Phase transitions in the eigenvalues & eigenvectors of perturbed random matrices

Raj Rao Nadakuditi

Dept. of Electrical Engg. & Computer Science
http://www.eecs.umich.edu/~rajnrao
Joint work with Florent Benaych-Georges
Numerical Experiment

- $G$ = Gaussian random matrix
  - $G = \text{randn}(n,n)$ or $G = \text{sign}(\text{randn}(n,n))$

- $X_n = \frac{G + G'}{\sqrt{2n}}$

- $\tilde{X}_n = X_n + P_n$
  - $P_n = \theta u u'$
  - $u$ is a fixed, non-random unit norm vector

Question: Largest eigenvalue? Corresponding eigenvector? Variation with $\theta$?
• $\theta = 4, \ n = 500$
• Bulk obeys semi-circle law on $[-2, 2]$
• Largest eig. $\approx 4.2$
• Clear phase transition @ $\theta = 1$ with increasing $n$
- Wannabe phase transition $\theta = 1$ with increasing $n$
- Norm-square of projection of largest (perturbed) eigenvector onto $v_n$
Theory

Theorem: Consider $\tilde{X}_n = X_n + \theta uu'$

$$\tilde{\lambda}_1 \xrightarrow{\text{a.s.}} \begin{cases} \theta + \frac{1}{\theta}, & \theta > 1 \\ 2, & \text{otherwise} \end{cases}$$

$$|\langle \tilde{u}_1, u \rangle|^2 \xrightarrow{\text{a.s.}} \begin{cases} \left(1 - \frac{1}{\theta^2}\right), & \theta > 1 \\ 0, & \text{otherwise} \end{cases}$$

- Eigenvalue result first due to Peche (2006), Peche-Feral (2007)
- Eigenvector result new
- Eigenvalues and eigenvectors are biased
Experiment 2

- $G = \text{Gaussian random matrix}$
  - $G = \text{randn}(n,m)$ or $G = \text{sign(\text{randn}(n,m))}$

- $X_n = \frac{GG'}{m}$

- $\tilde{X}_n = \sqrt{I + P_n X_n \sqrt{I + P_n}}$
  - $u$ is arbitrary unit norm vector
  - $P_n = \theta u u'$ is signal covariance matrix
  - $X_n$ models a noise-only sample covariance matrix
  - Motivated by additive linear models in statistics

Question: Largest eigenvalue? Corresponding eigenvector? Variation with $\theta$?
Theory

Theorem: Consider \( \tilde{X}_n = \sqrt{I + P_n} X_n \sqrt{I + P_n} \)

\[
\tilde{\lambda}_1 \xrightarrow{\text{a.s.}} \begin{cases} 
(\theta + 1) \left(1 + \frac{c}{\theta}\right), & \theta > \sqrt{c} \\
(1 + \sqrt{c})^2, & \text{otherwise}
\end{cases}
\]

\[
|\langle \tilde{u}_1, u \rangle|^2 \xrightarrow{\text{a.s.}} \begin{cases} 
\frac{\theta^2 - c}{\theta^2 + c \theta}, & \theta > \sqrt{c} \\
0, & \text{otherwise}
\end{cases}
\]

- Eigenvalue result due to Baik-Ben Arous-Peche (2005), Baik-Silverstein (2006)
- Eigenvector result first due to Paul (2007) and others since then
- Eigenvalues and eigenvectors are biased
- \( \theta \sim \text{SNR}, \ c = \lim \frac{n}{m} \)
Main message

Question that motivated this work:

- How does limit depend on the random matrix and perturbative model?

Answer we provide in this talk:

- Problem **solved** in great generality with very transparent proof

- Closed form expressions for location of phase transition
  - Limit = f(noise eigen-spectrum, perturbative model)

- "Spiked" free probability
Outline

• Why study the Phase transition?
• Analytical expressions for the phase transition
• Sketch of the proof
• Discussion
Why study the phase transition?

- Idea that signals lie in a low dimensional subspace relative to noise
- Eigen-analysis based dimensionality reduction exploit this fact
- Efficient algorithms exist (SVD and their fast variants)
- When PCA works well: (near)-optimality + strong performance guarantees

**Engineering motivation:** When will PCA fail? Can it made better?

- Massive data sets, “large \( p \) small \( n \)” type problems make this important
- Phase transitions provide basis for comparing within-class and out-of-class algorithms
- Analogous to breakdown-point work in sparse approximation theory (Donoho, Stodden, Tanner)
Outline

• Why study the Phase transition?

• Analytical expressions for the Phase transition

• Sketch of the proof

• Discussion
Definitions and assumptions

Spectral measure: Eigenvalues of $X_n$ are $\lambda_1, \ldots, \lambda_n$:

$$
\mu_{X_n} = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i}
$$

Assumptions:

1. $\mu_{X_n} \xrightarrow{a.s.} \mu_X$
2. $\text{supp } \mu_X$ compactly supported on $[a, b]$
3. $\max(\text{eig}) \xrightarrow{a.s.} b$
Perturbative model I

\[ \tilde{X}_n = \sum_{i=1}^{k} \theta_i u_i u_i' + X_n \]

Assumptions:

- \( X_n \) is symmetric, unitarily invariant random matrix with \( n \) real eigenvalues
- \( \theta_1, \ldots, \theta_k > 0 \)
- \( X_n = Q \Lambda Q' \) where \( Q \) is a Haar distributed unitary (or orthogonal) matrix
- \( U \) is a non-random orthogonal or unitary matrix (independent of \( Q \))
- \( u_1, \ldots, u_k \) are the \( k \) columns of \( U \)
Phase transition of largest eigenvalues

**Theorem** [Benaych-Georges and N.]: As $n \to \infty$,

$$\lambda_i(\tilde{X}_n) \xrightarrow{\text{a.s.}} \begin{cases} G^{-1}_\mu(1/\theta_i) & \text{if } 1/\theta_i < G_\mu(b^+), \\ b & \text{otherwise,} \end{cases}$$

- Critical threshold depends explicitly on spectral measure of “noise”

**Cauchy transform of $\mu$:**

$$G_\mu(z) = \int \frac{1}{z-y} d\mu(y) \quad \text{for } z \notin \text{supp } \mu_X.$$
Phase transition of eigenvectors

**Theorem** [Benaych-Georges and N.]: As $n \to \infty$, for $\theta > \theta_c$:

\[
\langle \tilde{u}_i, \ker(\theta_i I_n - P_n) \rangle \xrightarrow{\text{a.s.}} - \frac{1}{\theta_i^2 G_\mu'(\rho)}
\]

\[
\langle \tilde{u}_i, \bigoplus_{j \neq i} \ker(\theta_j I_n - P_n) \rangle \xrightarrow{\text{a.s.}} 0,
\]

- $\rho = G_\mu^{-1}(1/\theta_i)$ is the corresponding eigenvalue limit

**Theorem**: As $n \to \infty$, for $\theta \leq \theta_c$:

\[
\langle \tilde{u}_i, \ker(\theta_i I_n - P_n) \rangle \xrightarrow{\text{a.s.}} 0
\]

- Assumption of eigenvalue repulsion required
The result graphically

(a) Eigenvalue: $\theta > \theta_c$

(b) Eigenvector: $\theta > \theta_c$
The result graphically

(c) Eigenvalue: $\theta \leq \theta_c$

(d) Eigenvector: $\theta \leq \theta_c$
Perturbative model II

\[ \tilde{X}_n = (\sum_{i=1}^{k} \theta_i u_i u'_i + I) X_n \]

Equivalently (via a similarity transformation)

\[ \tilde{X}_n = \sqrt{\sum_{i=1}^{k} \theta_i u_i u'_i + I} X_n \sqrt{\sum_{i=1}^{k} \theta_i u_i u'_i + I} \]

Assumptions:

- \( X_n \) is symmetric, unitarily invariant random matrix with \( n \) real eigenvalues
- \( \theta_1, \ldots, \theta_k > 0 \)
- \( X_n = Q \Lambda Q' \) where \( Q \) is a Haar distributed unitary (or orthogonal) matrix
- \( U \) is a non-random orthogonal or unitary matrix (independent of \( X \))
- \( u_1, \ldots, u_k \) are the \( k \) columns of \( U \)
Phase transition of largest eigenvalues

**Theorem** [Benaych-Georges and N.]: As \( n \to \infty \),

\[
\lambda_i(\tilde{X}_n) \xrightarrow{a.s.} \begin{cases} T^{-1}_\mu(1/\theta_i) & \text{if } 1/\theta_i < T_\mu(b^+), \\ b & \text{otherwise,} \end{cases}
\]

- Critical threshold depends explicitly on spectral measure of “noise”

**T-transform of \( \mu \):**

\[
T_\mu(z) = \int \frac{t}{z-t} d\mu_X(t) \quad \text{for } z \notin \text{supp } \mu_X,
\]
Phase transition of eigenvectors

Theorem [Benaych-Georges and N.]: As $n \to \infty$, for $\theta > \theta_c$:

\[
\langle \tilde{u}_i, \ker(\theta_i I_n - P_n) \rangle^2 \xrightarrow{\text{a.s.}} \frac{-1}{\theta_i^2 \rho T'_{\mu X}(\rho) + \theta_i},
\]

\[
\langle \tilde{u}_i, \bigoplus_{j \neq i} \ker(\theta_j I_n - P_n) \rangle \xrightarrow{\text{a.s.}} 0,
\]

- $\rho = T_{\mu}^{-1}(1/\theta_i)$ is the corresponding eigenvalue limit

Theorem: As $n \to \infty$, for $\theta \leq \theta_c$:

\[
\langle \tilde{u}_i, \ker(\theta_i I_n - P_n) \rangle \xrightarrow{\text{a.s.}} 0
\]

- Assumption of eigenvalue repulsion required
Perturbative model III

\[ \tilde{X}_n = \sum_{i=1}^{k} \theta_i u_i v_i' + X_n \]

Assumptions:

- \( X_n \) is \( n \times m \) bi-unitarily invariant random matrix (\( n \leq m \)) with \( n \) singular values
- \( \theta_1, \ldots, \theta_k > 0 \)
- \( X_n = Q\Lambda W' \) where \( Q \) and \( W \) are Haar distributed unitary (or orthogonal) matrices
- \( U \) and \( V \) are non-random unitary matrices (independent of \( Q \) and \( W \))
- \( u_1, \ldots, u_k \) and \( v_1, \ldots, v_k \) are \( k \) columns of \( U \) and \( V \)
Phase transition of largest singular values

Theorem [Benaych-Georges and N.]: As $n \to \infty$, 

$$
\sigma_i(\tilde{X}_n) \xrightarrow{a.s.} \begin{cases} 
D_{\mu_X}^{-1}(c, 1/\theta_i^2) & \text{if } 1/\theta_i^2 < D_{\mu_X}(c, b^+), \\
b & \text{otherwise},
\end{cases}
$$

- Critical threshold depends explicitly on spectral measure of “noise”

D-transform of $\mu$:

$$
D_{\mu_X}(c, z) = \left[ \int \frac{z}{z^2 - t^2} d\mu(t) \right] \cdot \left[ c \int \frac{z}{z^2 - t^2} d\mu(t) + \frac{1 - c}{z} \right] \quad \text{for } z \notin \text{supp } \mu_X,
$$
Phase transition of singular vectors

Theorem: As $n, m \to \infty$, $n/m \to c$, for $\theta \leq \theta_c$:

\[
\langle u_i, \ker(\theta_i^2 I_n - P_n P_n^*) \rangle^2 \xrightarrow{\text{a.s.}} \frac{-2\varphi_{\mu X}(\rho)}{\theta_i^2 \partial_z D_{\mu X}(c, \rho)},
\]

\[
\langle v_i, \ker(\theta_i^2 I_m - P_n P_n^*) \rangle^2 \xrightarrow{\text{a.s.}} \frac{-2\varphi_{\tilde{\mu} X}(\rho)}{\theta_i^2 \partial_z D_{\mu X}(c, \rho)},
\]

- Here $\rho = D_{\mu X}^{-1}(c, 1/\theta_i^2)$ is the limit of $\theta_i$
- $\tilde{\mu}_X = c\mu_X + (1 - c)\delta_0$ and $\varphi_{\mu}(z) = \int \frac{z}{z^2 - t^2} d\mu(t)$

Theorem: As $n, m \to \infty$, for $\theta \leq \theta_c$:

\[
\langle \tilde{u}_i, \ker(\theta_i I_n - P_n P_n') \rangle \xrightarrow{\text{a.s.}} 0
\]

\[
\langle \tilde{v}_i, \ker(\theta_i I_m - P_n' P_n) \rangle \xrightarrow{\text{a.s.}} 0
\]
Outline

• Why study the Phase transition?

• Analytical expressions for the Phase transition

• Sketch of the proof

• Discussion
Phase transition of eigenvalues

Theorem: As \( n \to \infty \),
\[
\tilde{X}_n = \sum_{i=1}^{k} \theta_i u_i u'_i + X
\]
\[
\lambda_i(\tilde{X}_n) \xrightarrow{\text{a.s.}} \begin{cases} 
G_{\mu}^{-1}(1/\theta_i) & \text{if } 1/\theta_i < G_{\mu}(b^+), \\
b & \text{otherwise,}
\end{cases}
\]

• Critical threshold depends explicitly on spectral measure of “noise”

Cauchy transform of \( \mu \):

\[
G_{\mu}(z) = \int \frac{1}{z - y} d\mu(y) \quad \text{for } z \notin \text{supp } \mu_X.
\]
Consider $\tilde{X} = X + \theta uu'$. 

**Fact:** $z \neq \lambda(X)$ is an eigenvalue of $\tilde{X}$ if and only if:

- 1 is an eigenvalue of $(zI - X)^{-1}\theta uu'$
- This is equivalent to requiring $u'(zI - X)^{-1}u\theta = 1$
- Assuming $X = Q\Lambda Q'$ and letting $v = Q'u$ gives us

**Master (Secular) equation:** Eigenvalues $z$ of $\tilde{X}$ satisfy

$$\sum_{i=1}^{n} \frac{|v_i|^2}{z - \lambda_i} = \frac{1}{\theta}$$


Proof Step 1 - Master Equation derivation

Consider $\tilde{X} = X + \theta uu'$.

Fact: $z \neq \lambda(X)$ is an eigenvalue of $\tilde{X}$ if and only if:

$$\sum_{i=1}^{n} \frac{|v_i|^2}{z - \lambda_i} = \frac{1}{\theta}$$

Define weighted measure $\mu_n = \sum_{i=1}^{n} |v_i|^2 \delta_{\lambda_i}$ then:

$$G_{\mu_n}(z) = 1/\theta$$

Cauchy transform of $\mu$:

$$G_{\mu}(z) = \int \frac{1}{z - y} d\mu(y) \quad \text{for} \ z \in \mathbb{C}^+ \setminus \mathbb{R}.$$
Eigenvalues of $\tilde{X}$: Graphically

- Eigenvalues of $\tilde{X}$ satisfy $G_{\mu_n}(z) = 1/\theta$
Proof Step 2: “Smoothing” due to randomization

Master (Secular) equation: Eigenvalues $z$ of $\tilde{X}$ satisfy

$$G_{\mu_n}(z) = 1/\theta$$

- Weighted measure $\mu_n = \sum_{i=1}^{n} |v_i|^2 \delta_{\lambda_k}$

Argument:

- Recall that $Q$ is isotropically random and $v = Q'u$
- $\Rightarrow v$ is uniformly distributed on the sphere
- $\Rightarrow |v_k|^2 \approx 1/n$ with high probability for large $n$
- $\mu_n \xrightarrow{a.s.} \mu$ and $G_{\mu_n}(z) \xrightarrow{a.s.} G_{\mu}(z)$
- $\Rightarrow z \approx G_{\mu}^{-1}(1/\theta)$
**Origin of Phase transition: Largest eigenvalue when** $\theta > \theta_c$

When $1/\theta < G_{\mu_X}(b)$, $\lambda_1(\tilde{X}) \to \rho = G_{\mu_X}^{-1}(\frac{1}{\theta})$. 


Origin of Phase transition: Largest eigenvalue when $\theta \leq \theta_c$

When $1/\theta > G_{\mu_X}(b)$, $\lambda_1(\tilde{X}) \rightarrow b$. 
Eigenvector Proof Step 1 - Master Equation derivation

Consider $\tilde{X} = X + \theta uu'$.

**Fact:** If $z \neq \lambda(X)$ is an eigenvalue of $\tilde{X}$ then:

- $(X + \theta uu')\tilde{u} = z\tilde{u}$ for eigenvector $\tilde{u}$
- This is equivalent to requiring $(zI - X)\tilde{u} = \theta(u'\tilde{u})u$
- Note that $u'\tilde{u}$ is a scalar so above is exact expression for eigenvector!
- Assuming $X = Q\Lambda Q'$ and letting $v = Q'u$ gives us

**Master (Secular) equation:** Eigenvector $\tilde{u}$ with eigenvalue $z \neq \lambda(X)$ given by

$$\tilde{u} = \frac{Q(zI - \Lambda)^{-1}v}{\sqrt{v'(zI - \Lambda)^{-2}v}}$$
Proof Step 1 - Master equation derivation

Consider $\tilde{X} = X + \theta uu'$. 

Fact: Eigenvector $\tilde{u}$ with eigenvalue $z \neq \lambda(X)$ given by:

$$\tilde{u} = \frac{Q(zI - \Lambda)^{-1}v}{\sqrt{v'(zI - \Lambda)^{-2}v}}$$

Projection of eigenvector: (Recall $v = Q'u$)

$$|\langle \tilde{u}, u \rangle|^2 = \frac{(v'(zI - \Lambda)^{-1}v)^2}{v'(zI - \Lambda)^{-2}v} = \frac{G^2_{\mu n}(z)}{G'_{\mu n}(z)}$$
Proof Step 2 - “Smoothing” due to randomization

Projection of eigenvector:

\[ |\langle \tilde{u}, u \rangle|^2 = \frac{(v'(zI - \Lambda)^{-1}v)^2}{v'(zI - \Lambda)^{-2}v} = \frac{G_{\mu n}^2(z)}{G'_{\mu n}(z)} \]

Argument:

- Recall that \( Q \) is isotropically random and \( v = Q' u \)
  \( \Rightarrow \) \( v \) is uniformly distributed on the sphere
- \( \Rightarrow |v_k|^2 \approx 1/n \) with high probability for large \( n \)
- \( \mu_n \xrightarrow{a.s.} \mu \) and \( G_{\mu n}(z) \xrightarrow{a.s.} G_\mu(z) \)
  \( \Rightarrow |\langle \tilde{u}, u \rangle|^2 \approx \frac{G_\mu^2(z)}{G'_\mu(z)} \)
Origin of Phase transition: Largest eigenvalue when \( \theta > \theta_c \)

When \( 1/\theta < G_{\mu X}(b) \), \( \lambda_1(\tilde{X}) \to \rho = G_{\mu X}^{-1}(\frac{1}{\theta}) \).

\[ \Rightarrow |\langle \tilde{u}, u \rangle|^2 = \frac{G_{\mu}^2(\rho)}{G_{\mu}'(\rho)} \to \frac{1}{\theta^2 G_{\mu}'(\rho)} \]
Origin of Phase transition: Largest eigenvalue when $\theta \leq \theta_c$

When $1/\theta > G_{\mu \times}(b)$, $\lambda_1(\tilde{X}) \to b$.

$\Rightarrow |\langle \tilde{u}, u \rangle|^2 = \frac{G_{\mu \times}(\rho)}{G_{\mu \times}(\rho)} \to \frac{1}{\theta^2 G_{\mu \times}'(b)} \to 0$, if $G_{\mu \times}'(b^+) \to \infty$
Outline

• Why study the Phase transition?
• Analytical expressions for the Phase transition
• Sketch of the proof
• Discussion ⇐
Connection with free probability theory

Recall for $\tilde{X}_n = X + P$

**Theorem**: As $n \to \infty$,

\[
\lambda_i(\tilde{X}_n) \text{ a.s.} \to \begin{cases} 
G^{-1}_\mu(1/\theta_i) & \text{if } 1/\theta_i < G_\mu(b^+), \\
b & \text{otherwise}, 
\end{cases}
\]

- $G^{-1}_\mu(\cdot)$ related to the non-commutative analogue of the log-Fourier transform

Cauchy transform of $\mu$:

\[
G_\mu(z) = \int \frac{1}{z - y} d\mu(y) \quad \text{for } z \in \mathbb{C}^+ \setminus \mathbb{R}.
\]
Extensions

- \( \max(\text{eig}(X+P)) \) ✓
- \( \max(\text{eig}(X(I+P))) \) ✓
- \( \max(\text{svd}(X+P)) \) ✓
- \( \min(.) \) ✓
- \( \text{bulk}(.) \) ✓
- Randomized compressions ✓
- Haar-like perturbations ✓
- “Concentrated” random perturbations (any application?)
Main message

• Closed form expressions for the phase transition
  – Limit explicitly dependent on:
    ∗ Noise eigen-spectrum
    ∗ Perturbative model via appropriate integral transform

• Broad generality of result
  – Well-beyond Wigner, Wishart, Jacobi models in literature
  – Permits unified treatment

• "Spiked" free probability

http://arxiv.org/abs/0910.2120
Joint work with Florent Benaych-Georges