IRS Paris January 25, 2023

# Free probabilities in action: the spectral method for phase retrieval

Marc Potters[1]

[1] *Capital Fund Management, Paris*

# References


- Yue M. Lu and Gen Li, *Phase transitions of spectral initialization for high-dimensional nonconvex estimation*, 2020
- Wangyu Luo, Wael Alghamdi and Yue M. Lu, *Optimal Spectral Initialization for Signal Recovery With Applications to Phase Retrieval*, 2019
- MP and Jean-Philippe Bouchaud, *A First Course in Random Matrix Theory*, 2020 [+ Erratum].
- Federico Ricci-Tersenghi (private communication)

# Phase Retrieval

Let an unknown vector $\mathbf{x} \in \mathbb{R}^N$ be "probed" with $T$ vectors $\mathbf{a}_k$,

$y_k = |\mathbf{a}_k^T \mathbf{x}|^2$ with $k = 1, \cdots, T$

Vectors $\mathbf{x}$ and $\mathbf{a}_k$ are real or complex.

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left( \sum_k \left| |\mathbf{a}_k^T \mathbf{x}|^2 - y_k \right|^2 \right). \tag{1}$$

Gradient descent needs a good starting point

Problem is phase invariant ($\mathbf{a}_k \to -\mathbf{a}_k$), a linear method cannot work.

Use the phase invariant projectors $\mathbf{a}_k \mathbf{a}_k^T$

# Spectral Method

$$\mathbf{K} = \frac{1}{T}\sum_{k=1}^{T} f(y_k)\,\mathbf{a}_k\mathbf{a}_k^T, \tag{2}$$

$$\mathbf{G}_{\mathbf{K}}(z) = (z\mathbf{1} - \mathbf{K})^{-1} = \sum_{\ell=1}^{N} \frac{\mathbf{v}_\ell\mathbf{v}_\ell^T}{z - \lambda_\ell} \tag{3}$$

Rotational Invariance, assume $\mathbf{x} = \mathbf{e}_1$

$$\mathbf{K} = \begin{pmatrix} c_1 & \mathbf{w}^T \\ \mathbf{w} & \mathbf{M} \end{pmatrix}, \tag{4}$$

$$[\mathbf{G}_{\mathbf{K}}(z)]_{11} = \frac{1}{z - c_1 - \mathbf{w}^T\mathbf{G}_{\mathbf{M}}(z)\mathbf{w}} \tag{5}$$

$$\boxed{\lambda_1 = c_1 + \mathbf{w}^T\mathbf{G}_{\mathbf{M}}(\lambda_1)\mathbf{w} \equiv h(\lambda_1)} \tag{6}$$

# Overlap with Largest Eigenvector

$$[\mathbf{G}_{\mathbf{K}}(z)]_{11} = \frac{1}{z - h(z)} \tag{7}$$

$$\varrho := \frac{|\mathbf{v}_1^T \mathbf{x}|^2}{|\mathbf{x}|^2} = |\mathbf{v}_1^T \mathbf{e}_1|^2 = \lim_{z \to \lambda_1} \frac{z - \lambda_1}{z - h(z)}. \tag{8}$$

$$\boxed{\varrho = \frac{1}{1 - h'(\lambda_1)}} \quad \text{with} \quad \lambda_1 = h(\lambda_1) \tag{9}$$

# Computing $h(z)$

$$c_1 = \frac{1}{T}\sum_{k=1}^{T} f(y_k)([\mathbf{a}_k]_1)^2 \overset{T\to\infty}{\to} \left\langle f(y)([\mathbf{a}]_1)^2 \right\rangle . \tag{10}$$

$$\begin{aligned}
\mathbf{w}^T \mathbf{G}_{\mathbf{M}}(z)\mathbf{w} &= \sum_{k,\ell=1}^{T} \frac{1}{T^2} f(y_k) f(y_\ell)[\mathbf{a}_k]_1[\mathbf{a}_\ell]_1 \sum_{i,j>1}^{N} [\mathbf{a}_k]_i [\mathbf{G}_{\mathbf{M}}(z)]_{ij}[\mathbf{a}_\ell]_j \\
&= \sum_{k=1}^{T} \frac{1}{T^2} f^2(y_k)([\mathbf{a}_k]_1)^2 \sum_{i,j>1}^{N} [\mathbf{a}_k]_i [\mathbf{G}_{\mathbf{M}}(z)]_{ij}[\mathbf{a}_k]_j \qquad (11)\\
&\overset{T\to\infty}{\to} q\tau\left(\frac{\mathbf{H}\mathbf{D}_1\mathbf{H}^T}{T}\mathbf{G}_{\mathbf{M}}(z)\right), \quad q = N/T, \quad [\mathbf{H}]_{ik} = [\mathbf{a}_k]_i \quad i > 1 \qquad (12)
\end{aligned}$$

$$\boxed{h(z) = c_1 + q\tau\left(\mathbf{M}_1 \mathbf{G}_{\mathbf{M}}(z)\right)} \tag{13}$$

$$\mathbf{M} = \frac{1}{T}\mathbf{H}\mathbf{D}_0\mathbf{H}^T \quad \mathbf{M}_1 = \frac{1}{T}\mathbf{H}\mathbf{D}_1\mathbf{H}^T \quad [\mathbf{D}_0]_{kk} = f(y_k) \quad [\mathbf{D}_1]_{kk} = f^2(y_k)([\mathbf{a}_k]_1)^2 \tag{14}$$

# Free Probabilities

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \quad \text{or} \quad \mathbf{C} = \mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2} \tag{15}$$

$$R_{\mathbf{C}}(g) = R_{\mathbf{A}}(g) + R_{\mathbf{B}}(g) \quad \text{or} \quad S_{\mathbf{C}}(t) = S_{\mathbf{A}}(t)S_{\mathbf{B}}(t) \tag{16}$$

when **A** and **B** are *free*.

Large random matrices are free if their eigenvectors are unrelated. In particular if say **B** is a rotationally invariant matrix independent of **A**.

$$T_{\mathbf{A}}(z) = \mathbf{A}(z\mathbf{1} - \mathbf{A})^{-1} \quad \mathfrak{t}_{\mathbf{A}}(z) = \tau\left(T_{\mathbf{A}}(z)\right) \tag{17}$$

$$S_{\mathbf{A}}(t) = \frac{t+1}{t z_{\mathbf{A}}(t)} \tag{18}$$

# Subordination Relations

$$\mathbf{C} = \mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2} \tag{19}$$

$$\mathfrak{t}_{\mathbf{C}}(z) = \mathfrak{t}_{\mathbf{A}}\left(S_{\mathbf{B}}(\mathfrak{t}_{\mathbf{C}}(z))z\right) \tag{20}$$

$$\mathbb{E}\left[T_C(z)\right]_{\mathbf{B}} = T_{\mathbf{A}}\left(S_{\mathbf{B}}(\mathfrak{t}_{\mathbf{C}}(z))z\right) \tag{21}$$

or more explicitly

$$\mathbb{E}\left[\mathbf{C}(z\mathbf{1} - \mathbf{C})^{-1}\right]_{\mathbf{B}} = \mathbf{A}(Z(z)\mathbf{1} - \mathbf{A})^{-1} \quad Z(z) = S_{\mathbf{B}}(\mathfrak{t}_{\mathbf{C}}(z))z \tag{22}$$

# Wishart Matrices

$$\mathbf{W} = \frac{1}{T}\mathbf{H}\mathbf{H}^T \text{ with } q = N/T \tag{23}$$

$$S_{\mathbf{W}}(t) = \frac{1}{1+qt} \qquad \Longleftrightarrow \qquad zt = (t+1)(1+qt) \tag{24}$$

Colored Wisharts

$$\mathbf{E}_{\mathbf{C}} = \mathbf{C}^{1/2}\mathbf{W}\mathbf{C}^{1/2} \quad \text{and} \quad \mathbf{E}_{\mathbf{\Sigma}} = \frac{1}{T}\mathbf{H}\mathbf{\Sigma}\mathbf{H}^T \tag{25}$$

$$\mathfrak{t}_{\mathbf{E}_{\mathbf{C}}}(z) = \mathfrak{t}_{\mathbf{E}_{\mathbf{C}}}\left(\frac{z}{1+q\mathfrak{t}_{\mathbf{E}_{\mathbf{C}}}(z)}\right) \quad \text{and} \quad q\mathfrak{t}_{\mathbf{E}_{\mathbf{\Sigma}}}(z) = \mathfrak{t}_{\mathbf{E}_{\mathbf{C}}}\left(\frac{z}{q+q\mathfrak{t}_{\mathbf{E}_{\mathbf{\Sigma}}}(z)}\right) \tag{26}$$

$$\boxed{\mathbb{E}\left[\mathbf{E}_{\mathbf{C}}(z\mathbf{1}-\mathbf{E}_{\mathbf{C}})^{-1}\right]_{\mathbf{W}} = \mathbf{C}(Z(z)\mathbf{1}-\mathbf{C})^{-1}} \text{ with } Z(z) = \frac{z}{1+q\mathfrak{t}_{\mathbf{E}_{\mathbf{C}}}(z)} \tag{27}$$

# Back to Phase Retrieval

$$\frac{1}{N}\operatorname{Tr}\left(\frac{\mathbf{H}\mathbf{D}_1\mathbf{H}^T}{T}\left(z\mathbf{1}-\frac{\mathbf{H}\mathbf{D}_0\mathbf{H}^T}{T}\right)^{-1}\right)=\frac{1}{N}\operatorname{Tr}\left(\mathbf{D}_1\mathbf{F}(z)\right), \tag{28}$$

$$\begin{aligned}\mathbf{F}(z)&=\frac{1}{T}\mathbf{H}^T\left(z\mathbf{1}-\frac{\mathbf{H}\mathbf{D}_0\mathbf{H}^T}{T}\right)^{-1}\mathbf{H}\\&=\sum_{n=0}^{\infty}\frac{1}{z^{n+1}}\frac{1}{T}\mathbf{H}^T\left(\frac{\mathbf{H}\mathbf{D}_0\mathbf{H}^T}{T}\right)^{n}\mathbf{H}\\&=\mathbf{D}_0^{-1/2}\left[\tilde{\mathbf{M}}\left(z\mathbf{1}-\tilde{\mathbf{M}}\right)^{-1}\right]\mathbf{D}_0^{-1/2}\quad\text{with}\quad\tilde{\mathbf{M}}:=q\mathbf{D}_0^{1/2}\mathbf{H}\mathbf{H}^T\mathbf{D}_0^{1/2}/N.\end{aligned} \tag{29}$$

$$\mathbb{E}\left[\mathbf{F}(z)\right]=\mathbf{D}_0^{-1/2}\left[q\mathbf{D}_0\left(\frac{z}{1+\tilde{q}\mathfrak{t}_{\tilde{\mathbf{M}}}(z)}\mathbf{1}-q\mathbf{D}_0\right)^{-1}\right]\mathbf{D}_0^{-1/2}. \tag{30}$$

Note that $\tilde{q}\mathfrak{t}_{\tilde{\mathbf{M}}}(z)=\mathfrak{t}_{\mathbf{M}}(z)$

# Phase Retrieval Subordination Relation

$$h(z) = c_1 + \frac{1}{T} \operatorname{Tr}\left(\mathbf{D}\left(Z(z)\mathbf{1} - q\mathbf{D}_0\right)^{-1}\right) \tag{31}$$

$$h(z) = \left\langle \frac{a^2 f(y)}{Z(z) - f(a^2)} \right\rangle, \tag{32}$$

$$Z(z) = \frac{z}{q + q\mathfrak{t}_{\mathbf{M}}(z)} \tag{33}$$

$$q\mathfrak{t}_{\mathbf{M}}(z) = \mathfrak{t}_{\mathbf{D}_0}(Z(z)) \tag{34}$$

$$= \left\langle \frac{f(y)}{Z(z) - f(a^2)} \right\rangle. \tag{35}$$

# Parametric solution

$$\varrho = \frac{1}{1 - h'(\lambda_1)} \quad \text{with} \quad \lambda_1 = h(\lambda_1) \tag{36}$$

3 equations for 3 unknowns ($\lambda_1$,$h$ and $Z$). Eliminate $\lambda_1$, work with $Z(\lambda_1) \equiv Z$

$$\lambda_1 = h(\lambda_1) = Z(\lambda_1)(q + q\mathfrak{t}_{\mathbf{M}}(z)) \equiv \psi_q(Z) \tag{37}$$

$$h(Z) = \left\langle \frac{a^2 f(y)}{Z - f(a^2)} \right\rangle \quad , \quad \psi_q(Z) = q + \left\langle \frac{f(y)}{Z - f(a^2)} \right\rangle \tag{38}$$

$$\varrho = \frac{\psi_q'(Z)}{\psi_q'(Z) - h'(Z)} \quad \text{with} \quad \psi_q(Z) = h(Z) \tag{39}$$

# Explicit Results in the noiseless case $(y = a^2)$

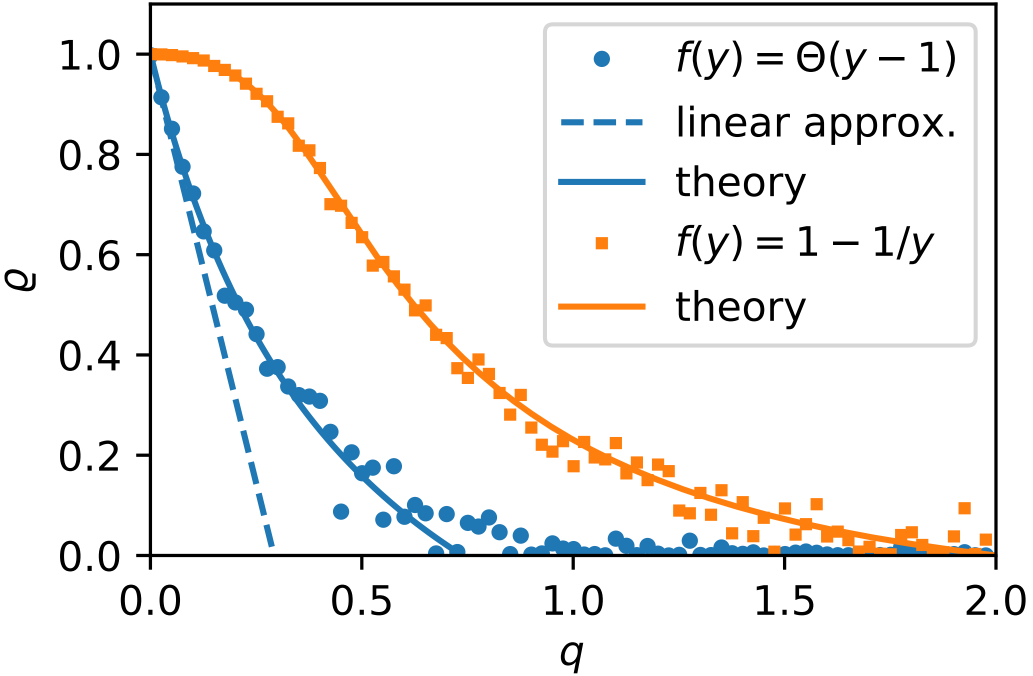
For $f(y) = \Theta(y-1)$:

$$\varrho = 1 - \frac{m_1 q}{(c_2 - m_1)^2 + (c_2 - m_1)q} \quad q < q^*, \tag{40}$$

Luo, Alghamdi and Lu have shown that $f(y) = 1 - 1/y$ is optimal for all $q$. In this case

$$q(Z) = \sqrt{\frac{\pi}{2}} Z^2 \frac{e^{\frac{1}{2(Z-1)}} \operatorname{erfc}\left(1/\sqrt{2(Z-1)}\right)}{(Z-1)^{5/2}} - \frac{Z}{(Z-1)^2} \quad , \quad \varrho(Z) = \frac{Z-1}{Z}. \quad \text{for } Z > 1 \tag{41}$$

# Theory vs Numerical simulations




Overlap $\varrho := |\mathbf{v}_1^T \mathbf{x}|^2 / |\mathbf{x}|^2$ between the largest eigenvector and the true signal as a function of $q := N/T$ for two functions: the simple $f(y) = \Theta(y-1)$ and the optimal $f(y) = 1 - 1/y$. Each dot correspond to a single matrix of aspect ratio $q$ and $NT = 10^7$.