Concentration of Covariance Matrices for for Distributions with $2 + \epsilon$ moments

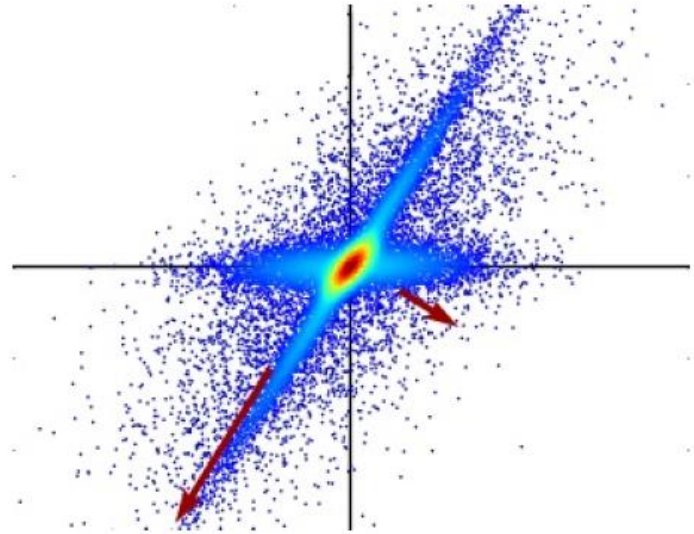
Nikhil Srivastava (Berkeley)
Roman Vershynin (Michigan)

Covariance Matrices

Random Vector $X \in \mathbf{R}^n$

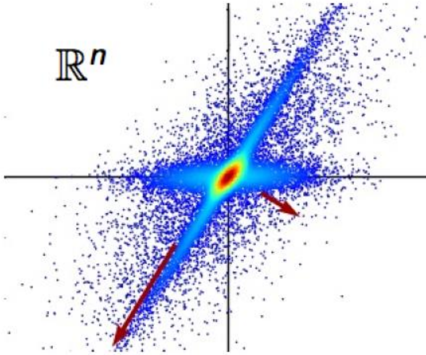
Covariance Matrix

$$\Sigma = \mathbf{E} X X^T$$



Variances $u^T \Sigma u = \mathbf{E}\langle u, X\rangle^2 \quad u \in S^{n-1}$

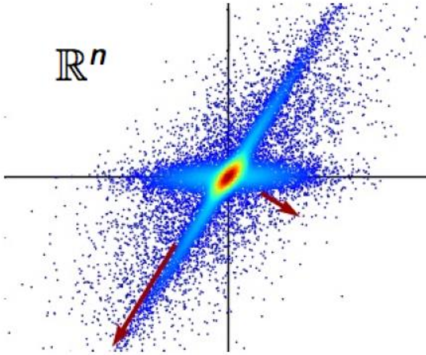
Covariance Estimation



Goal: Estimate $\sum$ given i.i.d. $X_1, \ldots, X_q$

Want: $(1-\epsilon)\Sigma \preceq \frac{1}{q}\sum_{i \leq q} X_i X_i^T \preceq (1+\epsilon)\Sigma$

Covariance Estimation



Goal: Estimate $\sum$ given i.i.d. $X_1, \ldots, X_q$

Want: $(1-\epsilon)\Sigma \preceq \frac{1}{q}\sum_{i\leq q} X_i X_i^T \preceq (1+\epsilon)\Sigma$

Question: How many samples $q = q(n, \epsilon)$ do we need?

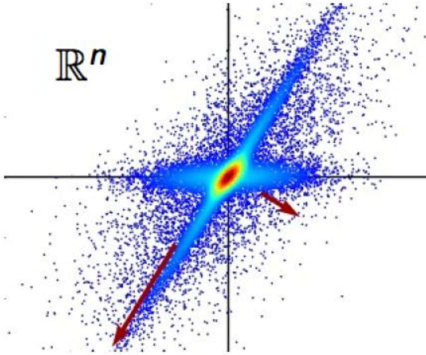
Applications

- Volume Computation [Kannan-Lovasz-Simonovits'95]
- Low Rank Approx [Rudelson-Vershynin'07]
- Graph Sparsification [Spielman-S '08]
- Sparse Approximation/Compressed Sensing
- Matrix Completion [Candes-Recht '09]

…

Like nonasymptotic Bai-Yin for matrices with independent rows.

Covariance Estimation



Goal: Estimate $\sum$ given i.i.d. $X_1, \ldots, X_q$

Want: $(1-\epsilon)\Sigma \preceq \frac{1}{q}\sum_{i\leq q} X_i X_i^T \preceq (1+\epsilon)\Sigma$

Question: How many samples $q = q(n, \epsilon)$ do we need?

Isotropic Position

Want: $(1-\epsilon)\Sigma \preceq \frac{1}{q}\sum_{i\leq q} X_i X_i^T \preceq (1+\epsilon)\Sigma$

Sufficient to handle $\Sigma = I$ **isotropic position.**

$$\mathbf{E}\langle u, X\rangle^2 = 1 \quad \forall u \qquad \mathbf{E}\|X\|^2 = n$$

Isotropic Position

Want: $(1-\epsilon)\Sigma \preceq \frac{1}{q}\sum_{i\leq q} X_i X_i^T \preceq (1+\epsilon)\Sigma$

Sufficient to handle $\Sigma = I$ **isotropic position.**

$$\mathbf{E}\langle u, X\rangle^2 = 1 \quad \forall u \qquad\qquad \mathbf{E}\|X\|^2 = n$$

Reduction: $X_i' = \Sigma^{-1/2} X_i$ **and**

$$(1-\epsilon)I \preceq \frac{1}{q}\sum_{i\leq q} X_i' X_i'^T \preceq (1+\epsilon)I$$

Isotropic Position

Want: $$\left\| \frac{1}{q} \sum_{i \le q} X_i X_i^T - I \right\|_2 \le \epsilon$$

Sufficient to handle $\Sigma = I$ **isotropic position.**

$$\mathbf{E}\langle u, X\rangle^2 = 1 \quad \forall u \qquad \mathbf{E}\|X\|^2 = n$$

Given $\mathbf{E} X X^T = I$ how large is **q**?

[KLS,B,...Rudelson'99]

Isotropic random $X \in \mathbf{R}^n$ with $\|X\|_2 \le O(\sqrt{n})$

If $q = \Omega(n \log n/\epsilon^2)$

Then $\left\| \frac{1}{q} \sum_{i \le q} X_i X_i^T - I \right\|_2 \le \epsilon$ whp.

[Rudelson'99]

Isotropic random $X \in \mathbf{R}^n$ with $\|X\|_2 \leq O(\sqrt{n})$

If $q = \Omega(n \log n / \epsilon^2)$

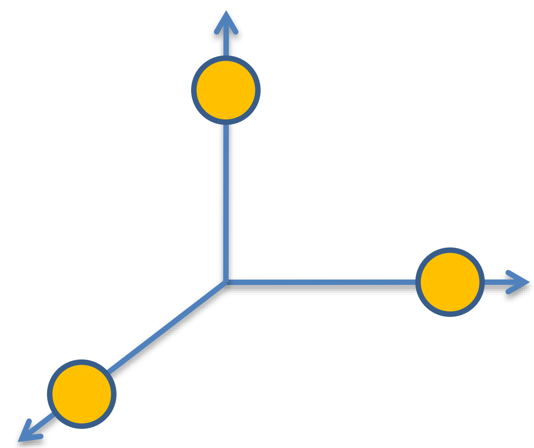
Then $\left\| \frac{1}{q} \sum_{i \leq q} X_i X_i^T - I \right\|_2 \leq \epsilon$ whp.

Tight example:

$$X = \sqrt{n} e_i \qquad \text{w. prob. } 1/n$$

$$\mathbf{E} X X^T = (1/n) \sum_{i \leq n} n e_i e_i^T = I$$

$$\Sigma_q(i,i) = \text{num. of balls in bin } i$$



[Rudelson'99]

Isotropic random $X \in \mathbf{R}^n$ with $\|X\|_2 \leq O(\sqrt{n})$

If $q = \Omega(n \log n / \epsilon^2)$

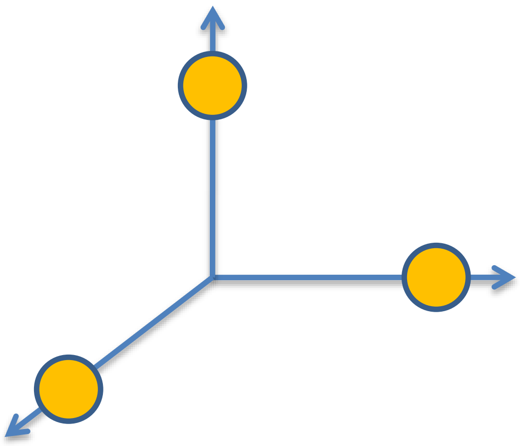
Then $\left\| \frac{1}{q} \sum_{i \leq q} X_i X_i^T - I \right\|_2 \leq \epsilon$ whp.

Tight example:

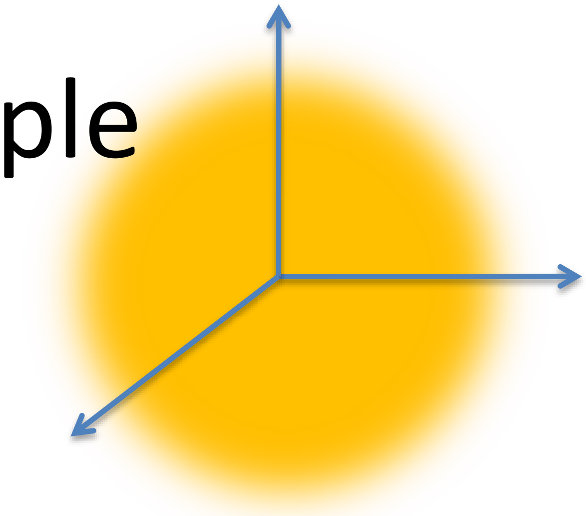
$X = \sqrt{n} e_i$ w. prob. $1/n$

$\mathbf{E} X X^T = (1/n) \sum_{i \leq n} n e_i e_i^T = I$

$\Sigma_q(i,i) = \text{num. of balls in bin } i$



A Good Example



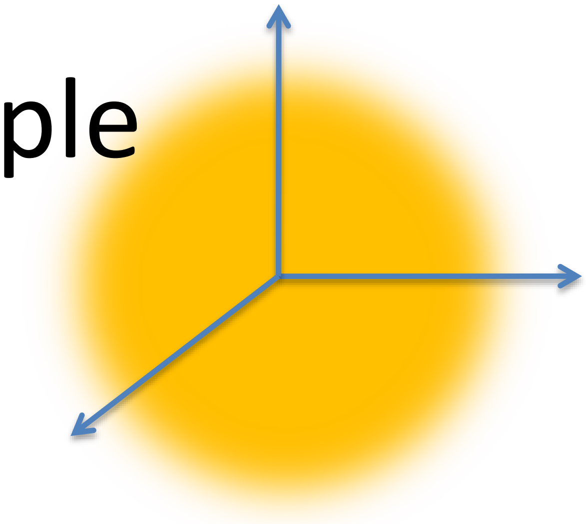
Standard Gaussian vector:

$$X \sim \mathcal{N}(0, I)$$

For any fixed direction

$$u \in S^{n-1} \qquad \langle u, X \rangle \sim \mathcal{N}(0, 1)$$

A Good Example



Standard Gaussian vector:

$$X \sim \mathcal{N}(0, I)$$

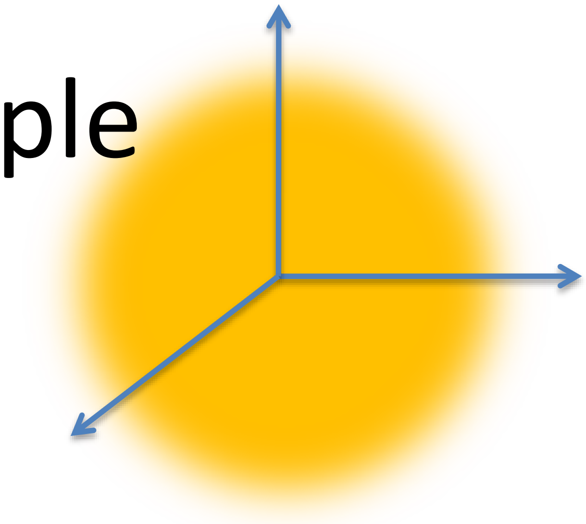
For any fixed direction

$$u \in S^{n-1} \qquad \langle u, X \rangle \sim \mathcal{N}(0, 1)$$

So for independent $\mathbf{X_1}$, … $\mathbf{X_q}$

$$u^T \Sigma_q u = \frac{1}{q} \sum_{i \leq q} \langle u, X_i \rangle^2 \sim \chi^2(q)$$

A Good Example



Standard Gaussian vector:

$$X \sim \mathcal{N}(0, I)$$

For any fixed direction

$$u \in S^{n-1} \qquad \langle u, X \rangle \sim \mathcal{N}(0, 1)$$

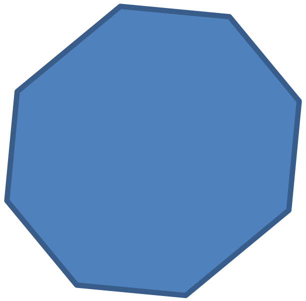
So for independent **X_1, … X_q**

$$u^T \Sigma_q u = \frac{1}{q} \sum_{i \leq q} \langle u, X_i \rangle^2 \sim \chi^2(q)$$

$$\mathbf{P}(|\frac{1}{q} \sum_{i \leq q} \langle u, X_i \rangle^2 - 1| > \epsilon) \leq \exp(-q\epsilon^2)$$

$$\text{Take } q \gg n/\epsilon^2, \text{ union bound.}$$

Convex Bodies [KLS'95]

$$X \sim$$ 

More generally **sub-exponential X**:

$$\forall u \in S^{n-1} : \quad \mathbf{P}(|\langle u, X\rangle| > t) \leq C\exp(-t)$$

[ALPT'09/ALLPT'11] $q = O(n/\epsilon^2)$ whp,
provided $\|X\|_2 \leq O(\sqrt{n})$

Heavier Tails

$$\forall u \in S^{n-1}: \quad \mathbf{P}(|\langle u, X\rangle| > t) \leq Ct^{-p})$$

[Vershynin'11] $q = O(n(\log\log n)^2)$

for **p>4** and $\|X\|_2 \leq O(\sqrt{n})$

[Mendelson-Paouris'12] $q = O(n/\epsilon^2)$

for **p>8** and something like $\|X\|_2 \leq O(\sqrt{n})$

Heavier Tails

[S-Vershynin'12] Suppose isotropic **X** satisfies:

1D $\forall u \in S^{n-1} : \quad \mathbf{P}(|\langle u, X\rangle| > t) \leq C/t^{2+\eta}$

Heavier Tails

[S-Vershynin'12] Suppose isotropic **X** satisfies:

1D $\forall u \in S^{n-1}: \quad \mathbf{P}(|\langle u, X\rangle| > t) \leq C/t^{2+\eta}$

cf. $\mathbf{E}\langle u, X\rangle^2 = 1$

Heavier Tails

[S-Vershynin'12] Suppose isotropic **X** satisfies:

1D $\forall u \in S^{n-1} : \quad \mathbf{P}(|\langle u, X\rangle| > t) \le C/t^{2+\eta}$

kD $\forall \Pi \quad \mathbf{P}(\|\Pi X\|_2 > t) \le C/t^{2+\eta}, \quad t > C\sqrt{\mathrm{rank}(\Pi)}$

cf. $\mathbf{E}\|\Pi X\|^2 = \mathrm{rank}(\Pi)$

Heavier Tails

[S-Vershynin'12] Suppose isotropic **X** satisfies:

1D $\forall u \in S^{n-1}: \quad \mathbf{P}(|\langle u, X\rangle| > t) \leq C/t^{2+\eta}$

kD $\forall \Pi \quad \mathbf{P}(\|\Pi X\|_2 > t) \leq C/t^{2+\eta}, \quad t > C\sqrt{\mathrm{rank}(\Pi)}$

Then $\mathbf{E}\left\|\frac{1}{q}\sum_{i\leq q} X_i X_i^T - I\right\|_2 \leq \epsilon$ for $q = O(n/\epsilon^{2+\frac{2}{\eta}})$

Heavier Tails

[S-Vershynin'12] Suppose isotropic **X** satisfies:

1D $\forall u \in S^{n-1}: \quad \mathbf{P}(|\langle u, X\rangle| > t) \le C/t^{2+\eta}$

kD $\forall \Pi \quad \mathbf{P}(\|\Pi X\|_2 > t) \le C/t^{2+\eta}, \quad t > C\sqrt{\mathrm{rank}(\Pi)}$

Then $\mathbf{E}\left\|\frac{1}{q}\sum_{i\le q} X_i X_i^T - I\right\|_2 \le \epsilon$ for $q = O(n/\epsilon^{2+\frac{2}{\eta}})$

Includes: log-concave **X** by [Paouris '07]

$$\mathbf{P}(\|X\| > t) \le \exp(-t), \quad t > C\sqrt{n}.$$

Heavier Tails

[S-Vershynin'12] Suppose isotropic **X** satisfies:

1D $\forall u \in S^{n-1} : \quad \mathbf{P}(|\langle u, X\rangle| > t) \leq C/t^{2+\eta}$

kD $\forall \Pi \quad \mathbf{P}(\|\Pi X\|_2 > t) \leq C/t^{2+\eta}, \quad t > C\sqrt{\text{rank}(\Pi)}$

Then $\mathbf{E}\left\| \frac{1}{q}\sum_{i\leq q} X_i X_i^T - I \right\|_2 \leq \epsilon$ for $q = O(n/\epsilon^{2+\frac{2}{\eta}})$

Includes: log-concave **X** by [Paouris '07]

product **X** with bdd $4+\eta$ moments

cf. [Latala'05]

Heavier Tails

[S-Vershynin'12] Suppose isotropic **X** satisfies:

1D $\forall u \in S^{n-1}: \quad \mathbf{P}(|\langle u, X\rangle| > t) \leq C/t^{2+\eta}$

kD $\forall \Pi \quad \mathbf{P}(\|\Pi X\|_2 > t) \leq C/t^{2+\eta}, \quad t > C\sqrt{\mathrm{rank}(\Pi)}$

Then $\mathbf{E}\left\|\frac{1}{q}\sum_{i\leq q} X_i X_i^T - I\right\|_2 \leq \epsilon$ for $q = O(n/\epsilon^{2+\frac{2}{\eta}})$

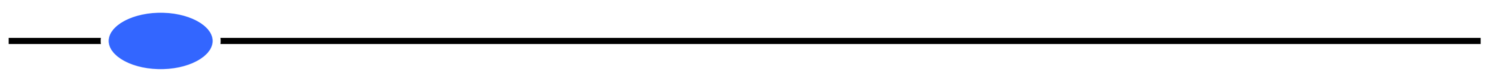
Lower edge is easier: Only require 1D for

$$\mathbf{E}\lambda_{min}(\Sigma_q) \geq 1 - \epsilon.$$

Sketch of the proof

Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$



$$A_0 = 0$$

Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$

$$A_1 = X_1 X_1^T$$

Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$

$$A_2 = X_1 X_1^T + X_2 X_2^T$$

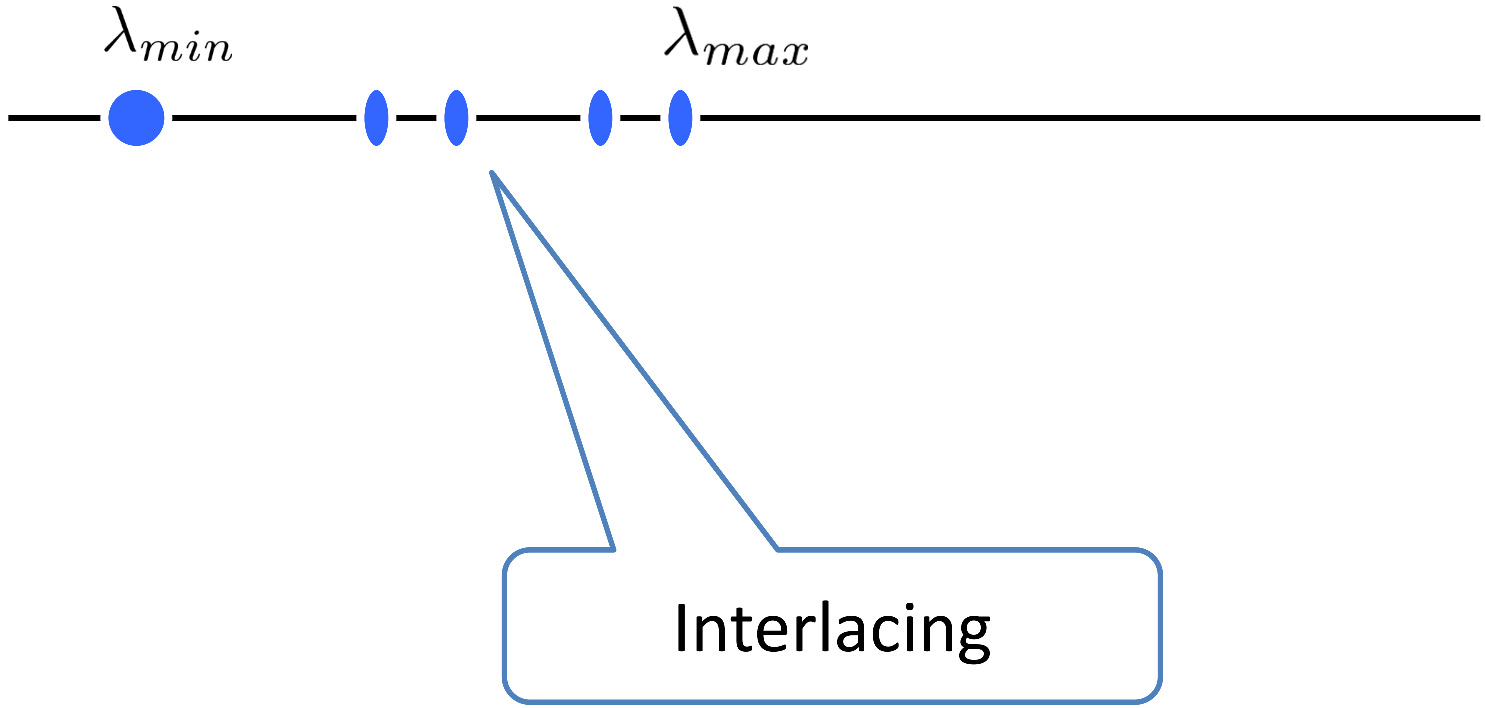
Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$

$$A_3 = X_1 X_1^T + X_2 X_2^T + X_3 X_3^T$$

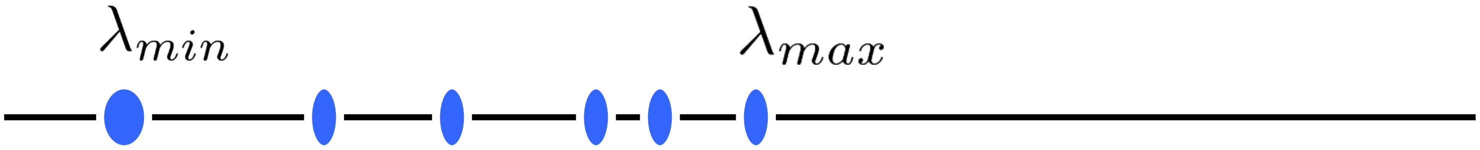
Basic Picture

$$A_k = \sum_{i \le k} X_i X_i^T = A_{k-1} + X_k X_k^T$$



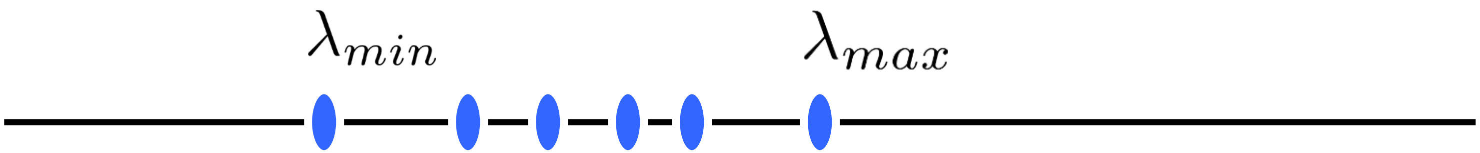
Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$



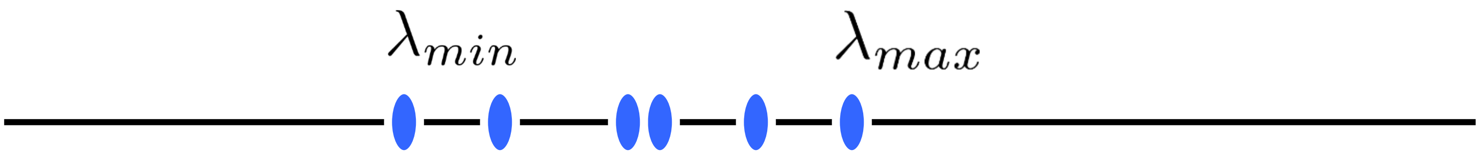
Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$



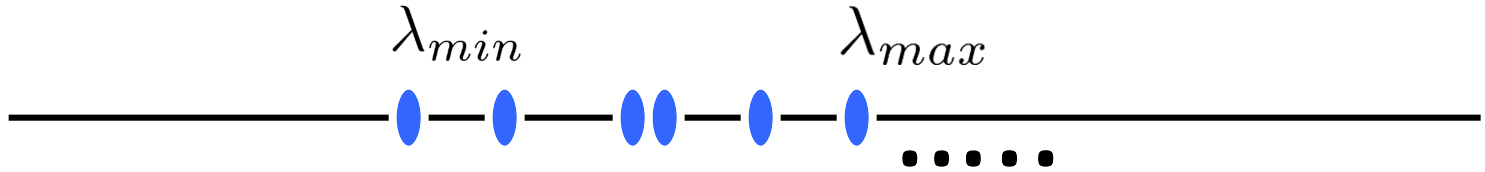
Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$



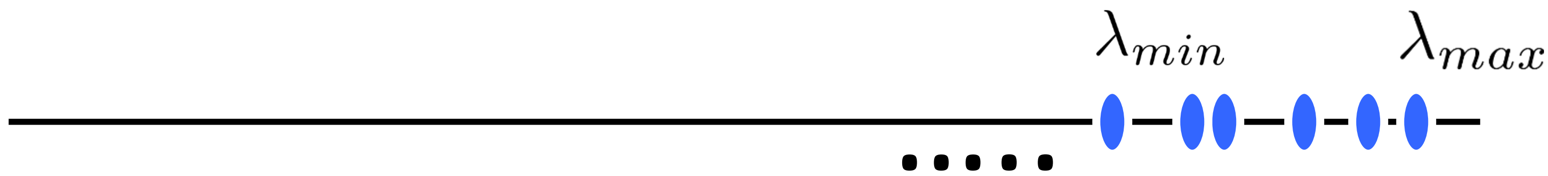
Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$



Basic Picture

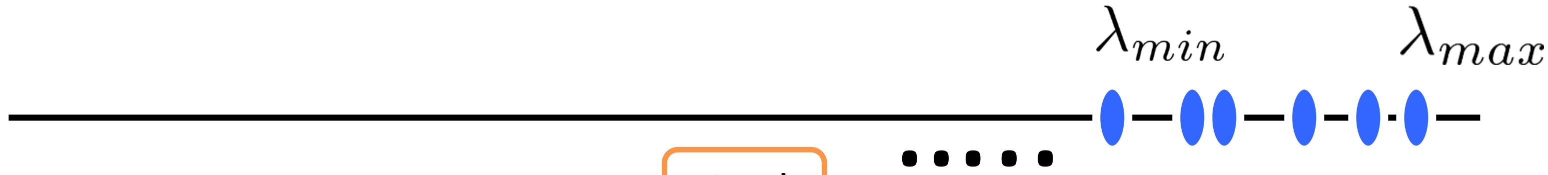
$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$



$$A_q = X_1 X_1^T + X_2 X_2^T \ldots X_q X_q^T$$

Basic Picture

$$A_k = \sum_{i \le k} X_i X_i^T = A_{k-1} + X_k X_k^T$$



Goal

$$(1-\epsilon)q \le \mathbf{E}\lambda \le (1+\epsilon)q$$

$$A_q = X_1 X_1^T + X_2 X_2^T \ldots X_q X_q^T$$

Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$

Goal

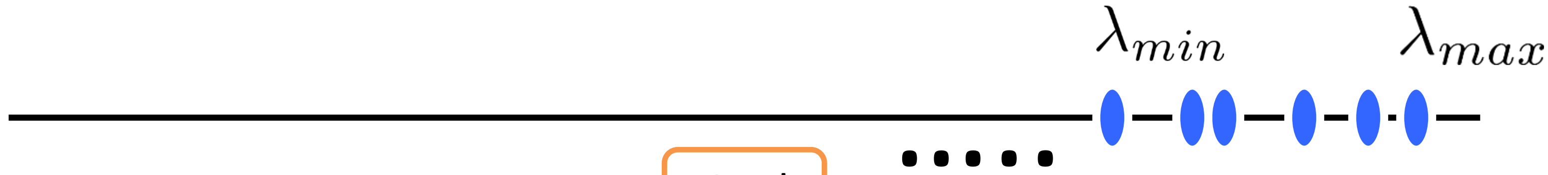
$$(1-\epsilon)q \leq \mathbf{E}\lambda \leq (1+\epsilon)q$$

$$A_q = X_1 X_1^T + X_2 X_2^T \ldots X_q X_q^T$$

$$\begin{aligned} \mathbf{E}\mathrm{Tr}(A_k) &= \mathrm{Tr}(A_{k-1}) + \mathrm{Tr}(\mathbf{E} X_k X_k^T) \\ &= \mathrm{Tr}(A_{k-1}) + n \end{aligned}$$

Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$



Goal

$$(1-\epsilon)q \leq \mathbf{E}\lambda \leq (1+\epsilon)q$$

$$A_q = X_1 X_1^T + X_2 X_2^T \ldots X_q X_q^T$$

$$\mathbf{E}\lambda_{avg}(A_k) = \lambda_{avg}(A_{k-1}) + 1$$

Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$

Goal

$$(1-\epsilon)q \leq \mathbf{E}\lambda \leq (1+\epsilon)q$$

$$A_q = X_1 X_1^T + X_2 X_2^T \ldots X_q X_q^T$$

$$\mathbf{E}\lambda_{max}(A_k) \leq \lambda_{max}(A_{k-1}) + O(1)?$$

$$\mathbf{E}\lambda_{avg}(A_k) = \lambda_{avg}(A_{k-1}) + 1$$

Basic Picture

$$A_k = \sum_{i \leq k} X_i X_i^T = A_{k-1} + X_k X_k^T$$

Goal

$$(1-\epsilon)q \leq \mathbf{E}\lambda \leq (1+\epsilon)q$$

$$A_q = X_1 X_1^T + X_2 X_2^T \ldots X_q X_q^T$$

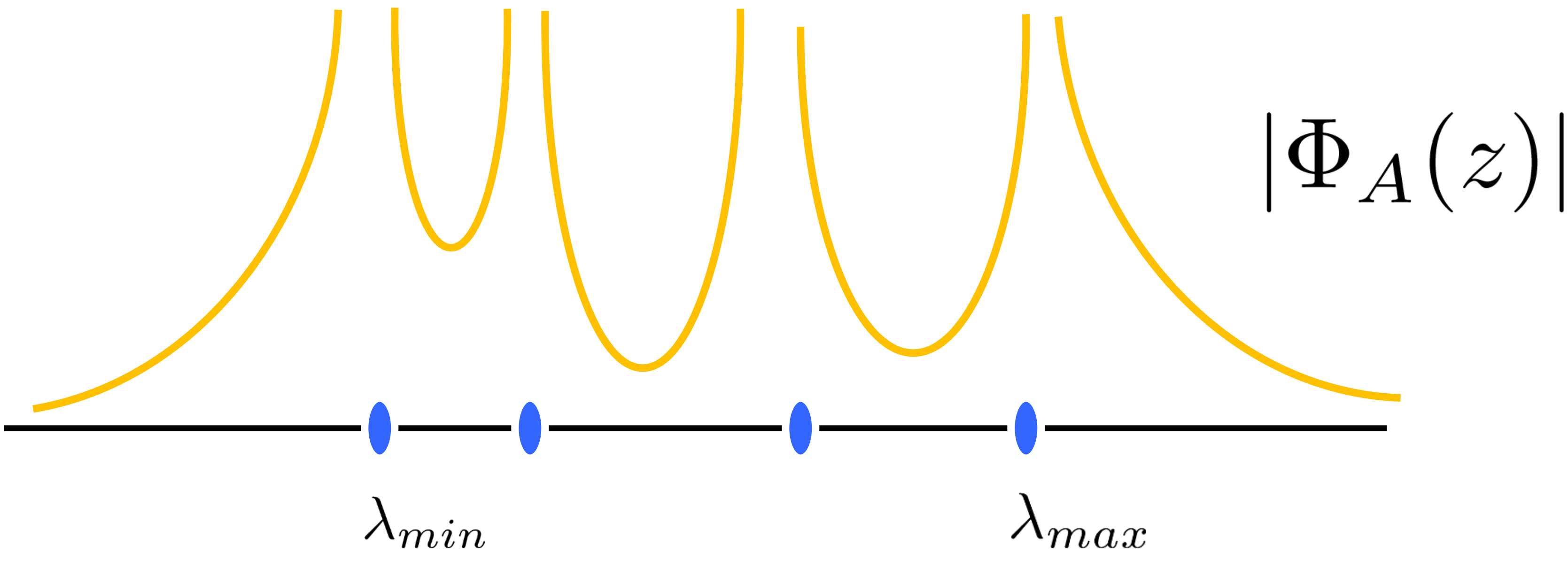
$$\mathbf{E}\lambda_{max}(A_k) \leq \lambda_{max}(A_{k-1}) + O(1)?$$

$$\mathbf{E}\lambda_{avg}(A_k) = \lambda_{avg}(A_{k-1}) + 1$$

False

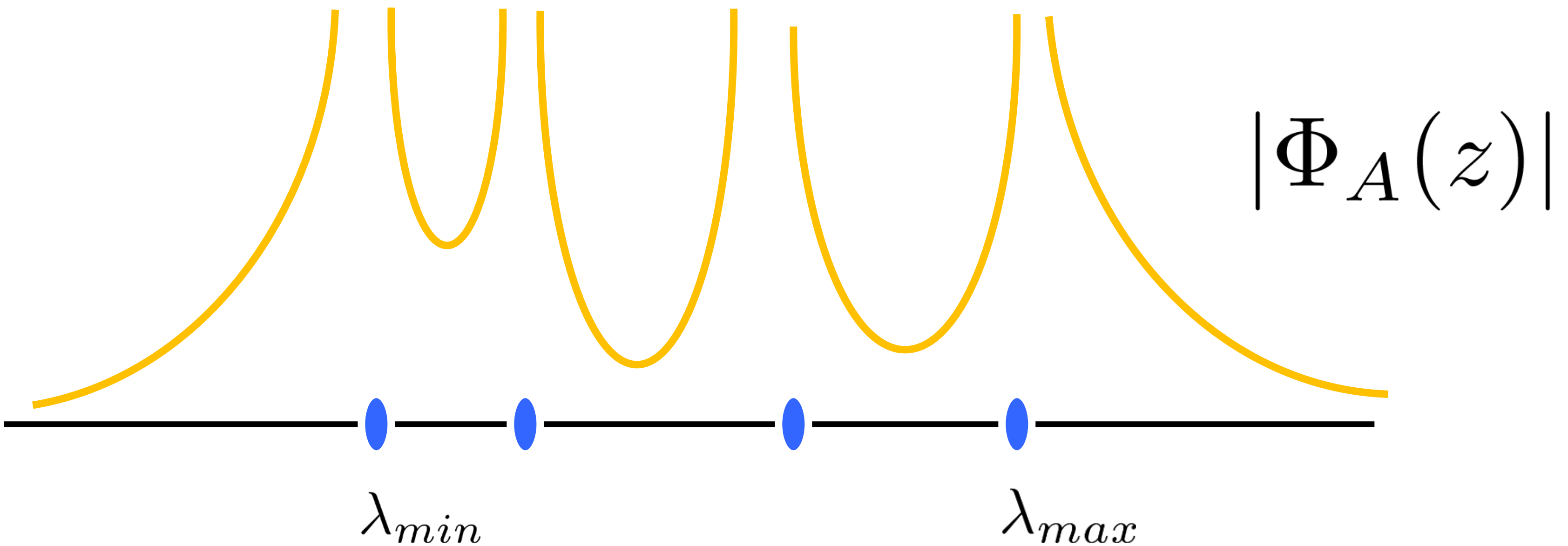
Softening the Edges

$$\Phi_A(z) = \mathrm{Tr}(zI - A)^{-1} = \sum_i \frac{1}{z-\lambda_i}$$



Softening the Edges

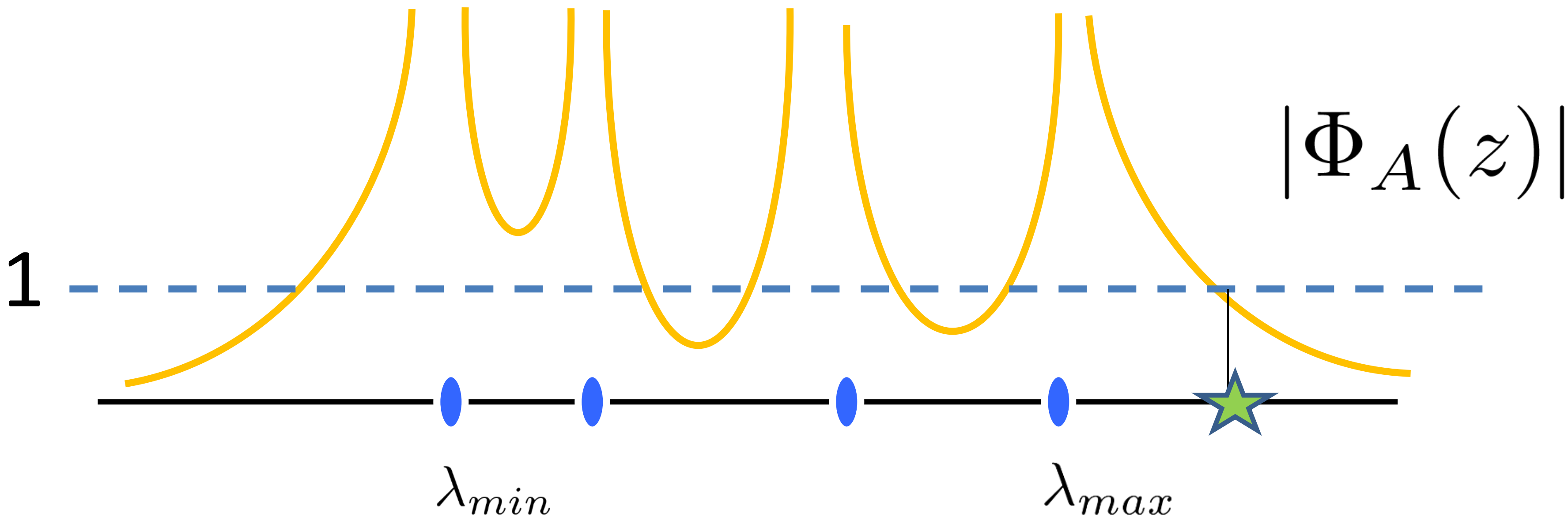
$$\Phi_A(z) = \mathrm{Tr}(zI - A)^{-1} = \sum_i \frac{1}{z - \lambda_i}$$



$$\lambda_{max} = \max\{z : \Phi_A(z) = \infty\}$$

Softening the Edges

$$\Phi_A(z) = \mathrm{Tr}(zI - A)^{-1} = \sum_i \frac{1}{z-\lambda_i}$$

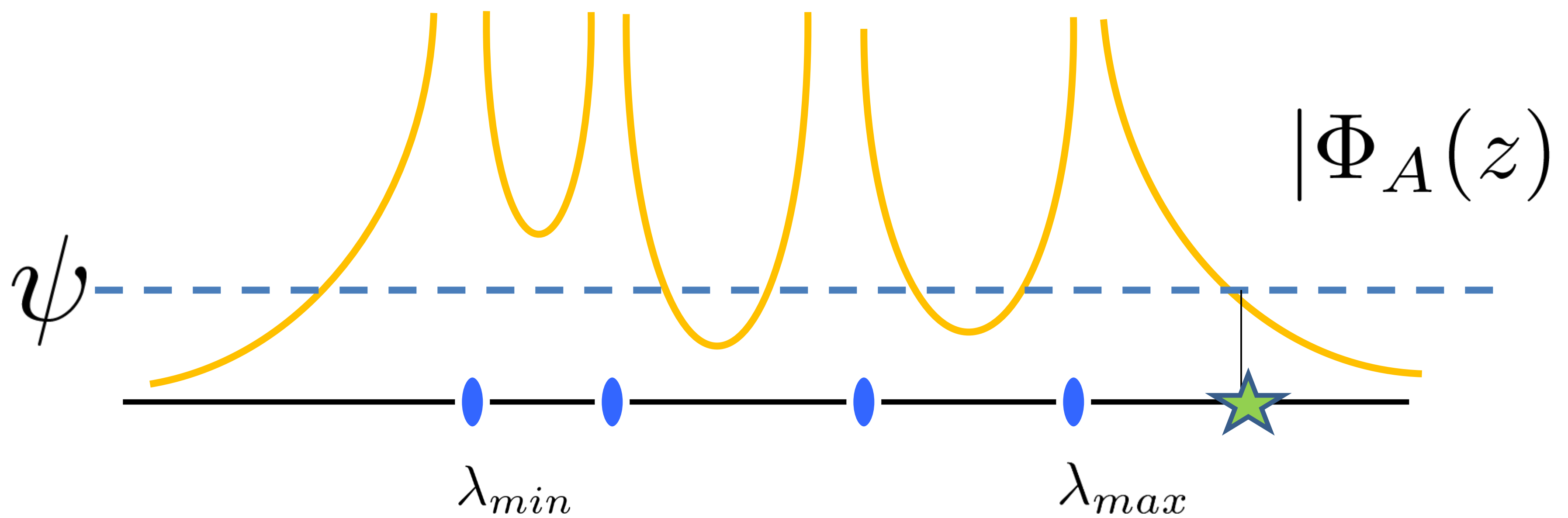


$$\lambda_{max} = \max\{z : \Phi_A(z) = \infty\}$$

$$\boxed{s_{max} := \max\{z : \Phi_A(z) = 1\}}$$

Softening the Edges

$$\Phi_A(z) = \mathrm{Tr}(zI - A)^{-1} = \sum_i \frac{1}{z-\lambda_i}$$

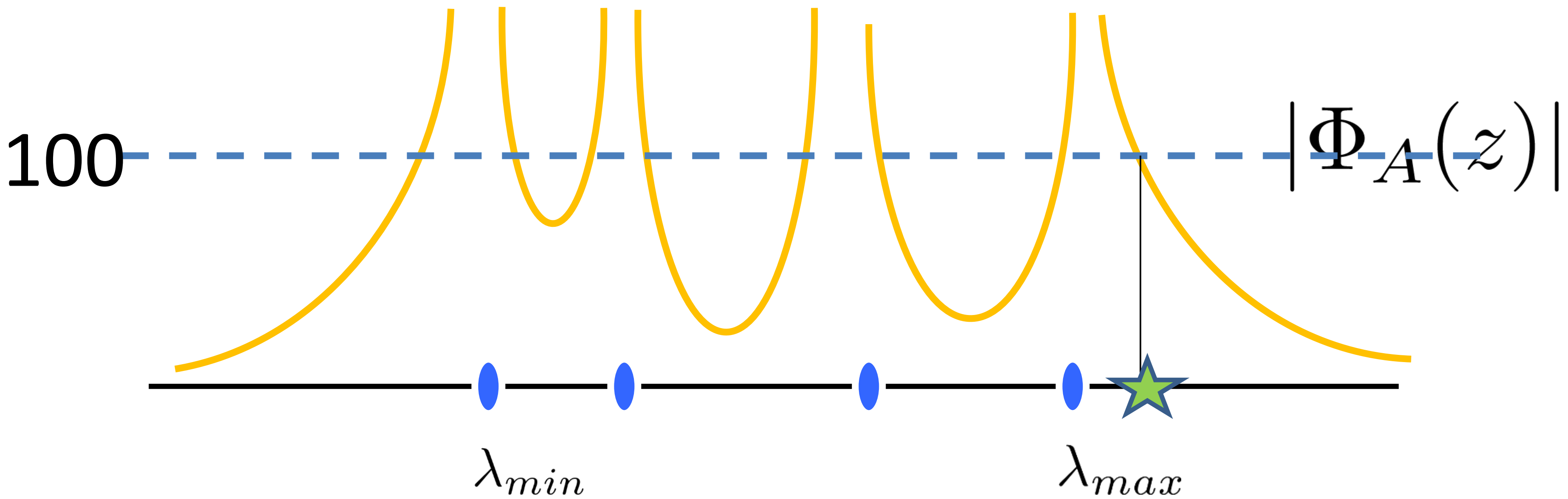


$$\lambda_{max} = \max\{z : \Phi_A(z) = \infty\}$$

$$\boxed{s_{max} := \max\{z : \Phi_A(z) = \psi\}}$$

Softening the Edges

$$\Phi_A(z) = \mathrm{Tr}(zI - A)^{-1} = \sum_i \frac{1}{z-\lambda_i}$$

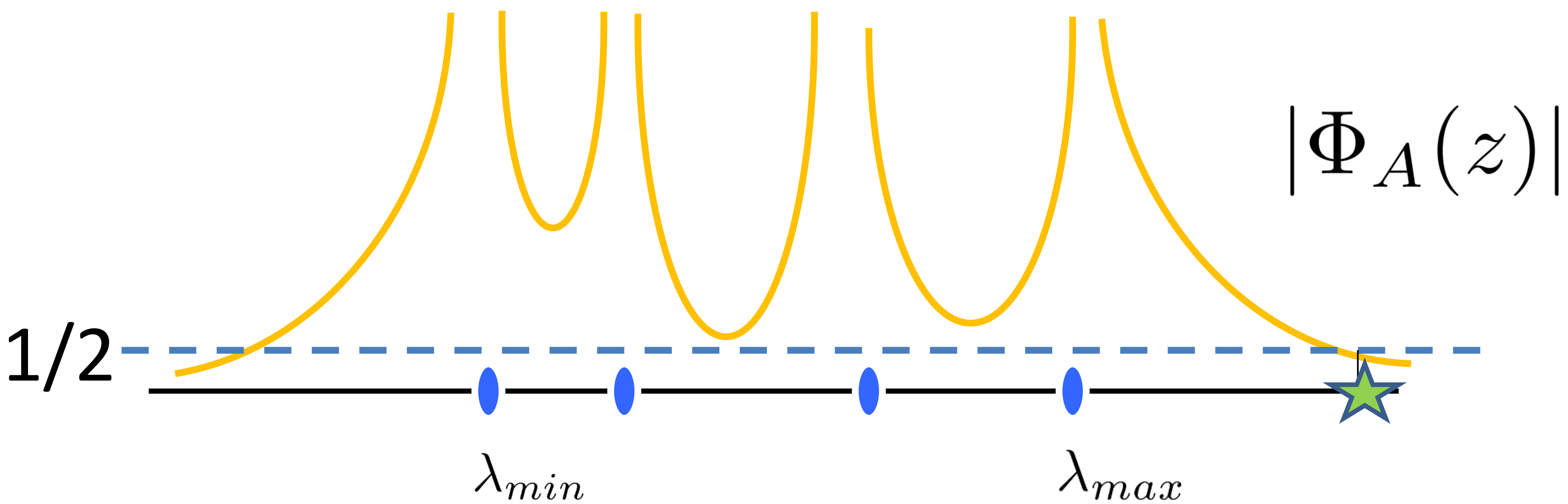


$$\lambda_{max} = \max\{z : \Phi_A(z) = \infty\}$$

$$\boxed{s_{max} := \max\{z : \Phi_A(z) = \psi\}}$$

Softening the Edges

$$\Phi_A(z) = \mathrm{Tr}(zI - A)^{-1} = \sum_i \frac{1}{z - \lambda_i}$$

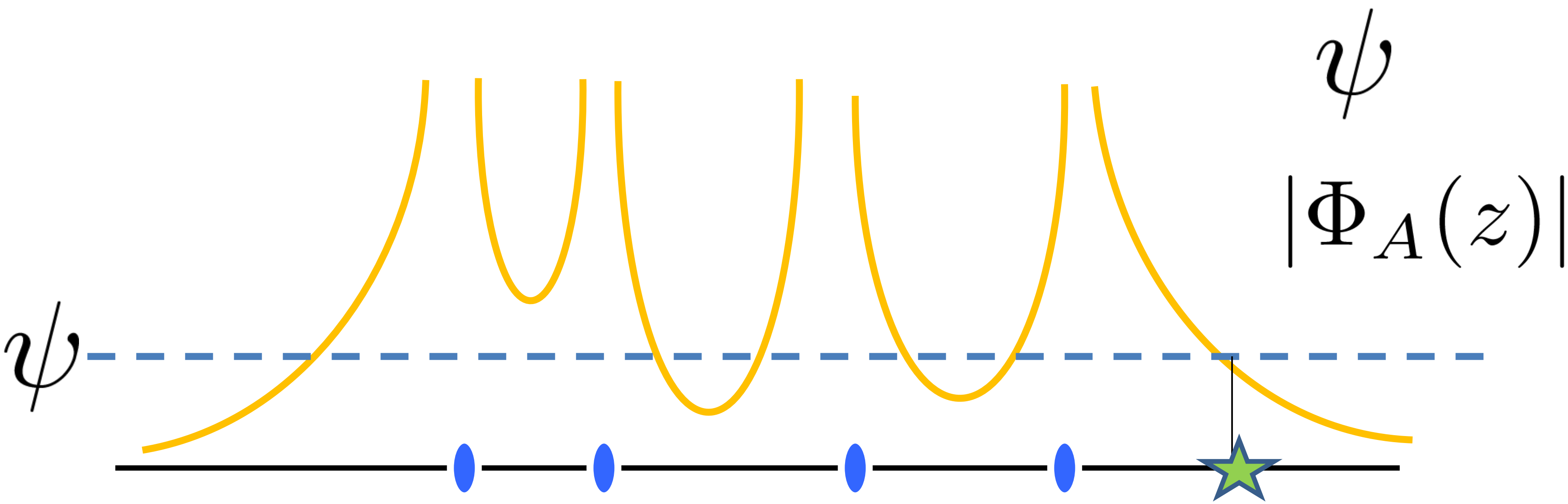


$$\lambda_{max} = \max\{z : \Phi_A(z) = \infty\}$$

$$\boxed{s_{max} := \max\{z : \Phi_A(z) = \psi\}}$$

Softening the Edges

$$\Phi_A(z) = \mathrm{Tr}(zI - A)^{-1}$$

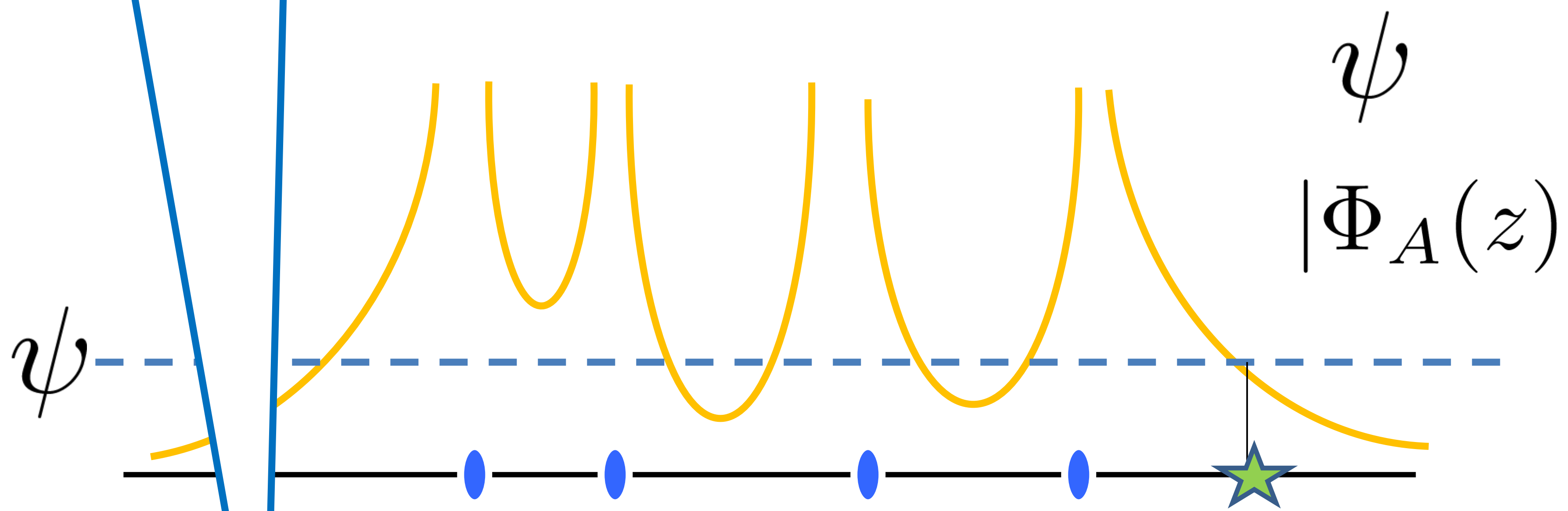


$$s_{max} := \max\{z : \Phi_A(z) = \psi\}$$

ng the Edges

Well-behaved + easy to manipulate

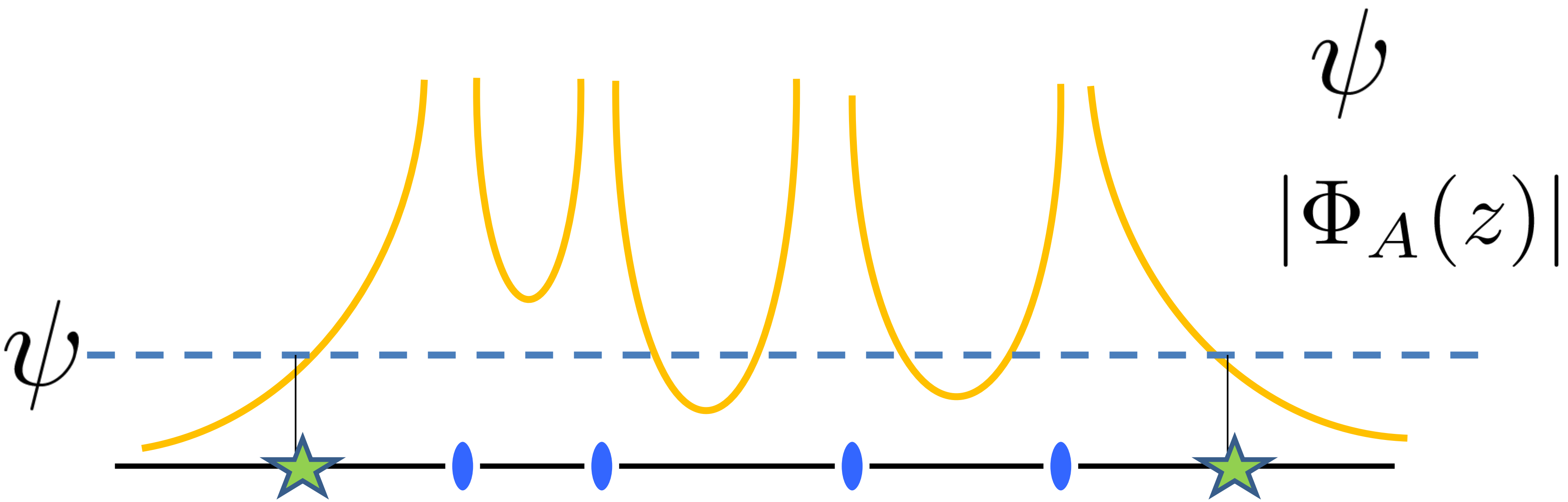
$$\Phi_A(z) = \mathrm{Tr}(zI - A)^{-1}$$



$$s_{max} := \max\{z : \Phi_A(z) = \psi\}$$

Softening the Edges

$$\Phi_A(z) = \mathrm{Tr}(zI - A)^{-1}$$



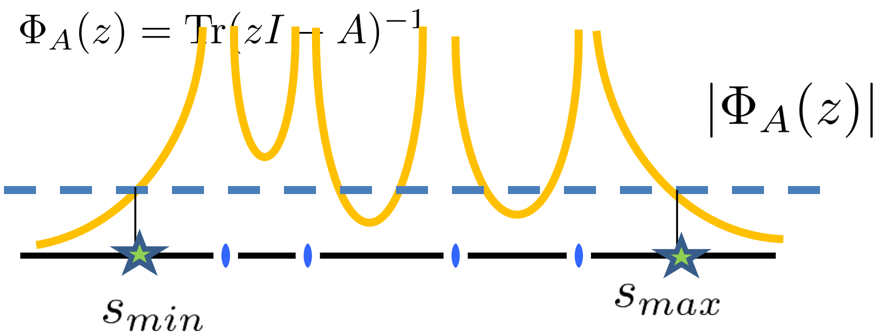
$$s_{min} := \min\{z : |\Phi_A(z)| = \psi\}$$

$$s_{max} := \max\{z : \Phi_A(z) = \psi\}$$

Inverse Stieltjes Transform Lemma

$$s_{min} := \min\{z : |\Phi_A(z)| = \psi\}$$

$$s_{max} := \max\{z : \Phi_A(z) = \psi\}$$



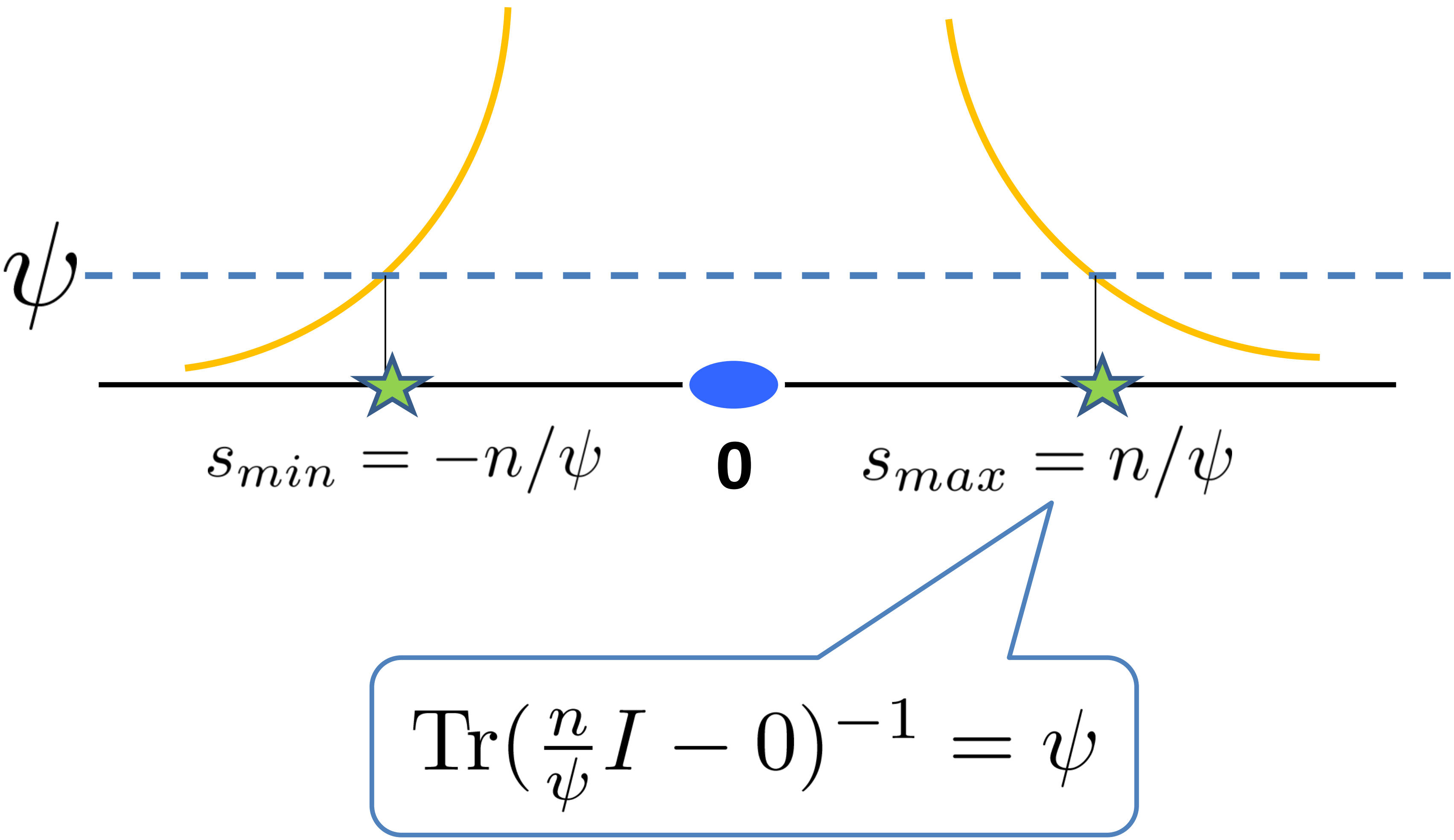
Main Lemma.

$A \succeq 0$, X **regular** isotropic random vector.
For all $\epsilon > 0$ there is $\psi = \psi(\epsilon)$ with

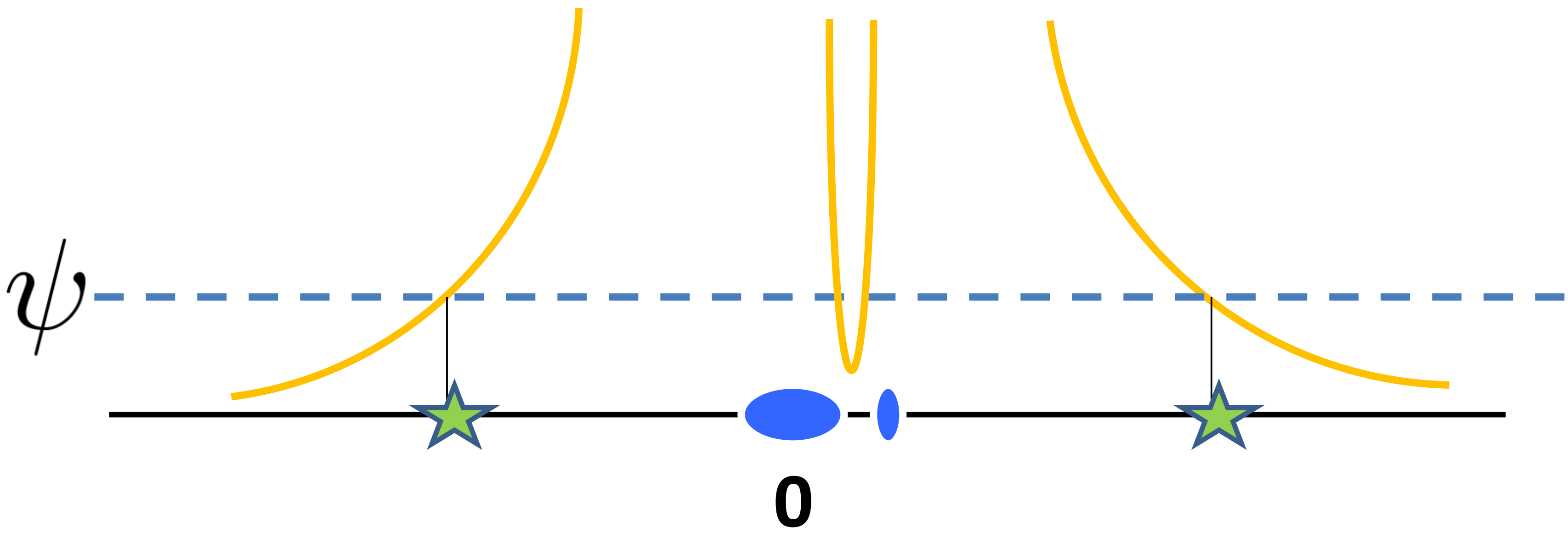
$$\mathbf{E} s_{max}(A + XX^T) \leq s_{max}(A) + 1 + \epsilon$$

$$\mathbf{E} s_{min}(A + XX^T) \geq s_{min}(A) + 1 - \epsilon.$$

Initial Positions

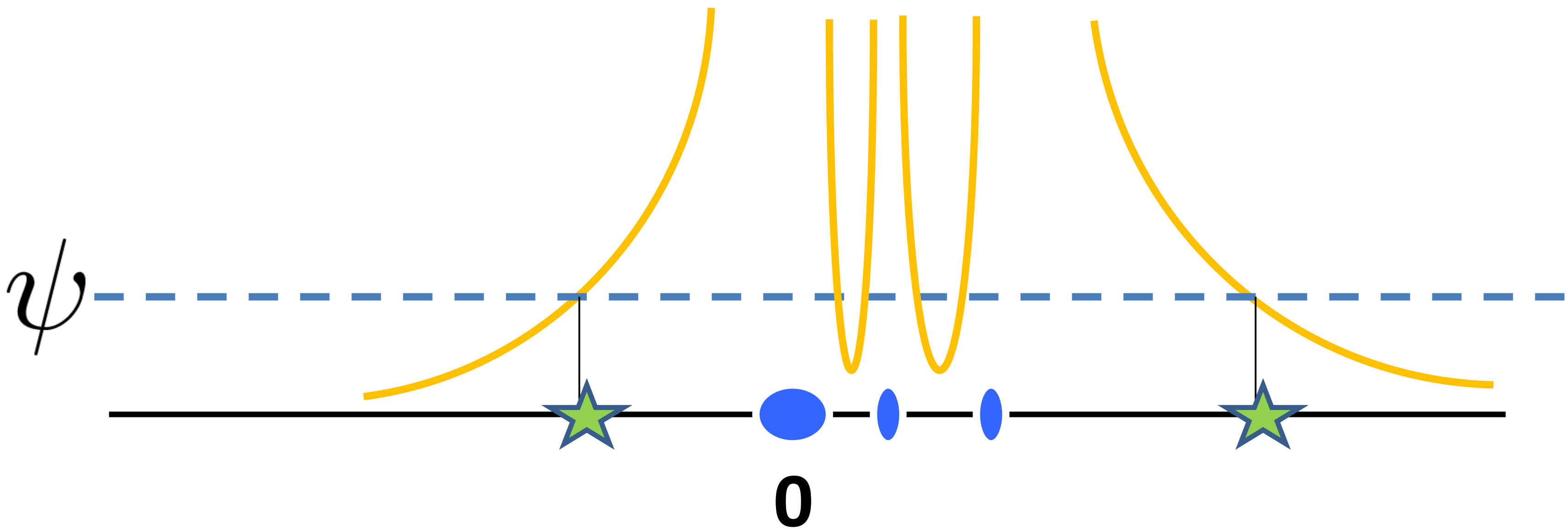


The Sensitivity Tradeoff



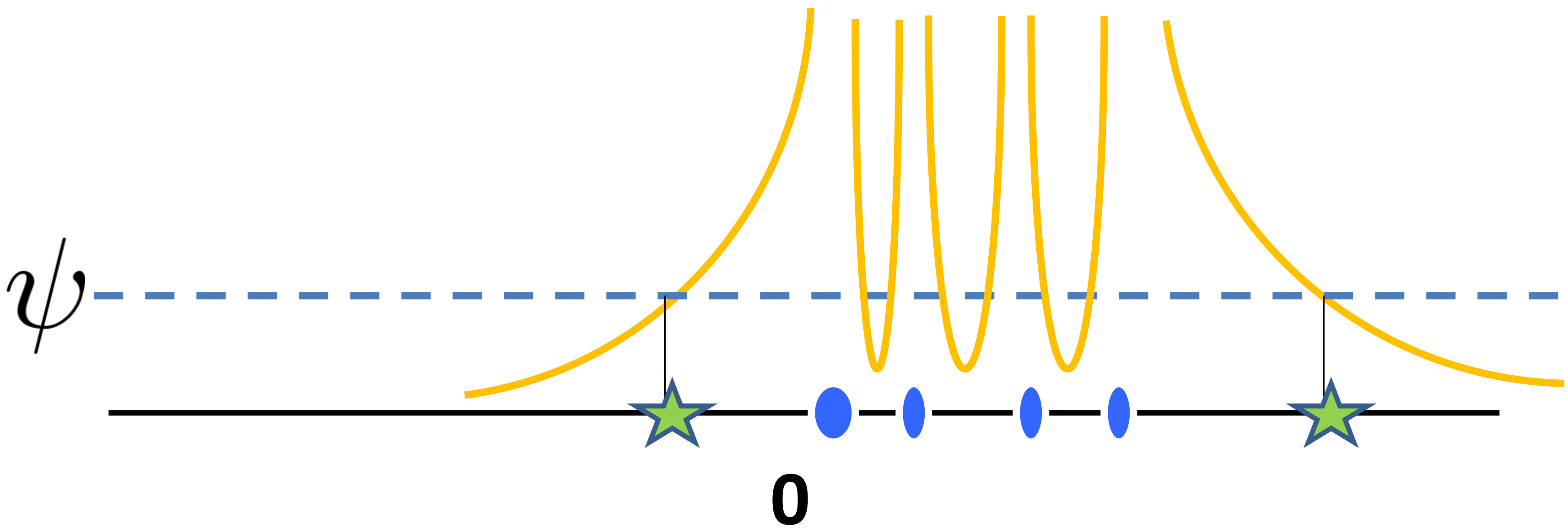
$$A_k = A_{k-1} + X_k X_k^T$$

The Sensitivity Tradeoff



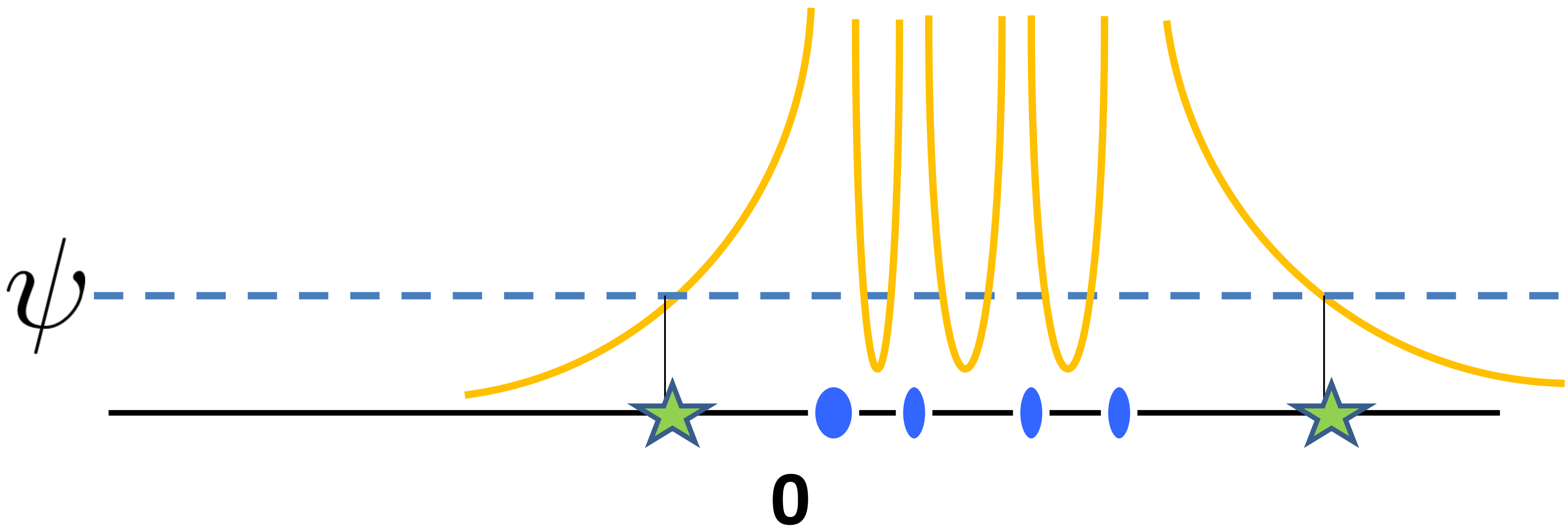
$$A_k = A_{k-1} + X_k X_k^T$$

The Sensitivity Tradeoff



$$A_k = A_{k-1} + X_k X_k^T$$

The Sensitivity Tradeoff

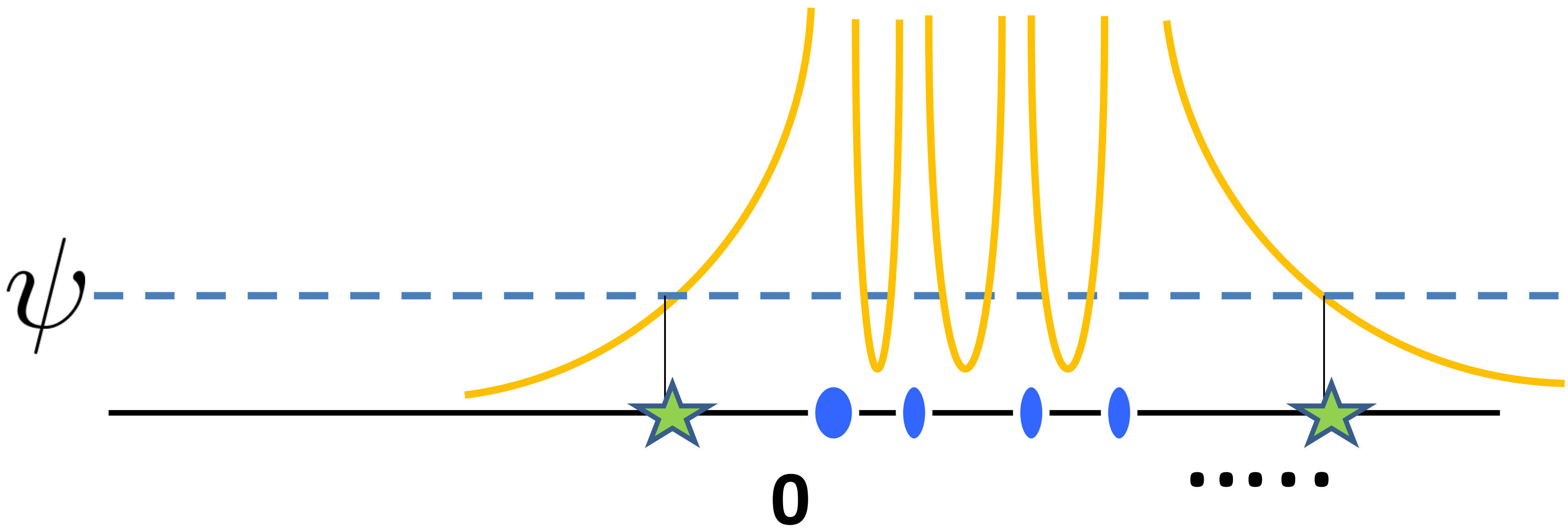


$$A_k = A_{k-1} + X_k X_k^T$$

Lemma.

$$\mathbf{E} s_{min}(A_k) \geq -n/\psi + k(1-\epsilon)$$
$$\mathbf{E} s_{max}(A_k) \leq n/\psi + k(1+\epsilon).$$

The Sensitivity Tradeoff

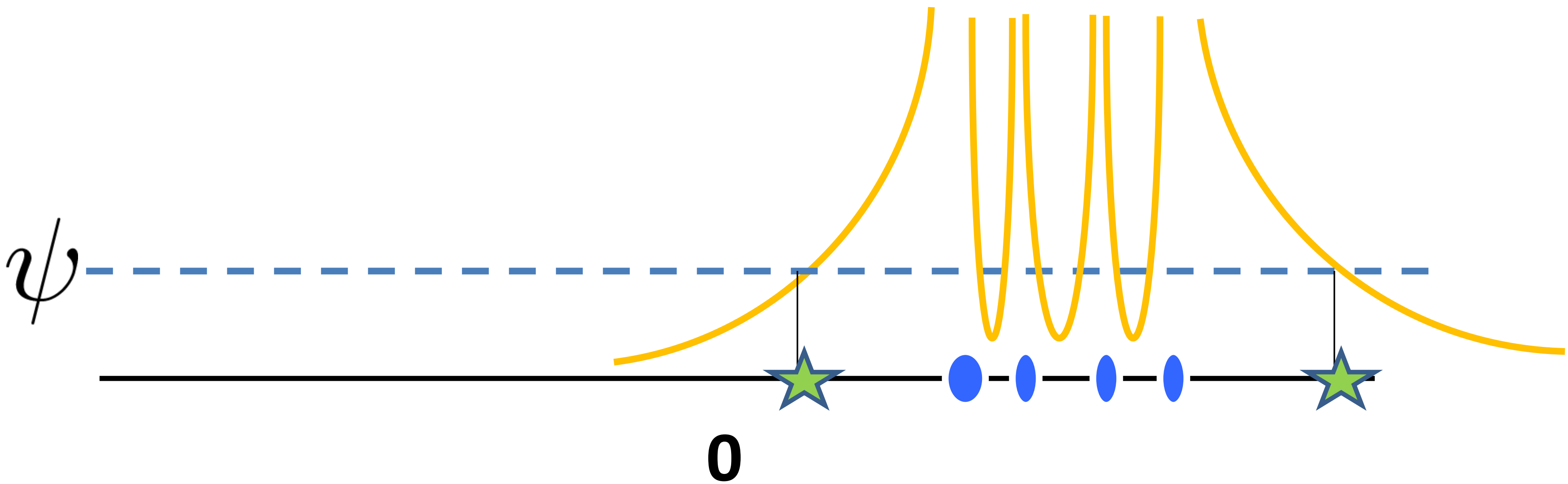


$$A_k = A_{k-1} + X_k X_k^T$$

Lemma.

$$\mathbf{E} s_{min}(A_k) \geq -n/\psi + k(1-\epsilon)$$
$$\mathbf{E} s_{max}(A_k) \leq n/\psi + k(1+\epsilon).$$

The Sensitivity Tradeoff

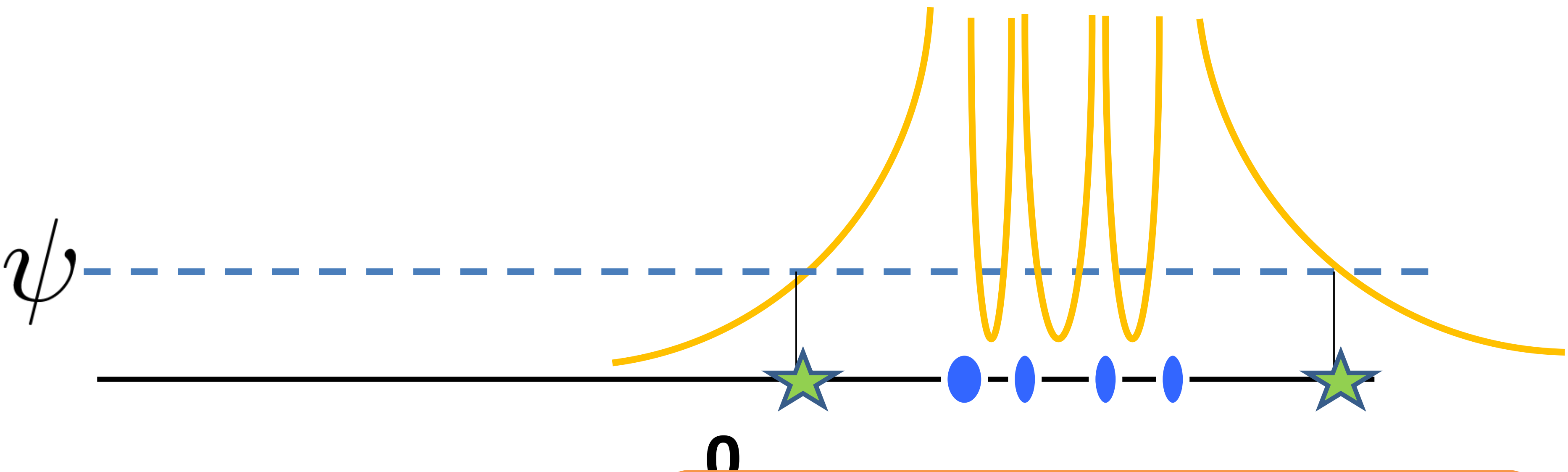


$$A_q = X_1 X_1^T + \ldots, X_q X_q^T$$

After **k** steps

$$\mathbf{E}s_{min}(A_k) \geq -n/\psi + k(1-\epsilon)$$
$$\mathbf{E}s_{max}(A_k) \leq n/\psi + k(1+\epsilon).$$

The Sensitivity Tradeoff



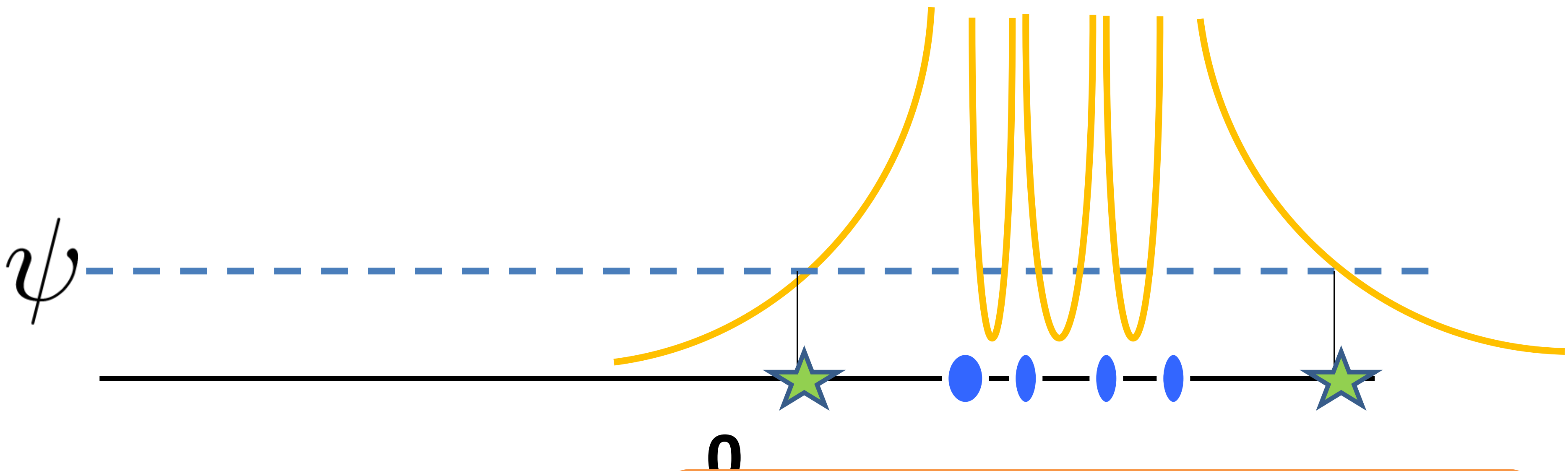
$$A_q = X_1 \ldots$$

Set $q > \frac{n}{\psi\epsilon} = c(\epsilon)n.$

After **q** steps

$$\mathbf{E}s_{min}(A_k) \geq -n/\psi + q(1-\epsilon)$$
$$\mathbf{E}s_{max}(A_k) \leq n/\psi + q(1+\epsilon).$$

The Sensitivity Tradeoff



$$A_q = X_1$$

Set $q > \frac{n}{\psi \epsilon} = c(\epsilon) n.$

After **q** steps

$$\mathbf{E} s_{min}(A_k) \geq -n/\psi + q(1-\epsilon)$$
$$\mathbf{E} s_{max}(A_k) \leq n/\psi + q(1+\epsilon).$$

$\square$

Proof of the Main Lemma

$A \succeq 0$, X **regular** isotropic random vector.
For all $\epsilon > 0$ there is $\psi = \psi(\epsilon)$ with

$$\mathbf{E} s_{max}(A + XX^T) \leq s_{max}(A) + 1 + \epsilon$$

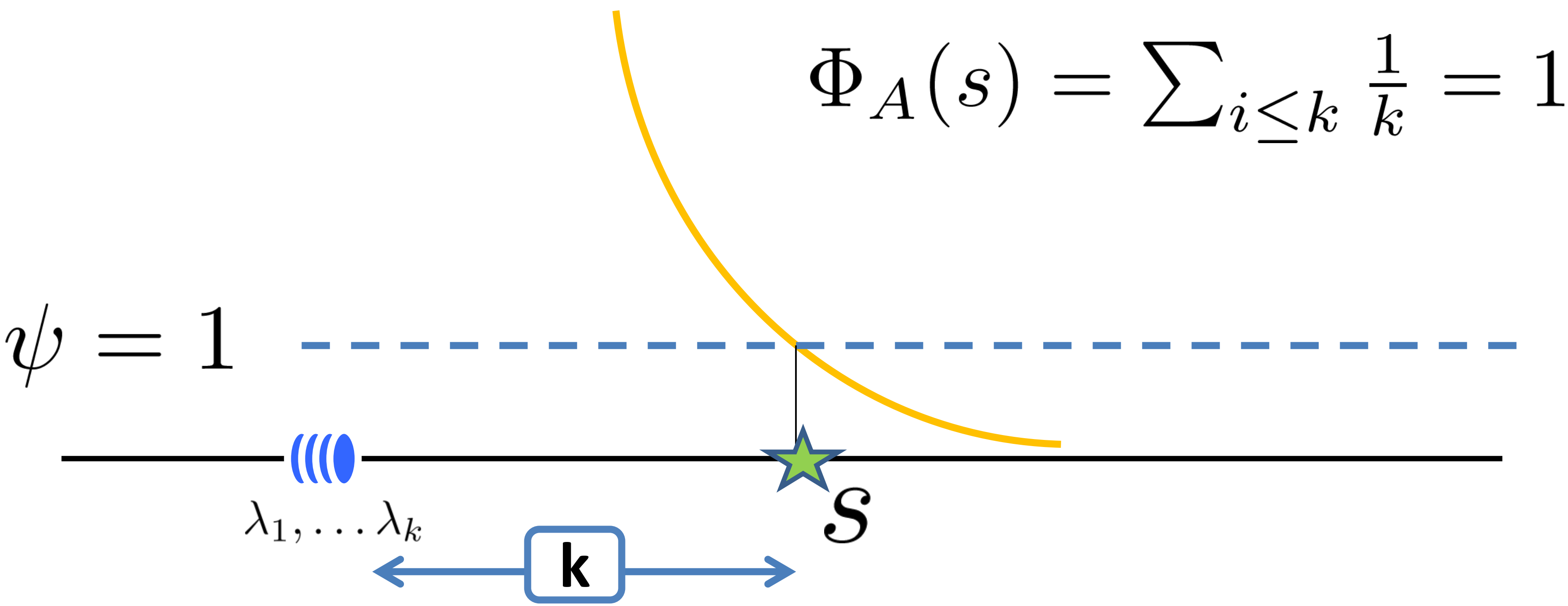
$$\mathbf{E} s_{min}(A + XX^T) \geq s_{min}(A) + 1 - \epsilon.$$

Need to solve inverse problem:

$$\mathbf{E}_X \{\max \quad z : \mathrm{Tr}(zI - A - XX^T)^{-1} = \psi\}$$

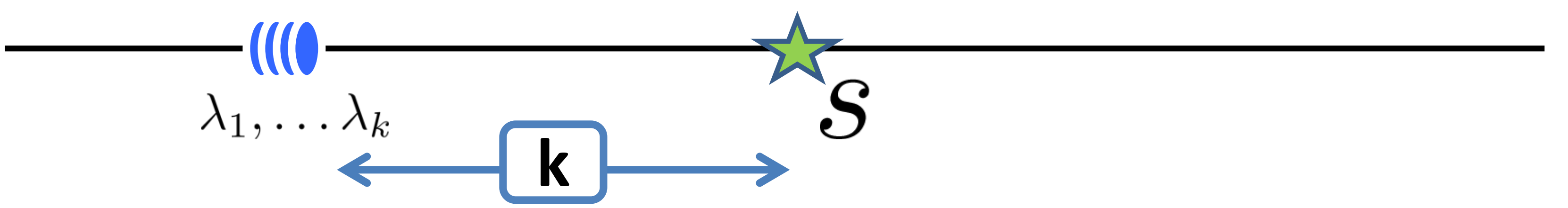
Origin of $\boxed{\text{kD}}$

A special case: **A** has **k** eigenvalues at distance **k**



Origin of $\boxed{\text{kD}}$

A special case: **A** has **k** eigenvalues at distance **k**

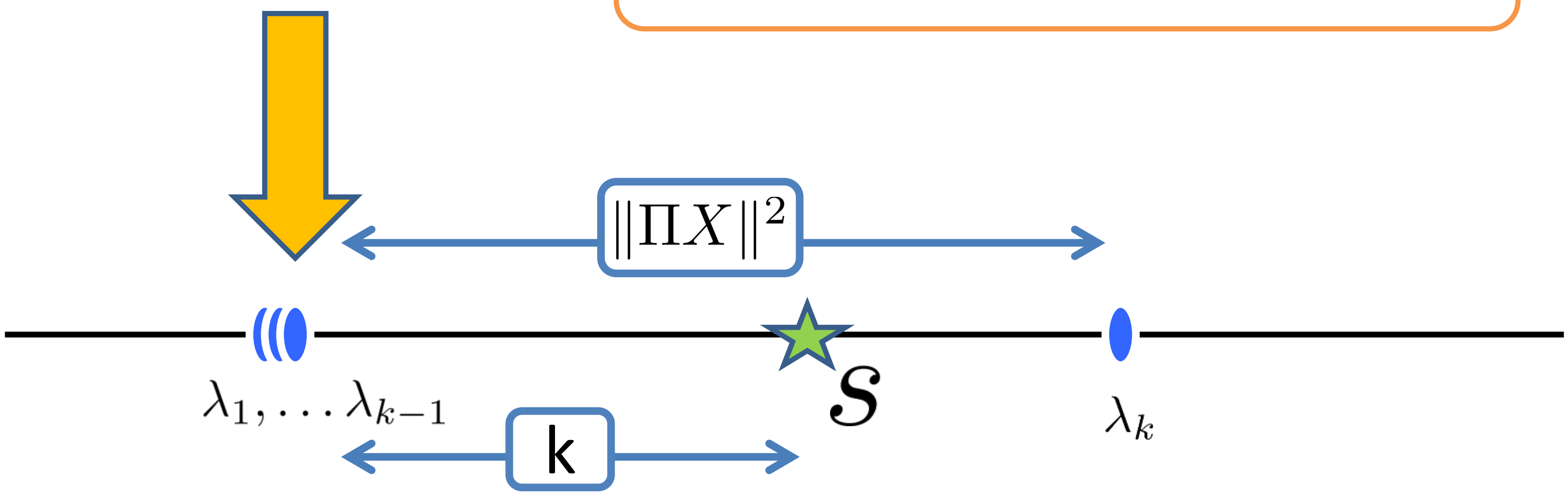


Origin of $\boxed{\text{kD}}$

A special case: **A** has **k** eigenvalues at distance **k**

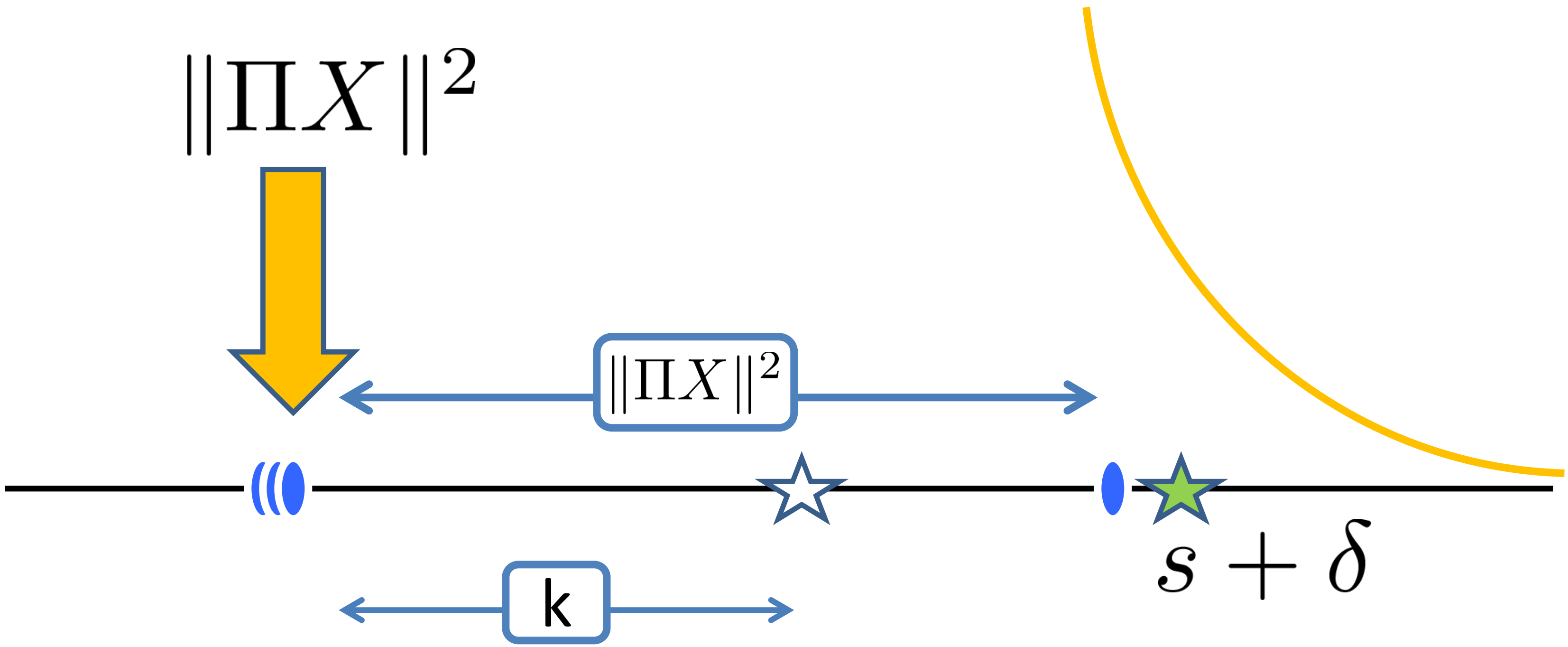
$$\|\Pi X\|^2$$

$$\boxed{\Pi = \mathrm{Proj}(u_1, \ldots, u_k)}$$



Origin of $\boxed{\text{kD}}$

A special case: **A** has **k** eigenvalues at distance **k**

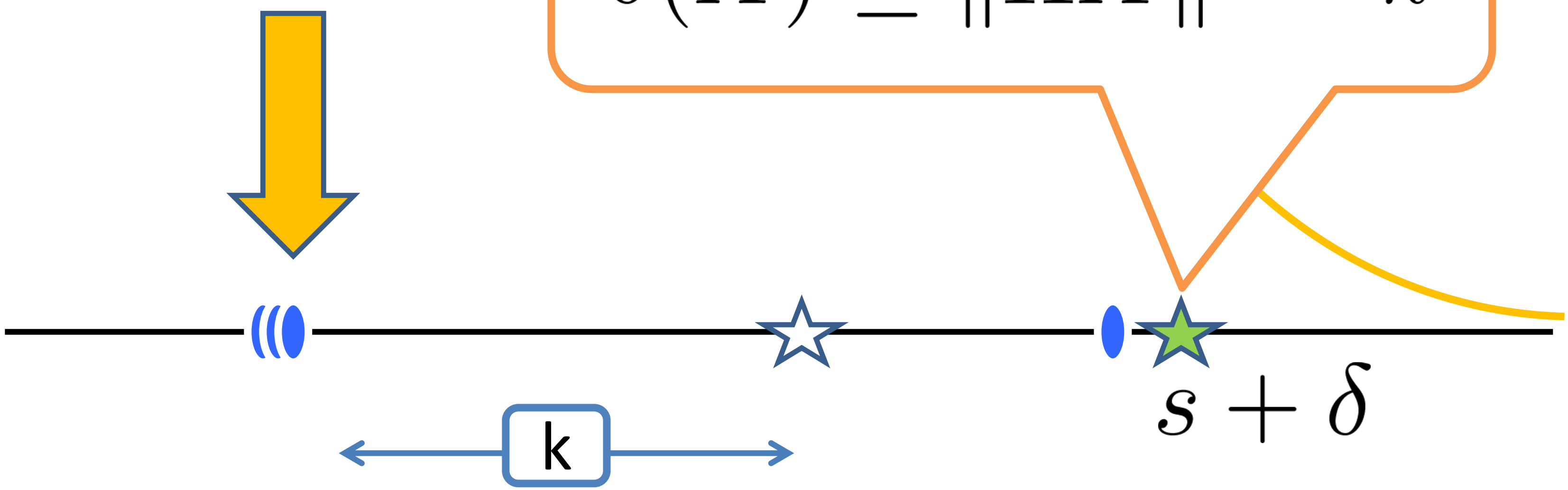


Origin of $\boxed{\text{kD}}$

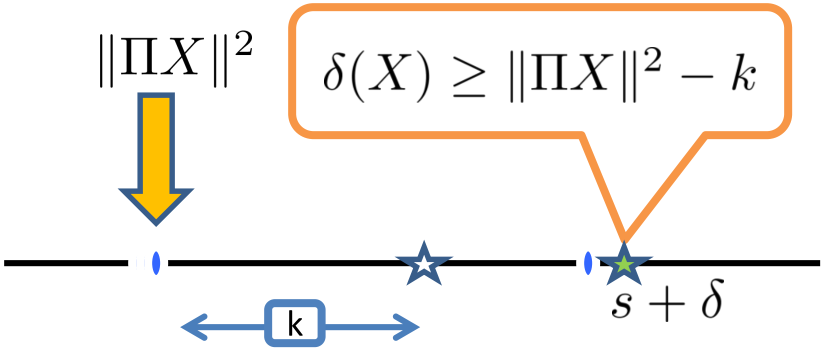
A special case: **A** has **k** eigenvalues at distance **k**

$$\|\Pi X\|^2$$

$$\delta(X) \geq \|\Pi X\|^2 - k$$

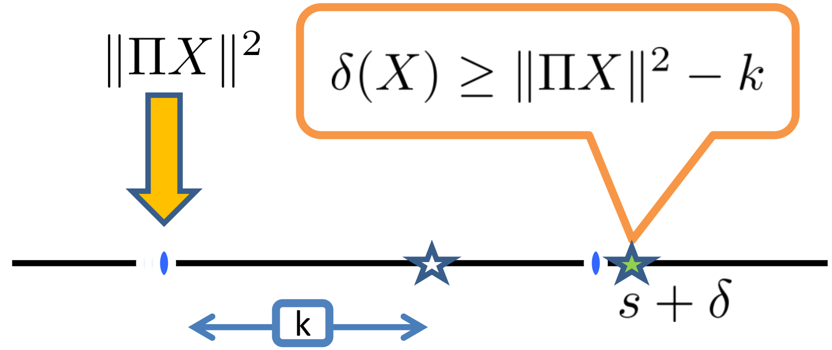


Origin of $\boxed{\text{kD}}$



$$\mathbf{E}\delta(X) \geq \mathbf{E}(\|\Pi X\|_2^2 - k)_+$$

Origin of kD



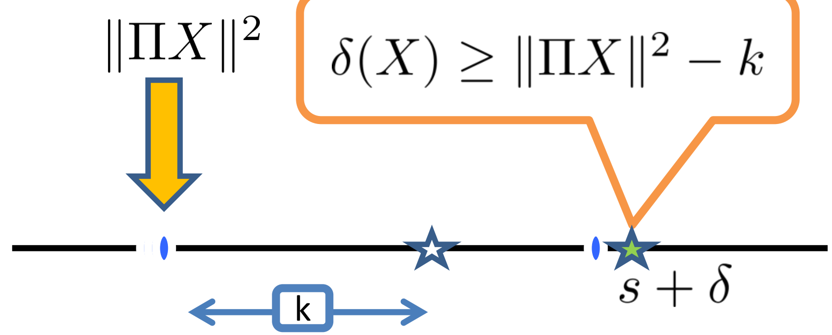
$$\mathbf{E}\delta(X) \geq \mathbf{E}(\|\Pi X\|_2^2 - k)_+$$

$$= \int_0^\infty \mathbf{P}(\|\Pi X\|^2 \geq k + t)\,dt$$

Origin of $\boxed{\text{kD}}$

Want $\mathbf{E}\delta(X) = O(1)$

$\delta(X) \geq \|\Pi X\|^2 - k$



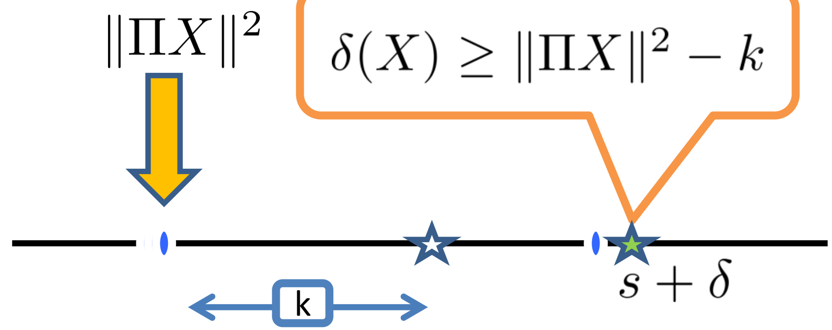
$$\mathbf{E}\delta(X) \geq \mathbf{E}(\|\Pi X\|_2^2 - k)_+$$

$$= \int_0^\infty \mathbf{P}(\|\Pi X\|^2 \geq k + t)\,dt$$

Origin of $\boxed{\text{kD}}$

Want $\mathbf{E}\delta(X) = O(1)$

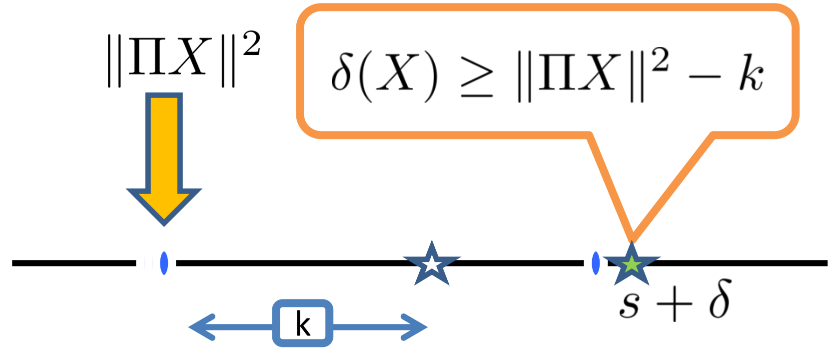
$\delta(X) \geq \|\Pi X\|^2 - k$



$$\mathbf{E}\delta(X) \geq \mathbf{E}(\|\Pi X\|_2^2 - k)_+$$
$$= \int_0^\infty \mathbf{P}(\|\Pi X\|^2 \geq k + t)dt$$

Need $\mathbf{P}(\|\Pi X\|^2 > k + t) \leq C/t^{1+\eta}$.

Origin of kD



$$\mathbf{E}\delta(X) \geq \mathbf{E}(\|\Pi X\|_2^2 - k)_+$$

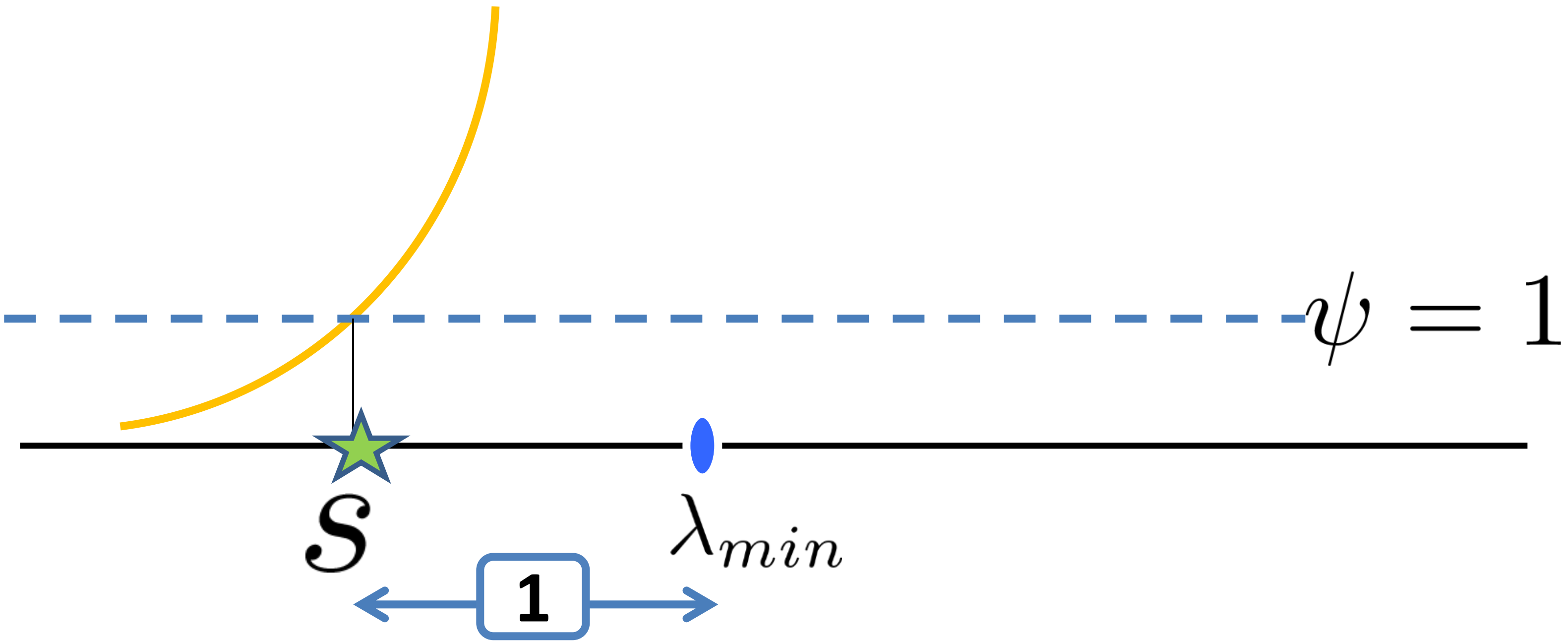
$$= \int_0^\infty \mathbf{P}(\|\Pi X\|^2 \geq k + t)dt$$

kD

Need $\mathbf{P}(\|\Pi X\| > t) \leq C/t^{2+\eta}, t > k.$

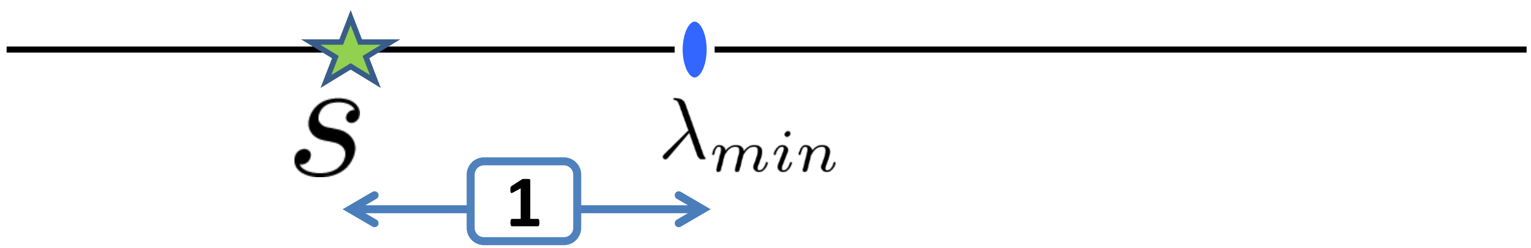
Origin of $\boxed{\text{1D}}$

A has **1** eigenvalue at distance **1**



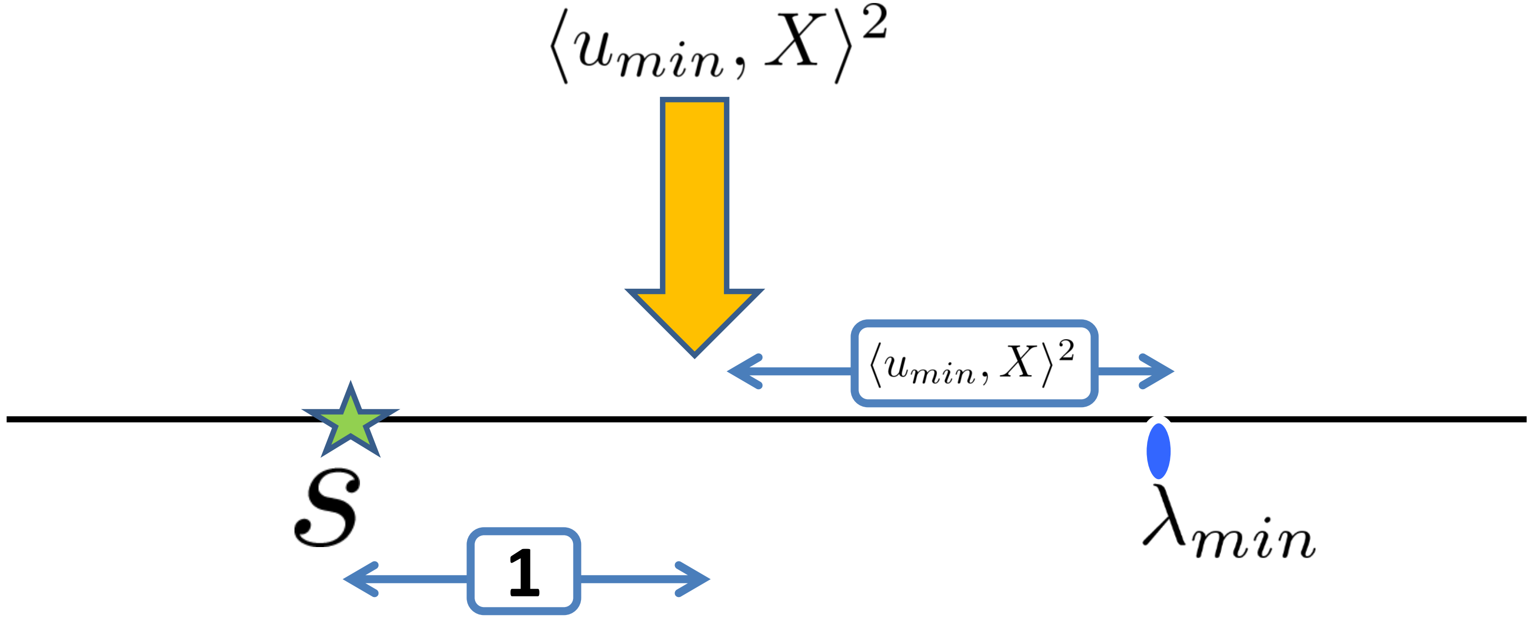
Origin of $\boxed{\text{1D}}$

A has **1** eigenvalue at distance **1**



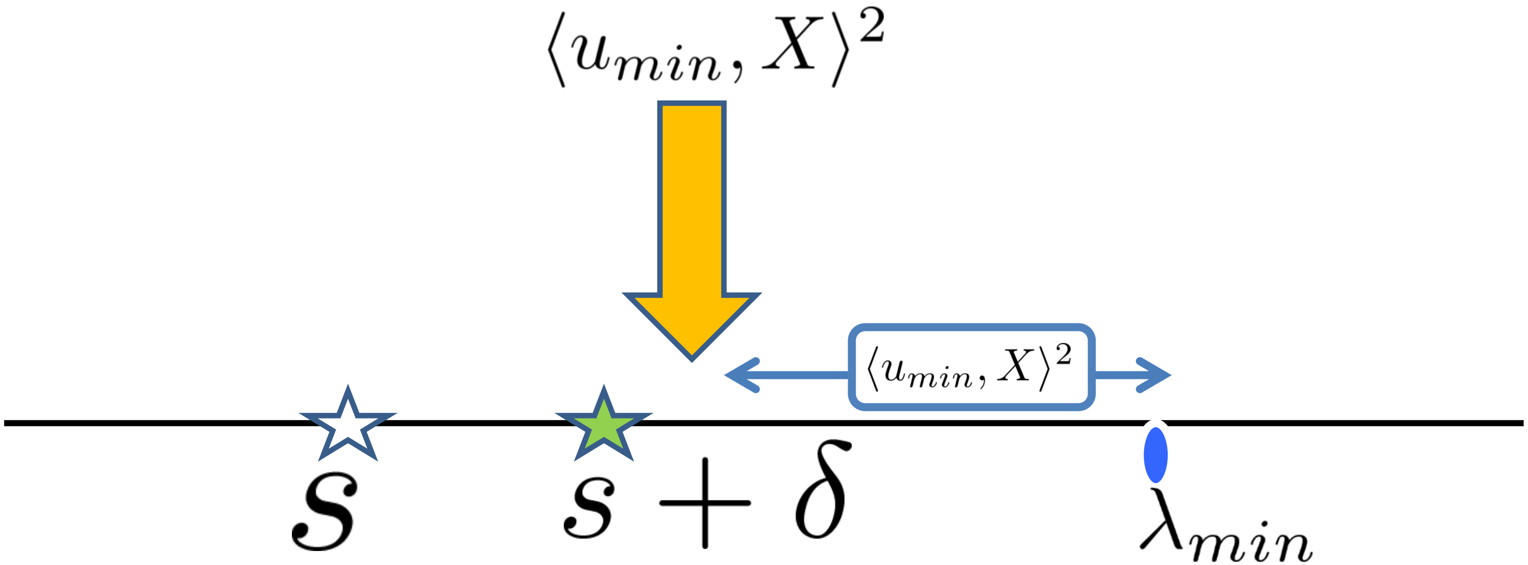
Origin of $\boxed{\text{1D}}$

A has **1** eigenvalue at distance **1**



Origin of $\boxed{\text{1D}}$

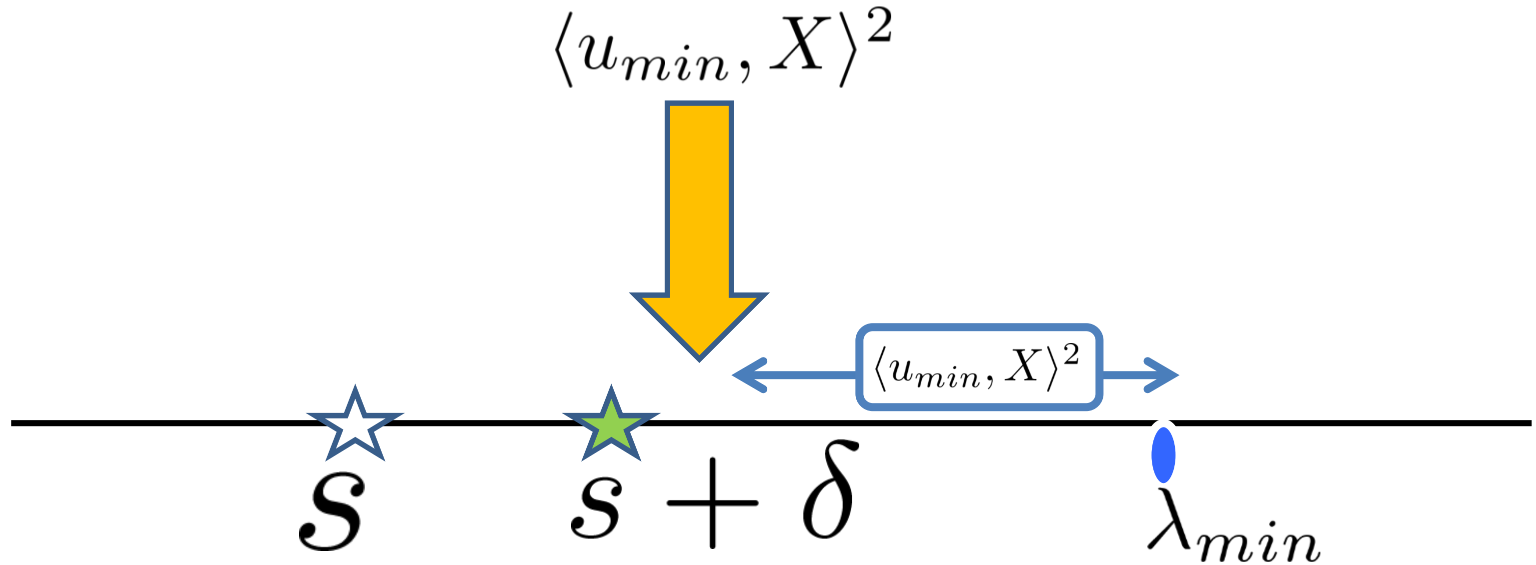
A has **1** eigenvalue at distance **1**



$$\delta(X) = \langle u_{min}, X\rangle^2, \text{ want } \geq 1 - \epsilon \text{ wcp.}$$

Origin of $\boxed{\text{1D}}$

A has **1** eigenvalue at distance **1**



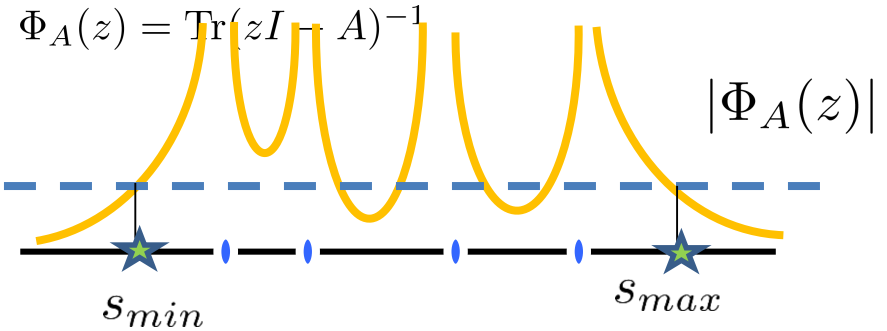
$$\delta(X) = \langle u_{min}, X\rangle^2, \text{ want } \geq 1 - \epsilon \text{ wcp.}$$

$$\boxed{\text{Control } \mathbf{E}|\langle u_{min}, X\rangle|^{2+\eta}.} \quad \boxed{\text{1D}}$$

Main Lemma

$$s_{min} := \min\{z : |\Phi_A(z)| = \psi\}$$

$$s_{max} := \max\{z : \Phi_A(z) = \psi\}$$



Main Lemma.

$A \succeq 0$, X **regular** isotropic random vector.
For all $\epsilon > 0$ there is $\psi = \psi(\epsilon)$ with

$$\mathbf{E}s_{max}(A + XX^T) \leq s_{max}(A) + 1 + \epsilon$$

$$\mathbf{E}s_{min}(A + XX^T) \geq s_{min}(A) + 1 - \epsilon.$$

Some Technical Details

Proof of the Main Lemma

$A \succeq 0$, X **regular** isotropic random vector.
For all $\epsilon > 0$ there is $\psi = \psi(\epsilon)$ with

$$\mathbf{E} s_{max}(A + XX^T) \leq s_{max}(A) + 1 + \epsilon$$

$$\mathbf{E} s_{min}(A + XX^T) \geq s_{min}(A) + 1 - \epsilon.$$

For fixed **A**, **X**, how do we certify

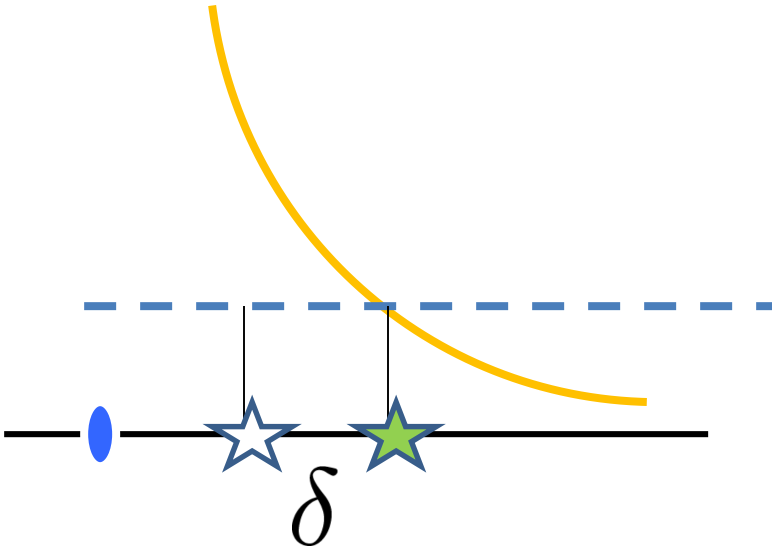
$$s_{max}(A + XX^T) \leq s_{max}(A) + \delta \text{ ?}$$

Upper Edge Shifts

Let $s_{max}(A) = s.$

$$s_{max}(A + XX^T) \leq s + \delta$$

$$\iff \Phi_{A+XX^T}(s + \delta) \leq \psi$$



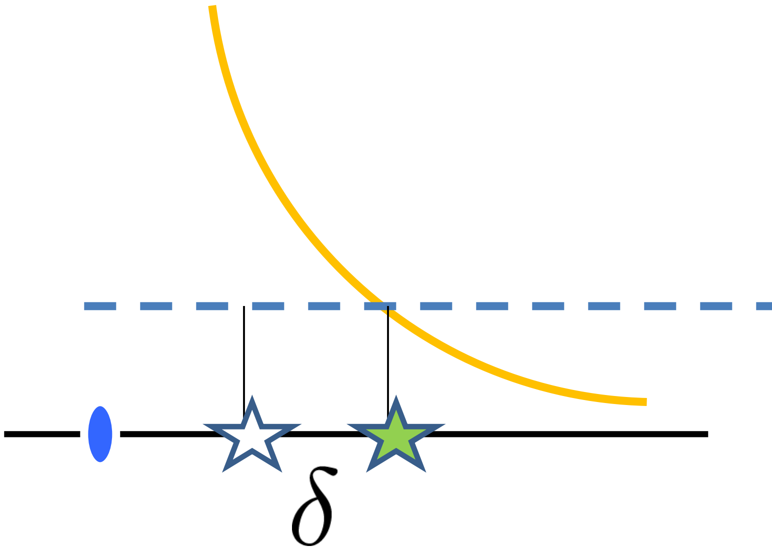
Upper Edge Shifts

Let $s_{max}(A) = s.$

$s_{max}(A + XX^T) \leq s + \delta$

$\iff \Phi_{A+XX^T}(s+\delta) \leq \psi$

$\iff \Phi_{A+XX^T}(s+\delta) \leq \Phi_A(s)$



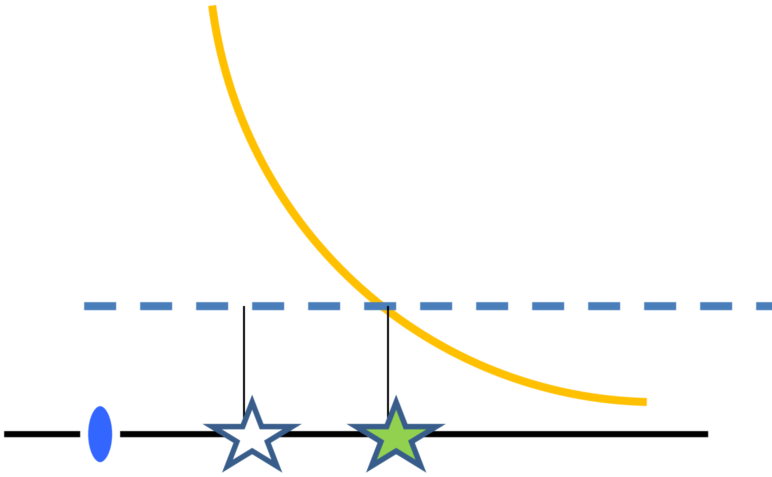
Upper Edge Shifts

Let $s_{max}(A) = s.$

$s_{max}(A + XX^T) \leq s + \delta$

$\iff \Phi_{A+XX^T}(s+\delta) \leq \psi$

$\iff \Phi_{A+XX^T}(s+\delta) \leq \Phi_A(s)$



$\iff \mathrm{Tr}(s + \delta - A - XX^T)^{-1} \leq \mathrm{Tr}(s - A)^{-1}$

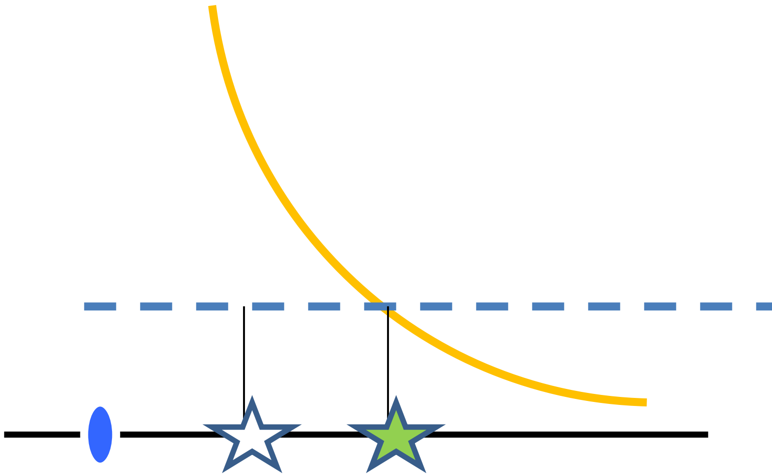
Upper Edge Shifts

Let $s_{max}(A) = s.$

$s_{max}(A + XX^T) \leq s + \delta$

$\iff \Phi_{A+XX^T}(s+\delta) \leq \psi$

$\iff \Phi_{A+XX^T}(s+\delta) \leq \Phi_A(s)$



$$\iff \mathrm{Tr}(s + \delta - A - XX^T)^{-1} \leq \mathrm{Tr}(s - A)^{-1}$$

Sherman Morrisson Formula

$$(A - XX^T)^{-1} = A^{-1} + \frac{A^{-1}XX^TA^{-1}}{1 - X^TA^{-1}X}.$$

Upper Edge Shifts

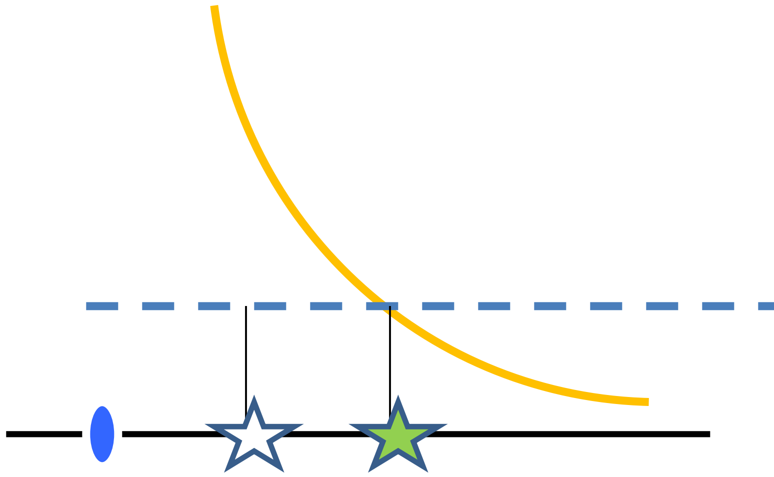
Let $s_{max}(A) = s.$

$$s_{max}(A + XX^T) \leq s + \delta$$

$$\iff \Phi_{A+XX^T}(s+\delta) \leq \psi$$

$$\iff \Phi_{A+XX^T}(s+\delta) \leq \Phi_A(s)$$

$$\iff \frac{X^T(s+\delta-A)^{-2}X}{\delta \mathrm{Tr}(s+\delta-A)^{-2}} + X^T(s+\delta-A)^{-1}X \leq 1.$$



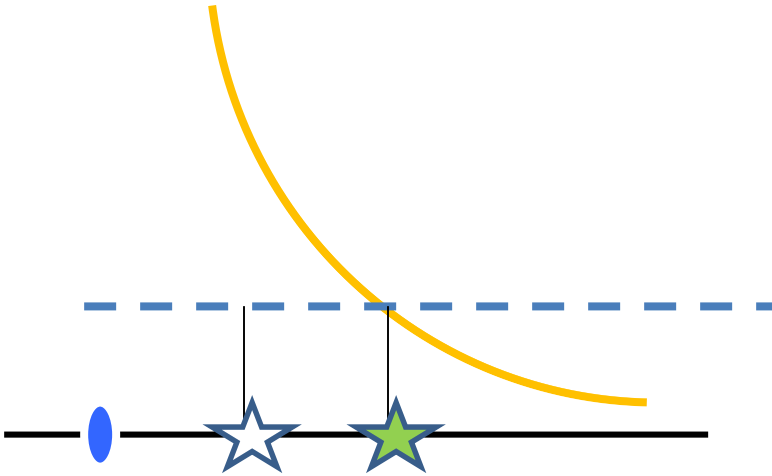
Upper Edge Shifts

Let $s_{max}(A) = s.$

$$s_{max}(A + XX^T) \leq s + \delta$$

$$\iff \Phi_{A+XX^T}(s+\delta) \leq \psi$$

$$\iff \Phi_{A+XX^T}(s+\delta) \leq \Phi_A(s)$$



$$\iff \frac{X^T(s+\delta-A)^{-2}X}{\delta \mathrm{Tr}(s+\delta-A)^{-2}} + X^T(s+\delta-A)^{-1}X \leq 1.$$

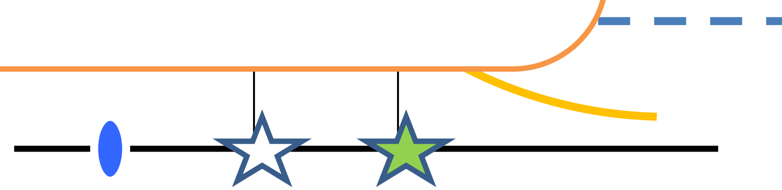
Quadratic form in **X**, $Q(\delta, X)$
decreasing in δ

$$\delta(X) := \min\{\delta : Q(\delta, X) \leq 1\}$$

"Inverse Problem"

Need to bound

$$\mathbf{E}_X s_{max}(A + XX^T) = s_{max}(A) + \mathbf{E}_X \delta(X)$$



$$s_{max}(A + XX^T) \leq s + \delta$$

$$\iff \frac{X^T(s+\delta-A)^{-2}X}{\delta \mathrm{Tr}(s+\delta-A)^{-2}} + X^T(s+\delta-A)^{-1}X \leq 1.$$

Quadratic form in **X**, $Q(\delta, X)$

decreasing in δ

"Heuristic" bound on $\mathbf{E}\delta(X)$

$$s_{max}(A + XX^T) \leq s_{max}(A) + \delta$$

$$\Updownarrow$$

$$\frac{X^T(s+\delta-A)^{-2}X}{\delta \mathrm{Tr}(s+\delta-A)^{-2}} + X^T(s+\delta-A)^{-1}X \leq 1.$$

“Heuristic” bound on $\mathbf{E}\delta(X)$

$$\mathbf{E}s_{max}(A + XX^T) \leq s_{max}(A) + \delta$$

$$\Updownarrow$$

$$\mathbf{E}\frac{X^T(s+\delta-A)^{-2}X}{\delta \mathrm{Tr}(s+\delta-A)^{-2}} + \mathbf{E}X^T(s+\delta-A)^{-1}X \leq 1.$$

“Heuristic” bound on $\mathbf{E}\delta(X)$

$$\mathbf{E}s_{max}(A + XX^T) \leq s_{max}(A) + \delta$$

$$\Updownarrow$$

$$\frac{(s + \delta - A)^{-2} \bullet \mathbf{E}XX^T}{\delta \mathrm{Tr}(s + \delta - A)^{-2}} + (s + \delta - A)^{-1} \bullet \mathbf{E}XX^T \leq 1.$$

"Heuristic" bound on $\mathbf{E}\delta(X)$

$$\mathbf{E}s_{max}(A + XX^T) \leq s_{max}(A) + \delta$$

$$\frac{(s + \delta - A)^{-2} \bullet \mathbf{E}XX^T}{\delta \mathrm{Tr}(s + \delta - A)^{-2}} + (s + \delta - A)^{-1} \bullet \mathbf{E}XX^T \leq 1.$$

$$\boxed{\mathbf{E}XX^T = I}$$

"Heuristic" bound on $\mathbf{E}\delta(X)$

$$\mathbf{E}s_{max}(A + XX^T) \leq s_{max}(A) + \delta$$

$$\Updownarrow$$

$$\frac{\mathrm{Tr}(s+\delta-A)^{-2}}{\delta\mathrm{Tr}(s+\delta-A)^{-2}} + \mathrm{Tr}(s+\delta-A)^{-1} \leq 1.$$

"Heuristic" bound on $\mathbf{E}\delta(X)$

$$\mathbf{E}s_{max}(A + XX^T) \leq s_{max}(A) + \delta$$

$$\frac{1}{\delta} + \mathrm{Tr}(s + \delta - A)^{-1} \leq 1.$$

Sensitivity Bound

“Heuristic” bound on $\mathbf{E}\delta(X)$

$$\mathbf{E}s_{max}(A + XX^T) \leq s_{max}(A) + \delta$$

$$\Uparrow$$

$$\frac{1}{\delta} + \mathrm{Tr}(s - A)^{-1} \leq 1.$$

“Heuristic” bound on $\mathbf{E}\delta(X)$

$$\mathbf{E}s_{max}(A + XX^T) \leq s_{max}(A) + \delta$$

$$\Uparrow$$

$$\frac{1}{\delta} + \psi \leq 1.$$

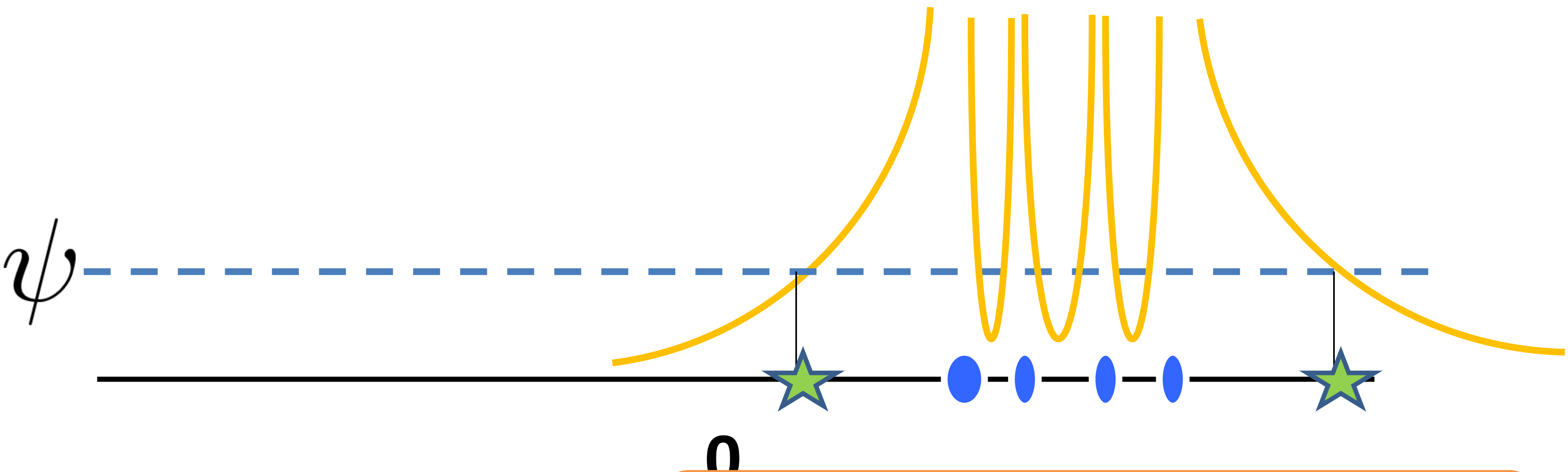
“Heuristic” bound on $\mathbf{E}\delta(X)$

$$\mathbf{E}s_{max}(A + XX^T) \leq s_{max}(A) + \delta$$

$$\frac{1}{\delta} + \psi \leq 1.$$

$$\psi \leq \frac{\delta - 1}{\delta} \approx \epsilon, \qquad \delta = 1 + \epsilon$$

The Sensitivity Tradeoff



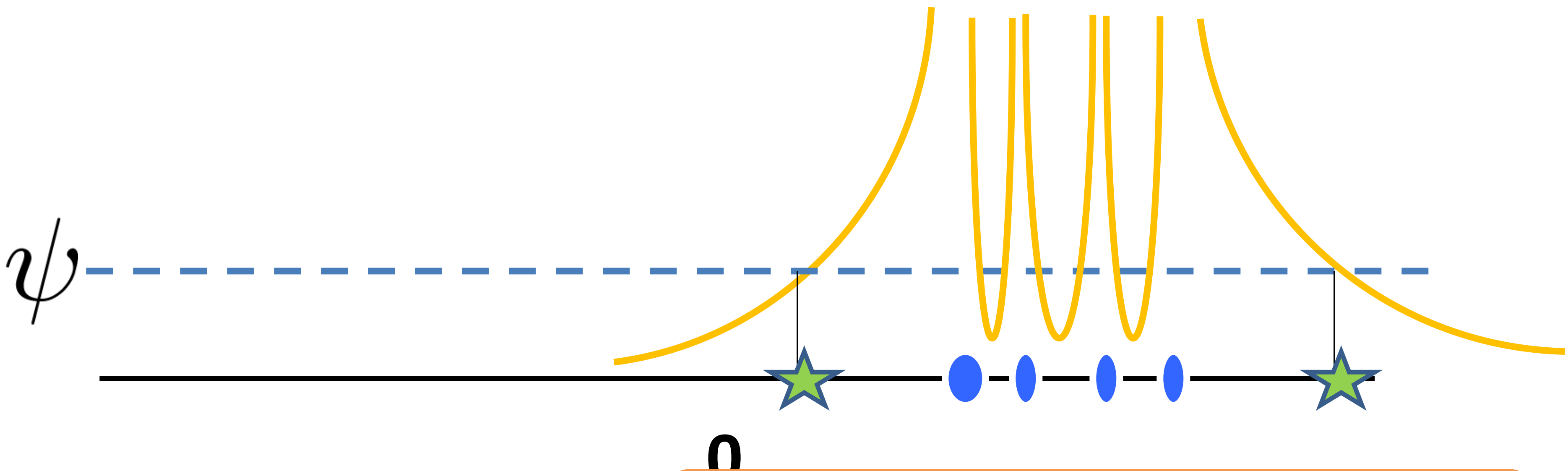
$A_q = X_1$

Set $q > \frac{n}{\psi \epsilon} = c(\epsilon) n.$

After **q** steps

$$\mathbf{E} s_{min}(A_k) \leq -n/\psi + q(1-\epsilon)$$
$$\mathbf{E} s_{max}(A_k) \geq n/\psi + q(1+\epsilon).$$

The Sensitivity Tradeoff



$$A_q = X_1$$

Set $q > \frac{n}{\psi\epsilon} = n/\epsilon^2.$

After **q** steps

$$\mathbf{E}s_{min}(A_k) \leq -n/\psi + q(1-\epsilon)$$
$$\mathbf{E}s_{max}(A_k) \geq n/\psi + q(1+\epsilon).$$

Nonsense?

Need to bound $\mathbf{E}_X \delta(X)$ for δ satisfying

$$Q(\delta, X) = \frac{X^T (s + \delta - A)^{-2} X}{\delta \mathrm{Tr}(s + \delta - A)^{-2}} + X^T (s + \delta - A)^{-1} X \leq 1.$$

Insufficient to bound "in expectation"

Nonsense?

Need to bound $\mathbf{E}_X \delta(X)$ for δ satisfying

$$Q(\delta, X) = \frac{X^T(s+\delta-A)^{-2}X}{\delta \mathrm{Tr}(s+\delta-A)^{-2}} + X^T(s+\delta-A)^{-1}X \leq 1.$$

If **X** is Gaussian: $Q(\delta, X) \approx \mathbf{E}Q(\delta, X)$ whp.

so Heuristic calculation is accurate.

Nonsense?

Need to bound $\mathbf{E}_X \delta(X)$ for δ satisfying

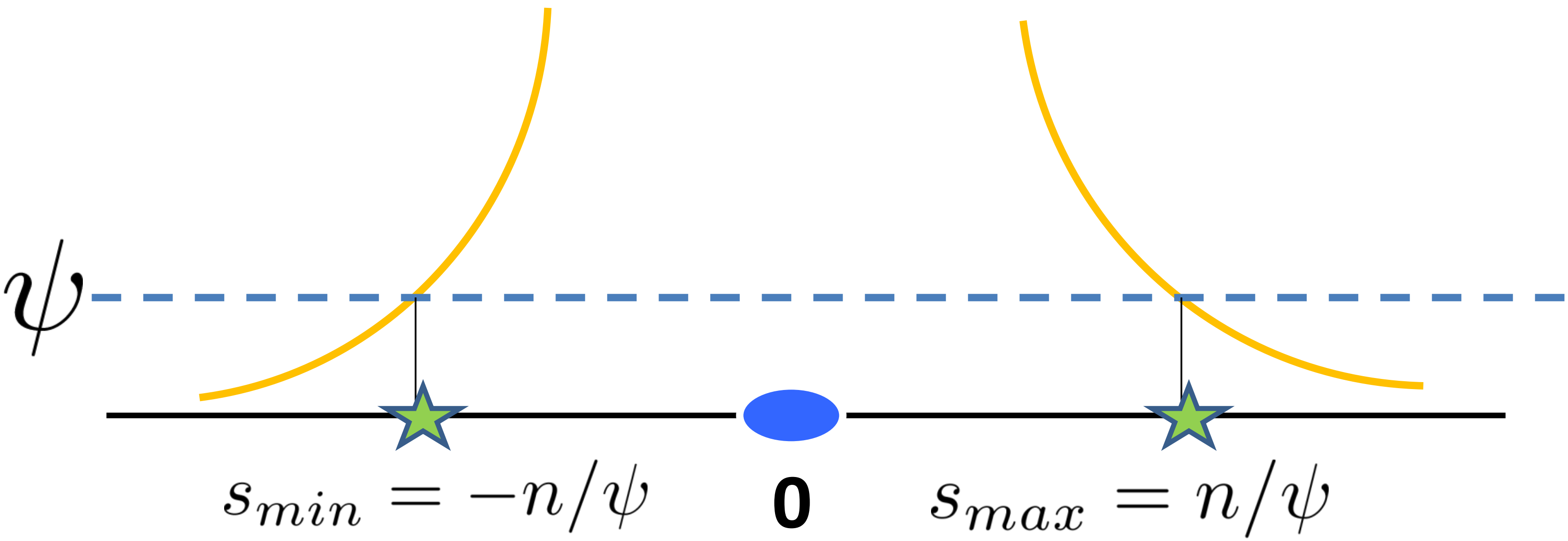
$$Q(\delta, X) = \frac{X^T(s+\delta-A)^{-2}X}{\delta \mathrm{Tr}(s+\delta-A)^{-2}} + X^T(s+\delta-A)^{-1}X \leq 1.$$

If **X** is Gaussian: $Q(\delta, X) \approx \mathbf{E}Q(\delta, X)$ whp.

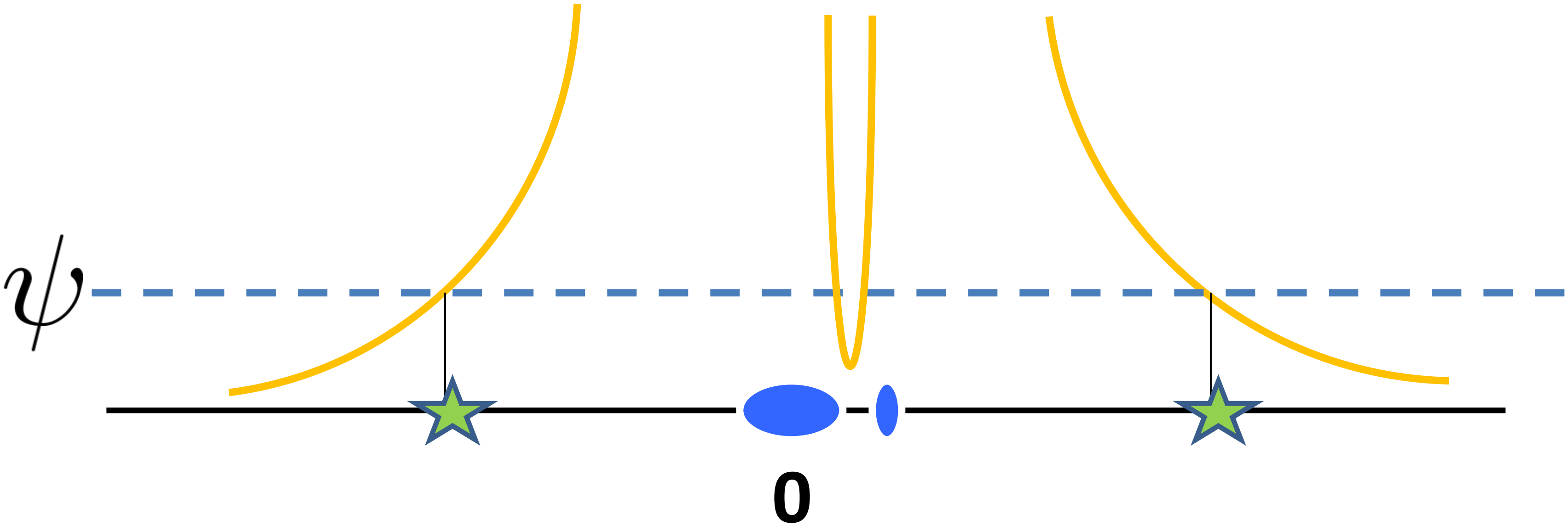
so Heuristic calculation is accurate.

Turns out $\boxed{\text{1D}}$ + $\boxed{\text{kD}}$ is enough.

End of the Proof

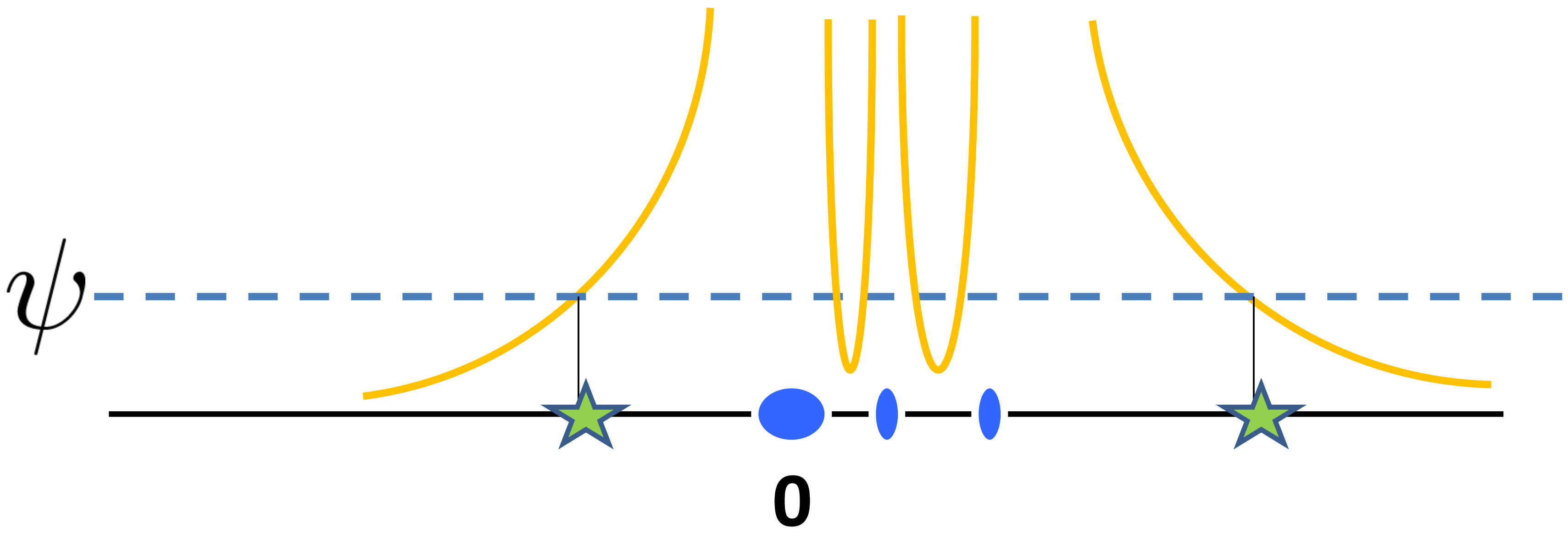


End of the Proof



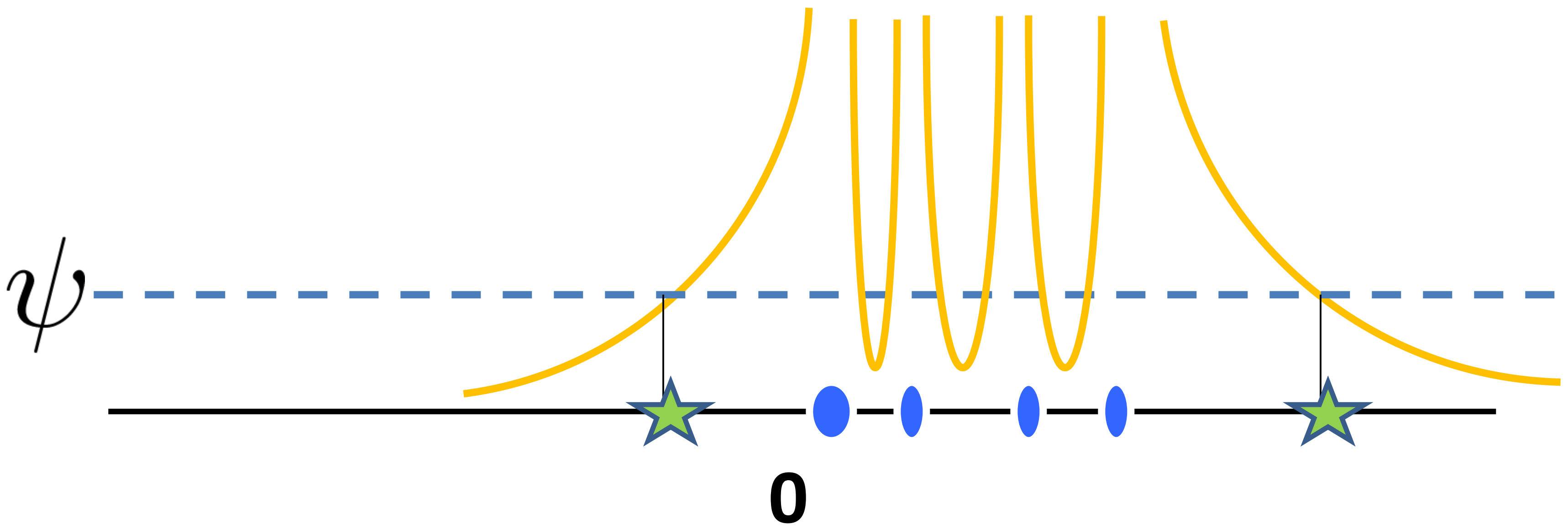
$$A_k = A_{k-1} + X_k X_k^T$$

End of the Proof



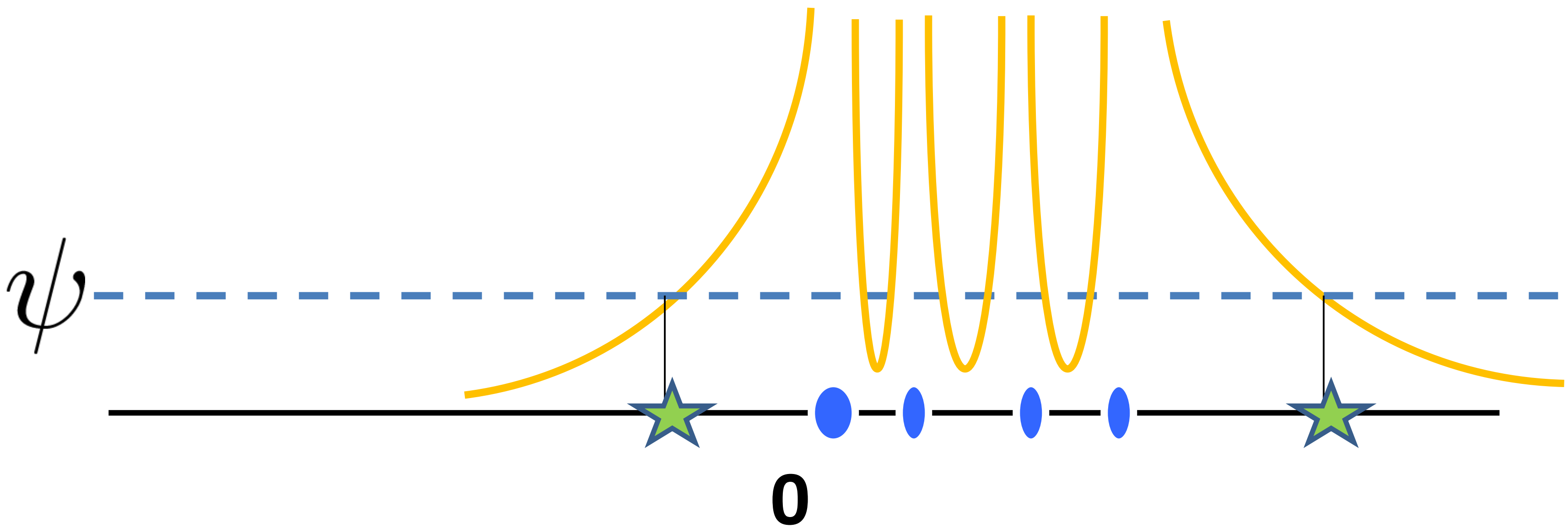
$$A_k = A_{k-1} + X_k X_k^T$$

End of the Proof



$$A_k = A_{k-1} + X_k X_k^T$$

End of the Proof

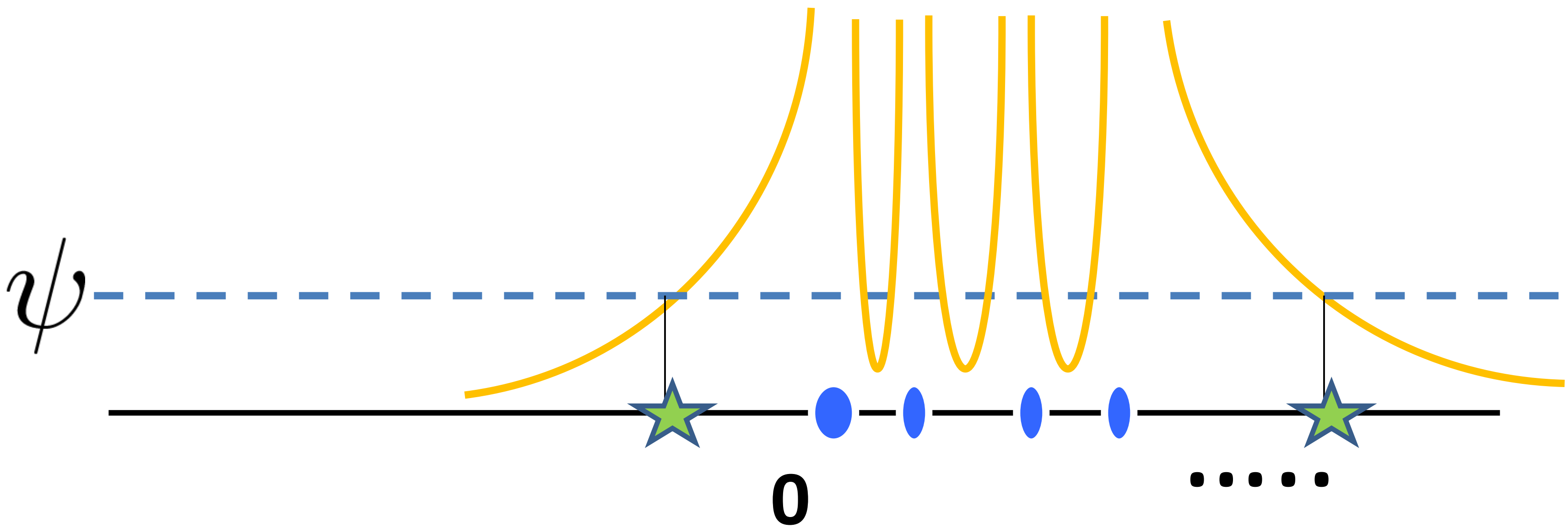


$$A_k = A_{k-1} + X_k X_k^T$$

Lemma.

$$\mathbf{E} s_{min}(A_k) \leq -n/\psi + k(1-\epsilon)$$
$$\mathbf{E} s_{max}(A_k) \geq n/\psi + k(1+\epsilon).$$

End of the Proof

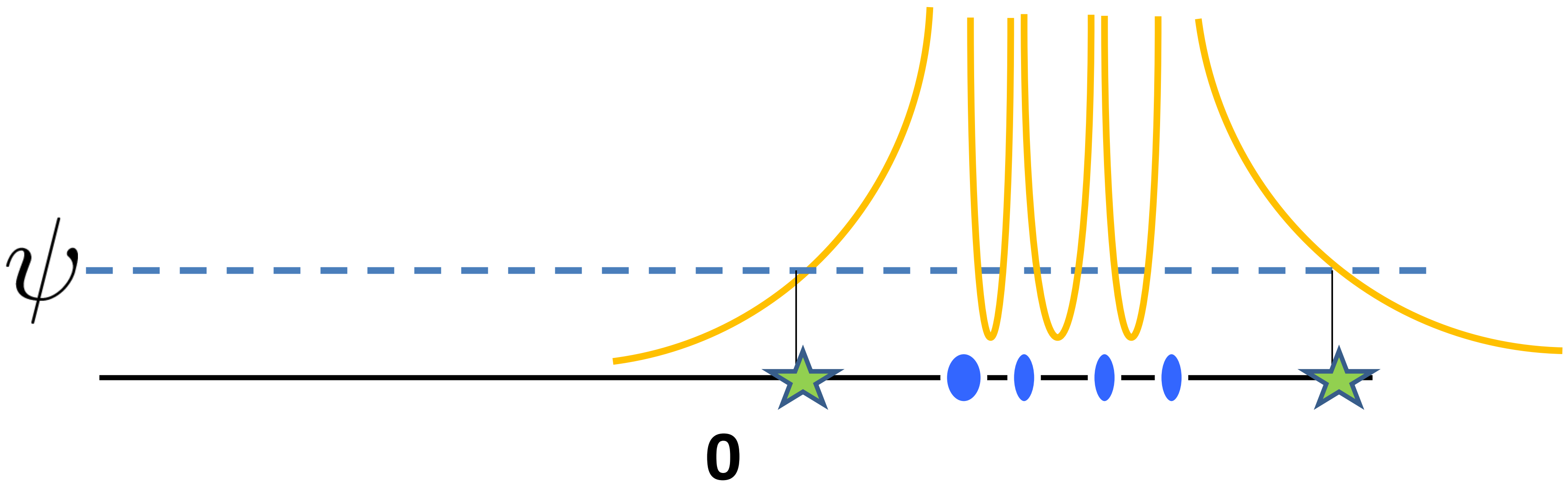


$$A_k = A_{k-1} + X_k X_k^T$$

Lemma.

$$\mathbf{E} s_{min}(A_k) \leq -n/\psi + k(1-\epsilon)$$
$$\mathbf{E} s_{max}(A_k) \geq n/\psi + k(1+\epsilon).$$

End of the Proof

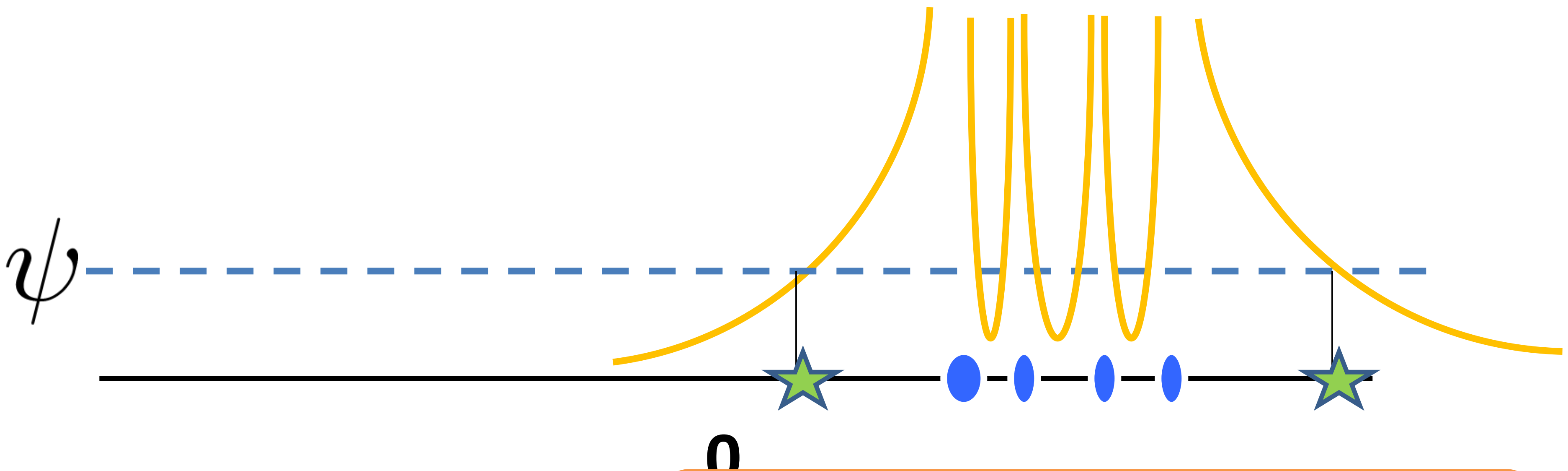


$$A_q = X_1 X_1^T + \ldots, X_q X_q^T$$

After **q** steps

$$\mathbf{E} s_{min}(A_k) \leq -n/\psi + q(1-\epsilon)$$
$$\mathbf{E} s_{max}(A_k) \geq n/\psi + q(1+\epsilon).$$

End of the Proof



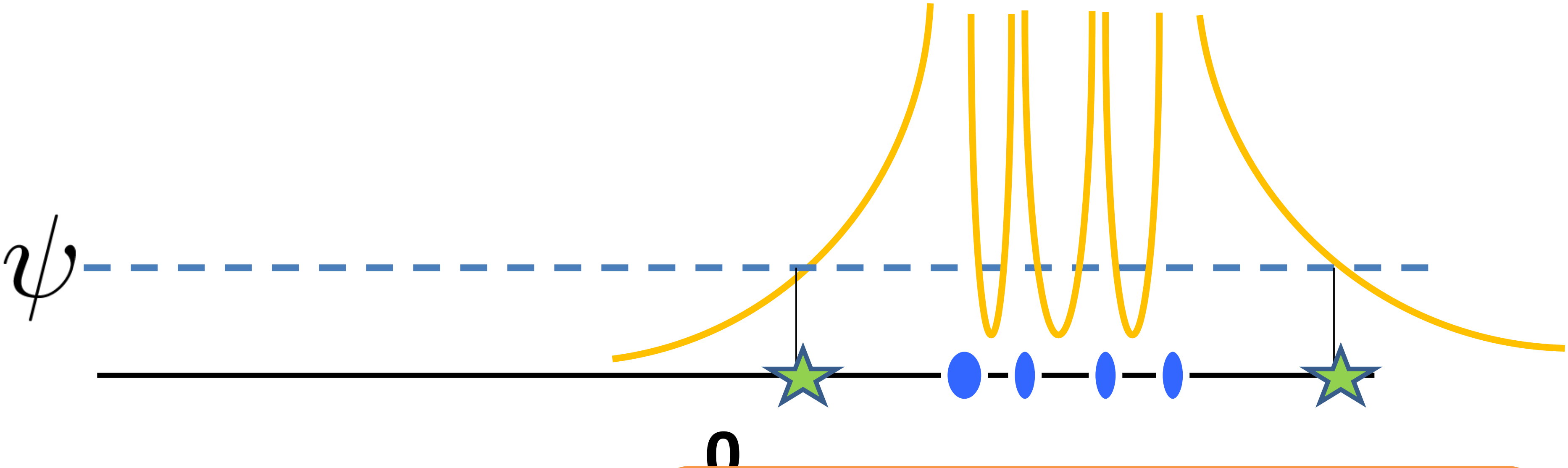
$$A_q = X_1$$

Set $q > \frac{n}{\psi\epsilon} = c(\epsilon)n.$

After **q** steps

$$\mathbf{E}s_{min}(A_k) \leq -n/\psi + q(1-\epsilon)$$
$$\mathbf{E}s_{max}(A_k) \geq n/\psi + q(1+\epsilon).$$

End of the Proof



$A_q = X_1$

Set $q > \frac{n}{\psi\epsilon} = c(\epsilon)n.$

After **q** steps

$$\mathbf{E}s_{min}(A_k) \leq -n/\psi + q(1-\epsilon)$$
$$\mathbf{E}s_{max}(A_k) \geq n/\psi + q(1+\epsilon).$$

$\square$

Open Questions

More delicate results? (fluctuations of top eigenvalue,…)

Preserving higher marginals

[Rudelson-Guedon'07, Vershynin'10]

Extension to higher rank matrices