Introduction to random matrices and tensors and application to Principal Component Analysis

Vasily Sazonov

University of Graz, Austria

FWF
Der Wissenschaftsfonds.

Matrices are everywhere in physics, randomness is also everywhere.

Random matrices first appeared in statistics, generalizing the chi-squared law to multivariate random data (Wishart, 1928) and are applicable to:

- the description of the energy spectra of heavy nuclei (Wigner, 1950s),
- the large N_c limit of QCD ('t Hooft, 1970s),
- random surfaces, 2d quantum gravity
- data analysis
- transport in disordered systems
- string theory
- number theory
- biology
- ....

Matrices are everywhere in physics, randomness is also everywhere.

Random matrices first appeared in statistics, generalizing the chi-squared law to multivariate random data (Wishart, 1928) and are applicable to:

- the description of the energy spectra of heavy nuclei (Wigner, 1950s),
- the large N_c limit of QCD ('t Hooft, 1970s),
- random surfaces, 2d quantum gravity
- data analysis
- transport in disordered systems
- string theory
- number theory
- biology
- ....

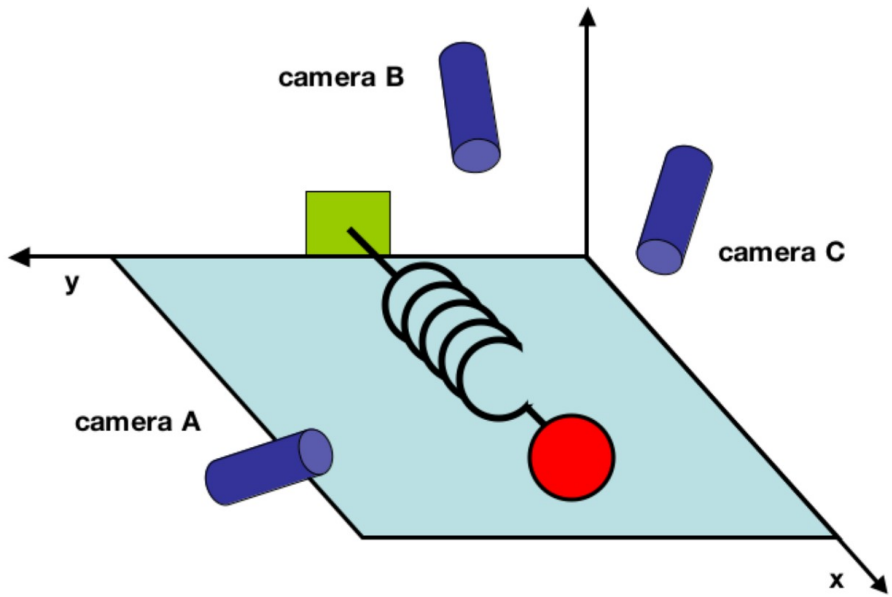
References

- General Introduction to Random Matrices:
  - G. Livan, M. Novaes, P. Vivo, Introduction to Random Matrices: Theory and Practice, Springer, (2018)
  - Marino, M., Chern-Simons theory, matrix models, and topological strings, Int.Ser.Monogr.Phys. 131 (2005) 1-197
  - Wikipedia
- Spiked Matrix Model:
  - D. Paul, A. Aue, Random matrix theory in statistics: A review, Journal of Statistical Planning and Inference, 150 (2014), 1-29
  - E. Brézin, S. Hikami, Random Matrix Theory with an External Source.
  - V. Rivasseau, talk at 2dn, OIST Mini Symposium "Holographic Tensors";
- Random Tensors and Spiked Tensor Model:
  - R. Gurau, Random tensors, Oxford, (2017)
  - Valentina Ros, Gerard Ben Arous, Giulio Biroli, Chiara Cammarota, Complex energy landscapes in spiked-tensor and simple glassy models: ruggedness, arrangements of local minima and phase transitions, Phys. Rev. X 9, 011003 (2019)

In the following the pictures, formulas and the logic of the material presenting are adopted from these sources without further referring to them.

Principal component analysis (PCA)

In experiments scientists frequently face with the clouded, unclear and even redundant data. This is not a trivial problem, but rather a fundamental obstacle in empirical science. Consider, the following toy example:



Principal component analysis (PCA)

- Clearly, in the example above the positions of cameras were not chosen appropriately.
- To describe the system we need measurements just in one dimension instead of 2×3.
- In addition, the signal can be contaminated by noise.

To fight these problems PCA provides:

- a simple, non-parametric method of extracting relevant information from confusing data sets
- computing the most meaningful basis to re-express a noisy data set.

Principal component analysis (PCA)

- Consider a $N \times M$ matrix X representing the initial data, where N is the number of measurement types and M is the number of samples.
- The corresponding covariance matrix

$$C_X := \frac{1}{N-1} X X^T$$

  is a square symmetric $N \times N$ matrix and contains:
  - the variance of particular measurement types in the diagonal terms;
  - the covariance between measurement types in the off-diagonal terms.
- In the diagonal terms, by assumption, large (small) values correspond to interesting dynamics (or noise).
- In the off-diagonal terms large (small) values correspond to high (low) redundancy.

Principal component analysis (PCA)

- We transformation the initial data X by some orthonormal matrix P

$$Y = PX$$

- In such a way that the covariance

$$C_Y := \frac{1}{N-1} YY^T$$

is diagonalized.

Principal component analysis (PCA)

- The covariance matrix can be written as

$$C_Y := \frac{1}{N-1} P(XX^T)P^T$$

  Consequently:
  - the variance of particular measurement types in the diagonal terms;
  - the covariance between measurement types in the off-diagonal terms.
- The principal components of X are the eigenvectors of XX^T, or the rows of P.

PCA: singular value decomposition

- The eigenvalue decomposition of XX^T reduces to the singular value decomposition of matrix X.
- $\sigma \geq 0$ is an singular value of matrix X, if there exist left and right singular vectors a_σ and b_σ, such that

$$\mathrm{X}\, a_\sigma = \sigma b_\sigma^T; \;\; b_\sigma\, \mathrm{X} = \sigma a_\sigma^T$$

- This means, we have to find the best low rank approximation for the matrix X

$$X_{ij} = a_i \otimes b_j$$

Random matrix model for PCA

- The data can be contaminated by noise

$$X_{ij} = a_i \otimes b_j + \eta H_{ij}$$

  with H_{ij} being noise and η being the noise to signal ratio.
- The term ηH_{ij} can be modeled by the RANDOM MATRIX
- Is it always possible to detect the signal?
- What is the critical ratio η_c

Ensembles of Random Matrices

The most studied random matrix ensembles are the Gaussian ensembles:

- The Gaussian unitary ensemble GUE(N) is described by the Gaussian measure

$$\frac{1}{Z_{\mathrm{GUE}(N)}} e^{-\frac{N}{2}\mathrm{tr}H^2}$$

here H is Hermitian matrix $N \times N$. The Gaussian measure is invariant under the unitary transformations:

$$H \to UHU^{\dagger}$$

- The Gaussian orthogonal ensemble GOE(N) is described by the Gaussian measure invariant under orthogonal conjugation

$$\frac{1}{Z_{\mathrm{GOE}(N)}} e^{-\frac{N}{4}\mathrm{tr}H^2}, \quad H \to \mathcal{O}H\mathcal{O}^{T}$$

- The Gaussian symplectic ensemble GSE(N) is defined by the Gaussian measure on the space Hermitian quaternionic matrices.

Eigenvalues

For many purposes it is highly important to know the statistics of eigenvalues. The first step is to write probability density in terms of eigenvalues.
The joint probability density for the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_N$ of GUE/GOE/GSE is given by

$$\frac{1}{Z_{\beta,N}} \prod_{k=1}^{N} e^{-\frac{\beta N}{4}\lambda_k^2} \prod_{i<j} |\lambda_j - \lambda_i|^\beta$$

where the Dyson index, $\beta = 1$ for GOE, $\beta = 2$ for GUE, and $\beta = 4$ for GSE, counts the number of real components per matrix element.
The Vandermonde determinant $|\lambda_j - \lambda_i|^\beta$ comes from the Jacobian of the change of variables like $H \to U^\dagger \Lambda U$.

Meaning of the eigenvalue distribution

One can understand the ensembles of eigenvalues of Gaussian matrices as a Coulomb gas. Indeed, we can rewrite the joined probability density as

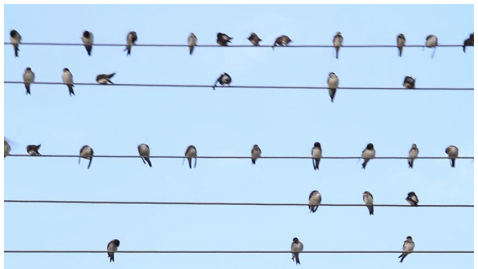
$$\frac{1}{Z_{\beta,N}} \prod_{k=1}^{N} e^{-\beta N^2 \mathcal{V}(\lambda)},$$

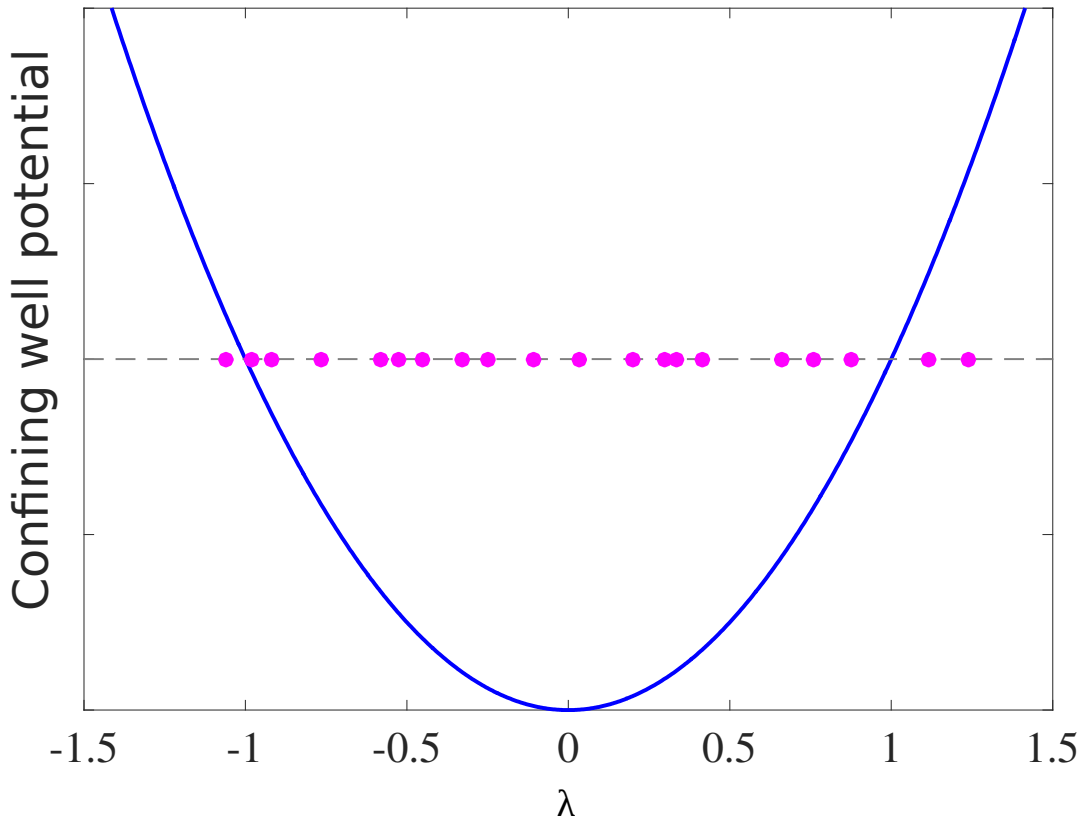
where the *energy* term in the exponent is

$$\mathcal{V}(\lambda) := \frac{1}{4N} \sum_i \lambda_i^2 - \frac{1}{2N^2} \sum_{i<j} \log |\lambda_j - \lambda_i|$$

The Gibbs-Boltzman weight $e^{-\beta N^2 \mathcal{V}(\lambda)}$ corresponds to a thermodynamical fluid of particles with positions $\{\lambda_1, ...\lambda_N\}$ on a line, in equilibrium at "inverse temperature" β under the effect of competing interactions: a quadratic single particle potential and repulsive all-to-all logarithmic term.

Meaning of the eigenvalue distribution





Distance between eigenvalues - Wigner's surmise

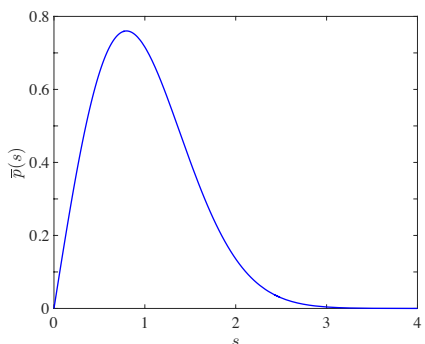
Consider a 2×2 GOE matrix $H_s = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}$ What is the pdf $p(s)$ of the spacing $s = \lambda_2 - \lambda_1$ between its two eigenvalues $(\lambda_2 > \lambda_1)$? λ_1, λ_2 are the roots of the characteristic polynomial and

$$\lambda_{1,2} = \left(x_1 + x_2 \pm \sqrt{(x_1 - x_2)^2 + 4x_3^2} \right) /2 \,, \quad s = \sqrt{(x_1 - x_2)^2 + 4x_3^2}$$

By definition, we have

$$p(s) = \int_{-\infty}^{\infty} dx_1 dx_2 dx_3 \frac{e^{-\frac{1}{2}x_1^2}}{\sqrt{2\pi}} \frac{e^{-\frac{1}{2}x_2^2}}{\sqrt{2\pi}} \frac{e^{-x_3^2}}{\sqrt{\pi}} \delta \left(s - \sqrt{(x_1 - x_2)^2 + 4x_3^2} \right)$$

$$= \boxed{\frac{s}{2} e^{-s^2/4}}$$



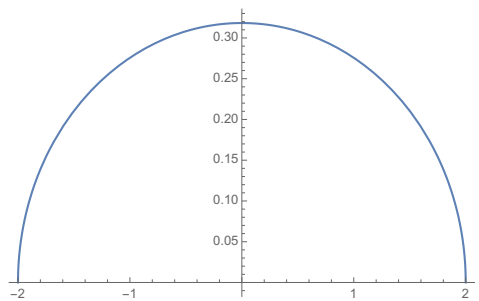
Wigner's semicircle law

Consider an ensemble of random matrices H, then the *average spectral density* defined as

$$\rho(z) = \langle \frac{1}{N} \mathrm{Tr} \delta(H - z) \rangle$$

in the large N limit is given by the semicircle (semi-ellipse) distribution

$$\rho_{SC} = \frac{1_{z \in [-2,2]}}{2\pi} \sqrt{4 - z^2}$$



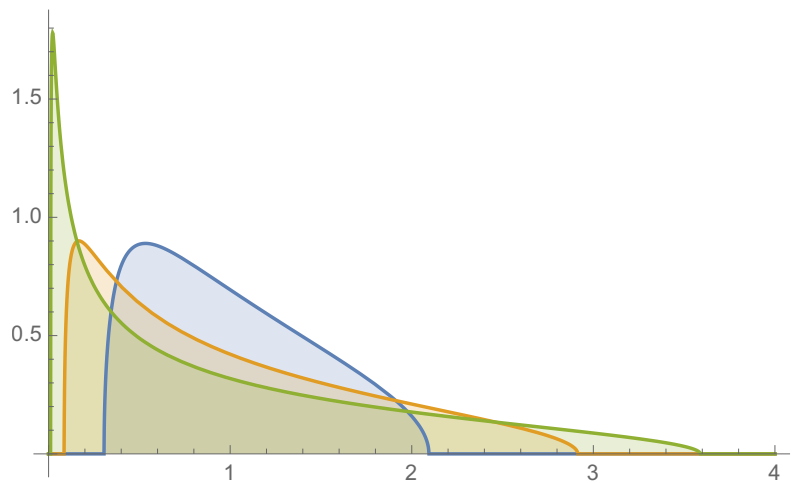
We will see more details about it considering the Spiked Matrix Model.

Marchenko–Pastur distribution

Marchenko–Pastur distribution is the limiting distribution of eigenvalues of Wishart matrices

$$Y_N = \frac{1}{N} X X^\dagger$$

constructed from $N \times M$ matrices X as both the matrix dimensions tend to infinity with ratio $\frac{N}{M} = \lambda$. The probability density function is well-defined for $\lambda \leq 1$. When $\lambda > 1$, $N - M$ eigenvalues of Y a zero.
The plot is for $\lambda = 0.2, 0.5, 0.8$.



The Spiked Matrix Model

- Signal: a deterministic rank one matrix $H_0 = \lambda u_i \bar{u}_j$ where $u \in \mathbb{C}^N$ is a normalized vector $\|u\| = 1$
- Noise: a N by N GUE random matrix H

We observe $M = H_0 + H$ and the goal is to recover u. λ allows to control the noise: λ large means small noise and vice-versa.

In practice in data analysis rectangular real matrices typically occur. So we have to use GOE's instead of GUE's, singular values and Marcenko-Pastur distribution instead of Wigner-Dyson etc... but this is detail.

Wigner-Dyson and Tracy-Widom laws for $\lambda = 0$

The spectral density of M is at leading order in $1/N$ the Wigner semi-circle law $\rho(z) = \frac{1_{z\in[-2,2]}}{2\pi}.\sqrt{4-z^2}$ with edge spectrum at $|z| = 2$.

Almost surely the largest eigenvalue is $\lambda_c = 2$ and the distribution of λ_c is the Tracy-Widom law

$$\lim_{N\to\infty} P(N^{2/3}\lambda_c \leq x) = F_2^{TW}(x).$$

Rank-one perturbation

At $\lambda \neq 0$ the $N \to \infty$ limit of the spectral density of M always remains the Wigner semi-circle law. However the statistics of the largest eigenvalue λ_c undergoes a sharp phase transition:

- if $\lambda < 1$ essentially nothing changes for the value λ_c which stays a.s. at the edge 2 of the Wigner semi-circle with a Tracy-Widom distribution.
- if $\lambda > 1$, we have $\lambda_c(\lambda) = \lambda + \frac{1}{\lambda}$, and the statistics of λ_c becomes that of a Gaussian error function.

Computation of λ_c, I

This type of results have a long history relying in particular on the Harish-Chandra-Itzykson-Zuber formula. Recently mathematical theorems by Ben Arous, Guionnet, Maïda, Péché and many others have established rigorously and in great detail many variants.

Here I shall give only a short version, namely the computation which shows that $\lambda_c(\lambda) = \lambda + \frac{1}{\lambda}$ (and not $\lambda_c = \lambda$).

Computation of λ_c, II

The normalized density of states is $\rho(z) = \frac{1}{\pi} \lim_{\epsilon \to 0} \Im G(z + i\epsilon)$, where

$$G(z) := \langle \frac{1}{N} \mathrm{Tr} \frac{1}{z - H - H_0} \rangle_H$$

Representing $G(z)$ as Fourier transform of the 'evolution' operator $U_{H_0}(t) := \langle e^{itH_0} \rangle_H$ like

$$G(z) = \int_0^\infty dt\, e^{-itz} U_{H_0}(t)\,,$$

employing Harish-Chandra-Itzykson-Zuber formula for integrating out gauge degrees of freedom, rewriting the result as contour integral,

Computation of λ_c, III

we arriwe at

$$\frac{\partial G}{\partial z} = \oint \frac{du}{2\pi i} \frac{1}{u + G_0(u) - z}$$

We have now to specify the contour of integration in the complex u-plane. It surrounds all the eigenvalues of H_0 and we have to determine the location of the zeroes of the denominator with respect to this contour. Let us return to the discrete form for the equation

$$u + G_0(u) = z$$

i.e.

$$u + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{u - \epsilon_i} = z$$

which posseses $(N + 1)$ real or complex roots in the u-plane. For z real and large, N of these roots are close to the ϵ_i and one, which will be denoted $\hat{u}(z)$, goes to infinity with z as

$$\hat{u}(z) = z - \frac{1}{z} + O\left(\frac{1}{z^2}\right)$$

Computation of λ_c, IV

For large z, the contour encloses all the roots of

$$u + \frac{1}{N}\sum_{i}^{N}\frac{1}{u-\epsilon_i} = z$$

except $\hat{u}(z)$. When z decreases the contour should not be crossed by any other root of the equation, therefore it is defined by the requirement that only one root remains at its exterior. Then it is easier to calculate the integral for $\frac{\partial G}{\partial z}$ by taking the residues of the singularities outside of the contour, rather than the N poles enclosed by this contour. There are two of them outside; one is $\hat{u}(z)$ and the other one is at infinity (since for large u, $G_0(u)$ vanishes). Taking these two singularities we obtain

$$\frac{\partial G}{\partial z} = 1 - \frac{1}{1+\frac{dG_0}{d\hat{u}(z)}} = 1 - \frac{d\hat{u}(z)}{dz}$$

The integration gives

$$G(z) = z - \hat{u}(z)$$

Computation of λ_c, V

- For $\lambda = 0$ we have $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{u-\varepsilon_i} = \frac{1}{u}$, hence the N-independent equation

$$\hat{u}_0 + \frac{1}{\hat{u}_0} = z <=> \hat{u}_0^2 - z\hat{u}_0 + 1 = 0.$$

  It solves to the Catalan function $\hat{u}_0 = \frac{1}{2}(z + \sqrt{z^2-4})$ so that $G(z) = z - \hat{u}_0 = \frac{1}{2}(z - \sqrt{z^2-4})$, from which Wigner's law easily follows.

- For $\lambda \neq 0$ we have $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{u-\varepsilon_i} = \frac{1}{u} + \frac{1}{N}[\frac{1}{u-\lambda} - \frac{1}{u}]$, hence the equation

$$\hat{u} + \frac{1}{\hat{u}} + \frac{\lambda}{N\hat{u}(\hat{u}-\lambda)} = z.$$

  Expanding $\hat{u} = \hat{u}_0 + \frac{1}{N}\hat{u}_1$, and defining $F(x) = x^2 - zx + 1$, we have to solve for

$$F(\hat{u}_0 + \frac{\hat{u}_1}{N}) + \frac{\lambda}{N(\hat{u}_0 + \frac{\hat{u}_1}{N} - \lambda)} = 0$$

Computation of λ_c, VI

At first order in $1/N$ it gives

$$F(\hat{u}_0) = 0, \quad F'(\hat{u}_0)\frac{\hat{u}_1}{N} = -\frac{\lambda}{N(\hat{u}_0 - \lambda)}$$

and since $F'(x) = 2x - z$ and $\hat{u}_0 = \frac{1}{2}(z + \sqrt{z^2 - 4})$ we get

$$\hat{u}_1 = -\frac{\lambda}{(2\hat{u}_0 - z)(\hat{u}_0 - \lambda)} = -\frac{2\lambda}{\sqrt{z^2 - 4}(z + \sqrt{z^2 - 4} - 2\lambda)}$$

Remark that the denominator can have only a single zero $\lambda_c > 2$, and only for $\lambda > 1$, at $\lambda_c = \lambda + \frac{1}{\lambda}$, since then $\sqrt{z^2 - 4}|_{z=\lambda_c} = \lambda - \frac{1}{\lambda} > 0$.

Moreover the residue is $-\frac{2\lambda}{\sqrt{z^2-4}+z}|_{z=\lambda_c} = -1$, so that the residue of $G = z - \hat{u}$ is $+1$. It proves that $\lambda > 1$ is the threshold at which a single eigenvalue $\lambda_c = \lambda + \frac{1}{\lambda}$ pops out of Wigner's semi-circle law.

Data recovery

Data recovery uses the random flow governed by the Hamiltonian or cost-function $f =< v, Mv >$.

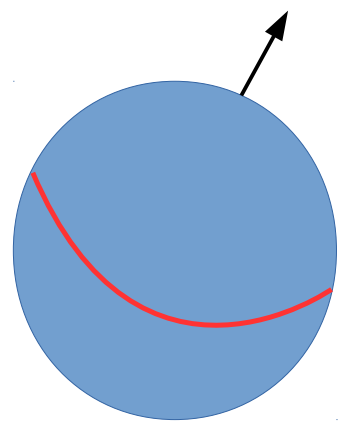
The corresponding flow $\dot{v} = Mv$ is linear with exact solution $v = e^{Mt} v_0$.

Following the previous analysis, this flow also undergoes a sharp transition between two regimes:

- if $\lambda > 1$ iterative methods (raising M to a high power) detects quickly the highest eigenvalue λ_c and its eigenvector u_c. The latter has $O(1)$ overlap with the true signal initial u, allowing easily its recovery.
- if $\lambda < 1$ recovery is essentially impossible in practice. Indeed the random eigenvalues of H essentially translate into uninformative critical points of f, and for N large the flow remains for a hopelessly long time trapped around the equator E_N of the S_N sphere orthogonal to u (unless further constraints on u are known)

Spiked tensor model

- The spiked tensor model was intensively studied in the recent time by Richard E., Montanari A., Perry A., Wein A., Bandeira A., Lesieur T., Miolane L., Lelarge M., Krzakala F., Zdeborova L, Ben Arous G.,....
- It is build for estimating a rank-one deformation of a symmetric random Gaussian tensor $T = rx^{\otimes p} + J$, $\sum x_i^2 = N$ with the Hamiltonian (cost function) defined as $H = -(T, \phi^{\otimes p})$, $\sum \phi_i^2 = N$.
- The problem can be viewed as finding of the spike on the multidimensional 'spiky' sphere.
- When $N \to \infty$, infinitely many components of ϕ_i are orthogonal to the spike x_i.
- Increasing p, we get ∂H,..,$\partial^{(p-1)} H = 0$.



Spiked Tensor Model

- Two possible goals:
  - to detect the existence of the spike;
  - to estimate it.
- The current knowledge can be roughly summarized as:
  - An ideal estimator achieves strictly positive correlation with the unknown vector starting from a critical signal-to-noise ratio $r_1 = O(1)$.
  - At the same time no polynomial-time algorithm is known providing the unknown vector unless $r_2 \geq CN^{(p-2)/4}$.
- The Langevin dynamics and gradient descent belong to the most popular algorithms for recovering unknown vector.
- Here we study the dynamics with random initial conditions.

Dynamics with random initial conditions

Consider the equation

$$\partial_t \phi_i(t) = -\frac{\delta H}{\delta \phi_i}(t)\,, \quad \phi_i(t_0) = \varphi_i\,,$$

The statistics of initial conditions is given by

$$C_{ij} = \langle \varphi_i \varphi_j \rangle_\varphi = \delta_{ij} \chi_N(j),$$

$$\sum_{j=1}^{N} \chi_N(j) = N\,,$$

from where follows the spherical constraint (in average) for the initial conditions:

$$\sum_{j=1}^{N} \langle \varphi_j^2 \rangle = N\,.$$

We would like to use the standard Martin-Sigga-Rose procedure to study the correlation functions of the process, but there is no explicit dependence on the initial conditions φ_i.

Random initial conditions

We equivalently rewrite the differential equation in the integral form as

$$\phi_i(t) = \varphi_i - \int_{t_0}^{t} d\tau \, \frac{\delta H}{\delta \phi_i}(\tau) \, .$$

If $\phi^0(t)$ is a solution, we can formally write the average of some operator $O[\phi^0]$ as

$$\begin{aligned} \langle O(\phi^0) \rangle_\varphi &= \langle \int D\phi \, O[\phi] \, \delta(\phi - \phi^0) \rangle_\varphi \\ &= \langle \int D\phi \int D\hat{\phi} \int D\bar{\xi} \int D\xi \, O[\phi] \, e^{-S_1} \rangle_\varphi \, , \end{aligned}$$

where

$$\begin{aligned} S_1 \quad := \quad & \int_{t_0}^{t_e} dt \left(i\hat{\phi}_i(t) \Big[\phi_i(t) - \int_{t_0}^{t} d\tau \, \frac{\delta H}{\delta \phi_i}(\tau) - \varphi_i \Big] \right) \\ - \quad & \int_{t_0}^{t_e} dt \int_{t_0}^{t} dt' \, \bar{\xi}_i(t) \Big[\frac{\delta^2 H}{\delta \phi_i \delta \phi_j}(t') \Big] \xi_j(t') \end{aligned}$$

Random initial conditions

Averaging over initial conditions gives

$$\langle O \rangle_{\varphi} = \int D\phi \int D\hat{\phi} \int D\bar{\xi} \int D\xi \, O \, e^{-S_2}$$

with

$$\begin{aligned} S_2 \quad := \quad & \sum_{i=1}^{N} \frac{\chi_N(i)}{2} \Big(\int_{t_0}^{t_e} dt \, \hat{\phi}_i(t) \Big)^2 - i \int_{t_0}^{t_e} dt \, \hat{\phi}_i(t) \phi_i(t) \\ - \quad & i \int_{t_0}^{t_e} dt \int_{t_0}^{t} dt' \, \hat{\phi}_i \frac{\delta H}{\delta \phi_i}(t') \\ - \quad & \int_{t_0}^{t_e} dt \int_{t_0}^{t} dt' \, \bar{\xi}_i(t) \Big[\frac{\delta^2 H}{\delta \phi_i \delta \phi_j}(t') \Big] \xi_j(t') \, . \end{aligned}$$

Super-field notations

Let us define a superfield, bi-local in time

$$\Phi_i(\bar{\theta}, \theta, t, t') := \phi_i(t') + \bar{\theta}\xi_i(t') + \bar{\xi}_i(t)\theta + i\hat{\phi}_i(t)\bar{\theta}\theta$$

and the integration in the superspace as

$$\int d1 := \int_{t_0}^{t_e} dt_1 \int_{t_0}^{t_1} dt_1' \int d\theta_1 \int d\bar{\theta}_1 \,,$$

$$\bar{\theta}\theta = -\theta\bar{\theta}\,, \quad \int d\theta = \int d\bar{\theta} = 0\,, \quad \int d\theta\theta = \int d\bar{\theta}\bar{\theta} = 1\,.$$

Super-field notations

Then, everything what is not a part of the gradient term, can be schematically grouped into the term

$$\sum_{i,j=1}^{N} \int d1 \int d2\, \Phi_i(1) K_{ij}(1,2) \Phi_j(2)\,,$$

The sum of other terms in the action S_2 is equal to

$$\int d1\, H(\Phi(1))\,.$$

The translation in the variable $\bar{\theta} \to \bar{\theta} + \epsilon$ generates a BRST-type of SUSY.

Spiked tensor model

The spiked tensor model is defined by the Hamiltonian

$$\begin{aligned} H &:= -\sum_{i_1<..<i_p} J_{i_1..i_p} \cdot \phi_{i_1}\phi_{i_2} \cdot ... \cdot \phi_{i_p} \\ &- \frac{r}{2N^{p-1}} \sum_{i_1,..,i_p} x_{i_1} \cdot ... \cdot x_{i_p} \cdot \phi_{i_1} \cdot ... \cdot \phi_{i_p} + \frac{\mu}{2}\Big(\sum_{i=1}^{N} \phi_i^2 - N\Big)\,, \end{aligned}$$

where x_i is a signal and $\frac{r}{2N^{p-1}}$ is its strength, $\sum_{i=1}^{N} x_i^2 = N$, μ is a Lagrange multiplier implementing the spherical constraint $\sum_{i=1}^{N} \phi_i^2 = N$, $J_{i_1..i_p}$ is a symmetric tensor noise with Gaussian statistics

$$\overline{J^2_{i_1..i_p}} = \frac{J^2 p!}{2N^{p-1}}\,.$$

Spiked tensor model: super-field notations

The average of an operator O in superspace notations reads as

$$\begin{aligned}\langle O \rangle &= \int D\Phi\, O\, \exp\Big[- \int d1 \int d2 \sum_{i=1}^{N} \Phi_i(1) G_0^{-1}(1,2) \Phi_i(2) \\ &- H_J[\Phi] - H_x[\Phi] \Big] ,\end{aligned}$$

where

$$\begin{aligned} G_0^{-1}(1,2) &:= \Big[\frac{1}{(t_1 - t_0)(t_2 - t_0)} \frac{\chi_N(i)}{2} \\ &+ \delta(1-2)\delta(t_1 - t_1')\theta_1 \partial_{\theta_2} + \delta(1-2)\frac{\mu(t_1')}{2} \Big] , \end{aligned}$$

Spiked tensor model: super-field notations

H_J is the disordered Hamiltonian

$$H_J := \int d1 \sum_{i_1<..<i_p} J_{i_1..i_p} \Phi_{i_1}(1)\Phi_{i_2}(1) \cdot ... \cdot \Phi_{i_p}(1)\,,$$

the term H_x contains the signal part

$$\begin{aligned} H_x &:= \frac{r}{2N^{p-1}} \int d1 \sum_{i_1,..,i_p} x_{i_1} x_{i_2} \cdot ... \cdot x_{i_p} \Phi_{i_1}(1)\Phi_{i_2}(1) \cdot ... \cdot \Phi_{i_p}(1) \\ &= \frac{r}{2} N \int d1 \Big(\frac{1}{N} \sum_i x_i \Phi_i(1)\Big)^p . \end{aligned}$$

Integrating out the disorder leads to the substitution of the term H_J by

$$H_{\int J} := \frac{J^2 p!}{2N^{p-1}} \int d1 \int d2 \sum_{i_1<..<i_p} \Phi_{i_1}(1)\Phi_{i_1}(2) \cdot ... \cdot \Phi_{i_p}(1)\Phi_{i_p}(2)\,.$$

Large N limit

Now we introduce two collective fields

$$Q(1,2) = \frac{1}{N}\sum_i \Phi_i(1)\Phi_i(2)\,,$$

$$R(1) = \frac{1}{N}\sum_i x_i\Phi_i(1)\,,$$

by making use of

$$\begin{aligned} 1 &= \int DQ \int D\hat{Q}\exp\Big[\frac{1}{2}\int d1\int d2\,\big(NQ(1,2)\hat{Q}(1,2) \\ &- \hat{Q}(1,2)\sum_i \Phi_i(1)\Phi_i(2)\big)\Big]\,, \\ 1 &= \int DR\int D\hat{R}\exp\Big[\int d1\,\big(NR(1)\hat{R}(1) - \hat{R}(1)\sum_i x_i\Phi_i(1)\big)\Big]\,. \end{aligned}$$

Large N limit

The expectation value an operator O can be written as

$$\begin{aligned}\langle O\rangle &= \int D\Phi\, DQ\, D\hat{Q}\, DR\, D\hat{R}\, O\, \exp\Big[-\int d1 \int d2 \\ &\quad \Big\{\sum_{i=1}^{N} \Phi_i(1)\big[G_0^{-1}(1,2)+\frac{1}{2}\hat{Q}(1,2)\big]\Phi_i(2) \\ &- \frac{N}{2}\big[J^2 Q^{\bullet p}(1,2)+\hat{Q}(1,2)Q(1,2)\big]\Big\} \\ &- N\int d1\,\big[\frac{r}{2}R^p(1)+\hat{R}(1)R(1)\big]+\int d1\,\hat{R}(1)\sum_i x_i\Phi_i(1)\Big],\end{aligned}$$

where the sign $\bullet$ indicates ordinary multiplication (not a convolution).

Large N limit

Integrating out initial degrees of freedom Φ_i, making a shift

$$\bar{Q}(1,2)/2 := \Big[G_0^{-1}(1,2) + \frac{1}{2}\hat{Q}(1,2)\Big],$$

and integrating over $\hat{R}$, we obtain

$$\begin{aligned}
\langle O \rangle \quad \sim \quad & \int DQ\, D\bar{Q}\, DR\, \widetilde{O}\, \exp\Big[-\frac{N}{2}\int d1 \int d2 \\
& \big[(G_0^{-1}(1,2) - \bar{Q}(1,2))Q(1,2) - J^2 Q^{\bullet p}(1,2)\big] \\
- \quad & \frac{N}{2} r \int d1\, R^p(1) - \frac{N}{2}\int d1\, d2\, R(1)\bar{Q}(1,2)R(2) \\
- \quad & (\frac{N}{2} - 1)\mathrm{Tr}\log \bar{Q}(1,2)\Big].
\end{aligned}$$

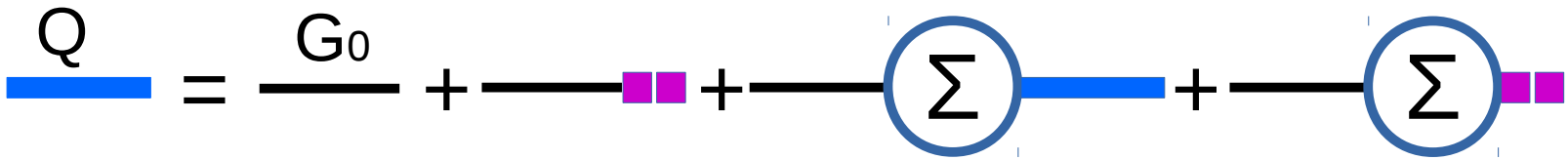
Saddle-point equations

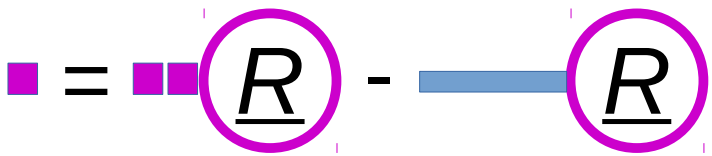
Then, the saddle-point equations are given by

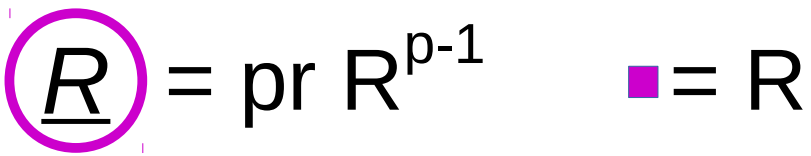
$$\bar{Q}(1,2) = (Q(1,2) - R(1)R(2))^{-1}$$
$$pQ(1,2) = K^{-1}(1,2) + R(1)R(2)$$
$$+ \int d3\, d4\, K^{-1}(1,4)(Q(3,4) - R(3)R(4))^{-1} Q^{\bullet(p-1)}(3,2)\,,$$
$$r\, p\, R^{p-1}(1) + \int d2\, (Q(1,2) - R(1)R(2))^{-1} R(2) = 0\,.$$

Saddle-point equations

Then, the saddle-point equations are given by

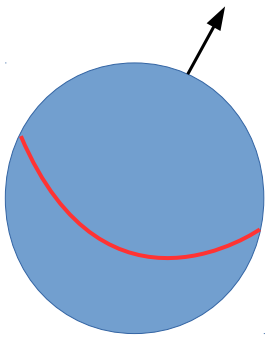






Saddle-point equations

For all r, there exist a solution of Saddle-point equations such that $R = 0$ and Q obeys melonic equation of the p-spin model.



When $\|R(1)R(2)\| \gg \|Q(1,2)\|$, the initial approximation for the R function might be found from

$$\int d1\, R^p(1) = \frac{1}{rp}\,.$$

NOVEMBER
4-6, 2019
PARIS
CENTRE CULTUREL IRLANDAIS

1st French–German Meeting in Physics, Mathematics & Artificial Intelligence Theory

SPEAKERS

Sylvain Carrozza
Romain Couillet
Razvan Gurau
Fred Hamprecht
Ullrich Köthe
Sven Krippendorf
Florent Krzakala
Philippe Loubaton
Valentina Ros
Fabian Rühle
Christoph Schnörr
Julian Urban
Sebastian Wetzel
Lenka Zdeborová

SCIENTIFIC COMMITTEE

Giulio Biroli
Ralf Klessen
Marc Mézard
Jan Pawlowski
Vincent Rivasseau
Manfred Salmhofer
Mohamed Tamaazousti
Adrian Tanasa

ORGANIZATION COMMITTEE

Giulio Biroli
Harold Erbin
Vincent Lahoche
Vincent Rivasseau
Mohamed Tamaazousti

SCAN ME!

list CEA tech · CARNOT · SysNum Sensors to decision · université de BORDEAUX · STRUCTURES CLUSTER OF EXCELLENCE · UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386 · LPT Orsay · Service pour la science et la technologie · DAAD · université PARIS-SACLAY