School "Optimization & Algorithms" - Toulouse 2024

# HIGH-DIMENSIONAL RANDOM LANDSCAPES

Valentina Ros
valentina.ros@cnrs.fr

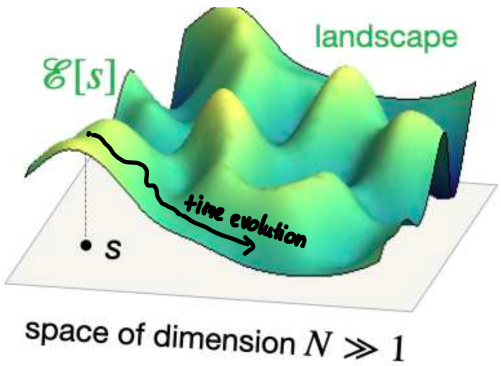
Laboratoire de Physique Théorique et Modèles Statistiques

[If you find typos please let me know!]

Exercises are at the end of these notes.

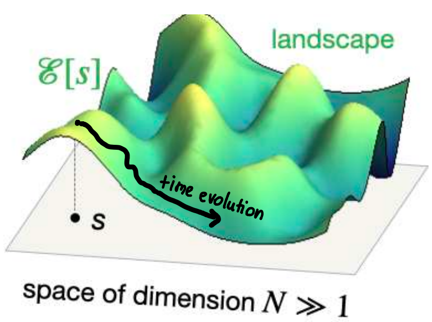
**WHAT:** High-D random landscapes are functions of many variables $\mathcal{E}[\vec{s}]$, $\vec{s} = (s_1, \dots, s_N)$ with $N \gg 1$, which are random, with given $P[\mathcal{E}[\vec{s}]]$ (in the following, Gaussian)

**WHY:** Many 'complex systems' are inherently high-dimensional. They evolve trying to optimize some function (fitness, energy, cost...). Function encodes complex interactions between constituents, often modelled with random variables.

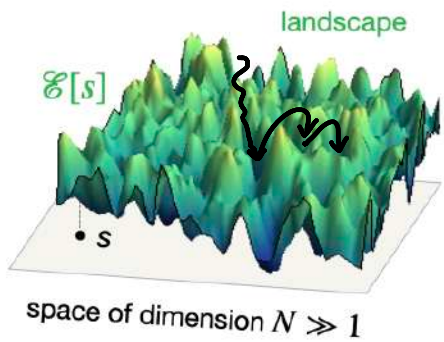



What to expect from this optimization processes in high-D, typically (i.e., with high probability)?

**HOW:** Characterize landscapes structure & its dynamical exploration using tools of stat physics ($N \gg 1$) of disordered systems:
random matrix theory, saddle-point & large-N limits, large-deviations, replica tricks, Kac-Rice counting formulas....




Scenario 1: "smooth" landscape.




Scenario 2: "rugged" landscape.

# HIGH-D RANDOM LANDSCAPES

## PART I: QUADRATIC HIGH-D LANDSCAPES

### WHY: an example from high-D inference

An 'easy' inference problem – From denoising to landscapes – Questions & strategy

### HOW: Random Matrix Theory

From landscapes back to random matrices – Basic RMT facts

### WHAT: Ground state, landscape, dynamics

Recovering the signal – A landscape of saddles – DMFT & beyond

## PART II: RUGGED HIGH-D LANDSCAPES

### WHY: another example from high-D inference

A 'hard' inference problem: noisy tensors – Landscape problem, & complexity

### HOW: Kac-Rice formalism

Averages vs typical values, and replicas – Kac-Rice formula(s) – Computing the complexity: 3 steps – The annealed complexity

### WHAT: Ground state, landscape, dynamics

Recovering the signal – A landscape of minima – DMFT. And beyond?

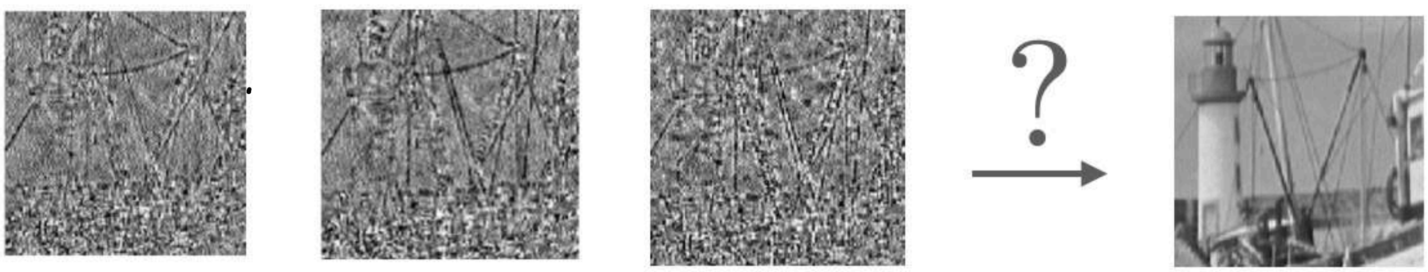
# PART I

## Quadratic high-D landscapes

# I.1 WHY: AN INFERENCE EXAMPLE

## An 'easy' inference problem: noisy matrices

Inference problem: measure a 'signal' corrupted by noise. Combining measurements, can recover information on signal?



(figure adapted from the web)

Denoising of matrices ("spiked" matrices): JOHNSTONE 2001

$$\boxed{\hat{M} = \underbrace{r}_{\text{signal strength}} \, \underbrace{\frac{\vec{v}\,\vec{v}^T}{N}}_{\text{signal}} + \underbrace{\hat{J}}_{\text{noise (randomness)}}}$$

size $N \times N$, $r \geq 0$

$N \gg 1$

■ SIGNAL $\vec{v}$: vector of norm $\|\vec{v}\|^2 = N$, $\vec{v} = (v_1, \dots, v_N)$
Unknown. Quenched (fixed). Independent of $\hat{J}$.

■ NOISE $\hat{J}$: matrix with random, symmetric ($J_{ij} = J_{ji}$) entries, $N \times N$. Gaussian statistics:

$$\langle J_{ij} \rangle = 0, \quad \langle J_{ij}^2 \rangle = \frac{\sigma^2}{N}(1 + \delta_{ij})$$

Probability to observe one instance of $\hat{J}$:

$$P_N(\hat{J})\, d\hat{J} = A_N \, e^{-\frac{N}{2\sigma^2}\sum_{i<j} J_{ij}^2 - \frac{N}{4\sigma^2}\sum_{i=1}^{N} J_{ii}^2} \prod_{i\leq j} dJ_{ij}$$

$$= A_N \, e^{-\frac{N}{4\sigma^2}\mathrm{Tr}(\hat{J}^2)} \prod_{i\leq j} dJ_{ij} \qquad A_N = \frac{1}{2^N}\left(\frac{N}{2\pi\sigma^2}\right)^{\frac{N(N+1)}{2}}$$

"GAUSSIAN ORTHOGONAL ENSEMBLE" = rotationally invariant ensemble. $\hat{O}$ rotation ($\hat{O}\hat{O}^T = \hat{1}$).
Matrix $\hat{J}$ in new basis: $\hat{J}_R = \hat{O}\hat{J}\hat{O}^T$.
Rotationally invariant means: $\hat{J}$ has same prob. as $\hat{J}_R = \hat{O}\hat{J}\hat{O}^T$: $\hat{J} \overset{\text{in law}}{\sim} \hat{J}_R$

Notice: Same eigenvalues, eigenvectors $\vec{u}_R = \hat{O}\vec{u}$. The evector of $\hat{J}$ has same distribution as any other vector obtained from it with rotation $\Rightarrow$ uniformly distributed vector on sphere.

## From denoising to landscapes

Estimator (guess) of $\vec{v}$: $\vec{S}_{gs} = \underset{\|\vec{s}\|^2 = N}{\operatorname{argmax}} \; \vec{s}^T \cdot \hat{M} \vec{s}$

this is "maximum likelihood estimator" of the signal $\vec{v}$.

### Maximum-likelihood

$$\hat{M} = r \frac{\vec{v}\vec{v}^T}{N} + \hat{J}$$

$\hat{M}$ - observation
$\vec{v}$ - unkown signal
$\hat{J}$ - iid gaussians

Bayes formula:

$$\underbrace{P(\vec{s}\,|\,\hat{M})}_{\text{POSTERIOR}} = \underbrace{P_0(\vec{s})}_{\text{PRIOR}} \underbrace{P(\hat{M}\,|\,\vec{s})}_{\text{LIKELIHOOD}} \frac{1}{P(\hat{M})} = P_0(\vec{s}) \frac{e^{-\frac{N}{4\sigma^2}\sum_{i,j}\left(M_{ij} - \frac{r}{N}s_i s_j\right)^2}}{Z(\hat{M})}$$

$$\mathcal{L}(\vec{s}\,|\,\hat{M}) = \log P(\hat{M}\,|\,\vec{s}) = -\frac{N}{2\sigma^2} \sum_{i\le j} \left(M_{ij} - \frac{r}{N} s_i s_j\right)^2 \left(\frac{1}{1+\delta_{ij}}\right) + \ell(\hat{M})$$

"log-likelihood"

The maximum-likelihood estimator is the vector that maximizes the log-likelihood.

If we know $\|\vec{v}\|^2 = N$, we can assume $\|\vec{s}\|^2 = N$ and thus the estimator is minimizing

$$\sim \sum_{i,j=1}^{N} \left( M_{ij} - \frac{r}{N} s_i s_j \right)^2 = \sum_{i,j=1}^{N} M_{ij}^2 + \frac{r^2}{N^2} \|\vec{s}\|^4 - \frac{2r}{N} \sum_{ij} M_{ij} s_i s_j$$

$$\Longrightarrow \vec{s}_{gs} = \underset{\|\vec{s}\|^2 = N}{\arg\max}\ \vec{s} \cdot \hat{M} \vec{s}$$

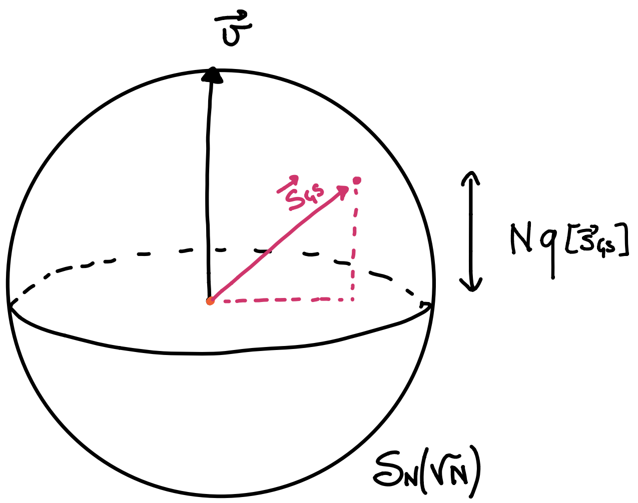
$s_{gs}$ is also the ground state of the energy landscape:

$$\mathcal{E}[\vec{s}] = -\frac{1}{2} \sum_{i,j=1}^{N} s_i M_{ij} s_j = -\frac{1}{2} \sum_{i,j=1}^{N} \left[ \underbrace{J_{ij} s_i s_j}_{\text{random isotropic}} + \underbrace{rN \left( \frac{\vec{v} \cdot \vec{s}}{N} \right)^2}_{\text{deterministic biased towards } \vec{v}} \right]$$

defined on $S_N(\sqrt{N}) = \{ \vec{s} : \|\vec{s}\|^2 = N \}$

Finding the estimator $\Longleftrightarrow$ solving optimization problem for random landscape $\mathcal{E}[\vec{s}]$.

Define the "OVERLAP WITH SIGNAL"




$$q_N[\vec{s}] = \frac{\vec{s}\cdot\vec{v}}{N}$$

$$q[\vec{s}] = \lim_{N\to\infty} q_N[\vec{s}]$$

$r=0$: random, fully-connected interactions between $s_i$ : "pure spherical $p=2$ model"
Isotropic statistics

$$\langle \mathcal{E}[\vec{s}]\rangle = 0$$

$$\langle \mathcal{E}[\vec{s}]\,\mathcal{E}[\vec{s}\,']\rangle = \frac{N}{2}\sigma^2\left(\frac{\vec{s}\cdot\vec{s}\,'}{N}\right)^2 \Longrightarrow$$

by entropy, expect $\vec{s}_{gs} \perp \vec{v}$ for $r=0$ (see below)

$\sigma=0$: the points in the vicinity of $\vec{v}$ are favored energetically, $\vec{s}_{gs} = \vec{v}$

Competition leads to transitions in $r/\sigma$ (signal-to-noise ratio) when $N\to\infty$.

# High-D geometry: typical values of overlaps

Let $\vec{v}$ be fixed vector $\|\vec{v}\|^2 = N$. Assume $\vec{s}$ uniformly taken on sphere. Then typical value of $\left(\frac{\vec{v}\cdot\vec{s}}{N}\right) \xrightarrow{N\to\infty} 0$. With overwhelming probability, two vectors are orthogonal when $N \to \infty$.

Indeed:

Basis-independent. Choose basis in which $\vec{v} = \sqrt{N}(0,0,0,1)$

$$\left\langle \left(\frac{\vec{v}\cdot\vec{s}}{N}\right)^2 \right\rangle = \frac{1}{|S_N(\sqrt{N})|} \int_{S_N(\sqrt{N})} \prod_{i=1}^{N} ds_i \left(\frac{\vec{v}\cdot\vec{s}}{N}\right)^2 = \frac{1}{|S_N|} \int_{S_N(\sqrt{N})} \prod_{i=1}^{N} ds_i \, \frac{s_N^2}{N} \; (=)$$

Surface of $(N-1)$-dimensional sphere of radius $R$, in $\mathbb{R}^N$:

$$|S_N(R)| = \frac{2\pi^{N/2}}{\Gamma(N/2)} R^{N-1} \qquad (*)$$

Rescale: $\vec{\sigma} = \vec{s}/\sqrt{N}$:

$$(=)\; N^{N/2} \int_{S_N(1)} \prod_{i=1}^{N} d\sigma_i \, \sigma_N^2 = N^{N/2-1} \int d\sigma_N \, \sigma_N^2 \underbrace{\left[ \int \prod_{i=1}^{N-1} d\sigma_i \, \delta\left( \sum_{i=1}^{N-1} \sigma_i^2 - [1-\sigma_N^2] \right) \right]}_{|S_{N-1}(\sqrt{1-\sigma_N^2})|}$$

$$\overset{N\gg 1}{\approx} \int d\sigma_N \, \sigma_N^2 \, \frac{e^{\frac{N}{2}\log[2\pi e(1-\sigma_N^2)] + o(N)}}{e^{N/2 \log(2\pi e)}} \xrightarrow[N\to\infty]{\text{SADDLE POINT: } \sigma_N^{SP} = 0} 0$$

Notice:

Could do this for all components by rotational invariance: all $\sigma_i^2$ are statistically equivalent $\Rightarrow \sum_{i=1}^{N} \sigma_i^2 \approx N \cdot \langle \sigma_i^2 \rangle = 1 \Rightarrow \langle \sigma_i^2 \rangle \approx 1/N$
$\Rightarrow \langle \left(\frac{v \cdot S}{N}\right)^2 \rangle \approx 1/N$

Indeed, setting $\sigma_N^2 = a/N$ and using (*):

$$\left\langle \left(\frac{\vec{S} \cdot \vec{v}}{N}\right)^2 \right\rangle \stackrel{\circ}{=} \frac{N^{N/2-1}}{|S_N(\sqrt{N})|} \frac{2\pi^{\frac{N-1}{2}}}{\Gamma\left(\frac{N-1}{2}\right)} \int \frac{da}{2\sqrt{Na}} \left(\frac{a}{N}\right) \left(1-\frac{a}{N}\right)^{\frac{N-2}{2}}$$

Since $\left(1-\frac{a}{N}\right)^{N/2} \xrightarrow{N \to \infty} e^{-a/2}$,

$$\stackrel{\circ}{=} C_N \underbrace{\frac{N^{N/2} \pi^{N/2}}{\Gamma\left(\frac{N-1}{2}\right) |S_N(\sqrt{N})|}}_{\text{exponential cancel!}} \int da \, f_N(a)$$

# Questions & strategy

▶ Three questions:

[Q1] RECOVERY QUESTION ($T=0$ EQUILIBRIUM)

for which values of $r/\sigma$ is $\vec{s}_{gs}$ informative of signal $\vec{v}$, i.e. "close" to $\vec{v}$ in configuration space?

For $N\to\infty$, $\quad q[\vec{s}_{gs}]>0 \quad$ ("magnetization")

[Q2] LANDSCAPE QUESTION (METASTABLE STATES)

are there many local minima / stationary points at higher energy? How far from $\vec{s}_{gs}$? How far from $\vec{v}$?

In following, "many" $=\mathcal{O}(e^N)$

[Q3] ALGORITHMIC QUESTION (DYNAMICS)

founding $\vec{s}_{gs}$ with (local) optimization algorithms (gradient descent / Langevin: $\frac{d\vec{s}(t)}{dt}=-\nabla_{\perp}\mathcal{E}[\vec{s}]+\sqrt{2T}\,\vec{\eta}(t)$) is

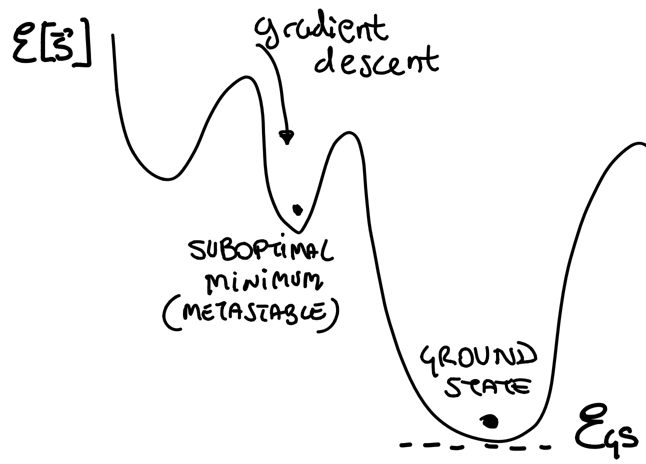
easy: timescales $\tau_{typ}\sim\mathcal{O}(N^{\alpha})$, or

hard: timescales $\tau_{typ}\sim\mathcal{O}(e^N)$?

gradient on the sphere

Q2 and Q3 related: optimization hard when many metastable states/local minima in which system gets stuck!

↳ glassiness




▶ The strategy:

Study the typical distribution of stationary points $\vec{s}^*$ : $\nabla_\perp \mathcal{E}[\vec{s}^*]=0$ as a function of:

(i) energy density $\epsilon_N[\vec{s}^*]= \mathcal{E}[\vec{s}^*]/N$

(ii) stability / local curvature (minima, saddles)

curvature = eigenvalues of Hessian $\nabla_\perp^2 \mathcal{E}[\vec{s}]$ $\left(\sim \frac{\partial^2 \mathcal{E}[\vec{s}]}{\partial s_i \partial s_j}\right)$

index $K_N[\vec{s}] = \{\#$ negative evalues $\nabla_\perp^2 \mathcal{E}[\vec{s}]\}$

(minima: all evalues positive, $K=0$)

(iii) geometry: overlap with signal $q_N[\vec{s}^*] = (\vec{s}^* \cdot \vec{v})/N$

Q1. Properties of global minimum

Q2/Q3. Properties of local minima

Notice: here "typical" means: happening with probability $\mathbb{P} \longrightarrow 1$ when $N \to \infty$.

"rare" means happening with $\mathbb{P} \xrightarrow{N \to \infty} 0$.

▶ We shall see:

Quadratic landscape $\mathcal{E}[\vec{s}]$: Can aswer all the questions when $N \to \infty$,
using Random matrix theory. Describe what happens typically (= with large probability) when $N$ large.
More complicated landscapes: PART II.

▶ Comment: Q1 and Q3 depend on the estimator and on the algorithm chosen. Here we discuss maximum likelihood (with spherical prior), and Langevin dynamics, and derive a recovery and algorithmic threshold for them.
"Information-theoretic threshold" (minimal $r/\sigma$ above which it is information theoretically possible to detect the signal) can be smaller then the recovery threshold predicted by ML.

# I.2 HOW: RANDOM MATRIX THEORY

## From landscapes back to random matrices

Consider a fixed realization of $\hat{M} \to$ of landscape $\mathcal{E}[\vec{s}]$

KOSTERLITZ, THOULESS, JONES 1976

Implement spherical constraint:

$$\mathcal{E}_\lambda[\vec{s}] = -\frac{1}{2}\sum_{i,j=1}^{N} M_{ij}\, s_i s_j + \frac{\lambda}{2}\left(\sum_{i=1}^{N} s_i^2 - N\right)$$

Stationary points $(\vec{s}^*, \lambda^*)$ satisfy:

$$\begin{cases} \dfrac{\partial \mathcal{E}_\lambda[\vec{s}^*]}{\partial s_i} = -\displaystyle\sum_{j=1}^{N} M_{ij}\, s_j^* + \lambda^* s_i^* = 0 & \forall i = 1, .., N \\ \dfrac{\partial \mathcal{E}_\lambda[\vec{s}^*]}{\partial \lambda} = \displaystyle\sum_i [s_i^*]^2 - N = 0 \end{cases}$$

The first equation is evalue equation for $\hat{M}$: $\hat{M}\vec{s}^* = \lambda^* \vec{s}^*$

If $\{\vec{u}_\alpha, \lambda_\alpha\}$ are evectors/evalues of $\hat{M}$ for $\alpha = 1, .., N$, then:

$\vec{s}_\alpha = \pm\sqrt{N}\,\vec{u}_\alpha$ are stationary points of $\mathcal{E}[\vec{s}]$: $2N$ of them!
(notice symmetry bc. quadratic function)

Properties:

(i) Energy. Multiply first equation by $s_i^*$, sum & use second one:

$$\sum_{i,j} s_i^* M_{ij} s_j^* = \lambda^* \cdot N \implies \lambda^* = -\frac{2\mathcal{E}[s^*]}{N} = -2\underbrace{\epsilon[s^*]}_{\text{energy density}}$$

$\implies$ The $\vec{s}_\alpha$ have energy density

$$\epsilon_N[\vec{s}_\alpha] = -\frac{\lambda_\alpha}{2}$$

(ii) Stability. minima, saddles?

Hessian: $\quad \nabla^2 \mathcal{E}_\lambda[s^*] = -M_{ij} + \lambda^*$

At stationary point $\vec{s}^\alpha$: $\nabla^2\mathcal{E}_\lambda[\vec{s}^\alpha] = -(\hat{M} - \lambda_\alpha \hat{1})$

The eigenvalues of $\hat{M}$ are $\lambda_1 \leq \dots \leq \lambda_N$. The eigenvalues of $\nabla^2\mathcal{E}_\lambda[\vec{s}^\alpha]$ are $\underbrace{-(\lambda_1 - \lambda_\alpha)}_{\text{positive if } \alpha > 1}, \underbrace{-(\lambda_2 - \lambda_\alpha)}_{\text{positive if } \alpha > 2} \cdots$

One zero eigenvalue (due to spherical constraint), $(\alpha - 1)$ positive and $N - \alpha$ negative: Stationary points $\vec{s}_\alpha$ are saddles of index $\kappa_N[s_\alpha] = N - \alpha$

Ground state: $\alpha = N$. Global minimum $(\kappa = 0)$

$\Rightarrow$ For each realization of randomness $\hat{J}$, $\mathcal{E}[\vec{s}]$ has $2N$ stationary points; their energy distribution is related to eigenvalue distribution of $\hat{M}$.

Statistical properties when $N \gg 1$ determined by Random Matrix Theory (RMT).

## Notation: gradients & Hessians on sphere

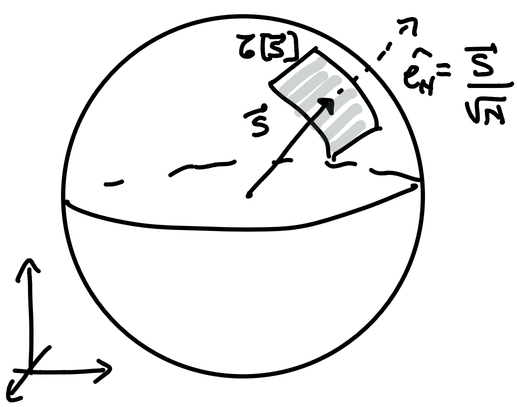
$\nabla \mathcal{E}[s] = \left(\frac{\partial \mathcal{E}}{\partial s_i}\right)_{i=1}^{N}$ gradient in $\mathbb{R}^N$

Lagrange multiplier $\lambda^*$ subtracts the radial component:

$\lambda^* = -\nabla \mathcal{E}[\vec{s}] \cdot \vec{s}$ , $\nabla \mathcal{E}_\lambda[\vec{s}] = \nabla \mathcal{E}[\vec{s}] - \underbrace{\left(\frac{\nabla \mathcal{E}[\vec{s}] \cdot \vec{s}}{N}\right)\vec{s}}_{\text{radial component}}$

Choose basis vectors such that

$\vec{e}_\alpha = \begin{cases} \perp \vec{s} & \alpha = 1, \dots, N-1 \\ \vec{e}_N = \vec{s}/\sqrt{N} & \end{cases}$ ← Spanning tangent plane $\mathcal{T}[\vec{s}]$




In this basis:

$\nabla \mathcal{E}_\lambda[\vec{s}] = \begin{pmatrix} \nabla_\perp \mathcal{E}[\vec{s}] \\ 0 \end{pmatrix}$ (N-1) - dim "gradient on the sphere".

Similarly, Hessian on the sphere $\nabla_\perp^2 \mathcal{E}[\vec{s}]$ is the $(N-1) \times (N-1)$ matrix $\frac{\partial^2 \mathcal{E}[\vec{s}]}{\partial s_i \partial s_j} + \lambda^*[\vec{s}]\hat{1}$ projected on $\mathcal{T}[\vec{s}]$

# Some facts in Random Matrix Theory (RMT)

- The results below hold true for rank-1 perturbed GOE matrices of the type:

$$\hat{M} = \hat{J} + \hat{R} = \hat{J} + r\,\vec{w}\,\vec{w}^T \qquad \left(\vec{w} = \vec{v}/\sqrt{N},\ \|\vec{w}\| = 1\right)$$

$\hat{J}$ = GOE matrix: both a Wigner matrix (real, symmetric, iid entries) & rotationally-invariant ($J \overset{\text{in law}}{\sim} J_R = O J O^T$)
Normalized so that spectrum in bounded interval when $N \to \infty$.

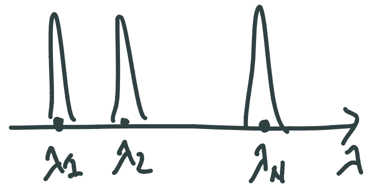
$\hat{R}$ = deterministic, rank-one matrix with 1 evalue equal to $r$, and $(N-1)$ zero eigenvalues. Independent of $\hat{J}$
Perturbation to GOE! "Spike"

- Some results have some degree of universality: can be generalized to other matrix ensambles, or perturbations of higher rank (finite in N)

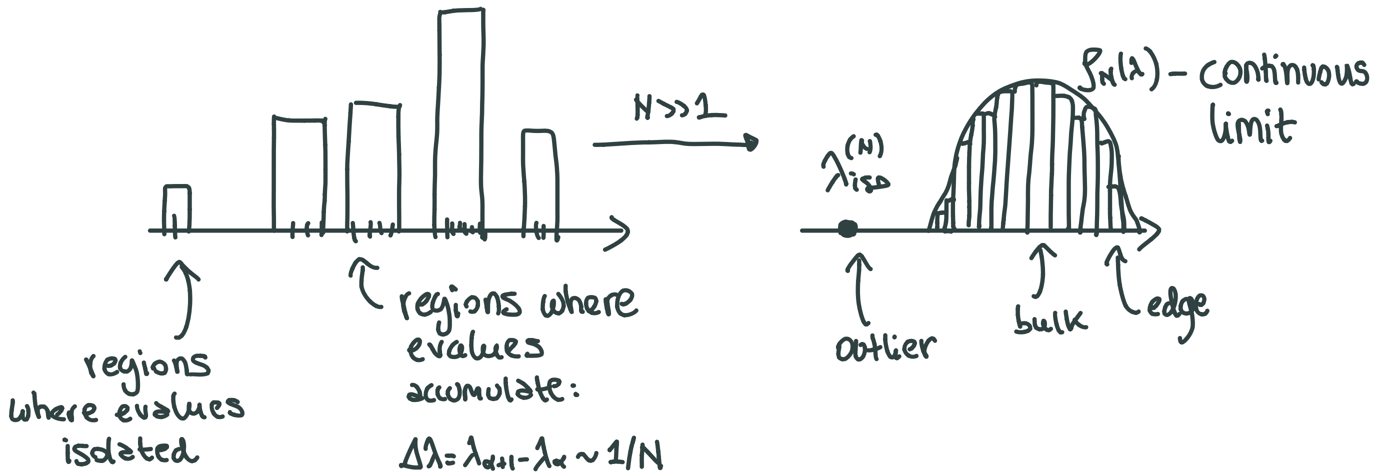
- Eigensystem: $\{\lambda_\alpha, u_\alpha\}_{\alpha=1}^N$. In this section, averages are w.r.t. distribution of $\hat{M}$: $\langle \cdot \rangle = \int d\hat{M}\, P(\hat{M}) \cdot$
Assume $\lambda_1 \leq \ldots \leq \lambda_N$, and $\|u_\alpha\| = 1$.

## ▶ The eigenvalue distribution: density, & outliers

N finite: $\nu_N(\lambda) = \frac{1}{N}\sum_{\alpha=1}^{N}\delta(\lambda-\lambda_\alpha)$




Typical scenario when N increases:




$$\boxed{\nu_N(\lambda) \overset{N\gg1}{\simeq} \underbrace{\rho_N(\lambda)}_{\text{Density}} + \underbrace{\frac{1}{N}\delta\left(\lambda-\lambda_{iso}^{(N)}\right)}_{\text{Outliers}}}$$

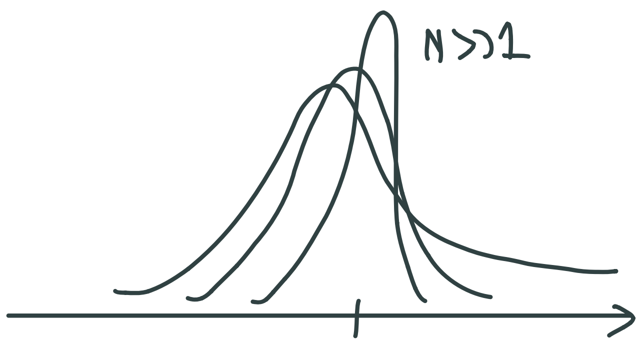
**The density.** Where evalues accumulate, distribution described by continuous density: $\rho_N(\lambda)$.

$$\mathbb{P}\left(\lambda_\alpha \in [\lambda,\lambda+\delta\lambda]\right) \overset{N\gg1}{\simeq} \rho_N(\lambda)\,d\lambda$$

have $O(N)$ evalues around $\lambda$, separated by $\mathcal{O}(1/N)$ in bulk, or $\mathcal{O}(1/N^\alpha)$ at edges.

Facts: (a) Density $\rho_N(\lambda)$ is self-averaging

$$\lim_{N\to\infty} \underbrace{\rho_N(\lambda)}_{\text{random function}} = \underbrace{\rho_\infty(\lambda)}_{\text{deterministic}} = \lim_{N\to\infty} \langle \rho_N(\lambda) \rangle$$




(b) Can be obtained from Stieltjes transform:

$$g_N(z) = \int \frac{d\nu_N(\lambda)}{z-\lambda} = \frac{1}{N} \sum_\alpha \frac{1}{z-\lambda_\alpha} = \frac{1}{N} \operatorname{tr} \underbrace{\left(\frac{1}{z-M}\right)}_{\text{resolvent}}$$

This function is singular when $z \to \lambda_\alpha$ (poles)
Define it away from real axis, e.g. $z \in \mathbb{C}^-$, $z = E - i\eta$
(then: analytically continue).
$\lim_{N\to\infty} g_N(z) = g_\infty(z)$ also self-averaging

(c) When $N \to \infty$, poles accumulate into branch-cut. The discontinuity at the cut is related to $\rho_\infty(\lambda)$:

$$\rho_\infty(\lambda) = \lim_{\eta \downarrow 0} \frac{1}{\pi} \operatorname{Im}\left\{ g_\infty(\lambda - i\eta) \right\}$$

Isolated eigenvalues. Isolated poles of $g_N(z)$, contributing to order $1/N$.

They also "concentrate": $\lim_{N\to\infty} \lambda_{iso}^{(N)} = \lambda_{iso}^{\infty}$

Questions:
- $\rho_\infty(\lambda)$, typical value of $\lambda_{iso}^{(N)}$ when $N \to \infty$?
- typical fluctuations at $N$ large & finite?
- atypical fluctuations: large deviations

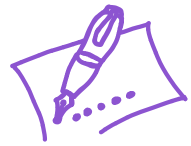
Books:

POTTERS, BOUCHAUD - A first course in random matrix theory, 2021

MEHTA - Random Matrices, 2004

## ▶ Typical values: the density $\rho_\infty(\lambda)$.

Can be studied with REPLICA METHOD → EXERCISE 1



One finds that:

(1) The finite rank perturbation $\hat{R}$ does not affect the density of $\hat{M}$, that is the same as the one of $\hat{J}$.

$$\lim_{N\to\infty} \rho_N(\lambda; r) = \rho_\infty(\lambda; r=0)$$ (effect of rank-1 perturbation disappears for $N\to\infty$)

(2) When $\hat{J}$ is Gaussian, $\langle J_{ij}^2 \rangle = \frac{\sigma^2}{N}(1+\delta_{ij})$, then:

The Stieltjes transform satisfies a self-consistent equation:

$$\boxed{\sigma^2 g_\infty^2(z) - z\, g_\infty(z) + 1 = 0} \qquad z \notin \text{spectrum}$$

(3) This is solved by: $g_{sc}(z) = \dfrac{z - z\sqrt{1 - 4\sigma^2/z}}{2\sigma^2}$ choice of branch?

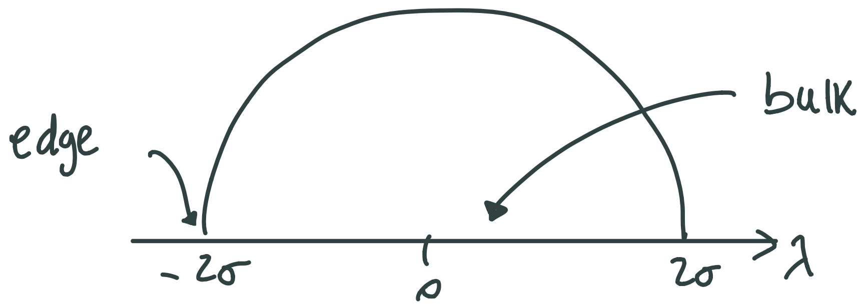
Continuation to real axis: $z \to \lambda$

$$g_{sc}(\lambda) = \frac{\lambda - \text{sign}(\lambda)\sqrt{\lambda^2 - 4\sigma^2}}{2\sigma^2} \qquad \lambda \notin [-2\sigma, 2\sigma]$$

Choise of branch guarantees $\lim_{|\lambda|\to\infty} g_{sc}(\lambda) = 0 \quad \left(g_{sc}(z) \sim \frac{1}{z}\right)$

By inversion formula:

$$\boxed{\rho_\infty(\lambda) = \rho_{sc}(\lambda) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2-\lambda^2}\ \mathbb{1}_{\lambda\in[-2\sigma,2\sigma]}}$$




- Universality of $\rho_{sc}(\lambda)$: it is the limiting density for a large class of matrices of the Wigner type: symmetric, with iid entries not necessarily Gaussian, finite second moment.

  ERDÖS – Universality of Wigner random matrices: a survey of results, 2010

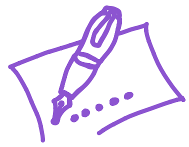
  BENAYCH-GEORGES & KNOWLES – Lectures on the local semicircle law for Wigner matrices, 2019

  Also spectrum of Laplacian of random graphs. (adjacency matrix), Burgers equation....

- $\hat{R}$ can have larger rank, not scaling with N (finite rank)

## ▶ Typical values: the isolated evalue(s) \ evector(s)

The $1/N$ contibutions to $g_N(\lambda)$ can be studied in a large-$N$ expansion $\longrightarrow$ **EXERCISE 2**



One finds that:

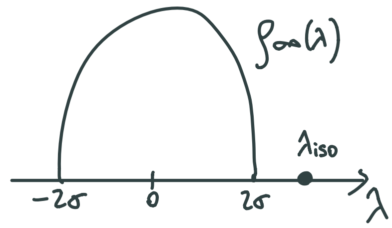
(1) For $\hat{R}=0$, There are no isolated eigenvalues

$$\lim_{N\to\infty} \lambda_1 = -2\sigma \quad \text{minimal eigenvalue} \qquad \text{(almost surely)}$$
$$\lim_{N\to\infty} \lambda_N = 2\sigma \quad \text{maximal eigenvalue}$$

(2) When $N\to\infty$, a transition in maximal evalue when $r=r_c=\sigma$ (notice: smaller than radius $2\sigma$)

$$\boxed{\lim_{N\to\infty} \lambda_N = \begin{cases} 2\sigma & r\leq r_c=\sigma \\ \frac{\sigma^2}{r}+r & r>r_c=\sigma \end{cases} \qquad \text{(almost surely)}}$$

For $r<r_c$, same behavior as for $r=0$: largest evalue sticks to boundary. For $r>r_c$, the largest eigenvalue is isolated:

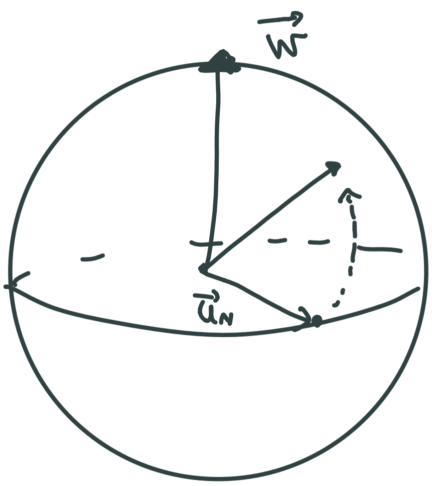



$r > r_c = \sigma$
$(N \to \infty)$

KOSTERLITZ, THOULESS, JONES 1976
PÉCHÉ 2006

(3) The eigenvector $\vec{u}_N$ when $r \geq r_c$ acquires macroscopic projection on $\vec{w} = \vec{v}/\sqrt{N}$

Then:
$$\boxed{\lim_{N\to\infty} (\vec{u}_N \cdot \vec{w})^2 = \begin{cases} 0 & \text{if } r \leq r_c \\ 1-(\sigma/r)^2 & r \geq r_c \end{cases}}$$




While all other eigenvectors such that $(\vec{u}_\alpha \cdot \vec{w})^2 = 0 \quad \alpha \neq N$

This can be seen as a "LOCALIZATION" transition.

- For $r = 0$, consistent with rotational invariance:

eigenvectors of $\hat{J}$ like random vectors on sphere (statistically), and $\vec{w}$ is independent of $\hat{J}$.

As in calculation above,

$$\langle (\vec{u}_\alpha \cdot \vec{w})^2 \rangle = \int \prod_{i=1}^{N} du_\alpha^i \, \delta(\|\vec{u}_\alpha\| - 1)\, (\vec{u}_\alpha \cdot \vec{w})^2 \sim \frac{1}{N} \xrightarrow{N \gg 1} 0$$

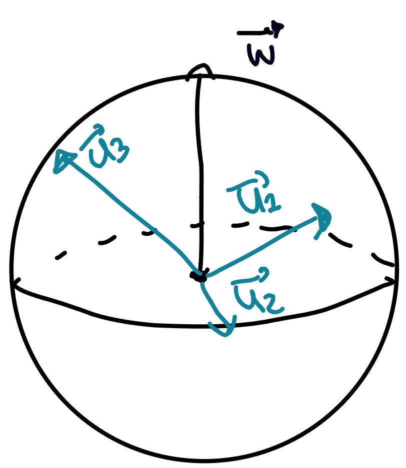
(Use this in Exercise 2): two arbitrary vectors on sphere are typically orthogonal when $N \to \infty$.

$\vec{w}$ is DELOCALIZED in basis $\vec{u}_\alpha$: overlap is of same order of magnitude for all $\alpha$, no special direction.

Terminology from quantum problems, where $\vec{u}_\alpha$ and $\vec{w}$ eigenvectors of local operators (QM is linear).

[CONNECTED NOTIONS: QUANTUM CHAOS, FREE PROBABILITY.]

- When $r > 0$, isotropy broken in direction $\vec{w}$. For $r > r_c$, $\vec{w}$ localized in basis $\vec{u}_\alpha$!




Measure of localization in a basis $\vec{u}_\alpha$: IPR, or HERFINDAHL INDEX:

$$\mathrm{IPR} = \sum_{\alpha=1}^{N} (\vec{w} \cdot \vec{u}_\alpha)^4 \Big/ \sum_{\alpha=1}^{N} (\vec{w} \cdot \vec{u}_\alpha)^2$$

non-zero in localized phase

$$IPR = \begin{cases} \sum_{\alpha=1}^{N} \left(\frac{1}{N}\right)^2 \sim \frac{1}{N} \xrightarrow{N\to\infty} 0 & r \le r_c \\ \sum_{\alpha=1}^{N-1} \left(\frac{1}{N}\right)^2 + \mathcal{O}(1) \xrightarrow{N\to\infty} \mathcal{O}(1) & r > r_c \end{cases}$$

It is also an instance of CONDENSATION (sum over many elements dominated by $\mathcal{O}(1)$ terms) $\rightarrow$ see EXERCISE 3

Generalizations:

- The above is true if $\hat{J}$ is extracted from a rotationally invariant ensemble (not necessarily Gaussian), with density $\rho_\infty(\lambda)$ supported in $[a,b]$. Then one can show that almost surely:

$$\lim_{N\to\infty} \lambda_N = \begin{cases} b & r \le r_c = 1/g_\infty(b) \\ g_\infty^{-1}\left(\frac{1}{r}\right) & r > r_c = 1/g_\infty(b) \end{cases}$$

$$\lim_{N\to\infty} (\vec{w}\cdot\vec{u}_N)^2 = \begin{cases} 0 \quad \text{if} & r \le r_c \\ \frac{-1}{r^2 g'_\infty(\lambda_{iso})} & r \ge r_c \end{cases}$$

PÉCHÉ 2006
BENAYCH-GEORGES & NADAKUDITI 2011

One can recover the GOE expressions from these general ones

Important thing: $\hat{R}$ is independent ("free") of $\hat{J}$.

CAPITAINE, DONATI-MARTIN 2016

- can be generalized to perturbations $\hat{R}$ with rank $n > 1$: $n$ transitions, potentially $n$ isolated eigenvalues. One $r_c$ for each of them.

## Finite-N fluctuations: small deviations

Above results describe $N \to \infty$ limit, when things are self-averaging / concentrate.

At finite N: fluctuations. Things are distributed.

Fluctuations of smallest eigenvalue?

- Transition at $r = r_c$ becomes a crossover.

Critical regime: $\omega = N^{2/3}(r - r_c) \Rightarrow r = r_c + N^{-2/3}\omega$

$$\begin{cases} \text{If } (r_c - r) \gg N^{-2/3} : \text{subcritical} \\ \text{If } (r - r_c) \gg N^{-2/3} : \text{supercritical} \end{cases}$$

BEN AROUS, BAIK, PECHE 2005
PECHÉ 2006
BLOEMENTAL, VIRÁG, 2013

(See example in figure below.)

Scaling $N^{2/3}$ of critical window: BBP 2005 for complex Wishart, but conjectured to be general.

- Subcritical regime $(r>0)$

$$\lambda_N \overset{N\gg 1}{\approx} 2\sigma + N^{-2/3}\sigma\,\xi_{TW}$$

$\xi_{TW}$ = variable with Tracy-Widom distibution $P_{TW}$ $(\beta=1)$

This means

$$\lim_{N\to\infty} P\left(\frac{N^{2/3}(\lambda_N - 2\sigma)}{\sigma}\right) = P_{TW}$$

TRACY, WIDOM 1994

FORRESTER 1993

The gap between eigenvalues at edge is $O(N^{-2/3})$ in subcritical regime

BAIK, LEE 2017

- Supercritical regime:

$$\lambda_N \overset{N\gg 1}{\approx} \lambda_{iso}^{\infty} + N^{-1/2}\sqrt{2\sigma^2\left(1-\frac{\sigma^2}{r^2}\right)}\,\xi_{Gauss}$$

$\xi_{Gauss}$ = random v. with Gaussian distrib.

$$\lim_{N\to\infty} P\left(\frac{N^{1/2}(\lambda_{iso} - \lambda_{iso}^{\infty})}{\sqrt{2\sigma^2(1-\sigma^2/r^2)}}\right) = P_{gauss}$$

## Critical regime ($r = r_c + N^{-2/3}\omega\sigma$)

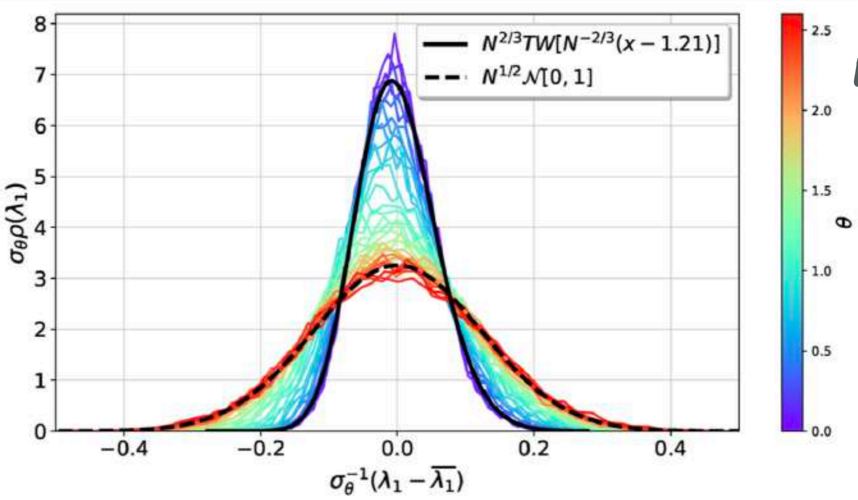
$\lambda_N \overset{N\gg 1}{\approx} 2\sigma + N^{-2/3} X_\omega$ $\quad$ $X_\omega$ random variable with distribution $p_\omega$

Such that: $p_\omega \xrightarrow{\omega\to\infty}$ gauss

$p_\omega \xrightarrow{\omega\to-\infty}$ tracy-widom

BLOEMENTAL, VIRÁG, 2013

**Figure 1.** Scaled probability density distributions of an ensemble of $10^4$ spike random matrices with $N = 100$. The distributions are centered relative to the ensemble average $\overline{\lambda_1}$ and $\sigma_\theta$ stands for the predicted standard deviation when $\theta > 1$. The centered TW distribution $TW$ (15) and the normal distribution $\mathcal{N}[0,1]$ (16) have been scaled similarly to the data.




Crossover in distribution of $\lambda_N$, from Tracy-Widom to gaussian [here $\theta$ denotes $r/\sigma$]

PIMENTA, STARIOLO 2023

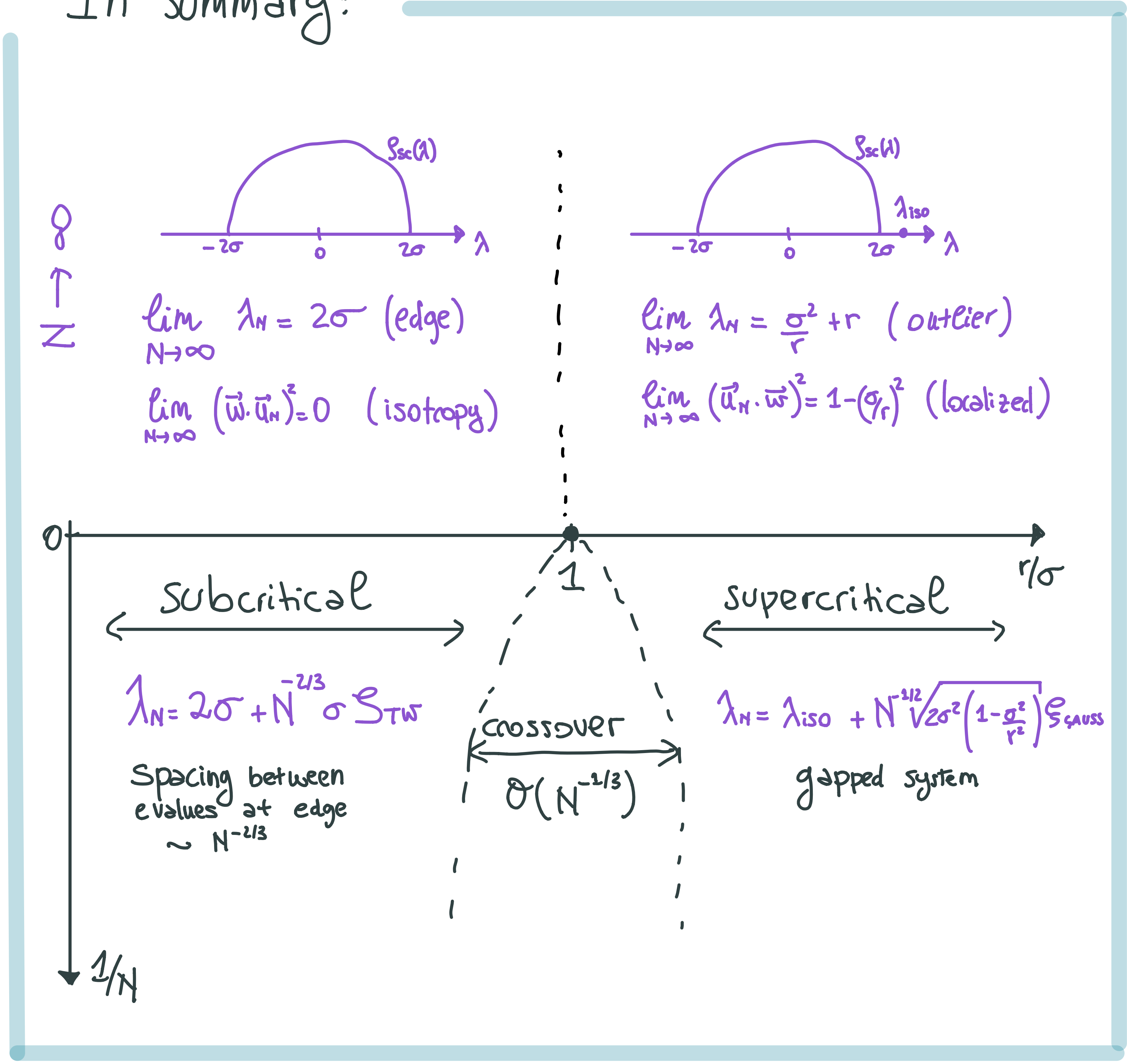
## The Tracy-Widom distribution appears in a huge variety of contexts: Universality.

"KPZ (Kardar, Parisi, Zhang) universality class"

▣ At $r_c$, also a transition on the scaling of the fluctuations of largest evalue, not just on its typical value $\Rightarrow$ "BBP transition".

BAIK, BEN AROUS, PÉCHÉ 2005

In summary:



## Finite N fluctuations: large deviations

Joint evalue-evector projection distribution

$$\mathcal{P}_N\left(\{\lambda_\alpha, \xi_\alpha\}\right) = \frac{e^{-N\sum_\alpha f(\lambda_\alpha, \xi_\alpha)}}{Z_N} \prod_{\alpha=1}^{N} \theta(\lambda_{\alpha+1}-\lambda_\alpha) \underbrace{\prod_{\alpha<\beta} |\lambda_\beta - \lambda_\alpha|}_{\text{Vandermonde}} \times \delta\left(\sum_{\alpha=1}^{N} \xi_\alpha - 1\right) \frac{1}{\sqrt{\xi_\alpha}}$$

where $f(\lambda_\alpha, \xi_\alpha) = \frac{1}{4\sigma^2}\left(\lambda_\alpha^2 - 2r\lambda_\alpha \xi_\alpha\right)$

$$\xi_\alpha = \left(\vec{u}_\alpha \cdot \frac{\vec{v}}{\sqrt{N}}\right)^2 = \left(\vec{u}_\alpha \cdot \vec{w}\right)^2$$

- $r=0$ [spectrum $\hat{\gamma}$]: decoupling of evalues & evectors proj.

The eigenvalues alone distributed as:

$$\mathcal{P}_N\left(\{\mu_1 \leq \cdots \leq \mu_N\}\right) = \frac{N!}{Z_N(\sigma)} \prod_{i=1}^{N} \left(e^{-\frac{N\mu_i^2}{4\sigma^2}} \theta(\mu_{i+1}-\mu_i)\right) \prod_{i<j} |\mu_i - \mu_j|$$

$$Z_N(\sigma) = \sigma^{\frac{N(N+1)}{2}} e^{\frac{3}{2}N\log 2} \left(\frac{2}{N}\right)^{\frac{N(N+1)}{4}} \prod_{i=1}^{N} \Gamma\left(1+\frac{i}{2}\right)$$

The eigenvectors have statistics of random unit vectors: setting $q_\alpha = \sqrt{\xi_\alpha}$, then:

$$\mathcal{P}_N\left(\{q_\alpha\}_{\alpha=1}^N\right) = C_N\, \delta\left(\sum_{\alpha=1}^N q_\alpha^2 - 1\right)$$

rotational invariance: evectors of $\hat{J}$ and $\hat{J}_R$ are equally probable

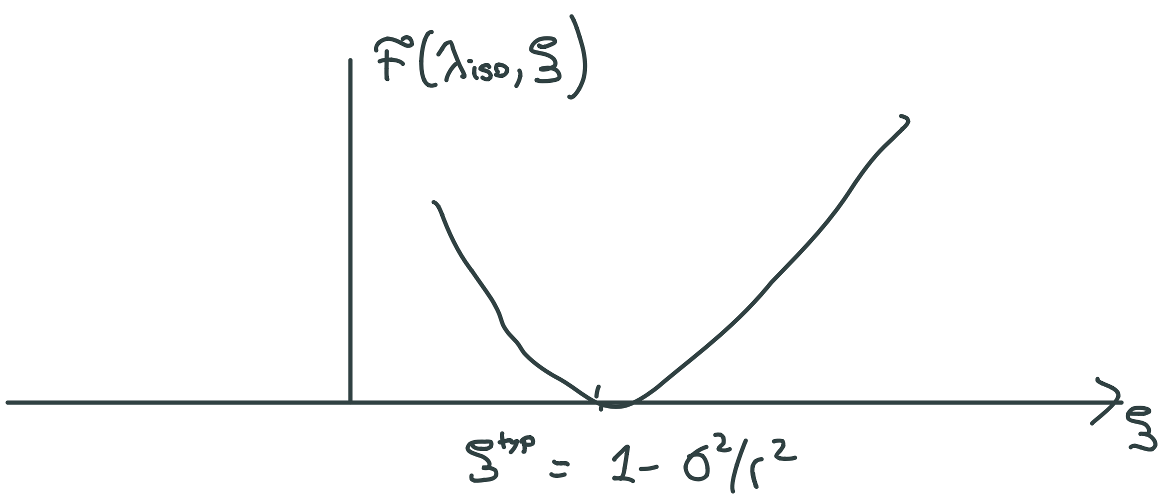
- For $r>0$: coupling of evalues & evector projection! This coupling can "pull" some eigenvector (the extremal) towards $\vec{w}$ when $r>r_c$.

- From $\mathcal{P}_N(\{\lambda_\alpha, \xi_\alpha\})$, can get the joint large deviation probability of $\lambda_N, \xi_N$ – maximal eigenvalue/vector

$$P_{LD}(\lambda_N, \xi_N) \sim e^{-N F(\lambda_N, \xi_N)}$$

BIROLI, GUIONNET 2019

probability of $\mathcal{O}(1)$ deviations from typical, asyptotic $N\to\infty$ value.



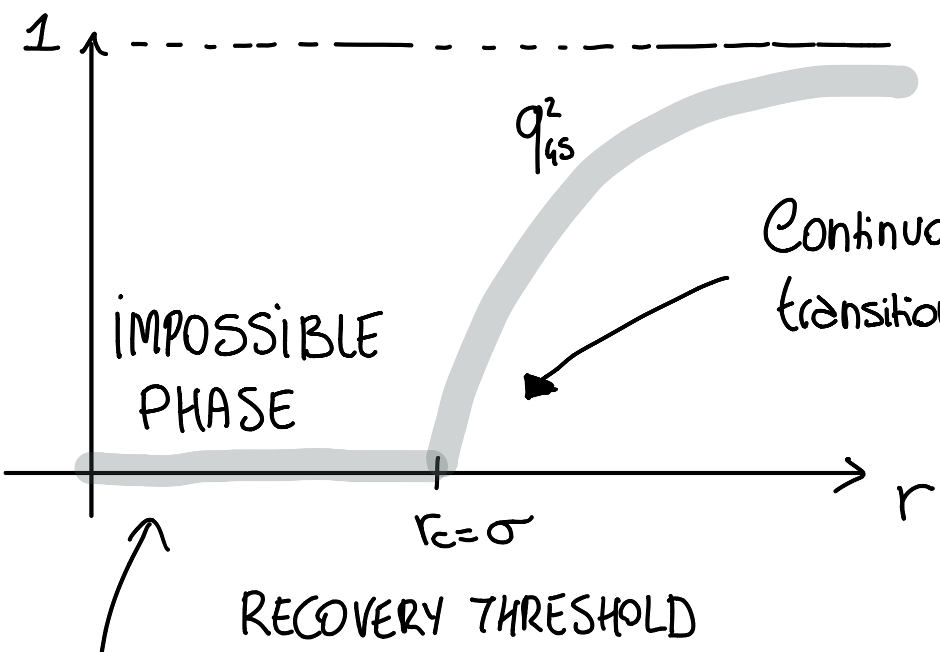
# I.3 WHAT: GS, LANDSCAPE, DYNAMICS

Back to the inference problem...

## Q1: Recovering the signal

Q1: when is $\vec{S}_{GS}$ informative, i.e. $q_{GS} > 0$?

A sharp transition when $N \to \infty$: informative for $r > r_c$




here, even if I am able to find $\vec{S}_{GS}$, I would get no info on $\vec{v}$ because $\vec{S}_{GS}$ is uncorrelated to it.

Comments:

- The transition in the ground state could be found also from thermodynamics, studying the $\beta\to\infty$ limit of:

$$Z_\beta = \int_{S_N(\sqrt{N})} d\vec{s}\, e^{-\beta \mathcal{E}[\vec{s}]} = \int d\vec{s}\, d\lambda\, e^{-\beta \mathcal{E}[s] - \beta\frac{\lambda}{2}\left(\sum_{i=1}^{N} s_i^2 - N\right)}$$

Thermodynamically, the zero-temperature transition at $r = r_c = \sigma$ is a transition between a spin-glass phase at $r < r_c$, and a ferromagnetic phase at $r > r_c$.

At $T > 0$: phenomenology of condensation $\Rightarrow$ EXERCISE 3!

KOSTERLITZ, THOULESS, JONES 1976

CUGLIANDOLO LECTURE NOTES CARGESE 2020

- Critical threshold for maximum likelihood is also "detection threshold" when $\vec{v}$ has gaussian or rademacher prior: below $r_c$, no estimator distinguishes between pure noise (GOE) and spiked matrices.

PERRY, WEIN, BANDEIRA, MOITRA 2018

# Q2: A landscape of saddles

- Stationary points above ground state.

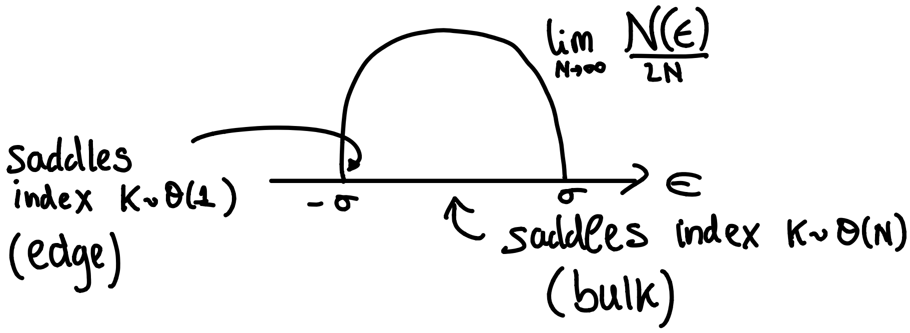
  $\mathcal{N}_N(\epsilon) = \#$ stationary points with $\mathcal{E}_N[s^*] = \epsilon$

  is a self-averaging random variable such that:

$$\boxed{\lim_{N\to\infty} \frac{\mathcal{N}_N(\epsilon)}{2N} = \lim_{N\to\infty} \frac{\langle \mathcal{N}_N(\epsilon)\rangle}{2N} = \rho_{sc}(2\epsilon)}$$

GOE density

- All stationary points (except GS) are saddles with negative directions of curvature: most have index $k \sim \mathcal{O}(N)$: NO trapping local minima!




- All these saddles have $q_N[s_\alpha] = \left(\frac{s_\alpha \cdot w}{\sqrt{N}}\right) \xrightarrow{N\to\infty} 0$. Can study scaled overlap $\Phi = N \cdot \mathbb{E}\left[\left(\frac{s_\alpha \cdot v}{N}\right)^2\right]$

  BUN, BOUCHAUD, POTTERS 2018

- Expect optimization not to be "hard" ($z \not\sim e^N$)

# Q3: Dynamics: DMFT, & beyond

Consider simplest algorithm: gradient descent.
(Langevin with $T \to 0$)

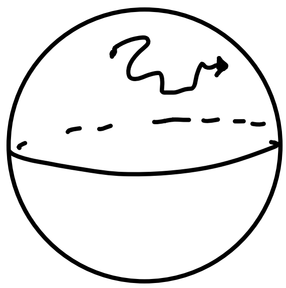
$$\frac{dS_i(t)}{dt} = -\sum_j J_{ij} S_j(t) - \lambda(t) S_i(t) + \sqrt{2T}\,\eta_i(t)$$

need to enforce to stay on sphere at each time

Gaussian white noise

$$\langle \eta_i(t)\eta_j(t')\rangle = \delta_{ij}\,\delta(t-t')$$

Mimicks coupling to degrees of freedom equilibrated at T



When $T \to 0$ (no noise), expect convergence to $T=0$ equilibrium state, the ground state $\vec{S}_{GS} = \pm\sqrt{N}\,\vec{u}_N$, when $t \to \infty$; how to take $N \to \infty$? Relevant timescales?

## Large-time and large-N limit: how?

(1) Mean-field dynamics: take $N \to \infty$ before, then $t \to \infty$.

Fully-connected models with randomness can be described by DMFT ('Dynamical Mean Field Theory')

Why? ▶ Dynamics becomes self-averaging when $N \to \infty$: properties of trajectories for different realizations of $\mathcal{E}[s]$ become deterministic, converge to average value [can replace average over $\mathcal{E}$ with $N \to \infty$]

eg. energy density: $\lim_{N\to\infty} \frac{\mathcal{E}[s(t)]}{N} = \lim_{N\to\infty} \left\langle \frac{\mathcal{E}[s(t)]}{N} \right\rangle$

▶ These properties are one and two-point functions in time, for which have closed eqs, "DMFT equations"

$\mathcal{E}(t) \leftarrow$ time-dependent energy

$C(t,t') = \frac{1}{N} \sum_{i=1}^{N} s_i(t) s_i(t') \longleftarrow$ correlation function

$R(t,t') = \frac{1}{N} \sum_{i=1}^{N} \left. \frac{\delta s_i(t)}{\delta \varepsilon_i(t')} \right|_{\varepsilon=0} \longleftarrow$ response function

$\Rightarrow$ Used in many contexts: CUGLIANDOLO 2023 (Annual review of condensed matter physics)

(2) Beyond mean-field: dynamics for N large but finite.

DIFFICULT PROBLEM!

Often, fluctuations matter, no self-averagingness

Quantities are distributed.

Averages & typical values are different.

$\Longrightarrow$ This model (for $r=0$) is a rare case in which dynamics can be studied in both regimes, using Random Matrix Theory.

- $T=0$. In the eigenbasis $s_\alpha = (\vec{s} \cdot \vec{u}_\alpha)$

$$\frac{ds_\alpha}{dt}(t) = -\left[\lambda_\alpha + \lambda(t)\right] s_\alpha(t)$$

couples all different $\alpha$.
makes the equations non-linear.

- $T=0$, dynamics should converge to $\vec{s}_{gs} = \pm\sqrt{N}\,\vec{u}_N$.
Study convergence by excess energy:

$$\Delta\epsilon_N(t) = \left(\frac{\mathcal{E}(t)}{N} - \epsilon_{gs}\right) = \frac{1}{2}\,\frac{\sum_{\alpha\neq N}(\lambda_N-\lambda_\alpha)\,e^{-2(\lambda_N-\lambda_\alpha)t}}{1+\sum_{\alpha\neq N} e^{-2(\lambda_N-\lambda_\alpha)t}} = f(\{\lambda_N-\lambda_\alpha\})$$

for random initial conditions $s_\alpha(t=0)=1 \quad \forall\alpha$.

## Short times, large times, dynamical crossovers

$$\Delta E_N(t) \overset{t\gg 1}{\approx} g_N \, e^{-2t g_N} \qquad g_N = \lambda_N - \lambda_{N-1} \ \text{`gap'}$$

Natural time where "probe" finite-N, energy scales where discreteness of spectrum matters:

$$\tau_{dync} \sim 1/g_N$$

such that:
$$\begin{cases} t \ll \tau_{dync} : & \text{dynamics looks as if } N\to\infty \text{ (DMFT-like)} \\ t \gg \tau_{dync} : & \text{finite-N dynamics} \end{cases}$$

The fluctuations of the gap $g_N$ are of the same order as those of maximal eigenvalue. Recall RMT detour:

$$1/g_N \sim \begin{cases} \mathcal{O}(N^{2/3}) & \text{subcritical regime } r \lesssim r_c \text{ (Tracy-Widom)} \\ ? & \text{critical regime } |r-r_c| \sim \mathcal{O}(N^{-1/3}) \\ \mathcal{O}(N^0) & \text{supercritical regime: } r \gtrsim r_c \text{ system is GAPPED!} \end{cases}$$

$\mathcal{O}(N^0)$ is more precisely $\sim \log N$ D'ASCOLI, REFINETTI, BIROLI 2022

# The mean-field dynamics: $N \to \infty$

One finds in this time regime:

$$\lim_{N\to\infty} \langle \Delta\epsilon_N(t) \rangle = \upsilon_{MF}(t) \overset{t\gg 1}{\approx} \begin{cases} \frac{3\sigma}{8t} & r \le r_c \\ \frac{3\sigma}{8t} + \left[-\sigma - \epsilon_{gs}\right] & r \ge r_c \end{cases}$$

Slow, algebraic decay to the energy density of the ground state ($\epsilon_{gs} = -\sigma$) when $r \le r_c$, and to the same energy (which is no longer the ground state) when $r > r_c$.

(1) Dynamics is always out-of-equilibrium in this regime. It is, in fact, glassy:

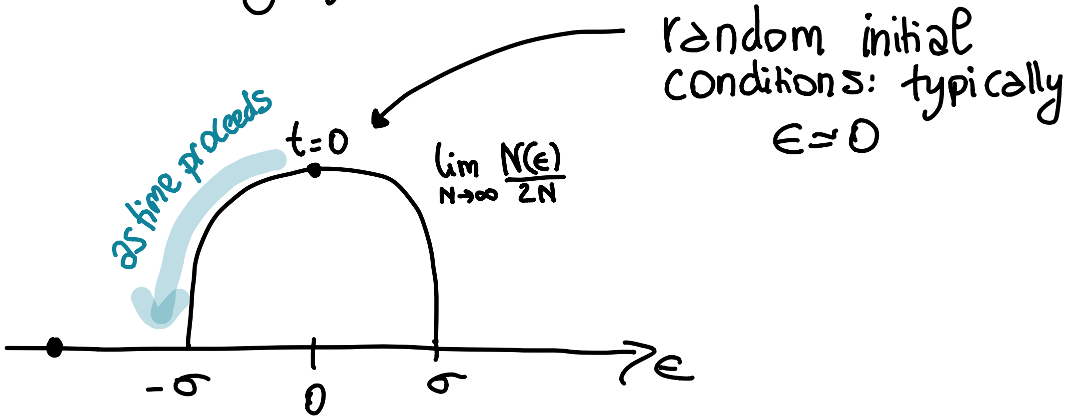
- $C(t,t') \neq c(t-t')$, modified FDT
- separation of timescales in $t-t'$
- aging & weak ergodicity breaking

"Aging": dynamics slower and slower as system becomes older (i.e., as time proceeds)

CUGLIANDOLO & DEAN 1995

BEN AROUS, DEMBO, GUIONNET 2001 (math)

(2) Landscape interpretation: in these timescales,
Probe landscape at extensive energies above $\epsilon_{gs}$, $\Delta\epsilon_N(t) \sim \Theta(1)$.
Region of landscape dominated by saddles, with density described by $\rho_{sc}(-2\epsilon)$:




Why slowing down? no trapping by local minima (there are not!), but slow decay due to decreasing number of negative directions of saddles (decreasing index).

DMFT, $N\to\infty$ dynamics probes the bulk on $\rho_{sc}(2\epsilon)$.

## The finite-N dynamics, $N \gg 1$

▶ The subcritical regime ($r \leq r_c$): dynamics as for $r=0$.
Crossover time $\tau_{dyn} \sim N^{2/3}$.

$$\langle \Delta\epsilon_N(t) \rangle \sim \begin{cases} v_{MF}(t) & t \ll N^{2/3} = \tau_{dyn} \\ N^{-2/3}\, v_{NMF}(t\, N^{-2/3}) & t \gg N^{2/3} = \tau_{dyn} \end{cases}$$

For $t \ll N^{2/3}$, the system explores extensive energies above $\epsilon_{gs}$.
Dynamics is self-averaging, captured by mean-field (DMFT)

For $t \gg N^{2/3}$, system explore intensive energies above $\epsilon_{gs}$.
Dynamics not self-averaging, not captured by mean-field.

Consider $t \gg N^{2/3}$

- System explores intensive energies on top of $\epsilon_{gs}$: Sensitive to statistics of extreme values and gaps $g_N$.
- Dynamics not self-averaging: $\langle \Delta\epsilon_N(t) \rangle$ dominated by realization where gap atypically small.
- The distribution of $g_N$ is known! PERRET, SCHEHR 2015

$$\begin{cases} P(N^{2/3} g_N) \sim b\, N^{2/3} g_N & N^{2/3} g_N \to 0 \quad \text{(small gaps)} \\ P(N^{2/3} g_N) \sim e^{-2/3 (N^{2/3} g_N)^{3/2}} & N^{2/3} g_N \to \infty \quad \text{(large gaps)} \end{cases}$$

$$\Longrightarrow \langle \Delta_N \epsilon(t) \rangle \sim N^{-2/3} f_0(t N^{-2/3}) \qquad f_0(x) \sim \begin{cases} \dfrac{3\sigma}{8x} & x \to 0 \\ \dfrac{a\sigma}{x^3} & x \to \infty \end{cases}$$

scaling function known!

FYODOROV, PERRET, SCHEHR 2015
BARBIER, PIMENTA, CUGLIANDOLO, STARIOLO 2021

▶ The supercritical regime: in this case system is gapped: for $t \gg \tau_{dyn} \sim \log N$, the system is able to reach the vicinity of $S_{gs}$ and to relax to it exponentially (as in ferromagnetic systems):

$$\langle \Delta\epsilon_N(t)\rangle \sim \begin{cases} v_{MF}(t) & t \ll \log N = \tau_{dyn} \\ \frac{C_r}{t^{3/2}}\, e^{-2t\left|\frac{\sigma^2}{r} + r - 2\sigma\right|} & t \gg \log N = \tau_{dyn} \end{cases}$$

$\uparrow$ gap

▶ The critical regime $(|r - r_c| \sim \mathcal{O}(N^{-2/3}))$: open problem!

$$\langle \Delta\epsilon_N(t)\rangle \overset{t \gg 1}{\sim} \int_0^\infty dg_N\, p(g_N)\, g_N\, e^{-2g_N t}$$

$\lambda_N$ and $\lambda_{N-1}$ strongly corelated, distribution $p(g_N)$ unknown.

From numerics, $p(g_N) \overset{g \ll 1}{\sim} g^{a(r,N)}$ PIMENTA, STARIOLO 2023

giving:
$$\langle \Delta\epsilon_N(t)\rangle \overset{t \gg 1}{\sim} \begin{cases} v_{MF}(t) \sim e^{-2t|\lambda_{iso}^\infty - 2\sigma|} & t \ll N^{2/3} \\ N^{2/3} f_a(t N^{-2/3}) & t \gg N^{2/3} \end{cases}$$

# PART II

## Rugged high-D landscapes

# II.1 WHY: A HIGH-D INFERENCE EXAMPLE

## A 'hard' inference problem: noisy tensors

Beyond matices? Tensors! MONTANARI, RICHARD 2014

$$M_{i_1 i_2 i_3 \dots i_p} = \frac{r}{N^{p-1}} v_{i_1} \dots v_{i_p} + J_{i_1 \dots i_p} \qquad (p>2)$$

$J_{i_1 \dots i_p}$ symmetric, iid gaussian $\langle J^2_{i_1 \dots i_p} \rangle = \frac{p!\, \tilde{\sigma}^2}{N^{p-1}}$

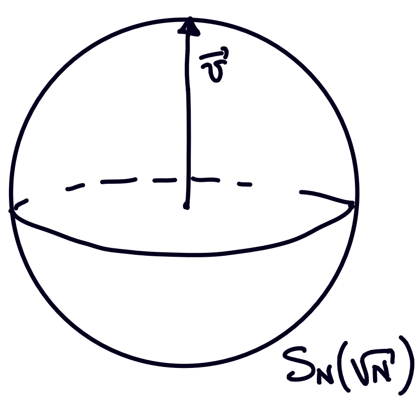
Energy landscape:

$$\mathcal{E}[\vec{s}] = - \sum_{i_1 \le i_2 \dots \le i_p} J_{i_1 \dots i_p} s_{i_1} \dots s_{i_p} - r N \left( \frac{\vec{v} \cdot \vec{s}}{N} \right)^p$$

Again, fully-connected random interactions.

$$\langle \mathcal{E}[\vec{s}] \rangle = -r N \left( \frac{\vec{s} \cdot \vec{v}}{N} \right)^p$$

$$\langle \mathcal{E}[\vec{s}]\, \mathcal{E}[\vec{s}'] \rangle_c = \tilde{\sigma}^2 N \left( \frac{\vec{s} \cdot \vec{s}'}{N} \right)^p$$




Here: no spectrum. Also, landscape at $r=0$ much different...

# Landscape problem & complexity

Same questions as above, same approach: study stationary points.

$$\mathcal{E}_\lambda[\vec{s}] = -\sum_{i_1\le i_2 \dots \le i_p} M_{i_1\dots i_p} s_{i_1}\cdots s_{i_p} + \frac{\lambda}{2}\left(\sum_{i=1}^{N} s_i^2 - N\right)$$

$$\begin{cases} \dfrac{\partial \mathcal{E}_\lambda[\vec{s}^*]}{\partial s_i} = -\displaystyle\sum_{i_2\le \dots \le i_p} M_{i i_2\dots i_p} s^*_{i_2}\cdots s^*_{i_p} + \lambda^* s^*_i \\ \dfrac{\partial \mathcal{E}_\lambda[\vec{s}^*]}{\partial \lambda} = \displaystyle\sum_{i=1}^{N} (s_i^*)^2 - N = 0 \end{cases}$$

As before, multiply first equation by $s_i$, sum & use second equation:

$$\lambda^* = -\frac{1}{N}\left(\sum_i \frac{\partial \mathcal{E}[\vec{s}^*]}{\partial s_i}\cdot s_i\right) = -p\,\frac{\mathcal{E}[\vec{s}^*]}{N} = -p\,\epsilon[\vec{s}^*]$$

However, first equation non-linear: how many solutions? Introduce the random variable

$$\boxed{\mathcal{N}_N(\epsilon, q) = \#\text{ stationary points } \vec{s}^* \text{ with } \epsilon_N[\vec{s}^*] = \epsilon \text{ and } q_N[\vec{s}^*] = \frac{\vec{s}\cdot\vec{v}}{N} = q.}$$

■ Quadratic landscape ($p=2$):

$\mathcal{N}_N(\epsilon)$ is $\mathcal{O}(N)$ when $N \gg 1$

Self-averaging: $\lim_{N\to\infty} \frac{\mathcal{N}_N(\epsilon)}{2N} = \lim_{N\to\infty} \frac{1}{2N} \langle \mathcal{N}_N(\epsilon) \rangle = \rho_{sc}(-2\epsilon)$

■ Landscape for $p>2$:

$\mathcal{N}_N(\epsilon, q)$ is $\mathcal{O}(e^N)$: $\mathcal{N}_N(\epsilon,q) \sim e^{N\Sigma_N(\epsilon,q)}$

$\mathcal{N}_N(\epsilon,q)$ not self-averaging

but $\Sigma_N(\epsilon,q)$ is:

$$\lim_{N\to\infty} \Sigma_N(\epsilon,q) = \lim_{N\to\infty} \langle \Sigma_N(\epsilon,q) \rangle = \Sigma_\infty(\epsilon,q)$$

$$\boxed{\Rightarrow \Sigma_\infty(\epsilon,q) = \lim_{N\to\infty} \frac{1}{N} \langle \log \mathcal{N}_N(\epsilon,q) \rangle \quad \text{"COMPLEXITY"}}$$

## ■ Averages vs typical values, and replicas

Means that typically when $N\to\infty$ (with probability $\to 1$):

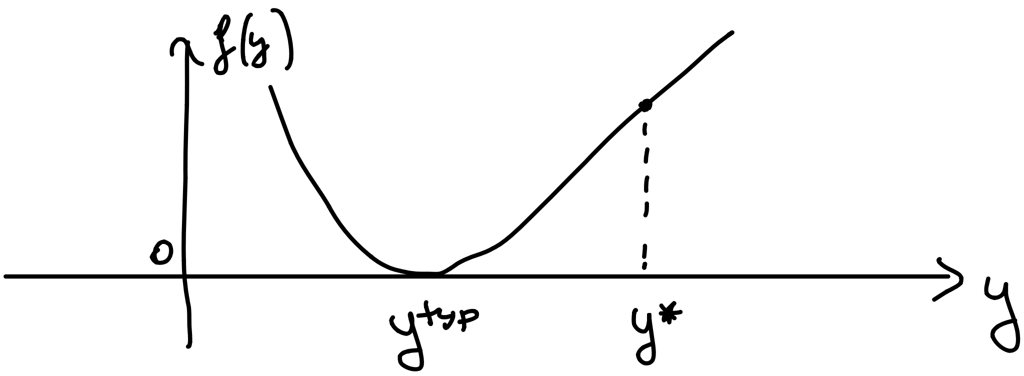
$[\mathcal{N}(\epsilon,q)]_{typ} \sim e^{N\Sigma_\infty(\epsilon,q)}$ (most probable value of $\mathcal{N}$)

But most-probable value is different from the average value: $\langle \mathcal{N}(\epsilon,q) \rangle \not\sim e^{N\Sigma_\infty(\epsilon,q)}$

## Average vs typical values: example.

Assume $X_N$ is a random variable scaling as $X_N \sim e^N$: means that $Y_N = \frac{\log X_N}{N}$ has a limiting distibution when $N \to \infty$.

Assume that when $N \gg 1$, distribution of $Y_N$ takes large-deviation form: $P_{Y_N}(y) \sim e^{-N f(y) + o(N)}$.




Then, typical value of $X_N$ is:

$$[X_N]^{typ} \sim e^{N y^{typ}} \quad \text{where } y^{typ} \text{ such that } f'(y^{typ}) = 0 = f(y^{typ}).$$

On the other hand:

$$\langle X_N \rangle = \int dy\, P_{Y_N}(y)\, e^{Ny} = \int dy\, e^{N[y - f(y)] + o(N)} \simeq e^{N[y^* - f(y^*)]}$$

and $y^*$ such that $f'(y^*) = 1$. Saddle point approximation

Since $y^* \neq y^{typ}$, $f(y^*) > 0$: $y^*$ is exponentially rare, but controls the average: average "dominated" by rare realizations of random variable!

Message: to characterize what happens typically (with large probability) when $N \gg 1$ need:

"QUENCHED CALCULATION" $\qquad \Sigma_\infty(\epsilon, q) = \lim_{N\to\infty} \frac{1}{N} \langle \log \mathcal{N}_N(\epsilon, q) \rangle$

But this is hard; requires tricks like REPLICAS:

$$\langle \log \mathcal{N} \rangle = \lim_{\omega \to 0} \frac{\langle \mathcal{N}^\omega \rangle - 1}{\omega}$$

$\omega$-th moment of $\mathcal{N}$

analytic continuation

In the following, we perform instead:

"ANNEALED APPROXIMATION" $\qquad \Sigma_A(\epsilon, q) = \lim_{N\to\infty} \frac{1}{N} \log \langle \mathcal{N}(\epsilon, q) \rangle$

It holds $\Sigma_A(\epsilon, q) \geq \Sigma_\infty(\epsilon, q) \Rightarrow \langle \mathcal{N}_N \rangle \gg [\mathcal{N}_N]_{typ}$

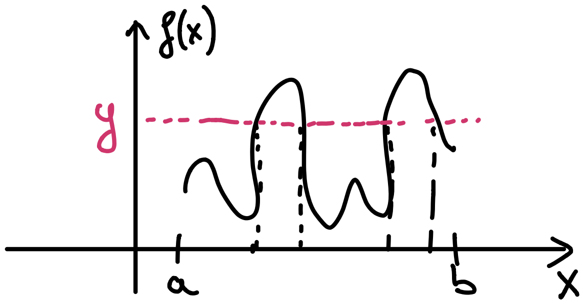
For the quenched calculation of the complexity in this model: ROS, BEN AROUS, BIROLI, CAMMAROTA 2018

# II 2. HOW: KAC-RICE FORMALISM

## Kac-Rice formula(s)

Kac-Rice formula = formula for average (or higher moments) of number of solutions of random equations.

### Counting formulas: example.




$f(x)$ random function in $[a,b]$.
How many solutions of $f(x)=y$ ?

$$N(y) = \int_a^b dx\, \delta\left(x - f^{-1}(y)\right) = \int_a^b dx\, \frac{1}{\left|\frac{d}{dy} f^{-1}(y)\right|}\, \delta\left(y - f(x)\right)$$

$$= \int_a^b dx\, |f'(x)|\, \delta\left(y - f(x)\right)$$

$|f'(x)| =$ Jacobian

In higher dimension: $\vec{x} \in \mathcal{I} \subset \mathbb{R}^d$, $f(\vec{x}) = \vec{y} \in \mathbb{R}^d$

$$N(\vec{y}) = \int_{\mathcal{I}} d\vec{x} \prod_{i=1}^{d} \delta\left(f_i(\vec{x}) - y_i\right) \left|\det\left(\frac{\partial f_i(\vec{x})}{\partial x_j}\right)_{ij}\right|$$

## ▶ Kac-Rice formula: stationary points of landscapes

Count solutions of $\vec{\nabla}_\perp \mathcal{E}[\vec{s}]=0$, $\mathcal{E}[\vec{s}]=N\epsilon$, $\vec{s}\cdot\vec{v}=Nq$

Then:

$$\mathcal{N}(\epsilon,q)=\int_{S_N(\sqrt{N})} d\vec{s}\,\left|\det \nabla_\perp^2 \mathcal{E}[\vec{s}]\right|\,\delta(\vec{\nabla}_\perp \mathcal{E}[\vec{s}])\,\delta(\mathcal{E}[\vec{s}]-N\epsilon)\,\delta(\vec{s}\cdot\vec{v}-Nq)$$

Take average $\Rightarrow$ Kac-Rice formula.

$$\langle \mathcal{N}(\epsilon,q)\rangle=\int_{S_N(\sqrt{N})} d\vec{s}\,\delta(\vec{s}\cdot\vec{v}-Nq)\,\underbrace{\left\langle\left|\det \nabla_\perp^2 \mathcal{E}[\vec{s}]\right|\right\rangle_{\substack{\nabla_\perp\mathcal{E}=0\\ \mathcal{E}=N\epsilon}}}_{\text{average conditioned to } \nabla_\perp \mathcal{E}[\vec{s}]=0 \text{ and } \mathcal{E}[\vec{s}]=N\epsilon}\;\underbrace{P_{\nabla_\perp\mathcal{E},\mathcal{E}}(\vec{0},N\epsilon)}_{\text{joint density of } (\nabla_\perp\mathcal{E},\mathcal{E}) \text{ evaluated at } (\vec{0},N\epsilon)}$$

BRAY, MOORE 1980

CAVAGNA, GIARDINA, PARISI 1998

FYODOROV 2013

BEN AROUS, AUFFINGER, CERNY 2010 (math)

## Computing the complexity: 3 steps

The calculation is done in 3 steps, & uses 3 main ingredients:

(1) GAUSSIANITY

The functions $\mathcal{E}[\vec{s}]$, $\frac{\partial \mathcal{E}}{\partial s_i}[\vec{s}]$, $\frac{\partial^2 \mathcal{E}}{\partial s_i \partial s_j}[\vec{s}]$ are Gaussian: to get distribution, need only averages & covariances.

Can be computed explicitly (TRY! see below for hints)

Doing so, one finds:

(F1) $\nabla_\perp \mathcal{E}[\vec{s}]$ independent of $\mathcal{E}[\vec{s}]$ and $\nabla_\perp^2 \mathcal{E}[\vec{s}]$.

Consequences:

▶ $$P_{\nabla_\perp \mathcal{E}[\vec{s}], \mathcal{E}[\vec{s}]}(\vec{0}, N\epsilon) = P_{\nabla_\perp \mathcal{E}[\vec{s}]}(\vec{0})\, P_{\mathcal{E}[\vec{s}]}(N\epsilon)$$

factorization: two gaussians, known explicitly.

▶ $$\left\langle \left| \det \nabla_\perp^2 \mathcal{E}[\vec{s}] \right| \right\rangle_{\substack{\nabla_\perp \mathcal{E}[\vec{s}]=0 \\ \mathcal{E}=N\epsilon}} = \left\langle \left| \det \nabla_\perp^2 \mathcal{E}[\vec{s}] \right| \right\rangle_{\mathcal{E}=N\epsilon}$$

Statistics of Hessian at stationary point is same as at any point of same energy.

(F2) The $(N-1)\times(N-1)$ matrix $\nabla_\perp^2 \mathcal{E}$ conditioned to $\mathcal{E}=N\epsilon$ has the same statistics as matrices:

$$\hat{M}[\vec{s}] = \hat{J} - p\epsilon \hat{1} - r_{eff}\left[q_N[\vec{s}]\right] \vec{w}_\perp \vec{w}_\perp^T$$

spherical constraint

rank-1 perturbation

Vector $\vec{w}$ is projection of $\vec{v}$ on tangent plane $\tau[\vec{s}]$

where $\hat{J}$ is a GOE: $\langle J_{ij}\rangle = 0$, $\langle J_{ij}^2\rangle = \frac{p(p-1)\tilde{\sigma}^2}{N}(1+\delta_{ij})$

$r_{eff}(q) = r\,p(p-1)\,q^{p-2}(1-q^2)$, $\|w_\perp\|^2 = 1$.

## (2) ISOTROPY

There is only one special direction in the sphere, that is $\vec{v}$. All averages & convariances, and so the joint distribution of $\mathcal{E}[\vec{s}]$, $\nabla_\perp \mathcal{E}[\vec{s}]$, $\nabla_\perp^2 \mathcal{E}[\vec{s}]$ depend on $\vec{s}$ only via $q_N[\vec{s}] = \left(\frac{\vec{s}\cdot\vec{v}}{N}\right) \longrightarrow$ (see above!)

Consequences: for all $\vec{s}$ such that $q_N[\vec{s}] = q$

$$P_{\nabla_\perp \mathcal{E}[\vec{s}]}(\vec{0}) \rightarrow P_1(q) = (2\pi p\,\tilde{\sigma}^2)^{-\left(\frac{N-1}{2}\right)} e^{-\frac{N}{2\tilde{\sigma}^2} p r^2 q^{2p-2}(1-q^2)}$$

$$P_{\mathcal{E}}(N\epsilon) \rightarrow P_2(\epsilon, q) = \sqrt{\frac{N}{2\pi\tilde{\sigma}^2}}\, e^{-\frac{N}{2\sigma^2}(\epsilon + r q^p)^2}$$

And $\left\langle \left|\det \nabla_\perp^2 \mathcal{E}[\vec{s}]\right| \right\rangle_{\mathcal{E}[\vec{s}]=N\epsilon} := \mathcal{D}_N(\epsilon, q)$

Therefore:

$$\langle \mathcal{N}(\epsilon, q)\rangle = \int_{S_N(\sqrt{N})} d\vec{s}\, \delta(\bar{s}\cdot\bar{v} - Nq) \left\langle \left|\det \nabla_\perp^2 \mathcal{E}[\vec{s}]\right| \right\rangle_{\mathcal{E}=N\epsilon} P_{\nabla_\perp \mathcal{E}}(\vec{0})\, p_{\mathcal{E}}(N\epsilon)$$

$$= \mathcal{D}_N(\epsilon, q)\, p_1(q)\, p_2(\epsilon, q)\, V_N(q)$$

where $V_N(q) = \int_{S_N(\sqrt{N})} d\vec{s}\ \delta(Nq - \bar{s}\cdot\bar{v})$ Volume of the sub-sphere

Can show that $V_N(q) \overset{N\gg 1}{\sim} e^{N/2 \log[2\pi e(1-q^2)] + o(N)}$

Example: 2d rotationally-invariant function

$$I = \int_{\mathbb{R}^2} ds_1\, ds_2\, f\left(\sqrt{s_1^2+s_2^2}\right) \delta\left(\sqrt{s_1^2+s_2^2} - q\right) = \int_0^{2\pi} d\theta \int_0^\infty dr\, r\, f(r)\, \delta(r-q)$$

$$= (2\pi q)\, f(q) = V(q)\cdot f(q)$$

## (3) LARGE-N AND RANDOM MATRIX THEORY

$$D_N(\epsilon, q) = \left\langle \left| \det\left( \hat{J} - p\epsilon \hat{1} - r_{eff}(q)\, \vec{\omega}_1 \vec{\omega}_1^T \right) \right| \right\rangle$$

Call $\lambda_2 \leq \ldots \leq \lambda_N = \{\lambda_\alpha\}_{\alpha=1}^{M=N-1}$ evalues of $\hat{J} - r_{eff}(q)\vec{\omega}_1\vec{\omega}_1^T$

Then:

$$D_N(\epsilon, q) = \left\langle \prod_{\alpha=1}^{M} |\lambda_\alpha - p\epsilon| \right\rangle = \left\langle e^{\sum_{\alpha=1}^{M} \log|\lambda_\alpha - p\epsilon|} \right\rangle$$

$$= \left\langle e^{M\int d\nu_M(\lambda) \log|\lambda - p\epsilon|} \right\rangle$$

where $\nu_M(\lambda) = \frac{1}{M}\sum_{\alpha=1}^{M} \delta(\lambda - \lambda_\alpha)$ $\qquad M = N-1$

Recall facts in RMT (part I):

- The leading order contibution when $N \gg 1$ is given by the continuous part of $\nu_N(\lambda)$, the density:

$$D_N \approx \left\langle e^{N\int d\lambda\, \rho_N(\lambda) \log|\lambda - p\epsilon| + o(N)} \right\rangle$$

The average $\langle \cdot \rangle$ becomes average over $\mathcal{P}_N(\{\rho\})\mathcal{D}\rho$

- The density $\rho_N(\lambda)$ is self-averaging, and $\rho_\infty(\lambda)$ does not depend on $r$ and is the semicircular law $\rho_{sc}(\lambda)$ with $\sigma^2 \to \tilde{\sigma}^2 p(p-1)$.

$$D_N \overset{N\gg 1}{\simeq} e^{N\int d\lambda\, \rho_{sc}(\lambda)\log|\lambda - p\epsilon| + o(N)}$$

$\Longrightarrow$ This integral can be done explicitly:

$$\int d\lambda \frac{1}{2\pi p(p-1)\tilde{\sigma}^2}\sqrt{4p(p-1)\tilde{\sigma}^2-\lambda^2}\,\log|\lambda-p\epsilon| =$$

$$= \int \frac{d\mu}{\pi}\sqrt{2-\mu^2}\,\log\left|\sqrt{2p(p-1)\tilde{\sigma}^2}\,\mu - p\varepsilon\right|$$

$$= \log\sqrt{2p(p-1)\tilde{\sigma}^2} + I\left(\frac{p\epsilon}{\sqrt{2p(p-1)\tilde{\sigma}^2}}\right)$$

$$I(y) = \int d\mu \frac{\sqrt{2-\mu^2}}{\pi}\log|\mu - y|$$

$$= \begin{cases} \dfrac{y^2-1}{2} + \dfrac{y}{2}\sqrt{y^2-2} + \log\left(\dfrac{-y+\sqrt{y^2-2}}{2}\right) & y \le -\sqrt{2} \\ \dfrac{y^2}{2} - \dfrac{1}{2}(1+\log 2) & y > -\sqrt{2} \end{cases}$$

# Computing distributions: example

Consider the unconstrained gradient: $\nabla \mathcal{E}[s] = \left(\frac{\partial \mathcal{E}}{\partial s_i}\right)_{i=1}^{N}$

Then:

$$\left\langle \frac{\partial \mathcal{E}}{\partial s_i}[\vec{s}] \right\rangle = -r N p \left(\frac{v \cdot s}{N}\right)^{p-1} \frac{v_i}{N}$$

while:

$$\left\langle \frac{\partial \mathcal{E}[\vec{s}]}{\partial s_i} \frac{\partial \mathcal{E}[\vec{s}']}{\partial s'_j} \right\rangle_c = \sum_{k_1=1}^{p} \sum_{k_2=1}^{p} \sum_{\substack{i_1 \leq \dots \leq i_p \\ j_1 \leq \dots \leq j_p}} \langle J_{i_1 i_2 \dots i_p} J_{j_1 j_2 \dots j_p} \rangle \, \delta_{i_{k_1}, i} \, \delta_{j_{k_2}, j} \, s_{i_1} \dots \cancel{s_{i_{k_1}}} \dots s_{i_p} \times$$

$$\times \, s'_{j_1} \dots \cancel{s'_{j_{k_2}}} \dots s'_{j_p}$$

Using that $\langle J_{i_1 \dots i_p} J_{j_1 \dots j_p} \rangle = \frac{p! \, \tilde{\sigma}^2}{N^{p-1}} \prod_{n=1}^{p} \delta_{i_n j_n}$,

$$= \sum_{k_1=1}^{p} \sum_{k_2=1}^{p} \frac{p! \, \tilde{\sigma}^2}{N^{p-1}} \frac{1}{p!} \sum_{i_1, \dots, i_p} \delta_{i_{k_1}, i} \, \delta_{i_{k_2}, j} \, s_{i_1} \dots \cancel{s_{i_{k_2}}} \dots s_{i_p} \times$$

$$\times \, s'_{i_1} \dots \cancel{s'_{i_{k_2}}} \dots s'_{i_p}$$

Distinguishing the case $k_1 = k_2$ ($p$ of them) and $k_1 \neq k_2$ ($p \cdot (p-1)$ of them) one gets:

$$\left\langle \frac{\partial \mathcal{E}[\vec{s}]}{\partial s_i} \frac{\partial \mathcal{E}[\vec{s}']}{\partial s'_j} \right\rangle_c = \tilde{\sigma}^2 \left\{ p \, \delta_{ij} \left(\frac{s \cdot s'}{N}\right)^{p-1} + p(p-1) \frac{s'_i s_j}{N} \left(\frac{s \cdot s'}{N}\right)^{p-2} \right\}$$

Now, $\nabla_\perp \mathcal{E}[\vec{s}]$ is the projections of $\nabla \mathcal{E}[\vec{s}]$ on the space orthogonal to $\vec{s}$, i.e. on the tangent plane $\tau[\vec{s}]$. Choosing $\vec{e}_\alpha[\vec{s}]$ a basis of $\tau[\vec{s}]$, one has $\vec{e}_\alpha \cdot \vec{s} = 0$. Thus:

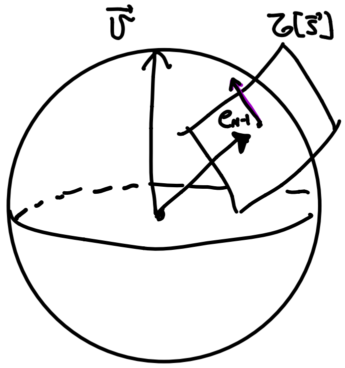
$$\langle (\nabla_\perp \mathcal{E}[\vec{s}])_\alpha \rangle = \langle (\nabla \mathcal{E}[\vec{s}] \cdot \vec{e}_\alpha) \rangle = -r N p \left(\frac{v \cdot s}{N}\right)^{p-1} \left(\frac{v \cdot e_\alpha}{N}\right)$$

And:

$$\langle (\nabla_\perp \mathcal{E}[\vec{s}])_\alpha (\nabla_\perp \mathcal{E}[\vec{s}])_\beta \rangle_c = \langle (\nabla \mathcal{E}[\vec{s}] \cdot e_\alpha[\vec{s}])(\nabla \mathcal{E}[s] \cdot e_\beta[\vec{s}]) \rangle =$$

$$= \tilde{\sigma}^2 p \, \delta_{\alpha\beta} + p(p-1) \underbrace{\left(\frac{s \cdot e_\alpha}{\sqrt{N}}\right)}_{=0} \underbrace{\left(\frac{s \cdot e_\beta}{\sqrt{N}}\right)}_{=0} = p \tilde{\sigma}^2 \delta_{\alpha\beta}$$

In the annealed calculation, all distributions depend on $\vec{s}$ only via $q[\vec{s}] = \frac{\vec{v} \cdot \vec{s}}{N}$ : $\vec{v}$ is the only 'special direction' on the sphere, that breaks isotropy. It is convenient to choose, for each $\vec{s}$, this basis on the tangent plane:




$$\vec{e}_{N-1}[\vec{s}] = \frac{1}{\sqrt{N(1-q^2)}} \left( \vec{v} - q[\vec{s}] \, \vec{s} \right)$$

$$\vec{e}_\alpha[\vec{s}] \perp \{\vec{v}, \vec{s}\} \qquad \alpha = 1, .., N-2$$

$\Rightarrow$ only $(\nabla_\perp \mathcal{E})_{N-1}$ and $(\nabla_\perp^2 \mathcal{E})_{\alpha\, N-1}$ or $(\nabla_\perp^2 \mathcal{E})_{N-1, \alpha}$ will have a q-dependent distribution

$$\langle \nabla \mathcal{E}[\vec{s}] \cdot \vec{e}_\alpha \rangle = r N p \left(\frac{Q[\vec{s}]}{N}\right)^{p-1} \left(\frac{\vec{v}}{N} \cdot \vec{e}_\alpha\right) = \begin{cases} 0 & \alpha < N-1 \\ \neq 0 & \alpha = N-1 \end{cases}$$

$$\langle (\nabla_\perp \mathcal{E}[\vec{s}])_\alpha (\nabla_\perp \mathcal{E}[\vec{s}])_\beta \rangle_c = p \tilde{\sigma}^2 \delta_{\alpha\beta}$$

# The annealed complexity

Combine all terms:

$$\langle \mathcal{N}(\epsilon,q)\rangle = V_N(q)\, D_N(\epsilon,q)\, P_1(q)\, P_2(N\epsilon) = e^{N\Sigma_A(\epsilon,q)+o(N)}$$

$$\boxed{\begin{aligned}\Sigma_A(\epsilon,q) = \frac{1}{2}\log\left[2e(p-1)(1-q^2)\right] - \frac{p}{2\tilde{\sigma}^2} r^2 q^{2p-2}(1-q^2) \\ -\frac{1}{2\tilde{\sigma}^2}\left(\epsilon + r q^p\right)^2 + I\left(\sqrt{\frac{p}{2(p-1)\tilde{\sigma}^2}}\,\epsilon\right)\end{aligned}}$$

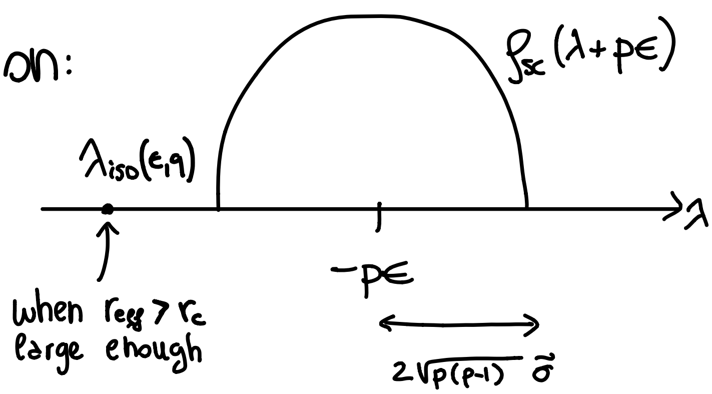
This gives distribution of stationary points in energy and geometry (overlap with $\vec{v}$), on average.

What about stability?

- The Hessian at a stationary point with $(\epsilon,q)$ is a rank-1 perturbed, shifted GOE:

$$\nabla_\perp^2 \mathcal{E}\Big|_{\epsilon,q} \overset{\text{law}}{\sim} \hat{J} - p\epsilon\,\hat{1} - r_{\text{eff}}(q)\,\vec{w}_\perp \vec{w}_\perp^T$$

The eigenvalue distribution:



Local minima have all eigenvalues positive. For bulk, need:

$$-p\epsilon > 2\sqrt{p(p-1)}\,\tilde{\sigma} \implies \epsilon < \epsilon_{th} = -2\tilde{\sigma}\sqrt{\frac{p-1}{p}}$$

$\epsilon_{th}$ = "threshold energy". Also, $\lambda_{iso}(\epsilon,q) > 0$.

## The $p \to 2$ limit of $\Sigma_A(\epsilon,q)$

The annealed complexity is maximal at $q=0$.
We set $\Sigma_A(\epsilon) = \Sigma_A(\epsilon, q=0)$.

Recall that $\langle J^2_{i_1 \dots i_p} \rangle = \frac{p!\,\tilde{\sigma}^2}{N}$ while in PART I we set $\langle J_{ij}^2 \rangle = \frac{\sigma^2}{2}(1+\delta_{ij})$. To be consistent, $\tilde{\sigma}^2 = \frac{\sigma^2}{2}$

$$\Sigma_A(\epsilon) \xrightarrow{p=2} \frac{1}{2}\log(2e) - \frac{\epsilon^2}{\sigma^2} + I\left(\sqrt{2/\sigma^2}\,\epsilon\right)$$

Then, given that $\epsilon > -\sigma$:

$$\longrightarrow \frac{1}{2}\log(2e) - \frac{\epsilon^2}{\sigma^2} + \frac{\epsilon^2}{\sigma^2} - \frac{1}{2} - \frac{\log 2}{2} = 0$$

Consistently with the fact that for $p=2$ there are <u>not</u> exponentially-many stationary points.
One can use the Kac-Rice formula to get the results of PART I: exercise 4!

## ▨ The quenched calculation: what would change?

One needs to compute higher moments $\langle \mathcal{N}_N^{\omega}(\epsilon,q)\rangle$ with $\omega = 2,3,4\dots$ and $\omega \to 0$!

One can use Kac-Rice formulas, too, for higher moments: need to consider $\omega$ points on sphere: $\vec{s}^{\,a}$ with $a=1,\dots,\omega$. The fields $\mathcal{E}[\vec{s}^{\,a}], \nabla_\perp \mathcal{E}[\vec{s}^{\,a}], \nabla_\perp^2 \mathcal{E}[\vec{s}^{\,a}]$ are correlated.

$$\langle \mathcal{N}^{\omega}(\epsilon,q)\rangle = \int_{S_N(\sqrt{N})} \prod_{a=1}^{\omega} d\vec{s}^{\,a}\, \delta\left(\vec{s}^{\,a}\cdot\vec{v} - Nq\right) \, P_{\nabla_\perp \mathcal{E}[\vec{s}^{\,a}],\, \mathcal{E}[\vec{s}^{\,a}]}(\vec{0}, N\epsilon) \left\langle \prod_{a=1}^{\omega} \left| \det \nabla_\perp^2 \mathcal{E}[\vec{s}^{\,a}]\right| \right\rangle_{\substack{\nabla_\perp \mathcal{E}[\vec{s}^{\,a}]=0 \\ \mathcal{E}[\vec{s}^{\,a}]=N\epsilon}}$$

Joint distribution of $\omega$ $(N-1)$-dim vectors $\nabla_\perp \mathcal{E}[\vec{s}^{\,a}]$ and $\omega$ scalars $\mathcal{E}[\vec{s}^{\,a}]$

joint expectation value, conditioned to $\omega$ vectors and functions

Some consequences of correlations:

(i) No decoupling: $\nabla_\perp \mathcal{E}[\vec{s}^{\,a}]$ for fixed $a$ is independent of $\mathcal{E}[\vec{s}^{\,a}], \nabla_\perp^2 \mathcal{E}[\vec{s}^{\,a}]$, but not of $\mathcal{E}[\vec{s}^{\,b}], \nabla_\perp^2 \mathcal{E}[\vec{s}^{\,b}]$ at $b \neq a$.

Consequences: (1) need to compute joint distributions, (2) the expectation of Hessians is a problem of coupled random matrices.

What helps: Still Gaussian for (1), and large-$N$ for (2).

(ii) Distributions depend not only on $q_N[\vec{s}^a] = \left(\frac{\vec{s}^a \cdot \vec{v}}{N}\right)$,

but also on mutual overlaps $Q_N[\vec{s}^a, \vec{s}^b] = \left(\frac{\vec{s}^a \cdot \vec{s}^b}{N}\right)$:

Consequence: no longer 1 special direction, but $\varpi$ of them.
What helps: Still, huge dimensionality reduction!
From $N \cdot \varpi$ variables $s_i^a$ to $\frac{\varpi(\varpi-1)}{2} + \varpi$ ones,
the $Q_N[\vec{s}^a, \vec{s}^b]$ and $q_N[\vec{s}^a]$. Because fully-connected.

(iii) The conditional distribution of the Hessian at one point $\vec{s}^a$ is still that of a perturbed GOE, but finite-rank perturbations are more complicated: both additive & multiplicative, and not "free" (in the sense of free probability).

WHY: multiplicative perturbations due to conditioning to $\nabla_\perp \mathcal{E}[\vec{s}^b]$ with $b \neq a$.

Consequence: calculation of isolated evalues is more involved; What helps: perturbation is still of finite-rank.

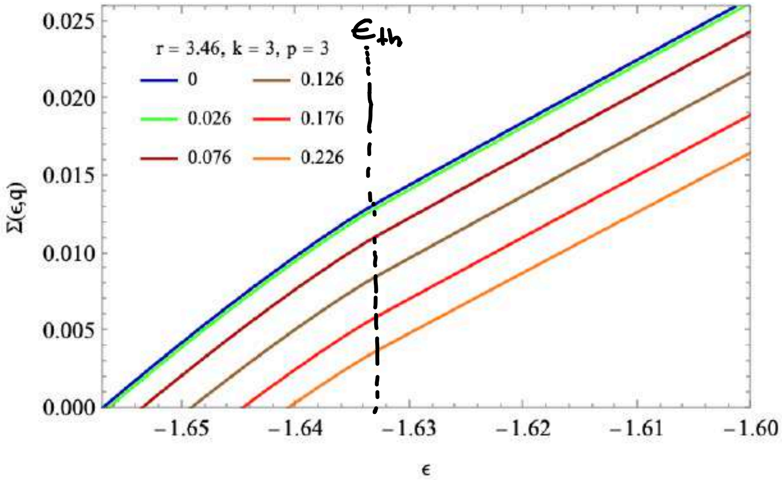
To see comparisons between quenched & annealed, see

ROS, BEN AROUS, BIROLI, CAMMAROTA 2018

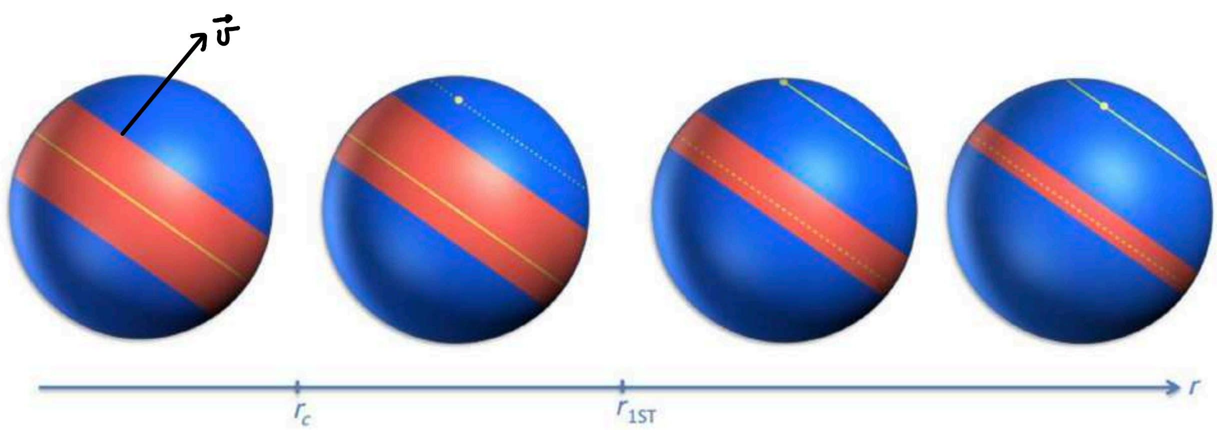
# II 3. WHAT: GS, LANDSCAPE, DYNAMICS

Back to the inference problem. Here, summarize results of quenched calculation:

- Quenched complexity curves ($\gamma = 1$) fixed r




- Landscape's evolution with r: regions where $\Sigma_\infty(\epsilon, q) > 0$ for some $\epsilon$ (in red), and $q[S_{GS}]$ (yellow).

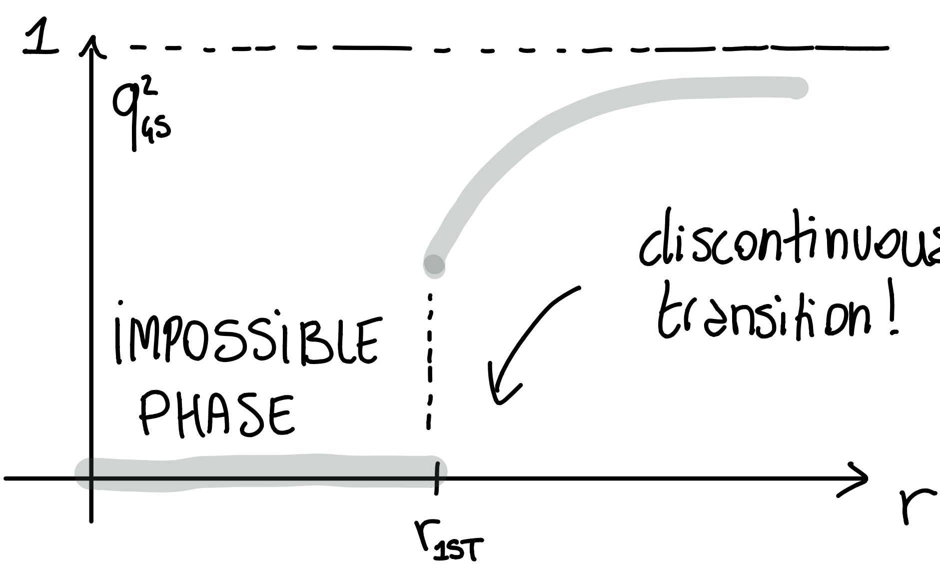



$\hookrightarrow$ an isolated (local) minimum appears close to signal $\vec{v}$.

# Recovering the signal

Q1: when is $\vec{\xi}_{GS}$ informative, i.e. $q_{GS} > 0$ ?

A sharp transition when $N \to \infty$ at some $r = r_{1ST}$




RECOVERY THRESHOLD

Differences with respect to $p=2$: the transition is discontinuous, first order!
As for $p=2$: could be obtained with thermodynamic calculation for $\beta \to \infty$

GILLIN SHERRINGTON 2000

# A Landscape of minima

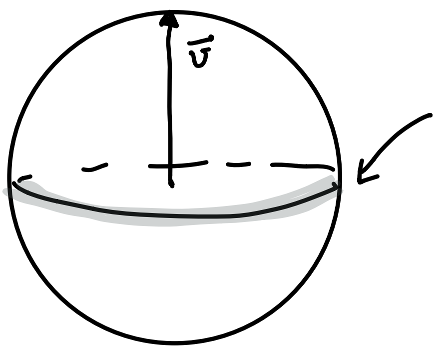
- Most stationary points are un-informative of $\vec{v}$:

(Neglecting isolated minimum at high overlap)

Optimize over $q$ : $\Sigma_\infty(\epsilon, q)$ maximal at $q=0$:

$$\Sigma_\infty(\epsilon) = \max_q \Sigma_\infty(\epsilon, q) = \frac{1}{2}\log\left[2e(p-1)\right] - \frac{\epsilon^2}{2\tilde{\sigma}^2} + I\left(\sqrt{\frac{p}{2(p-1)\tilde{\sigma}^2}}\,\epsilon\right)$$

does not depend on $r$! Also, $\Sigma_\infty(\epsilon, q=0) = \Sigma_A(\epsilon, q=0)$




exponential majority of stationary points is orthogonal to the signal! (Not informative)

- Exponentially many local minima!

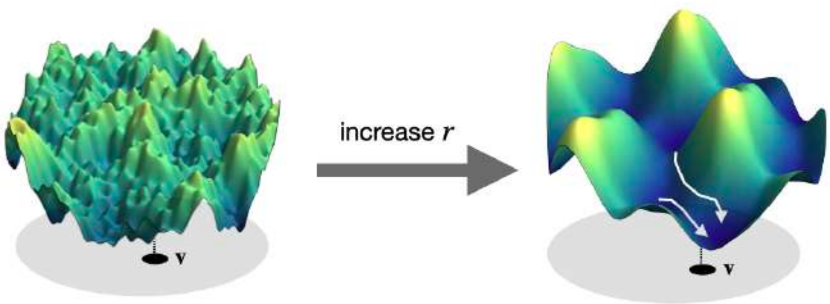
Recall Hessian (annealed calculation)

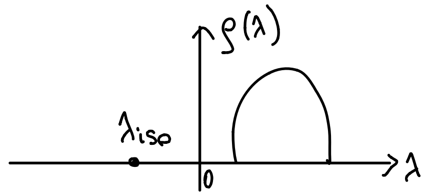
$$\nabla_\perp^2 \mathcal{E}\Big|_{\epsilon, q} \overset{\text{law}}{\sim} \hat{J} - p\epsilon\,\hat{1} - r_{\text{eff}}(q)\,\vec{w}_\perp \vec{w}_\perp^T$$

Local minima: $\epsilon < \epsilon_{th}$, $\lambda_{iso}(\epsilon, q) > 0$.

$q=0$ : $r_{eff}(q=0)=0$. No isolated evalue. Exponentially-many local minima $\epsilon<\epsilon_{th}$: trapping states for dynamics!

$q>0$ : when $r$ large enough, generate isolated evalue, that can become negative : minima → saddle transitions




Spectrum of rank-1 saddle

- "Topological trivialization?" How strong should $r$ be to destabilize also minima at equator? Need $r \sim N^{\alpha}$:

$$r_{eff} = r\,p(p-1)\left(\frac{\vec{s}\cdot\vec{v}}{N}\right)^{p-2}\left(1-\left(\frac{\vec{s}\cdot\vec{v}}{N}\right)^{2}\right) \sim r\left(\frac{1}{\sqrt{N}}\right)^{p-2}$$

$$\Longrightarrow \alpha = \frac{p-2}{2}$$

# Dynamics: DMFT. And beyond?

'Easy' phase: for $r \sim N^{\alpha}$ with $\alpha > \alpha_c = \frac{p-2}{2}$, gradient descent converges to $\vec{s}_{GS}$ in times $O(N^0)$.
BEN AROUS, GHEISSARI, JAGANNATH 2020

'Hard' phase $r \sim O(1)$: dynamics from random initial conditions stuck in high-entropy $q=0$ region, the equator. Here landscape is as if $r=0$.

## The dynamics at $r=0$: "short times".

Described by DMFT ($N \to \infty$ before $t \to \infty$)
Excess energy does not decay to zero as for $p=2$, but converges to finite value:

$$\lim_{N\to\infty} \Delta_N \epsilon(t) = \lim_{N\to\infty} \left(\epsilon_N(t) - \epsilon_{GS}\right) = -\tilde{\sigma}^2 p \int_0^t C^{p-1}(t,s)\, R(t,s)\, ds - \epsilon_{GS}$$

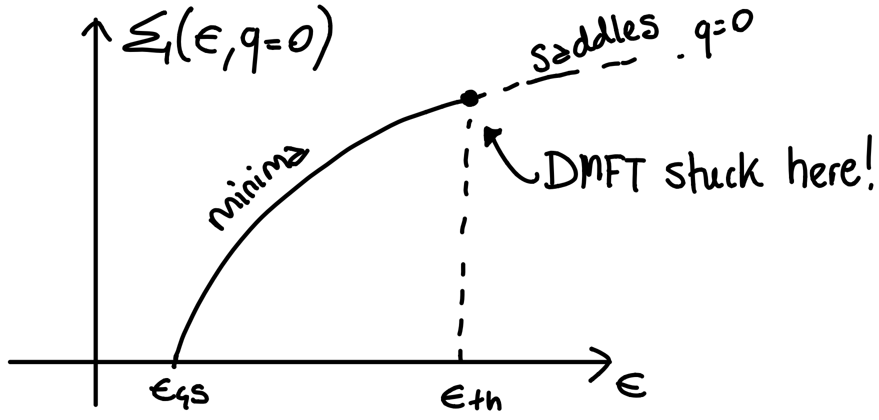
When $t \to \infty$, converges to finite value:

$$\lim_{t\to\infty} \lim_{N\to\infty} \Delta_N \epsilon(t) = \epsilon_{th} - \epsilon_{GS} > 0$$

Never reach the GS energy density in these timescales. Out-of-equilibrium glassy dynamics, aging. CUGLIANDOLO, KURCHAN 1993

BOUCHAUD, CUGLIANDOLO, KURCHAN, MEZARD 1997 (review)

Landscape interpretation?




$\Longrightarrow$ gradient descent gets stuck at energies of the highest-energy minima, that are exponentially numerous.

CUGLIANDOLO, KURCHAN 1993

SELLKE 2024 (math)

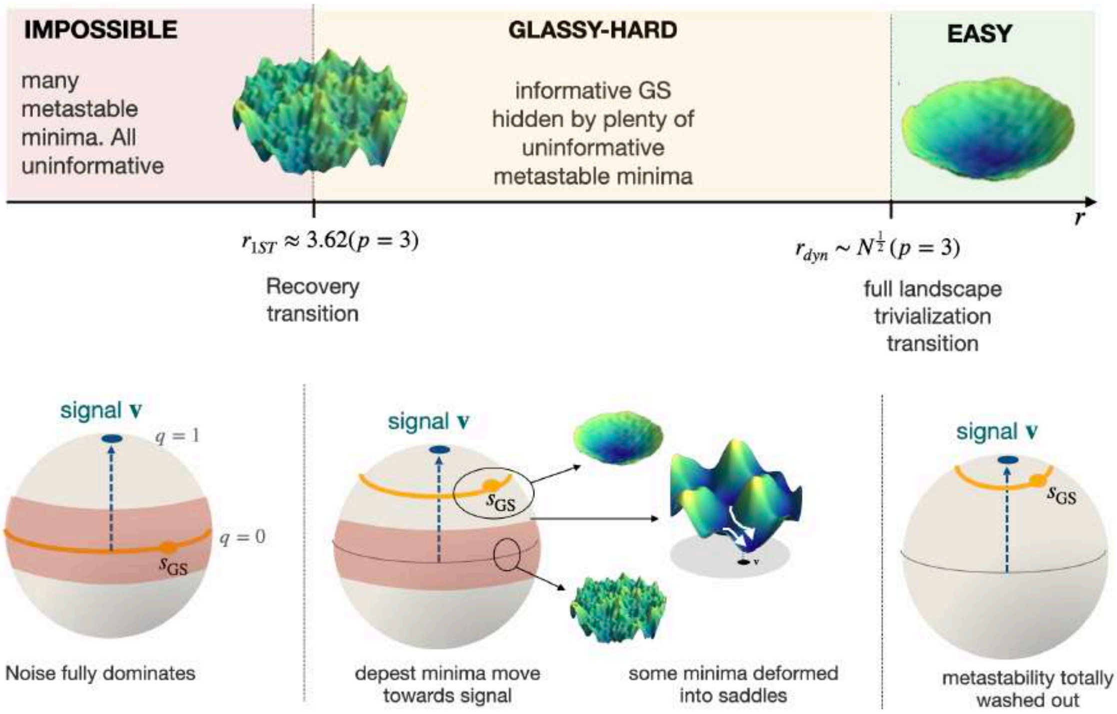
▶ The dynamics at $r=0$: "long times".

For $p=2$, equilibration timescales $\sim \mathcal{O}(N^{2/3})$.

For $p\geq 3$, expect timescales $\sim \mathcal{O}(e^N)$: system has to escape from trapping minima crossing energy barriers $\Delta E \sim \mathcal{O}(N) \Rightarrow$ ACTIVATED DYNAMICS.

This regime of the dynamics is open problem!

# In summary




- The ground-state becomes correlated with $\vec{v}$ for $r > r_{1ST}$
- Exponentially-many local minima for all values of r. Those closer to $\vec{v}$ become saddles when r increases, those at equator remain minima.
- Optimization is hard: system trapped by metastable states. Mean-field dynamics studied a lot for r=0. Dynamics at finite N is open problem.

# Key Tools

- [x] RMT: Stieltjes transform, expansion resolvent in $1/N$, density & isolated evalues
- [x] BBP transition, typical values, finite-N fluctuations, scalin
- [x] mean-field dynamics, DMFT, crossovers
- [x] Landscape complexity, quenched vs annealed, replica trick, Kac-Rice formalism

# Spiked GOE: eigenvalues density and outliers

[Ref: Bouchaud, Potters, A First Course in Random Matrix Theory, Cambridge University Press 2020].

Take the $N \times N$ matrix $\hat{M} = \hat{J} + \hat{R}$, where $\hat{J}$ is a GOE matrix with $\langle J_{ij} \rangle = 0$ and $\langle J_{ij}^2 \rangle = \frac{\sigma^2}{N}(1 + \delta_{ij})$, while $\hat{R} = r\vec{w}\vec{w}^T$ is a rank-1 perturbation, with $||\vec{w}||^2 = 1$. Call $\lambda_\alpha$ with $\alpha = 1, \cdots, N$ the eigenvalues of $\hat{M}$, and call $\vec{u}_\alpha$ the corresponding eigenvectors. The resolvent of $\hat{M}$ is

$$\hat{G}_{\hat{M}}(z) = \frac{1}{z\hat{1} - \hat{M}} = \sum_{\alpha=1}^{N} \frac{\vec{u}_\alpha \vec{u}_\alpha^T}{z - \lambda_\alpha}$$

The goal of these two exercises is to derive the self-consistent equations for the Stieltjes transform of $\hat{M}$, and for its isolated eigenvalue.

### Exercise 1. Replica calculation of the Stieltjes transform.

The starting point of the calculation is the Gaussian identity :

$$\left(\frac{1}{z\hat{1} - \hat{M}}\right)_{ij} = \frac{1}{\mathcal{Z}} \int \prod_{i=1}^{N} \frac{d\psi_i}{\sqrt{2\pi}} \psi_i \psi_j e^{-\frac{1}{2}\sum_{i,j=1}^{N} \psi_i (z\hat{1} - \hat{M})_{ij} \psi_j}, \quad \mathcal{Z} = \int \prod_{i=1}^{N} \frac{d\psi_i}{\sqrt{2\pi}} e^{-\frac{1}{2}\sum_{i,j=1}^{N} \psi_i (z\hat{1} - \hat{M})_{ij} \psi_j}$$

We wish to take the average of this expression with respect to the matrix $\hat{M}$. However, averaging the partition function in the denominator makes the calculation potentially difficult; to proceed, we make use of the replica trick to write

$$\mathcal{Z}^{-1} = \lim_{n \to 0} \mathcal{Z}^{n-1}.$$

We then follow the standard steps of replica calculations, see below.

(i) **From randomness to coupled replicas.** Using the replica trick, justify why $(z\hat{1} - \hat{M})^{-1} = \lim_{n \to 0} I_{ij}^{(n)}$ where

$$I_{ij}^{(n)} = \int \prod_{a=1}^{n} \prod_{i=1}^{N} \frac{d\psi_i^a}{\sqrt{2\pi}} \psi_i^1 \psi_j^1 e^{-\frac{1}{2}\sum_{a=1}^{n} \sum_{i,j=1}^{N} \psi_i^a (z\hat{1} - \hat{J} - r\vec{w}\vec{w}^T)_{ij} \psi_j^a}$$

Take the average of this expression with respect to $J_{ij}$, and show that

$$\langle I_{ij}^{(n)} \rangle = \int \prod_{a=1}^{n} \prod_{i=1}^{N} \frac{d\psi_i^a}{\sqrt{2\pi}} \psi_i^1 \psi_j^1 e^{-\frac{1}{2}\sum_{a=1}^{n} \sum_{i,j=1}^{N} \psi_i^a (z\delta_{ij} - rw_i w_j) \psi_j^a} e^{\frac{\sigma^2}{4N} \sum_{a,b} \left(\sum_{i=1}^{N} \psi_i^a \psi_i^b\right)^2}.$$

Now one has an expression without randomness, in which the replicated variables $\psi^a$ are coupled with each others.

(ii) **Hubbard–Stratonovich.** We would like now to perform the integral over the variables $\psi_i^a$; however, this integral contains quartic terms in the exponent; in order to turn such an integral into a Gaussian one, we perform a Hubbard-Stratonovich transformation: we introduce the order parameters

$$Q_{ab}[\psi] = \frac{1}{N} \sum_{i=1}^{N} \psi_i^a \psi_i^b \quad a \leq b$$

and write the integral as

$$\int \prod_{a=1}^{n} \prod_{i=1}^{N} \frac{d\psi_i^a}{\sqrt{2\pi}} \cdots \to N^{\frac{n(n+1)}{2}} \int \prod_{a \leq b} dQ_{ab} \int \prod_{a=1}^{n} \prod_{i=1}^{N} \frac{d\psi_i^a}{\sqrt{2\pi}} \prod_{a \leq b} \delta\left(NQ_{ab} - \sum_{i=1}^{N} \psi_i^a \psi_i^b\right) \cdots$$

Show that using the integral representation of the delta distributions

$$\delta\left(NQ_{ab}-\sum_{i=1}^{N}\psi_i^a\psi_i^b\right)=\int\frac{d\lambda_{ab}}{2\pi}e^{i\lambda_{ab}\left[NQ_{ab}-\sum_{i=1}^{N}\psi_i^a\psi_i^b\right]}$$

and introducing the $n\times n$ matrix $\Lambda$ with components $\Lambda_{ab}=2\lambda_{aa}\delta_{ab}+\lambda_{ab}(1-\delta_{ab})$ and the $N\times N$ matrix $A$ with components $A_{ij}=z\delta_{ij}+rw_iw_j$, the average can be cast in the following form:

$$\langle I_{ij}^{(n)}\rangle=N^{\frac{n(n+1)}{2}}\int\prod_{a\leq b}dQ_{ab}d\lambda_{ab}e^{\frac{N\sigma^2}{4}\mathrm{Tr}_n[Q^2]+\frac{N}{2}\mathrm{Tr}_n[i\Lambda\, Q]}f_N[Q,\vec{w}] \tag{1}$$

with

$$f_N[Q,\vec{w}]=\int\prod_{a=1}^{n}\prod_{i=1}^{N}\frac{d\psi_i^a}{\sqrt{2\pi}}\psi_i^1\psi_j^1e^{-\frac{1}{2}\sum_{a,b}\sum_{i,j}\psi_i^a\left[\hat{1}_N\otimes i\Lambda+A\otimes\hat{1}_n\right]_{ij}^{ab}\psi_j^b}.$$

(iii) **Gaussian integration.** Performing the Gaussian integral, show that

$$\langle I_{ij}^{(n)}\rangle=\delta_{ij}\int\prod_{a\leq b}dQ_{ab}d\lambda_{ab}e^{\frac{N}{2}A_N[Q,i\Lambda]}\left[(A\otimes 1_n+1_N\otimes i\Lambda)^{-1}\right]_{ij}^{11}$$

$$A_N[Q,i\Lambda]=\frac{\sigma^2}{2}\mathrm{Tr}_n[Q^2]+\mathrm{Tr}_n[i\Lambda\, Q]-\frac{1}{N}\mathrm{Tr}_{nN}[\log\left(A\otimes 1_n+1_N\otimes i\Lambda\right)]$$

*Hint.* Use that $\int\prod_{i=1}^{d}\frac{dx_i}{\sqrt{2\pi}}x_lx_me^{-\frac{1}{2}\vec{x}\cdot\hat{K}\vec{x}}=\hat{K}^{-1}_{lm}|\det K|^{-1}$ and that $\log|\det K|=\mathrm{Tr}\log K$.

(iv) **Saddle point.** The integral can now be computed with a saddle point approximation. Show that the saddle point equations for the matrices $Q$ and $i\Lambda$ read

$$i\Lambda=-\sigma^2Q,\qquad Q=\frac{1}{N}\mathrm{Tr}_{nN}\left[\frac{1}{A\otimes 1_n+1_N\otimes i\Lambda}\right]$$

Show that, plugging the first into the second and assuming that the matrices $\Lambda, Q$ are diagonal and replica symmetric, i.e. $Q_{ab}=\delta_{ab}g$ and $\lambda_{ab}=\delta_{ab}\lambda$, one reduces to a single equation for $g$ which reads

$$g=\frac{1}{N}\mathrm{Tr}_N\left[\frac{1}{(z-\sigma^2g)\hat{1}_N-r\vec{w}\vec{w}^T}\right]$$

Using that

$$\langle(z\hat{1}-\hat{M})^{-1}\rangle=\lim_{n\to 0}\langle I_{ij}^{(n)}\rangle=\left[\left(A\otimes 1_n-\sigma^2g1_N\otimes 1_n\right)^{-1}\right]_{ij}^{11},$$

justify why $g$ is the Stieljes transform of the matrix $M$. Show that expanding $g=g_\infty+g_1/N+\cdots$, the leading order term satisfies the equation

$$g_\infty^{-1}=z-\sigma^2g_\infty.$$

## Exercise 2. The isolated eigenvalue and eigenvector.

(i) Show that if $\hat{A}$ is a matrix and $\vec{v},\vec{u}$ are vectors, then

$$(\hat{A}+\vec{u}\vec{v}^T)^{-1}=\hat{A}^{-1}-\frac{A^{-1}\vec{u}\vec{v}^TA^{-1}}{1+\vec{v}\cdot A^{-1}\vec{u}}.$$

Use this formula (Shermann-Morrison formula) to get an expression for $\hat{G}_{\hat{M}}(z)$.

(ii) The isolated eigenvalue is a pole of the resolvent operator $\hat{G}_{\hat{M}}(z)$, which is real and such that $\lambda_{\rm iso}>2\sigma$. Using that $\lambda_{\rm iso}$ does not belong to the spectrum of the unperturbed matrix $\hat{J}$, show that it solves the equation

$$r\vec{w}\cdot G_{\hat{J}}(\lambda_{\rm iso})\vec{w}=1.$$

(iii) Using that $\hat{J}$ and $\vec{w}$ are independent and that typically $\vec{w}$ is *delocalized* in the eigenbasis of $\hat{J}$, show that

$$\vec{w} \cdot G_{\hat{J}}(\lambda_{\rm iso})\vec{w} \overset{N\to\infty}{\longrightarrow} g_{\rm sc}(\lambda_{\rm iso})$$

where $g_{\rm sc}(\lambda)$ is the Stieltijes transform of the GOE matrix $\hat{J}$.

(iv) Using the self-consistent equation satisfied by $g_{\rm sc}(\lambda)$, derive the expression of the inverse function $g_{\rm sc}^{-1}$ and determine its domain; use it to show that

$$\lambda_{\rm iso} = \frac{\sigma^2}{r} + r \qquad r \geq \sigma.$$

(v) The eigenvectors projections $\xi_\alpha = (\vec{w} \cdot \vec{u}_\alpha)^2$ can be obtained from the resolvent as residues of the poles:

$$\xi_\alpha = \lim_{\lambda \to \lambda_\alpha} (\lambda - \lambda_\alpha)\vec{w} \cdot G_{\hat{M}}(\lambda)\vec{w}$$

Use this to show that if $\alpha = N$ labels the isolated eigenvalue, then

$$\xi_N = -\frac{1}{r^2 g'_{\rm sc}(\lambda_{\rm iso})} = 1 - \frac{\sigma^2}{r^2}.$$

*Hint.* Use that if $\lim_{\lambda\to\lambda_0} f(\lambda) = 0 = \lim_{\lambda\to\lambda_0} g(\lambda)$, then $\lim_{\lambda\to\lambda_0} \frac{f(\lambda)}{g(\lambda)} = \lim_{\lambda\to\lambda_0} \frac{f'(\lambda)}{g'(\lambda)}$.

# Condensation transition

[Ref: Kosterlitz, Thouless, Jones, *Spherical Model of a Spin-Glass*, PRL 36 (1976)].

The matrix denoising problem is formulated in terms of the ground state of the energy lansdcape:

$$\mathcal{E}[\vec{s}] = -\frac{1}{2}\sum_{ij} s_i (J_{ij} + r v_i v_j) s_j, \qquad ||\vec{s}||^2 = N = ||\vec{v}||^2, \qquad \hat{J} \sim GOE$$

The behavior of the ground state can be characterized by studying the thermodynamics of the system in the limit $\beta \to \infty$, through the partition function:

$$\mathcal{Z}_\beta = \int_{S_N(\sqrt{N})} d\vec{s} e^{-\beta \mathcal{E}[\vec{s}]}, \qquad S_N(\sqrt{N}) = \left\{\vec{s} : ||\vec{s}||^2 = N\right\}$$

As a function of temperature, this model exhibits a transition at a critical temperature $T_c(r)$, which can be interpreted as a *condensation transition* (like in BEC physics).

### Exercise 3. Thermodynamics of the model

(i) Call $\lambda_\alpha$ ($\lambda_1 \leq \lambda_2 \leq \cdots \lambda_N$) the eigenvalues of $\hat{M} = \hat{J} + \hat{R}$, and $\vec{u}_\alpha$ the corresponding eigenvectors. Call $s_\alpha = \vec{s} \cdot \vec{u}_\alpha$. Show that the partition function can be written as

$$\mathcal{Z}_\beta = \int d\lambda \int \prod_{\alpha=1}^{N} ds_\alpha e^{\frac{\beta}{2}\left[\sum_\alpha \lambda_\alpha s_\alpha^2 - \lambda(\sum_\alpha s_\alpha^2 - N)\right]}$$

(ii) Show that the thermal expectation value of the mode occupations is

$$\langle s_\gamma^2 \rangle = \frac{1}{\mathcal{Z}_\beta} \int d\lambda \int \prod_{\alpha=1}^{N} ds_\alpha\, s_\gamma^2\, e^{-\frac{\beta}{2}\left[-\sum_\alpha \lambda_\alpha s_\alpha^2 + \lambda(\sum_\alpha s_\alpha^2 - N)\right]} = \frac{1}{\beta(\lambda^* - \lambda_\gamma)}$$

where $\lambda^* > \lambda_\gamma$ for all $\gamma$ is fixed by the equation

$$\sum_{\gamma=1}^{N} \langle s_\gamma^2 \rangle = N = \sum_{\gamma=1}^{N} \frac{1}{\beta(\lambda^* - \lambda_\gamma)}$$

(iii) The matrix $\hat{M}$ is a spiked GOE. Take $r < r_c = \sigma$. Justify why for large $N$ the equation for $\lambda^*$ becomes:

$$\beta = g_{\rm sc}(\lambda^*) \qquad \lambda^* > 2\sigma$$

where $g_{\rm sc}(\lambda^*)$ is the Stieltjies transform of the GOE; show that there is a critical temperature $\beta_c = \sigma^{-1}$ and compute the solution $\lambda^*$ for $\beta < \beta_c$. Show that at $\beta_c$, $\lambda^*$ attains its maximal possible value. Show that at low temperature $\beta > \beta_c$ the equation can be solved assuming *condensation* of the fluctuations in the lowest-energy mode:

$$\frac{1}{N}\langle s_N^2 \rangle = 1 - \frac{1}{\beta\sigma}$$

This condensation transition corresponds also to a transition between a paramagnet at high temperature, and a spin-glass at low temperature.

(iv) Consider now $r > r_c = \sigma$, when the maximal eigenvalue is $\lambda_N = \lambda_{\rm iso} = \frac{\sigma^2}{r} + r$; justify why now the critical temperature is $\beta_c = 1/r$, and a solution of the equation for $\lambda^*$ (with $\lambda^* > \lambda_\gamma$) exists for $\beta < \beta_c$. Show that for $\beta > \beta_c$ it must hold

$$\frac{1}{N}\langle s_N^2 \rangle = \frac{1}{N}\langle s_{\rm iso}^2 \rangle = 1 - \frac{1}{\beta r}$$

In this regime, the condensation transition coincides with a transition between a paramagnet at high temperature, and a ferromagnet at low temperature.

# Exercise 1 - solution

## Stieltjes transform with replica method

(i) The normalization $Z$ is an integral over the variables $\varphi_i$. Writing:

$$Z^{n-1} = \left[ \int \prod_{i=1}^{N} \frac{d\varphi_i}{\sqrt{2\pi}} \cdots \right]^{n-1} = \left[ \int \prod_{i=1}^{N} \frac{d\varphi_i^{(2)}}{\sqrt{2\pi}} \cdots \right] \cdots \left[ \int \prod_{i=1}^{N} \frac{d\varphi_i^{(n)}}{\sqrt{2\pi}} \cdots \right]$$

we can set:

$$\lim_{n\to 0} Z^{n-1} \int \prod_{i=1}^{N} \frac{d\varphi_i}{\sqrt{2\pi}} \varphi_i \varphi_j \, e^{-\frac{1}{2}\sum_{ij} \varphi_i (z-M)_{ij} \varphi_j} =$$

$$= \lim_{n\to 0} \int \prod_{i=1}^{N} \frac{d\varphi_i}{\sqrt{2\pi}} \varphi_i \varphi_j \, e^{-\frac{1}{2}\sum_{ij} \varphi_i (z-M)_{ij} \varphi_j} \int \prod_{a=2}^{n} \prod_{i=1}^{N} \frac{d\varphi_i^{(a)}}{\sqrt{2\pi}} \cdots$$

label these variables as $\varphi_i^{(1)}$

$$= \lim_{n\to 0} \int \prod_{a=1}^{n} \prod_{i=1}^{N} \frac{d\varphi_i^{(a)}}{\sqrt{2\pi}} \varphi_i^{(1)} \varphi_j^{(1)} \, e^{-\frac{1}{2}\sum_{a=1}^{n}\sum_{ij} \varphi_i^{a} (z-M)_{ij} \varphi_j^{a}}$$

$$= \lim_{n\to 0} I_{ij}^{(n)}$$

(ii) Using the integral representation of $\delta(\cdot)$, we obtain:

$$\langle I_{ij}^{(n)} \rangle = \int \prod_{a=1}^{n} \prod_{i=1}^{N} \frac{d\psi_i^{(a)}}{\sqrt{2\pi}} \cdot N^{\frac{n(n+1)}{2}} \int \prod_{a\le b} dQ_{ab} \int \prod_{a\le b} \frac{d\lambda_{ab}}{2\pi} \, e^{i \sum_{a\le b} \lambda_{ab} \left( N Q_{ab} - \sum_i \psi_i^a \psi_i^b \right)} \times$$

$$\times \, \psi_i^1 \psi_j^1 \, e^{-\frac{1}{2} \sum_a \sum_{ij} \psi_i^a \underbrace{(z\delta_{ij} - r w_i w_j)}_{A_{ij}} \psi_j^a} \times$$

$$\times \, e^{\frac{\sigma^2}{4N} \sum_{a,b} \Big( \underbrace{\sum_i \psi_i^a \psi_j^b}_{N Q_{ab}} \Big)^2}$$

← exchange order integration

$$= N^{\frac{n(n+1)}{2}} \int \prod_{a\le b} dQ_{ab} \int \prod_{a\le b} \frac{d\lambda_{ab}}{2\pi} \, e^{\overbrace{iN\left( \frac{1}{2} \sum_{a\ne b} Q_{ab}\lambda_{ab} + \sum_a Q_{aa}\lambda_{aa} \right)}^{(*)}}$$

$$\times \, e^{\frac{\sigma^2 N}{4} \sum_{a,b} Q_{ab}^2} \times \int \prod_{a=1}^{n} \prod_{i=1}^{N} \frac{d\psi_i^a}{\sqrt{2\pi}} \, \psi_i^1 \psi_j^1$$

$$e^{\underbrace{-i\left( \frac{1}{2} \sum_{a\ne b} \lambda_{ab} \sum_i \psi_i^a \psi_i^b + \sum_a \lambda_{aa} \sum_i \psi_i^a \psi_i^a \right)}_{(**)}} \, e^{-\frac{1}{2} \sum_a \sum_{ij} \psi_i^a A_{ij} \psi_j^a}$$

Introducing $\Lambda_{ab} = 2\lambda_{aa}\delta_{ab} + \lambda_{ab}(1-\delta_{ab})$ and the trace $\mathrm{tr}_n[O] = \sum_{a=1}^{n} O_{aa}$, we can rewrite

$$(*) = \frac{N}{2} \mathrm{tr}_n [Q \cdot i\Lambda]$$

and

$$(**) = -\frac{1}{2}\sum_{ij}\sum_{ab}\psi_i^a\left[1_N\otimes i\Lambda\right]_{ij}^{ab}\psi_j^b \qquad \text{where } 1_N = \overset{N\times N}{\begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \\ & & & \ddots \end{pmatrix}}$$

Moreover, $\sum_{a,b} Q_{ab}^2 = \mathrm{tr}_n[Q^2]$.

(iii) The integral over the $\psi_i^a$ is now gaussian.

Using that for an arbitrary (positive-definite) matrix $K_{ij}$ it holds

$$\int \prod_{i=1}^N dx_i \, e^{-\frac{1}{2}\sum_{ij} x_i K_{ij} x_j} x_\ell x_m = \frac{(K^{-1})_{\ell m}(2\pi)^{N/2}}{|\det K|}$$

and that $\log|\det K| = \mathrm{tr}\log|K|$,

we get:

$$\int \prod_{a=1}^n \prod_{i=1}^N \frac{d\psi_i^a}{\sqrt{2\pi}}\psi_i^1\psi_j^1 e^{-\frac{1}{2}\sum_{a,b}\sum_{ij}\psi_i^a\left[A_{ij}\delta_{ab} + \delta_{ij}(i\Lambda)_{ab}\right]\psi_j^b} =$$

$$= (K^{-1})_{ij}^{11} e^{-\mathrm{tr}\log K} \qquad \text{where } \overset{Nn\times Nn}{K} = A\otimes 1_n + 1_N \otimes i\Lambda$$

Combining everything, one gets the final expression

(iv) The saddle point equations are obtained taking the variation of

$$A_N[Q, i\Lambda] = \frac{\sigma^2}{2} \sum_{a,b} Q_{ab}^2 + \sum_{a,b} (i\Lambda)_{ab} Q_{ab} - \frac{1}{N} \mathrm{Tr} \log \left( A \otimes 1_n + 1_N \otimes i\Lambda \right)$$

$$\frac{\delta A_N}{\delta Q_{ab}} = \sigma^2 Q_{ab} + i\Lambda_{ab} = 0 \qquad \Longrightarrow \quad i\Lambda = -\sigma^2 Q$$

$$\frac{\delta A_N}{\delta \Lambda_{ab}} = Q_{ab} - \frac{1}{N} \mathrm{tr}_N \left( \frac{1}{A \otimes 1_n + 1_N \otimes i\Lambda} \right)_{ab} = 0$$

$$\Longrightarrow Q = \frac{1}{N} \mathrm{tr}_N \left( \frac{1}{A \otimes 1_n + 1_N \otimes i\Lambda} \right) = \frac{1}{N} \mathrm{tr}_N \left( \frac{1}{A \otimes 1_n - \sigma^2 1_N \otimes Q} \right)$$

If $Q = \begin{pmatrix} g & & \\ & \ddots & \\ & & g \end{pmatrix}$, then componentwise:

$$g = \frac{1}{N} \mathrm{tr}_N \left( \frac{1}{z - r w w^T - \sigma^2 g} \right)$$

To compute the trace, one can choose a basis $e_\alpha$ such that $e_1 = w$, $e_\alpha \perp w \;\; \forall \alpha = 2, \dots, N$. Then:

$$g = \frac{1}{N} (N-1) \frac{1}{z - \sigma^2 g} + \frac{1}{N} \frac{1}{z - r - \sigma^2 g} = \frac{1}{z - \sigma^2 g} + \mathcal{O}(1/N)$$

$$\Longrightarrow g_\infty = \frac{1}{z - \sigma^2 g_\infty} \Longrightarrow \sigma^2 g_\infty - z g_\infty + 1 = 0.$$

# Exercise 2 - Solution

## isolated evalue/evector of spiked GOE matrix

(i) One has $(A+uv^T)^{-1} = \left(A\,[1+A^{-1}uv^T]\right)^{-1} = (1+A^{-1}uv^T)^{-1}A^{-1}$

Using the formal expansion:

$$(1+A^{-1}uv^T)^{-1} = 1 - A^{-1}uv^T + A^{-1}uv^T A^{-1}uv^T + \dots$$

leads to

$$(A+uv^T)^{-1} = A^{-1} - A^{-1}uv^T A^{-1} + A^{-1}u\underbrace{(v^T A^{-1} u)}_{\text{number}} v^T A^{-1} + \dots$$

Calling $X = v^T A^{-1} u$ and resumming the series:

$$(A+uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1+X}$$

In the case of the rank-1 perturbation with $\vec{u} = -\frac{r}{N}\vec{w}, \vec{v} = \vec{w}$ and $\hat{A} = z\hat{1} - \hat{J}$ we get

$$\hat{G}_M(z) = (z-\hat{M})^{-1} = \hat{G}_J(z) + r\,\frac{\hat{G}_J(z)\,\vec{w}\vec{w}^T\,\hat{G}_J(z)}{1-r\,\vec{w}.\hat{G}_J(z)\vec{w}} \quad (*)$$

(ii) The eigenvalues of $\hat{M}$ are poles of $\hat{G}_M(z)$.
If $\lambda_{iso}$ is an outlier, it is not a pole of $\hat{G}_J(z)$, because it does not belong to the spectrum of $\hat{J}$ that is the semicircle in $[-2\sigma, 2\sigma]$.

To be a pole of $\hat{g}_M(z)$ and not of $\hat{G}_J(z)$, $\lambda_{iso}$ must be a zero of the denominator of the second term in $(*)$:

$$1 - r\,\vec{w}\cdot\hat{G}_J(\lambda_{iso})\vec{w} = 0$$

(iii) The fact that $\vec{w}$ is "delocalized" in the basis of eigenstates of $\hat{J}$, which I call $\vec{e}_\alpha$, implies that typically

$$(\vec{w}\cdot\vec{e}_\alpha)^2 \sim 1/N \qquad N \gg 1.$$

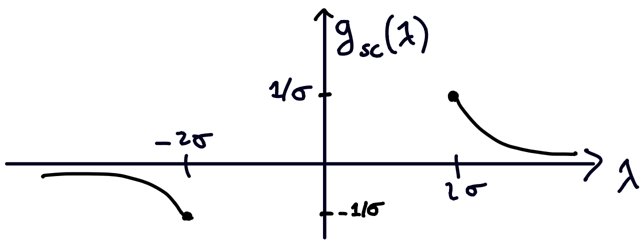
The scalar product $\vec{w}\cdot\hat{G}_J(\lambda)\vec{w}$ can be expanded in the eigenbasis of $\hat{J}$, and one gets:

$$\vec{w}\cdot\hat{G}_J(\lambda)\cdot\vec{w} = \sum_{\beta=1}^{N}(\vec{e}_\beta\cdot\vec{w})^2\left[\hat{G}_J(z)\right]_{\beta\beta} \overset{N\gg 1}{\sim} \frac{1}{N}\sum_{\beta=1}^{N}\left[\hat{G}_J(z)\right]_{\beta\beta}$$

The last term is the normalized trace of the resolvent, i.e. the Stieltjes transform. Therefore:

$$\lim_{N\to\infty} \vec{w}\cdot\hat{G}_J(\lambda)\cdot\vec{w} = g_{sc}(\lambda)$$

(iv) The function $g_{sc}(\lambda)$ has the following behavior on the real axis:



The function is invertible only if $y \in [-1/\sigma, 1/\sigma]$.

The expression for $g_{sc}^{-1}$ can be more easily obtained from the self-consistent equation:

$$\sigma^2 g_{sc}^2(z) - z\, g_{sc}(z) - 1 = 0$$

$$\Longrightarrow z = \sigma^2 g_{sc}(z) + \frac{1}{g_{sc}(z)} \Rightarrow g_{sc}^{-1}(y) = \sigma^2 y + 1/y$$

The equation for $\lambda_{iso}$ reads: $g_{sc}(\lambda_{iso}) = 1/r$.

It admits a solution only for $1/r \in [-\frac{1}{\sigma}, \frac{1}{\sigma}]$.

Meaning that $r \geqslant \sigma$ for $r > 0$.

In this case, $\lambda_{iso} = g_{sc}^{-1}(1/r) = \frac{\sigma^2}{r} + r$

(v) Using the decomposition of $g_{\hat{M}}$ in its eigenbasis $(\lambda_\alpha, \vec{u}_\alpha)_{\alpha=1}^N$

$$\hat{g}_M(z) = \sum_{\beta=1}^N \frac{\vec{u}_\beta \vec{u}_\beta^T}{z - \lambda_\beta} \Rightarrow \vec{w}^T \hat{g}_M \vec{w} = \sum_{\beta=1}^N \frac{\xi_\beta^2}{z - \lambda_\beta}$$

Then obviously if $z \to \lambda_\alpha$ is an isolated pole,

$$\xi_\alpha^2 = \lim_{\lambda \to \lambda_\alpha} \sum_{\beta=1}^N \frac{(\lambda - \lambda_\alpha)}{(\lambda - \lambda_\beta)} \xi_\beta^2$$

We use again the expression (*). Since $\lambda_{iso}$ is not a pole of $\hat{J}$, the first term will not contribute to the residue and so:

$$\xi_N = \lim_{\lambda \to \lambda_{iso}} \frac{r\left(\vec{w}^T \hat{G}_J(\lambda) \cdot \vec{w}\right)^2}{1 - r\, \vec{w} \cdot \hat{G}_J(\lambda) \vec{w}} (\lambda - \lambda_{iso})$$

$N \gg 1$

$$\simeq \lim_{\lambda \to \lambda_{iso}} (\lambda - \lambda_{iso}) \frac{r\, g_{sc}^2(\lambda)}{1 - r\, g_{sc}(\lambda)}$$

$$= \lim_{\lambda \to \lambda_{iso}} \frac{(\lambda - \lambda_{iso})}{1 - r g_{sc}(\lambda)} \cdot g_{sc}(\lambda_{iso})$$

When $\lambda \to \lambda_{iso}$, $1 - r g_{sc}(\lambda) \to 0$ and thus the limit gives 0/0: One has to compute it by taking the derivative of both numerator & denominator

$$\lim_{\lambda \to \lambda_{iso}} \frac{(\lambda - \lambda_{iso})}{1 - r g_{sc}(\lambda)} g_{sc}(\lambda_{iso}) = g_{sc}(\lambda_{iso}) \lim_{\lambda \to \lambda_{iso}} \frac{-1}{r g'_{sc}(\lambda)}$$

Using that $g_{sc}(\lambda_{iso}) = 1/r$, one gets: $\xi_N = -\dfrac{1}{r^2 g'_{sc}(\lambda_{iso})}$

To make this more explicit,
convenient to take the self-consistent eq. for $g_{sc}(\lambda)$ and derive it:

$$2\sigma^2 g'_{sc}(z)\, g_{sc}(z) - g_{sc}(z) - z\, g'_{sc}(z) = 0$$

$$(2\sigma^2 g_{sc} - z)\, g'_{sc} = g_{sc} \Rightarrow \frac{g_{sc}}{g'_{sc}} = 2\sigma^2 g_{sc} - z$$

At $z = \lambda_{iso}$,

$$\xi_N = -\frac{1}{r}\left(\frac{g_{sc}}{g'_{sc}}\right) = -\frac{1}{r}\left(2\sigma^2 g_{sc}(\lambda_{iso}) - \lambda_{iso}\right)$$

$$= \frac{-2\sigma^2}{r^2} + \frac{1}{r}\left(\frac{\sigma^2}{r^2} + r\right) = 1 - \sigma^2/r^2$$

# Exercise 3 - Solution

## Thermodynamics and the condensation transition

(i) One has:

$$Z_\beta = \int_{S_N(\sqrt{N})} d\vec{s}\; e^{\frac{\beta}{2}\sum_{i,j=1}^{N} s_i \underbrace{(\gamma_{ij} + r\, v_i v_j)}_{M_{ij}} s_j} = \int d\vec{s}\, d\lambda\; e^{\beta/2 \sum_{ij} s_i M_{ij} s_j - \frac{\beta\lambda}{2}\left(\sum_i s_i^2 - N\right)}$$

implement spherical constraint.

Performing the change of basis, one gets:

$$Z_\beta = \int d\lambda \int \prod_{\alpha=1}^{N} ds_\alpha \;\; e^{\frac{\beta}{2}\sum_\alpha \lambda_\alpha s_\alpha^2 - \frac{\beta\lambda}{2}\left(\sum_\alpha s_\alpha^2 - N\right)}$$

(ii) The average:

$$\langle s_\gamma^2 \rangle = \frac{1}{Z_\beta}\int d\lambda\, e^{\frac{\beta\lambda N}{2}} \int \prod_{\alpha\neq\gamma} e^{\frac{\beta}{2}\lambda_\alpha s_\alpha^2 - \frac{\beta\lambda}{2} s_\alpha^2}\, ds_\alpha \int ds_\gamma\, s_\gamma^2\, e^{\frac{\beta}{2}\lambda_\gamma s_\gamma^2 - \frac{\beta\lambda}{2} s_\gamma^2}$$

$$= \frac{1}{Z_\beta}\int d\lambda\; e^{+\frac{\beta\lambda N}{2}} \left(\frac{2\pi}{\beta}\right)^{N/2} \left(\prod_{\alpha=1}^{N} \frac{1}{\lambda-\lambda_\alpha}\right)^{1/2} \frac{1}{\beta(\lambda-\lambda_\gamma)} \qquad (*)$$

Assuming $\lambda > \lambda_\alpha \;\; \forall \alpha$.

The integral over $\lambda$ can be performed with a saddle point when $N \gg 1$, optimizing

$$f(\lambda) = \lambda\beta - \frac{1}{N}\sum_{\alpha=1}^{N} \log(\lambda - \lambda_\alpha)$$

$$f'(\lambda)\Big|_{\lambda=\lambda^*} = 0 \Rightarrow \beta = \frac{1}{N}\sum_{\alpha=1}^{N} \frac{1}{\lambda^* - \lambda_\alpha}$$

Plugging this in (*) and simplifying the exponential terms in numerator with those in $Z_\beta$, one gets

$$\langle S_\gamma^2 \rangle = \frac{1}{\beta(\lambda^* - \lambda_\gamma)}$$

with $\lambda^*$ solving:

$$N = \sum_{\gamma=1}^{N} \frac{1}{\beta(\lambda^* - \lambda_\gamma)} = \sum_{\gamma=1}^{N} \langle S_\gamma^2 \rangle$$

(iii) For $r < r_c = \sigma$, there is no isolated eigenvalue and the spectrum of $\hat{M}$ has an eigenvalue density that tends to the semicircle $\rho_{sc}(\lambda)$ when $N \to \infty$. Thus:

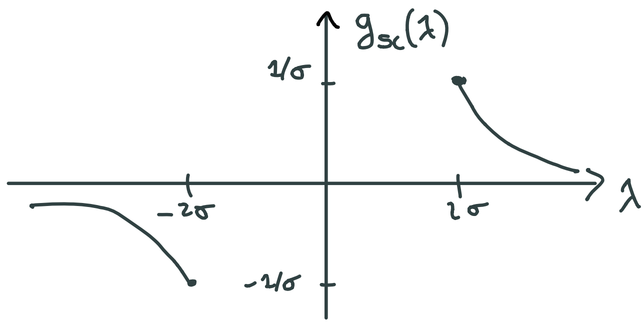
$$\frac{1}{N}\sum_\gamma \frac{1}{\lambda^* - \lambda_\gamma} \overset{N \gg 1}{\sim} \int d\lambda \frac{\rho_{sc}(\lambda)}{\lambda^* - \lambda_\gamma} = g_{sc}(\lambda^*)$$

The equation for $\lambda^*$ becomes:

$$g_{sc}(\lambda^*) = \beta \quad \text{for } \lambda^* > \lambda_N = 2\sigma$$

This can be solved only for $\beta < \beta_c = 1/\sigma$, and in this case

$$\lambda^* = \frac{1}{\beta} + \sigma^2 \beta$$




At $\beta \to \beta_c$, $\lambda^* \to 2\sigma$, that is the boundary value of the domain where the saddle point can be taken. For $\beta > \beta_c$, the saddle point sticks to the boundary: $\lambda^* = 2\sigma$

This is a freezing transition: it signals the transition to a glass phase.

Then the equation for $\lambda^*$ is solved assuming condensation in the lowest energy mode

$$\langle S_N^2 \rangle \sim \mathcal{O}(N)$$

In particular: $1 = \frac{1}{\beta} \underbrace{g_{sc}(\lambda^* = 2\sigma)}_{1/\sigma} + \frac{1}{N}\langle S_N^2 \rangle$

$$\Longrightarrow \frac{1}{N}\langle S_N^2 \rangle = 1 - 1/\sigma\beta .$$

(iv) For $r > r_c = \sigma$, $\lambda_N = \lambda_{iso} = \frac{\sigma^2}{r} + r > 2\sigma$. is the maximal value that $\lambda^*$ can take. Since $g_{sc}(\lambda)$ is monotonically decreasing, the maximal $\beta$ for which a solution to $\beta = g_{sc}(\lambda^*)$ can be found is the $\beta$ such that: $\beta = g_{sc}(\lambda_{iso})$

Recalling that $g_{sc}(\lambda_{iso}) = 1/r$, one has $\beta_c = 1/r$.

For $\beta > \beta_c$, it must hold:

$$1 = \frac{1}{\beta} g_{sc}(\lambda_{iso}) + \frac{1}{N}\langle S_N^2 \rangle \Rightarrow \frac{1}{N}\langle S_N^2 \rangle = 1 - 1/\beta r .$$

Phase transitions in temperature: