Cavity method for low-rank matrix factorization

Computing the information-theoretic threshold

March 4, 2017

Marc Lelarge & Léo Miolane
INRIA, ENS

Spiked Wigner model

$$\underbrace{\mathbf{Y}}_{\text{observations}} = \sqrt{\frac{\lambda}{n}}\,\underbrace{\mathbf{X}\mathbf{X}^\intercal}_{\text{signal}} + \underbrace{\mathbf{Z}}_{\text{noise}}$$

- $\mathbf{X}$: vector of dimension n with entries i.i.d. P_0 (discrete prior)
- $Z_{i,j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Universality: Krzakala et al., 2016; Lesieur et al., 2015a
- λ: signal-to-noise ratio.

The prior P_0 can encode many interesting problems!

MMSE and information-theoretic threshold

Goal

$$\begin{aligned}\mathrm{MMSE}_n &= \min_{\hat{\theta}} \frac{1}{n^2}\mathbb{E}\left\|\mathbf{X}\mathbf{X}^\intercal - \hat{\theta}(\mathbf{Y})\right\|^2 \\ &= \frac{1}{n^2}\sum_{1\le i,j\le n}(X_iX_j - \mathbb{E}[X_iX_j|\mathbf{Y}])^2 \le \underbrace{\mathbb{E}[X^2]^2 - \mathbb{E}[X]^4}_{\text{Dummy MMSE}}\end{aligned}$$

MMSE and information-theoretic threshold

Goal

$$\mathrm{MMSE}_n = \min_{\hat{\theta}} \frac{1}{n^2}\mathbb{E}\left\|\mathbf{X}\mathbf{X}^\intercal - \hat{\theta}(\mathbf{Y})\right\|^2$$
$$= \frac{1}{n^2}\sum_{1\le i,j\le n}(X_iX_j - \mathbb{E}[X_iX_j|\mathbf{Y}])^2 \le \underbrace{\mathbb{E}[X^2]^2 - \mathbb{E}[X]^4}_{\text{Dummy MMSE}}$$

Information-theoretic threshold

1. Compute $\lim_{n\to\infty} \mathrm{MMSE}_n$
2. Deduce the information-theoretic threshold, i.e. the critical value λ_c such that
   - if $\lambda > \lambda_c$, $\lim_{n\to\infty} \mathrm{MMSE}_n <$ Dummy MMSE
   - if $\lambda < \lambda_c$, $\lim_{n\to\infty} \mathrm{MMSE}_n =$ Dummy MMSE

Contents

Part 1.

Motivation

Naive PCA

B.B.P. phase transition

- The matrix $\mathbf{Y}/\sqrt{n} = \sqrt{\lambda}\mathbf{X}\mathbf{X}^\intercal/n + \mathbf{Z}/\sqrt{n}$ is a perturbed low-rank matrix.
- Estimate $\mathbf{X}$ using the eigenvector $\hat{\mathbf{x}}_n$ associated with the largest eigenvalue μ_n.

Naive PCA

B.B.P. phase transition

- The matrix $\mathbf{Y}/\sqrt{n} = \sqrt{\lambda}\mathbf{X}\mathbf{X}^\intercal/n + \mathbf{Z}/\sqrt{n}$ is a perturbed low-rank matrix.
- Estimate $\mathbf{X}$ using the eigenvector $\hat{\mathbf{x}}_n$ associated with the largest eigenvalue μ_n.

B.B.P. phase transition

- if $\lambda \leq 1$ $\begin{cases} \mu_n & \longrightarrow 2 \\ (\mathbf{X} \cdot \hat{\mathbf{x}}_n)^2 & \longrightarrow 0 \end{cases}$
- if $\lambda > 1$ $\begin{cases} \mu_n & \longrightarrow \sqrt{\lambda} + \frac{1}{\sqrt{\lambda}} > 2 \\ (\mathbf{X} \cdot \hat{\mathbf{x}}_n)^2 & \longrightarrow 1 - 1/\lambda > 0 \end{cases}$

Baik et al., 2005; Benaych-Georges and Nadakuditi, 2011

Sparse PCA

Exploiting sparsity

- Setting: $P_0 = \text{Ber}(\epsilon)$.
- Can we beat naive PCA under this sparsity assumption?

Some previous works

- Impossible to exactly recover the support of $\mathbf{X}$ (Amini and Wainwright, 2008).
- AMP achieves MMSE for $\epsilon > \epsilon^*$ (Rangan and Fletcher, 2012, Deshpande and Montanari, 2014).
- AMP is conjectured to be sub-optimal below ϵ^* (Lesieur et al., 2015b).

What is the best achievable performance in the $\epsilon < \epsilon^*$ regime?

Community detection in the stochastic block model

Asymmetric case

Asymmetric SBM

- n people are divided in two classes of sizes pn and $(1-p)n$.
- Connection matrix $\mathbf{M} = \frac{d}{n}\begin{pmatrix} a & b \\ b & c \end{pmatrix}$
- Constant average degree: $pa + (1-p)b = pb + (1-p)c = 1$
- Signal-to-noise ratio $\lambda = d(1-b)^2$

Community detection in the stochastic block model

Asymmetric case

Asymmetric SBM

- n people are divided in two classes of sizes pn and $(1-p)n$.
- Connection matrix $\mathbf{M} = \frac{d}{n}\begin{pmatrix} a & b \\ b & c \end{pmatrix}$
- Constant average degree: $pa + (1-p)b = pb + (1-p)c = 1$
- Signal-to-noise ratio $\lambda = d(1-b)^2$

Equivalence (Deshpande and Abbe, 2016)

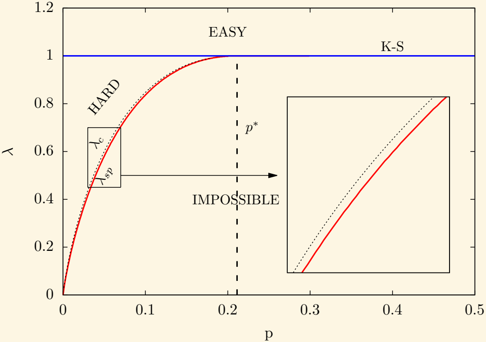
When $n, d \to \infty$ with λ fixed, the community detection problem is equivalent to

$$\mathbf{Y} = \sqrt{\frac{\lambda}{n}}\mathbf{X}\mathbf{X}^\intercal + \mathbf{Z}$$

where $\begin{cases} P_0(\sqrt{(1-p)/p}) & = p \\ P_0(-\sqrt{p/(1-p)}) & = 1-p \end{cases}$

Phase diagram for asymmetric SBM

Apparition of a "hard region"



Caltagirone et al., 2016

Some previous works

- For $p = 1/2$, KS bound by Massoulié, 2014; Mossel et al., 2015.
- Easy above KS, Bordenave et al., 2015.
- Possible but hard below KS, Neeman and Netrapalli, 2014.

We would like to find the information-theoretic threshold.

Part 2.

Locating the information-theoretic threshold

Available proof techniques

- Random matrix theory. Provides bounds. Not expected to be tight in presence of hard phase.
- Second moment computations and contiguity. Provides a finer description: "detection/recovery" (Banks et al., 2016; Perry et al., 2016). May not be tight in our regime.
- Message-passing algorithms: provide tight bounds in absence of hard phase. In presence of hard phase, need for spatial coupling techniques (Barbier et al., 2016).

Connection with statistical physics

A planted spin glass model

- Compute the MMSE for $\mathbf{Y} = \sqrt{\frac{\lambda}{n}}\mathbf{X}\mathbf{X}^\intercal + \mathbf{Z}$

Connection with statistical physics

A planted spin glass model

- Compute the MMSE for $\mathbf{Y} = \sqrt{\frac{\lambda}{n}}\mathbf{X}\mathbf{X}^\intercal + \mathbf{Z}$
- Study the posterior $\mathbb{P}(\mathbf{x} \mid \mathbf{Y}) = \frac{1}{Z_n} P_0(\mathbf{x}) \exp(H_n(\mathbf{x}))$ where

$$
\begin{aligned}
H_n(\mathbf{x}) &= \sum_{i<j} \sqrt{\frac{\lambda}{n}} Y_{i,j} x_i x_j - \frac{\lambda}{2n} x_i^2 x_j^2 \\
&= \sum_{i<j} \underbrace{\sqrt{\frac{\lambda}{n}} Z_{i,j} x_i x_j}_{\text{SK}} + \underbrace{\frac{\lambda}{n} X_i X_j x_i x_j - \frac{\lambda}{2n} x_i^2 x_j^2}_{\text{planting}}
\end{aligned}
$$

Connection with statistical physics

A planted spin glass model

- Compute the MMSE for $\mathbf{Y} = \sqrt{\frac{\lambda}{n}}\mathbf{X}\mathbf{X}^\intercal + \mathbf{Z}$
- Study the posterior $\mathbb{P}(\mathbf{x} \,|\, \mathbf{Y}) = \frac{1}{Z_n} P_0(\mathbf{x}) \exp(H_n(\mathbf{x}))$ where

$$
\begin{aligned}
H_n(\mathbf{x}) &= \sum_{i<j} \sqrt{\frac{\lambda}{n}} Y_{i,j} x_i x_j - \frac{\lambda}{2n} x_i^2 x_j^2 \\
&= \sum_{i<j} \underbrace{\sqrt{\frac{\lambda}{n}} Z_{i,j} x_i x_j}_{\text{SK}} + \underbrace{\frac{\lambda}{n} X_i X_j x_i x_j - \frac{\lambda}{2n} x_i^2 x_j^2}_{\text{planting}}
\end{aligned}
$$

- Compute the limit of the free energy $F_n = \frac{1}{n} \mathbb{E} \log Z_n$ because

$$
\text{Constant} - F_n = \frac{1}{n} I(\mathbf{X}; \mathbf{Y}) \xrightarrow[\text{I-MMSE theorem}]{\partial \lambda} \mathrm{MMSE}
$$

Replica symmetric formula

The scalar channel

Lesieur et al., 2015a conjectured that the problem is characterized par the scalar channel:

$$Y_0 = \sqrt{\gamma} X_0 + Z_0$$

and the scalar free energy: $\mathcal{F}(\gamma) = \mathbb{E}\left[\log \sum_{x_0} P_0(x_0) e^{\sqrt{\gamma} Y_0 x_0 - \frac{\gamma}{2} x_0^2}\right]$

Replica symmetric formula

The scalar channel

Lesieur et al., 2015a conjectured that the problem is characterized par the scalar channel:

$$Y_0 = \sqrt{\gamma} X_0 + Z_0$$

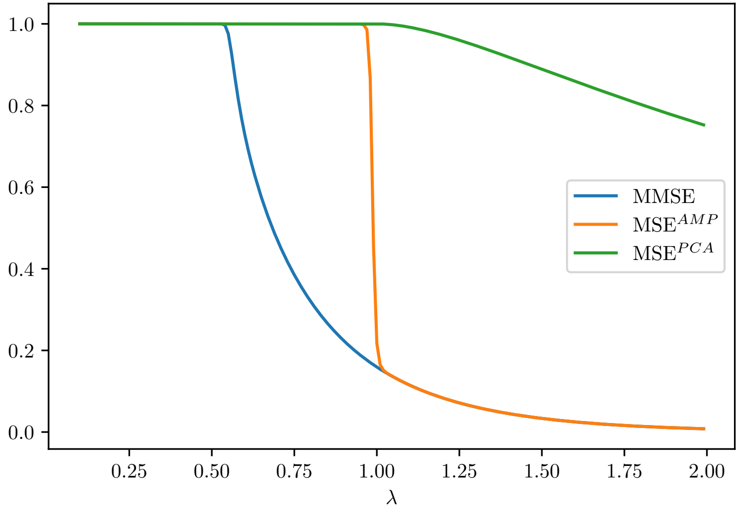
and the scalar free energy: $\mathcal{F}(\gamma) = \mathbb{E}\left[\log \sum_{x_0} P_0(x_0) e^{\sqrt{\gamma} Y_0 x_0 - \frac{\gamma}{2} x_0^2}\right]$

Replica symmetric formula

$$F_n \xrightarrow[n\to\infty]{} \sup_{q \geq 0} \mathcal{F}(\lambda q) - \frac{\lambda}{4} q^2$$

$$\mathrm{MMSE}_n \xrightarrow[n\to\infty]{} \mathbb{E}_{P_0}[X^2]^2 - q^*(\lambda)^2$$

Proved by Barbier et al., 2016 using spatial coupling techniques.



MMSE, $\mathrm{MSE}^{\mathrm{PCA}}$ and $\mathrm{MSE}^{\mathrm{AMP}}$, asymmetric SBM: $p = 0.05$.

Part 3.

Proof ideas

Proof ideas

Upper and lower bounds

- Lower bound: Guerra's interpolation technique. Adapted in Korada and Macris, 2009; Krzakala et al., 2016.

$$\begin{cases} \mathbf{Y} & = \sqrt{t} & \sqrt{\lambda/n} & \mathbf{X}\mathbf{X}^\intercal & + & \mathbf{Z} \\ \mathbf{Y}' & = \sqrt{1-t} & \sqrt{\lambda} & \mathbf{X} & + & \mathbf{Z}' \end{cases}$$

- Upper bound: Cavity computations (Mézard et al., 1987). Aizenman-Sims-Starr scheme Aizenman et al., 2003; Talagrand, 2010.

Proof ideas

Upper and lower bounds

- Lower bound: Guerra's interpolation technique. Adapted in Korada and Macris, 2009; Krzakala et al., 2016.

$$\begin{cases} \mathbf{Y} & = \sqrt{t} & \sqrt{\lambda/n} & \mathbf{X}\mathbf{X}^\intercal & + & \mathbf{Z} \\ \mathbf{Y}' & = \sqrt{1-t} & \sqrt{\lambda} & \mathbf{X} & + & \mathbf{Z}' \end{cases}$$

- Upper bound: Cavity computations (Mézard et al., 1987). Aizenman-Sims-Starr scheme Aizenman et al., 2003; Talagrand, 2010.

We now prove the upper bound: $\limsup_{n\to\infty} F_n \leq \sup_{q\geq 0} \mathcal{F}(\lambda q) - \frac{\lambda}{4} q^2$.

Overlap concentration

A general principle

Magic lemma: Montanari, 2008

Revealing a small fraction $\mathbf{X}^*$ of the planted solution forces the correlations to decay:

$$\frac{1}{n^2} \sum_{1 \leq i,j \leq n} I\big(X_i ; X_j \,\big|\, \mathbf{Y}, \mathbf{X}^*\big) \xrightarrow[n \to \infty]{} 0$$

Overlap concentration

A general principle

Magic lemma: Montanari, 2008

Revealing a small fraction $\mathbf{X}^*$ of the planted solution forces the correlations to decay:

$$\frac{1}{n^2} \sum_{1 \leq i,j \leq n} I\big(X_i\,;X_j \,\big|\, \mathbf{Y}, \mathbf{X}^*\big) \xrightarrow[n\to\infty]{} 0$$

- Gibbs measure $\langle \cdot \rangle = \mathbb{P}(\mathbf{x} \,|\, \mathbf{Y}, \mathbf{X}^*)$
- Let $\mathbf{x}^{(\mathbf{1})}, \mathbf{x}^{(\mathbf{2})} \sim \langle \cdot \rangle$ be two replicas.
- Overlap $\mathbf{x}^{(\mathbf{1})} \cdot \mathbf{x}^{(\mathbf{2})} = \frac{1}{n} \sum_{i=1}^{n} x_i^{(1)} x_i^{(2)}$ and mean overlap $Q = \big\langle \mathbf{x}^{(\mathbf{1})} \cdot \mathbf{x}^{(\mathbf{2})} \big\rangle$.

Overlap concentration

A general principle

Magic lemma: Montanari, 2008

Revealing a small fraction $\mathbf{X}^*$ of the planted solution forces the correlations to decay:

$$\frac{1}{n^2} \sum_{1 \leq i,j \leq n} I\big(X_i \,; X_j \,\big|\, \mathbf{Y}, \mathbf{X}^*\big) \xrightarrow[n\to\infty]{} 0$$

- Gibbs measure $\langle \cdot \rangle = \mathbb{P}(\mathbf{x} \,|\, \mathbf{Y}, \mathbf{X}^*)$
- Let $\mathbf{x}^{(\mathbf{1})}, \mathbf{x}^{(\mathbf{2})} \sim \langle \cdot \rangle$ be two replicas.
- Overlap $\mathbf{x}^{(\mathbf{1})} \cdot \mathbf{x}^{(\mathbf{2})} = \frac{1}{n} \sum_{i=1}^{n} x_i^{(1)} x_i^{(2)}$ and mean overlap $Q = \big\langle \mathbf{x}^{(\mathbf{1})} \cdot \mathbf{x}^{(\mathbf{2})} \big\rangle$.

Consequence: under a small perturbation of the inference model:

$$\mathbb{E}\Big\langle (\mathbf{x}^{(\mathbf{1})} \cdot \mathbf{x}^{(\mathbf{2})} - Q)^2 \Big\rangle \,,\; \mathbb{E}\Big\langle (\mathbf{x} \cdot \mathbf{X} - Q)^2 \Big\rangle \xrightarrow[n\to\infty]{} 0$$

Cavity computations

Aizenman-Sims-Starr scheme

- **Idea**: compare the system with $n+1$ variables (spins) to the system with n variables.
- Study the influence of the n first variables **over the last one** $x' = x_{n+1}$.

Cavity computations

Aizenman-Sims-Starr scheme

- **Idea**: compare the system with $n+1$ variables (spins) to the system with n variables.
- Study the influence of the n first variables **over the last one** $x' = x_{n+1}$.

$$H_{n+1}(\mathbf{x}, x') = \sum_{1 \le i < j \le n+1} \sqrt{\frac{\lambda}{n+1}} Z_{i,j} x_i x_j + \frac{\lambda}{n+1} x_i x_j X_i X_j - \frac{\lambda}{2(n+1)} x_i^2 x_j^2$$

(Cheat)
$$\simeq H_n(\mathbf{x}) + x' \sqrt{\frac{\lambda}{n}} \sum_{i=1}^{n} x_i Z'_i + x' X' \frac{\lambda}{n} \sum_{i=1}^{n} x_i X_i - (x')^2 \frac{\lambda}{2n} \sum_{i=1}^{n} x_i^2$$

$$= H_n(\mathbf{x}) + \sqrt{\lambda} x' z(\mathbf{x}) + \lambda x' X' (\mathbf{x} \cdot \mathbf{X}) - \frac{\lambda}{2} (x')^2 (\mathbf{x} \cdot \mathbf{x})$$

Cavity computations

Aizenman-Sims-Starr scheme

- **Idea**: compare the system with $n+1$ variables (spins) to the system with n variables.
- Study the influence of the n first variables **over the last one** $x' = x_{n+1}$.

$$H_{n+1}(\mathbf{x}, x') = \sum_{1 \leq i < j \leq n+1} \sqrt{\frac{\lambda}{n+1}} Z_{i,j} x_i x_j + \frac{\lambda}{n+1} x_i x_j X_i X_j - \frac{\lambda}{2(n+1)} x_i^2 x_j^2$$

$$\textbf{(Cheat)} \quad \simeq H_n(\mathbf{x}) + x' \sqrt{\frac{\lambda}{n}} \sum_{i=1}^{n} x_i Z_i' + x' X' \frac{\lambda}{n} \sum_{i=1}^{n} x_i X_i - (x')^2 \frac{\lambda}{2n} \sum_{i=1}^{n} x_i^2$$

$$= H_n(\mathbf{x}) + \sqrt{\lambda} x' z(\mathbf{x}) + \lambda x' X' (\mathbf{x} \cdot \mathbf{X}) - \frac{\lambda}{2} (x')^2 (\mathbf{x} \cdot \mathbf{x})$$

Remark that $F_n = \frac{1}{n} \sum_{k=0}^{n-1} \Big(\mathbb{E} \log Z_{k+1} - \mathbb{E} \log Z_k \Big)$

- We have therefore to compute the limit of

$$\mathbb{E}\log Z_{n+1} - \mathbb{E}\log Z_n \simeq \mathbb{E}\left[\log \frac{\sum P_0(\mathbf{x}, x')e^{H_{n+1}(\mathbf{x},x')}}{\sum P_0(\mathbf{x})e^{H_n(\mathbf{x})}}\right]$$

$$\simeq \mathbb{E}\log\left\langle \sum_{x'} P_0(x')\exp\left(\sqrt{\lambda}z(\mathbf{x})x' + \lambda(\mathbf{x}\cdot\mathbf{X})x'X' - \frac{\lambda}{2}(\mathbf{x}\cdot\mathbf{x})x'^2\right)\right\rangle$$

- We have therefore to compute the limit of

$$\mathbb{E}\log Z_{n+1} - \mathbb{E}\log Z_n \simeq \mathbb{E}\left[\log\frac{\sum P_0(\mathbf{x},x')e^{H_{n+1}(\mathbf{x},x')}}{\sum P_0(\mathbf{x})e^{H_n(\mathbf{x})}}\right]$$

$$\simeq \mathbb{E}\log\left\langle\sum_{x'} P_0(x')\exp\left(\sqrt{\lambda}z(\mathbf{x})x' + \lambda(\mathbf{x}\cdot\mathbf{X})x'X' - \frac{\lambda}{2}(\mathbf{x}\cdot\mathbf{x})x'^2\right)\right\rangle$$

- Remember

$$\mathcal{F}(\lambda Q) = \mathbb{E}\left[\log\sum_{x_0} P_0(x_0)\exp\left(\sqrt{\lambda Q}Z_0x_0 + \lambda Q x_0X_0 - \frac{\lambda Q}{2}x_0^2\right)\right]$$

- We have therefore to compute the limit of

$$\mathbb{E}\log Z_{n+1} - \mathbb{E}\log Z_n \simeq \mathbb{E}\left[\log \frac{\sum P_0(\mathbf{x}, x')e^{H_{n+1}(\mathbf{x},x')}}{\sum P_0(\mathbf{x})e^{H_n(\mathbf{x})}}\right]$$

$$\simeq \mathbb{E}\log\left\langle \sum_{x'} P_0(x')\exp\left(\sqrt{\lambda}z(\mathbf{x})x' + \lambda(\mathbf{x}\cdot\mathbf{X})x'X' - \frac{\lambda}{2}(\mathbf{x}\cdot\mathbf{x})x'^2\right)\right\rangle$$

- Remember

$$\mathcal{F}(\lambda Q) = \mathbb{E}\left[\log\sum_{x_0} P_0(x_0)\exp\left(\sqrt{\lambda Q}Z_0x_0 + \lambda Q x_0 X_0 - \frac{\lambda Q}{2}x_0^2\right)\right]$$

- Using the overlap concentration:

$$\mathbb{E}\log Z_{n+1} - \mathbb{E}\log Z_n \simeq \mathbb{E}\left[\mathcal{F}(\lambda Q) \underbrace{- \quad \frac{\lambda}{4}\; Q^2}_{\text{because we cheated}}\right]$$

- Upper bound: $\limsup_{n\to\infty} F_n \leq \sup_{q\geq 0} \mathcal{F}(\lambda q) - \frac{\lambda}{4}q^2$ ∎

Conclusion

- Overlap concentration + cavity computations = robust arguments. See Coja-Oghlan et al., 2016.
- The non-symmetric case $\mathbf{Y} = \sqrt{\lambda/n}\mathbf{U}\mathbf{V}^\intercal + \mathbf{Z}$ can be treated (almost) the same way.
- Can be extended to low-rank tensor factorization, see Lesieur et al., 2017.
- Challenge: extension to extensive rank, see Christian Schmidt poster and Kabashima et al., 2016.

Thank you for your attention.

Any questions?

Part 4.

References

References I

- Aizenman, Michael, Robert Sims, and Shannon L Starr (2003). "Extended variational principle for the Sherrington-Kirkpatrick spin-glass model". In: *Physical Review B* 68.21, p. 214403.
- Amini, Arash A and Martin J Wainwright (2008). "High-dimensional analysis of semidefinite relaxations for sparse principal components". In: *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*. IEEE, pp. 2454–2458.
- Baik, Jinho, Gérard Ben Arous, and Sandrine Péché (2005). "Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices". In: *Annals of Probability*, pp. 1643–1697.
- Banks, Jess et al. (2016). "Information-theoretic bounds and phase transitions in clustering, sparse PCA, and submatrix localization". In: *arXiv preprint arXiv:1607.05222v2*.
- Barbier, Jean et al. (2016). "Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula". In: *Advances in Neural Information Processing Systems*, pp. 424–432.

References II

- Benaych-Georges, Florent and Raj Rao Nadakuditi (2011). “The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices”. In: *Advances in Mathematics* 227.1, pp. 494–521.
- Bordenave, Charles, Marc Lelarge, and Laurent Massoulié (2015). “Non-backtracking spectrum of random graphs: community detection and non-regular Ramanujan graphs”. In: *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*. IEEE, pp. 1347–1357.
- Caltagirone, Francesco, Marc Lelarge, and Léo Miolane (2016). “Recovering asymmetric communities in the stochastic block model”. In: *arXiv preprint arXiv:1610.03680*.
- Coja-Oghlan, Amin et al. (2016). “Information-theoretic thresholds from the cavity method”. In: *arXiv preprint arXiv:1611.00814*.
- Deshpande, Yash and Emmanuel Abbe (2016). “Asymptotic mutual information for the balanced binary stochastic block model”. In: *Information and Inference*, iaw017.
- Deshpande, Yash and Andrea Montanari (2014). “Information-theoretically optimal sparse PCA”. In: *2014 IEEE International Symposium on Information Theory*. IEEE, pp. 2197–2201.

References III

- Kabashima, Yoshiyuki et al. (2016). "Phase transitions and sample complexity in bayes-optimal matrix factorization". In: *IEEE Transactions on Information Theory* 62.7, pp. 4228–4265.
- Korada, Satish Babu and Nicolas Macris (2009). "Exact solution of the gauge symmetric p-spin glass model on a complete graph". In: *Journal of Statistical Physics* 136.2, pp. 205–230.
- Krzakala, Florent, Jiaming Xu, and Lenka Zdeborová (2016). "Mutual information in rank-one matrix estimation". In: *Information Theory Workshop (ITW), 2016 IEEE*. IEEE, pp. 71–75.
- Lesieur, Thibault, Florent Krzakala, and Lenka Zdeborová (2015a). "MMSE of probabilistic low-rank matrix estimation: Universality with respect to the output channel". In: *53rd Annual Allerton Conference on Communication, Control, and Computing, Allerton 2015, Allerton Park & Retreat Center, Monticello, IL, USA, September 29 - October 2, 2015*. IEEE, pp. 680–687. ISBN: 978-1-5090-1824-6. DOI: 10.1109/ALLERTON.2015.7447070. URL: http://dx.doi.org/10.1109/ALLERTON.2015.7447070.

References IV

- Lesieur, Thibault, Florent Krzakala, and Lenka Zdeborová (2015b). "Phase transitions in sparse PCA". In: *IEEE International Symposium on Information Theory, ISIT 2015, Hong Kong, China, June 14-19, 2015*. IEEE, pp. 1635–1639. ISBN: 978-1-4673-7704-1. DOI: 10.1109/ISIT.2015.7282733. URL: http://dx.doi.org/10.1109/ISIT.2015.7282733.
- Lesieur, Thibault et al. (2017). "Statistical and computational phase transitions in spiked tensor estimation". In: *arXiv preprint arXiv:1701.08010*.
- Massoulié, Laurent (2014). "Community detection thresholds and the weak Ramanujan property". In: *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*. ACM, pp. 694–703.
- Mézard, Marc, Giorgio Parisi, and Miguel Virasoro (1987). *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*. Vol. 9. World Scientific Publishing Co Inc.
- Montanari, Andrea (2008). "Estimating random variables from random sparse observations". In: *European Transactions on Telecommunications* 19.4, pp. 385–403.

References V

- Mossel, Elchanan, Joe Neeman, and Allan Sly (2015). "Reconstruction and estimation in the planted partition model". In: *Probability Theory and Related Fields* 162.3-4, pp. 431–461.
- Neeman, Joe and Praneeth Netrapalli (2014). "Non-Reconstructability in the Stochastic Block Model". In: *arXiv preprint arXiv:1404.6304*.
- Perry, Amelia et al. (2016). "Optimality and Sub-optimality of PCA for Spiked Random Matrices and Synchronization". In: *arXiv preprint arXiv:1609.05573*.
- Rangan, Sundeep and Alyson K Fletcher (2012). "Iterative estimation of constrained rank-one matrices in noise". In: *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*. IEEE, pp. 1246–1250.
- Talagrand, Michel (2010). *Mean field models for spin glasses: Volume I: Basic examples*. Vol. 54. Springer Science & Business Media.