ for all $a = 1, \dots, n$ and $\beta \neq 1$. As a consequence, at the saddle point solution the secondary replicas σ^1 are coupled only to the original reference configuration σ^0 and not to its replicas. It is simple to check that this implies that to leading order in N , the expectation value (B2) is identical to its annealed version, which is obtained averaging separately the numerator and the denominator. This means that we can write

$$\begin{aligned}
 M_n(\epsilon_1, q|\epsilon_0) &:= \frac{1}{\mathbb{E}[\mathcal{N}(\epsilon_0)]} \mathbb{E} \left[\int_{\mathcal{S}_N(1)} d\sigma^0 \omega(\sigma^0) \int_{\mathcal{S}_N(1)} \prod_{a=1}^n d\sigma^a \omega_{\epsilon_1, q}(\sigma^a|\sigma^0) \right] \\
 &= \mathbb{E} \left[\int_{\mathcal{S}_N(1)} \prod_{a=1}^n d\sigma^a \omega_{\epsilon_1, q}(\sigma^a|\sigma^0) \mid \begin{array}{l} h(\sigma^0) = \sqrt{2N}\epsilon_0 \\ \mathbf{g}(\sigma^0) = \mathbf{0} \end{array} \right] \\
 &= \int_{\mathcal{S}_N(1)} \prod_{a=1}^n d\sigma^a \mathbb{E} \left[\prod_{a=1}^n \omega_{\epsilon_1, q}(\sigma^a|\sigma^0) \mid \begin{array}{l} h(\sigma^0) = \sqrt{2N}\epsilon_0 \\ \mathbf{g}(\sigma^0) = \mathbf{0} \end{array} \right] \\
 &= \int_{\mathcal{S}_N(1)} \prod_{a=1}^n d\sigma^a \mathbb{E} \left[\prod_{a=1}^n |\det \mathcal{H}(\sigma^a)| \mid \left\{ \begin{array}{l} h^a = \sqrt{2N}\epsilon_1, h^0 = \sqrt{2N}\epsilon_0 \\ \mathbf{g}^a = \mathbf{0}, \mathbf{g}^0 = \mathbf{0} \forall a = 1, \dots, n \end{array} \right\} \right] p_{\vec{\sigma}|\sigma^0}(\mathbf{0}, \epsilon_1)
 \end{aligned} \tag{B6}$$

where $\vec{\sigma} = (\sigma^1, \dots, \sigma^n)$ and

$$p_{\vec{\sigma}|\sigma^0}(\mathbf{0}, \epsilon_1) = \mathbb{E} \left[\prod_{a=1}^n \delta(h^a - \sqrt{2N}\epsilon_1) \delta(\mathbf{g}^a) \mid \begin{array}{l} h^0 = \sqrt{2N}\epsilon_0 \\ \mathbf{g}^0 = \mathbf{0} \end{array} \right]. \tag{B7}$$

The expectation value is now a function of the overlaps q_{ab}^1 , and its leading order term can be determined again with a saddle point calculation. Searching for a saddle point solution where all replicas have the same overlap, $q_{ab}^1 = q_1$ for all $a, b = 1, \dots, n$ and $a \neq b$, one finds the solution $q_1 = q^2$ [26], which is the smallest possible overlap between replicas that are all subject to the constraint of having overlap q with the reference configuration. It can be showed explicitly that this implies that the complexity $\Sigma(\epsilon_1, q|\epsilon_0)$ computed within the quenched formalism is the same as that obtained within the annealed framework, i.e. setting $n = 1$ in the formulas above. More precisely, one obtained the identity:

$$\Sigma(\epsilon_1, q|\epsilon_0) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[\log \mathcal{N}_{\sigma_0}(\epsilon_1, q|\epsilon_0)] = \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[\mathcal{N}_{\sigma_0}(\epsilon_1, q|\epsilon_0)].$$

The final formula giving the complexity reads [26]:

$$\Sigma(\epsilon_1, q|\epsilon_0) = \frac{1}{2} \log \left(\frac{p}{2} (\tilde{z} - \epsilon_1)^2 \right) + \frac{p(\epsilon_1^2 + \epsilon_1 \tilde{z})}{2(p-1)} + \frac{Q}{2}, \quad \tilde{z} = \sqrt{\epsilon_1^2 - \epsilon_{\text{th}}^2} \tag{B8}$$

where $\epsilon_{\text{th}} = -\sqrt{2(p-1)/p}$ and

$$Q = \log \left(\frac{1 - q^2}{1 - q^{2p-2}} \right) - 2(\epsilon_0^2 U_0(q) + \epsilon_0 \epsilon_1 U(q) + \epsilon_1^2 U_1(q)) \tag{B9}$$

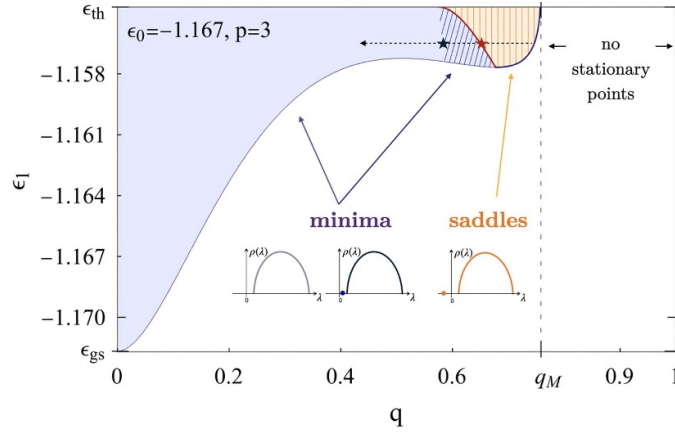


Figure 5. The plot shows in color the region in the plane (q, ϵ_1) where the complexity is positive, for $\epsilon_0 = -1.167$. The blue leftmost area corresponds to minima, the dashed blue zone to minima with an isolated eigenvalue in the Hessian spectrum, the rightmost (dashed) yellow zone to index-1 saddles. The stars mark the transitions in the properties of the Hessians of the stationary points, for fixed ϵ_1 . In the white area the complexity is vanishing.

with

$$\begin{aligned}
 U_0(q) &= \frac{q^{2p} [-q^{2p} + pq^2(1 - q^2) + q^4]}{q^{4p} - q^{2p} [(p-1)^2(1 + q^4) - 2p(p-2)q^2] + q^4}, \\
 U(q) &= \frac{2q^{3p} [1 - p(1 - q^2)] - 2q^{p+4}}{q^{4p} - ((p-1)^2(1 + q^4) - 2(p-2)pq^2)q^{2p} + q^4}, \\
 U_1(q) &= \frac{q^4 - q^{2p} - pq^{2p} [(p-1)q^4 + (3-2p)q^2 + p-2]}{q^{4p} - q^{2p} [(p-1)^2(1 + q^4) - 2p(p-2)q^2] + q^4}.
 \end{aligned} \tag{B10}$$

The results of this calculation are summarized in figure 5 for a representative value of $\epsilon_0 = -1.167$ [28]. For fixed $\epsilon_0 < \epsilon_{th}$, one finds that $\Sigma(\epsilon_1, q|\epsilon_0)$ is positive only for q smaller than a threshold value q_M , which depends on ϵ_0 . For fixed $q < q_M$, the complexity is positive for ϵ_1 belonging to a finite range, that is shown as a colored area (blue and yellow) in figure 5. By looking at the statistics of the Hessian of the stationary points at given q, ϵ_1 , one can discriminate whether they are stable local minima (such that all the eigenvalues of the Hessian are typically positive) or saddles with index k , i.e. with an Hessian with k negative eigenvalues. As we recall below, the statistic of the Hessian matrices is the same as that of GOE matrices perturbed with additive and multiplicative low-rank perturbations, and shifted by the quantity $p\epsilon_1$. Fixing the energy ϵ_1 and decreasing the overlap q (see the dashed arrow in figure 5), one finds that the typical stationary points are first saddles of index $k = 1$ (yellow area), then minima with an isolated but positive mode (blue dashed area), and finally minima with an Hessian spectrum described simply by a shifted semicircular law, without any isolated mode (blue area). The transitions between these different populations of stationary points are marked with stars in the figure. For $\epsilon > \epsilon_{th}$, all stationary points are saddles of extensive index $k \sim \mathcal{O}(N)$.

In the following, we select pairs of stationary points with parameters ϵ_a, q which belong to the colored region, i.e. which are such that the corresponding complexity $\Sigma(\epsilon_1, q | \epsilon_0) \geq 0$.

Appendix C. The energy profile along a path, and the case $p = 3$

Consider a path on the hypersphere, parametrized as in equation (4). Let $h[\gamma; f] \equiv h(\sigma[\gamma; f])$ denote the energy profile along the path. It holds:

$$h[\gamma; f] = \frac{1}{\sqrt{p!}} \sum_{i_1, \dots, i_p} a_{i_1 \dots i_p} \sum_{k_1=0}^p \sum_{k_2=0}^{k_1} \binom{p}{k_1} \binom{k_1}{k_2} f^{p-k_1} \beta^{k_1-k_2} \gamma^{k_2} \mathbf{v}_{i_1} \dots \mathbf{v}_{i_{p-k_1}} \sigma_{i_{p-k_1+1}}^0 \dots \sigma_{i_{p-k_2}}^0 \sigma_{i_{p-k_2+1}}^1 \dots \sigma_{i_p}^0. \quad (\text{C1})$$

In the special case $p = 3$, this function depends only on the gradients and the Hessians of the energy field at the initial and final configurations σ^a , $a = 0, 1$. Indeed, for $p = 3$ the above expansion reduces to:

$$h[\gamma; f] \stackrel{p=3}{=} \gamma^3 h(\sigma^1) + \beta^3 h(\sigma^0) + f^3 h(\mathbf{v}) + \gamma^2 \beta \nabla h(\sigma^1) \cdot \sigma^0 + \beta^2 \gamma \nabla h(\sigma^0) \cdot \sigma^1 + \gamma^2 f \nabla h(\sigma^1) \cdot \mathbf{v} + \beta^2 f \nabla h(\sigma^0) \cdot \mathbf{v} + \gamma \beta f \mathbf{v} \cdot \nabla^2 h(\sigma^0) \cdot \sigma^1 + \frac{f^2}{2} (\gamma \mathbf{v} \cdot \nabla^2 h(\sigma^1) \cdot \mathbf{v} + \beta \mathbf{v} \cdot \nabla^2 h(\sigma^0) \cdot \mathbf{v}). \quad (\text{C2})$$

We are interested in the case in which the beginning and end points of the path are stationary points satisfying $\mathbf{g}(\sigma^a) = 0$, and $\sigma^0 \cdot \sigma^1 = q$. Moreover, we assume that the vector \mathbf{v} is chosen in such a way that $\mathbf{v} \perp \sigma^0, \sigma^1$. The identity (A7) then implies that

$$h[\gamma; f] = (\gamma^3 + 3\gamma^2 \beta q) h(\sigma^1) + (\beta^3 + 3\beta^2 \gamma q) h(\sigma^0) + f^3 h(\mathbf{v}) + \gamma \beta f \mathbf{v} \cdot \nabla^2 h(\sigma^0) \cdot \sigma^1 + \frac{f^2}{2} (\gamma \mathbf{v} \cdot \nabla^2 h(\sigma^1) \cdot \mathbf{v} + \beta \mathbf{v} \cdot \nabla^2 h(\sigma^0) \cdot \mathbf{v}). \quad (\text{C3})$$

From this equation, equation (11) follows by using the block structure of the conditional Hessian matrices described below, as well as the fact that $h(\mathbf{v}) = 0$ typically. We remark that this expression is at fixed random couplings. In the rest of the paper, we restrict to the case $p = 3$ and average the energy profile over all configurations σ^0 and σ^1 satisfying

$$\begin{aligned} \mathbf{g}(\sigma^0) = \mathbf{0} \quad \& \quad \mathbf{g}(\sigma^1) = \mathbf{0} \\ h(\sigma^0) = \sqrt{2N}\epsilon_0 \quad \& \quad h(\sigma^1) = \sqrt{2N}\epsilon_1 \\ \sigma^0 \cdot \sigma^1 = q, \end{aligned} \quad (\text{C4})$$

as well as over the distribution of the random couplings $a_{i_1 \dots i_p}$.

Appendix D. Averages: quenched vs annealed energy profiles

It follows from equation (C3) that the energy profile along a path is completely determined by the local properties at the two configurations σ^a for $a = 0, 1$, if \mathbf{v} is also chosen to be a local property that depends only in the statistics of the landscape at σ^0 or σ^1 . In the following, we assume that \mathbf{v} is correlated with the softest mode of the Hessian matrix at either σ^0 or

σ^1 . We now consider the average of this function, over the family of stationary points σ^a of energy density ϵ_a and overlap q , and over all possible realizations of the random landscape. We consider two different averaging prescriptions, which we refer to as *quenched* and *annealed* in analogy with [29].

The quenched energy profile. The quenched averaging protocol corresponds to the following procedure: (i) for fixed realization of the random landscape, select σ^0 among the exponentially many stationary points of energy density ϵ_0 ; (ii) select σ^1 among the stationary points of energy density ϵ_1 that are at overlap q with the previously selected σ^0 ; (iii) evaluate the energy profile along the path, under the condition that \mathbf{v} has a projection on the minimal Hessian eigenvector at the start or at the end; (iv) average the profile over the population of stationary points σ^1 with parameters ϵ_1, q , (v) average over all σ^0 , (vi) average over the realizations of the landscape. We recall that in order to satisfy the spherical constraint, all integrals over configurations are taken over $\mathcal{S}_N(1)$. Formally, this prescription corresponds to:

$$h^{(Q)}[\gamma;f] := \mathbb{E} \left[\frac{1}{\mathcal{N}(\epsilon_0)} \int_{\mathcal{S}_N(1)} d\sigma^0 \omega_{\epsilon_0}(\sigma^0) \frac{1}{\mathcal{N}_{\sigma^0}(\epsilon_1, q|\epsilon_0)} \int_{\mathcal{S}_N(1)} d\sigma^1 \omega_{\epsilon_1, q}(\sigma^1|\sigma^0) h[\gamma;f] \right] \quad (D1)$$

where all the quantities above have already been defined in appendix B. In (D1), the random landscape appears both in the numerator and in the denominator; in order to perform the average, one has to resort to the replica trick to treat both denominators, through the formula $x^{-1} = \lim_{n \rightarrow 0} x^{n-1}$. Since there are two denominators, two sets of replicas should be exploited. However, due to the fact that the typical overlap between stationary points σ^0 in the pure spherical p -spin is zero [15], the flat average over σ^0 can be reproduced by conditioning the average over the landscape to realizations for which the configuration σ^0 is stationary, with a given energy density ϵ_0 (see appendix B). This implies:

$$h^{(Q)}[\gamma;f] = \mathbb{E} \left[\frac{1}{\mathcal{N}_{\sigma^0}(\epsilon_1, q|\epsilon_0)} \int_{\mathcal{S}_N(1)} d\sigma^1 \omega_{\epsilon_1, q}(\sigma^1|\sigma^0) h[\gamma;f] \mid \begin{matrix} h(\sigma^0) = \sqrt{2N}\epsilon_0 \\ \mathbf{g}(\sigma^0) = \mathbf{0} \end{matrix} \right]. \quad (D2)$$

By means of the replica trick, this average can be re-written as:

$$h^{(Q)}[\gamma;f] = \lim_{n \rightarrow 0} \int_{\mathcal{S}_N(1)} \prod_{k=1}^n d\sigma^k \mathbb{E} \left[\prod_{k=1}^n \omega_{\epsilon_1, q}(\sigma^k|\sigma^0) h[\gamma;f] \mid \begin{matrix} h(\sigma^0) = \sqrt{2N}\epsilon_0 \\ \mathbf{g}(\sigma^0) = \mathbf{0} \end{matrix} \right]. \quad (D3)$$

The delta functions inside $\omega_{\epsilon_1, q}(\sigma^1|\sigma^0)$ can be replaced by a conditioning in the average. To simplify the notation, for all the $n+1$ configurations σ^k with $k=0, 1, \dots, n$ we define the gradient vectors $\mathbf{g}^k \equiv \mathbf{g}(\sigma^k)$, the Hessian matrices $\mathcal{H}^k \equiv \mathcal{H}(\sigma^k)$, and the rescaled energy functional $h^k \equiv h(\sigma^k)$. This leads to:

$$h^{(Q)}[\gamma;f] = \lim_{n \rightarrow 0} \int_{\mathcal{S}_N(1)} \prod_{k=1}^n d\sigma^k \delta(\sigma^k \cdot \sigma^0 - q) \times \mathbb{E} \left[\prod_{k=1}^n |\det \mathcal{H}(\sigma^k)| h[\gamma;f] \mid \left\{ \begin{matrix} h^k = \sqrt{2N}\epsilon_1, h^0 = \sqrt{2N}\epsilon_0 \\ \mathbf{g}^k = \mathbf{0}, \mathbf{g}^0 = \mathbf{0} \forall k = 1, \dots, n \end{matrix} \right\} \right] p_{\bar{\sigma}|\sigma^0}(\mathbf{0}, \epsilon_1), \quad (D4)$$

where $\vec{\sigma} = (\sigma^1, \dots, \sigma^n)$ and $p_{\vec{\sigma}|\sigma^0}(\mathbf{0}, \epsilon_1)$, the joint density function of the gradients \vec{g}^k and fields \vec{h}^k , induced by the distribution of the couplings and conditioned to $\mathbf{g}^0 = \mathbf{0}$ and $h^0 = \sqrt{2N}\epsilon_0$, have already been defined in equation (B7)

We define:

$$\mathbb{E}_n^{\text{cond}}[\cdot] \equiv \mathbb{E}\left[\cdot \mid \left\{ \begin{array}{l} h^k = \sqrt{2N}\epsilon_1, h^0 = \sqrt{2N}\epsilon_0 \\ \mathbf{g}^k = \mathbf{0}, \mathbf{g}^0 = \mathbf{0} \forall k = 1, \dots, n \end{array} \right\}\right]. \quad (\text{D5})$$

Plugging (C3) into (D4) and implementing the conditioning, we find:

$$\begin{aligned} h^{(Q)}[\gamma; f] &= (\gamma^3 + 3\gamma^2\beta q) \sqrt{2N}\epsilon_1 + (\beta^3 + 3\beta^2\gamma q) \sqrt{2N}\epsilon_0 \\ &+ f^3(\gamma) \lim_{n \rightarrow 0} \int_{S_N(1)} \prod_{k=1}^n d\sigma^k \delta(\sigma^k \cdot \sigma^0 - q) \\ &\times \mathbb{E}_n^{\text{cond}} \left[\prod_{k=1}^n |\det \mathcal{H}(\sigma^k)| h(\mathbf{v}) \right] p_{\vec{\sigma}|\sigma^0}(\mathbf{0}, \epsilon_1) \\ &+ \gamma\beta[\gamma; f] f(\gamma) \lim_{n \rightarrow 0} \int_{S_N(1)} \prod_{k=1}^n d\sigma^k \delta(\sigma^k \cdot \sigma^0 - q) \\ &\times \mathbb{E}_n^{\text{cond}} \left[\prod_{k=1}^n |\det \mathcal{H}(\sigma^k)| \mathbf{v} \cdot \nabla^2 h(\sigma^0) \cdot \sigma^1 \right] p_{\vec{\sigma}|\sigma^0}(\mathbf{0}, \epsilon_1) \\ &+ \frac{f^2(\gamma)}{2} \lim_{n \rightarrow 0} \int_{S_N(1)} \prod_{k=1}^n d\sigma^k \delta(\sigma^k \cdot \sigma^0 - q) \\ &\times \mathbb{E}_n^{\text{cond}} \left[\prod_{k=1}^n |\det \mathcal{H}(\sigma^k)| (\gamma \mathbf{v} \cdot \nabla^2 h(\sigma^1) \cdot \mathbf{v} + \beta \mathbf{v} \cdot \nabla^2 h(\sigma^0) \cdot \mathbf{v}) \right] p_{\vec{\sigma}|\sigma^0}(\mathbf{0}, \epsilon_1), \end{aligned} \quad (\text{D6})$$

where the Riemannian and unconstrained Hessian matrices are related by (A8). The above equation contains the conditional average of the product of n determinants, multiplied by those terms in the energy profile h that depend on the matrices $\mathcal{H}(\sigma^1)$ and $\mathcal{H}(\sigma^0)$. All the matrices $\mathcal{H}(\sigma^\alpha)$ with $\alpha = 0, 1, \dots, n$ are coupled, and each one has to be conditioned to the gradients and energies of the $n + 1$ configurations. The correct way to proceed is to determine the expectation value for fixed n , and subsequently take $n \rightarrow 0$. In the language of the Franz-Parisi potential [29], this is a *quenched* calculation, since the two configurations σ^0 and σ^1 are not on the same footing: σ^0 can be considered as a reference configuration that is selected first, and σ^1 as a secondary one.

The annealed energy profile. In the annealed case, the two configurations σ^0 and σ^1 are treated on the same footing; the averaging protocol corresponds to the following procedure: (i) for fixed realization of the random landscape, select σ^0 and σ^1 among the exponentially many stationary points of energy density ϵ_0 and ϵ_1 , at mutual overlap q ; (ii) proceed as in the quenched case. Formally, this is equivalent to taking equation (D2) and factorizing the average of the numerator and denominator, getting:

$$\begin{aligned}
h^{(A)}[\gamma;f] &= \frac{1}{\mathbb{E}[\mathcal{N}_{\sigma^0}(\epsilon_1, q|\epsilon_0)]} \int_{\mathcal{S}_N(1)} d\boldsymbol{\sigma}^1 \delta(\boldsymbol{\sigma}^1 \cdot \boldsymbol{\sigma}^0 - q) \\
&\times \mathbb{E} \left[|\det \mathcal{H}(\boldsymbol{\sigma}^1)| h[\gamma, f] \mid \left\{ \begin{array}{l} h^1 = \sqrt{2N}\epsilon_1, h^0 = \sqrt{2N}\epsilon_0 \\ \mathbf{g}^\alpha = \mathbf{0} \forall \alpha = 0, 1 \end{array} \right\} \right] p_{\boldsymbol{\sigma}^1|\boldsymbol{\sigma}^0}(\mathbf{0}, \epsilon_1),
\end{aligned} \tag{D7}$$

where now

$$p_{\boldsymbol{\sigma}^1|\boldsymbol{\sigma}^0}(\mathbf{0}, \epsilon_1) = \mathbb{E} \left[\delta \left(h^1 - \sqrt{2N}\epsilon_1 \right) \delta(\mathbf{g}^1) \mid \begin{array}{l} h^0 = \sqrt{2N}\epsilon_0 \\ \mathbf{g}^0 = \mathbf{0} \end{array} \right]. \tag{D8}$$

One sees that the numerator is the same as (D4) with $n = 1$. Plugging (C3) into this expression and implementing the conditioning, we find in this case:

$$\begin{aligned}
h^{(A)}[\gamma;f] &= (\gamma^3 + 3\gamma^2\beta q) \sqrt{2N}\epsilon_1 + (\beta^3 + 3\beta^2\gamma q) \sqrt{2N}\epsilon_0 \\
&+ \frac{f^3(\gamma)}{\mathbb{E}[\mathcal{N}_{\sigma^0}(\epsilon_1, q|\epsilon_0)]} \int_{\mathcal{S}_N(1)} d\boldsymbol{\sigma}^1 \delta(\boldsymbol{\sigma}^1 \cdot \boldsymbol{\sigma}^0 - q) \\
&\times \mathbb{E}_1^{\text{cond}} [|\det \mathcal{H}(\boldsymbol{\sigma}^1)| h(\mathbf{v})] p_{\boldsymbol{\sigma}^1|\boldsymbol{\sigma}^0}(\mathbf{0}, \epsilon_1) \\
&+ \frac{\gamma\beta[\gamma;f]f(\gamma)}{\mathbb{E}[\mathcal{N}_{\sigma^0}(\epsilon_1, q|\epsilon_0)]} \int_{\mathcal{S}_N(1)} d\boldsymbol{\sigma}^1 \delta(\boldsymbol{\sigma}^1 \cdot \boldsymbol{\sigma}^0 - q) \\
&\times \mathbb{E}_1^{\text{cond}} [|\det \mathcal{H}(\boldsymbol{\sigma}^1)| \mathbf{v} \cdot \nabla^2 h(\boldsymbol{\sigma}^0) \cdot \boldsymbol{\sigma}^1] p_{\boldsymbol{\sigma}^1|\boldsymbol{\sigma}^0}(\mathbf{0}, \epsilon_1) \\
&+ \frac{f^2(\gamma)}{2\mathbb{E}[\mathcal{N}_{\sigma^0}(\epsilon_1, q|\epsilon_0)]} \int_{\mathcal{S}_N(1)} d\boldsymbol{\sigma}^1 \delta(\boldsymbol{\sigma}^1 \cdot \boldsymbol{\sigma}^0 - q) \\
&\times \mathbb{E}_1^{\text{cond}} [|\det \mathcal{H}(\boldsymbol{\sigma}^1)| (\gamma \mathbf{v} \cdot \nabla^2 h(\boldsymbol{\sigma}^1) \cdot \mathbf{v} + \beta \mathbf{v} \cdot \nabla^2 h(\boldsymbol{\sigma}^0) \cdot \mathbf{v})] p_{\boldsymbol{\sigma}^1|\boldsymbol{\sigma}^0}(\mathbf{0}, \epsilon_1).
\end{aligned} \tag{D9}$$

In this case, one has to deal with a pair of matrices, conditioned to only two gradients and energies. In the language of the Franz-Parisi potential, this is an *annealed* calculation.

Appendix E. Statistics of the Hessian matrices

In this appendix, we discuss the statistical distribution of the unconstrained Hessian matrices $\frac{1}{\sqrt{N-1}} \nabla^2 h(\boldsymbol{\sigma}^a)$ with $a = 0, 1$, subject to the conditioning on the energies and gradients of the various replicas. As we have seen in appendix A, the unconstrained Hessian and the Riemannian Hessian are easily related by a projection and shift.

E.1. Statistics of the Hessians: the annealed setup

E.1.1. Matrix distribution after conditioning. In this case, one has to determine the joint distributions of the two Hessian matrices $\nabla^2 h(\boldsymbol{\sigma}^a)$, each one conditioned to $\mathbf{g}(\boldsymbol{\sigma}^a) = \mathbf{0}$ and $h(\boldsymbol{\sigma}^a) = \sqrt{2N}\epsilon_a$ for $a = 0, 1$. This conditional, joint distribution has been determined in [24, 26]: each

matrix $\nabla^2 h(\sigma^a)$, expressed in its own tangent plane basis \mathcal{B}^a introduced in appendix A, has the following block structure:

$$\frac{1}{\sqrt{N-1}} \nabla^2 h(\sigma^a) = \begin{pmatrix} & & & m_{1N-1}^a & & 0 \\ & \mathbf{B}^a & & \vdots & & 0 \\ & & & m_{N-2N-1}^a & & 0 \\ m_{1N-1}^a & \cdots & m_{N-2N-1}^a & m_{N-1N-1}^a + \mu_a & & 0 \\ 0 & 0 & 0 & 0 & & \sqrt{\frac{2N}{N-1}} p(p-1) \epsilon_a \end{pmatrix} \in \mathbb{R}^{N \times N}. \tag{E1}$$

The entries in the $(N-2) \times (N-2)$ blocks \mathbf{B}^a are independent of the entries m_{iN-1}^a , and correlated with each others as follows:

$$\begin{aligned} \mathbb{E} [B_{ij}^a B_{kl}^b] &= \left(\delta_{ab} \frac{\sigma^2}{N-1} + (1 - \delta_{ab}) \frac{\sigma_H^2}{N-1} \right) (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \\ \mathbb{E} [m_{iM}^a m_{jM}^b] &= \left(\delta_{ab} \frac{\Delta^2}{N-1} + (1 - \delta_{ab}) \frac{\Delta_h^2}{N-1} \right) \delta_{ik} \quad i, j < N-1, \\ \mathbb{E} [m_{N-1N-1}^a m_{N-1N-1}^b] &= \frac{v_{ab}}{N-1}, \quad a, b \in \{0, 1\}. \end{aligned} \tag{E2}$$

Therefore, the blocks \mathbf{B}^a are $(N-2) \times (N-2)$ matrices with GOE (Gaussian Orthogonal Ensemble) statistics, with rescaled variance $\sigma^2(N-2)/(N-1)$. The two blocks are coupled component-wise. The parameters $\sigma, \sigma_H, \Delta, \Delta_h$ and μ_a are functions of p and of the parameters ϵ_a, q , and read [26]:

$$\begin{aligned} \sigma^2 &= p(p-1) \\ \sigma_H^2 &= p(p-1)q^{p-2} \\ \sigma_W^2 &= \sigma^2 - \sigma_H^2 = p(p-1)(1 - q^{p-2}) \\ \Delta^2 &= p(p-1) \left[1 - \frac{(p-1)(1-q^2)q^{2p-4}}{1 - q^{2p-2}} \right] \\ \Delta_h^2 &= p(p-1)q^{p-3}(-1) \left[1 - \frac{(p-1)(1-q^2)}{1 - q^{2p-2}} \right] \\ \mu_1 &= \mu(\epsilon_0, \epsilon_1), \quad \mu_0 = \mu(\epsilon_1, \epsilon_0) \end{aligned} \tag{E3}$$

where:

$$\begin{aligned} \mu(x, y) &\equiv \sqrt{2}p(p-1)(1-q^2) \\ &\times \frac{[q^4 - (p-1)q^{2p} + (p-2)q^{2p+2}]x - [q^{3p} + (p-2)q^{p+2} - (p-1)q^{p+4}]y}{q^{6-p} + q^{3p+2} - q^{p+2} [(p-1)^2(1+q^4) - 2p(p-2)q^2]}. \end{aligned} \tag{E4}$$

The conditional distribution of the Riemannian Hessians therefore reads as follows:

$$\frac{\mathcal{H}(\sigma^a)}{\sqrt{N-1}} = \begin{pmatrix} & & & m_{1N-1}^a \\ & & & \vdots \\ & \mathbf{B}^a & & \\ & & & m_{N-2N-1}^a \\ m_{1N-1}^a & \cdots & m_{N-2N-1}^a & m_{N-1N-1}^a \end{pmatrix} + \mu_a \mathbf{e}_{N-1}^a [\mathbf{e}_{N-1}^a]^T - \sqrt{\frac{2N}{N-1}} p \epsilon_a \in \mathbb{R}^{N-1 \times N-1}. \tag{E5}$$

For later convenience, we also define the matrix:

$$\frac{\tilde{\mathcal{H}}_A(\sigma^a)}{\sqrt{N-1}} = \begin{pmatrix} & & & m_{1N-1}^a \\ & & & \vdots \\ & \mathbf{B}^a & & \\ & & & m_{N-2N-1}^a \\ m_{1N-1}^a & \cdots & m_{N-2N-1}^a & m_{N-1N-1}^a + \mu_a \end{pmatrix} \in \mathbb{R}^{N-1 \times N-1}. \tag{E6}$$

where A stands for annealed. These matrices can be thought of as matrices with a GOE statistics, perturbed with both an additive and multiplicative finite rank perturbation along the direction corresponding to the basis vector $\mathbf{e}_{N-1}(\sigma^a)$, and shifted by a diagonal matrix. Notice that in the annealed formalism the random matrix problem is symmetric, in the sense that the two conditioned Hessians have the same structure at both the configurations σ^0 and σ^1 .

Let us now consider specifically the case $p = 3$. One sees from (E3) that for $p = 3$ it holds $\Delta^2 = \Delta_h^2 = 6(1 - q^2)(1 + q^2)^{-1}$, which implies that $m_{iN-1}^0 = m_{iN-1}^1 \equiv m_{iN-1}$. Moreover, in this case $\mu_1 = 6\sqrt{2}q(\epsilon_0 - q\epsilon_1)(1 - q^2)^{-1}$ and $\mu_0 = 6\sqrt{2}q(\epsilon_1 - q\epsilon_0)(1 - q^2)^{-1}$.

E.1.2. Spectral statistics and isolated eigenvalue. The matrices (E5) have a block structure, with a $(N - 2) \times (N - 2)$ block \mathbf{B}^a with GOE statistics, and a special line and column. Let us neglect the term proportional to the identity, which corresponds to a constant shift of the whole spectrum, and work with $\tilde{\mathcal{H}}_A(\sigma^a)$. The spectral properties of these kind of matrices are discussed in detail in [33]. To leading order in the size of the matrix, the eigenvalue density $\rho_N^a(\lambda)$ of both matrices is not affected by the presence of the special line and column, and it just coincides with the eigenvalue density of the GOE block, i.e. it is given by the semicircular law

$$\rho_N^a(\lambda) = \rho_\sigma(\lambda) + \mathcal{O}\left(\frac{1}{N}\right) = \frac{1}{2\pi\sigma^2} \sqrt{4\sigma^2 - \lambda^2} + \mathcal{O}\left(\frac{1}{N}\right), \quad \sigma^2 = p(p - 1). \tag{E7}$$

The presence of the special row and column can give rise to subleading contributions to the eigenvalue density: these contributions correspond to eigenvalues that do not belong to the support of the semicircular law (and are said to be ‘isolated’), and whose typical value depends on the parameters Δ, μ_a governing the statistics of the entries of the special row and column. As argued in [33], the fact that $\Delta \leq \sigma$ (as can be easily verified to be the case here), implies that only one isolated eigenvalue can exist for these matrices. Such eigenvalue exists whenever

$$|\mu_a| > \sigma \left(1 + \frac{\sigma^2 - \Delta^2}{\sigma^2}\right), \tag{E8}$$

and it is obtained as the real solution of the equation

$$\lambda - \mu_a - \Delta^2 \mathfrak{g}_\sigma(\lambda) = 0, \tag{E9}$$

where for z real such that $|z| > 2\sigma$ one has

$$\mathfrak{g}_\sigma(z) = \frac{1}{2\sigma^2} \left(z - \text{sign}(z) \sqrt{z^2 - 4\sigma^2} \right). \tag{E10}$$

The equation (E9) is obtained imposing that the resolvent of the matrix (E6), projected into the direction \mathbf{e}_{N-1}^a of the rank-1 perturbation, has a pole. Explicitly, the eigenvalue is given by:

$$\lambda_{\text{iso}}^a = \frac{2\mu_a\sigma^2 - \Delta^2\mu_a - \text{sign}(\mu_a)\Delta^2\sqrt{\mu_a^2 - 4(\sigma^2 - \Delta^2)}}{2(\sigma^2 - \Delta^2)}. \tag{E11}$$

Notice that for $\mu_a < 0$, this eigenvalue is negative and it coincides with the smallest one of the matrix, i.e. $\lambda_{\text{min}}^a = \lambda_{\text{iso}}^a$. Whenever the isolated eigenvalue exists, its eigenvector $\mathbf{e}_{\text{iso}}^a$ has a projection on the vector $\mathbf{e}_{N-1}(\sigma^a)$, corresponding to the special line and column of the matrix, which remains of $\mathcal{O}(1)$ when N is large. The typical value of this projection has been computed in [27, 33] and reads:

$$\left(\mathbf{e}_{\text{iso}}^a \cdot \mathbf{e}_{N-1}^a\right)^2 = \mathfrak{q}_{\sigma,\Delta}(\lambda_{\text{iso}}^a, \mu_a) \tag{E12}$$

where we introduced the function:

$$\mathfrak{q}_{\sigma,\Delta}(\lambda, \mu) := \text{sign}(\mu) \frac{\text{sign}(\lambda)\Delta^2\sqrt{\lambda^2 - 4\sigma^2} - \lambda(2\sigma^2 - \Delta^2) + 2\mu\sigma^2}{2\Delta^2\sqrt{\mu^2 - 4(\sigma^2 - \Delta^2)}}. \tag{E13}$$

It is shown in [33] that equation (E12) is indeed positive on the Right Hand Side, as it should be. Moreover, whenever the squared overlap is non-zero, the eigenvector is aligned with the direction of the finite-rank perturbation, meaning that $\mathbf{e}_{\text{iso}}^a \cdot \mathbf{e}_{N-1}^a > 0$. In the main text, we consider the case in which $\mu_a < 0$ and thus $\lambda_{\text{min}}^a = \lambda_{\text{iso}}^a$, and define $u^a = (\mathbf{e}_{\text{iso}}^a \cdot \mathbf{e}_{N-1}^a)^2$. The expression for these quantities in the case $a = 1$ and $p = 3$ are given in equation (13).

E.1.3. Change of basis. We recall that the matrices (E1) are expressed in the bases \mathcal{B}^a . It is sometimes convenient for the calculations to express both matrices in the same basis. This is achieved by performing a simple rotation, which involves only the two elements of the basis sets. We use the fact that

$$\begin{aligned} \boldsymbol{\sigma}^0 &= q\boldsymbol{\sigma}^1 - \sqrt{1-q^2}\mathbf{e}_{N-1}(\boldsymbol{\sigma}^1) & \mathbf{e}_{N-1}(\boldsymbol{\sigma}^0) &= -\sqrt{1-q^2}\boldsymbol{\sigma}^1 - q\mathbf{e}_{N-1}(\boldsymbol{\sigma}^1) \\ \boldsymbol{\sigma}^1 &= q\boldsymbol{\sigma}^0 - \sqrt{1-q^2}\mathbf{e}_{N-1}(\boldsymbol{\sigma}^0) & \mathbf{e}_{N-1}(\boldsymbol{\sigma}^1) &= -\sqrt{1-q^2}\boldsymbol{\sigma}^0 - q\mathbf{e}_{N-1}(\boldsymbol{\sigma}^0). \end{aligned} \tag{E14}$$

The matrix of change of basis $\mathcal{B}^0 \rightarrow \mathcal{B}^1$ reads:

$$\mathbf{R}_{0 \rightarrow 1} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 \\ & \ddots & & & 0 & 0 \\ & & \ddots & & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & \dots & \dots & 0 & -q & -\sqrt{1-q^2} \\ 0 & \dots & \dots & 0 & -\sqrt{1-q^2} & q \end{pmatrix} \in \mathbb{R}^{N \times N}, \tag{E15}$$

and it is equivalent to $\mathbf{R}_{1 \rightarrow 0}$. This implies that in the basis \mathcal{B}^1 we have:

$$\mathbf{R}_{0 \rightarrow 1}^{-1} \frac{\nabla^2 h(\boldsymbol{\sigma}^0)}{\sqrt{N-1}} \mathbf{R}_{0 \rightarrow 1} = \begin{pmatrix} & & & -qm_{1N-1}^0 & -\sqrt{1-q^2}m_{1N-1}^0 \\ & & & -qm_{iN-1}^0 & -\sqrt{1-q^2}m_{iN-1}^0 \\ & & \mathbf{B}^0 & & \\ & & & -qm_{N-2N-1}^0 & -\sqrt{1-q^2}m_{N-2N-1}^0 \\ -qm_{1N-1}^0 & \cdots & \cdots & -qm_{N-2N-1}^0 & q^2(m_{N-1N-1}^0 + \mu_0) + l_{N-1N-1}^0 & l_{N-1N}^0 \\ -\sqrt{1-q^2}m_{1N-1}^0 & \cdots & \cdots & -\sqrt{1-q^2}m_{N-2N-1}^0 & l_{N-1N}^0 & l_{NN}^0 \end{pmatrix}, \tag{E16}$$

where

$$\begin{aligned} l_{N-1N-1}^0 &= (1-q^2) \sqrt{\frac{2N}{N-1}} p(p-1) \epsilon_0, \\ l_{NN}^0 &= (1-q^2) (m_{N-1N-1}^0 + \mu_0) + q^2 \sqrt{\frac{2N}{N-1}} p(p-1) \epsilon_0, \\ l_{N-1N}^0 &= q\sqrt{1-q^2} \left(m_{N-1N-1}^0 + \mu_0 - \sqrt{\frac{2N}{N-1}} p(p-1) \epsilon_0 \right). \end{aligned} \tag{E17}$$

Defining the $(N-1) \times (N-1)$ matrix

$$\Pi_q^a = \mathbb{I} - (1-q) \mathbf{e}_{N-1}(\boldsymbol{\sigma}^a) \mathbf{e}_{N-1}^T(\boldsymbol{\sigma}^a) \tag{E18}$$

and using (E6), one can write the projection of this matrix onto the basis of $\tau[\boldsymbol{\sigma}^1]$ as

$$\mathbf{R}_{0 \rightarrow 1}^{-1} \frac{\nabla^2 h(\boldsymbol{\sigma}^0)}{\sqrt{N-1}} \mathbf{R}_{0 \rightarrow 1} \Big|_{\perp} = \Pi_q^1 \cdot \frac{\tilde{\mathcal{H}}_A(\boldsymbol{\sigma}^0)}{\sqrt{N-1}} \cdot \Pi_q^1 + l_{N-1N-1}^0 \mathbf{e}_{N-1}(\boldsymbol{\sigma}^1) \mathbf{e}_{N-1}^T(\boldsymbol{\sigma}^1) \tag{E19}$$

Notice that the matrix $\nabla^2 h(\boldsymbol{\sigma}^1)$ expressed in the basis \mathcal{B}^0 has exactly the same form, with the superscript 0 replaced by 1.

E.2. Statistics of the Hessians: the quenched setup

E.2.1. Matrix distribution after conditioning. In the quenched setting, replicas are introduced and one has to consider the joint distribution of $n+1$ matrices $\nabla^2 h(\boldsymbol{\sigma}^\alpha)$ with $\alpha = 0, 1, \dots, n$, conditioned to $\mathbf{g}(\boldsymbol{\sigma}^\alpha) = \mathbf{0}$ and $h(\boldsymbol{\sigma}^0) = \sqrt{2N}\epsilon_0$, $h(\boldsymbol{\sigma}^k) = \sqrt{2N}\epsilon_1$ with $k = 1, \dots, n$. The joint distribution of the matrices $\nabla^2 h(\boldsymbol{\sigma}^k)$ with $k = 1, \dots, n$ subject to the above conditioning is derived in [26]. We recall it in the following, discussing in addition the statistics of $\nabla^2 h(\boldsymbol{\sigma}^0)$ and its correlations with $\nabla^2 h(\boldsymbol{\sigma}^1)$, which are the two matrices appearing in the energy profile. In this section, $M := N - n$.

The conditional Hessians have a block structure when expressed in an appropriate basis. We make a particular choice of the vectors $\mathbf{e}_i(\boldsymbol{\sigma}^a)$ for $i = M, \dots, N-2$. As in appendix B, we define the overlap between the replicas as

$$q_l = \boldsymbol{\sigma}^k \cdot \boldsymbol{\sigma}^l \quad k, l = 1, \dots, n, \quad k \neq l. \tag{E20}$$

For $a = 1$, we choose the basis vectors in \mathcal{B}^1 as:

$$\begin{aligned}
 \mathbf{e}_M(\sigma^1) &= \frac{\sigma^2 - \sigma^3}{\sqrt{2(1 - q_1)}} \\
 \mathbf{e}_{M+1}(\sigma^1) &= \frac{\sigma^2 + \sigma^3 - 2\sigma^4}{\sqrt{3 \cdot 2(1 - q_1)}} \\
 \mathbf{e}_{M+2}(\sigma^1) &= \frac{\sigma^2 + \sigma^3 + \sigma^4 - 3\sigma^5}{\sqrt{4 \cdot 3(1 - q_1)}} \\
 &\dots \\
 \mathbf{e}_{N-3}(\sigma^1) &= \frac{\sigma^2 + \dots + \sigma^{n-1} - (n-2)\sigma^n}{\sqrt{(n-1)(n-2)(1 - q_1)}} \\
 \mathbf{e}_{N-2}(\sigma^1) &= \frac{\sigma^2 + \dots + \sigma^n - (n-1)q_1\sigma^1 + (n-1)q_1^{\frac{1-q_1}{1-q^2}}(q\sigma^1 - \sigma^0)}{\sqrt{(n-1)\left[1 + (n-2)q_1 - (n-1)q_1^2 - (n-1)q_1^2 \frac{1-q_1}{1-q^2}(1 - q_1)\right]}}
 \end{aligned} \tag{E21}$$

and, as specified in appendix A,

$$\mathbf{e}_{N-1}(\sigma^1) = \frac{q\sigma^1 - \sigma^0}{\sqrt{1 - q^2}}.$$

It is simple to check that all these vectors are all orthogonal to each others under the assumption (E20), and are orthogonal to σ^1 : therefore, they all belong to the tangent plane $\tau[\sigma^1]$. When using the notation \mathcal{B}^1 in the following, we assume this choice of basis vectors. Notice that the typical value of the overlap (E20) between stationary points (as it follows from the saddle-point calculation of the complexity recalled in appendix B) is $q_1 = q^2$, thus at the saddle point:

$$\mathbf{e}_{N-2}(\sigma^1) \Big|_{q_1=q^2} = \frac{\sigma^2 + \dots + \sigma^n - (n-1)q\sigma^0}{\sqrt{(n-1)[1 - q^2]}}. \tag{E22}$$

The choice of the basis vectors for σ^0 is analogous:

$$\mathbf{e}_i(\sigma^1) = \mathbf{e}_i(\sigma^0) \quad i = M, \dots, N-3 \tag{E23}$$

with

$$\mathbf{e}_{N-2}(\sigma^0) = \frac{\sigma^2 + \dots + \sigma^n - (n-1)q\sigma^0 + (n-1)\frac{q_1 - q^2}{1 - q^2}(q\sigma^0 - \sigma^1)}{\sqrt{(n-1)\left[1 - q^2 + (n-2)(q_1 - q^2) - (n-1)(1 - q^2)\left(\frac{q_1 - q^2}{1 - q^2}\right)^2 - 2\frac{(q_1 - q^2)^2}{1 - q^2}\right]}} \tag{E24}$$

and again

$$\mathbf{e}_{N-1}(\sigma^0) = \frac{q\sigma^0 - \sigma^1}{\sqrt{1 - q^2}}. \tag{E25}$$

Notice that when $q_1 = q^2$, the vector $\mathbf{e}_{N-2}(\boldsymbol{\sigma}^0)$ simplifies and becomes equivalent to $\mathbf{e}_{N-2}(\boldsymbol{\sigma}^1)$, and thus all the basis vectors are the same except for \mathbf{e}_{N-1} .

At variance with the annealed case, in the quenched setting the problem is not symmetric in the two configurations $\boldsymbol{\sigma}^a$. Consider first the Hessian at $\boldsymbol{\sigma}^1$. Its conditional distribution for finite, arbitrary n is given in [26]: in the basis \mathcal{B}^1 introduced just above, the unconstrained, conditioned Hessian takes the following form:

$$\frac{\nabla^2 h(\boldsymbol{\sigma}^1)}{\sqrt{N-1}} = \begin{pmatrix} & n_{1M}^1 & \cdots & n_{1N-2}^1 & m_{1N-1}^1 & 0 \\ & \mathbf{B}^1 & & & & \\ & \vdots & & \vdots & \vdots & \\ n_{1M}^1 & \cdots & n_{M-1M}^1 & n_{MM}^1 + \nu_1 & \cdots & n_{M-1N-2}^1 & m_{M-1N-1}^1 & 0 \\ \vdots & & & n_{M+1M+1}^1 + \nu_1 & & & m_{MN-1}^1 & 0 \\ \vdots & & & & & \ddots & & 0 \\ m_{1N-1}^1 & \cdots & m_{M-1N-1}^1 & m_{MN-1}^1 & \cdots & & m_{N-1N-1}^1 + \mu_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{2N}{N-1}} p(p-1) \epsilon_1 \end{pmatrix} \quad (\text{E26})$$

where \mathbf{B}^1 is a $(M-1) \times (M-1)$ block with the same statistics defined above in the annealed case. For $q_1 = q^2$ and $i, j < M$ and $M \leq k, l < N-1$ it holds:

$$\mathbb{E} [n_{ik}^1 n_{jl}^1] = \delta_{ij} \delta_{kl} p(p-1) \left[1 - \frac{(p-1)(1-q^2)q^{2p-4}}{1-q^{2p-2}} q^{2p-4} \right] \equiv \delta_{ij} \delta_{kl} \delta^2 \quad (\text{E27})$$

while for $i, j < M$

$$\mathbb{E} [m_{iN-1}^1 m_{jN-1}^1] = \delta_{ij} p(p-1) \left[1 - \frac{(p-1)(1-q^2)q^{2p-4}}{1-q^{2p-2}} \right] = \delta_{ij} \Delta^2, \quad (\text{E28})$$

where Δ^2 is the same as in (E3). The entries n_{ij}^1, m_{iN-1}^1 with $M \leq i, j < N-1$ form an $n \times n$ block with components that are correlated to each others, but independent from all the other entries of the matrix. We do not report the expression of the corresponding covariances, since it is not needed for the calculation below⁴. The diagonal terms in this $n \times n$ block contain deterministic contributions that arise from the conditioning, with μ_1 appearing already in (E4), and $\nu_1 \neq \mu_1$. The explicit expression of ν_1 is not reported here since it will not be needed in the remainder of the paper, but can be found in [26] (see equation (60) in appendix F of that

⁴ In the following, we are interested in determining the eigenvalue density of the Hessian matrices to order $1/N$, meaning that we aim at determining the bulk of the eigenvalues distribution, as well as the typical value of any isolated eigenvalue that may exist. It turns out that these quantities are insensitive to changes in the variance of $O(N^0)$ components of the matrix, and thus they do not depend on the covariances of the entries of the $n \times n$ block.

paper). As before, one finds that the conditional distribution of the Riemannian Hessians in the quenched case reads:

$$\frac{\mathcal{H}(\sigma^1)}{\sqrt{N-1}} = \frac{\tilde{\mathcal{H}}_Q(\sigma^1)}{\sqrt{N-1}} - \sqrt{\frac{2N}{N-1}} p \epsilon_1 \in \mathbb{R}^{N-1 \times N-1}, \quad (\text{E29})$$

where now:

$$\frac{\tilde{\mathcal{H}}_Q(\sigma^1)}{\sqrt{N-1}} = \begin{pmatrix} & & & n_{1M}^1 & \cdots & n_{1N-2}^1 & m_{1N-1}^1 \\ & \mathbf{B}^1 & & \vdots & & & \\ n_{1M}^1 & \cdots & n_{M-1M}^1 & n_{MM}^1 + \nu_1 & \cdots & n_{M-1N-2}^1 & m_{M-1N-1}^1 \\ \vdots & & & & n_{M+1M+1}^1 + \nu_1 & & m_{MN-1}^1 \\ \vdots & & & & & \ddots & \\ m_{1N-1}^1 & \cdots & m_{M-1N-1}^1 & m_{MN-1}^1 & \cdots & & m_{N-1N-1}^1 + \mu_1 \end{pmatrix}. \quad (\text{E30})$$

Here Q stands for quenched. In the quenched case, one has therefore n special lines, $n - 1$ out of which have exactly the same statistics controlled by the parameters δ^2 and ν_1 , while the last one has yet different averages μ_1 and variances Δ^2 . When $n \rightarrow 1$, only the last special line survives and one recovers the annealed Hessian distribution discussed above.

We now consider the Hessian at σ^0 . Its conditional statistics is not given explicitly in [26], but can be easily derived following the same reasoning of appendix C of that paper. In the basis \mathcal{B}^0 defined above, the unconstrained, conditioned Hessian takes the following form:

$$\frac{\nabla^2 h(\sigma^0)}{\sqrt{N-1}} = \begin{pmatrix} & & & n_{1M}^0 & \cdots & n_{1N-2}^0 & m_{1N-1}^0 & 0 \\ & \mathbf{B}^0 & & \vdots & & & & \\ n_{1M}^0 & \cdots & n_{M-1M}^0 & n_{MM}^0 + \mu_0 & \cdots & n_{M-1N-2}^0 & m_{M-1N-1}^0 & 0 \\ \vdots & & & & n_{M+1M+1}^0 + \mu_0 & & m_{MN-1}^0 & 0 \\ \vdots & & & & & \ddots & & 0 \\ m_{1N-1}^0 & \cdots & m_{M-1N-1}^0 & m_{MN-1}^0 & \cdots & & m_{N-1N-1}^0 + \mu_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{2N}{N-1}} p(p-1) \epsilon_0 \end{pmatrix}, \quad (\text{E31})$$

where \mathbf{B}^0 is again a $(M-1) \times (M-1)$ block with the same statistics defined above in the annealed case, but in this case for $q_1 = q^2$ and $i, j < M$ and $M \leq k, l < N-1$ it holds:

$$\mathbb{E} [n_{ik}^0 n_{jl}^0] = \delta_{ij} \delta_{kl} p(p-1) \left[1 - \frac{(p-1)(1-q^2)q^{2p-4}}{1-q^{2p-2}} \right] = \delta_{kl} \mathbb{E} [m_{iN-1}^0 m_{jN-1}^0] = \delta_{ij} \delta_{kl} \Delta^2. \tag{E32}$$

Once more, the entries n_{ij}^0, m_{iN-1}^0 with $M \leq i, j < N-1$ form an $n \times n$ block with components correlated to each others, but independent from all the other entries of the matrix. Finally, μ_0 is obtained from μ_1 in (E4) by switching $\epsilon_0 \leftrightarrow \epsilon_1$. As it appears from these formulas, in the case of σ^0 the conditioning modifies the variances of the entries that belong to n special lines and columns of the matrix, in a way that is identical for all rows and columns. Notice that the covariance of the matrix elements n_{ij}^0 is different with respect to that of the elements n_{ij}^1 , and it coincides instead with that of the elements m_{iN-1}^0 . This is consistent with the fact that, for $n \rightarrow 1$, one should recover the annealed statistics described in the section above. As before, one finds that the conditional distribution of the Riemannian Hessians in the quenched case reads:

$$\frac{\mathcal{H}(\sigma^0)}{\sqrt{N-1}} = \frac{\tilde{\mathcal{H}}_Q(\sigma^0)}{\sqrt{N-1}} - \sqrt{\frac{2N}{N-1}} p \epsilon_1 \in \mathbb{R}^{N-1 \times N-1}, \tag{E33}$$

where:

$$\frac{\tilde{\mathcal{H}}_Q(\sigma^0)}{\sqrt{N-1}} = \begin{pmatrix} & & & n_{1M}^0 & \cdots & m_{1N-1}^0 \\ & & & \vdots & & \\ & \mathbf{B}^0 & & \vdots & & \\ n_{1M}^0 & \cdots & n_{M-1M}^0 & n_{MM}^0 + \mu_0 & \cdots & m_{M-1N-1}^0 \\ \vdots & & & & & m_{MN-1}^0 \\ m_{1N-1}^0 & \cdots & m_{M-1N-1}^0 & m_{MN-1}^0 & \cdots & m_{N-1N-1}^0 + \mu_0 \end{pmatrix}. \tag{E34}$$

Finally, we discuss the correlations between the entries of these two Hessian. This can be obtained with an extension of the calculation described in appendix C of [26]. As recalled in appendix B, at the saddle point one finds that the overlap between replicas is $q_1 = q^2$. Then, for $i, j < M$ and $M \leq k, l < N-1$ it holds:

$$\mathbb{E} [n_{ik}^0 n_{jl}^1] = \delta_{ij} \delta_{kl} p(p-1) q^{p-2} \left[1 - (p-1) \frac{q^{2p-4}}{1-q^{2p-2}} (1-q^2) \right] \equiv \delta_{ij} \delta_{kl} \delta_h^2, \tag{E35}$$

while

$$\mathbb{E} [m_{iN-1}^0 m_{jN-1}^1] = \delta_{ij} p(p-1) q^{p-3} (-1) \left[1 - (p-1) \frac{1}{1-q^{2p-2}} (1-q^2) \right] \equiv \delta_{ij} \Delta_h^2, \tag{E36}$$

where Δ_h^2 is the same appearing in (E3). Notice that δ_h^2 differs from δ^2 by a factor of q^{p-2} ; instead Δ_h^2 differs from Δ^2 by a factor of q^{p-3} , meaning that the two quantities coincide for $p=3$.

E.2.2. Spectral statistics and isolated eigenvalue. Let us discuss the spectral properties of the conditional Hessian in the case in which n is an integer parameter. As in the annealed case, we neglect the constant shifts. In the quenched setting, the distribution of the two matrices is not equivalent: we begin by discussing the shifted Hessian at σ^1 , defined in equation (E30). This matrix has $(n - 1)$ lines with identical statistics given by δ^2, ν_1 , and one last line with statistics controlled by Δ^2, μ_1 . For integer n the isolated eigenvalues, whenever they exist, are solutions of the equation:

$$[\lambda - \nu_1 - \delta^2 g_\sigma(\lambda)]^{n-1} [\lambda - \mu_1 - \Delta^2 g_\sigma(\lambda)] = 0, \tag{E37}$$

which is again obtained imposing that the resolvent of (E30) projected onto the n -dimensional subspace corresponding to the finite-rank perturbations has some poles. One sees that two different eigenvalues may exist, one of which with degeneracy $(n - 1)$. In the replica calculation, however, one is required to take the limit $n \rightarrow 0$ at the end of the calculation. It appears that in this limit, the first factor of (E37) goes into the denominator, and therefore the corresponding solution becomes a zero, and not a pole, of the resolvent. In the limit $n \rightarrow 0$, the equation (E37) can exhibit only one meaningful solution, whenever $\lambda - \mu_1 - \Delta^2 g_\sigma(\lambda) = 0$. The corresponding solution is given by (E11), as in the annealed case, and it exists whenever (E8) is satisfied. Notice that combining this argument with the discussion in the annealed case, one can also conclude that (E30) subtracted of the last element μ_1 has no isolated eigenvalue in the limit $n \rightarrow 0$.

We now consider the shifted Hessian at σ^0 , defined in equation (E34). In this case, there are n special lines with identical statistics. For integer n the isolated eigenvalue, whenever it exists, is the solution of the following equation:

$$[\lambda - \mu_0 - \Delta^2 g_\sigma(\lambda)]^n = 0, \tag{E38}$$

and has degeneracy n . In the limit $n \rightarrow 0$, however, this equation has no solution, and thus one recovers the spectrum of an unperturbed GOE matrix.

E.2.3. Change of basis. We remark that when $q_1 = q^2$, the matrix of change of basis $\mathcal{B}^0 \rightarrow \mathcal{B}^1$ is exactly the same as in the annealed case, (E15). Therefore one can write the projection of this matrix onto the basis of $\tau[\sigma^1]$ as

$$\mathbf{R}_{0 \rightarrow 1}^{-1} \frac{\nabla^2 h(\sigma^0)}{\sqrt{N-1}} \mathbf{R}_{0 \rightarrow 1} \Big|_{\perp} = \Pi_q^1 \cdot \frac{\tilde{\mathcal{H}}_Q(\sigma^0)}{\sqrt{N-1}} \cdot \Pi_q^1 + l_{N-1 N-1}^0 \mathbf{e}_{N-1}^1 [\mathbf{e}_{N-1}^1]^T \tag{E39}$$

where $l_{N-1 N-1}^0$ is the same as in (E17). The other change of basis is obtained in an analogous way.

Appendix F. The $p = 3$ energy profile: deformation along softest Hessian modes

In this appendix, we determine the expression of the profile (C3) under the assumption that the unconstrained Hessians have the structure and the statistics discussed above, which follows from conditioning on the properties of the various replicas σ^a . As before, we assume $p = 3$. We consider both the case in which \mathbf{v} is chosen to be aligned along the direction of the softest mode of the Hessian $\mathcal{H}(\sigma^0)$, and the one in which \mathbf{v} is chosen to be aligned along the direction of the softest mode of the Hessian $\mathcal{H}(\sigma^1)$.

F.1. Softest mode at σ^0

We consider the case in which \mathbf{v} is chosen to be aligned along the direction of the softest mode \mathbf{e}_{\min}^0 of the Hessian $\mathcal{H}(\sigma^0)$; we introduce the following notation for the minimal eigenvalue of the shifted Hessian, and for the component squared of the corresponding eigenvector along the direction \mathbf{e}_{N-1}^0 :

$$\frac{\tilde{\mathcal{H}}(\sigma^0)}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 = \lambda_{\min}^0 \mathbf{e}_{\min}^0, \quad u^0 = (\mathbf{e}_{\min}^0 \cdot \mathbf{e}_{N-1}^0)^2. \quad (\text{F1})$$

We recall from equation (9) the choice:

$$\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^0 = \frac{\mathbf{e}_{\min}^0 - \sqrt{u^0} \mathbf{e}_{N-1}^0}{\sqrt{1-u^0}}.$$

Plugging this expression, we find

$$\begin{aligned} \mathbf{v}_{\text{soft}}^0 \cdot \frac{\nabla^2 h(\sigma^0)}{\sqrt{N-1}} \cdot \sigma^1 &= \left(\frac{\mathbf{e}_{\min}^0 - \sqrt{u^0} \mathbf{e}_{N-1}^0}{\sqrt{1-u^0}} \right) \cdot \frac{\nabla^2 h(\sigma^0)}{\sqrt{N-1}} \cdot (q\sigma^0 - \sqrt{1-q^2} \mathbf{e}_{N-1}^0) \\ &= \sqrt{\frac{1-q^2}{1-u^0}} (\sqrt{u^0} \mathbf{e}_{N-1}^0 - \mathbf{e}_{\min}^0) \cdot \frac{\nabla^2 h(\sigma^0)}{\sqrt{N-1}} \cdot \mathbf{e}_{N-1}^0 \\ &= \sqrt{\frac{u^0(1-q^2)}{1-u^0}} (\mu_0 - \lambda_{\min}^0), \end{aligned} \quad (\text{F2})$$

and

$$\mathbf{v}_{\text{soft}}^0 \cdot \frac{\nabla^2 h(\sigma^0)}{\sqrt{N-1}} \cdot \mathbf{v}_{\text{soft}}^0 = \frac{1}{1-u^0} (\lambda_{\min}^0 + u^0 \mu_0 - 2u^0 \lambda_{\min}^0). \quad (\text{F3})$$

Using a variant of (E19) and (E39) applied to $\nabla^2 h(\sigma^1)$, and the fact that $\mathbf{v}_{\text{soft}}^0 \perp \mathbf{e}_{N-1}^0$, one also obtains

$$\begin{aligned} \mathbf{v}_{\text{soft}}^0 \cdot \frac{\nabla^2 h(\sigma^1)}{\sqrt{N-1}} \cdot \mathbf{v}_{\text{soft}}^0 &= \left(\frac{\mathbf{e}_{\min}^0 - \sqrt{u^0} \mathbf{e}_{N-1}^0}{\sqrt{1-u^0}} \right) \cdot \frac{\tilde{\mathcal{H}}_x^1}{\sqrt{N-1}} \cdot \left(\frac{\mathbf{e}_{\min}^0 - \sqrt{u^0} \mathbf{e}_{N-1}^0}{\sqrt{1-u^0}} \right) \\ &= \frac{1}{1-u^0} \left(\mathbf{e}_{\min}^0 \cdot \frac{\tilde{\mathcal{H}}_x^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 - 2\sqrt{u^0} \mathbf{e}_{N-1}^0 \cdot \frac{\tilde{\mathcal{H}}_x^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 + u^0 \mu_1 \right) \\ &= \frac{1}{1-u^0} \left(\mathbf{e}_{\min}^0 \cdot \frac{\tilde{\mathcal{H}}_x^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 - 2\sqrt{u^0} \sqrt{1-u^0} \mathbf{v}_{\text{soft}}^0 \cdot \mathbf{m}^1 - u^0 \mu_1 \right), \end{aligned} \quad (\text{F4})$$

where $x \in \{A, Q\}$ and $\mathbf{m}^1 = (m_{1N-1}^1, \dots, m_{M-1N-1}^1)$. Now, the definition of $\mathbf{v}_{\text{soft}}^0$ is such that $\mathbf{e}_{\min}^0 = \sqrt{1-u^0} \mathbf{v}_{\text{soft}}^0 + \sqrt{u^0} \mathbf{e}_{N-1}^0$, which implies

$$\lambda_{\min}^0 \sqrt{u^0} = \mathbf{e}_{N-1}^0 \cdot \frac{\tilde{\mathcal{H}}_x^0}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 = \sqrt{1-u^0} \mathbf{v}_{\text{soft}}^0 \cdot \mathbf{m}^0 + \sqrt{u^0} \mu_0, \quad (\text{F5})$$

where $\mathbf{m}^0 = (m_{1N-1}^0, \dots, m_{M-1N-1}^0)$. For $p = 3$, making use of the fact that $\mathbf{m}^0 = \mathbf{m}^1 \equiv \mathbf{m}$, one finds therefore $\sqrt{1 - u^0} \mathbf{v}_{\text{soft}}^0 \cdot \mathbf{m} = \sqrt{u^0} (\lambda_{\min}^0 - \mu_0)$, which plugged into (F4) gives:

$$\mathbf{v}_{\text{soft}}^0 \cdot \frac{\nabla^2 h(\boldsymbol{\sigma}^1)}{\sqrt{N-1}} \cdot \mathbf{v}_{\text{soft}}^0 = \frac{1}{1 - u^0} \left(\mathbf{e}_{\min}^0 \cdot \frac{\tilde{\mathcal{H}}_x^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 - 2u^0 (\lambda_{\min}^0 - \mu_0) - u^0 \mu_1 \right). \quad (\text{F6})$$

Therefore, when $\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^0$, the energy profile (C3), under the conditions that $\boldsymbol{\sigma}^a$ are stationary points with energy density ϵ_a , becomes equal to:

$$\begin{aligned} h^{(x)}[\gamma; f] \longrightarrow & (\gamma^3 + 3\gamma^2 \beta q) \sqrt{2N} \epsilon_1 + (\beta^3 + 3\beta^2 \gamma q) \sqrt{2N} \epsilon_0 + f^3 h(\mathbf{v}_{\text{soft}}^0) \\ & + \gamma \beta f \sqrt{N-1} \sqrt{\frac{u^0(1-q^2)}{1-u^0}} (\mu_0 - \lambda_{\min}^0) \\ & + \frac{f^2}{2} \sqrt{N-1} \left[\frac{\gamma}{1-u^0} \left(\mathbf{e}_{\min}^0 \cdot \frac{\tilde{\mathcal{H}}_x^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 - 2u^0 (\lambda_{\min}^0 - \mu_0) - u^0 \mu_1 \right) \right. \\ & \left. + \frac{\beta}{1-u^0} (\lambda_{\min}^0 + u^0 \mu_0 - 2u^0 \lambda_{\min}^0) \right]. \end{aligned} \quad (\text{F7})$$

This expression depends on four random variables: the minimal eigenvalue λ_{\min}^0 and its eigenvector component u^0 , the rescaled energy $h(\mathbf{v}_{\text{soft}}^0)$, and the quantity:

$$\chi_x^0 := \mathbf{e}_{\min}^0 \cdot \frac{\tilde{\mathcal{H}}_x^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 \quad (\text{F8})$$

where the subscript $x \in \{Q, A\}$ keeps track of whether we are performing the conditioning of the Hessian statistics under the annealed or the quenched assumptions. The average energy profile is obtained plugging (F7) into (D6) and (D9) and performing the corresponding averages. In the limit of large N , however, the quantities λ_{\min}^0, u^0 and χ^0 concentrate, meaning that their distribution converges around their typical value [27, 32]. In this limit, one can therefore replace the quantities in (F7) with their typical value (under the conditional distribution), and check that the remaining terms in the average factorize with the denominators in the annealed case, or converge to unity in the limit $n \rightarrow 0$ in the quenched case. Let us denote with $\lambda_{\text{typ}}^0, u_{\text{typ}}^0$ and $\chi_{x, \text{typ}}^0$ the typical values of the quantities, and use that the typical value of $h(\mathbf{v}_{\text{soft}}^0)$ is equal to zero⁵. Then we find that the energy profile when $\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^0$ reduces to:

$$\begin{aligned} \epsilon_{\mathbf{v}_{\text{soft}}^0}^{(x)}[\gamma; f] := \lim_{N \rightarrow \infty} \frac{h^{(x)}[\gamma; f]}{\sqrt{2N}} = & (\gamma^3 + 3\gamma^2 \beta q) \epsilon_1 + (\beta^3 + 3\beta^2 \gamma q) \epsilon_0 \\ & + \gamma \beta f \sqrt{\frac{u_{\text{typ}}^0(1-q^2)}{2(1-u_{\text{typ}}^0)}} (\mu_0 - \lambda_{\text{typ}}^0) + \frac{f^2}{2\sqrt{2}} \left[\frac{\beta}{1-u_{\text{typ}}^0} (\lambda_{\text{typ}}^0 + u_{\text{typ}}^0 \mu_0 - 2u_{\text{typ}}^0 \lambda_{\text{typ}}^0) \right. \\ & \left. - \frac{\gamma}{1-u_{\text{typ}}^0} (u_{\text{typ}}^0 \mu_1 + 2u_{\text{typ}}^0 (\lambda_{\text{typ}}^0 - \mu_0)) \right] + \frac{f^2}{2\sqrt{2}} \frac{\gamma}{1-u_{\text{typ}}^0} \chi_{x, \text{typ}}^0 \end{aligned}$$

(F9)

⁵ This follows from the fact that the vector $\mathbf{v}_{\text{soft}}^0$ has zero overlap with σ^a , and thus the random variable $h(\mathbf{v}_{\text{soft}}^0)$ is uncorrelated to any of the random variables on which we are conditioning in this calculation. As a consequence, $h(\mathbf{v}_{\text{soft}}^0)$ fluctuates around its average value, that is zero.

$x \in \{Q, A\}$, where we have

$$\begin{aligned} \chi_{Q,\text{typ}}^0 &= \lim_{n \rightarrow 0} \mathbb{E}_n^{\text{cond}} \left[\mathbf{e}_{\min}^0 \cdot \frac{\tilde{\mathcal{H}}_Q^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 \right], \\ \chi_{A,\text{typ}}^0 &= \mathbb{E}_1^{\text{cond}} \left[\mathbf{e}_{\min}^0 \cdot \frac{\tilde{\mathcal{H}}_A^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 \right]. \end{aligned} \tag{F10}$$

The typical values $\lambda_{\text{typ}}^0, u_{\text{typ}}^0$ are known, and discussed in appendix E. To get the energy profile, it remains to determine the quantities (F10). We discuss this in appendix G.

F2. Softest mode at σ^1

Let us now consider the case in which \mathbf{v} is chosen to be aligned along the direction of the softest mode of $\mathcal{H}(\sigma^1)$, i.e. in the same notation as above,

$$\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^1 = \frac{\mathbf{e}_{\min}^1 - \sqrt{u^1} \mathbf{e}_{N-1}^1}{\sqrt{1-u^1}}, \quad u^1 = (\mathbf{e}_{\min}^1 \cdot \mathbf{e}_{N-1}^1)^2, \quad \frac{\tilde{\mathcal{H}}(\sigma^1)}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^1 = \lambda_{\min}^1 \mathbf{e}_{\min}^1.$$

In this case, using that for $p=3$ one has $\nabla^2 h(\sigma^0) \cdot \sigma^1 = \nabla^2 h(\sigma^1) \cdot \sigma^0$, we find

$$\begin{aligned} \mathbf{v}_{\text{soft}}^1 \cdot \frac{\nabla^2 h(\sigma^0)}{\sqrt{N-1}} \cdot \sigma^1 &= \mathbf{v}_{\text{soft}}^1 \cdot \frac{\nabla^2 h(\sigma^1)}{\sqrt{N-1}} \cdot \sigma^0 \\ &= \left(\frac{\mathbf{e}_{\min}^1 - \sqrt{u^1} \mathbf{e}_{N-1}^1}{\sqrt{1-u^1}} \right) \cdot \frac{\nabla^2 h(\sigma^1)}{\sqrt{N-1}} \cdot (q\sigma^1 - \sqrt{1-q^2} \mathbf{e}_{N-1}^1) \\ &= \sqrt{\frac{1-q^2}{1-u^1}} (\sqrt{u^1} \mathbf{e}_{N-1}^1 - \mathbf{e}_{\min}^1) \cdot \frac{\nabla^2 h(\sigma^1)}{\sqrt{N-1}} \cdot \mathbf{e}_{N-1}^1 \\ &= \sqrt{\frac{u^1(1-q^2)}{1-u^1}} (\mu_1 - \lambda_{\min}^1), \end{aligned} \tag{F11}$$

and

$$\mathbf{v}_{\text{soft}}^1 \cdot \frac{\nabla^2 h(\sigma^1)}{\sqrt{N-1}} \cdot \mathbf{v}_{\text{soft}}^1 = \frac{1}{1-u^1} (\lambda_{\min}^1 + u^1 \mu_1 - 2u^1 \lambda_{\min}^1) \tag{F12}$$

and finally

$$\mathbf{v}_{\text{soft}}^1 \cdot \frac{\nabla^2 h(\sigma^0)}{\sqrt{N-1}} \cdot \mathbf{v}_{\text{soft}}^1 = \frac{1}{1-u^1} \left(\mathbf{e}_{\min}^1 \cdot \frac{\tilde{\mathcal{H}}_x^0}{\sqrt{M}} \cdot \mathbf{e}_{\min}^1 - 2u^1 (\lambda_{\min}^1 - \mu_1) - u^1 \mu_0 \right). \tag{F13}$$

Therefore, when $\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^1$, the energy profile (C3), under the conditions that σ^a are stationary points with energy density ϵ_a , becomes equal to (F7) by swapping the indices 0 and 1, and same for β and γ , in the terms that contain f . Introducing

$$\chi_x^1 := \mathbf{e}_{\min}^1 \cdot \frac{\tilde{\mathcal{H}}_x^0}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^1, \tag{F14}$$

as well as the typical values as before, we get that the energy profile in this case reads:

$$\epsilon_{\mathbf{v}_{\text{soft}}^1}^{(x)}[\gamma;f] := \lim_{N \rightarrow \infty} \frac{h^{(x)}[\gamma;f]}{\sqrt{2N}} = (\gamma^3 + 3\gamma^2\beta q)\epsilon_1 + (\beta^3 + 3\beta^2\gamma q)\epsilon_0 + \gamma\beta f \sqrt{\frac{u_{\text{typ}}^1(1-q^2)}{2(1-u_{\text{typ}}^1)}} (\mu_1 - \lambda_{\text{typ}}^1) + \frac{f^2}{2\sqrt{2}} \left[\frac{\gamma}{1-u_{\text{typ}}^1} (\lambda_{\text{typ}}^1 + u_{\text{typ}}^1\mu_1 - 2u_{\text{typ}}^1\lambda_{\text{typ}}^1) - \frac{\beta}{1-u_{\text{typ}}^1} (u_{\text{typ}}^1\mu_0 + 2u_{\text{typ}}^1(\lambda_{\text{typ}}^1 - \mu_1)) \right] + \frac{f^2}{2\sqrt{2}} \frac{\beta}{1-u_{\text{typ}}^1} \chi_{x,\text{typ}}^1 \tag{F15}$$

$x \in \{Q, A\}$, where now

$$\chi_{Q,\text{typ}}^1 = \lim_{n \rightarrow 0} \mathbb{E}_n^{\text{cond}} \left[\mathbf{e}_{\min}^1 \cdot \frac{\tilde{\mathcal{H}}_Q^0}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^1 \right], \tag{F16}$$

$$\chi_{A,\text{typ}}^1 = \mathbb{E}_1^{\text{cond}} \left[\mathbf{e}_{\min}^1 \cdot \frac{\tilde{\mathcal{H}}_A^0}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^1 \right].$$

Notice that despite the analogy of the expression, in the quenched setting the problem is not symmetric since the statistics of the Hessian matrices is different for $a=0$ and $a=1$, and therefore the typical values of λ_{\min}^a, u^a can be quite different for $a=0$ and $a=1$, as we commented in appendix E. We discuss the quantities (F16) in appendix G.

Appendix G. Eigenvectors overlaps

The manipulations of appendix F show that in order to determine the typical value of the energy profile, one has to compute the typical value of the terms (F8) and (F14). Using the eigenvalue decomposition $\tilde{\mathcal{H}}_x^a = \sum_{\alpha=1}^{N-1} \tilde{\lambda}_\alpha^a \mathbf{u}_\alpha^a [\mathbf{u}_\alpha^a]^T$, for $x \in \{A, Q\}$ and $a \in \{0, 1\}$, we can formally write

$$\chi_x^0 = \sum_{\alpha=1}^{N-1} \tilde{\lambda}_\alpha^0 [\mathbf{e}_{\min}(\boldsymbol{\sigma}^0) \cdot \mathbf{u}_\alpha^0]^2, \quad \chi_x^1 = \sum_{\alpha=1}^{N-1} \tilde{\lambda}_\alpha^1 [\mathbf{e}_{\min}(\boldsymbol{\sigma}^1) \cdot \mathbf{u}_\alpha^1]^2, \tag{G1}$$

where the dependence of the eigenvalue and eigenvector distribution on x is implicit. In the limit of large matrix size $N \gg 1$, the quantities χ_x^a with $a=0, 1$ converge to their average. Following [32, 33], we introduce the averaged squared overlap between eigenstates of two different, correlated matrices of size $(N-1) \times (N-1)$:

$$\Phi(\lambda_\alpha^0, \lambda_\beta^1) := (N-1) \mathbb{E} \left[\left(\mathbf{u}_{\lambda_\alpha^0} \cdot \mathbf{u}_{\lambda_\beta^1} \right)^2 \right], \tag{G2}$$

where the expectation is over the realizations of the random matrices. We remark that such expression can take different expressions whether there are zero, one or two isolated eigenvalues among $\lambda_\alpha^0, \lambda_\beta^1$ [33]. We define with $\tilde{\rho}_N^a(\lambda)$ the average eigenvalues density of the matrices $\tilde{\mathcal{H}}_x^a$ (the leading order term, as well as eventual subleading corrections corresponding to isolated eigenvalues of the matrix). Then the term to be determined can be formally written as:

$$\chi_{x,\text{typ}}^0 = \int \lambda \tilde{\rho}_N^0(\lambda) \Phi(\lambda_{\min}^0, \lambda) d\lambda, \quad \chi_{x,\text{typ}}^1 = \int \lambda \tilde{\rho}_N^1(\lambda) \Phi(\lambda_{\min}^1, \lambda) d\lambda. \tag{G3}$$

In order to compute these terms, one has to determine both the density $\tilde{\rho}_N^a(\lambda)$ of each matrix $\tilde{\mathcal{H}}_x^a$ (paying attention to the possible existence of isolated eigenvalues), as well as the typical overlaps between the eigenvectors of $\tilde{\mathcal{H}}_x^a$ and the minimal eigenvector of $\tilde{\mathcal{H}}_x^{b \neq a}$, which are correlated according to appendix E. In the following, we discuss the quenched and annealed case separately.

G.1. Eigenvectors overlaps: the annealed setting.

Without loss of generality, we focus on the case in which $\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^0$, and we have to determine the typical value

$$\chi_{A,\text{typ}}^0 = \mathbb{E}_1^{\text{cond}} \left[\mathbf{e}_{\min}^0 \cdot \frac{\tilde{\mathcal{H}}_A^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^0 \right], \tag{G4}$$

where \mathbf{e}_{\min}^0 is the eigenvector associated to the minimal eigenvalue of the shifted Riemannian Hessian $\tilde{\mathcal{H}}_A(\sigma^0)$ given in (E6). As we argued in appendix E.1, the matrix $\tilde{\mathcal{H}}_A^1$ can have one or zero isolated eigenvalues. On the contrary, while $\tilde{\mathcal{H}}_A^0$ could *a priori* present an isolated eigenvalue, it can be checked that for no value of $\epsilon_{\text{gs}} \leq \epsilon_0, \epsilon_1 \leq \epsilon_{\text{th}}$ and q , does μ_0 generate an isolated eigenvalue in the region of parameters in which the complexity $\Sigma(\epsilon_1, q | \epsilon_0) > 0$. Therefore, the minimal eigenvalue is at the edge of the semicircle, $\lambda_{\min}^0 = -2\sigma$. This implies that the rescaled overlap (G2) is of $\mathcal{O}(1)$ for all possible values of λ . References [32, 33]. In the limit of $N \rightarrow +\infty$ we have

$$\tilde{\rho}_N^1(\lambda) = \rho_\sigma(\lambda) + \frac{\Theta(|\mu_1| - 2\sigma + \Delta^2/\sigma)}{N} \delta(\lambda - \lambda_{\text{iso}}^1) + \mathcal{O}\left(\frac{1}{N^2}\right), \quad \rho_\sigma(\lambda) = \frac{\sqrt{4\sigma^2 - \lambda^2}}{2\pi\sigma^2}, \tag{G5}$$

where Θ is non-zero only whenever (E8) holds. Then we obtain

$$\begin{aligned} \chi_{\text{typ},A}^0 &= \int \lambda \rho_\sigma(\lambda) \Phi(\lambda_{\min}^0, \lambda) d\lambda + \frac{\Theta}{N} \lambda_{\text{iso}}^1 \Phi(\lambda_{\min}^0, \lambda_{\text{iso}}^1) + \mathcal{O}\left(\frac{1}{N^2}\right) \\ &= \int \lambda \rho_\sigma(\lambda) \lambda \Phi(\lambda_{\min}^0, \lambda) d\lambda + \mathcal{O}\left(\frac{1}{N}\right). \end{aligned} \tag{G6}$$

From [32, 33] one can get the expression for the rescaled, squared overlaps $\Phi(\lambda_{\min}^0, \lambda)$ between eigenvalues belonging to the continuous part of the distribution. In the specific case in which $\lambda_{\min}^0 = -2\sigma$, this reads

$$\Phi(-2\sigma, \lambda) = \frac{\sigma^4 \sigma_W^2 (2\sigma^2 - \sigma_W^2)}{(2\sigma^4 - 2\sigma^2 \sigma_W^2 + \sigma_W^4 + \sigma^3 \lambda - \sigma \sigma_W^2 \lambda)^2} + \mathcal{O}\left(\frac{1}{N}\right), \quad \lambda \in [-2\sigma, 2\sigma] \tag{G7}$$

where the parameters given in (E3) for general values of p . Therefore

$$\chi_{A,\text{typ}}^0 = \int_{-2\sigma}^{2\sigma} \lambda \frac{\sqrt{4\sigma^2 - \lambda^2}}{2\pi} \frac{\sigma^2 \sigma_W^2 (2\sigma^2 - \sigma_W^2)}{(2\sigma^4 - 2\sigma^2 \sigma_W^2 + \sigma_W^4 + \sigma^3 \lambda - \sigma \sigma_W^2 \lambda)^2} d\lambda + \mathcal{O}\left(\frac{1}{N}\right) \tag{G8}$$

$$= \frac{1}{2} \sigma^2 \sigma_W^2 (2\sigma^2 - \sigma_W^2) \frac{a}{b^3} \frac{c^2 + 2\sqrt{1-c^2} - 2}{\sqrt{1-c^2}} + \mathcal{O}\left(\frac{1}{N}\right) \tag{G9}$$

where

$$\begin{cases} a = 2\sigma^4 - 2\sigma^2\sigma_W^2 + \sigma_W^4 \\ b = \sigma^3 - \sigma\sigma_W^2 \\ c = 2\sigma b/a. \end{cases} \quad (\text{G10})$$

For $p = 3$ one has $\sigma = \sqrt{6}$ and $\sigma_W = \sqrt{6(1-q)}$ and this expression reduces to:

$$\chi_{A,\text{typ}}^0 = \frac{1}{2}\sigma^2\sigma_W^2(2\sigma^2 - \sigma_W^2) \frac{a}{b^3} \frac{c^2 + 2\sqrt{1-c^2} - 2}{\sqrt{1-c^2}} = -2\sqrt{6}q. \quad (\text{G11})$$

Let us now consider $\chi_{A,\text{typ}}^1$. As we have remarked above, there exist values of $\epsilon_0, \epsilon_1, q$ such that $\Sigma(\epsilon_1, q|\epsilon_0) > 0$ and such that the matrix $\tilde{\mathcal{H}}_A^1$ exhibits an isolated eigenvalue (meaning that μ_1 satisfies equation (E8)). It can be checked that for this regime of parameters μ_1 is negative, and so it is the isolated eigenvalue, given that the two quantities have the same sign [33]. Therefore, when the isolated eigenvalue exists, it is the minimal eigenvalue of $\tilde{\mathcal{H}}_A^1$. On the other hand, we find that μ_0 does not generate an isolated eigenvalue in the regime of parameters we are interested in. When λ_{min}^1 is isolated, i.e. when μ_1 satisfies equation (E8), to compute $\chi_{A,\text{typ}}^1$ we have use a different expression for the rescaled, squared overlap Φ . In general, we can write

$$\begin{aligned} \chi_{A,\text{typ}}^1 &= \int_{-2\sigma}^{2\sigma} \lambda \frac{\sqrt{4\sigma^2 - \lambda^2}}{2\pi\sigma^2} [\Phi(-2\sigma, \lambda) (1 - \Theta(|\mu_1| - 2\sigma + \Delta^2/\sigma)) \\ &\quad + \Phi(\lambda_{\text{iso}}^1, \lambda) \Theta(|\mu_1| - 2\sigma + \Delta^2/\sigma)] d\lambda + \mathcal{O}\left(\frac{1}{N}\right) \end{aligned} \quad (\text{G12})$$

where the expression for $\Phi(\lambda_{\text{iso}}^1, \lambda)$ has been computed in [33], and applied to our case reads

$$\begin{aligned} &\Phi(\lambda_{\text{iso}}^1, y) \\ &= q_{\sigma, \Delta}(\lambda_{\text{iso}}^1, \mu_1) \left[\frac{4\Delta^2\sigma^4}{\sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2}} \frac{bc - ad}{c^2 + d^2} - 4\sigma^4\Delta^4 \frac{b_1c_1e_1 - a_1d_1e_1 - a_1c_1f_1 - b_1d_1f_1}{(c_1^2 + d_1^2)(e_1^2 + f_1^2)} (4\sigma^2 - y^2) \right. \\ &\quad \left. - 8\sigma^4\Delta^2 \frac{c_1b_2 - a_2b_2\Delta^2 - a_2c_1\sigma_W^2 - \sigma_W^2\Delta^2}{(c_1^2 + \Delta^4)(b_2^2 + \sigma_W^4)} (4\sigma^2 - y^2) + \frac{\Delta^4 g_\sigma(\lambda_{\text{iso}}^1)}{(\lambda_{\text{iso}}^1 - \mu_1)} \frac{1}{a_3^2 + b_3^2} \right] \end{aligned} \quad (\text{G13})$$

with parameters

$$\begin{aligned}
 a &= -(4\sigma^2 - \lambda_{\text{iso}}^1 y) \\
 b &= \sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2} \\
 c &= (2\sigma^2 - \sigma_W^2)^2 (\lambda_{\text{iso}}^1 - y)^2 - 2\sigma_W^2 (2\sigma^2 - \sigma_W^2) \sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2} (\lambda_{\text{iso}}^1 - y) \\
 &\quad + \sigma_W^4 \left([\lambda_{\text{iso}}^1]^2 - 4\sigma^2 \right) + \sigma_W^4 (y^2 - 4\sigma^2) \\
 d &= -2\sigma_W^2 \left[-\sigma_W^2 \sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2} + (2\sigma^2 - \sigma_W^2) (\lambda_{\text{iso}}^1 - y) \right] \\
 a_1 &= (\lambda_{\text{iso}}^1 - y)^2 - 2\sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2} (y - \lambda_{\text{iso}}^1) + [\lambda_{\text{iso}}^1]^2 + y^2 - 8\sigma^2 \\
 b_1 &= -2 \left[y - \lambda_{\text{iso}}^1 - \sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2} \right] \\
 c_1 &= -2\mu_0\sigma^2 - \Delta^2 y + 2\sigma^2 y \\
 d_1 &= \Delta^2 \tag{G14} \\
 e_1 &= (\sigma_W^2 - 2\sigma^2)^2 (\lambda_{\text{iso}}^1 - y)^2 - 2\sigma_W^2 (2\sigma^2 - \sigma_W^2) \sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2} (\lambda_{\text{iso}}^1 - y) \\
 &\quad + \sigma_W^4 \left([\lambda_{\text{iso}}^1]^2 - 4\sigma^2 \right) + \sigma_W^4 (y^2 - 4\sigma^2) \\
 f_1 &= -2\sigma_W^2 \left[(2\sigma^2 - \sigma_W^2) (\lambda_{\text{iso}}^1 - y) - \sigma_W^2 \sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2} \right] \\
 a_2 &= \lambda_{\text{iso}}^1 - y + \sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2} \\
 b_2 &= (\sigma_W^2 - 2\sigma^2) (\lambda_{\text{iso}}^1 - y) + \sigma_W^2 \sqrt{[\lambda_{\text{iso}}^1]^2 - 4\sigma^2} \\
 a_3 &= y - \mu_0 - \frac{\Delta^2 y}{2\sigma^2} \\
 b_3 &= \frac{\Delta^2}{2\sigma^2},
 \end{aligned}$$

and where the function $q_{\sigma,\Delta}$ and the explicit expression of λ_{iso}^1 are given in (E13) and (E11).

G.2. Eigenvectors overlaps: the quenched setting.

We discuss first the case in which $\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^0$, and we have to determine the expectation

$$\chi_{Q,\text{typ}}^0 = \lim_{n \rightarrow 0} \mathbb{E}_n^{\text{cond}} \left[\mathbf{e}_{\text{min}}^0 \cdot \frac{\tilde{\mathcal{H}}_Q^1}{\sqrt{N-1}} \cdot \mathbf{e}_{\text{min}}^0 \right]. \tag{G15}$$

In this case $\mathbf{e}_{\text{min}}^0$ is the eigenvector associated to the minimal eigenvalue of the shifted Riemannian Hessian $\tilde{\mathcal{H}}_Q^0$, which together with $\tilde{\mathcal{H}}_Q^1$ is defined in equations (E30) and (E34) respectively. As argued in appendix E.2, in the limit $n \rightarrow 0$, the matrix $\tilde{\mathcal{H}}_Q^0$ has no isolated

eigenvalues, because the zeros of equation (E38) become poles in the limit $n \rightarrow 0$. On the contrary, in the limit $n \rightarrow 0$ the matrix $\tilde{\mathcal{H}}_Q^1$ could present an isolated eigenvalue depending on the strength of μ_1 . Hence, $\lambda_{\min}^0 = -2\sigma$, and by applying the same reasoning as in the annealed case, we are left with the fact that any possible contribution $\Phi(\lambda_{\min}^0, \lambda_{\text{iso}}^1)$ comes at an order of $1/N$ and can be neglected to leading order at large N . Hence, one ends up with

$$\chi_{Q,\text{typ}}^0 = \chi_{A,\text{typ}}^0.$$

We turn now to the case $\mathbf{v} \rightarrow \mathbf{v}_{\text{soft}}^1$, where we need to determine

$$\chi_{Q,\text{typ}}^1 = \lim_{n \rightarrow 0} \mathbb{E}_n^{\text{cond}} \left[\mathbf{e}_{\min}^1 \cdot \frac{\tilde{\mathcal{H}}_Q^0}{\sqrt{N-1}} \cdot \mathbf{e}_{\min}^1 \right]. \tag{G16}$$

This case is a slightly more complicated, since $\tilde{\mathcal{H}}_Q^1$ can display an isolated eigenvalue in the limit $n \rightarrow 0$, and thus one has to consider how the function $\Phi(\lambda_{\text{iso}}^1, \lambda)$ changes in the quenched case, for $\lambda \in [-2\sigma, 2\sigma]$. The way to proceed is to replicate the computation in [33] in the case of finite n , and then send $n \rightarrow 0$. It turns out that this calculation reproduces the result in equation (G13), implying that

$$\chi_{Q,\text{typ}}^1 = \chi_{A,\text{typ}}^1.$$

G.3. On quenched vs annealed.

As we recalled in appendix B, when computing the complexity $\Sigma(\epsilon_1, q|\epsilon_0)$ one obtains the identity (B8), which implies that the calculation performed within the quenched formalism (making use of the replica trick with $n \rightarrow 0$) gives the same result as that performed within the annealed formalism (which corresponds to $n = 1$). This follows from the fact that the typical value of the overlap parameter q_1 , giving the typical overlap between the replicas of the secondary configuration σ^1 , takes the particularly simple value $q_1 = q^2$. This identity does not imply *a priori* that the typical energy profile computed within the quenched prescription is the same as that computed within the annealed prescription, since the statistics of the conditional Hessians (in particular, the finite rank perturbations) is different in the two cases. We find however that the two prescriptions give the same result for the energy profile, too, to leading order in N . In fact, the main difference between the two prescriptions is in the statistics of the conditional Hessian $\mathcal{H}(\sigma^0)$: in the quenched case, this Hessian exhibits n special line and columns with identical statistical properties, whose effect vanishes as $n \rightarrow 0$. In this limit, one recovers the statistics of the unconditional Hessian. This is consistent with the fact that in the quenched setup the reference configuration σ^0 is selected first, independently of the secondary one σ^1 , and therefore the statistics of its Hessian should not be affected by the coupling to σ^1 . On the contrary, in the annealed framework the primary and secondary configurations are treated on the same footing, and the conditional distribution of the Hessian at σ^0 has a special line and column that depends on the correlation with σ^1 . The two prescriptions give different results whenever this special line affects the eigenvalue distribution of the Hessian by generating isolated eigenvalues, which are not there in the quenched case. However, we find that this circumstance is never realized for the values of parameters ϵ_a, q that are of interest (such that $\Sigma(\epsilon_1, q|\epsilon_0) > 0$, $\epsilon_{\text{gs}} \leq \epsilon_a \leq \epsilon_{\text{th}}$ and $q \in [0, 1]$). As a consequence, in the regime of parameters that we are interested in the typical energy profile is the same with both prescriptions. For this reason, we will now drop the Q or A pedices in the following.

Appendix H. Resulting energy profiles

We give in this appendix the final result for the typical energy profile, in the case $p=3$. We separate the cases in which \mathbf{v} is chosen to be related to the softest mode of the Hessian at the reference and secondary configuration, respectively.

Softest mode at σ^0 . As it follows from the discussion of appendices F and G, in this case $\lambda_{\text{typ}}^0 = -2\sigma = -2\sqrt{6}$; Moreover, $\mathbf{e}_{\min}(\sigma^0)$ is a random vector uniformly distributed on $\mathcal{S}_N(1)$, so that its typical overlap with $\mathbf{e}_{N-1}(\sigma^0)$ is $u_{\text{typ}}^0 = \mathbf{e}_{\min}(\sigma^0) \cdot \mathbf{e}_{N-1}(\sigma^0) = 0$. Using our previously found results, the energy profile (F9) simplifies into:

$$\epsilon_{\mathbf{v}_{\text{soft}}^0}[\gamma;f] = (\gamma^3 + 3\gamma^2\beta q) \epsilon_1 + (\beta^3 + 3\beta^2\gamma q) \epsilon_0 - \sqrt{3}f^2(\gamma) \beta - f^2(\gamma) \gamma\sqrt{3}q,$$

which corresponds to the equation (12) in the main text.

Softest mode at σ^1 . In this case there are two possibilities: if there is no isolated eigenvalue at $\tilde{\mathcal{H}}(\sigma^1)$, then $\lambda_{\min}^1 = -2\sigma$ and we obtain a similar expression as above:

$$\epsilon_{\mathbf{v}_{\text{soft}}^1}[\gamma;f] = (\gamma^3 + 3\gamma^2\beta q) \epsilon_1 + (\beta^3 + 3\beta^2\gamma q) \epsilon_0 - \sqrt{3}f^2(\gamma) \gamma - f^2(\gamma) \beta\sqrt{3}q.$$

If there is a negative isolated eigenvalue, then the formula changes, and equation (F15) becomes:

$$\begin{aligned} \epsilon_{\mathbf{v}_{\text{soft}}^1}[\gamma;f] &= (\gamma^3 + 3\gamma^2\beta q) \epsilon_1 + (\beta^3 + 3\beta^2\gamma q) \epsilon_0 + \gamma\beta f \sqrt{\frac{u_{\text{typ}}^1(1-q^2)}{2(1-u_{\text{typ}}^1)}} (\mu_1 - \lambda_{\text{typ}}^1) \\ &+ \frac{f^2}{2\sqrt{2}} \left[\frac{\gamma}{1-u_{\text{typ}}^1} (\lambda_{\text{typ}}^1 + u_{\text{typ}}^1\mu_1 - 2u_{\text{typ}}^1\lambda_{\text{typ}}^1) \right. \\ &\left. - \frac{\beta}{1-u_{\text{typ}}^1} (u_{\text{typ}}^1\mu_0 + 2u_{\text{typ}}^1(\lambda_{\text{typ}}^1 - \mu_1)) \right] + \frac{f^2}{2\sqrt{2}} \frac{\beta}{1-u_{\text{typ}}^1} \chi_{\text{iso}}^1 \end{aligned}$$

where we recall the values of the various quantities:

$$\begin{aligned} \lambda_{\text{typ}}^1 &= \frac{2\mu_1\sigma^2 - \Delta^2\mu_1 - \text{sign}(\mu_1)\Delta^2\sqrt{\mu_a^2 - 4(\sigma^2 - \Delta^2)}}{2(\sigma^2 - \Delta^2)} \\ u_{\text{typ}}^1 &= \mathfrak{q}_{\sigma,\Delta}(\lambda_{\text{typ}}^1, \mu_1) \\ \chi_{\text{iso}}^1 &:= \int_{-2\sigma}^{2\sigma} \lambda \frac{\sqrt{4\sigma^2 - \lambda^2}}{2\pi\sigma^2} \Phi(\lambda_{\text{typ}}^1, \lambda) d\lambda \end{aligned}$$

where the integral χ_{iso}^1 can be easily computed numerically. This expression coincides with equation (14) in the main text.

Appendix I. Gradient along the geodesic path

In this appendix, we derive the statistical distribution of the vector $\nabla h(\sigma[\gamma;f=0])$ at each configuration along the geodesic path, parametrized as (4) with $f(\gamma) = 0$. We denote each configuration along the path simply as $\sigma(\gamma) := \sigma[\gamma;0]$, and take $p=3$. For any fixed value of $\gamma \in (0, 1)$, plugging the expression for $\sigma(\gamma)$ inside the formula for ∇h one gets

$$\nabla h(\sigma(\gamma)) = \gamma^2 \nabla h(\sigma^1) + \beta^2 \nabla h(\sigma^0) + \gamma\beta \nabla^2 h(\sigma^0) \cdot \sigma^1, \quad (11)$$

which using (E14) and conditioning on the properties of σ^a reduces to:

$$\begin{aligned} \nabla h(\sigma(\gamma)) &= \gamma^2 3\sqrt{2N}\epsilon_1 \sigma^1 + \beta^2 3\sqrt{2N}\epsilon_0 \sigma^0 + \gamma\beta \nabla^2 h(\sigma^0) \cdot \left(q\sigma^0 - \sqrt{1-q^2} \mathbf{e}_{N-1}(\sigma^0) \right) \\ &= \gamma^2 3\sqrt{2N}\epsilon_1 \sigma^1 + \beta^2 3\sqrt{2N}\epsilon_0 \sigma^0 + \gamma\beta \left[6q\sqrt{2N}\epsilon_0 \sigma^0 - \sqrt{1-q^2} \nabla^2 h(\sigma^0) \cdot \mathbf{e}_{N-1}(\sigma^0) \right]. \end{aligned} \quad (12)$$

We now consider the vector $\mathbf{g}(\gamma) := \mathbf{g}(\sigma(\gamma))$, that is the projection of ∇h on the $(N-1)$ -dimensional tangent plane $\tau[\sigma(\gamma)]$. We choose a basis of this tangent plane in such a way that the first $N-2$ elements span the subspace orthogonal to both σ^0, σ^1 ; we denote them with $\mathbf{x}_1, \dots, \mathbf{x}_{N-2}$ as in appendix A. The remaining vector equals to

$$\mathbf{e}_{N-1}(\gamma) := \mathbf{e}_{N-1}(\sigma(\gamma)) = A(\sigma^0 + C\sigma^1), \quad A = (1 + 2Cq + C^2)^{-1/2}, \quad C = -\frac{\gamma q + \beta}{\gamma + \beta q}. \quad (13)$$

This vector is orthogonal to $\sigma(\gamma)$ and has unit norm. It is simple to check that it coincides with the tangent vector to the geodesic path at the point $\sigma(\gamma)$. Therefore, the gradient $\mathbf{g}(\gamma) = \mathbf{g}^{\parallel}(\gamma) + \mathbf{g}^{\perp}(\gamma)$, where $\mathbf{g}^{\parallel}(\gamma) = (\mathbf{g}(\gamma) \cdot \mathbf{e}_{N-1}(\gamma)) \mathbf{e}_{N-1}(\gamma)$ is the component tangent to the geodesic, while $\mathbf{g}^{\perp}(\gamma)$ is the orthogonal one. Using that

$$\mathbf{e}_{N-1}(\gamma) = (A + ACq)\sigma^0 - AC\sqrt{1-q^2}\mathbf{e}_{N-1}(\sigma^0) \quad (14)$$

we see that the tangent component equals to:

$$\begin{aligned} \nabla h(\sigma(\gamma)) \cdot \mathbf{e}_{N-1}(\gamma) &= 3\gamma^2 \sqrt{2N}\epsilon_1 (Aq + AC) + 3\beta^2 \sqrt{2N}\epsilon_0 (A + ACq) \\ &\quad + \gamma\beta \left[6q\sqrt{2N}\epsilon_0 (A + ACq) - \sqrt{1-q^2} \mathbf{e}_{N-1}(\gamma) \cdot \nabla^2 h(\sigma^0) \cdot \mathbf{e}_{N-1}(\sigma^0) \right]. \end{aligned} \quad (15)$$

One can check explicitly that this expression vanishes at the point γ where the geodesic energy profile reaches its maximum.

Consider now the orthogonal component of the gradient. In the notation of appendix E, we see that component-wise in the chosen local basis it holds for $i < N-1$:

$$\frac{\nabla h(\sigma(\gamma))}{\sqrt{N-1}} \cdot \mathbf{x}_i = -\gamma\beta \sqrt{1-q^2} \mathbf{x}_i \cdot \frac{\nabla^2 h(\sigma^0)}{\sqrt{N-1}} \cdot \mathbf{e}_{N-1}(\sigma^0) = -\gamma\beta \sqrt{1-q^2} m_{i,N-1}^0. \quad (16)$$

Therefore, the orthogonal component $\mathbf{g}^{\perp}(\gamma)$ is proportional to the vector that makes up the last column of the shifted Hessian $\tilde{\mathcal{H}}(\sigma^0)$ (neglecting the last component of the column). This vector is in general not vanishing at the point that corresponds to the maximum of the energy profile along the geodesic. We define the normalized vector

$$\begin{aligned} \mathbf{v}_{\text{Hess}} &= Z \sum_{i=1}^{N-2} \left[\mathbf{x}_i \cdot \tilde{\mathcal{H}}(\sigma^0) \cdot \mathbf{e}_{N-1}(\sigma^0) \right] \mathbf{x}_i \\ &= \frac{1}{\sqrt{\sum_{i=1}^{N-2} [m_{i,N-1}^0]^2}} (m_{1,N-1}^0, m_{2,N-1}^0, \dots, m_{N-2,N-1}^0, 0)^T, \end{aligned} \quad (17)$$

which is orthogonal to both σ^a (Z is the normalization factor). Plugging this into (11) and making use of the fact that the entries $m_{i,N-1}^a$ of the Hessians are uncorrelated to all other

entries of the Hessian matrices, we see that the third line in equation (11) vanishes on average, while in the second line we have:

$$\begin{aligned} \mathbb{E} \left[\mathbf{v}_{\text{Hess}} \cdot \frac{\tilde{\mathcal{H}}(\boldsymbol{\sigma}^0)}{\sqrt{2N}} \cdot \mathbf{e}_{N-1}(\boldsymbol{\sigma}^0) \right] &= \mathbb{E} \left[\sqrt{\frac{\sum_{i=1}^{N-2} [m_{i,N-1}^0]^2}{2N}} \right] \\ &= \sqrt{\frac{\Delta^2}{2}} + \mathcal{O}\left(\frac{1}{N}\right) \stackrel{p=3}{=} \sqrt{\frac{3(1-q^2)}{1+q^2}} + \mathcal{O}\left(\frac{1}{N}\right), \end{aligned} \quad (18)$$

which leads to equation (15) in the main text.

Finally, we determine the overlap between \mathbf{v}_{Hess} and $\mathbf{v}_{\text{soft}}^1$, to show that the two vectors are correlated when $u_{\text{typ}}^1 \neq 0$. It holds:

$$\begin{aligned} \mathbf{v}_{\text{Hess}} \cdot \mathbf{v}_{\text{soft}}^1 &= \mathbf{e}_{N-1}(\boldsymbol{\sigma}^0) \cdot \frac{\nabla^2 h(\boldsymbol{\sigma}^0)}{\sqrt{N-1}} \cdot \mathbf{v}_{\text{soft}}^1 = \frac{q\boldsymbol{\sigma}^0 - \boldsymbol{\sigma}^1}{\sqrt{1-q^2}} \cdot \frac{\nabla^2 h(\boldsymbol{\sigma}^0)}{\sqrt{N-1}} \cdot \mathbf{v}_{\text{soft}}^1 \\ &= -\frac{\boldsymbol{\sigma}^1}{\sqrt{1-q^2}} \cdot \frac{\nabla^2 h(\boldsymbol{\sigma}^0)}{\sqrt{N-1}} \cdot \mathbf{v}_{\text{soft}}^1, \end{aligned} \quad (19)$$

which is equivalent to (F11) and thus of $\mathcal{O}(1)$ and non-vanishing when $u_{\text{typ}}^1 \neq 0$, as claimed in the main text.

ORCID iDs

Alessandro Pacco  <https://orcid.org/0009-0001-7323-6917>

Giulio Biroli  <https://orcid.org/0000-0003-4115-2644>

Valentina Ros  <https://orcid.org/0000-0001-8189-8809>

References

- [1] Wales D J, Miller M A and Walsh T R 1998 Archetypal energy landscapes *Nature* **394** 758–60
- [2] Jónsson H, Mills G and Jacobsen K W 1998 Nudged elastic band method for finding minimum energy paths of transitions *Classical and Quantum Dynamics in Condensed Phase Simulations* (World Scientific) pp 385–404
- [3] Mauri E, Cocco S and Monasson R 2023 Mutational paths with sequence-based models of proteins: From sampling to mean-field characterization *Phys. Rev. Lett.* **130** 158402
- [4] Tian P and Best R B 2020 Exploring the sequence fitness landscape of a bridge between protein folds *PLoS Comput. Biol.* **16** e1008285
- [5] Draxler F, Veschgini K, Salmhofer M and Hamprecht F 2018 Essentially no barriers in neural network energy landscape *Proc. 35th Int. Conf. on Machine Learning (Proc. Machine Learning Research vol 80) (10–15 July 2018)* (PMLR) pp 1309–18
- [6] Daniel Freeman C and Bruna J 2016 Topology and geometry of half-rectified network optimization *Int. Conf. on Learning Representations*
- [7] Livio Annesi B, Lauditi C, Lucibello C, Malatesta E M, Perugini G, Pittorino F and Saglietti L 2023 The star-shaped space of solutions of the spherical negative perceptron (arXiv:2305.10623)
- [8] Mézard M, Parisi G and Angel Virasoro M 1987 *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications* vol 9 (World Scientific Publishing Company)
- [9] Berthier L and Biroli G 2011 Theoretical perspective on the glass transition and amorphous materials *Rev. Mod. Phys.* **83** 587
- [10] Parisi G, Urbani P and Zamponi F 2020 *Theory of Simple Glasses: Exact Solutions in Infinite Dimensions* (Cambridge University Press)

- [11] Debenedetti P G and Stillinger F H 2001 Supercooled liquids and the glass transition *Nature* **410** 259–67
- [12] Nelson O J, Luthey-Schulten Z and Wolynes P G 1997 Theory of protein folding: the energy landscape perspective *Annu. Rev. Phys. Chem.* **48** 545–600
- [13] Semerjian G 2008 On the freezing of variables in random constraint satisfaction problems *J. Stat. Phys.* **130** 251–93
- [14] Garipov T, Izmailov P, Podoprikhin D, Vetrov D P and Wilson A G 2018 Loss surfaces, mode connectivity and fast ensembling of dnns *Advances in Neural Information Processing Systems* p 31
- [15] Crisanti A and Sommers H-J 1992 The spherical p-spin interaction spin glass model: the statics *Z. Phys. B* **87** 341–54
- [16] Crisanti A, Horner H and Sommers H J 1993 The spherical p-spin interaction spin-glass model: the dynamics *Z. Phys. B* **92** 257–71
- [17] Ning X, Vitelli V, Liu A J and Nagel S R 2010 Anharmonic and quasi-localized vibrations in jammed solids—modes for mechanical failure *Europhys. Lett.* **90** 56001
- [18] Widmer-Cooper A, Perry H, Harrowell P and Reichman D R 2008 Irreversible reorganization in a supercooled liquid originates from localized soft modes *Nat. Phys.* **4** 711–5
- [19] Cavagna A, Giardina I and Parisi G 1998 Stationary points of the Thouless-Anderson-Palmer free energy *Phys. Rev. B* **57** 11251
- [20] Cavagna A, Giardina I and Parisi G 1997 An investigation of the hidden structure of states in a mean-field spin-glass model *J. Phys. A: Math. Gen.* **30** 7021
- [21] Cavagna A, Giardina I and Parisi G 1997 Structure of metastable states in spin glasses by means of a three replica potential *J. Phys. A: Math. Gen.* **30** 4449
- [22] Cavagna A, Garrahan J P and Giardina I 1999 Quenched complexity of the mean-field p-spin spherical model with external magnetic field *J. Phys. A: Math. Gen.* **32** 711
- [23] Auffinger A, Ben Arous G and Černý J 2013 Random matrices and complexity of spin glasses *Commun. Pure Appl. Math.* **66** 165–201
- [24] Subag E 2017 The complexity of spherical p-spin models—a second moment approach *Ann. Probab.* **45** 3385–450
- [25] Auffinger A and Gold J 2020 The number of saddles of the spherical p-spin model (arXiv:2007.09269)
- [26] Ros V, Biroli G and Cammarota C 2019 Complexity of energy barriers in mean-field glassy systems *Europhys. Lett.* **126** 20003
- [27] Ros V 2020 Distribution of rare saddles in the p-spin energy landscape *J. Phys. A: Math. Theor.* **53** 125002
- [28] Ros V, Biroli G and Cammarota C 2021 Dynamical instantons and activated processes in mean-field glass models *SciPost Phys.* **10** 002
- [29] Franz S and Parisi G 1998 Effective potential in glassy systems: theory and simulations *Physica A* **261** 317–39
- [30] Bolhuis P G, Chandler D, Dellago C and Geissler P L 2002 Transition path sampling: throwing ropes *Annu. Rev. Phys. Chem.* **53** 291–318
- [31] Bun J, Allez R, Bouchaud J-P and Potters M 2016 Rotational invariant estimator for general noisy matrices *IEEE Trans. Inf. Theory* **62** 7475–90
- [32] Bun J, Bouchaud J-P and Potters M 2018 Overlaps between eigenvectors of correlated random matrices *Phys. Rev. E* **98** 052145
- [33] Pacco A and Ros V 2023 Overlaps between eigenvectors of spiked, correlated random matrices: From matrix principal component analysis to random Gaussian landscapes *Phys. Rev. E* **108** 024145
- [34] Khomenko D, Reichman D R and Zamponi F 2021 Relationship between two-level systems and quasilocated normal modes in glasses *Phys. Rev. Mater.* **5** 055602
- [35] Wyart M, Nagel S R and Witten T A 2005 Geometric origin of excess low-frequency vibrational modes in weakly connected amorphous solids *Europhys. Lett.* **72** 486
- [36] Broderix K, Bhattacharya K K, Cavagna A, Zippelius A and Giardina I 2000 Energy landscape of a Lennard-Jones liquid: Statistics of stationary points *Phys. Rev. Lett.* **85** 5360
- [37] Angelani L, Di Leonardo R, Ruocco G, Scala A and Sciortino F 2000 Saddles in the energy landscape probed by supercooled liquids *Phys. Rev. Lett.* **85** 5356
- [38] Coslovich D, Ninarello A and Berthier L 2019 A localization transition underlies the mode-coupling crossover of glasses *SciPost Phys.* **7** 077

- [39] Heuer A 2008 Exploring the potential energy landscape of glass-forming systems: from inherent structures via metabasins to macroscopic transport *J. Phys.: Condens. Matter* **20** 373101
- [40] Baitty-Jesi M, Biroli G and Reichman D R 2021 Revisiting the concept of activation in supercooled liquids *Eur. Phys. J. E* **44** 77
- [41] Rizzo T 2021 Path integral approach unveils role of complex energy landscape for activated dynamics of glassy systems *Phys. Rev. B* **104** 094203
- [42] Folena G, Franz S and Ricci-Tersenghi F 2020 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
- [43] Kent-Dobias J 2024 Arrangement of nearby minima and saddles in the mixed spherical energy landscapes *SciPost Phys.* **16** 1