Random Matrices, Free Probability & The Replica Method

Ralf R. Müller

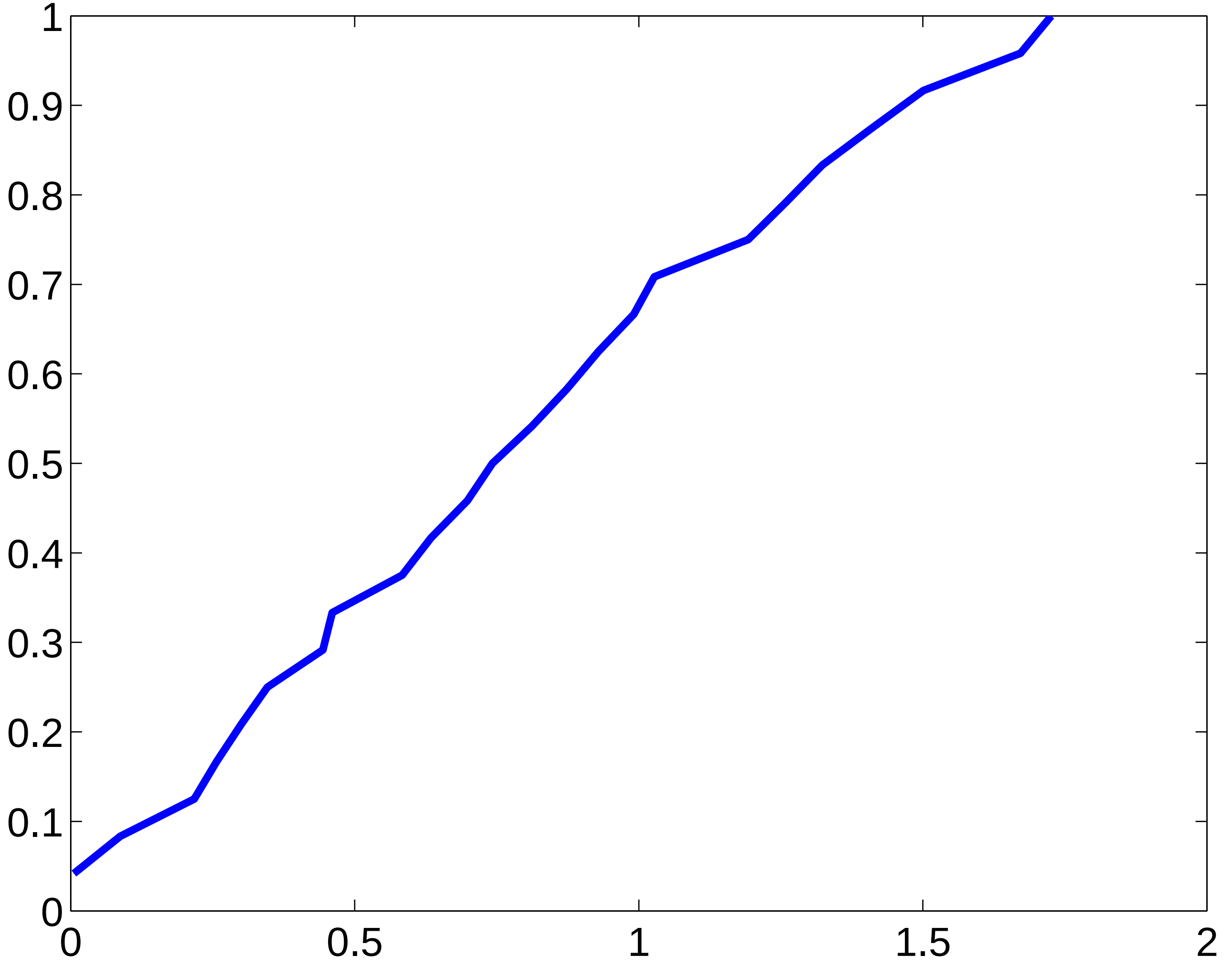
Vienna Telecommunications Research Centre (ftw.)

```
>>randn('state',0);
```

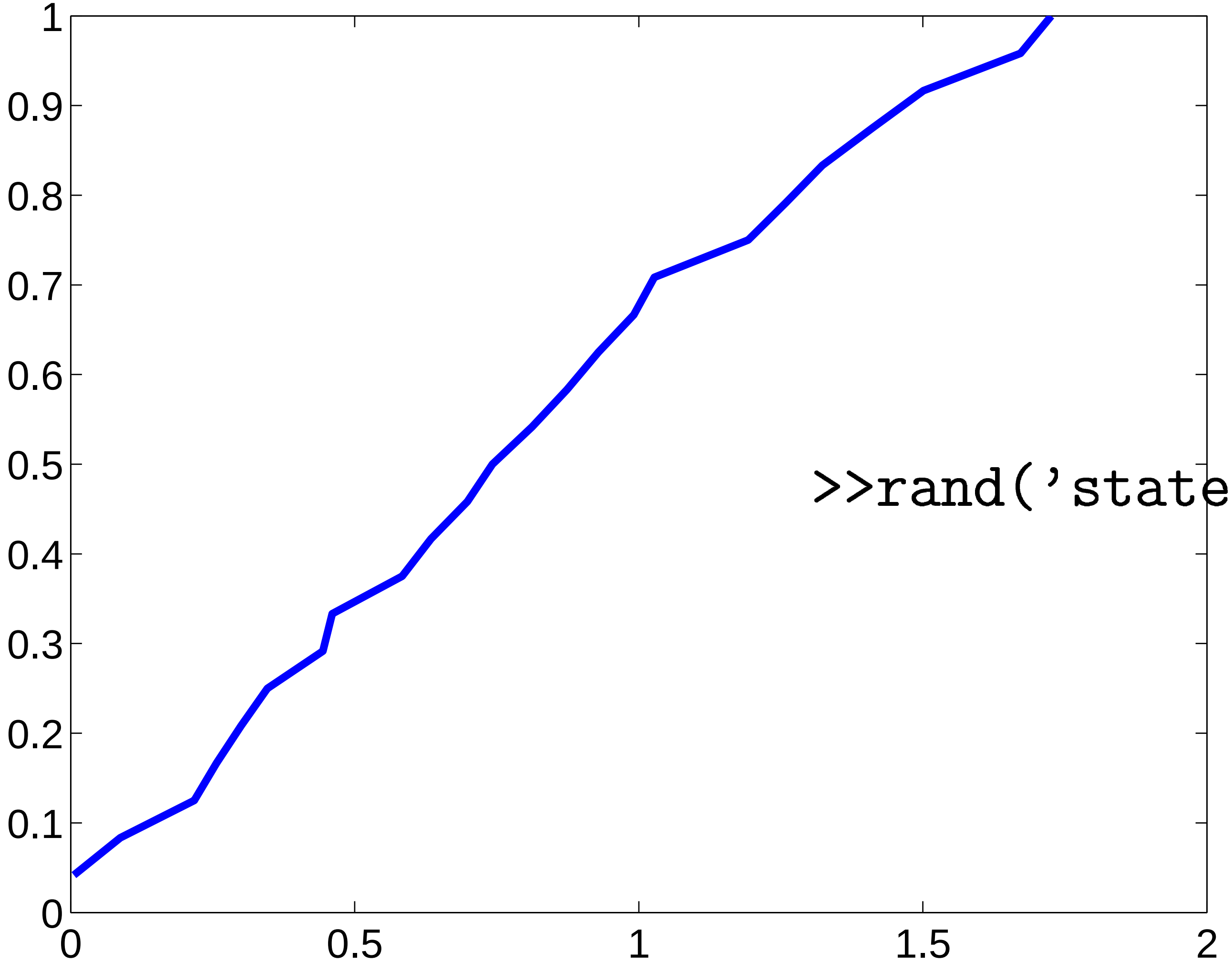
```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
```

```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
```

```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```

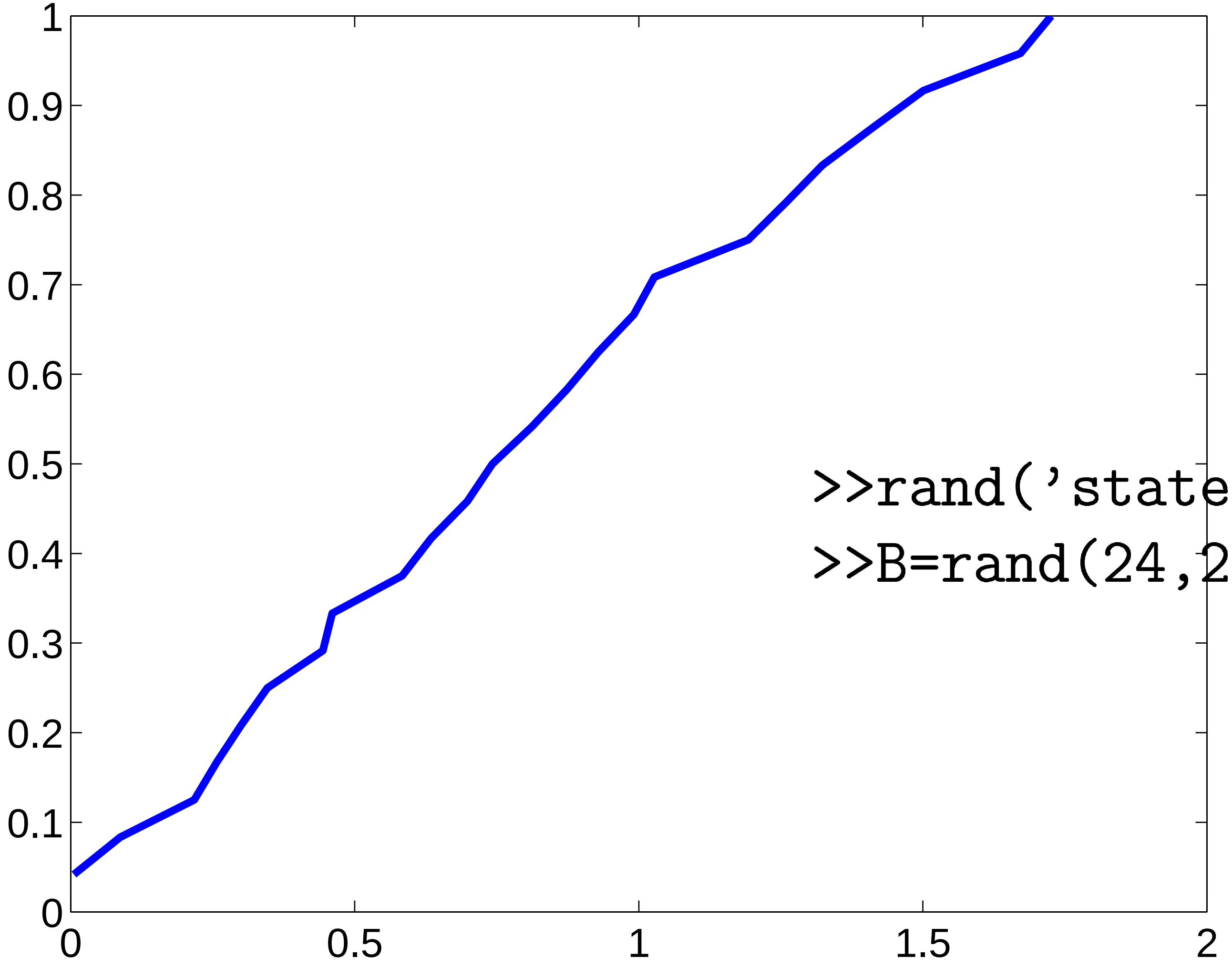



```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```



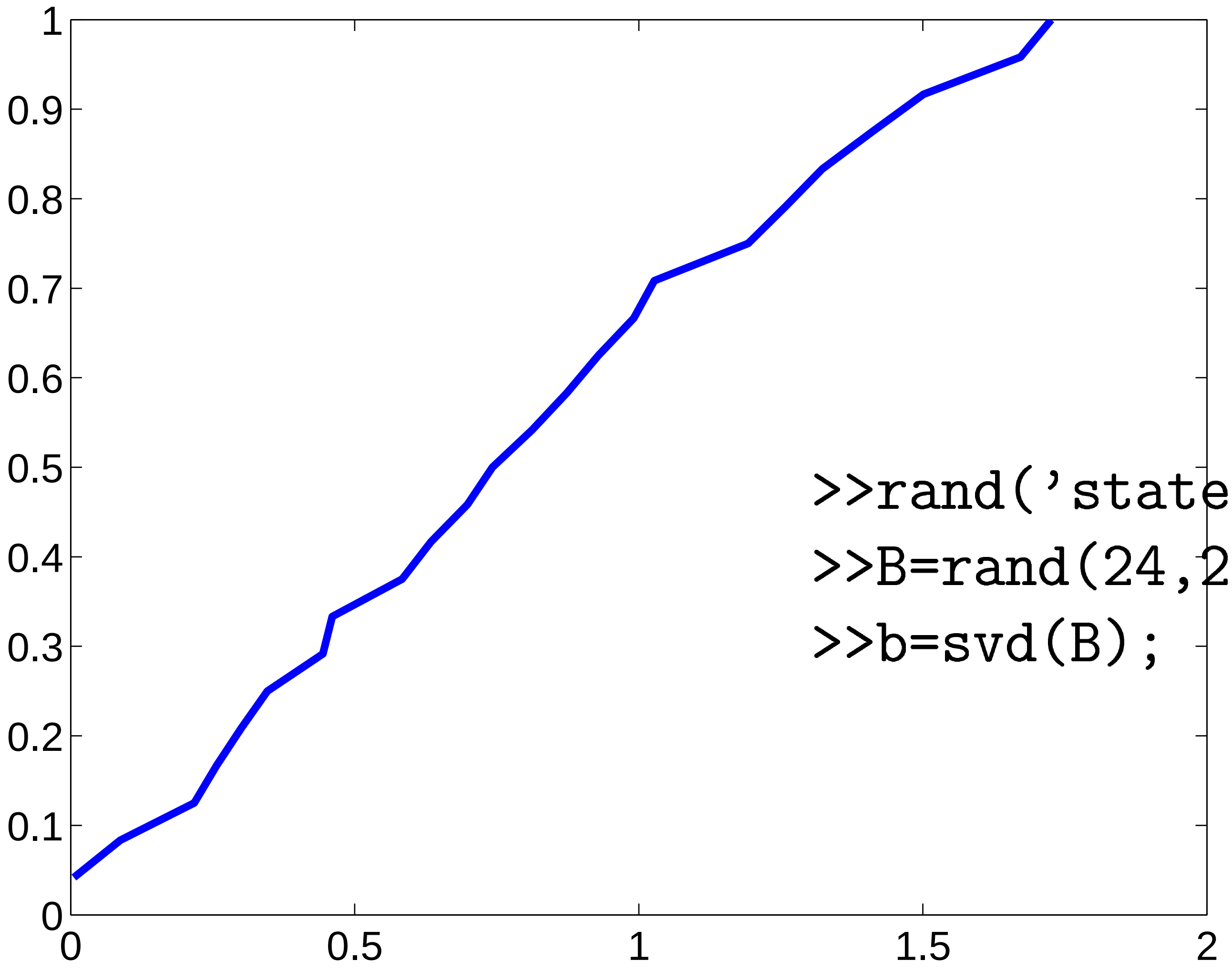

```
>>rand('state',0);
```

```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```



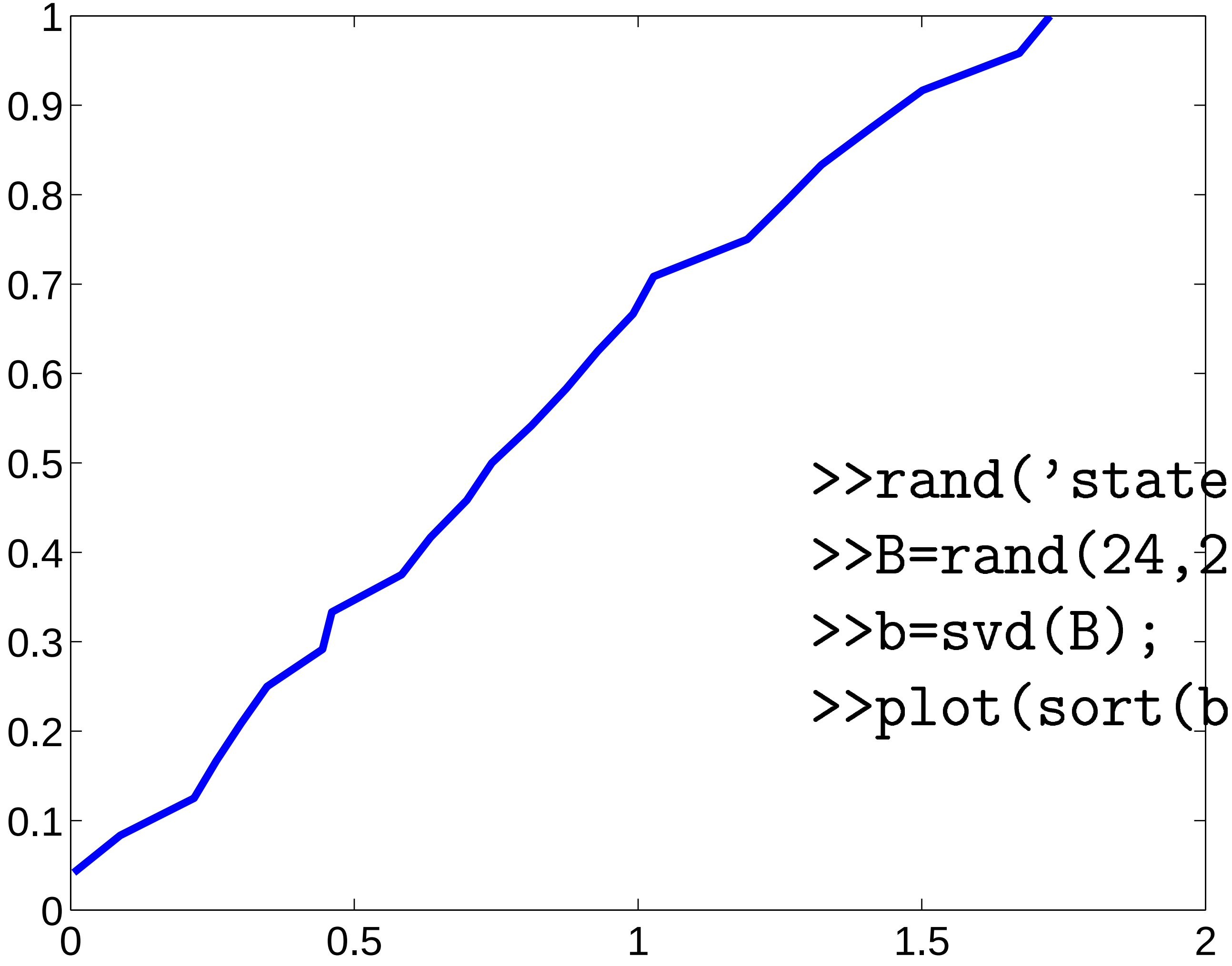

```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```

```
>>rand('state',0);
>>B=rand(24,24)/sqrt(2);
```



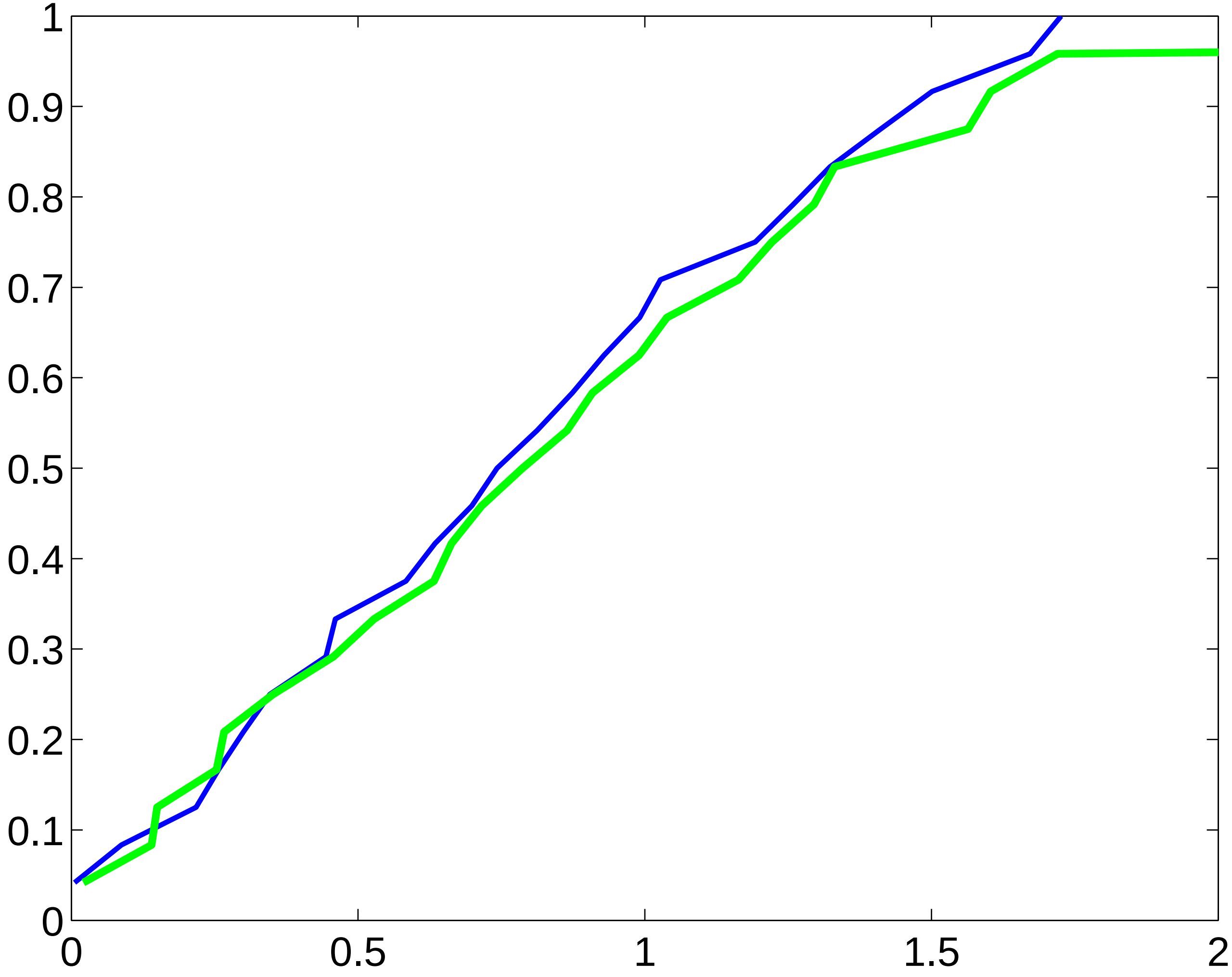

```
>>rand('state',0);
>>B=rand(24,24)/sqrt(2);
>>b=svd(B);
```

```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```



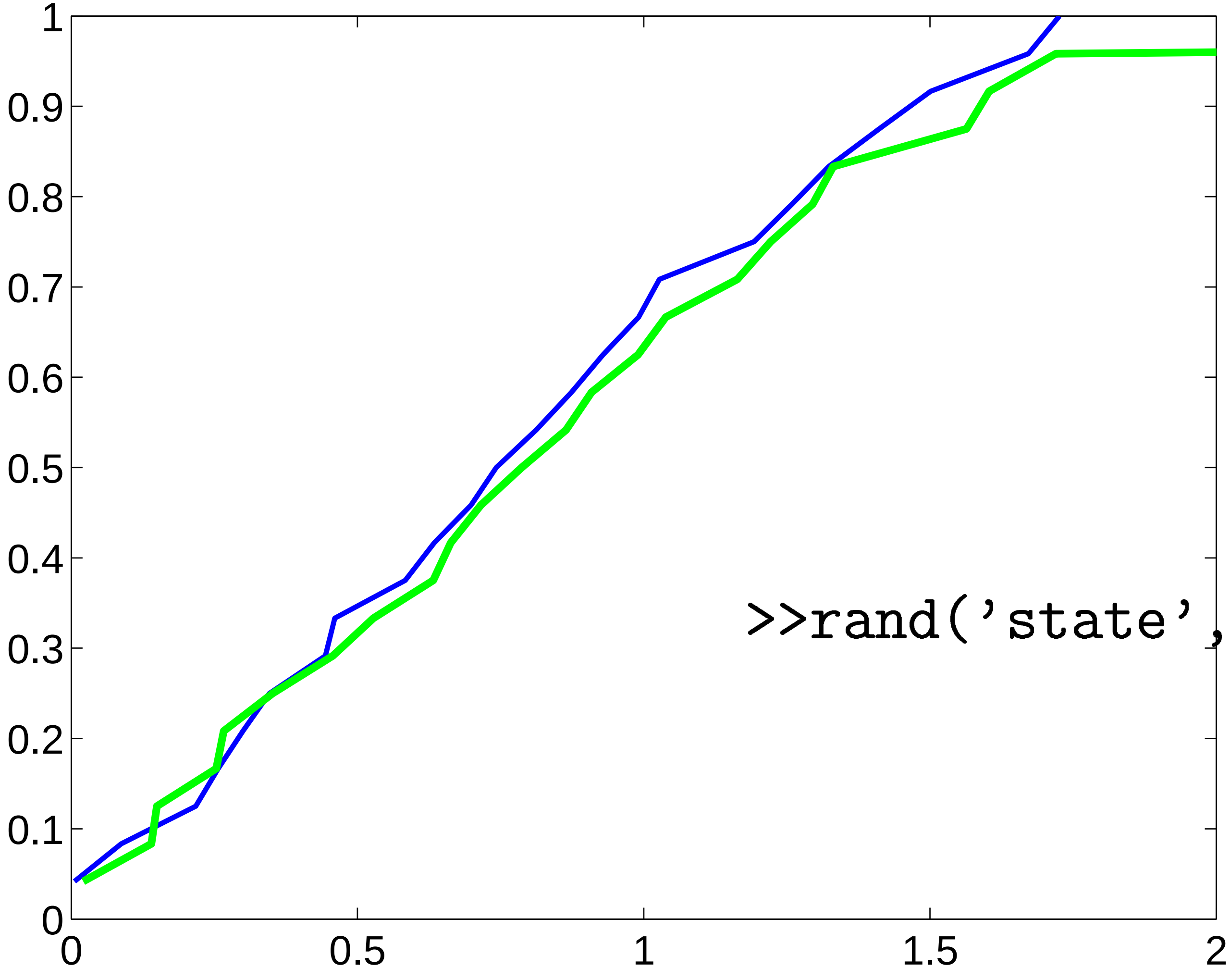

```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```

```
>>rand('state',0);
>>B=rand(24,24)/sqrt(2);
>>b=svd(B);
>>plot(sort(b),(1:24)/24)
```




```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```

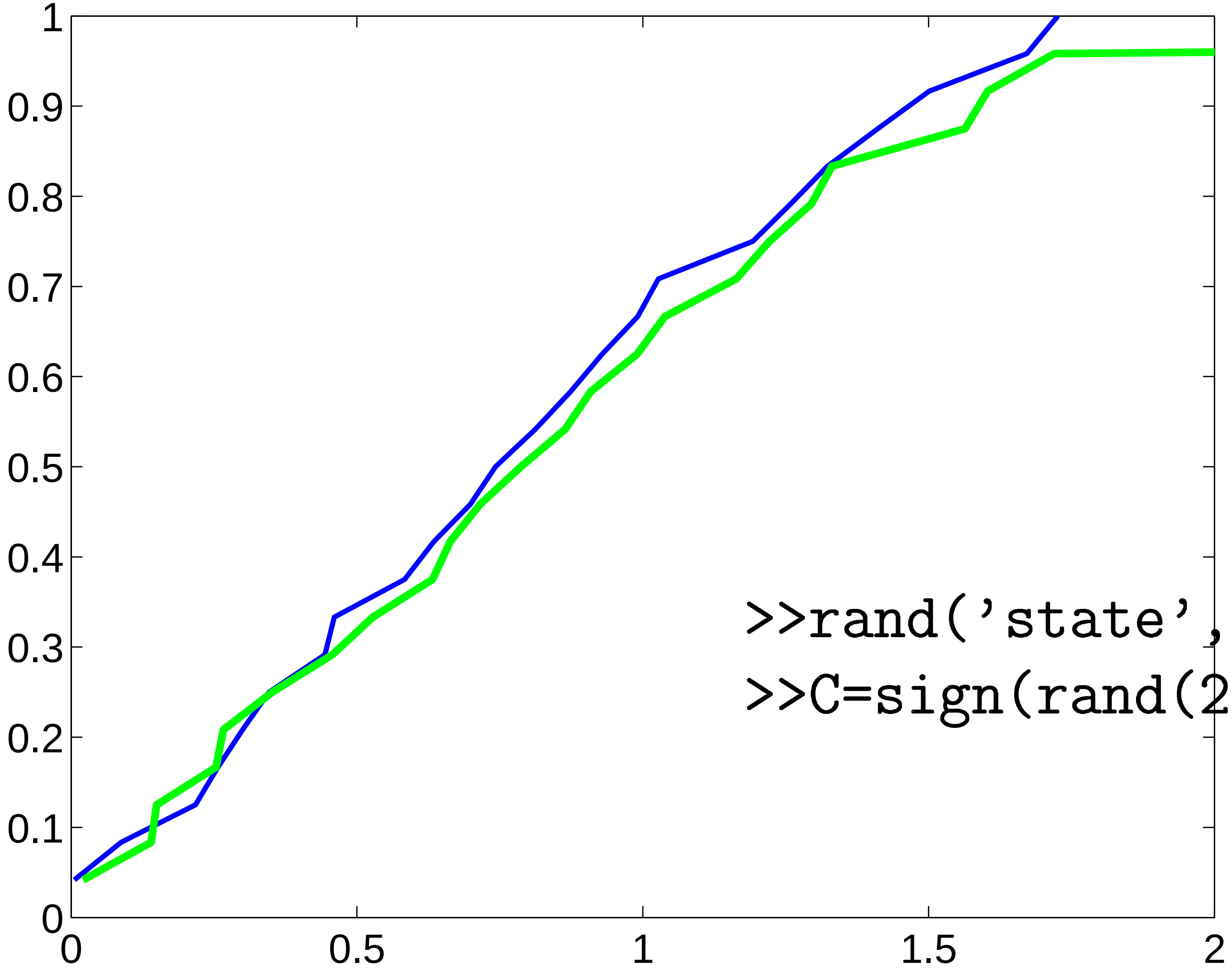
```
>>rand('state',0);
>>B=rand(24,24)/sqrt(2);
>>b=svd(B);
>>plot(sort(b),(1:24)/24)
```




```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```

```
>>rand('state',0);
>>B=rand(24,24)/sqrt(2);
>>b=svd(B);
>>plot(sort(b),(1:24)/24)
```

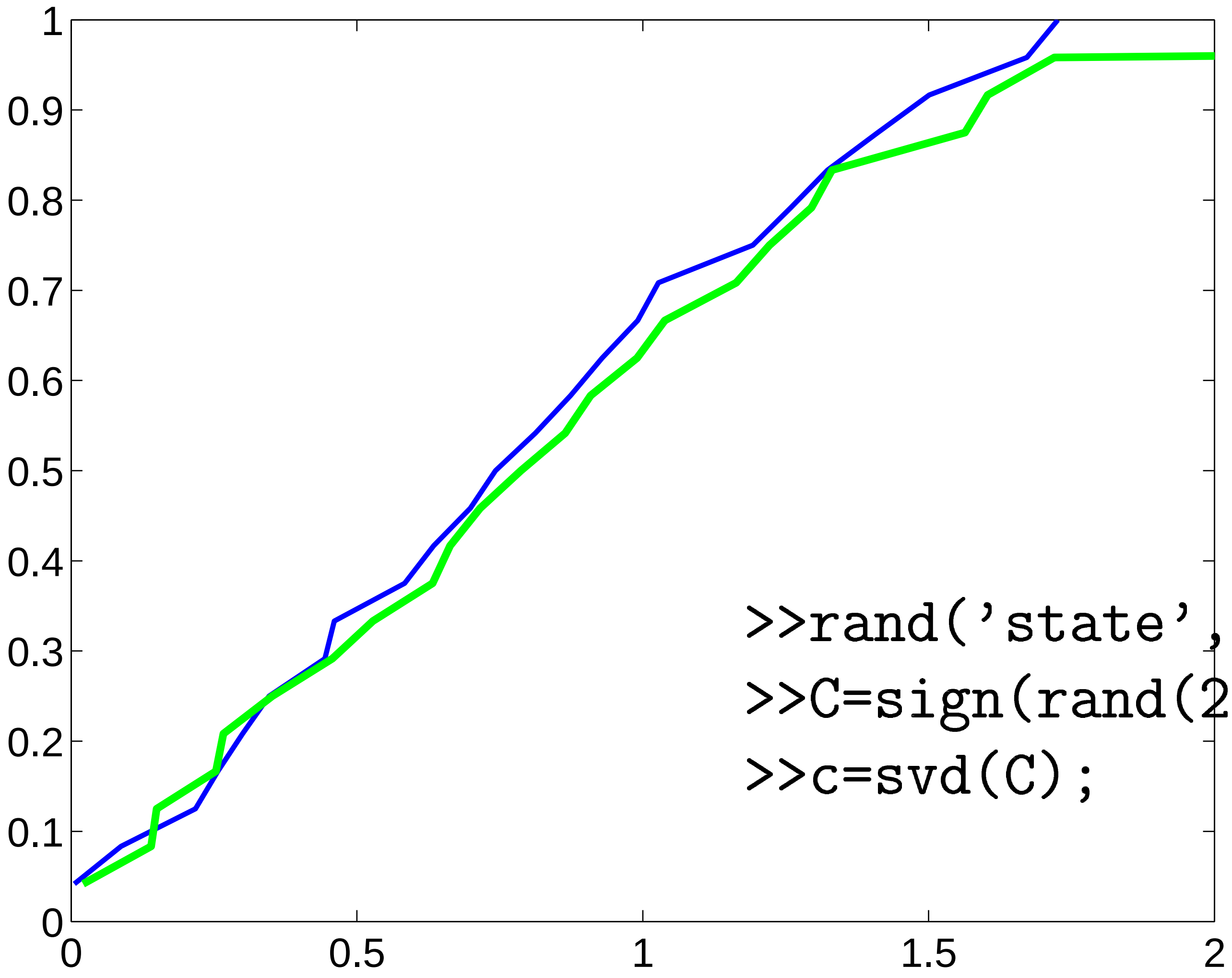
```
>>rand('state',0);
```




```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```

```
>>rand('state',0);
>>B=rand(24,24)/sqrt(2);
>>b=svd(B);
>>plot(sort(b),(1:24)/24)
```

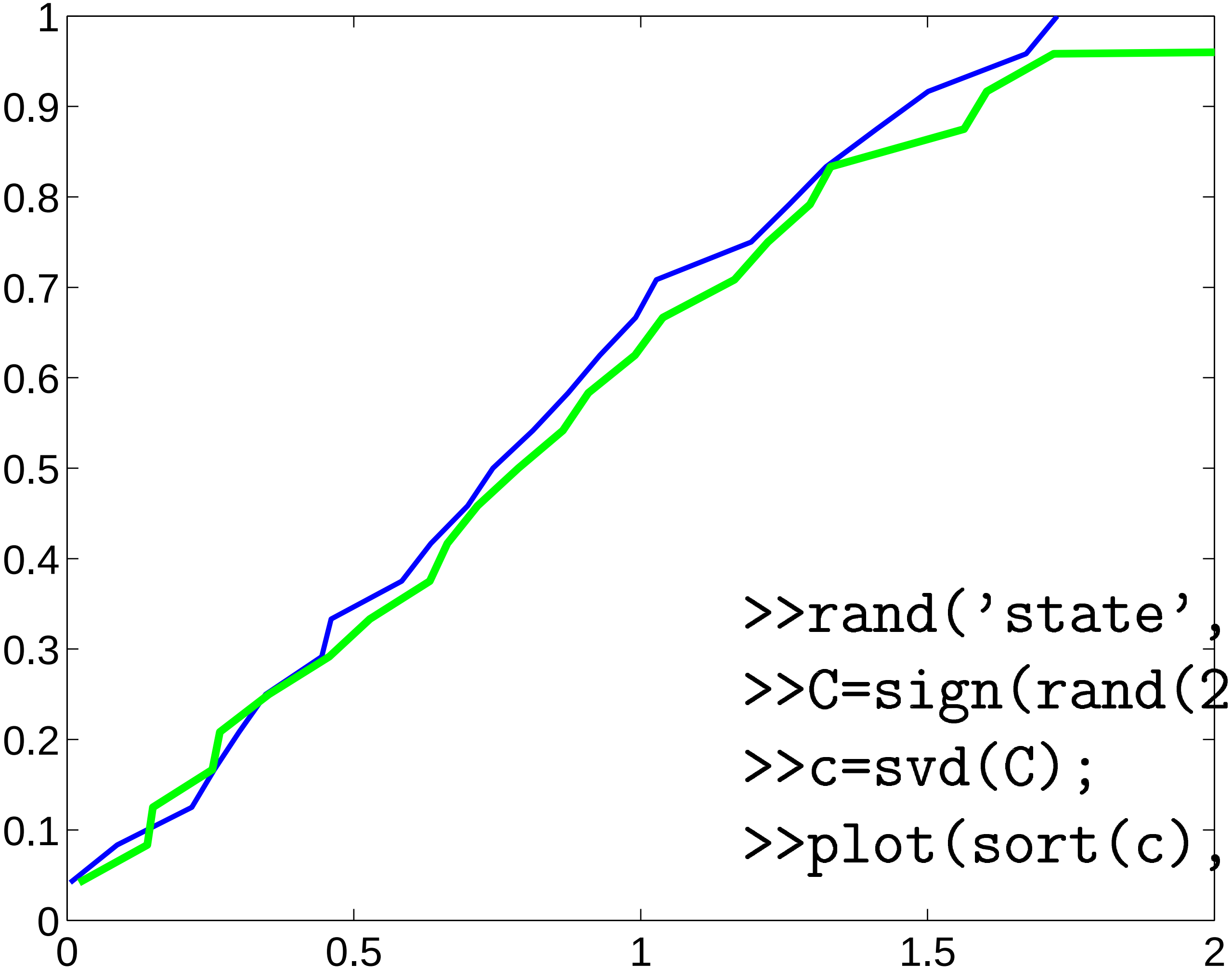
```
>>rand('state',0);
>>C=sign(rand(24,24)-.5)/sqrt(24);
```




```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```

```
>>rand('state',0);
>>B=rand(24,24)/sqrt(2);
>>b=svd(B);
>>plot(sort(b),(1:24)/24)
```

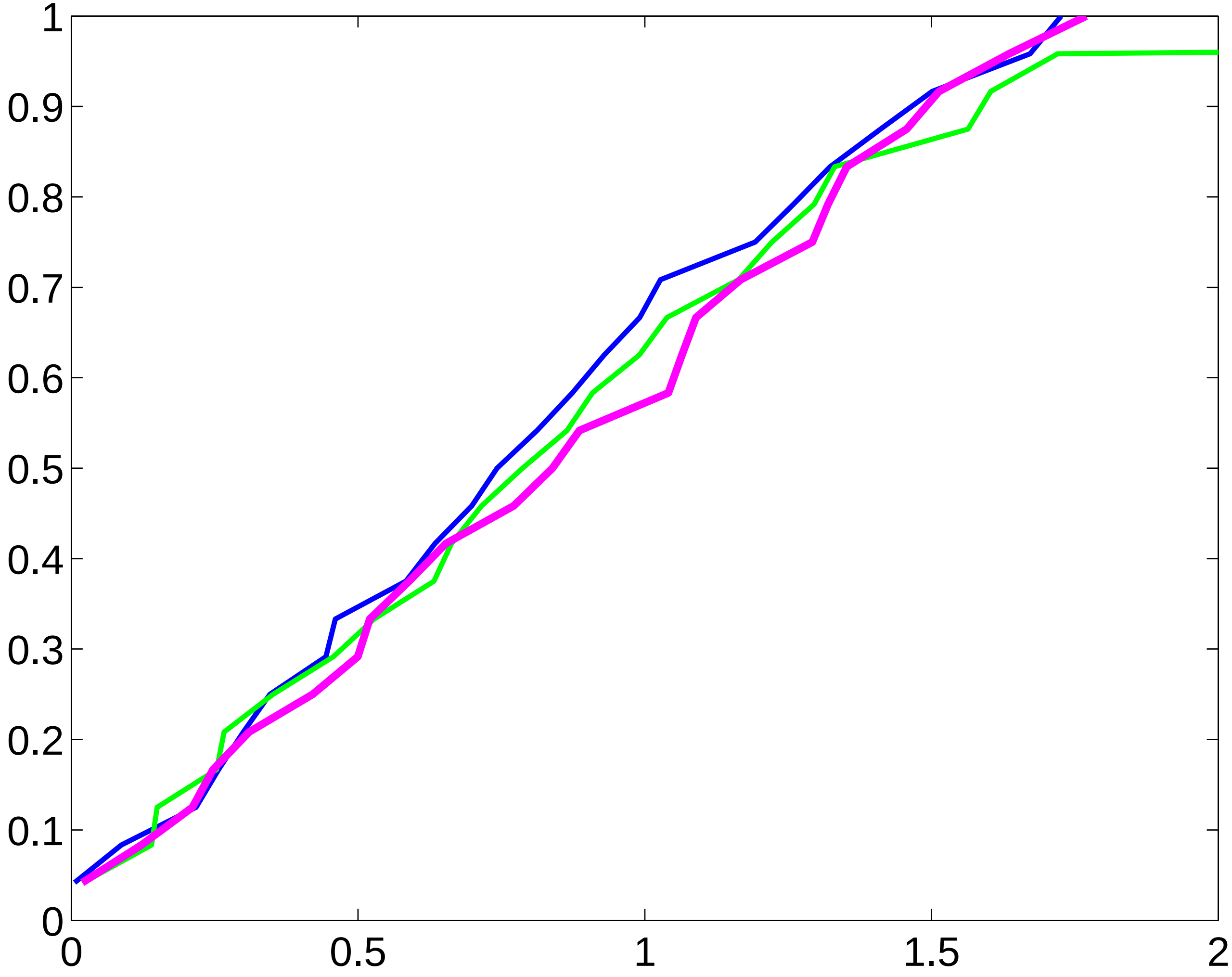
```
>>rand('state',0);
>>C=sign(rand(24,24)-.5)/sqrt(24);
>>c=svd(C);
```




```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)
```

```
>>rand('state',0);
>>B=rand(24,24)/sqrt(2);
>>b=svd(B);
>>plot(sort(b),(1:24)/24)
```

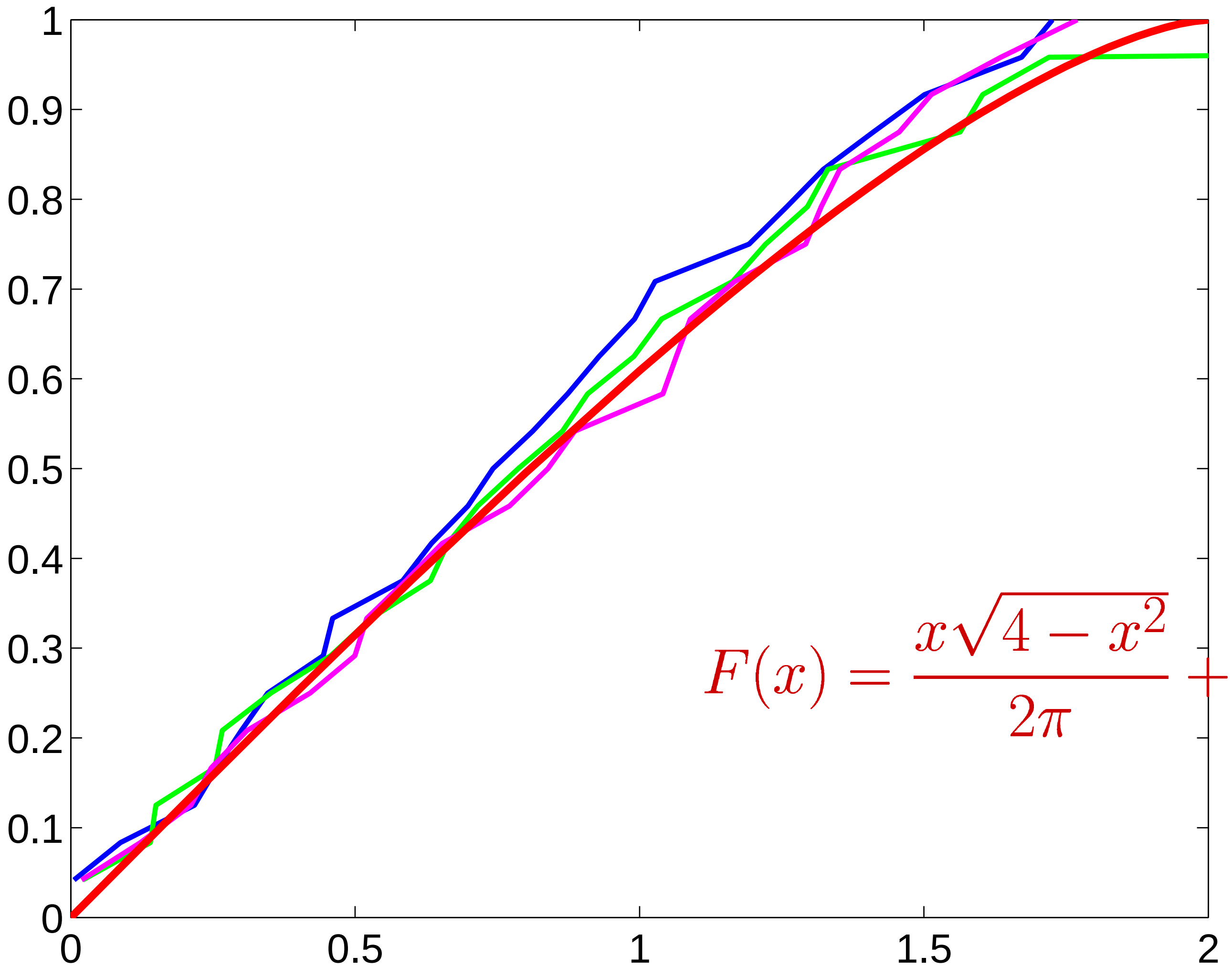
```
>>rand('state',0);
>>C=sign(rand(24,24)-.5)/sqrt(24);
>>c=svd(C);
>>plot(sort(c),(1:24)/24)
```



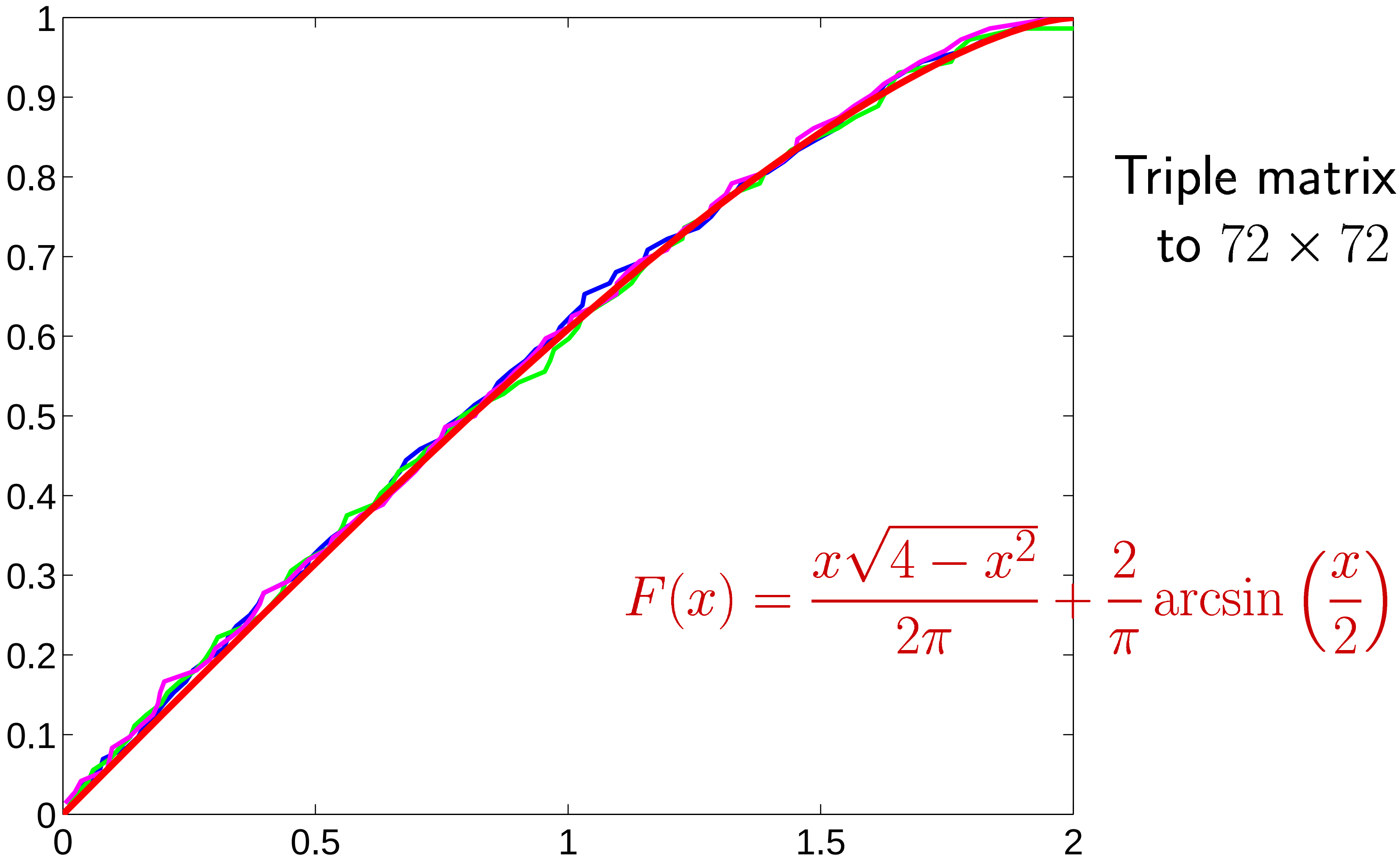

```
>>randn('state',0);
>>A=randn(24,24)/sqrt(24);
>>a=svd(A);
>>plot(sort(a),(1:24)/24)

>>rand('state',0);
>>B=rand(24,24)/sqrt(2);
>>b=svd(B);
>>plot(sort(b),(1:24)/24)

>>rand('state',0);
>>C=sign(rand(24,24)-.5)/sqrt(24);
>>c=svd(C);
>>plot(sort(c),(1:24)/24)
```

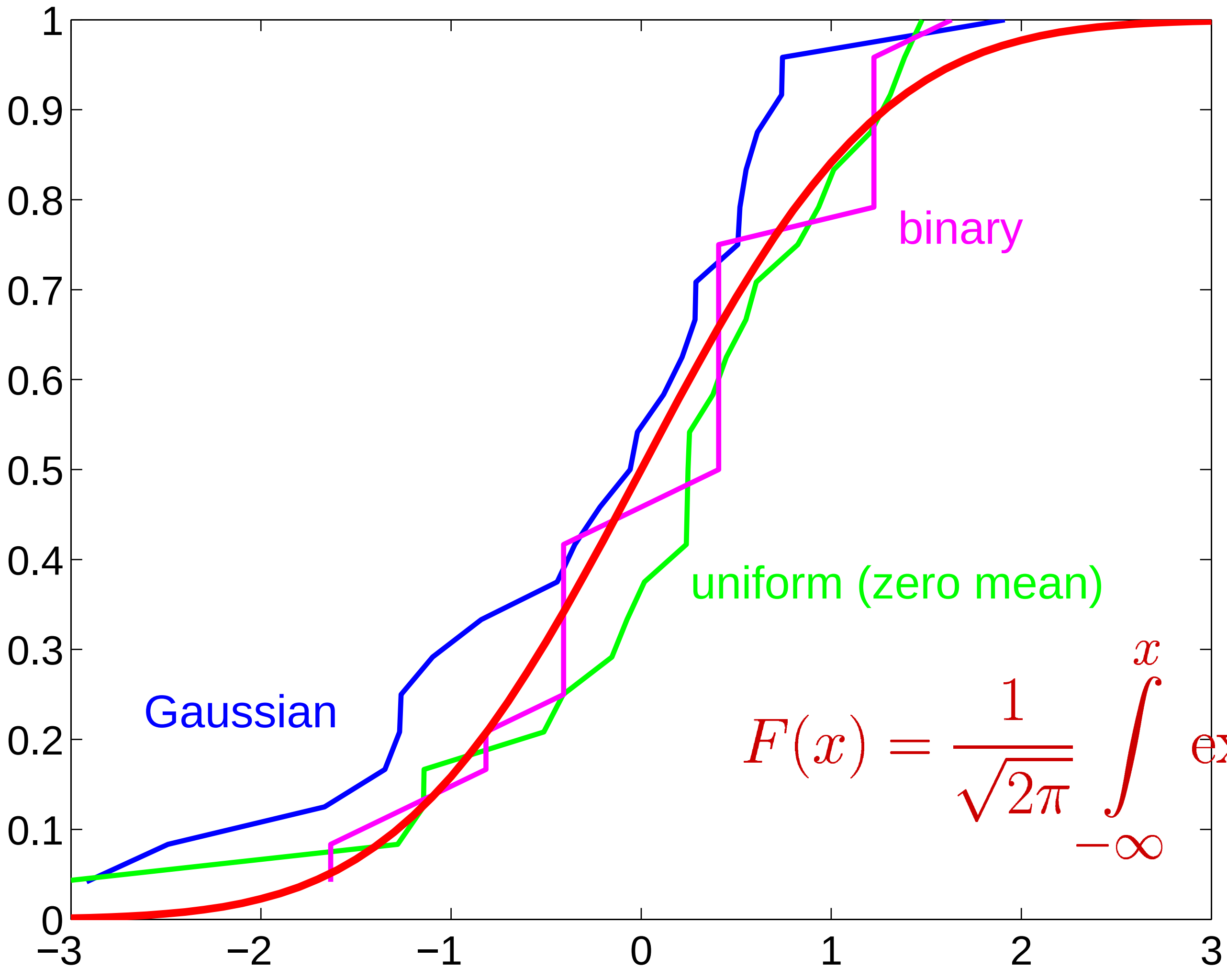


$$F(x) = \frac{x\sqrt{4-x^2}}{2\pi} + \frac{2}{\pi}\arcsin\left(\frac{x}{2}\right)$$



Triple matrix size to 72×72

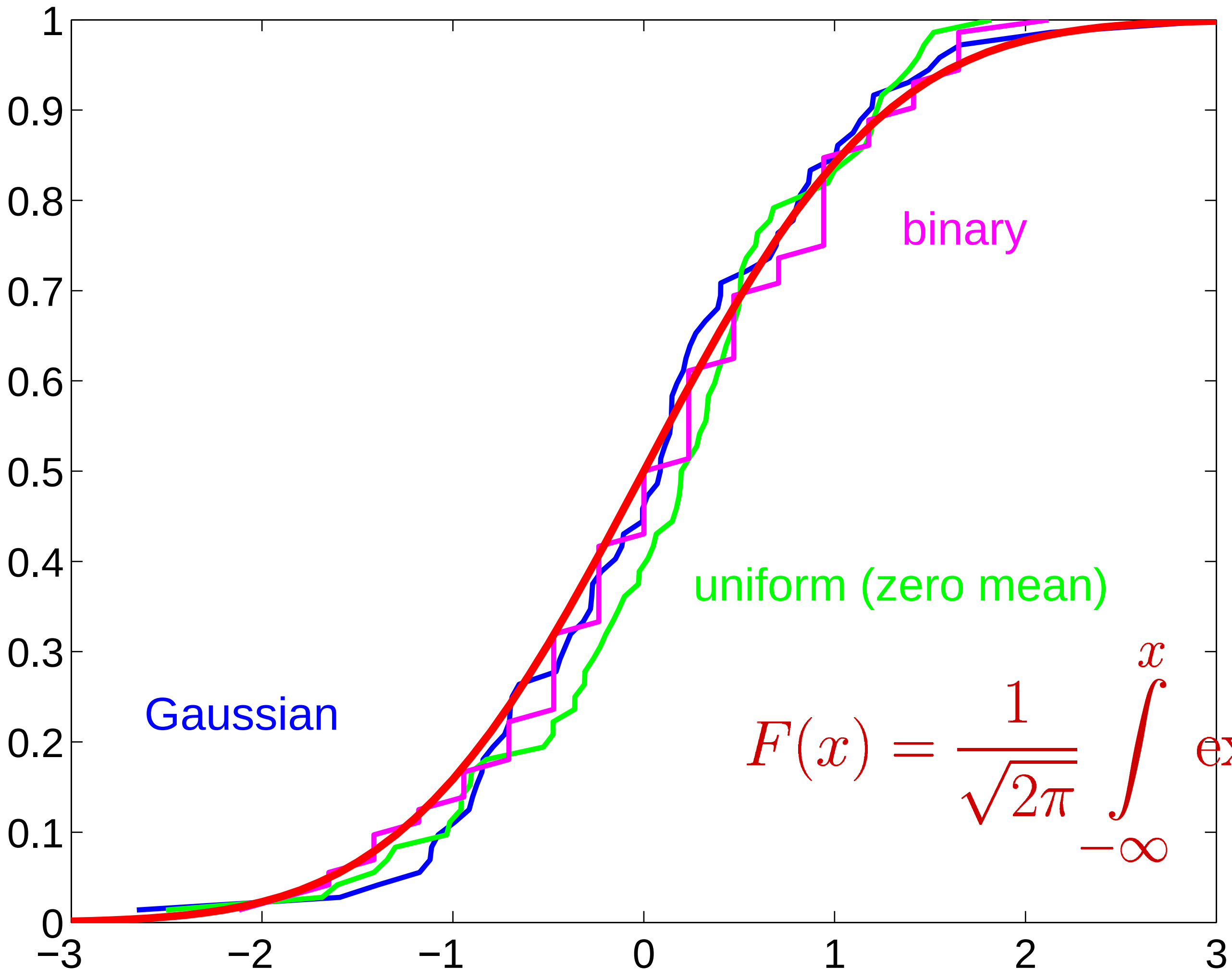
Change `svd(·)` *to* `sum(·)`*.*



Matrix size

24×24

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{\xi^2}{2}\right) d\xi$$



Matrix size

72×72

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{\xi^2}{2}\right) d\xi$$

Observations

- In both cases, the limit distribution did not depend on the distribution of the matrix entries.

- For `svd(·)` convergence is faster than for `sum(·)`.

- The limit distribution depends on the projection

$$f : \mathbb{R}^{K\times K} \mapsto \mathbb{R}^K.$$

Theories

Random Matrix Theory (RMT) considers the limit distributions for various projection functions f and various joint distributions of the matrix elements.

Free Probability Theory (FPT) considers the large random matrix as a single random operator and develops a probability theory for non-commutative operator algebras.

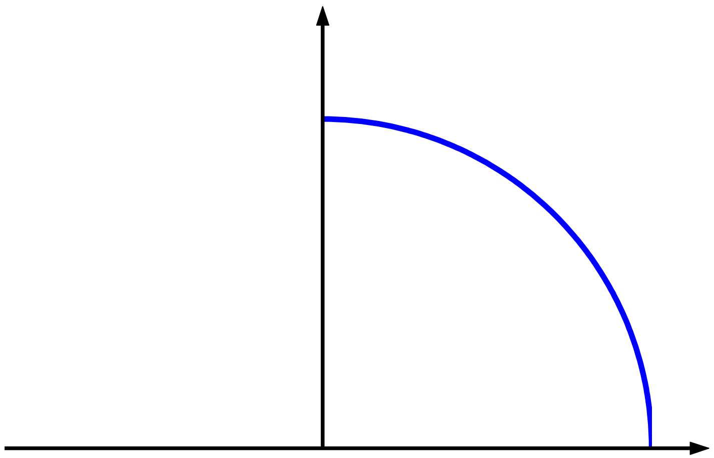
The two theories are relevant for communications engineering.

Quarter Circle Law

For a $K \times K$ random matrix with i.i.d. elements of variance $\frac{1}{K}$, the empirical distribution of the singular values converges almost surely to a deterministic limit distribution with density

$$f(x) = \frac{1}{\pi}\sqrt{4 - x^2} \qquad x \in [0; 2)$$

as $K \to \infty$.

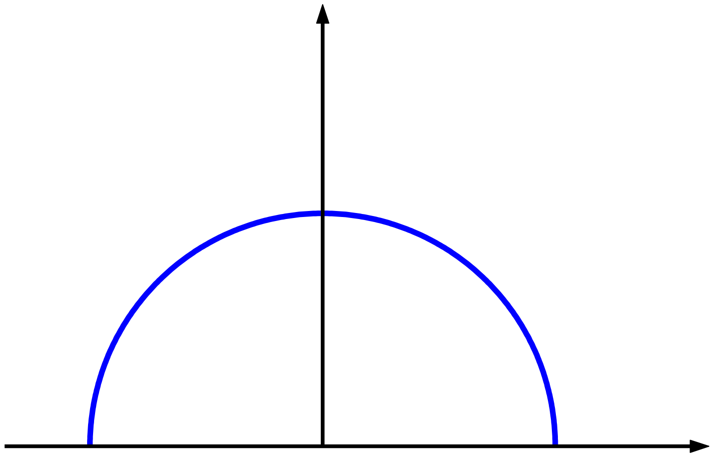


Semi-Circle Law

For a $K \times K$ random matrix $\boldsymbol{H}$ with i.i.d. zero mean elements of variance $\frac{1}{K}$, the empirical distribution of the eigenvalues of $\frac{1}{2}(\boldsymbol{H}+\boldsymbol{H}^\dagger)$ converges almost surely to a deterministic limit distribution with density

$$f(x) = \frac{1}{\pi}\sqrt{2-x^2} \qquad x \in (-\sqrt{2}; +\sqrt{2})$$

as $K \to \infty$.



Full Circle Law

For a $K \times K$ random matrix with i.i.d. zero-mean elements of variance $\frac{1}{K}$, the empirical distribution of the eigenvalues converges almost surely to a deterministic limit distribution with density

$$f(z) = \frac{1}{\pi} \qquad |z| < 1$$

as $K \to \infty$.

Uniform distribution on the complex unit disk.

Haar Distribution

For a $K \times K$ random matrix $\boldsymbol{H}$ with i.i.d. zero-mean Gaussian elements of finite variance, the empirical distribution of the eigenvalues of $\boldsymbol{H}(\boldsymbol{H}^\dagger \boldsymbol{H})^{-\frac{1}{2}}$ converges almost surely to a deterministic limit distribution with density

$$f(z) = \frac{1}{2\pi} \qquad |z| = 1$$

as $K \to \infty$.

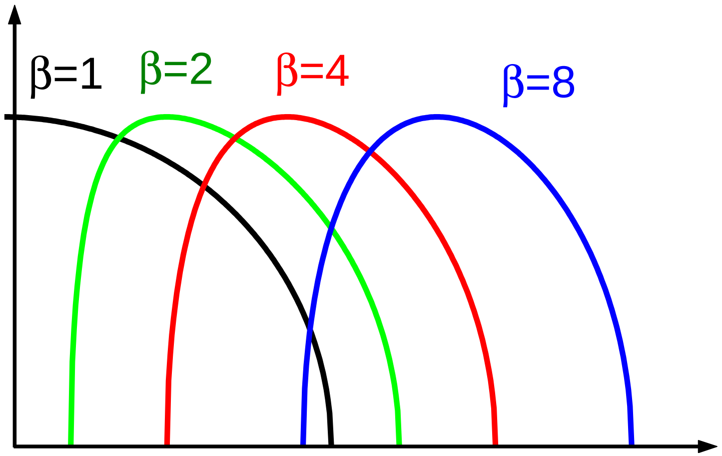
Uniform distribution on the complex unit circle.

Deformed Quarter Circle Law

For an $N \times K$, $N < K$ random matrix with i.i.d. elements of variance $\frac{1}{N}$, the empirical distribution of the singular values converges almost surely to a deterministic limit distribution with density

$$f(x) = \frac{\sqrt{4\beta - (x^2 - 1 - \beta)^2}}{\pi x} \qquad x \in (\sqrt{\beta} - 1; \sqrt{\beta} + 1)$$

as $K = \beta N \to \infty$.



The Stieltjes Transform

The densities for most other projections cannot be given in explicit form.
They are more easily characterized in terms of their Stieltjes transforms

$$G(s) \triangleq \int \frac{f(x)dx}{x-s} \qquad \mathrm{Im}(s) > 0.$$

The Stieltjes Transform

The densities for most other projections cannot be given in explicit form. They are more easily characterized in terms of their Stieltjes transforms

$$G(s) \triangleq \int \frac{f(x)dx}{x-s} \qquad \mathrm{Im}(s) > 0.$$

Girko's Law: The Stieltjes transform can be used to characterize the asymptotic singular value distribution of a random matrix with independent but arbitrarily distributed entries.

The Stieltjes Transform

The densities for most other projections cannot be given in explicit form. They are more easily characterized in terms of their Stieltjes transforms

$$G(s) \triangleq \int \frac{f(x)dx}{x-s} \qquad \mathrm{Im}(s) > 0.$$

Girko's Law: The Stieltjes transform can be used to characterize the asymptotic singular value distribution of a random matrix with independent but arbitrarily distributed entries.

Stieltjes Inversion Formula:

$$f(x) = \lim_{y \to 0+} \frac{1}{\pi} \mathrm{Im}\left[G(x + \mathrm{j}y)\right]$$

Classical vs. Free Probability

Random matrix is ensemble	Random matrix is a sample

Classical vs. Free Probability

Classical	Free
Random matrix is ensemble	Random matrix is a sample
Multiplication is commutative	Multiplication is *not* commutative

Classical vs. Free Probability

Classical	Free
Random matrix is ensemble	Random matrix is a sample
Multiplication is commutative	Multiplication is *not* commutative

Non-commutative joint moments:

$$\mathrm{E}\left\{\boldsymbol{A}^2\boldsymbol{B}^2\right\} \neq \mathrm{E}\left\{\boldsymbol{ABAB}\right\}$$

Classical vs. Free Probability

Random matrix is ensemble	Random matrix is a sample
Multiplication is commutative	Multiplication is *not* commutative

Non-commutative joint moments:

$$\mathrm{E}\left\{\boldsymbol{A}^2\boldsymbol{B}^2\right\} \neq \mathrm{E}\{\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}\boldsymbol{B}\}$$

Statistical independence cannot be defined without respect to the elements of the matrix.

Independence vs. Freeness

The Definition of Freeness

Define

$$\operatorname{Tr}(\boldsymbol{X}) = \lim_{K\to\infty} \frac{1}{K} \operatorname{trace}(\boldsymbol{X}).$$

Two random matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are asymptotically free, if for all $n > 0$ and any sequences of non-negative integers $(\alpha_1, \ldots, \alpha_n)$, $(\beta_1, \ldots, \beta_n)$ for which

$$\operatorname{Tr}(\boldsymbol{A}^{\alpha_1}) = \cdots = \operatorname{Tr}(\boldsymbol{A}^{\alpha_n}) = 0$$

and

$$\operatorname{Tr}\left(\boldsymbol{B}^{\beta_1}\right) = \cdots = \operatorname{Tr}\left(\boldsymbol{B}^{\beta_n}\right) = 0,$$

we have

$$\operatorname{Tr}\left(\boldsymbol{A}^{\alpha_1}\boldsymbol{B}^{\beta_1}\cdots\boldsymbol{A}^{\alpha_n}\boldsymbol{B}^{\beta_n}\right) = 0.$$

The generalization to more than two random variables is not straightforward.

Examples for Freeness

- Two independent Haar distributed random matrices are asymptotically free.
- Two i.i.d. Gaussian distributed random matrices are asymptotically free.
- A Haar or i.i.d. Gaussian distributed random matrix is asymptotically free from a constant matrix.

Caveat: There exist independent random matrices which are not asymptotically free and dependent random matrices which are.

Additive Free Convolution

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be free and

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}.$$

Additive Free Convolution

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be free and

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}.$$

Then,

$$R_{\boldsymbol{C}}(w) = R_{\boldsymbol{A}}(w) + R_{\boldsymbol{B}}(w).$$

Additive Free Convolution

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be free and

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}.$$

Then,

$$R_{\boldsymbol{C}}(w) = R_{\boldsymbol{A}}(w) + R_{\boldsymbol{B}}(w).$$

The R-transform is defined as

$$R(w) \stackrel{\triangle}{=} G^{-1}(-w) - \frac{1}{w}.$$

The R-transform linearizes additive free convolution.

Multiplicative Free Convolution

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be free and

$$\boldsymbol{D} = \boldsymbol{A}\boldsymbol{B}.$$

Then,

$$S_{\boldsymbol{D}}(z) = S_{\boldsymbol{A}}(z) S_{\boldsymbol{B}}(z).$$

The S-transform is defined as

$$S(z) \triangleq \frac{1+z}{z} \Upsilon^{-1}(z) \qquad \text{with} \qquad \Upsilon(s) \triangleq -1 - \frac{G^{-1}\left(\frac{1}{s}\right)}{s}.$$

The S-transform linearizes multiplicative free convolution.

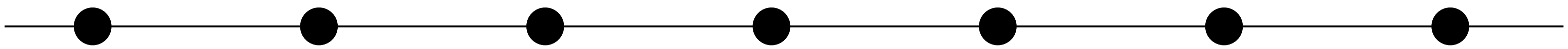
The Two Theories

. . . solve most engineering problems which are governed by the singular or eigenvalues of large random matrices:

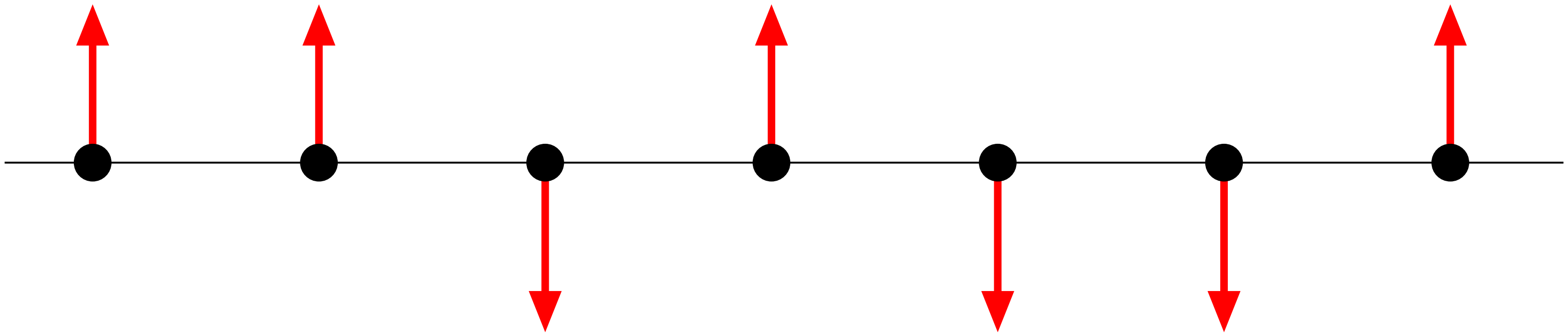
- Mutual information calculations
- SINR calculations of linear receivers
- Asymptotic design of linear multistage receivers
- Power control for iterative multiuser decoders

Research is ongoing in the math community and frequently new results come up.

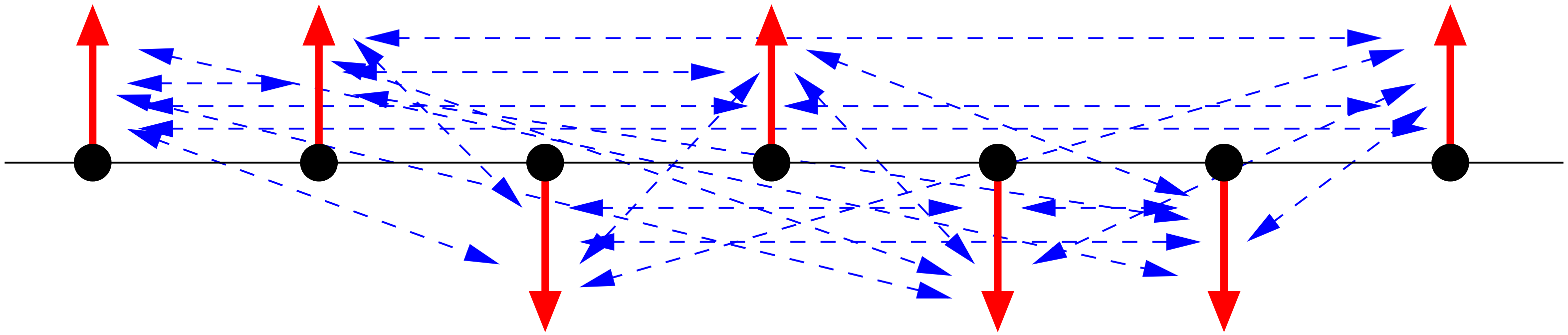
Spin Glasses



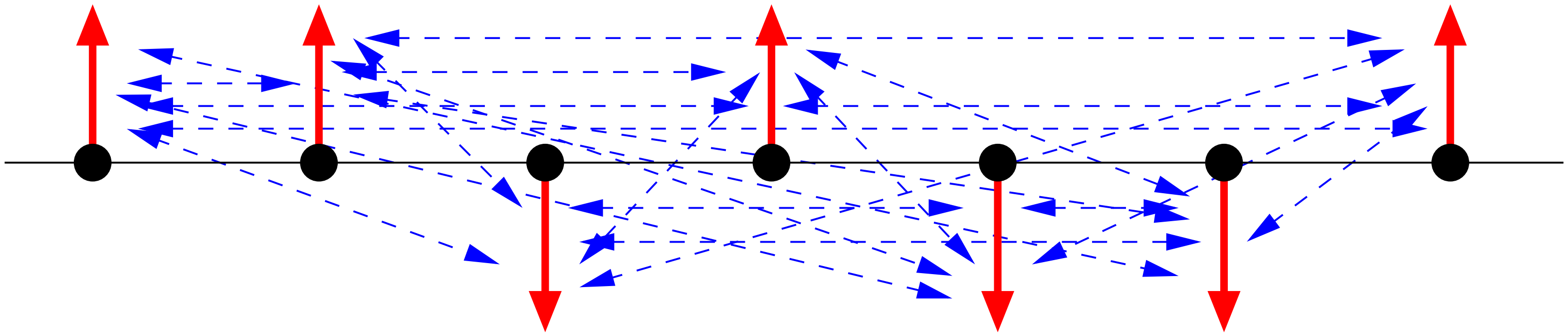
Spin Glasses



Spin Glasses



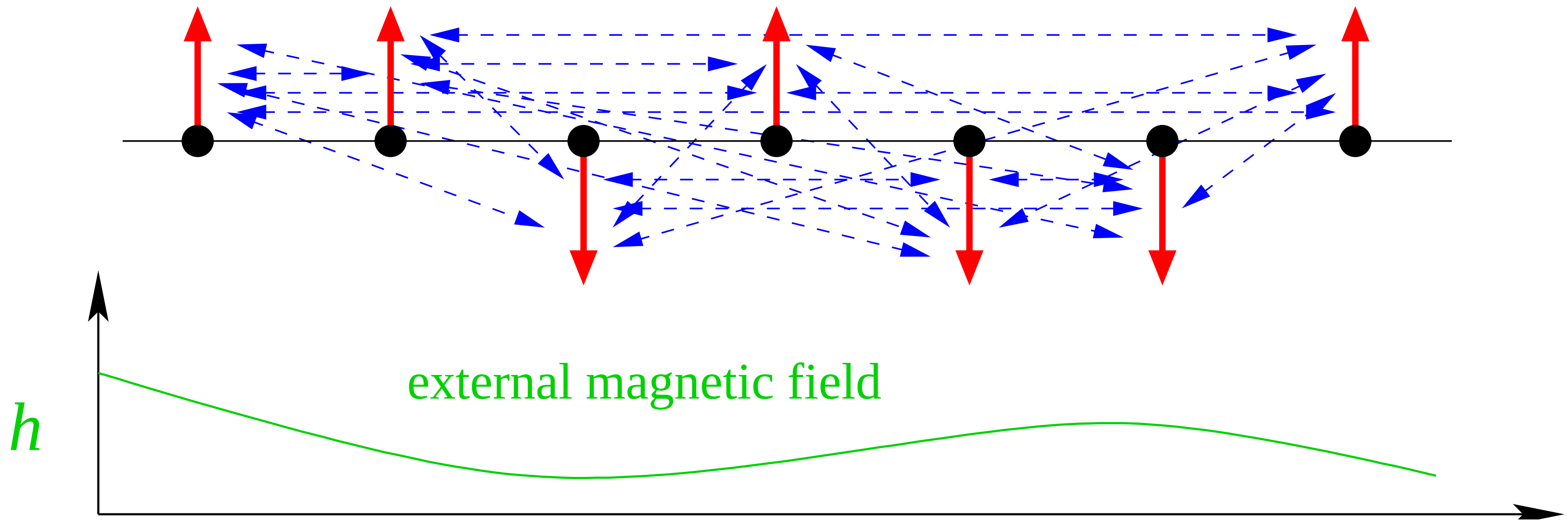
Spin Glasses



Energy function (Hamiltonian):

$$-\sum_{i}\sum_{j<i} r_{ij} x_i x_j$$

Spin Glasses



Energy function (Hamiltonian):

$$-\sum_i \sum_{j<i} r_{ij} x_i x_j - \sum_i h_i x_i$$

Optimal Detection of Vector Channel

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{n}$$

Optimal Detection of Vector Channel

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{n}$$

Best estimate for transmitted data:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x} \in \{\pm 1\}^K}{\operatorname{argmin}} ||\boldsymbol{y} - \boldsymbol{S}\boldsymbol{x}||$$

Optimal Detection of Vector Channel

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{n}$$

Best estimate for transmitted data:

$$\begin{aligned} \hat{\boldsymbol{x}} &= \underset{\boldsymbol{x}\in\{\pm 1\}^K}{\operatorname{argmin}} \; ||\boldsymbol{y} - \boldsymbol{S}\boldsymbol{x}|| \\ &= \underset{\boldsymbol{x}\in\{\pm 1\}^K}{\operatorname{argmin}} \; -\frac{1}{2}\boldsymbol{x}^\dagger \boldsymbol{R}\boldsymbol{x} - \boldsymbol{h}^\dagger \boldsymbol{x} + \boldsymbol{y}^\dagger \boldsymbol{y} \end{aligned}$$

with

$$\begin{aligned} \boldsymbol{R} &= -2\boldsymbol{S}^\dagger \boldsymbol{S} \\ \boldsymbol{h} &= 2\boldsymbol{S}^\dagger \boldsymbol{y} \end{aligned}$$

Optimal Detection of Vector Channel

$$y = Sx + n$$

Best estimate for transmitted data:

$$\begin{aligned} \hat{x} &= \operatorname*{argmin}_{x \in \{\pm 1\}^K} ||y - Sx|| \\ &= \operatorname*{argmin}_{x \in \{\pm 1\}^K} -\frac{1}{2} x^\dagger R x - h^\dagger x \qquad \text{with} \qquad \begin{aligned} R &= -2S^\dagger S \\ h &= 2S^\dagger y \end{aligned} \\ &= \operatorname*{argmin}_{x \in \{\pm 1\}^K} -\sum_i \sum_{j<i} r_{ij} x_i x_j - \sum_i h_i x_i - \frac{1}{2} \sum_i r_{ii} x_i^2 \end{aligned}$$

Optimal Detection of Vector Channel

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{n}$$

Best estimate for transmitted data:

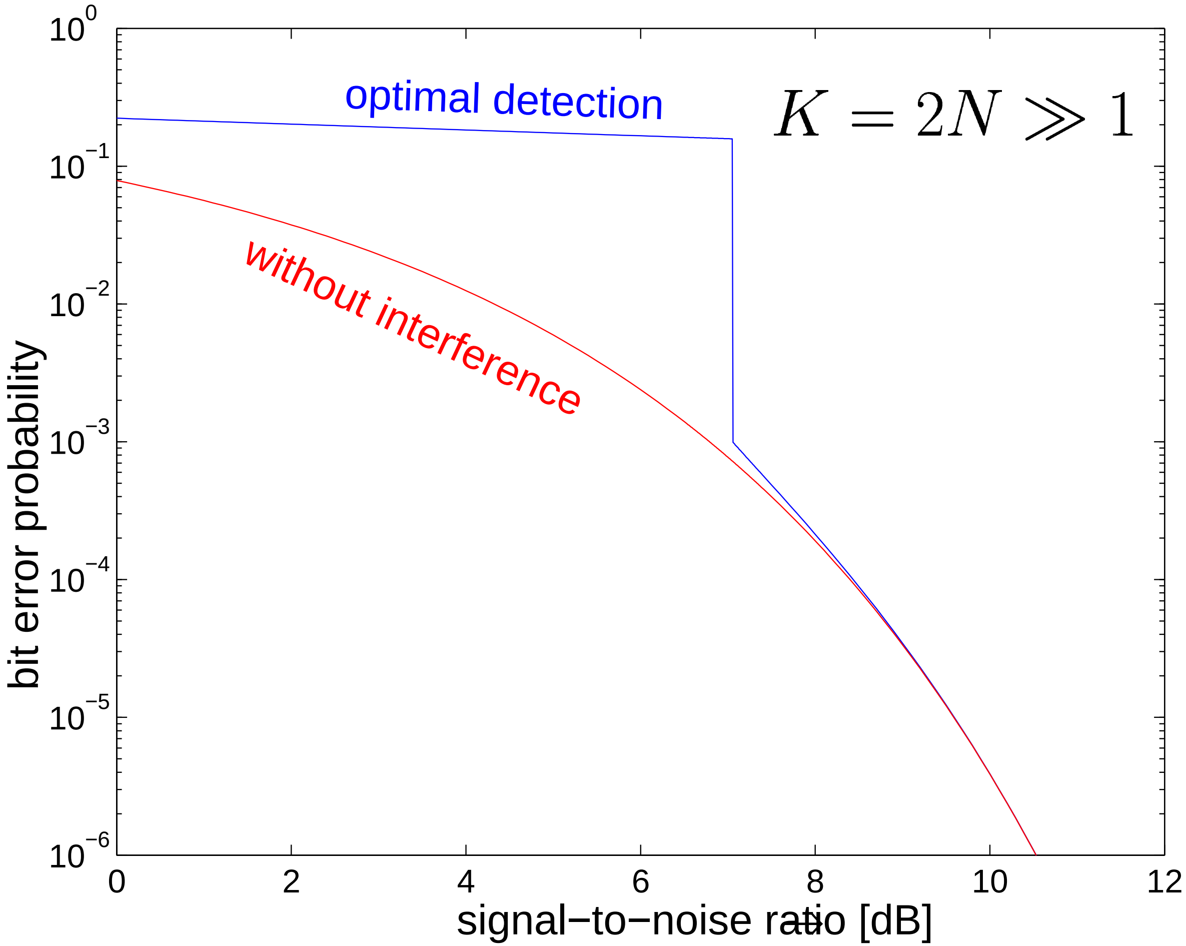
$$\begin{aligned}\hat{\boldsymbol{x}} &= \underset{\boldsymbol{x}\in\{\pm 1\}^K}{\operatorname{argmin}} \ ||\boldsymbol{y} - \boldsymbol{S}\boldsymbol{x}|| \\ &= \underset{\boldsymbol{x}\in\{\pm 1\}^K}{\operatorname{argmin}} \ -\frac{1}{2}\boldsymbol{x}^\dagger \boldsymbol{R}\boldsymbol{x} - \boldsymbol{h}^\dagger \boldsymbol{x} \\ &= \underset{\boldsymbol{x}\in\{\pm 1\}^K}{\operatorname{argmin}} \ -\sum_i \sum_{j<i} r_{ij} x_i x_j - \sum_i h_i x_i\end{aligned}$$

with

$$\begin{aligned}\boldsymbol{R} &= -2\boldsymbol{S}^\dagger \boldsymbol{S} \\ \boldsymbol{h} &= 2\boldsymbol{S}^\dagger \boldsymbol{y}\end{aligned}$$

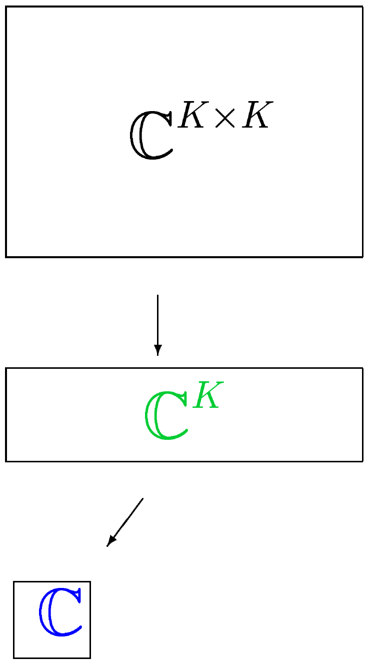
Minimization of the energy function of a spin glass!

A Phase Transition in Random CDMA



[Tanaka 2001]

Large Systems



$\mathcal{O}\left(K^2\right)$ interactions

$\uparrow$

$\mathcal{O}(K)$ microscopic objects

$\downarrow$

$\mathcal{O}(1)$ macroscopic variables

Macroscopic variables are self-averaging.

Boltzmann Distribution

The Thermodynamic Equilibrium maximizes the entropy

$$H(X) = -\sum_i \Pr(x_i) \log \Pr(x_i)$$

for given constant energy

$$E(X) = \sum_i ||x_i|| \Pr(x_i)$$

yielding the *Boltzmann distribution*

$$\Pr(x_i) = \frac{\mathrm{e}^{-\frac{1}{T}||x_i||}}{\sum_i \mathrm{e}^{-\frac{1}{T}||x_i||}}.$$

Free Energy

Since the energy is constant, we can minimize the free energy

$$F(X) \stackrel{\triangle}{=} E(X) - TH(X)$$

instead of maximizing entropy.

With the Boltzmann distribution, the free energy is given by

$$F(X) = -T \log \left[\sum_i \mathrm{e}^{-\frac{1}{T}||x_i||} \right].$$

The free energy is self-averaging.

Average Free Energy

When analyzing a random system, we evaluate the average free energy

Average Free Energy

When analyzing a random system, we evaluate the average free energy

as Shannon analyzed the average performance of all codes.

Free Energy for a Random Parameter

$$F(X|y_j) = \mathop{\mathrm{E}}_{Y} F(X|Y)$$

Free Energy for a Random Parameter

$$
\begin{aligned}
F(X|y_j) &= \underset{Y}{\mathrm{E}}\, F(X|Y) \\
&= -T\, \underset{Y}{\mathrm{E}} \log \left[\sum_i \mathrm{e}^{-\frac{1}{T}||x_i||} \right]
\end{aligned}
$$

The energy function depends on the random parameter y_j.

The expectation of a logarithm is a hard problem.

Replica Continuity

$$\mathop{\mathrm{E}}_{Y} \log(Y) = \lim_{n\to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} Y^n$$

Evalute n^{th} moments for integer n and assume analytic continuity for the limit.

Replica Continuity

$$\operatorname*{E}_{Y} \log(Y) = \lim_{n\to 0} \frac{\partial}{\partial n} \operatorname*{E}_{Y} Y^n$$

Evalute n^{th} moments for integer n and assume analytic continuity for the limit.

More general, we have

$$\operatorname*{E}_{Y} \log\left(\sum_i f(x_i, Y)\right) = \lim_{n\to 0} \frac{\partial}{\partial n} \operatorname*{E}_{Y} \left[\sum_i f(x_i, Y)\right]^n$$

Replica Continuity

$$\mathop{\mathrm{E}}_{Y} \log \left(\sum_i f(x_i, Y) \right) = \lim_{n \to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} \left[\sum_i f(x_i, Y) \right]^n$$

Replica Continuity

$$\mathop{\mathrm{E}}_{Y} \log\left(\sum_i f(x_i, Y)\right) = \lim_{n\to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} \left[\sum_i f(x_i, Y)\right]^n$$

With

$$\left(\textstyle\sum_i x_i\right)^n = \prod_{a=1}^{n} \textstyle\sum_i x_i^{(a)}$$

Replica Continuity

$$\mathop{\mathrm{E}}_{Y} \log \left(\sum_i f(x_i, Y) \right) = \lim_{n \to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} \left[\sum_i f(x_i, Y) \right]^n$$

With

$$\left(\sum_i x_i\right)^n = \prod_{a=1}^{n} \sum_i x_i^{(a)}$$

we finally get

$$\mathop{\mathrm{E}}_{Y} \log \left(\sum_i f(x_i, Y) \right) = \lim_{n \to 0} \frac{\partial}{\partial n} \mathop{\mathrm{E}}_{Y} \prod_{a=1}^{n} \sum_i f(x_i^{(a)}, Y)$$

Replica Symmetry

Throughout the calaculations, we solve integrals of the form

$$I = \frac{1}{K} \log \int \mathrm{e}^{Kf(x_1,x_2)} dx_1 dx_2$$

for $K \to \infty$ by *saddle point integration*.

Replica Symmetry

Throughout the calaculations, we solve integrals of the form

$$I = \frac{1}{K} \log \int \mathrm{e}^{K f(x_1, x_2)} dx_1 dx_2 \to \max_{x_1, x_2} f(x_1, x_2)$$

for $K \to \infty$ by *saddle point integration*.

Replica Symmetry

Throughout the calaculations, we solve integrals of the form

$$I = \frac{1}{K} \log \int \mathrm{e}^{K f(x_1, x_2)} dx_1 dx_2 \to \max_{x_1, x_2} f(x_1, x_2)$$

for $K \to \infty$ by *saddle point integration*.

If the maximization is too tedious, we assume *replica symmetry*:

$$\max_{x_1, x_2} f(x_1, x_2) = \max_{x} f(x, x)$$

Replica symmetry is a strong assumption and not always valid.

The Meaning of the Energy Function

The choice of the energy function is almost arbitrary:

$$||\cdot||:\mathbb{C}^K \mapsto \mathbb{R}^+$$

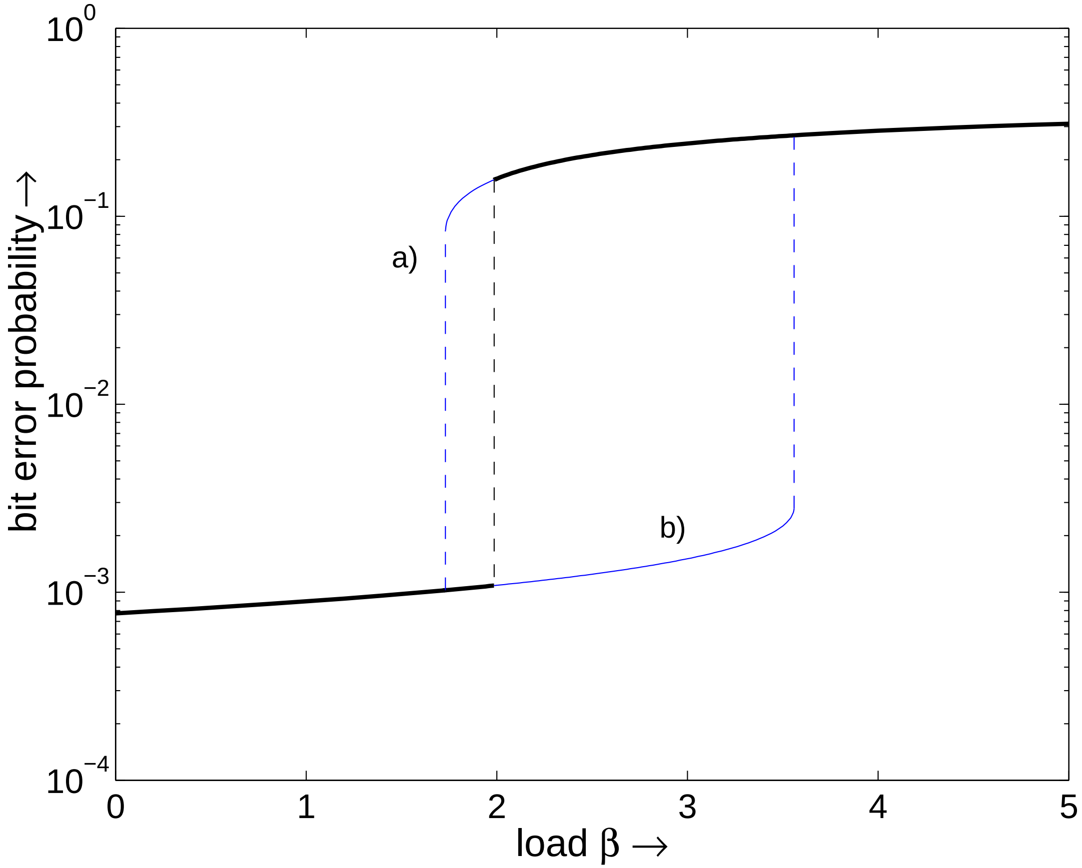
It is conveniently chosen as the metric of the receiver *(which can even be based on wrong or incomplete knowledge of the channel).*

The free choice of the energy function allows to analyze mismatched receivers.

Phase Transitions

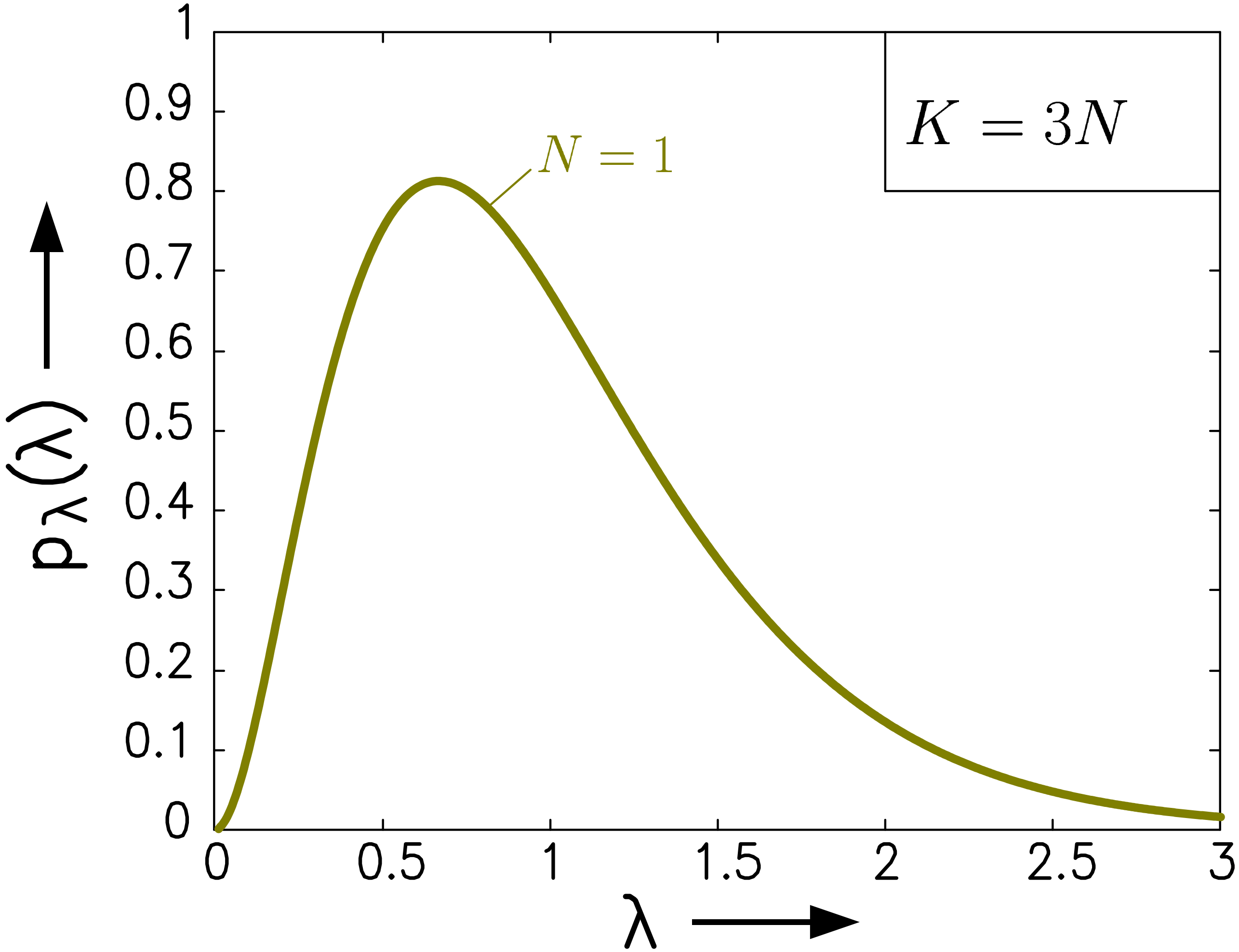
If the final equations allow for multiple solutions, the correct solution is identified by minimizing the free energy.

Phase Transitions and Neural Networks



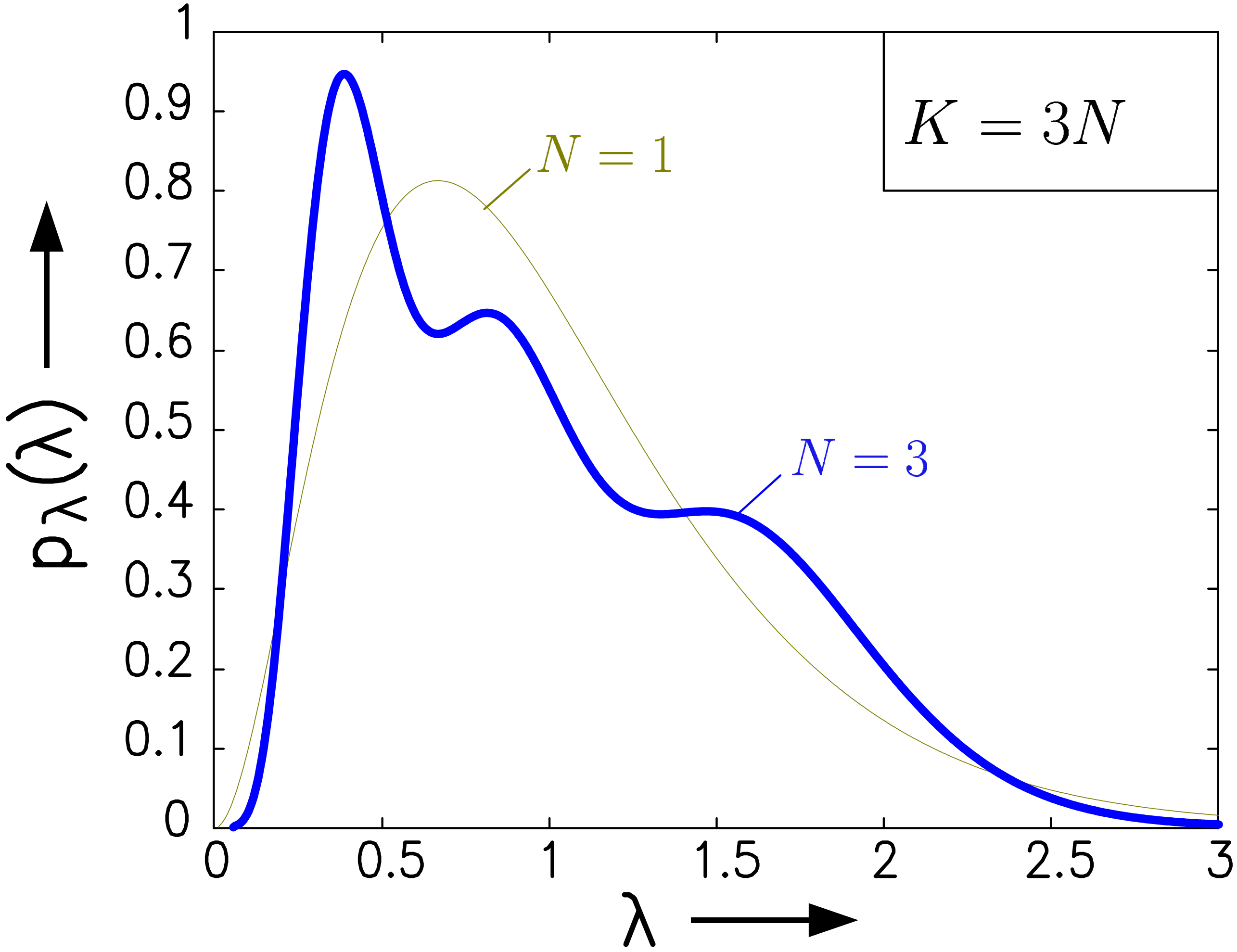
Statistical Physics have developed even more tools to analyze large random systems.

Random Matrix Theory



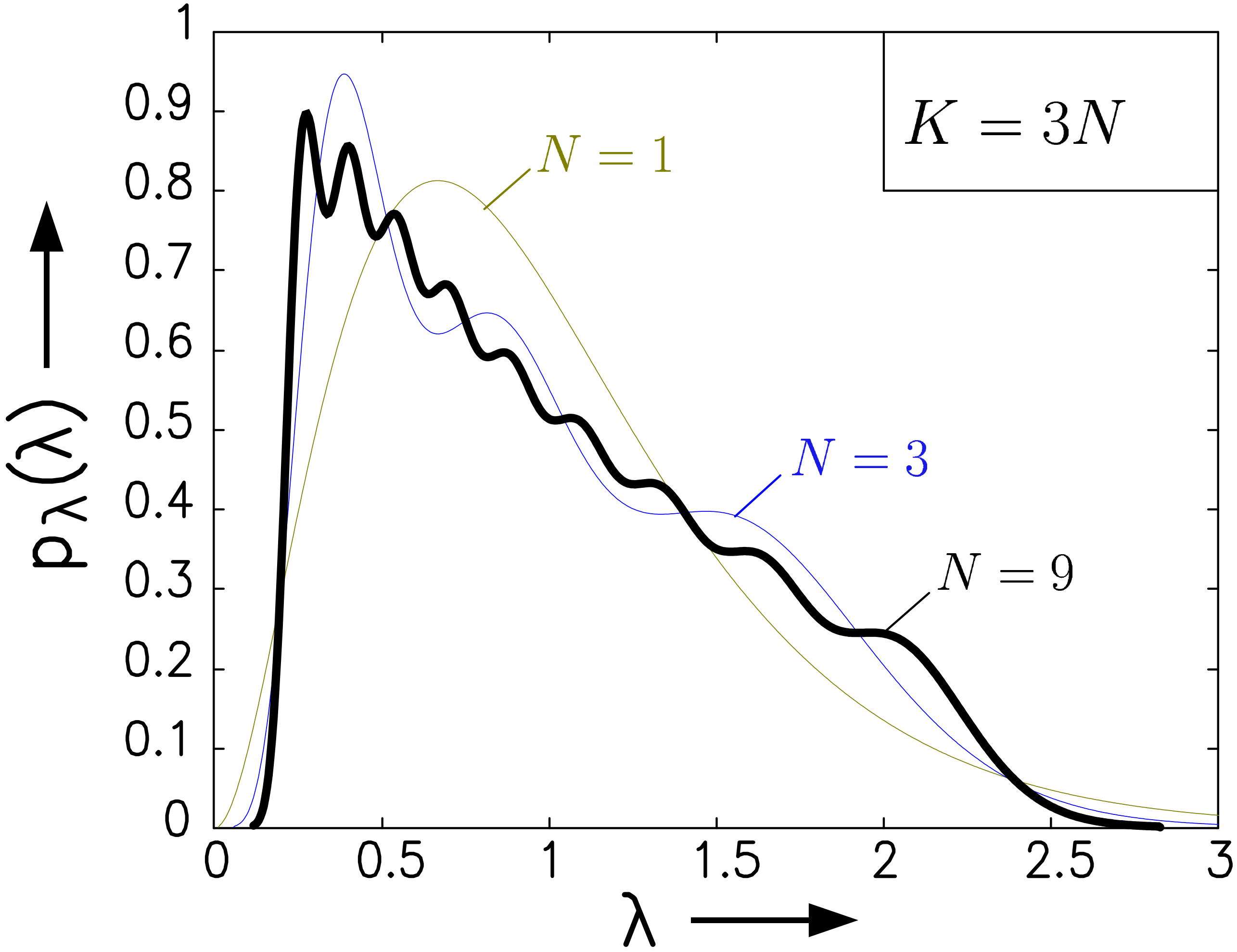
Eigenvalues of $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}$ for i.i.d. entries in the $N \times K$ matrix $\boldsymbol{S}$.

Random Matrix Theory



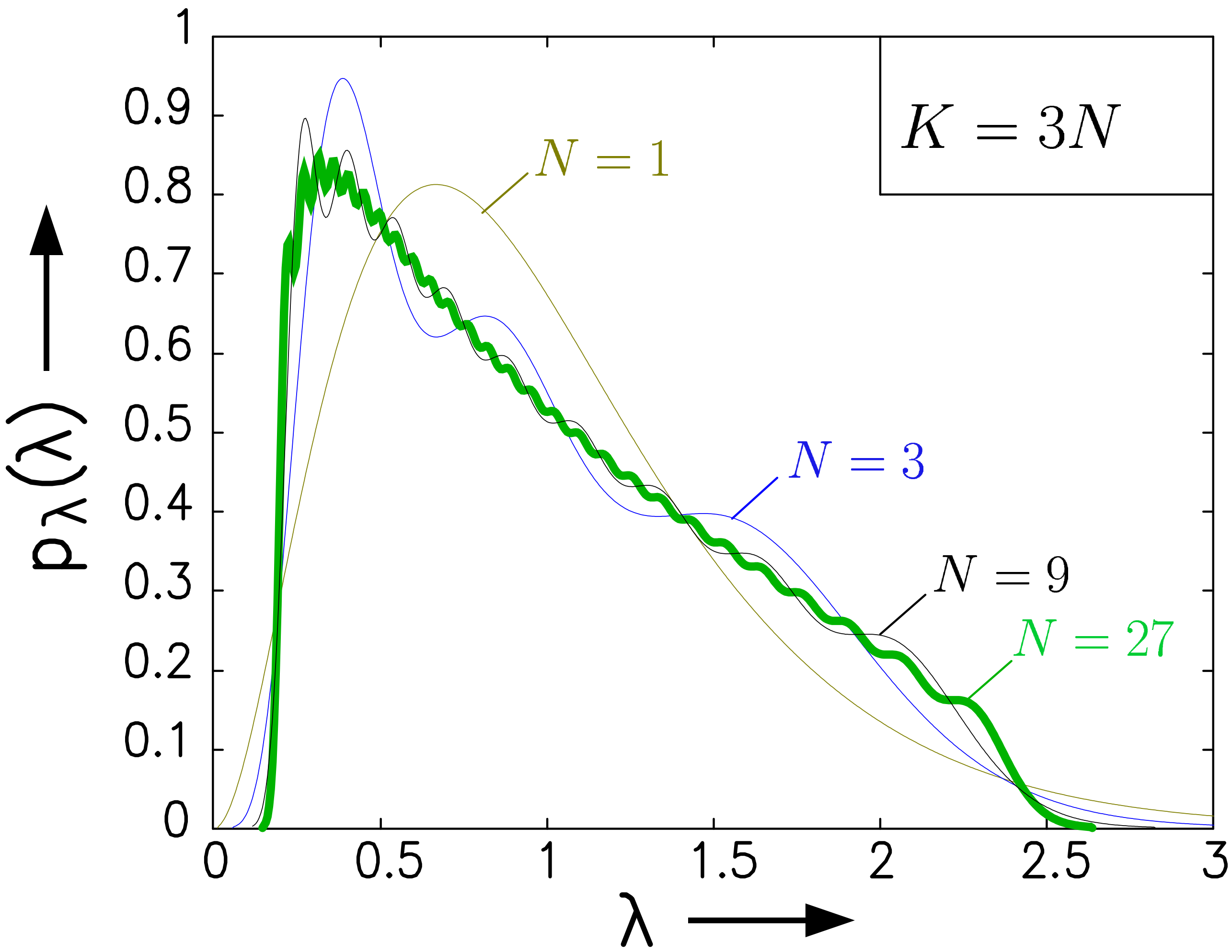
Eigenvalues of $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}$ for i.i.d. entries in the $N \times K$ matrix $\boldsymbol{S}$.

Random Matrix Theory



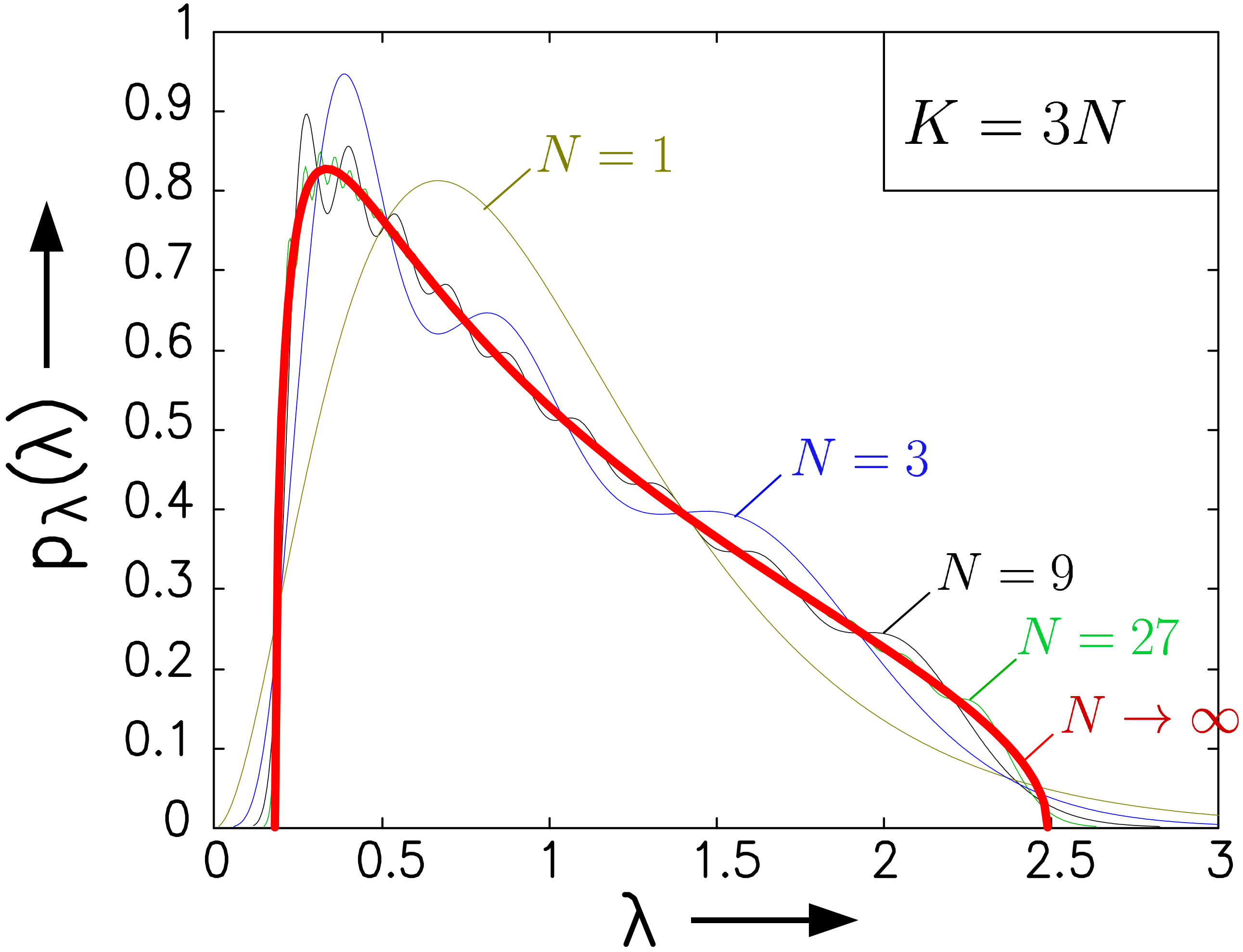
Eigenvalues of $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}$ for i.i.d. entries in the $N \times K$ matrix $\boldsymbol{S}$.

Random Matrix Theory



Eigenvalues of $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}$ for i.i.d. entries in the $N \times K$ matrix $\boldsymbol{S}$.

Random Matrix Theory



Eigenvalues of $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}$ for i.i.d. entries in the $N \times K$ matrix $\boldsymbol{S}$.