Melonic Non-linear Flows and the Spiked Tensor Model

Vincent Rivasseau
Laboratoire de Physique Théorique
CNRS UMR 8627, Université Paris-Sud

Mini Symposium "Holographic Tensors",
Okinawa Institute for Science and Technology, October 31 2018

Holographic Tensors

Two years ago E. Witten remarked the link between the SYK model of $NAdS_2/NCFT_1$ holographic correspondence and random tensors.

He proposed the Gurau-Witten action

$$I = \int dt \left(\frac{i}{2} \sum_i \psi_i \frac{d}{dt} \psi_i - i^{q/2} j \psi_0 \psi_1 \cdots \psi_D \right)$$

where ψ's are $D+1$ fermionic tensors and their index contractions follow the complete graph pattern of Gurau's colored tensor model and of Boulatov group field theory.

An uncolored, i.e. single tensor model with similar properties was soon developped by Klebanov and Tarnopolsky, based on the three dimensional $O(N)$ tensor interaction of Carrozza and Tanasa (2015).

Holographic Tensors

Two years ago E. Witten remarked the link between the SYK model of $NAdS_2/NCFT_1$ holographic correspondence and random tensors.

He proposed the Gurau-Witten action

$$I = \int dt \left(\frac{i}{2} \sum_i \psi_i \frac{d}{dt} \psi_i - i^{q/2} j \psi_0 \psi_1 \cdots \psi_D \right)$$

where ψ's are $D+1$ fermionic tensors and their index contractions follow the complete graph pattern of Gurau's colored tensor model and of Boulatov group field theory.

An uncolored, i.e. single tensor model with similar properties was soon developped by Klebanov and Tarnopolsky, based on the three dimensional $O(N)$ tensor interaction of Carrozza and Tanasa (2015).

Holographic Tensors

Two years ago E. Witten remarked the link between the SYK model of $NAdS_2/NCFT_1$ holographic correspondence and random tensors.

He proposed the Gurau-Witten action

$$I = \int dt \left(\frac{i}{2} \sum_i \psi_i \frac{d}{dt} \psi_i - i^{q/2} j \psi_0 \psi_1 \cdots \psi_D \right)$$

where ψ's are $D+1$ fermionic tensors and their index contractions follow the complete graph pattern of Gurau's colored tensor model and of Boulatov group field theory.

An uncolored, i.e. single tensor model with similar properties was soon developped by Klebanov and Tarnopolsky, based on the three dimensional $O(N)$ tensor interaction of Carrozza and Tanasa (2015).

Holographic Tensors

Two years ago E. Witten remarked the link between the SYK model of $NAdS_2/NCFT_1$ holographic correspondence and random tensors.

He proposed the Gurau-Witten action

$$I = \int dt \left(\frac{i}{2} \sum_i \psi_i \frac{d}{dt} \psi_i - i^{q/2} j \psi_0 \psi_1 \cdots \psi_D \right)$$

where ψ's are $D+1$ fermionic tensors and their index contractions follow the complete graph pattern of Gurau's colored tensor model and of Boulatov group field theory.

An uncolored, i.e. single tensor model with similar properties was soon developped by Klebanov and Tarnopolsky, based on the three dimensional $O(N)$ tensor interaction of Carrozza and Tanasa (2015).

Random Holography

- To understand the physics of the gravity side in this $NAdS_2/NCFT_1$ correspondence is a hot current topic.
- A main issue for the future is in my opinion to understand the link between the random geometric and the holographic aspects of random tensors.

Random Holography?

Random Holography

- To understand the physics of the gravity side in this $NAdS_2/NCFT_1$ correspondence is a hot current topic.
- A main issue for the future is in my opinion to understand the link between the random geometric and the holographic aspects of random tensors.

Random Holography?

Random Holography

- To understand the physics of the gravity side in this $NAdS_2/NCFT_1$ correspondence is a hot current topic.
- A main issue for the future is in my opinion to understand the link between the random geometric and the holographic aspects of random tensors.

Random Holography?

Random Holography

- To understand the physics of the gravity side in this $NAdS_2/NCFT_1$ correspondence is a hot current topic.
- A main issue for the future is in my opinion to understand the link between the random geometric and the holographic aspects of random tensors.

Random Holography?

Random Vectors, Matrices, Tensors

Random Vectors $\subset$ Random Matrices $\subset$ Random Tensors

- Each class is richer than the previous one, having more and more invariants
- Each class has universal aspects but different $1/N$ expansions
- Each class is connected to the discretized random geometric approach to quantum gravity, where roughly speaking rank $\simeq$ dimension
- Each class has or should have many concrete applications in probability theory, statistical mechanics, data analysis, deep learning...

Random Vectors, Matrices, Tensors

Random Vectors $\subset$ Random Matrices $\subset$ Random Tensors

- Each class is richer than the previous one, having more and more invariants
- Each class has universal aspects but different $1/N$ expansions
- Each class is connected to the discretized random geometric approach to quantum gravity, where roughly speaking rank $\simeq$ dimension
- Each class has or should have many concrete applications in probability theory, statistical mechanics, data analysis, deep learning...

Random Vectors, Matrices, Tensors

Random Vectors $\subset$ Random Matrices $\subset$ Random Tensors

- Each class is richer than the previous one, having more and more invariants
- Each class has universal aspects but different $1/N$ expansions
- Each class is connected to the discretized random geometric approach to quantum gravity, where roughly speaking rank $\simeq$ dimension
- Each class has or should have many concrete applications in probability theory, statistical mechanics, data analysis, deep learning...

Random Vectors, Matrices, Tensors

Random Vectors $\subset$ Random Matrices $\subset$ Random Tensors

- Each class is richer than the previous one, having more and more invariants
- Each class has universal aspects but different $1/N$ expansions
- Each class is connected to the discretized random geometric approach to quantum gravity, where roughly speaking rank $\simeq$ dimension
- Each class has or should have many concrete applications in probability theory, statistical mechanics, data analysis, deep learning...

Random Vectors, Matrices, Tensors

Random Vectors $\subset$ Random Matrices $\subset$ Random Tensors

- Each class is richer than the previous one, having more and more invariants
- Each class has universal aspects but different $1/N$ expansions
- Each class is connected to the discretized random geometric approach to quantum gravity, where roughly speaking rank $\simeq$ dimension
- Each class has or should have many concrete applications in probability theory, statistical mechanics, data analysis, deep learning...

Random Vectors, Matrices, Tensors

Random Vectors $\subset$ Random Matrices $\subset$ Random Tensors

- Each class is richer than the previous one, having more and more invariants
- Each class has universal aspects but different $1/N$ expansions
- Each class is connected to the discretized random geometric approach to quantum gravity, where roughly speaking rank $\simeq$ dimension
- Each class has or should have many concrete applications in probability theory, statistical mechanics, data analysis, deep learning...

Random Vectors, Matrices, Tensors

Random Vectors $\subset$ Random Matrices $\subset$ Random Tensors

- Each class is richer than the previous one, having more and more invariants
- Each class has universal aspects but different $1/N$ expansions
- Each class is connected to the discretized random geometric approach to quantum gravity, where roughly speaking rank $\simeq$ dimension
- Each class has or should have many concrete applications in probability theory, statistical mechanics, data analysis, deep learning...

Random Vectors, Matrices, Tensors

Random Vectors $\subset$ Random Matrices $\subset$ Random Tensors

- Each class is richer than the previous one, having more and more invariants
- Each class has universal aspects but different $1/N$ expansions
- Each class is connected to the discretized random geometric approach to quantum gravity, where roughly speaking rank $\simeq$ dimension
- Each class has or should have many concrete applications in probability theory, statistical mechanics, data analysis, deep learning...

Tensor-Vectors Models

Until now all applications of random tensors have been in the simplified form which we may call tensor-vectors models.

They correspond to saturate a single random tensor T of rank p with p (random) vectors u^i

$$H(T,u) = \sum_{i_1,\cdots,i_p=1}^{N} T_{i_1,\cdots,i_p} u^1_{i_1} \cdots u^p_{i_p}.$$

In the matrix case $p=2$ the action is therefore just the associated bilinear form. In the literature it goes under various names and variants

- p spin models (spin glasses, disordered systems)
- isotropic models (theory of Gaussian processes)
- spiked tensor model (data analysis)
- Sachdev-Ye-Kitaev model (condensed matter, quantum black holes...)

Tensor-Vectors Models

Until now all applications of random tensors have been in the simplified form which we may call tensor-vectors models.

They correspond to saturate a single random tensor T of rank p with p (random) vectors u^i

$$H(T,u) = \sum_{i_1,\cdots,i_p=1}^{N} T_{i_1,\cdots,i_p} u^1_{i_1} \cdot \cdots u^p_{i_p}.$$

In the matrix case $p = 2$ the action is therefore just the associated bilinear form. In the literature it goes under various names and variants

- p spin models (spin glasses, disordered systems)
- isotropic models (theory of Gaussian processes)
- spiked tensor model (data analysis)
- Sachdev-Ye-Kitaev model (condensed matter, quantum black holes...)

Tensor-Vectors Models

Until now all applications of random tensors have been in the simplified form which we may call tensor-vectors models.

They correspond to saturate a single random tensor T of rank p with p (random) vectors u^i

$$H(T, u) = \sum_{i_1, \cdots, i_p = 1}^{N} T_{i_1, \cdots, i_p} u^1_{i_1} \cdot \cdots u^p_{i_p}.$$

In the matrix case $p = 2$ the action is therefore just the associated bilinear form. In the literature it goes under various names and variants

- p spin models (spin glasses, disordered systems)
- isotropic models (theory of Gaussian processes)
- spiked tensor model (data analysis)
- Sachdev-Ye-Kitaev model (condensed matter, quantum black holes...)

Tensor-Vectors Models

Until now all applications of random tensors have been in the simplified form which we may call tensor-vectors models.

They correspond to saturate a single random tensor T of rank p with p (random) vectors u^i

$$H(T,u) = \sum_{i_1,\cdots,i_p=1}^{N} T_{i_1,\cdots,i_p} u^1_{i_1} \cdots u^p_{i_p}.$$

In the matrix case $p = 2$ the action is therefore just the associated bilinear form. In the literature it goes under various names and variants

- p spin models (spin glasses, disordered systems)
- isotropic models (theory of Gaussian processes)
- spiked tensor model (data analysis)
- Sachdev-Ye-Kitaev model (condensed matter, quantum black holes...)

Tensor-Vectors Models

Until now all applications of random tensors have been in the simplified form which we may call tensor-vectors models.

They correspond to saturate a single random tensor T of rank p with p (random) vectors u^i

$$H(T,u) = \sum_{i_1,\cdots,i_p=1}^{N} T_{i_1,\cdots,i_p} u^1_{i_1} \cdots u^p_{i_p}.$$

In the matrix case $p=2$ the action is therefore just the associated bilinear form. In the literature it goes under various names and variants

- p spin models (spin glasses, disordered systems)
- isotropic models (theory of Gaussian processes)
- spiked tensor model (data analysis)
- Sachdev-Ye-Kitaev model (condensed matter, quantum black holes...)

Tensor-Vectors Models

Until now all applications of random tensors have been in the simplified form which we may call tensor-vectors models.

They correspond to saturate a single random tensor T of rank p with p (random) vectors u^i

$$H(T, u) = \sum_{i_1, \cdots, i_p = 1}^{N} T_{i_1, \cdots, i_p} u^1_{i_1} \cdot \cdots u^p_{i_p}.$$

In the matrix case $p = 2$ the action is therefore just the associated bilinear form. In the literature it goes under various names and variants

- p spin models (spin glasses, disordered systems)
- isotropic models (theory of Gaussian processes)
- spiked tensor model (data analysis)
- Sachdev-Ye-Kitaev model (condensed matter, quantum black holes...)

Tensor-Vectors Models

Until now all applications of random tensors have been in the simplified form which we may call tensor-vectors models.

They correspond to saturate a single random tensor T of rank p with p (random) vectors u^i

$$H(T, u) = \sum_{i_1, \cdots, i_p = 1}^{N} T_{i_1, \cdots, i_p} u^1_{i_1} \cdot \cdots u^p_{i_p}.$$

In the matrix case $p = 2$ the action is therefore just the associated bilinear form. In the literature it goes under various names and variants

- p spin models (spin glasses, disordered systems)
- isotropic models (theory of Gaussian processes)
- spiked tensor model (data analysis)
- Sachdev-Ye-Kitaev model (condensed matter, quantum black holes...)

Tensor-Vectors Models

The rich random geometric aspects of tensor models and of their $1/N$ expansion is then missing; tensor-vectors essentially reduce to vector models, but with "higher order interactions"

- Matrix case: $\int e^{-M\cdot M+i\lambda M\cdot u^{\otimes 2}} \simeq e^{-\lambda^2 (u\cdot u)^2} =>$ snails
- Tensor case: $\int e^{-T\cdot T+i\lambda T\cdot u^{\otimes p}} \simeq e^{-\lambda^2 (u\cdot u)^p} =>$ melons

In tensor-vectors models melons dominate as a hint of the full tensor $1/N$ expansion, a bit like in matrix models Wigner's law can be considered a hint of 'tHooft $1/N$ expansion.

Tensor-Vectors Models

The rich random geometric aspects of tensor models and of their $1/N$ expansion is then missing; tensor-vectors essentially reduce to vector models, but with "higher order interactions"

- Matrix case: $\int e^{-M\cdot M + i\lambda M\cdot u^{\otimes 2}} \simeq e^{-\lambda^2 (u\cdot u)^2}$ => snails
- Tensor case: $\int e^{-T\cdot T + i\lambda T\cdot u^{\otimes p}} \simeq e^{-\lambda^2 (u\cdot u)^p}$ => melons

In tensor-vectors models melons dominate as a hint of the full tensor $1/N$ expansion, a bit like in matrix models Wigner's law can be considered a hint of 'tHooft $1/N$ expansion.

Tensor-Vectors Models

The rich random geometric aspects of tensor models and of their $1/N$ expansion is then missing; tensor-vectors essentially reduce to vector models, but with “higher order interactions”

- Matrix case: $\int e^{-M\cdot M + i\lambda M\cdot u^{\otimes 2}} \simeq e^{-\lambda^2 (u\cdot u)^2}$ => snails
- Tensor case: $\int e^{-T\cdot T + i\lambda T\cdot u^{\otimes p}} \simeq e^{-\lambda^2 (u\cdot u)^p}$ => melons

In tensor-vectors models melons dominate as a hint of the full tensor $1/N$ expansion, a bit like in matrix models Wigner's law can be considered a hint of 'tHooft $1/N$ expansion.

Tensor-Vectors Models

The rich random geometric aspects of tensor models and of their $1/N$ expansion is then missing; tensor-vectors essentially reduce to vector models, but with "higher order interactions"

- Matrix case: $\int e^{-M\cdot M + i\lambda M\cdot u^{\otimes 2}} \simeq e^{-\lambda^2 (u\cdot u)^2}$ => snails
- Tensor case: $\int e^{-T\cdot T + i\lambda T\cdot u^{\otimes p}} \simeq e^{-\lambda^2 (u\cdot u)^p}$ => melons

In tensor-vectors models melons dominate as a hint of the full tensor $1/N$ expansion, a bit like in matrix models Wigner's law can be considered a hint of 'tHooft $1/N$ expansion.

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.

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.

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.

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.

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 W 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

$$\lim_{N\to\infty} P(N^{2/3}\lambda_c \leq x) = F_2^{TW}(x). \tag{2.1}$$

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

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

$$\lim_{N\to\infty} P(N^{2/3}\lambda_c \leq x) = F_2^{TW}(x). \tag{2.1}$$

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

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

$$\lim_{N\to\infty} P(N^{2/3}\lambda_c \leq x) = F_2^{TW}(x). \tag{2.1}$$

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

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

$$\lim_{N\to\infty} P(N^{2/3}\lambda_c \leq x) = F_2^{TW}(x). \tag{2.1}$$

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.

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.

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.

A long history made short

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 $\lambda_c = \lambda$).

A long history made short

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, I

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

where

$$G(z) = < \frac{1}{N} \mathrm{Tr} \frac{1}{z - H_0 - H} >_H .$$

The next two slides are extracted from E. Brézin - S. Hikami *Correlations of Nearby Levels Induced by a Random Potential*, Nuclear Physics B **479**, 697-706 (1996).

Computation of λ_c, I

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

where

$$G(z) = < \frac{1}{N} \mathrm{Tr} \frac{1}{z - H_0 - H} >_H .$$

The next two slides are extracted from E. Brézin - S. Hikami *Correlations of Nearby Levels Induced by a Random Potential*, Nuclear Physics B **479**, 697-706 (1996).

Computation of λ_c, I

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

where

$$G(z) = < \frac{1}{N} \mathrm{Tr} \frac{1}{z - H_0 - H} >_H .$$

The next two slides are extracted from E. Brézin - S. Hikami *Correlations of Nearby Levels Induced by a Random Potential*, Nuclear Physics B **479**, 697-706 (1996).

Computation of λ_c, II

$$\frac{\partial G}{\partial z} = \oint \frac{du}{2\pi i} \frac{1}{u + G_0(u) - z} \tag{3.8}$$

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

$$u + G_0(u) = z \tag{3.9}$$

i.e.

$$u + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{u - \epsilon_i} = z \tag{3.10}$$

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

$$\hat{u}(z) = z - \frac{1}{z} + O\left(\frac{1}{z^2}\right) \tag{3.11}$$

Computation of λ_c, III

Therefore, for large z, the contour encloses all the roots of (3.10) 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. Therefore it is easier to calculate the integral (3.8) 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

$$\begin{aligned}\frac{\partial G}{\partial z} &= 1 - \frac{1}{1 + \frac{dG_0}{d\hat{u}(z)}} \\ &= 1 - \frac{d\hat{u}(z)}{dz} \qquad (3.12)\end{aligned}$$

The integration gives

$$G(z) = z - \hat{u}(z) \qquad (3.13)$$

Computation of λ_c, IV

Following these good authors, we have therefore to compute $\hat{u}(z)$.

- 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, IV

Following these good authors, we have therefore to compute $\hat{u}(z)$.

- 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, IV

Following these good authors, we have therefore to compute $\hat{u}(z)$.

- 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, IV

Following these good authors, we have therefore to compute $\hat{u}(z)$.

- 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, IV

Following these good authors, we have therefore to compute $\hat{u}(z)$.

- 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, V

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.

Computation of λ_c, V

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.

Computation of λ_c, V

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.

Computation of λ_c, V

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.

Computation of λ_c, V

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, or...good luck!)

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, or...good luck!)

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, or...good luck!)

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, or...good luck!)

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, or...good luck!)

The Spiked Tensor Model

- Signal: a deterministic tensor $T_0 = \lambda u^{\otimes p} \in [\mathbb{R}^N]^{\otimes p}$ where $u \in \mathbb{R}^N$ is a normalized vector $\|u\| = 1$
- Noise: a N^p Gaussian iid random tensor T (symmetric or not)

We observe $T = S_0 + S = \lambda u_{i_1} \cdots u_{i_p} + S_{i_1 \cdots i_p}$ and the goal is to recover u.

λ again allows to control the noise: λ large means small noise and vice-versa.

Since 2014 it has been studied intensively eg by Montanari, Richard, Ben Arous, Biroli and many coworkers.

The Spiked Tensor Model

- Signal: a deterministic tensor $T_0 = \lambda u^{\otimes p} \in [\mathbb{R}^N]^{\otimes p}$ where $u \in \mathbb{R}^N$ is a normalized vector $\|u\| = 1$
- Noise: a N^p Gaussian iid random tensor T (symmetric or not)

We observe $T = S_0 + S = \lambda u_{i_1} \cdots u_{i_p} + S_{i_1 \cdots i_p}$ and the goal is to recover u.

λ again allows to control the noise: λ large means small noise and vice-versa.

Since 2014 it has been studied intensively eg by Montanari, Richard, Ben Arous, Biroli and many coworkers.

The Spiked Tensor Model

- Signal: a deterministic tensor $T_0 = \lambda u^{\otimes p} \in [\mathbb{R}^N]^{\otimes p}$ where $u \in \mathbb{R}^N$ is a normalized vector $\|u\| = 1$
- Noise: a N^p Gaussian iid random tensor T (symmetric or not)

We observe $T = S_0 + S = \lambda u_{i_1} \cdots u_{i_p} + S_{i_1 \cdots i_p}$ and the goal is to recover u.

λ again allows to control the noise: λ large means small noise and vice-versa.

Since 2014 it has been studied intensively eg by Montanari, Richard, Ben Arous, Biroli and many coworkers.

The Spiked Tensor Model

- Signal: a deterministic tensor $T_0 = \lambda u^{\otimes p} \in [\mathbb{R}^N]^{\otimes p}$ where $u \in \mathbb{R}^N$ is a normalized vector $\|u\| = 1$
- Noise: a N^p Gaussian iid random tensor T (symmetric or not)

We observe $T = S_0 + S = \lambda u_{i_1} \cdots u_{i_p} + S_{i_1 \cdots i_p}$ and the goal is to recover u.

λ again allows to control the noise: λ large means small noise and vice-versa.

Since 2014 it has been studied intensively eg by Montanari, Richard, Ben Arous, Biroli and many coworkers.

The Spiked Tensor Model

- Signal: a deterministic tensor $T_0 = \lambda u^{\otimes p} \in [\mathbb{R}^N]^{\otimes p}$ where $u \in \mathbb{R}^N$ is a normalized vector $\|u\| = 1$
- Noise: a N^p Gaussian iid random tensor T (symmetric or not)

We observe $T = S_0 + S = \lambda u_{i_1} \cdots u_{i_p} + S_{i_1 \cdots i_p}$ and the goal is to recover u.

λ again allows to control the noise: λ large means small noise and vice-versa.

Since 2014 it has been studied intensively eg by Montanari, Richard, Ben Arous, Biroli and many coworkers.

The Spiked Tensor Model

- Signal: a deterministic tensor $T_0 = \lambda u^{\otimes p} \in [\mathbb{R}^N]^{\otimes p}$ where $u \in \mathbb{R}^N$ is a normalized vector $\|u\| = 1$
- Noise: a N^p Gaussian iid random tensor T (symmetric or not)

We observe $T = S_0 + S = \lambda u_{i_1} \cdots u_{i_p} + S_{i_1 \cdots i_p}$ and the goal is to recover u.

λ again allows to control the noise: λ large means small noise and vice-versa.

Since 2014 it has been studied intensively eg by Montanari, Richard, Ben Arous, Biroli and many coworkers.

Data Recovery

The cost function $f(v) =< v^{\otimes p} \cdot T >= \lambda(v \cdot u)^p + T \cdot v^{\otimes p}$ makes data recovery quite different than in the matrix case, since the associated flow $\dot{v} = v^{\otimes p-1} \cdot T$ is non-linear.

As we know, non-linear flows cannot be usually integrated out explicitly and on non-compact spaces they can diverge in finite time. For instance on $\mathbb{R}$

$$\dot{x} = \lambda x^p$$

for $p \geq 2$ trivially solves to

$$x_t = x_0[1 - (p-1)\lambda t x_0^{p-1}]^{-\frac{1}{p-1}} = x_0 \sum_{h \in \mathbb{N}} \frac{(\lambda t x_0^{p-1})^h}{h!} \prod_{k=1}^{h-1} [k(p-1)+1].$$

which converges for $|t| < \frac{1}{(q-1)\lambda x_0^{q-1}}$, but diverges after that, e.g. for λ and x_0 positive. The solution can be expanded graphically into (heap-ordered) trees.

Data Recovery

The cost function $f(v) =< v^{\otimes p} \cdot T >= \lambda(v \cdot u)^p + T \cdot v^{\otimes p}$ makes data recovery quite different than in the matrix case, since the associated flow $\dot{v} = v^{\otimes p-1} \cdot T$ is non-linear.

As we know, non-linear flows cannot be usually integrated out explicitly and on non-compact spaces they can diverge in finite time. For instance on $\mathbb{R}$

$$\dot{x} = \lambda x^p$$

for $p \geq 2$ trivially solves to

$$x_t = x_0[1 - (p-1)\lambda t x_0^{p-1}]^{-\frac{1}{p-1}} = x_0 \sum_{h \in \mathbb{N}} \frac{(\lambda t x_0^{p-1})^h}{h!} \prod_{k=1}^{h-1} [k(p-1)+1].$$

which converges for $|t| < \frac{1}{(q-1)\lambda x_0^{q-1}}$, but diverges after that, e.g. for λ and x_0 positive. The solution can be expanded graphically into (heap-ordered) trees.

Data Recovery

The cost function $f(v) =< v^{\otimes p} \cdot T >= \lambda(v \cdot u)^p + T \cdot v^{\otimes p}$ makes data recovery quite different than in the matrix case, since the associated flow $\dot{v} = v^{\otimes p-1} \cdot T$ is non-linear.

As we know, non-linear flows cannot be usually integrated out explicitly and on non-compact spaces they can diverge in finite time. For instance on $\mathbb{R}$

$$\dot{x} = \lambda x^p$$

for $p \geq 2$ trivially solves to

$$x_t = x_0[1-(p-1)\lambda t x_0^{p-1}]^{-\frac{1}{p-1}} = x_0 \sum_{h \in \mathbb{N}} \frac{(\lambda t x_0^{p-1})^h}{h!} \prod_{k=1}^{h-1} [k(p-1)+1].$$

which converges for $|t| < \frac{1}{(q-1)\lambda x_0^{q-1}}$, but diverges after that, e.g. for λ and x_0 positive. The solution can be expanded graphically into (heap-ordered) trees.

Data Recovery

The cost function $f(v) =< v^{\otimes p} \cdot T >= \lambda(v \cdot u)^p + T \cdot v^{\otimes p}$ makes data recovery quite different than in the matrix case, since the associated flow $\dot{v} = v^{\otimes p-1} \cdot T$ is non-linear.

As we know, non-linear flows cannot be usually integrated out explicitly and on non-compact spaces they can diverge in finite time. For instance on $\mathbb{R}$

$$\dot{x} = \lambda x^p$$

for $p \geq 2$ trivially solves to

$$x_t = x_0[1-(p-1)\lambda t x_0^{p-1}]^{-\frac{1}{p-1}} = x_0 \sum_{h\in\mathbb{N}} \frac{(\lambda t x_0^{p-1})^h}{h!} \prod_{k=1}^{h-1}[k(p-1)+1].$$

which converges for $|t| < \frac{1}{(q-1)\lambda x_0^{q-1}}$, but diverges after that, e.g. for λ and x_0 positive. The solution can be expanded graphically into (heap-ordered) trees.

Data Recovery

The cost function $f(v) =< v^{\otimes p} \cdot T >= \lambda(v \cdot u)^p + T \cdot v^{\otimes p}$ makes data recovery quite different than in the matrix case, since the associated flow $\dot{v} = v^{\otimes p-1} \cdot T$ is non-linear.

As we know, non-linear flows cannot be usually integrated out explicitly and on non-compact spaces they can diverge in finite time. For instance on $\mathbb{R}$

$$\dot{x} = \lambda x^p$$

for $p \geq 2$ trivially solves to

$$x_t = x_0[1 - (p-1)\lambda t x_0^{p-1}]^{-\frac{1}{p-1}} = x_0 \sum_{h \in \mathbb{N}} \frac{(\lambda t x_0^{p-1})^h}{h!} \prod_{k=1}^{h-1} [k(p-1)+1].$$

which converges for $|t| < \frac{1}{(q-1)\lambda x_0^{q-1}}$, but diverges after that, e.g. for λ and x_0 positive. The solution can be expanded graphically into (heap-ordered) trees.

Melonic Non-Linear Flows

Averaging over the random tensor T in the tree expansion of the non-linear flow, and over initial conditions or over the white noise in the Langevin version, one gets loops on the tree, hence a Feynman graph expansion of the flow.

One can check that in the large N limit melons dominate this Feynman graph expansion. The corresponding melonic approximation is known as the direct interaction approximation in the turbulence literature (going back to works of Kraichnan et al.) and as the mode coupling approximation in the spin-glass literature (see Bouchaud-Cugliandolo 1995).

I hope the study of random non-linear flows in this approximation and beyond could benefit from the progress on SYK and tensor models.

Melonic Non-Linear Flows

Averaging over the random tensor T in the tree expansion of the non-linear flow, and over initial conditions or over the white noise in the Langevin version, one gets loops on the tree, hence a Feynman graph expansion of the flow.

One can check that in the large N limit melons dominate this Feynman graph expansion. The corresponding melonic approximation is known as the direct interaction approximation in the turbulence literature (going back to works of Kraichnan et al.) and as the mode coupling approximation in the spin-glass literature (see Bouchaud-Cugliandolo 1995).

I hope the study of random non-linear flows in this approximation and beyond could benefit from the progress on SYK and tensor models.

Melonic Non-Linear Flows

Averaging over the random tensor T in the tree expansion of the non-linear flow, and over initial conditions or over the white noise in the Langevin version, one gets loops on the tree, hence a Feynman graph expansion of the flow.

One can check that in the large N limit melons dominate this Feynman graph expansion. The corresponding melonic approximation is known as the direct interaction approximation in the turbulence literature (going back to works of Kraichnan et al.) and as the mode coupling approximation in the spin-glass literature (see Bouchaud-Cugliandolo 1995).

I hope the study of random non-linear flows in this approximation and beyond could benefit from the progress on SYK and tensor models.

Data Recovery in the Spiked Tensor Case

Returning to the spiked tensor model, the experts recognize three different regimes (with constants $\lambda_c^1 > \lambda_c^2$ depending on the rank p)

- for $\lambda > \lambda_c^1(p)$ recovery is easy by iterative methods as in the previous $\lambda \geq 1$ matrix case;
- for $\lambda_c^2(p) < \lambda < \lambda_c^1(p)$ recovery is hard because dynamics is slow. This is the "glassy" regime one wants to better understand;
- for $\lambda < \lambda_c^2(p)$ recovery is quasi-impossible as in the previous $\lambda < 1$ matrix case.

Data Recovery in the Spiked Tensor Case

Returning to the spiked tensor model, the experts recognize three different regimes (with constants $\lambda_c^1 > \lambda_c^2$ depending on the rank p)

- for $\lambda > \lambda_c^1(p)$ recovery is easy by iterative methods as in the previous $\lambda \geq 1$ matrix case;
- for $\lambda_c^2(p) < \lambda < \lambda_c^1(p)$ recovery is hard because dynamics is slow. This is the "glassy" regime one wants to better understand;
- for $\lambda < \lambda_c^2(p)$ recovery is quasi-impossible as in the previous $\lambda < 1$ matrix case.

Data Recovery in the Spiked Tensor Case

Returning to the spiked tensor model, the experts recognize three different regimes (with constants $\lambda_c^1 > \lambda_c^2$ depending on the rank p)

- for $\lambda > \lambda_c^1(p)$ recovery is easy by iterative methods as in the previous $\lambda \geq 1$ matrix case;
- for $\lambda_c^2(p) < \lambda < \lambda_c^1(p)$ recovery is hard because dynamics is slow. This is the "glassy" regime one wants to better understand;
- for $\lambda < \lambda_c^2(p)$ recovery is quasi-impossible as in the previous $\lambda < 1$ matrix case.

Data Recovery in the Spiked Tensor Case

Returning to the spiked tensor model, the experts recognize three different regimes (with constants $\lambda_c^1 > \lambda_c^2$ depending on the rank p)

- for $\lambda > \lambda_c^1(p)$ recovery is easy by iterative methods as in the previous $\lambda \geq 1$ matrix case;
- for $\lambda_c^2(p) < \lambda < \lambda_c^1(p)$ recovery is hard because dynamics is slow. This is the "glassy" regime one wants to better understand;
- for $\lambda < \lambda_c^2(p)$ recovery is quasi-impossible as in the previous $\lambda < 1$ matrix case.

Landscape Complexity

The work of Montanari, Ben Arous et al on the spiked tensor model relies on random matrix theory.

For instance to detect u in a high rank tensor of rank p one can detect first a factorized q by r object with $q + r = p$ by matrix method, then iterate. This method called unfolding works especially well in k steps if $p = 2^k$.

One can also give asymptotic estimates on the number of critical points of the cost function f and on their indices. Indeed these depend only of two indices of the underlying random tensor, hence of a matrix. This is the reason why such data at large N are accessible through random matrix eigenvalues asymptotic estimates.

Landscape Complexity

The work of Montanari, Ben Arous et al on the spiked tensor model relies on random matrix theory.

For instance to detect u in a high rank tensor of rank p one can detect first a factorized q by r object with $q + r = p$ by matrix method, then iterate. This method called unfolding works especially well in k steps if $p = 2^k$.

One can also give asymptotic estimates on the number of critical points of the cost function f and on their indices. Indeed these depend only of two indices of the underlying random tensor, hence of a matrix. This is the reason why such data at large N are accessible through random matrix eigenvalues asymptotic estimates.

Landscape Complexity

The work of Montanari, Ben Arous et al on the spiked tensor model relies on random matrix theory.

For instance to detect u in a high rank tensor of rank p one can detect first a factorized q by r object with $q + r = p$ by matrix method, then iterate. This method called unfolding works especially well in k steps if $p = 2^k$.

One can also give asymptotic estimates on the number of critical points of the cost function f and on their indices. Indeed these depend only of two indices of the underlying random tensor, hence of a matrix. This is the reason why such data at large N are accessible through random matrix eigenvalues asymptotic estimates.

Landscape Complexity

More precisely consider the Hamiltonian $H(T,u)$ as a function on the sphere S_N. Its critical points are defined as zeroes of the differential ∇H, hence noted $\nabla H = 0$ and their index i is defined by the number of negative eigenvalues of the Hessian $\nabla\nabla H$ at these critical points.

In fact this Hessian is a random matrix and it's conditional expectation value exactly correspond to GOE expectation values. More precisely Aufflinger-Ben Arous-Cerny proved that for $B \subset \mathbb{R}$ the number $CRT_{N,k}(B)$ of critical points of index k such that $H(T,u) \in B$:

$$< Crt_{N,k}(B) >_T = 2\sqrt{\frac{2}{p}}(p-1)^{\frac{N}{2}} < e^{-N\frac{p-2}{2p}\lambda_k^2} 1\{\lambda_k \in \sqrt{\frac{p}{2p-2}}B\} >_{GOE(N)}$$

Landscape Complexity

More precisely consider the Hamiltonian $H(T, u)$ as a function on the sphere S_N. Its critical points are defined as zeroes of the differential ∇H, hence noted $\nabla H = 0$ and their index i is defined by the number of negative eigenvalues of the Hessian $\nabla\nabla H$ at these critical points.

In fact this Hessian is a random matrix and it's conditional expectation value exactly correspond to GOE expectation values. More precisely Aufflinger-Ben Arous-Cerny proved that for $B \subset \mathbb{R}$ the number $CRT_{N,k}(B)$ of critical points of index k such that $H(T, u) \in B$:

$$< Crt_{N,k}(B) >_T = 2\sqrt{\frac{2}{p}}(p-1)^{\frac{N}{2}} < e^{-N\frac{p-2}{2p}} \lambda_k^2 1\{\lambda_k \in \sqrt{\frac{p}{2p-2}}B\} >_{GOE(N)}$$

Landscape Complexity

More precisely consider the Hamiltonian $H(T,u)$ as a function on the sphere S_N. Its critical points are defined as zeroes of the differential ∇H, hence noted $\nabla H = 0$ and their index i is defined by the number of negative eigenvalues of the Hessian $\nabla\nabla H$ at these critical points.

In fact this Hessian is a random matrix and it's conditional expectation value exactly correspond to GOE expectation values. More precisely Aufflinger-Ben Arous-Cerny proved that for $B \subset \mathbb{R}$ the number $CRT_{N,k}(B)$ of critical points of index k such that $H(T,u) \in B$:

$$< Crt_{N,k}(B) >_T = 2\sqrt{\frac{2}{p}}(p-1)^{\frac{N}{2}} < e^{-N\frac{p-2}{2p}\lambda_k^2} 1\{\lambda_k \in \sqrt{\frac{p}{2p-2}}B\} >_{GOE(N)}$$

- These results on the complexity of the landscape of the spiked tensor model are undoubtedly spectacular!
- In principle they give information or at least strong hints about the slow dynamics of such models in the interesting glassy phase.
- However ultimately non-linear flows do not depend solely on the number and types of critical points but also on the depth of the barriers between them. This does not seem to me fully accessible with purely random matrix eigenvalues methods; random tensors contain more information.
- Hence it seems to me that in a long term perspective, tensor data analysis should be studied with truly tensorial methods, such as graphical and combinatorial methods.

- These results on the complexity of the landscape of the spiked tensor model are undoubtedly spectacular!
- In principle they give information or at least strong hints about the slow dynamics of such models in the interesting glassy phase.
- However ultimately non-linear flows do not depend solely on the number and types of critical points but also on the depth of the barriers between them. This does not seem to me fully accessible with purely random matrix eigenvalues methods; random tensors contain more information.
- Hence it seems to me that in a long term perspective, tensor data analysis should be studied with truly tensorial methods, such as graphical and combinatorial methods.

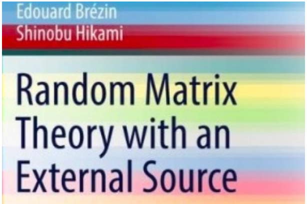
- These results on the complexity of the landscape of the spiked tensor model are undoubtedly spectacular!
- In principle they give information or at least strong hints about the slow dynamics of such models in the interesting glassy phase.
- However ultimately non-linear flows do not depend solely on the number and types of critical points but also on the depth of the barriers between them. This does not seem to me fully accessible with purely random matrix eigenvalues methods; random tensors contain more information.
- Hence it seems to me that in a long term perspective, tensor data analysis should be studied with truly tensorial methods, such as graphical and combinatorial methods.

- These results on the complexity of the landscape of the spiked tensor model are undoubtedly spectacular!
- In principle they give information or at least strong hints about the slow dynamics of such models in the interesting glassy phase.
- However ultimately non-linear flows do not depend solely on the number and types of critical points but also on the depth of the barriers between them. This does not seem to me fully accessible with purely random matrix eigenvalues methods; random tensors contain more information.
- Hence it seems to me that in a long term perspective, tensor data analysis should be studied with truly tensorial methods, such as graphical and combinatorial methods.

- These results on the complexity of the landscape of the spiked tensor model are undoubtedly spectacular!
- In principle they give information or at least strong hints about the slow dynamics of such models in the interesting glassy phase.
- However ultimately non-linear flows do not depend solely on the number and types of critical points but also on the depth of the barriers between them. This does not seem to me fully accessible with purely random matrix eigenvalues methods; random tensors contain more information.
- Hence it seems to me that in a long term perspective, tensor data analysis should be studied with truly tensorial methods, such as graphical and combinatorial methods.

Conclusion

- By studying random tensors you can reconcile quantum gravity research with real life applications.



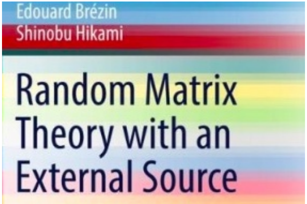
-
- I hope to have convinced some young participants at this 2018 OIST workshop to raise to the challenge of tensorial data analysis and for instance perhaps write in the future some book entitled

Random Tensor Theory with an External Source

Thanks for your attention!

Conclusion

- By studying random tensors you can reconcile quantum gravity research with real life applications.



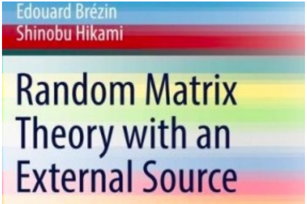
-
- I hope to have convinced some young participants at this 2018 OIST workshop to raise to the challenge of tensorial data analysis and for instance perhaps write in the future some book entitled

Random Tensor Theory with an External Source

Thanks for your attention!

Conclusion

- By studying random tensors you can reconcile quantum gravity research with real life applications.



-

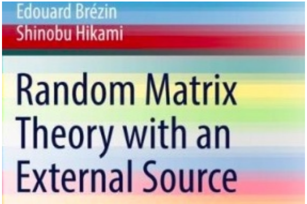
- I hope to have convinced some young participants at this 2018 OIST workshop to raise to the challenge of tensorial data analysis and for instance perhaps write in the future some book entitled

Random Tensor Theory with an External Source

Thanks for your attention!

Conclusion

- By studying random tensors you can reconcile quantum gravity research with real life applications.



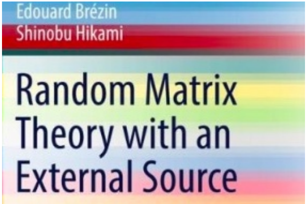
-
- I hope to have convinced some young participants at this 2018 OIST workshop to raise to the challenge of tensorial data analysis and for instance perhaps write in the future some book entitled

Random Tensor Theory with an External Source

Thanks for your attention!

Conclusion

- By studying random tensors you can reconcile quantum gravity research with real life applications.



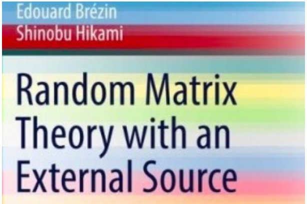
-
- I hope to have convinced some young participants at this 2018 OIST workshop to raise to the challenge of tensorial data analysis and for instance perhaps write in the future some book entitled

Random Tensor Theory with an External Source

Thanks for your attention!

Conclusion

- By studying random tensors you can reconcile quantum gravity research with real life applications.



-
- I hope to have convinced some young participants at this 2018 OIST workshop to raise to the challenge of tensorial data analysis and for instance perhaps write in the future some book entitled

Random Tensor Theory with an External Source

Thanks for your attention!