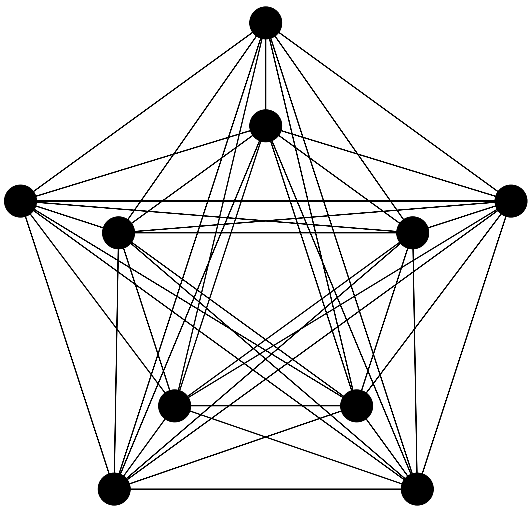
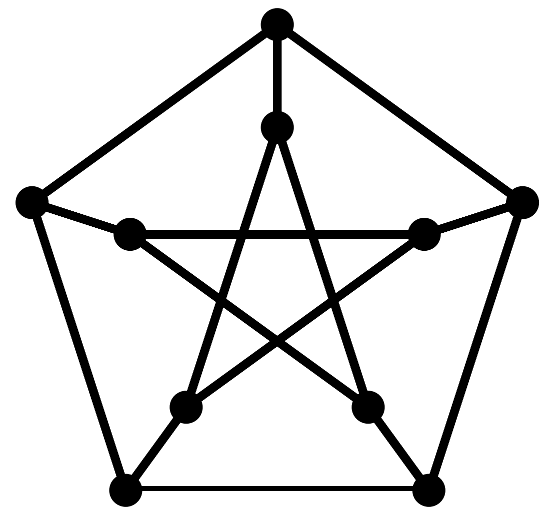
Sparsification of Graphs and Matrices





Daniel A. Spielman

Yale University

joint work with

Joshua Batson (MIT)

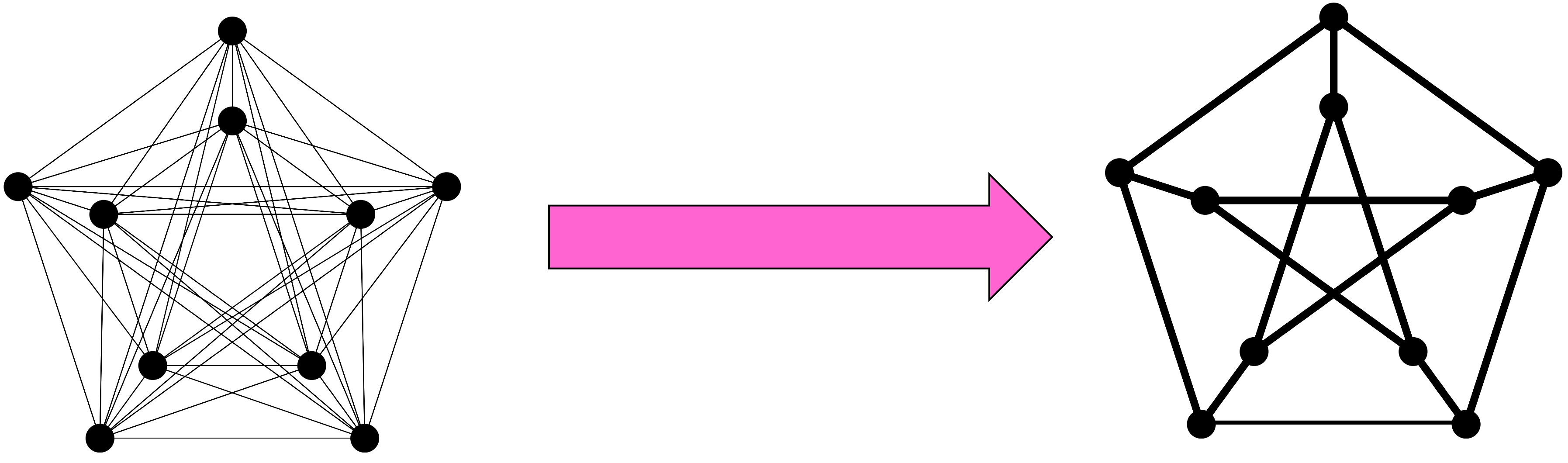
Nikhil Srivastava (MSR)

Shang-Hua Teng (USC)

HUJI, May 21, 2014

Objective of Sparsification:

Approximate any (weighted) graph by a sparse weighted graph.

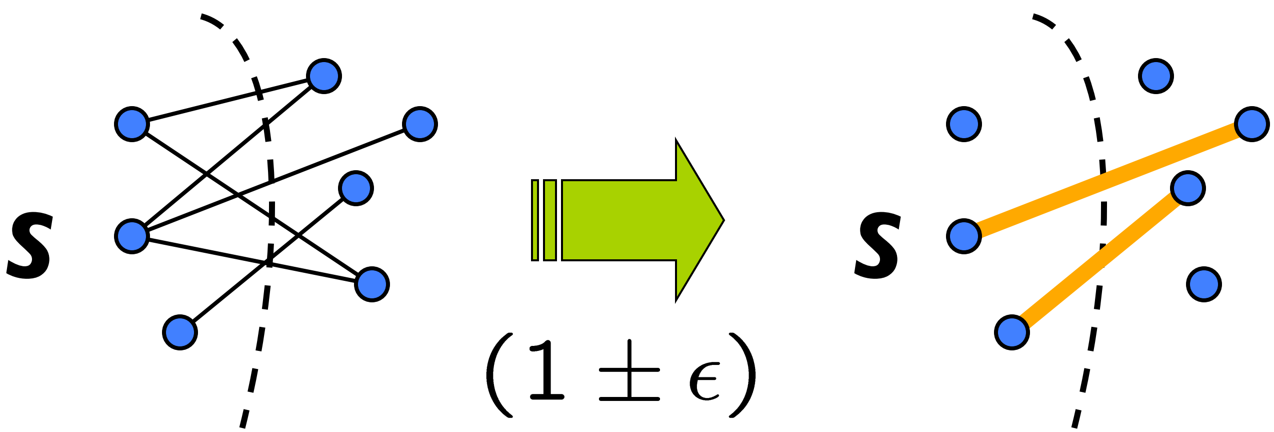


Objective of Sparsification:

Approximate any (weighted) graph by a sparse weighted graph.

Spanners - Preserve Distances [Chew '89]

Cut-Sparsifiers – preserve wt of edges leaving every set $S \subseteq V$ [Benczur-Karger '96]



Spectral Sparsification [S-Teng]

Approximate any (weighted) graph by a sparse weighted graph.

Graph ⟺ Laplacian

$$G = (V, E, w)$$

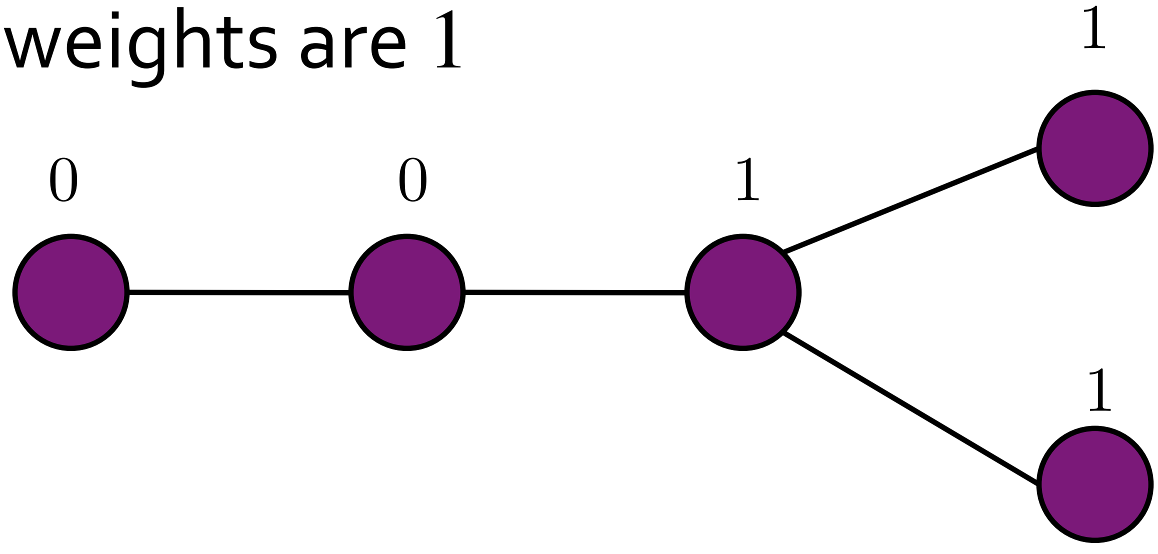
$$\sum_{(u,v)\in E} w_{u,v}(x(u) - x(v))^2$$

$$= x^T L_G x$$

Laplacian Quadratic Form, examples

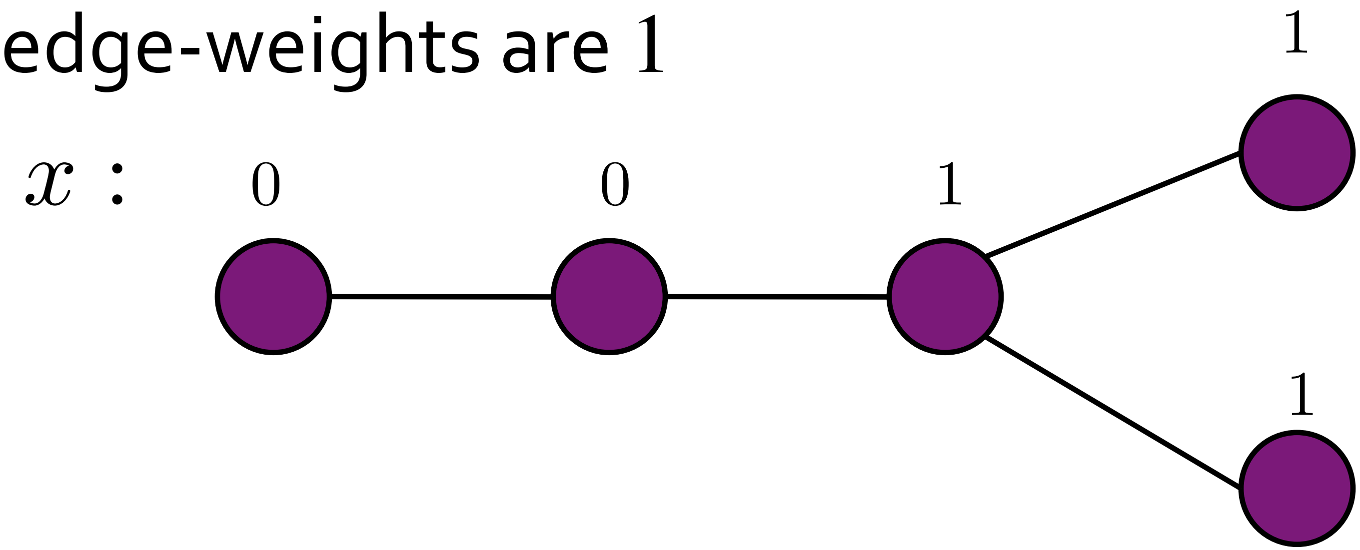
All edge-weights are 1

$x:$



Laplacian Quadratic Form, examples

All edge-weights are 1

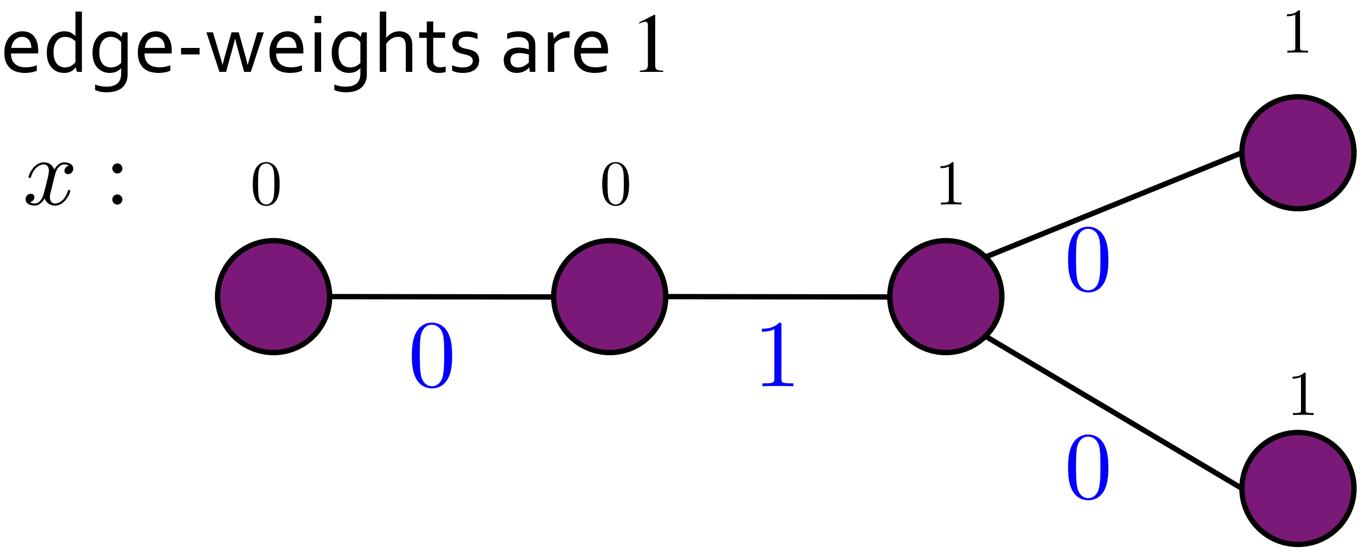


$$x^T L_G x = \sum_{(u,v)\in E} (x(u) - x(v))^2$$

= Sum of squares of differences across edges

Laplacian Quadratic Form, examples

All edge-weights are 1



$$x^T L_G x = \sum_{(u,v)\in E} (x(u) - x(v))^2$$

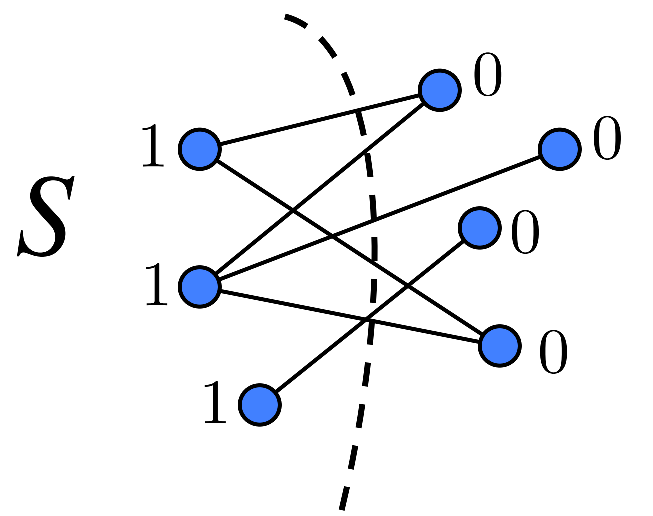
$=$ Sum of squares of differences across edges

$= 1$

Laplacian Quadratic Form, examples

When x is the characteristic vector of a set S, sum the weights of edges on the boundary of S

$$x^T L_G x = \sum_{\substack{(u,v)\in E \\ u\in S \\ v\notin S}} w_{u,v}$$

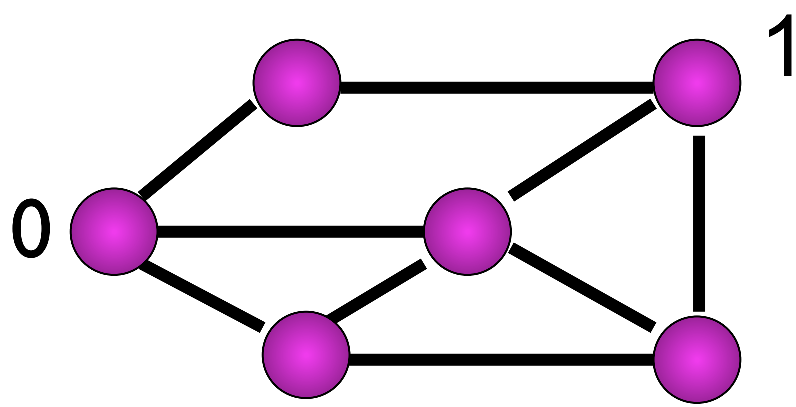


Learning on Graphs (Zhu-Ghahramani-Lafferty ’03)

Infer values of a function at all vertices from known values at a few vertices.

Minimize $$\sum_{(a,b)\in E} (x(a) - x(b))^2$$

Subject to known values

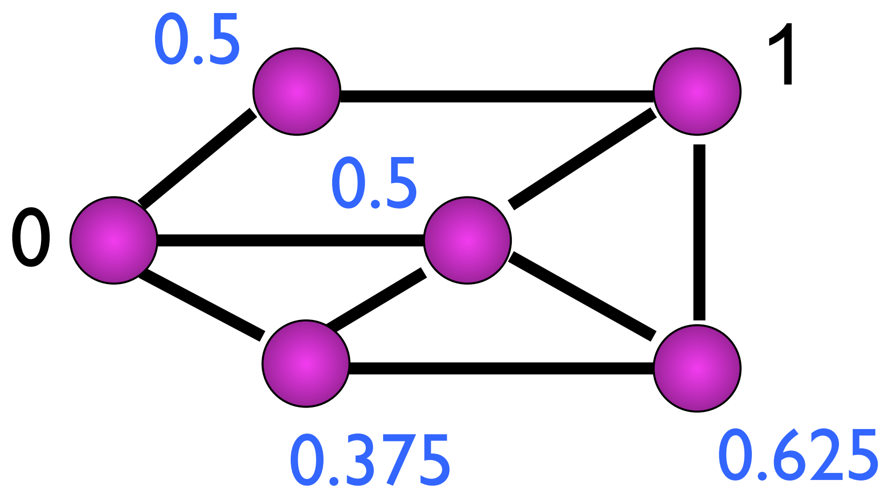


Learning on Graphs (Zhu-Ghahramani-Lafferty '03)

Infer values of a function at all vertices
from known values at a few vertices.

Minimize $$\sum_{(a,b)\in E} (x(a) - x(b))^2$$

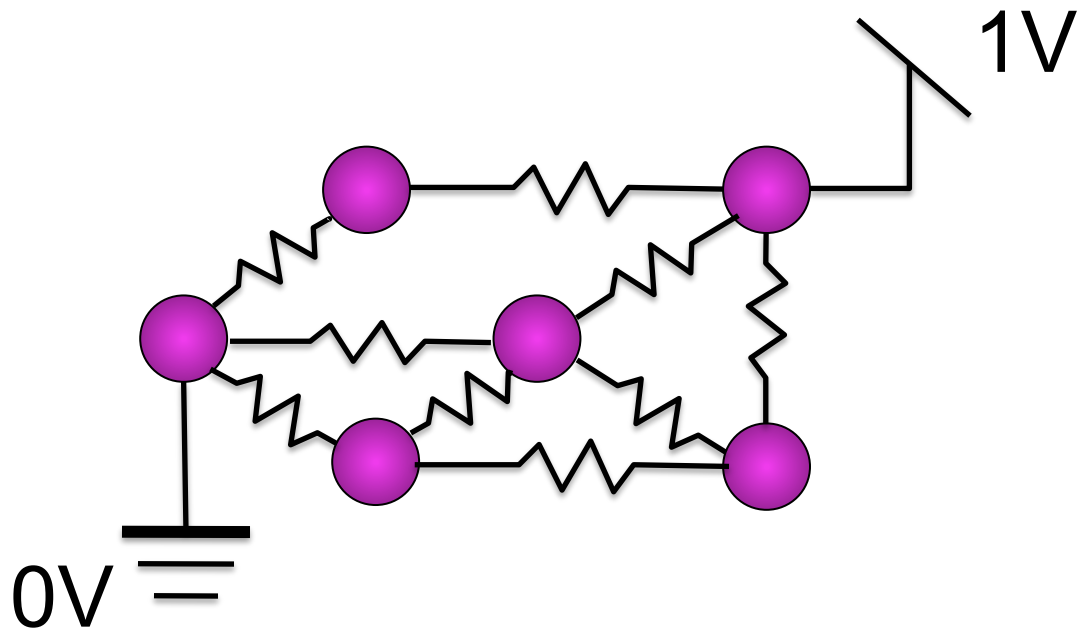
Subject to known values



Graphs as Resistor Networks

View edges as resistors connecting vertices

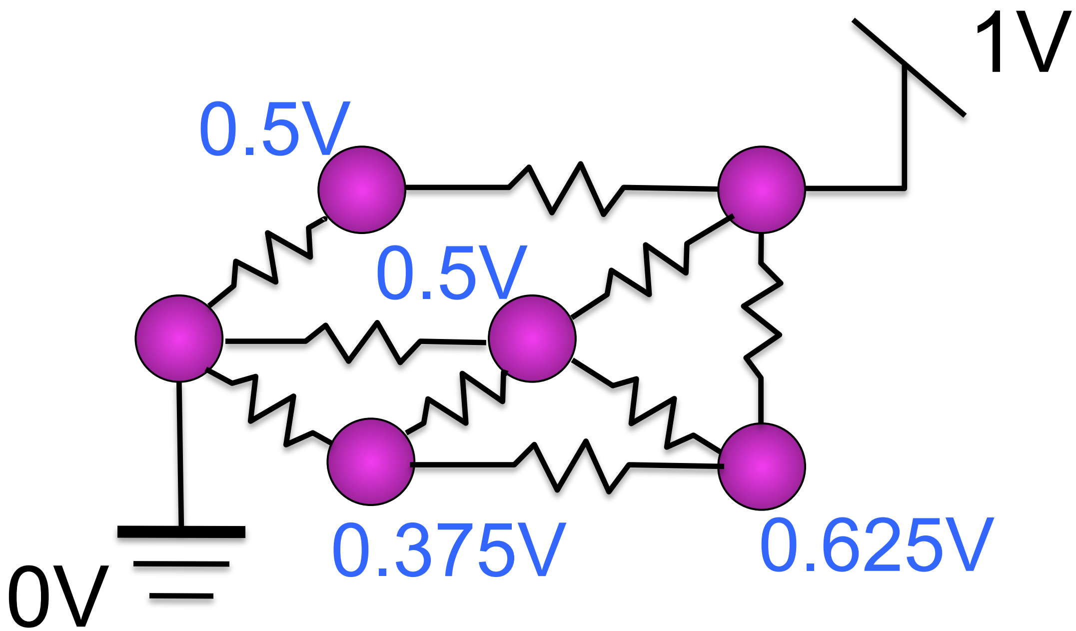
Apply voltages at some vertices.
Measure induced voltages and current flow.



Graphs as Resistor Networks

View edges as resistors connecting vertices

Apply voltages at some vertices.
Measure induced voltages and current flow.



Graphs as Resistor Networks

View edges as resistors connecting vertices

Apply voltages at some vertices.
Measure induced voltages and current flow.

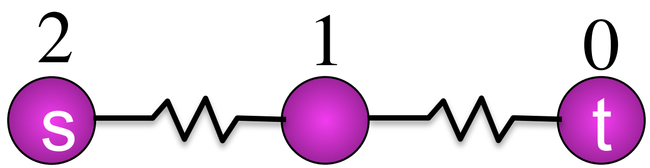
Induced voltages minimize

$$\sum_{(a,b)\in E} (v(a) - v(b))^2$$

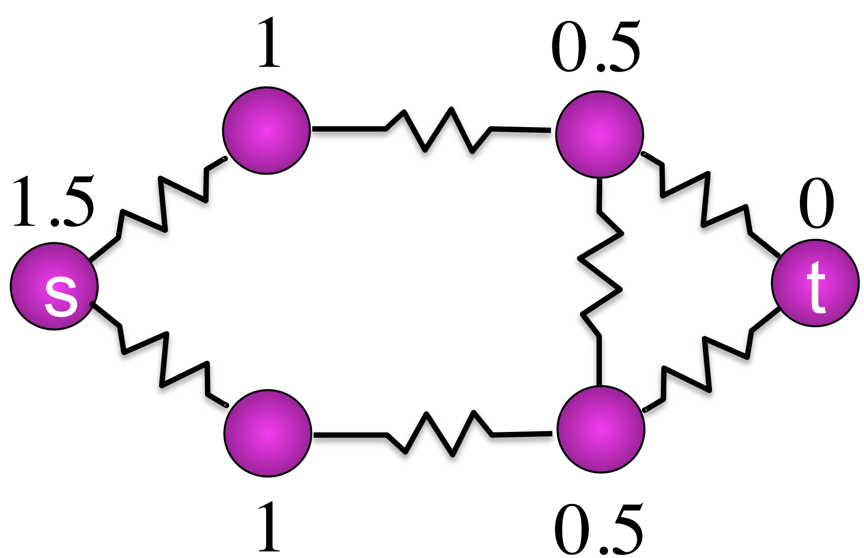
Subject to fixed voltages (by battery)

Graphs as Resistor Networks

Effective Resistance between s and t =
potential difference of unit flow



$$\mathrm{Reff}(s,t) = 2$$



$$\mathrm{Reff}(s,t) = 1.5$$

Laplacian Matrices

$$x^T L_G x = \sum_{(u,v)\in E} w_{u,v}(x(u) - x(v))^2$$

$$L_G = \sum_{(u,v)\in E} w_{u,v} L_{u,v}$$

E.g. $L_{1,2} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

$$= \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix}$$

Laplacian Matrices

$$x^T L_G x = \sum_{(u,v)\in E} w_{u,v}(x(u) - x(v))^2$$

$$L_G = \sum_{(u,v)\in E} w_{u,v} L_{u,v}$$

$$= \sum_{(u,v)\in E} w_{u,v}(b_{u,v} b_{u,v}^T)$$

Sum of outer products

where $\quad b_{u,v} = \delta_u - \delta_v$

Laplacian Matrices

$$x^T L_G x = \sum_{(u,v)\in E} w_{u,v}(x(u) - x(v))^2$$

$$L_G = \sum_{(u,v)\in E} w_{u,v} L_{u,v}$$

$$= \sum_{(u,v)\in E} w_{u,v}(b_{u,v} b_{u,v}^T)$$

Positive semidefinite

If connected, nullspace = Span($\mathbf{1}$)

Inequalities on Graphs and Matrices

For matrices M and $\widetilde{M}$

$$M \preccurlyeq \widetilde{M} \quad \text{if} \quad x^T M x \leq x^T \widetilde{M} x \quad \text{for all } x$$

For graphs $G = (V, E, w)$ and $H = (V, F, z)$

$$G \preccurlyeq H \quad \text{if} \quad L_G \preccurlyeq L_H$$

Inequalities on Graphs and Matrices

For matrices M and $\widetilde{M}$

$$M \preccurlyeq \widetilde{M} \quad \text{if} \quad x^T M x \leq x^T \widetilde{M} x \quad \text{for all } x$$

For graphs $G = (V, E, w)$ and $H = (V, F, z)$

$$G \preccurlyeq H \quad \text{if} \quad L_G \preccurlyeq L_H$$

$$G \preccurlyeq kH \quad \text{if} \quad L_G \preccurlyeq kL_H$$

Approximations of Graphs and Matrices

$$M \approx_\epsilon \widetilde{M} \quad \text{if} \quad \frac{1}{1+\epsilon} M \preccurlyeq \widetilde{M} \preccurlyeq (1+\epsilon) M$$

For graphs $G = (V, E, w)$ and $H = (V, F, z)$

$$G \approx_\epsilon H \quad \text{if} \quad L_G \approx_\epsilon L_H$$

Approximations of Graphs and Matrices

$$M \approx_\epsilon \widetilde{M} \qquad \text{if} \qquad \frac{1}{1+\epsilon} M \preccurlyeq \widetilde{M} \preccurlyeq (1+\epsilon) M$$

For graphs $G = (V, E, w)$ and $H = (V, F, z)$

$$G \approx_\epsilon H \qquad \text{if} \qquad L_G \approx_\epsilon L_H$$

That is, for all $x \in R^V$

$$\frac{1}{1+\epsilon} \leq \frac{\sum_{(u,v)\in F} z_{u,v}(x(u) - x(v))^2}{\sum_{(u,v)\in E} w_{u,v}(x(u) - x(v))^2} \leq 1+\epsilon$$

Implications of Approximation

$$G \approx_\epsilon H$$

Boundaries of sets are similar.

Effective resistances are similar.

L_H and L_G have similar eigenvalues

$$L_G^+ \approx_\epsilon L_H^+$$

Solutions to systems of linear equations are similar.

Spectral Sparsification [S-Teng]

For an input graph G with n vertices,

find a sparse graph H having $\tilde{O}(n)$ edges

so that $G \approx_\epsilon H$

Why?

Solving linear equations in Laplacian Matrices

key part of nearly-linear time algorithm

use for learning on graphs, maxflow, PDEs, …

Preserve Eigenvectors, Eigenvalues

and electrical properties

Generalize Expanders

Certifiable cut-sparsifiers

Approximations of Complete Graphs are Expanders

Expanders:

d-regular graphs on n vertices (n grows, d fixed)

every set of vertices has large boundary

random walks mix quickly

incredibly useful

Approximations of Complete Graphs are Expanders

Expanders:

d-regular graphs on n vertices (n grows, d fixed)

weak expanders: eigenvalues bounded from 0

strong expanders: all eigenvalues near d

Example: Approximating a Complete Graph

For G the complete graph on n verts.
all non-zero eigenvalues of L_G are n.

For $x \perp \mathbf{1}$, $\|x\| = 1$ $\qquad x^T L_G x = n$

Example: Approximating a Complete Graph

For G the complete graph on n verts.
all non-zero eigenvalues of L_G are n.

For $x \perp \mathbf{1}$, $\|x\| = 1$ $\quad x^T L_G x = n$

For H a d-regular strong expander,
all non-zero eigenvalues of L_H are close to d.

For $x \perp \mathbf{1}$, $\|x\| = 1$ $\quad x^T L_H x \sim d$

Example: Approximating a Complete Graph

For G the complete graph on n verts.
all non-zero eigenvalues of L_G are n.

For $x \perp \mathbf{1}$, $\|x\| = 1$ $\qquad x^T L_G x = n$

For H a d-regular strong expander,
all non-zero eigenvalues of L_H are close to d.

For $x \perp \mathbf{1}$, $\|x\| = 1$ $\qquad x^T L_H x \sim d$

$\frac{n}{d} H$ is a good approximation of G

Best Approximations of Complete Graphs

Ramanujan Expanders
[Margulis, Lubotzky-Phillips-Sarnak]

$$d - 2\sqrt{d-1} \leq \lambda(L_H) \leq d + 2\sqrt{d-1}$$

Best Approximations of Complete Graphs

Ramanujan Expanders

[Margulis, Lubotzky-Phillips-Sarnak]

$$d - 2\sqrt{d-1} \leq \lambda(L_H) \leq d + 2\sqrt{d-1}$$

Cannot do better if n grows while d is fixed

[Alon-Boppana]

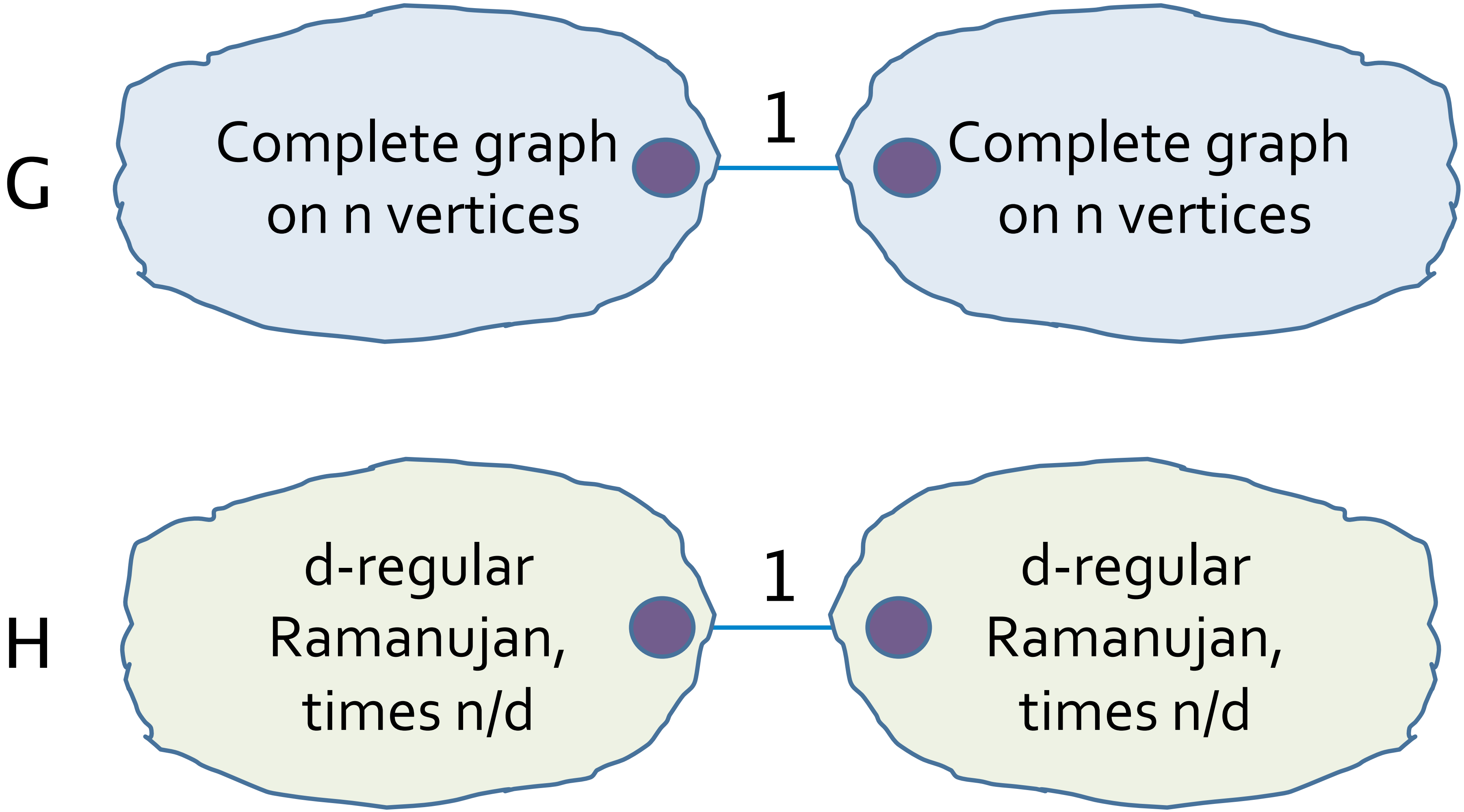
Best Approximations of Complete Graphs

Ramanujan Expanders
[Margulis, Lubotzky-Phillips-Sarnak]

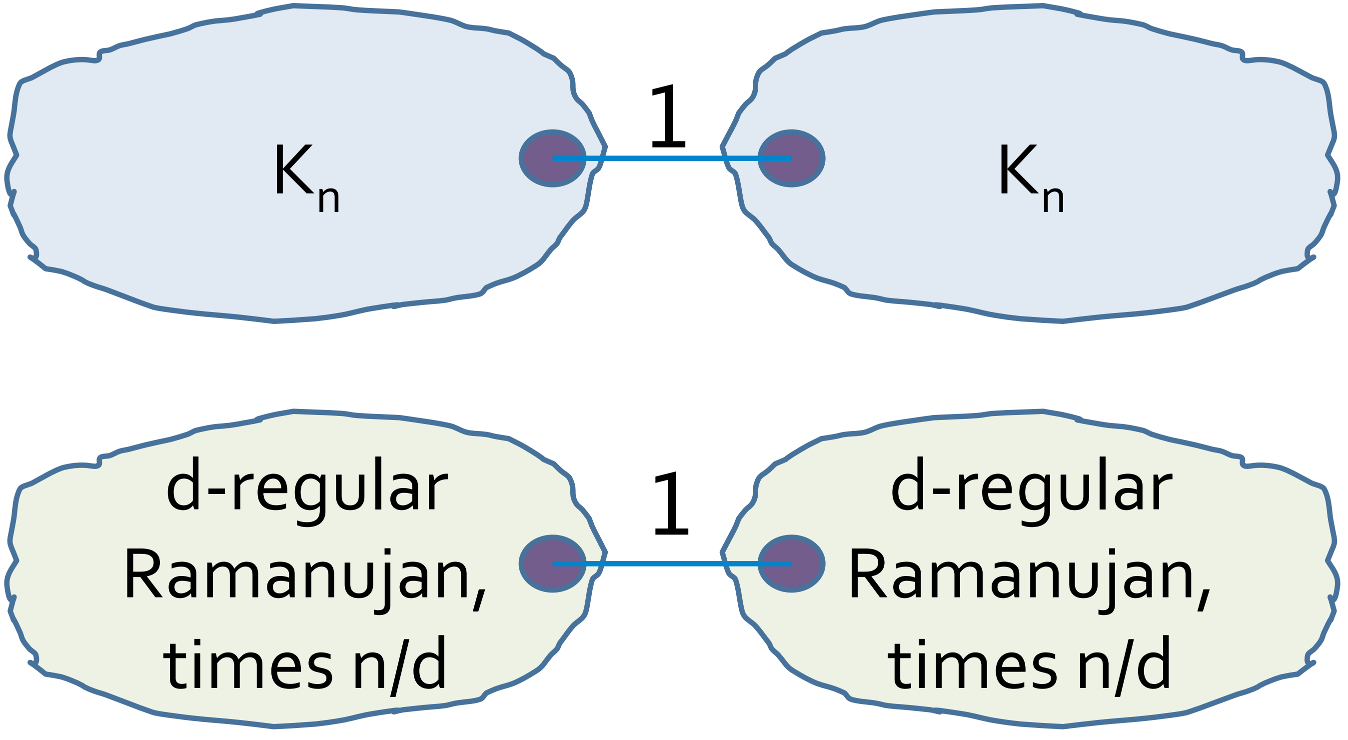
$$d - 2\sqrt{d-1} \leq \lambda(L_H) \leq d + 2\sqrt{d-1}$$

Can we approximate every graph this well?

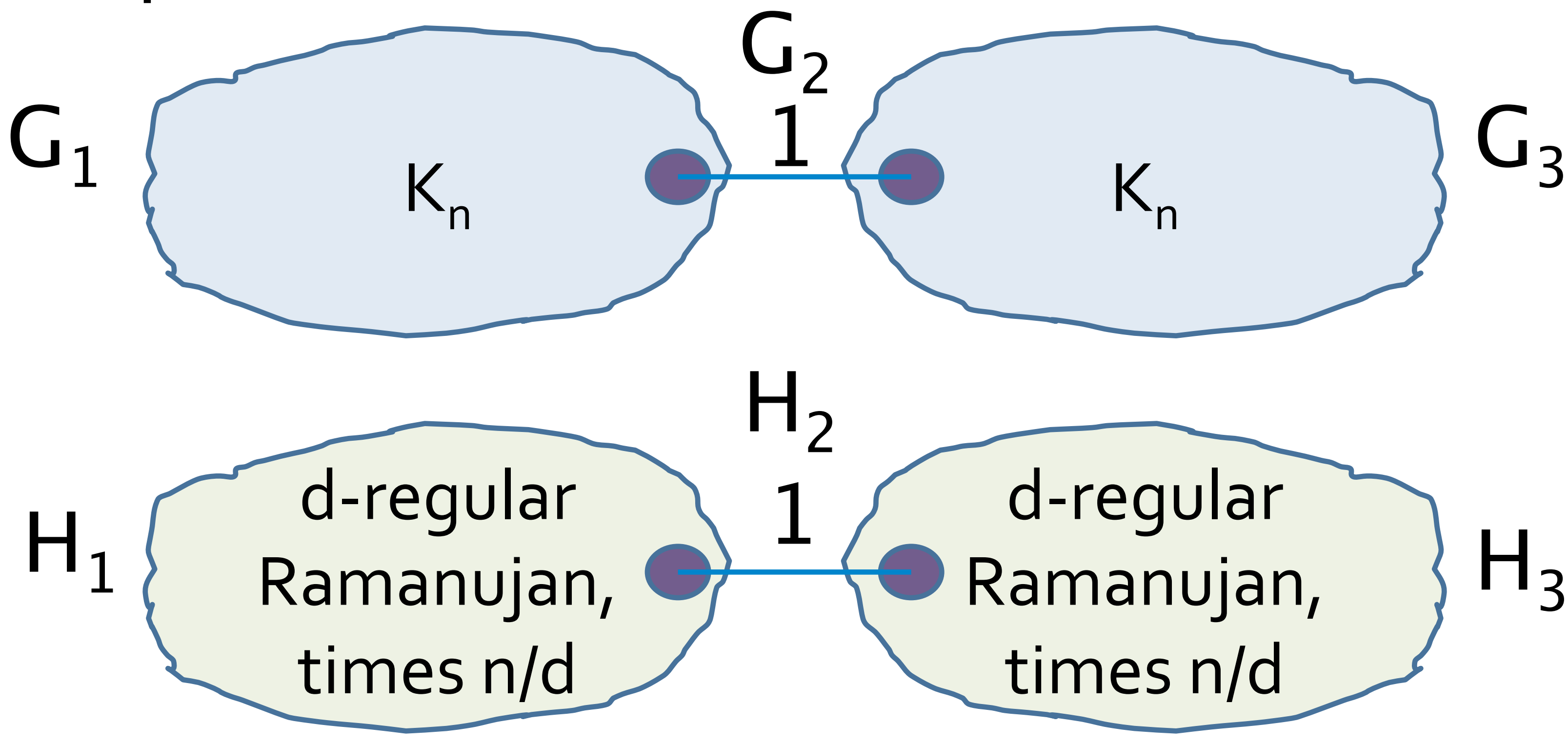
Example: Dumbbell



Example: Dumbbell



Example: Dumbbell



$$G = G_1 + G_2 + G_3$$

$$H = H_1 + H_2 + H_3$$

$$G_1 \preccurlyeq (1+\epsilon) H_1$$

$$G_2 = H_2$$

$$G_3 \preccurlyeq (1+\epsilon) H_3$$

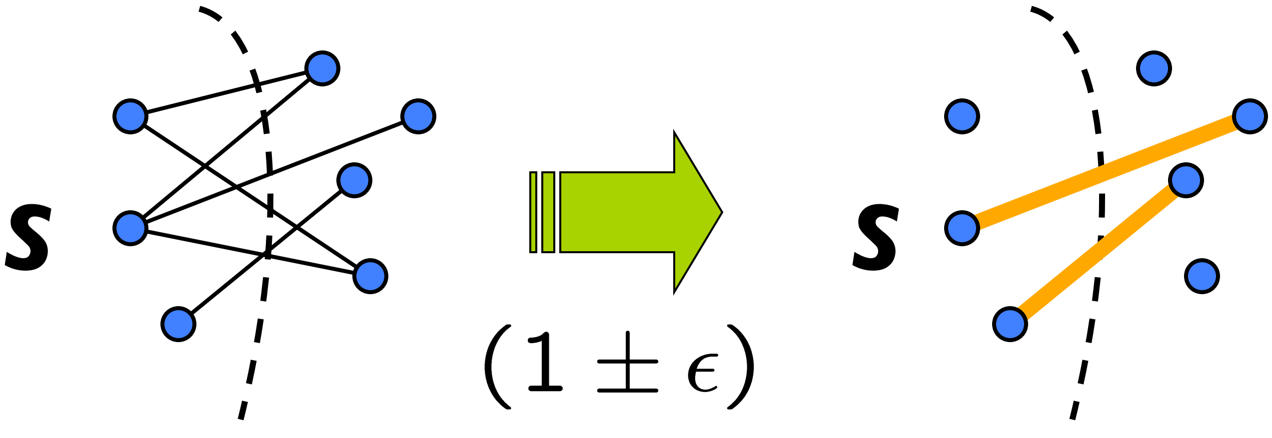
$$G \preccurlyeq (1+\epsilon) H$$

Cut-Sparsifiers [Benczur-Karger '96]

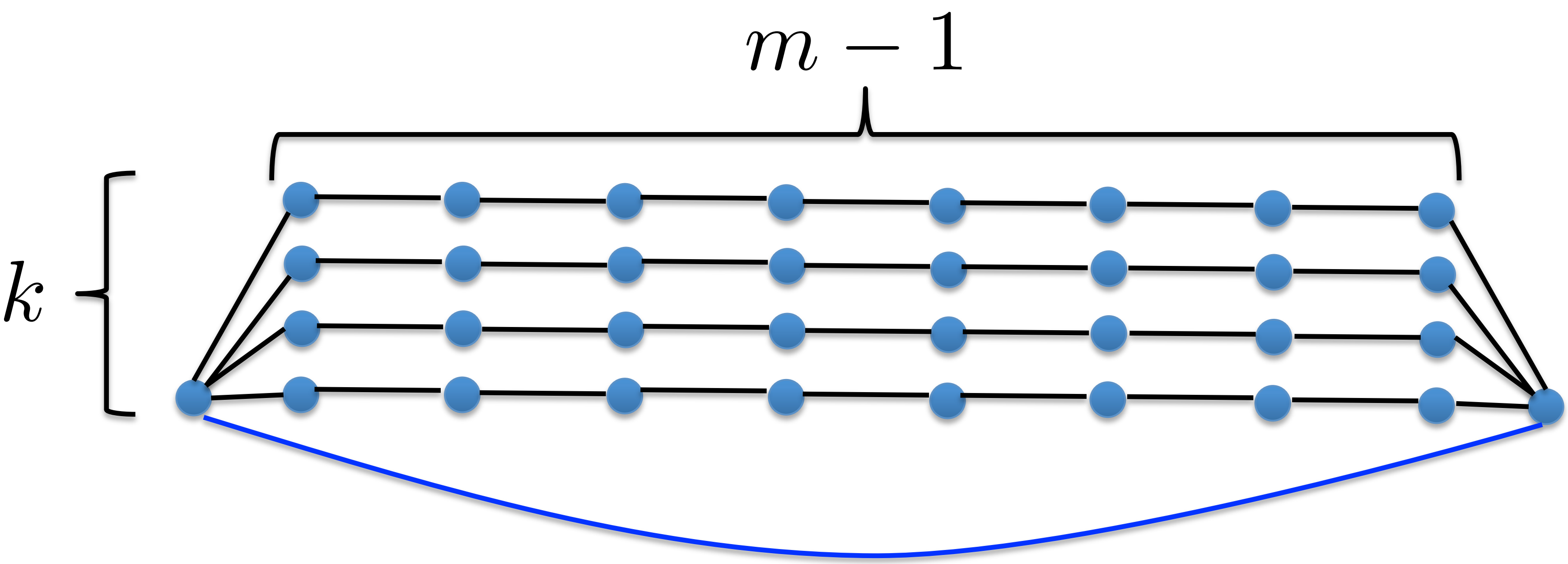
For every G, is an H with $O(n \log n/\epsilon^2)$ edges

for all $x \in \{0,1\}^n$

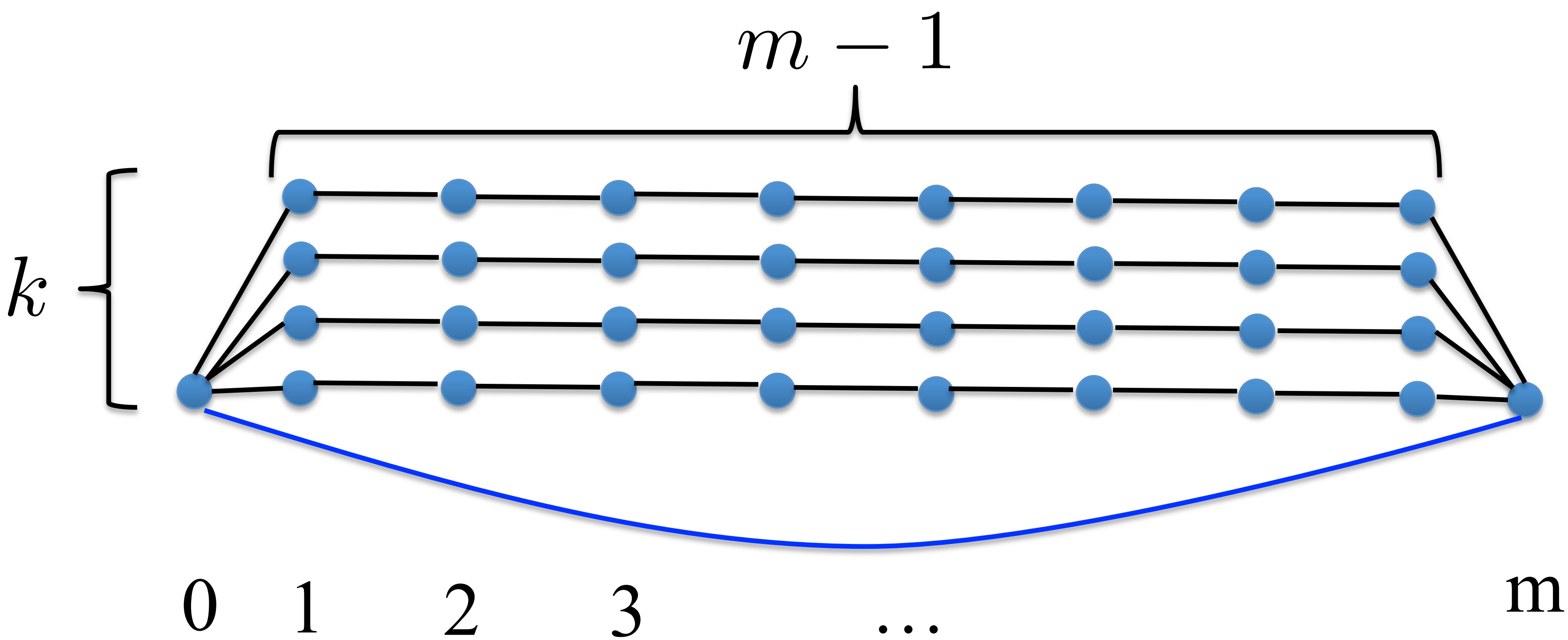
$$\frac{1}{1+\epsilon} \leq \frac{x^T L_H x}{x^T L_G x} \leq 1+\epsilon$$



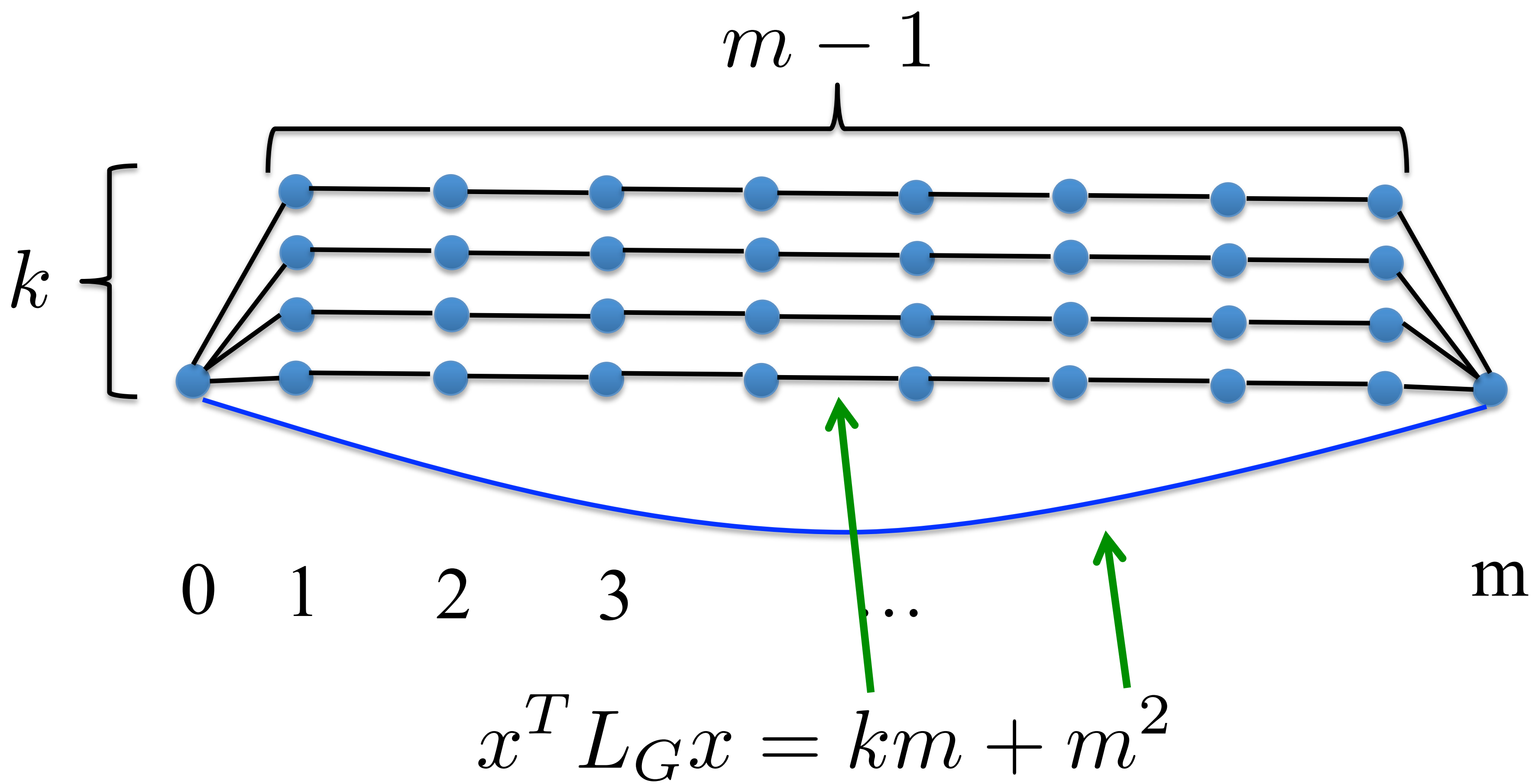
Cut-approximation is different



Cut-approximation is different



Cut-approximation is different



Need long edge if $k < m$

Main Theorems for Graphs

For every $G = (V, E, w)$, there is a $H = (V, F, z)$ s.t.

$$G \approx_{\epsilon} H \quad \text{and} \quad |F| \leq |V| (2 + \epsilon)^2 / \epsilon^2$$

Main Theorems for Graphs

For every $G = (V, E, w)$, there is a $H = (V, F, z)$ s.t.

$$G \approx_{\epsilon} H \quad \text{and} \quad |F| \leq |V| \, (2+\epsilon)^2 / \epsilon^2$$

Within a factor of 2 of the Ramanujan bound

Main Theorems for Graphs

For every $G = (V, E, w)$, there is a $H = (V, F, z)$ s.t.

$$G \approx_\epsilon H \qquad \text{and} \qquad |F| \leq |V| \, (2+\epsilon)^2/\epsilon^2$$

By careful random sampling, get

$$|F| \leq O\left(|V| \log |V| \, / \epsilon^2\right)$$

(S-Srivastava 08)

In time $O(|E| \log^2 |V| \log(1/\epsilon))$

(Koutis-Levin-Peng '12)

Sparsification by Random Sampling

Assign a probability $p_{u,v}$ to each edge *(u,v)*

Include edge *(u,v)* in H with probability $p_{u,v.}$

If include edge *(u,v)*, give it weight $w_{u,v}/p_{u,v}$

$$\mathbb{E}\left[\ L_H\ \right] = \sum_{(u,v)\in E} p_{u,v}(w_{u,v}/p_{u,v})L_{u,v} = L_G$$

Sparsification by Random Sampling

Choose $p_{u,v}$ to be $w_{u,v}$ times the effective resistance between u and v.

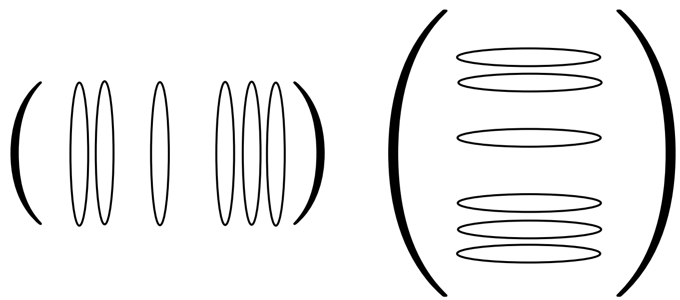
Low resistance between u and v means there are many alternate routes for current to flow and that the edge is not critical.

Proof by random matrix concentration bounds (Rudelson, Ahlswede-Winter, Tropp, etc.)

Matrix Sparsification

$$(M) = (\ B\ )\left(B^T\right)$$

$$(\widetilde{M}) =$$

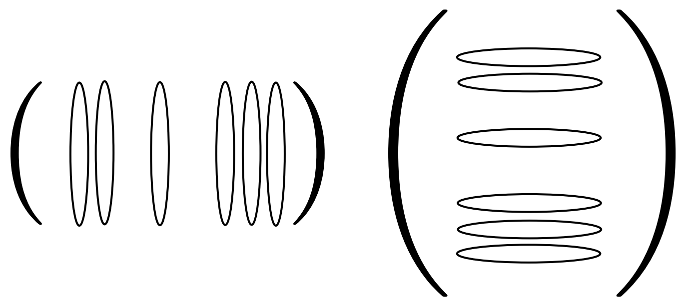


$$\frac{1}{(1+\epsilon)} M \preccurlyeq \widetilde{M} \preccurlyeq (1+\epsilon) M$$

Matrix Sparsification

$$(M) = (\ B\ )\left(B^T\right) = \sum_e b_e b_e^T$$

$$(\widetilde{M}) =$$



$$= \sum_e s_e b_e b_e^T$$

most $s_e = 0$

$$\frac{1}{(1+\epsilon)} M \preccurlyeq \widetilde{M} \preccurlyeq (1+\epsilon) M$$

Main Theorem (Batson-S-Srivastava)

For $M = \sum_e b_e b_e^T$, there exist s_e so that for

$$\widetilde{M} = \sum_e s_e b_e b_e^T$$

$$M \approx_\epsilon \widetilde{M}$$

and

at most $n(2+\epsilon)^2/\epsilon^2$ s_e are non-zero

Simplification of Matrix Sparsification

$$\frac{1}{(1+\epsilon)} M \preccurlyeq \widetilde{M} \preccurlyeq (1+\epsilon) M$$

is equivalent to

$$\frac{1}{(1+\epsilon)} I \preccurlyeq M^{-1/2} \widetilde{M} M^{-1/2} \preccurlyeq (1+\epsilon) I$$

Simplification of Matrix Sparsification

$$\frac{1}{(1+\epsilon)} I \preccurlyeq M^{-1/2} \widetilde{M} M^{-1/2} \preccurlyeq (1+\epsilon) I$$

Set $v_e = M^{-1/2} b_e$

$$\sum_e v_e v_e^T = I$$

"Decomposition of the identity"

$$\sum_e \langle u, v_e \rangle^2 = \|u\|^2$$

Simplification of Matrix Sparsification

$$\frac{1}{(1+\epsilon)} I \preccurlyeq M^{-1/2} \widetilde{M} M^{-1/2} \preccurlyeq (1+\epsilon) I$$

Set $v_e = M^{-1/2} b_e$

$$\sum_e v_e v_e^T = I$$

We need

$$\sum_e s_e v_e v_e^T \approx_\epsilon I$$

Simplification of Matrix Sparsification

$$\frac{1}{(1+\epsilon)} I \preccurlyeq M^{-1/2} \widetilde{M} M^{-1/2} \preccurlyeq (1+\epsilon) I$$

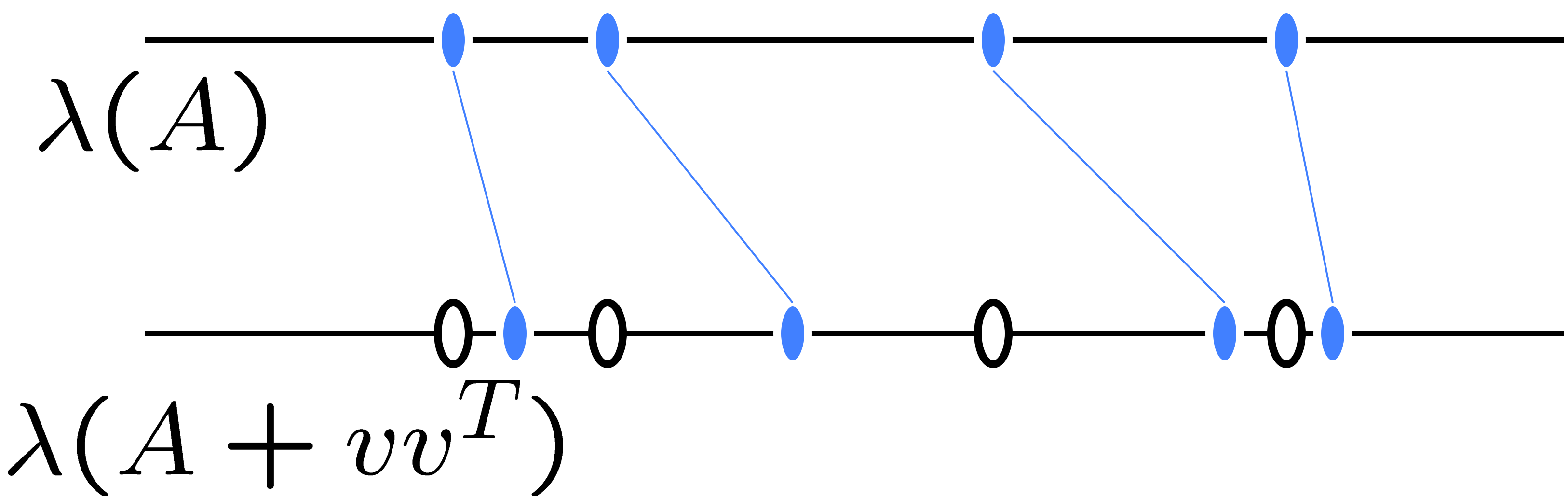
Set $v_e = M^{-1/2} b_e$ $\qquad \sum_e v_e v_e^T = I$

We need $\qquad \sum_e s_e v_e v_e^T \approx_\epsilon I$

Random sampling sets $p_e = \|v_e\|^2$

What happens when we add a vector?

Interlacing



More precisely

Characteristic Polynomial:

$$p_A(x) = \det(xI - A)$$

More precisely

Characteristic Polynomial:

$$p_A(x) = \det(xI - A)$$

Rank-one update:

$$p_{A+vv^T} = \left(1 + \sum_i \frac{<v,u_i>^2}{\lambda_i - x}\right) p_A$$

Where $Au_i = \lambda_i u_i$

More precisely

Characteristic Polynomial:

$$p_A(x) = \det(xI - A)$$

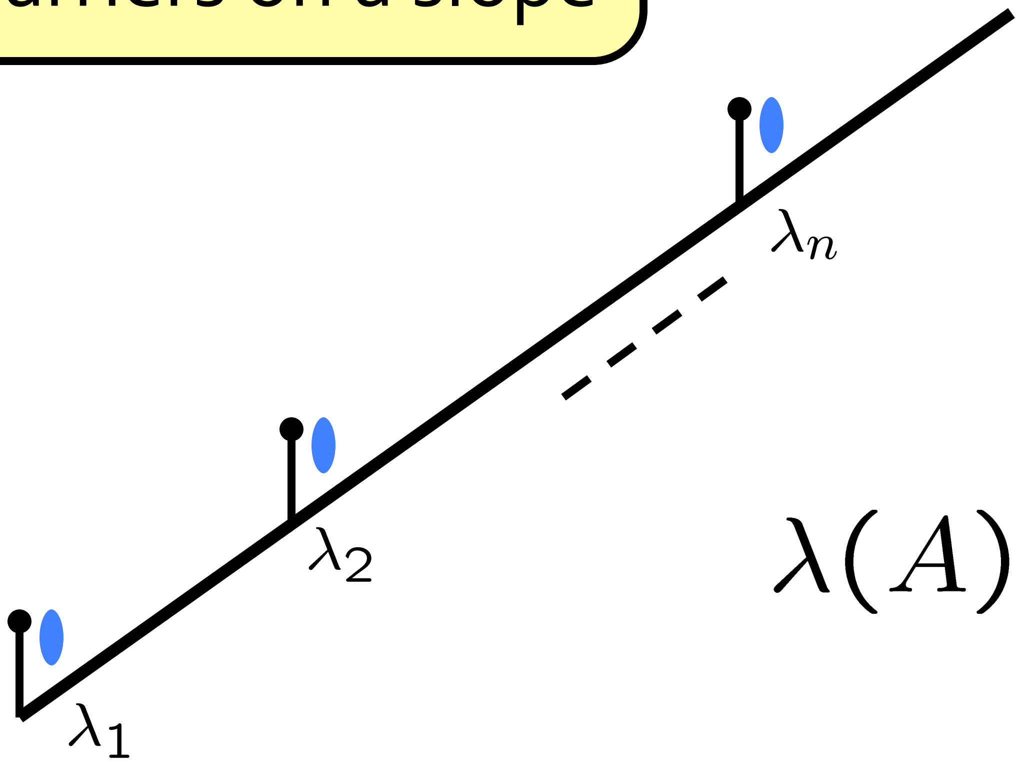
Rank-one update:

$$p_{A+vv^T} = \left(1 + \sum_i \frac{<v,u_i>^2}{\lambda_i - x}\right) p_A$$

$\lambda(A + vv^T)$ are zeros of

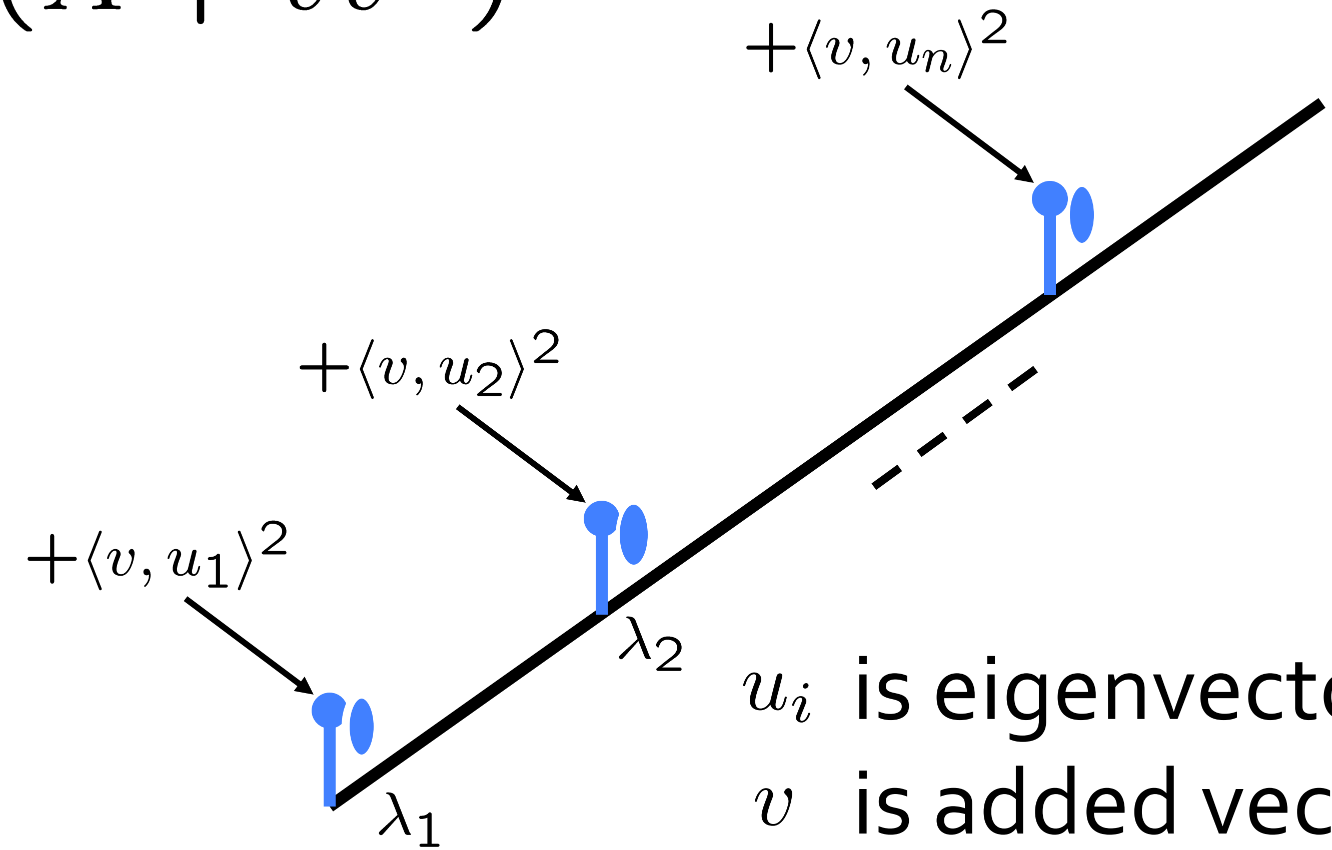
Physical model of interlacing

λ_i = positive unit charges resting at barriers on a slope



Physical model of interlacing

$$\lambda(A + vv^T)$$

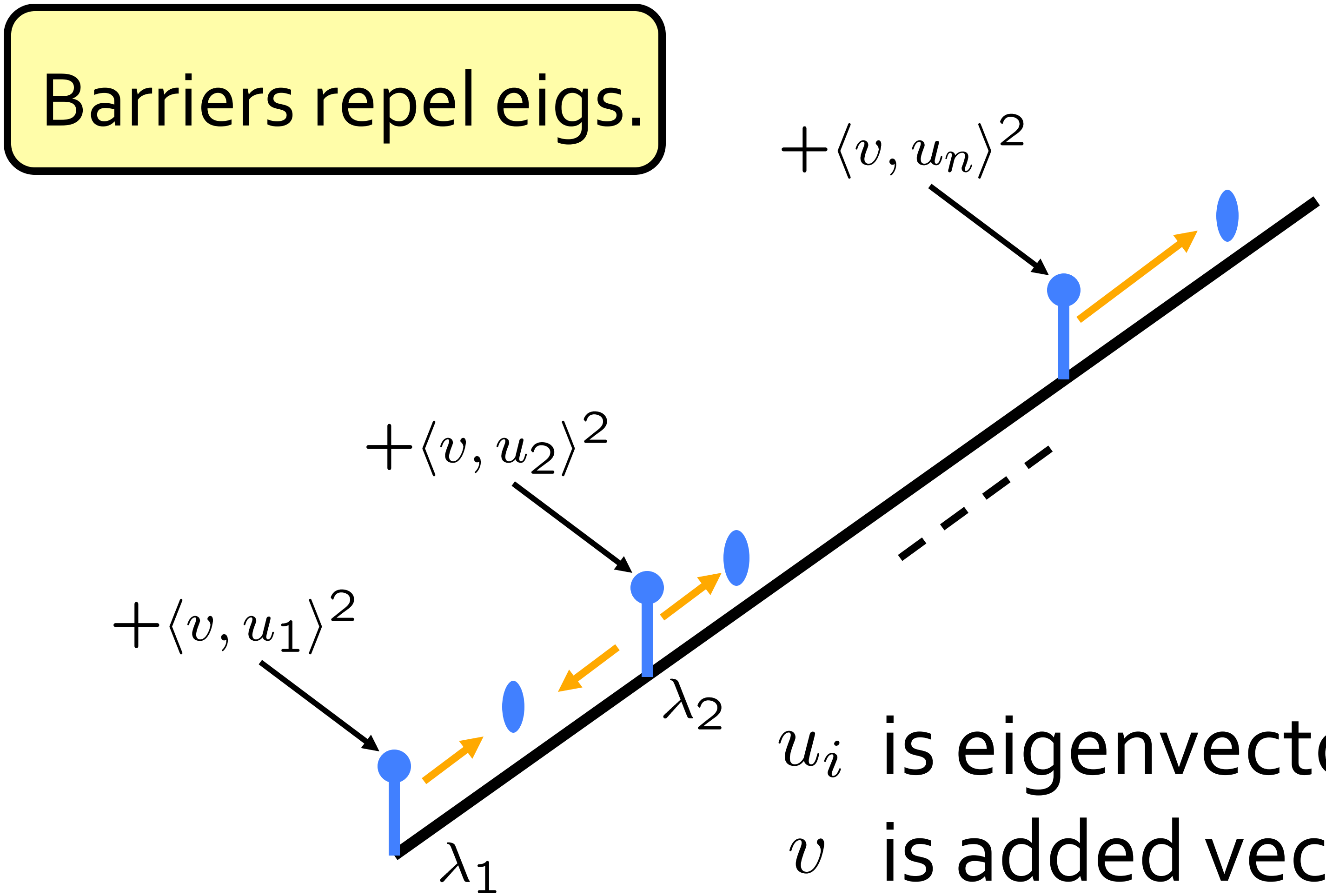


u_i is eigenvector

v is added vector

$\langle v, u_i \rangle^2$ charge on barrier

Physical model of interlacing

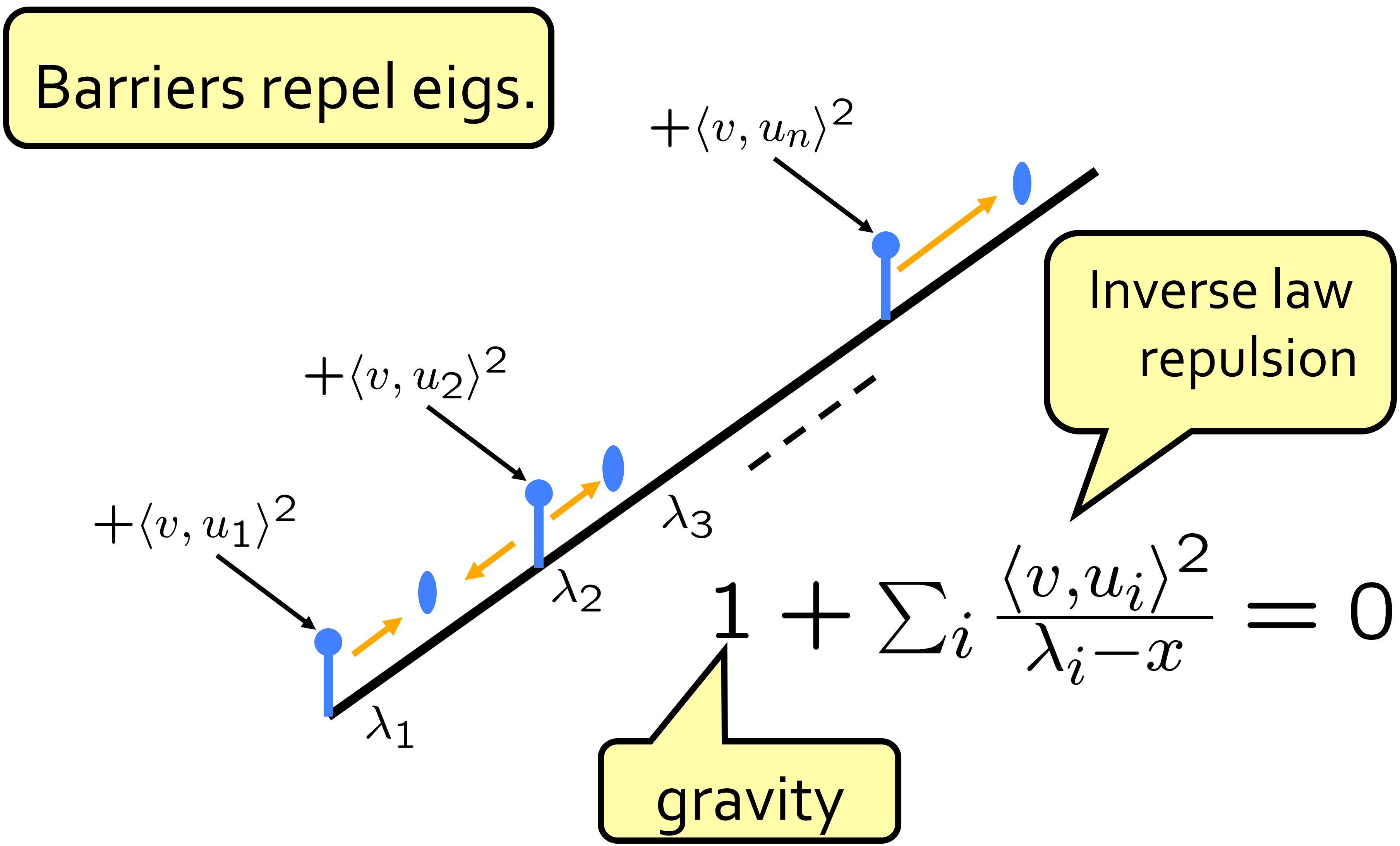


u_i is eigenvector

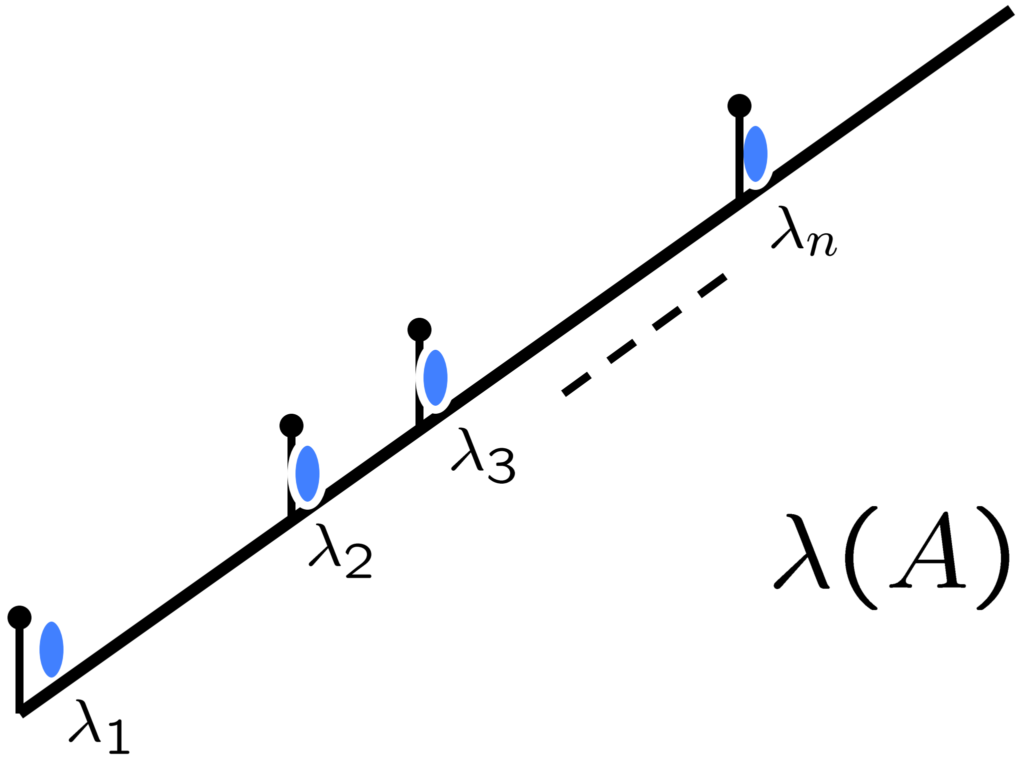
v is added vector

$\langle v, u_i \rangle^2$ charge on barrier

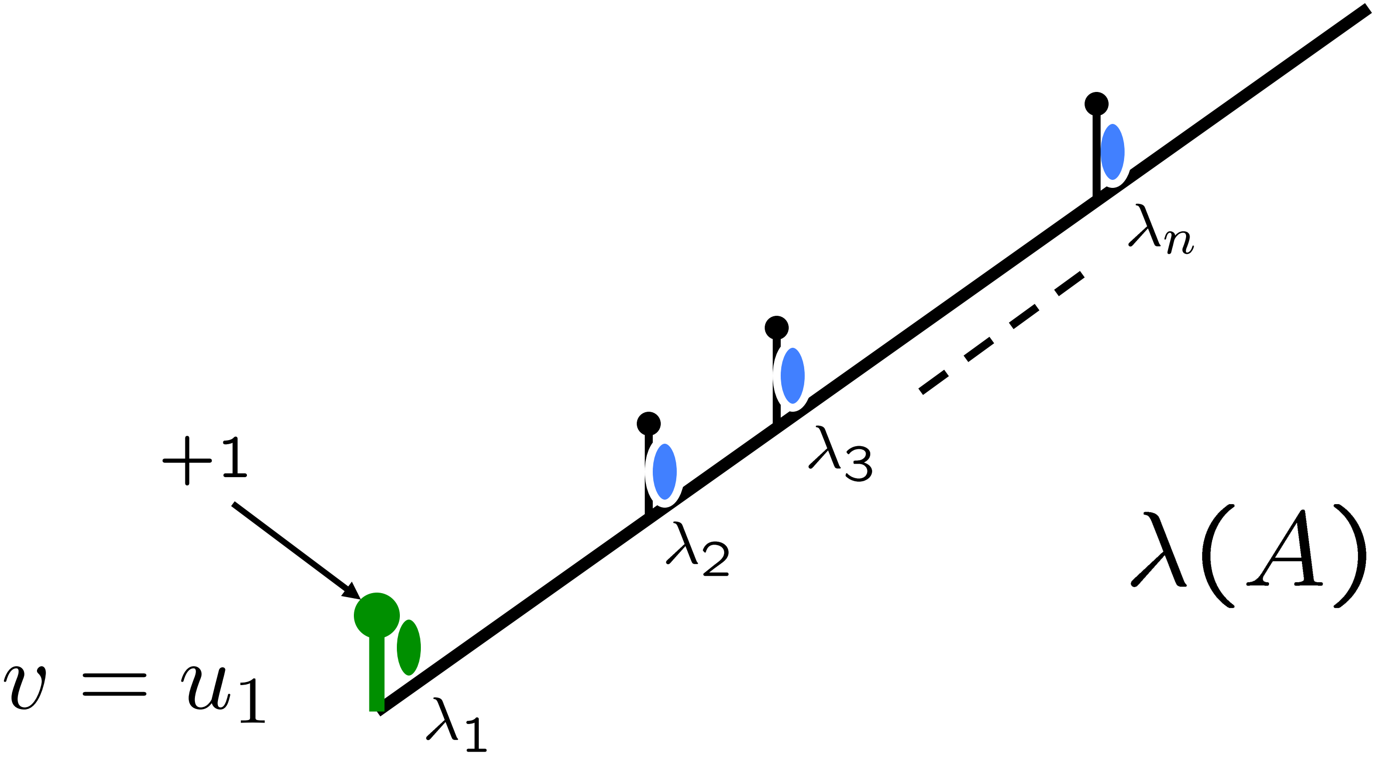
Physical model of interlacing



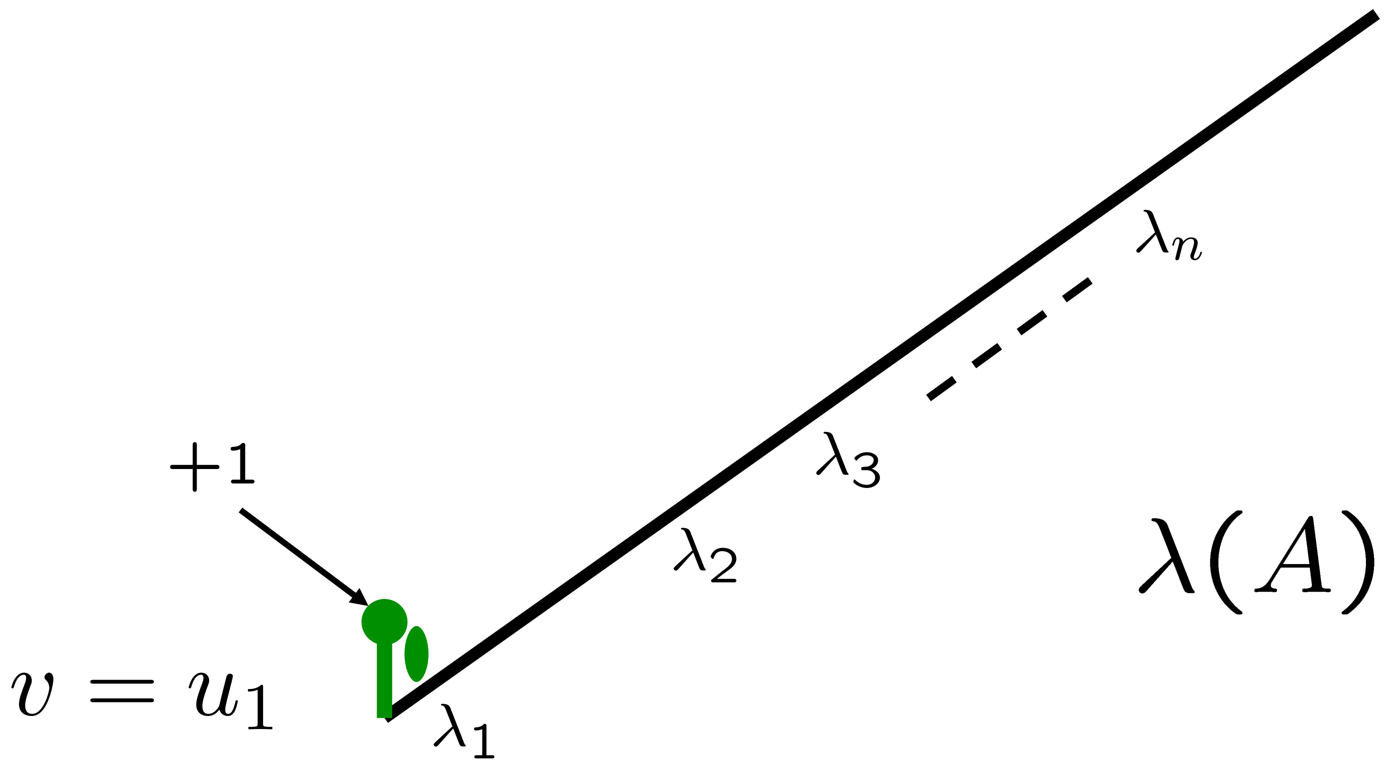
Examples



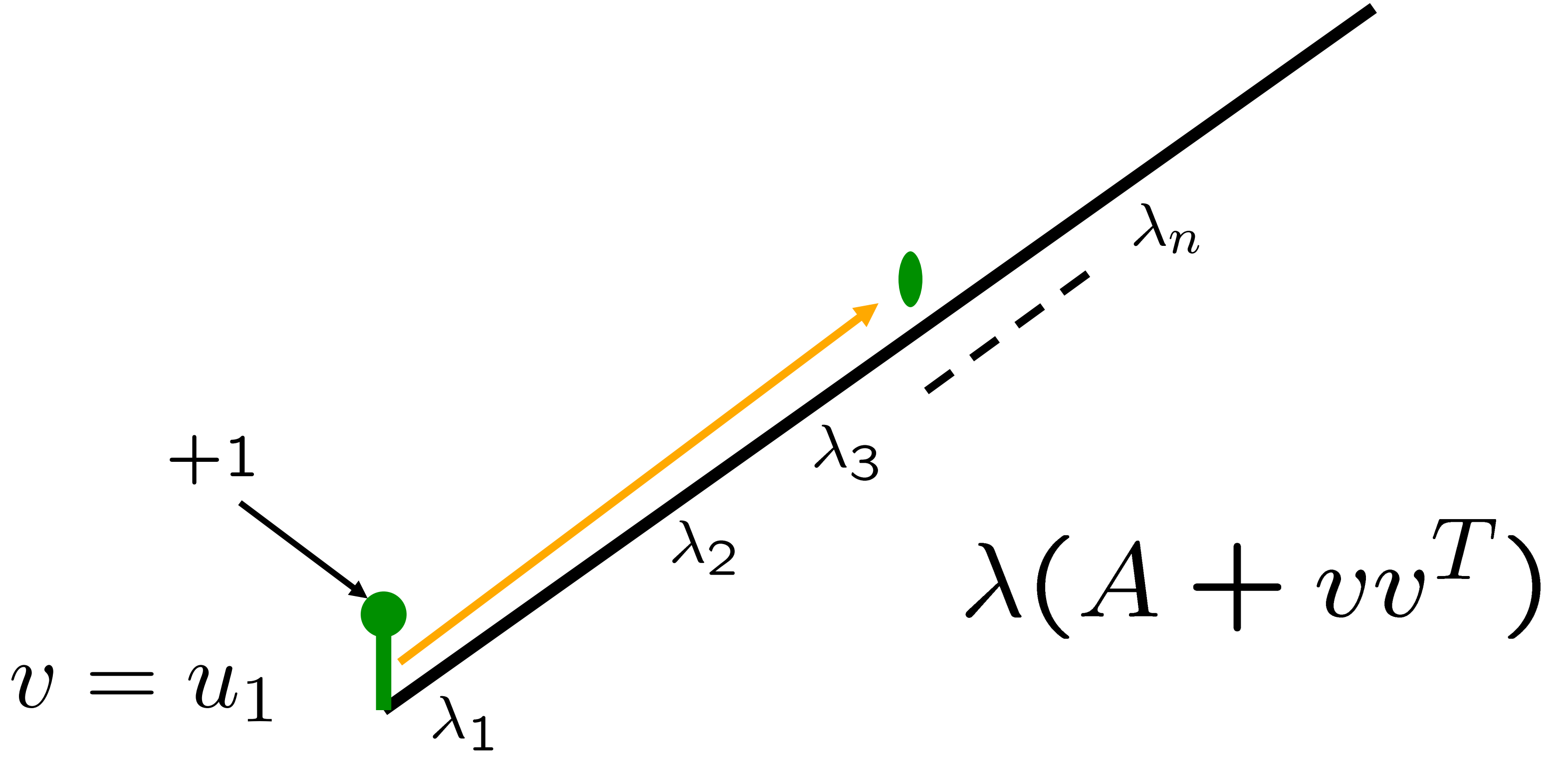
Ex1: All weight on u_1



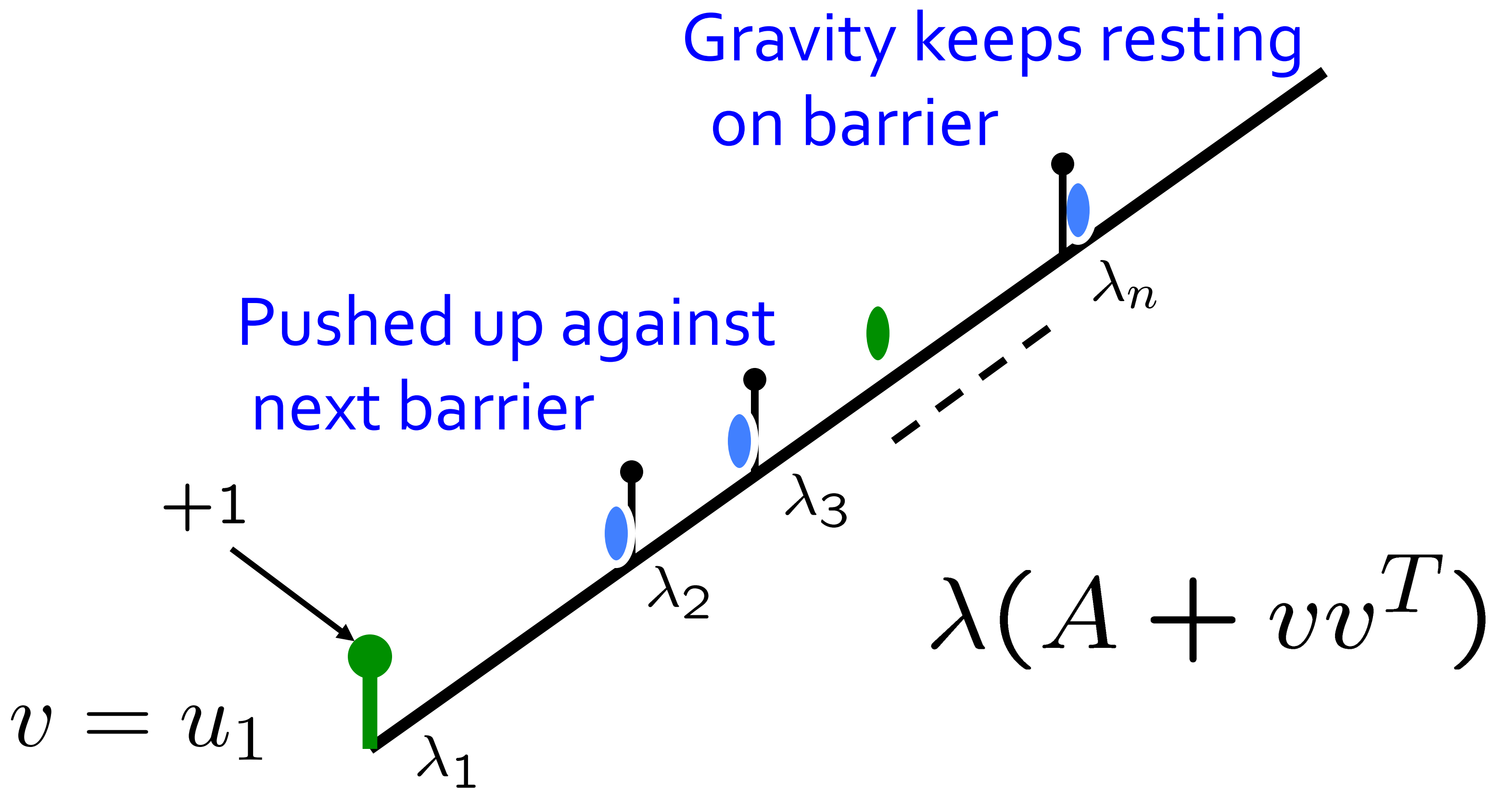
Ex1: All weight on u_1



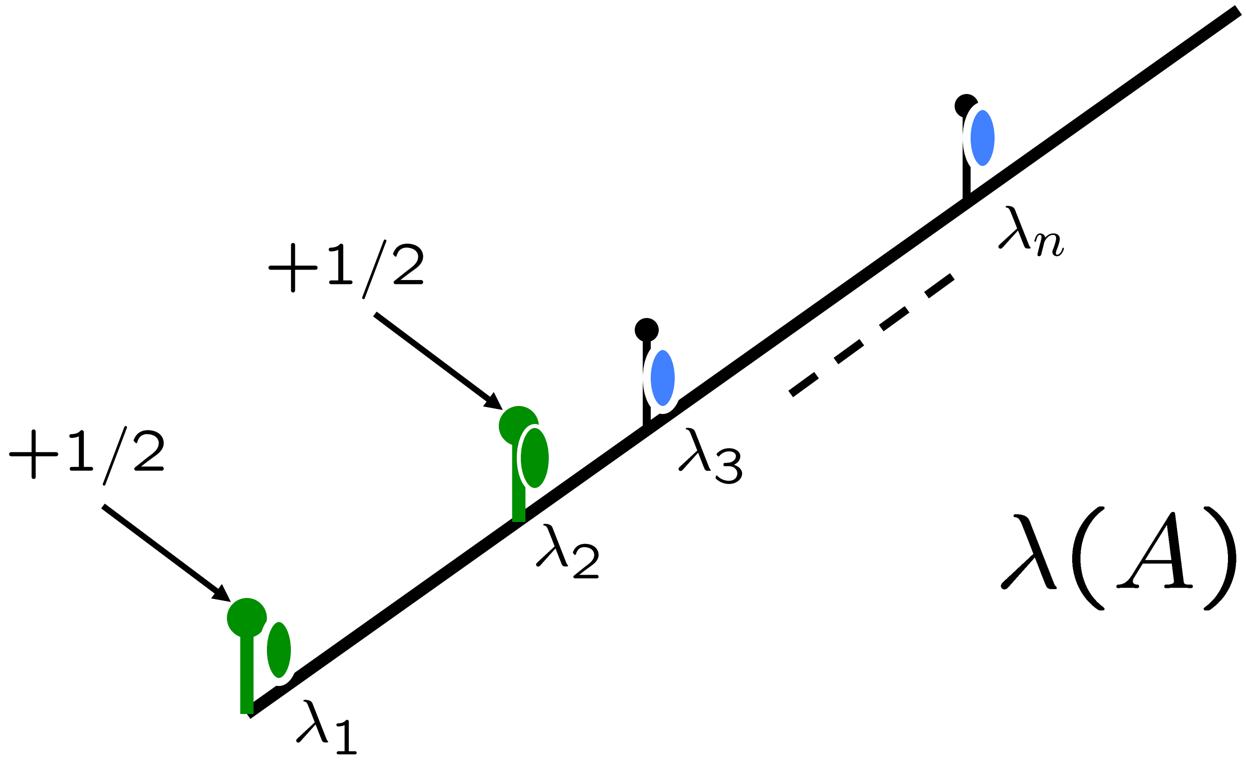
Ex1: All weight on u_1



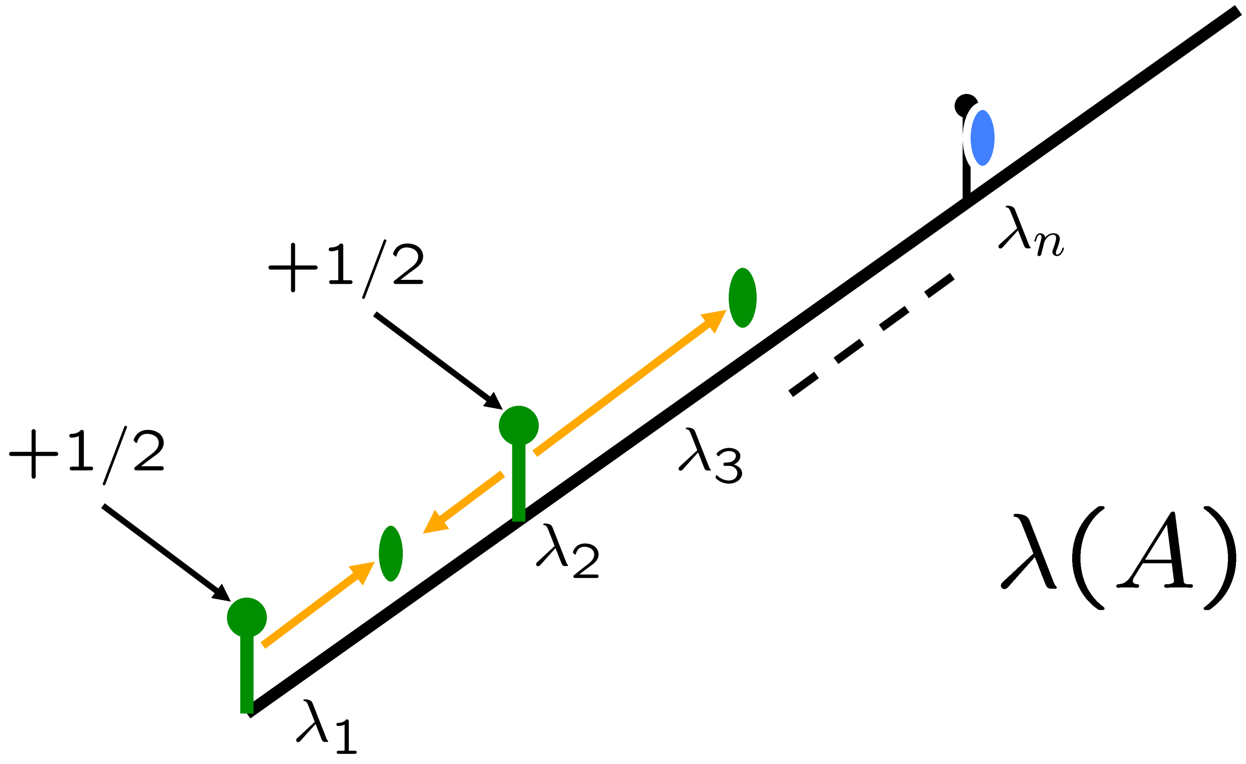
Ex1: All weight on u_1



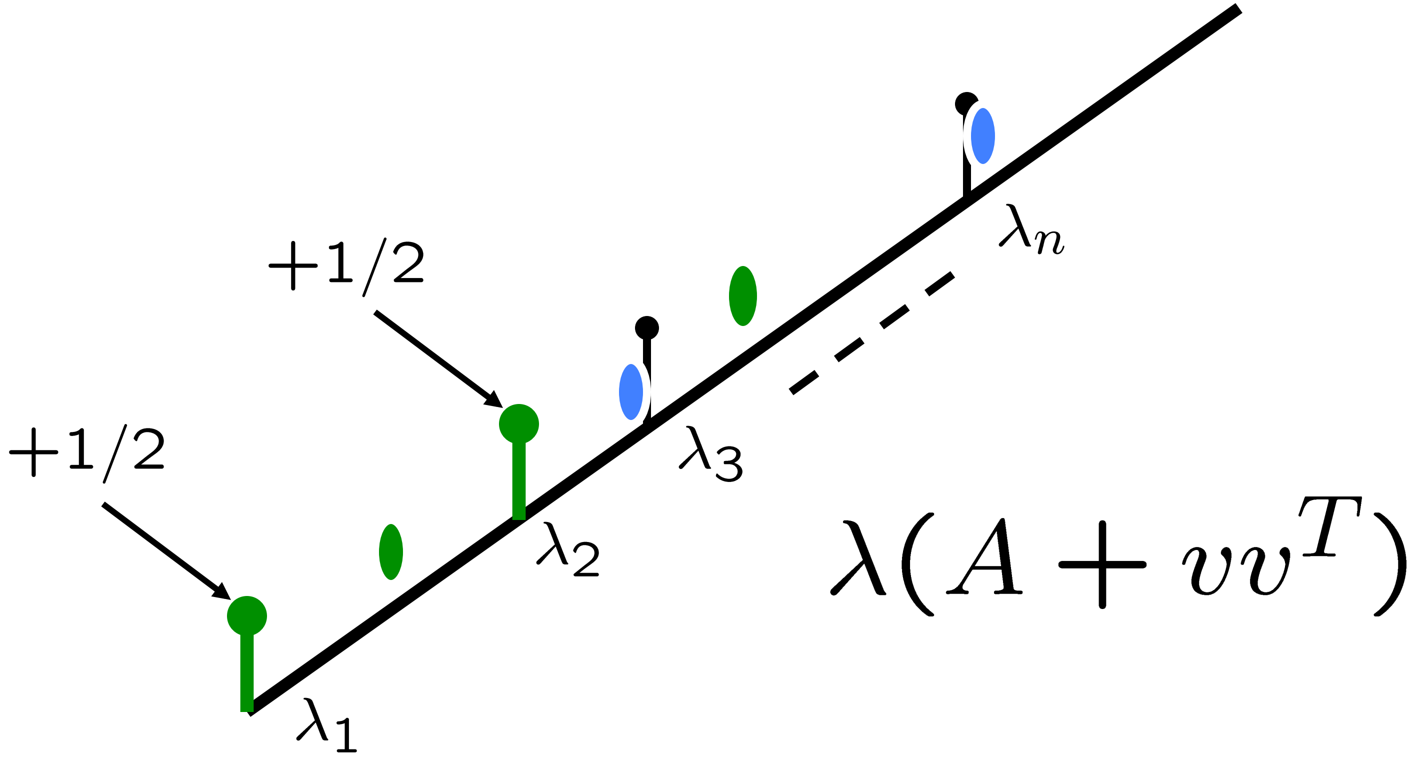
Ex2: Equal weight on u_1, u_2



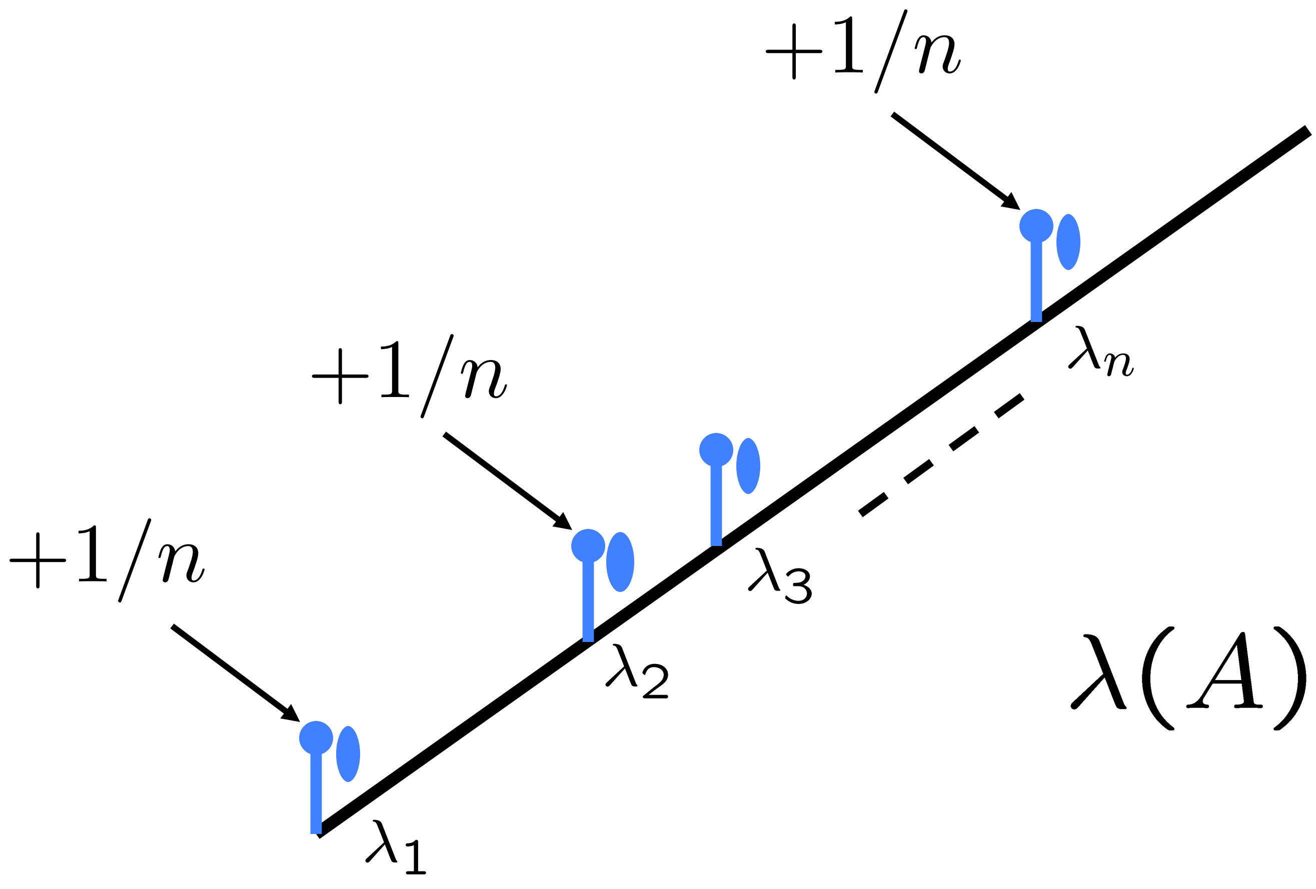
Ex2: Equal weight on u_1, u_2



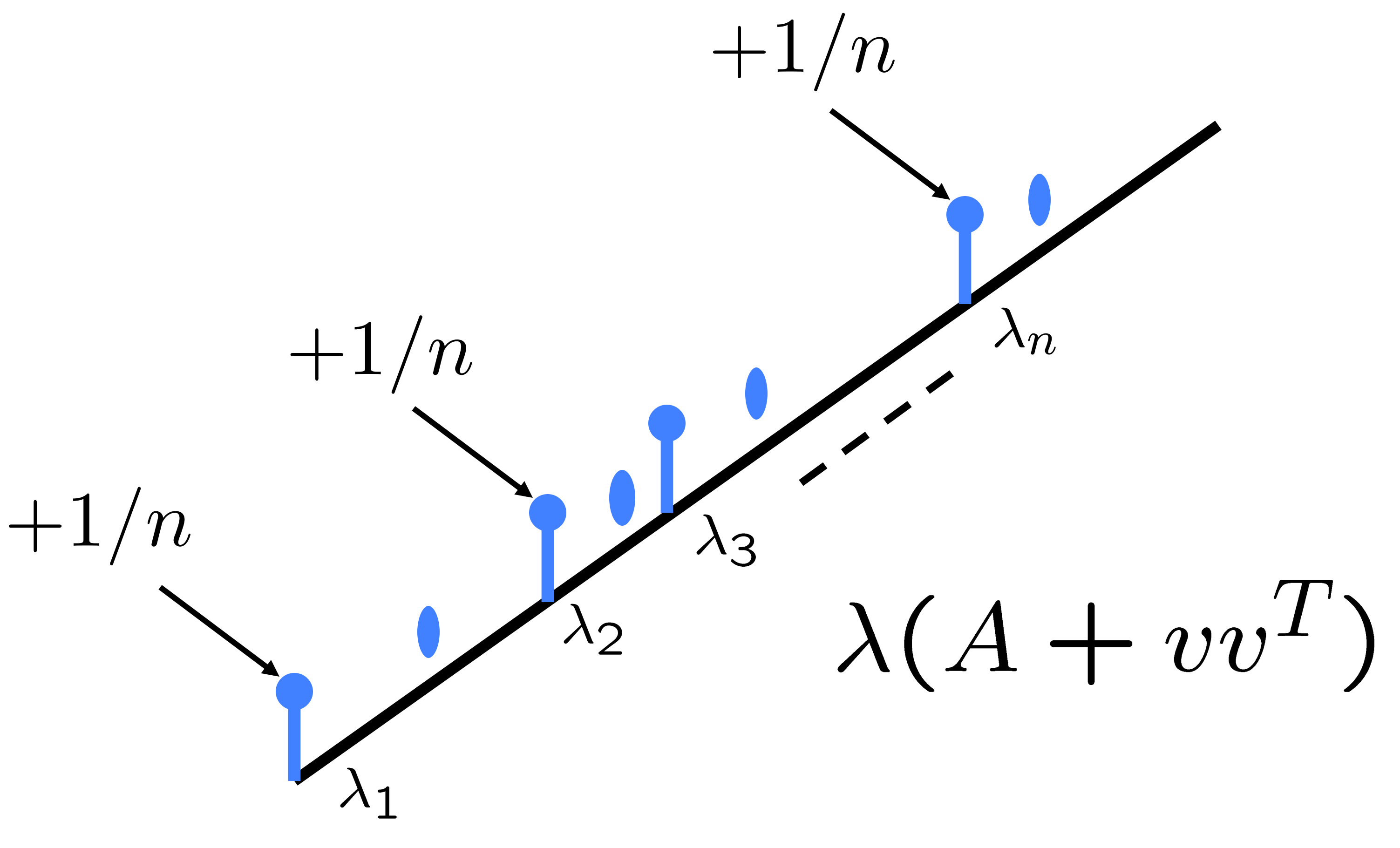
Ex2: Equal weight on u_1, u_2



Ex3: Equal weight on all $u_1, u_2, \ldots, u_n$



Ex3: Equal weight on all $u_1, u_2, \ldots, u_n$



Adding a random v_e

Because v_e are decomposition of identity,

$$\mathop{\mathbf{E}}_{e}\left[\ \langle v_e, u_i \rangle^2\ \right] = 1/m$$

$$\mathop{\mathbf{E}}_{e}\left[\ P_{A+v_e v_e^T}\ \right] = \left(1 + \frac{1}{m}\sum_i \frac{1}{\lambda_i - x}\right) P_A$$

$$= P_A - \frac{1}{m}\frac{d}{dx}P_A$$

Many random v_e

$$\mathop{\mathbb{E}}_{e}\left[\ P_{A+v_e v_e^T}(x)\ \right] = 1 - \frac{1}{m}\frac{d}{dx}P_A(x)$$

$$\mathop{\mathbb{E}}_{e_1,\ldots,e_k}\left[\ P_{v_{e_1}v_{e_1}^T+\cdots+v_{e_k}v_{e_k}^T}(x)\ \right] = \left(1 - \frac{1}{m}\frac{d}{dx}\right)^k x^n$$

Many random v_e

$$\mathop{\mathbb{E}}_{e}\left[\ P_{A+v_e v_e^T}(x)\ \right] = 1 - \frac{1}{m}\frac{d}{dx}P_A(x)$$

$$\mathop{\mathbb{E}}_{e_1,\dots,e_k}\left[\ P_{v_{e_1}v_{e_1}^T+\cdots+v_{e_k}v_{e_k}^T}(x)\ \right] = \left(1 - \frac{1}{m}\frac{d}{dx}\right)^k x^n$$

Is an associated Laguerre polynomial!

For $k = n/\epsilon^2$,
roots lie between $(1-\epsilon)^2\frac{n}{m\epsilon^2}$ and $(1+\epsilon)^2\frac{n}{m\epsilon^2}$

Matrix Sparsification Proof Sketch

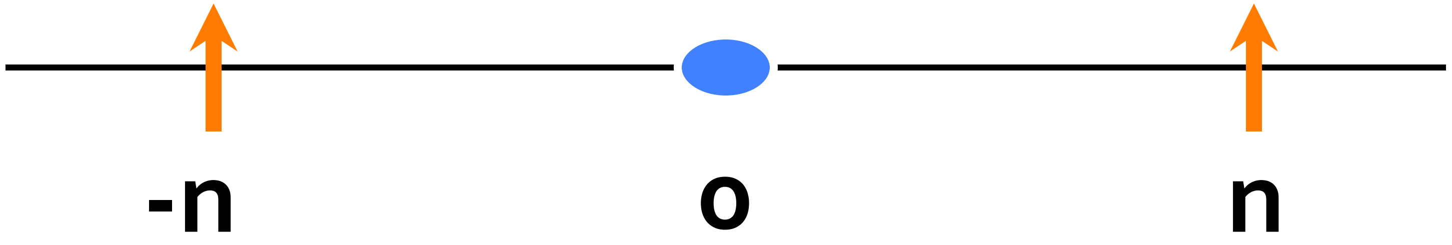
Have $\sum_e v_e v_e^T = I$ Want $\sum_e s_e v_e v_e^T \approx_\epsilon I$

Will do with $|\{e : s_e \neq 0\}| \leq 6n$

All eigenvalues between 1 and 13, $\epsilon \approx 2.6$

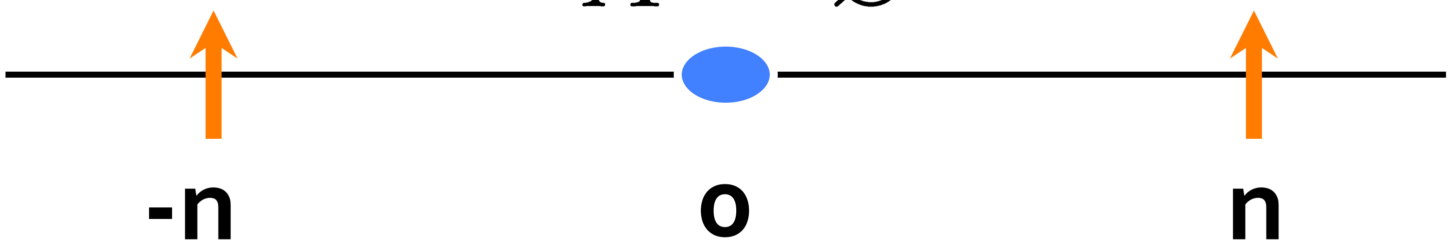
Broad outline: moving barriers

$$A = \varnothing$$



Step 1

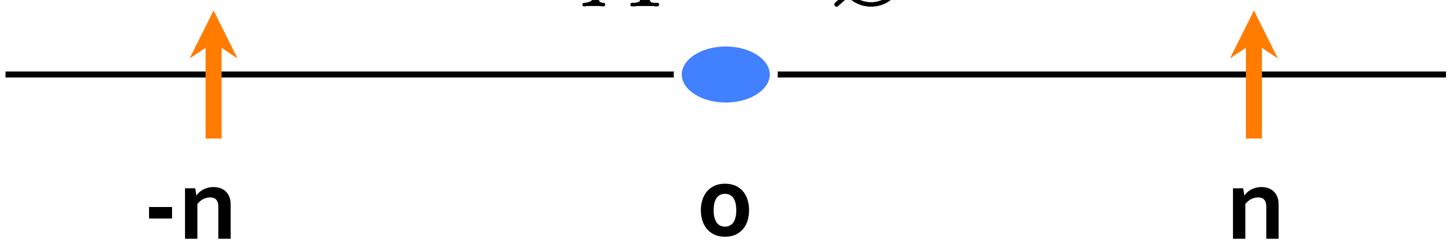
$$A = \varnothing$$



$$+\mathbf{v}\mathbf{v}^T \qquad \mathbf{v} \in \{v_e\}$$

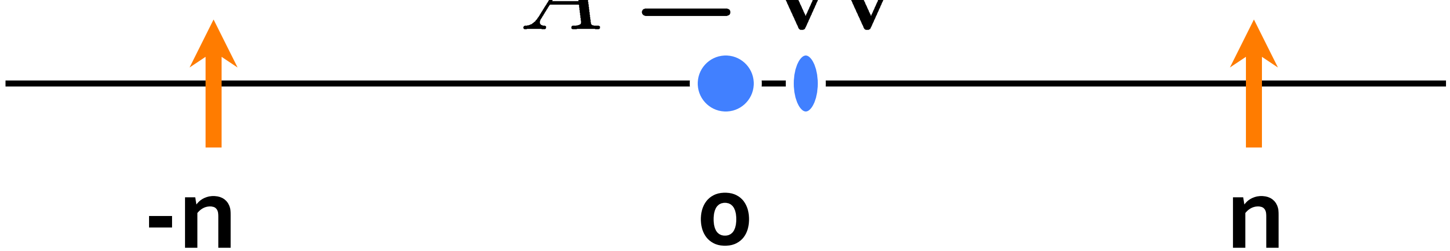
Step 1

$$A = \varnothing$$



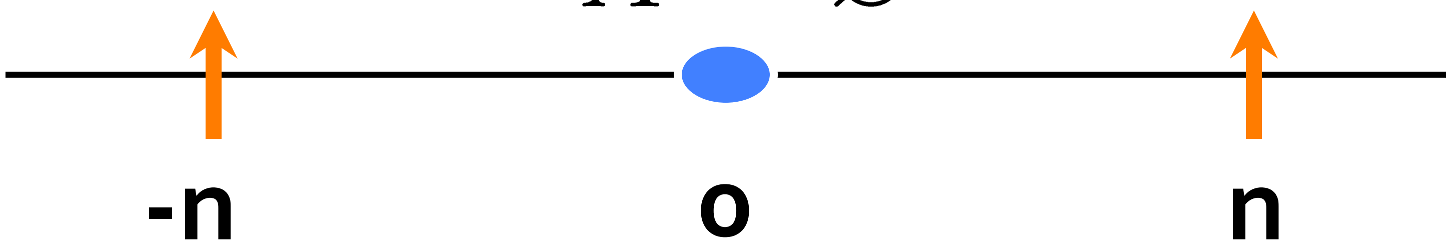
$$+\mathbf{v}\mathbf{v}^T \qquad \mathbf{v} \in \{v_e\}$$

$$A = \mathbf{v}\mathbf{v}^T$$



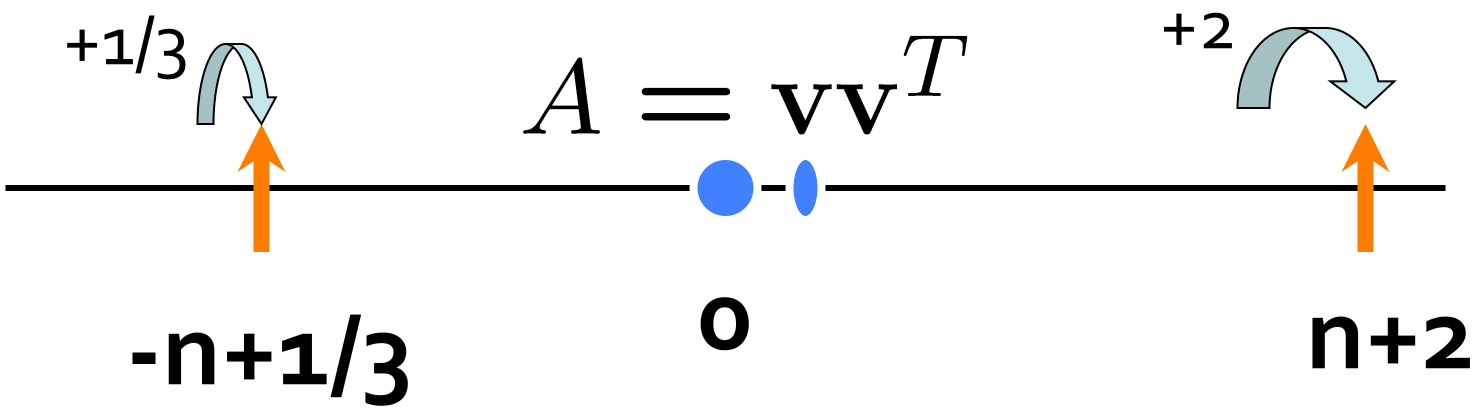
Step 1

$$A = \varnothing$$

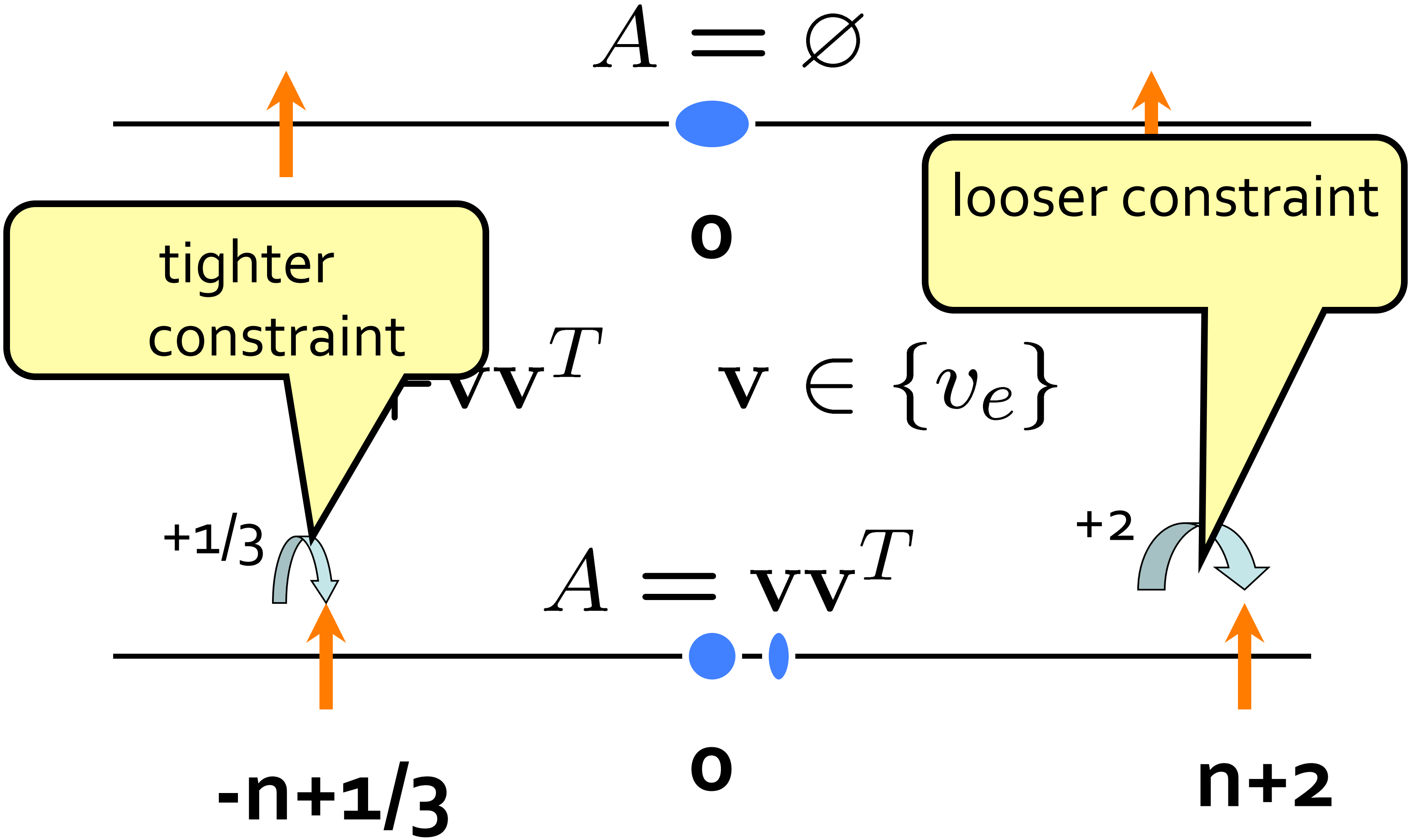


$$+\mathbf{v}\mathbf{v}^T \qquad \mathbf{v} \in \{v_e\}$$

$$A = \mathbf{v}\mathbf{v}^T$$

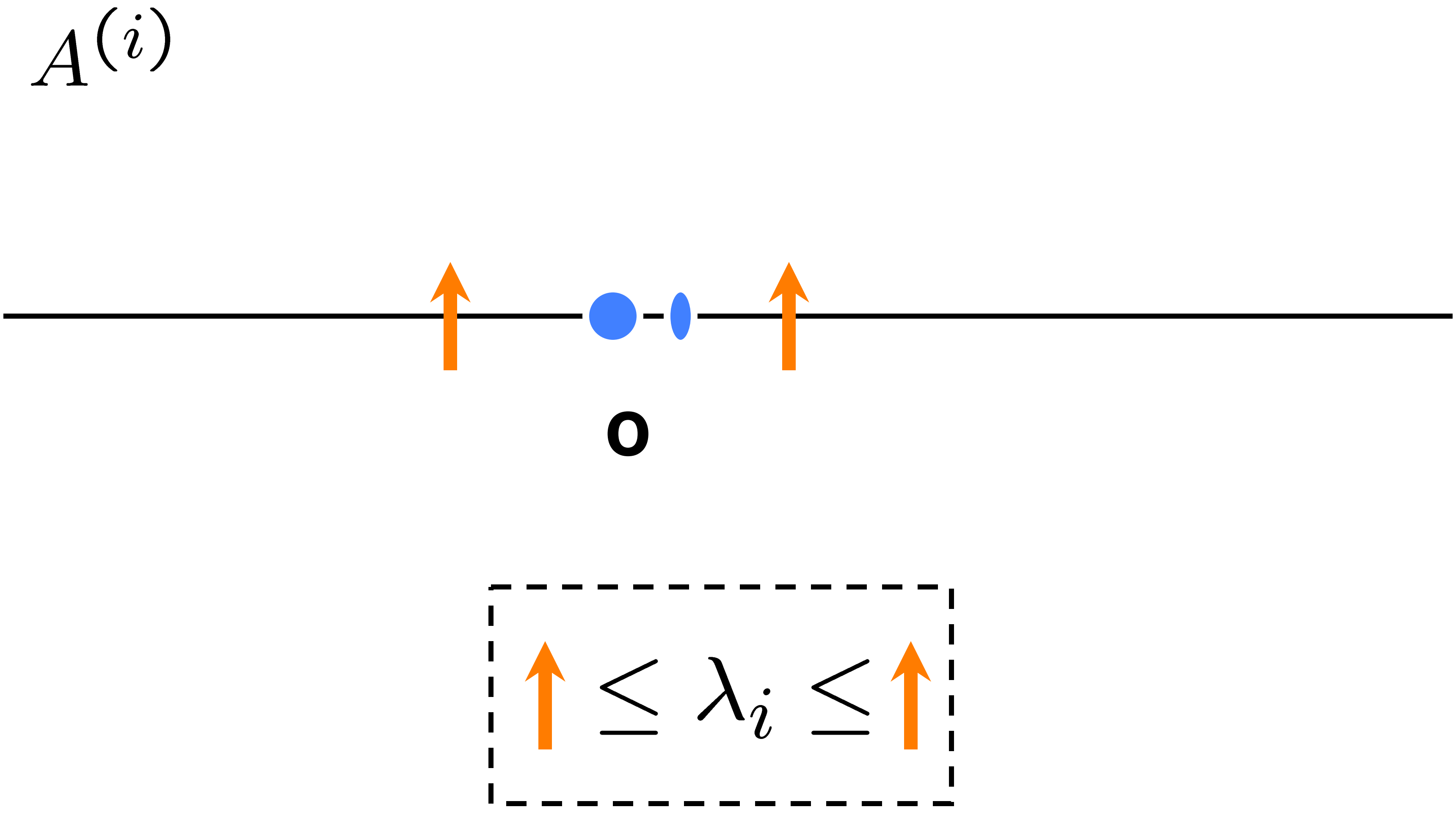


Step 1



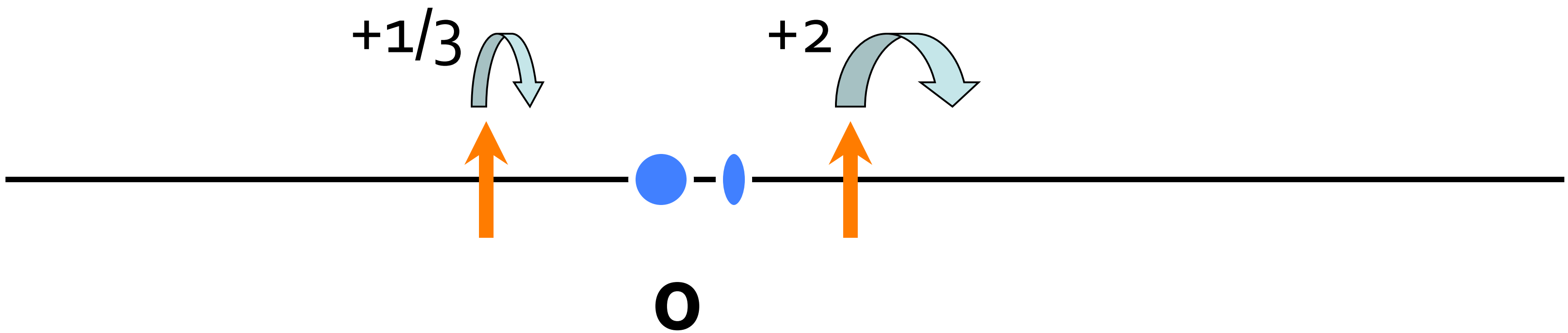
$A = \varnothing$
0
tighter constraint
looser constraint
$\mathbf{v} \in \{v_e\}$
+1/3
+2
$A = \mathbf{v}\mathbf{v}^T$
0
-n+1/3
n+2

Step i+1



Step i+1

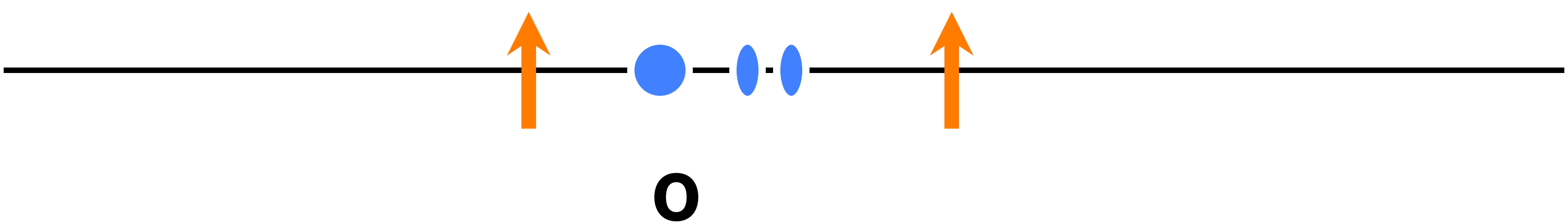
$$A^{(i)}$$

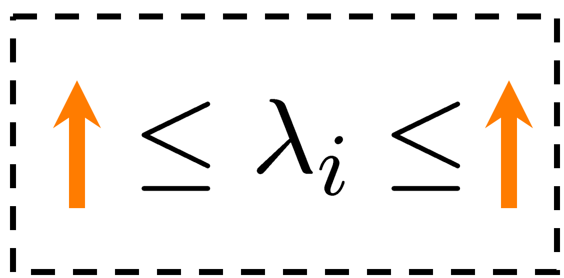


$$+\mathbf{v}\mathbf{v}^T$$

Step i+1

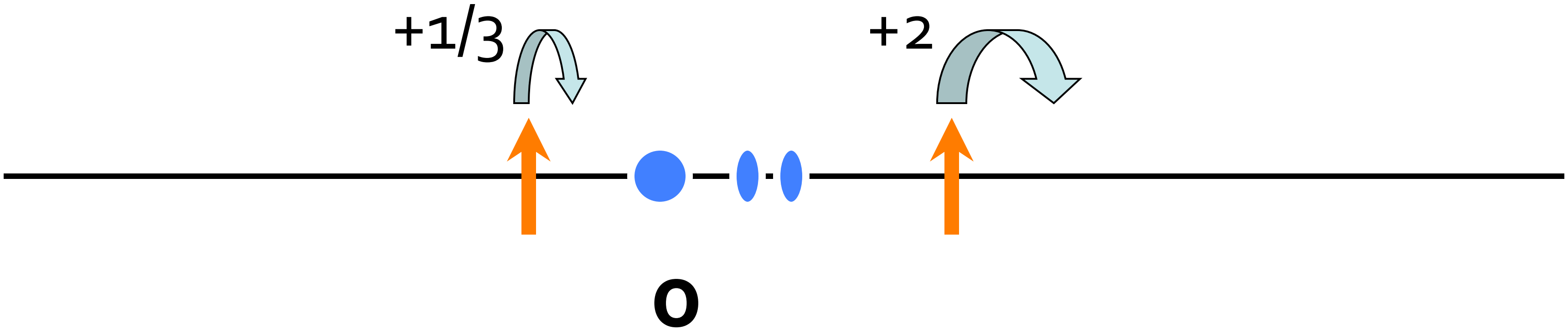
$$A^{(i)}, A^{(i+1)}$$





Step i+1

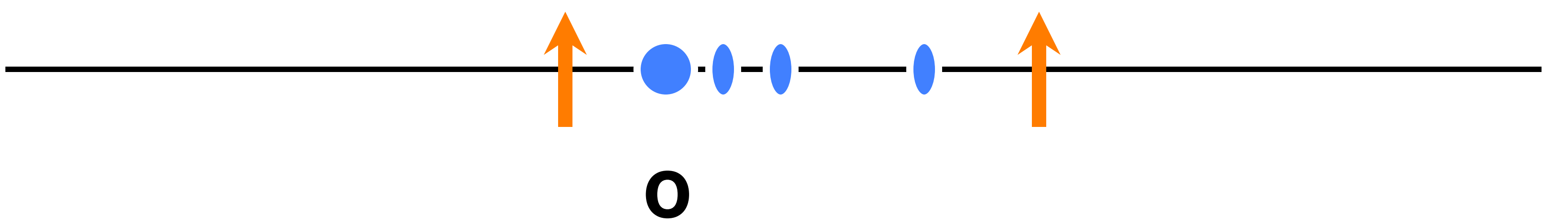
$$A^{(i)}, A^{(i+1)}$$

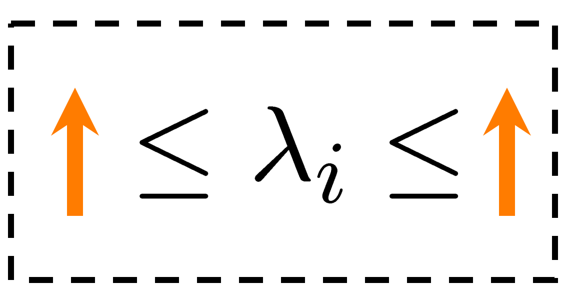


$$+\mathbf{v}\mathbf{v}^T$$

Step i+1

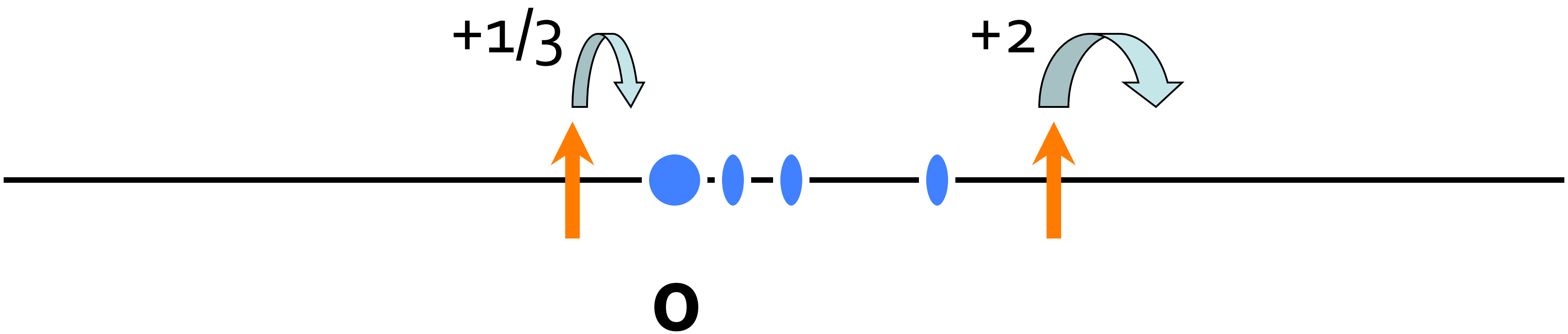
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}$$





Step i+1

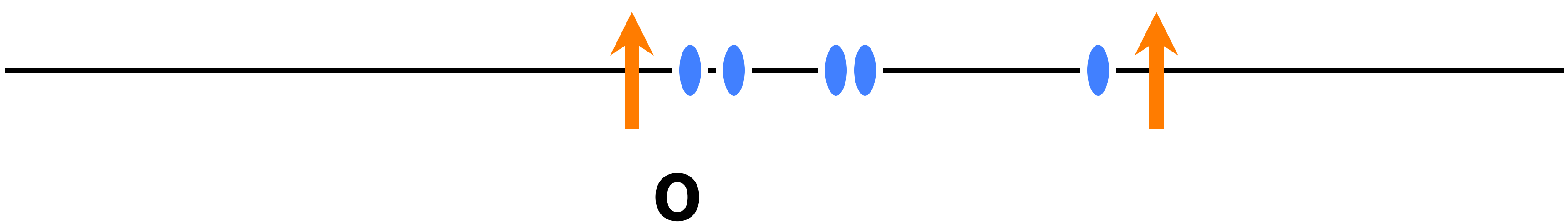
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}$$

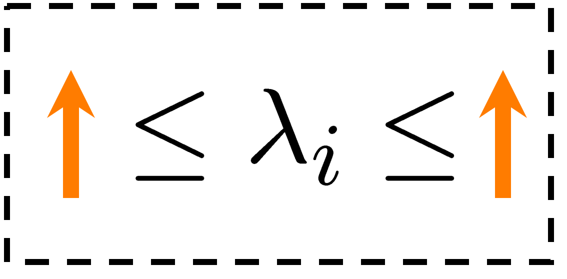


$$+\mathbf{v}\mathbf{v}^T$$

Step i+1

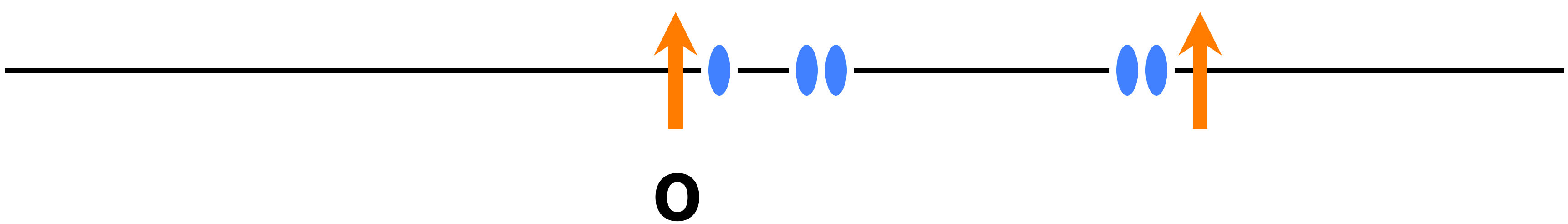
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}$$

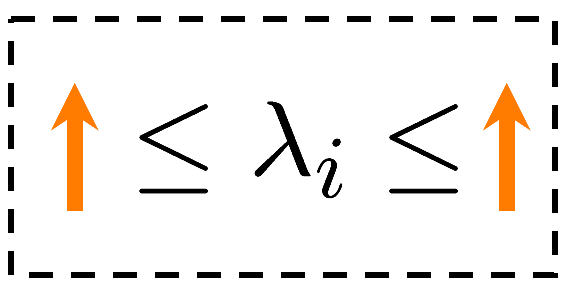




Step i+1

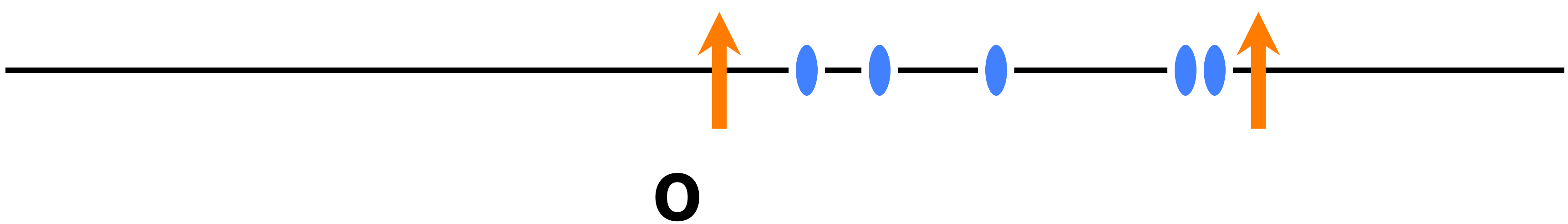
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$

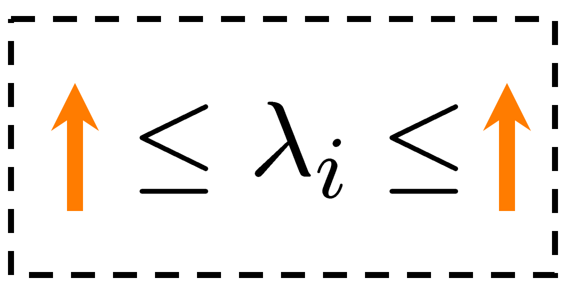




Step i+1

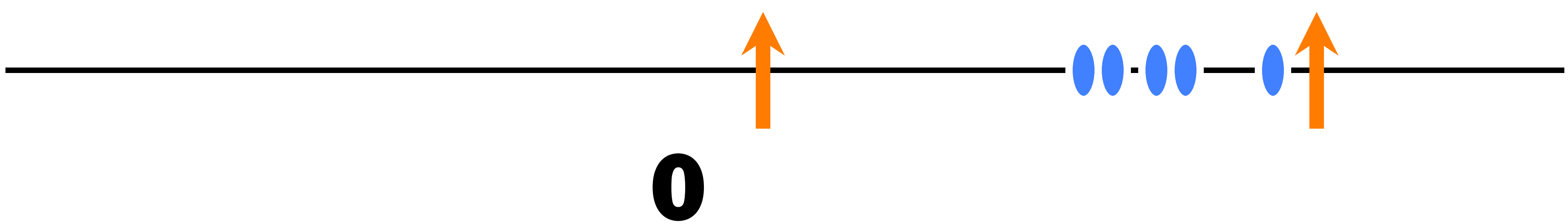
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$

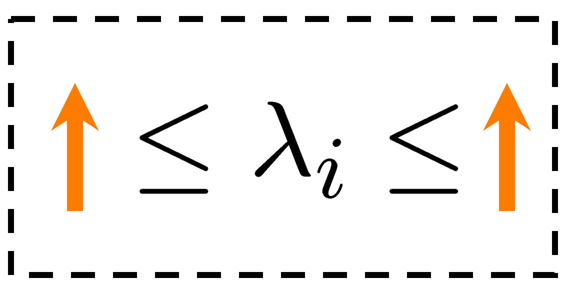




Step i+1

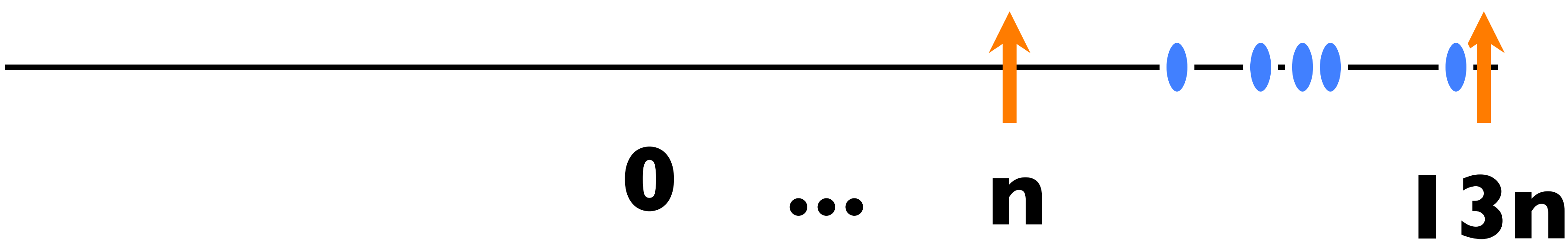
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$

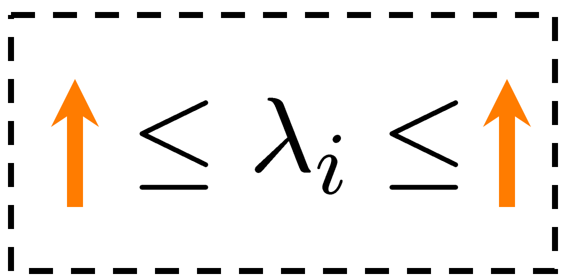




Step 6n

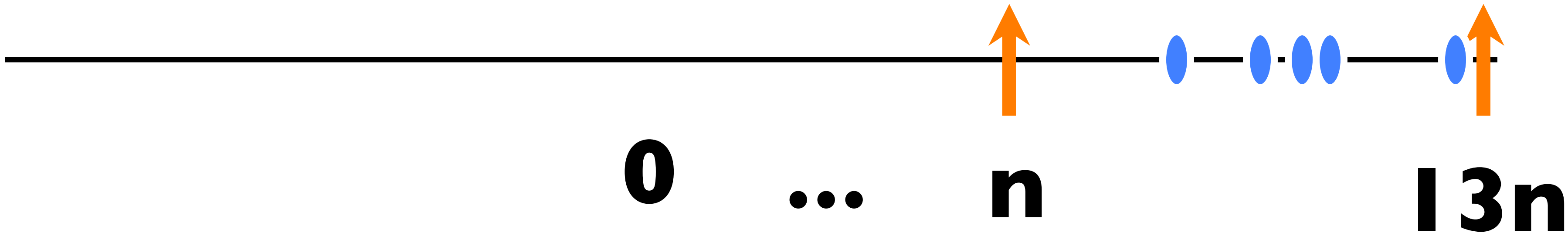
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$





Step 6n

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$



2.6-approximation with 6n vectors.

$\square$

Problem

need to show that an appropriate

$$v_e v_e^T$$

always exists.

Problem

need to show that an appropriate

$$v_e v_e^T$$

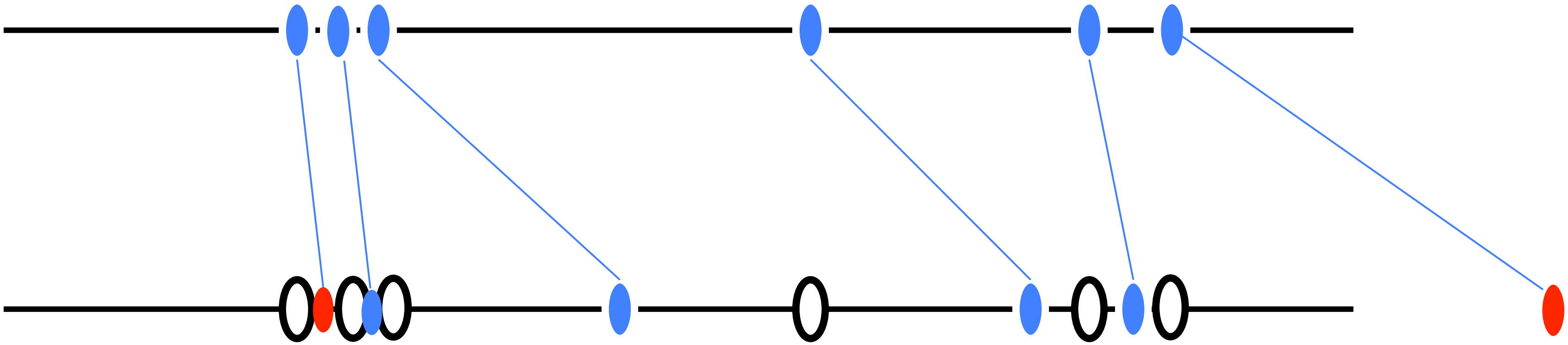
always exists.

$$\uparrow \leq \lambda_i \leq \uparrow$$

Is not strong enough for induction

Problems

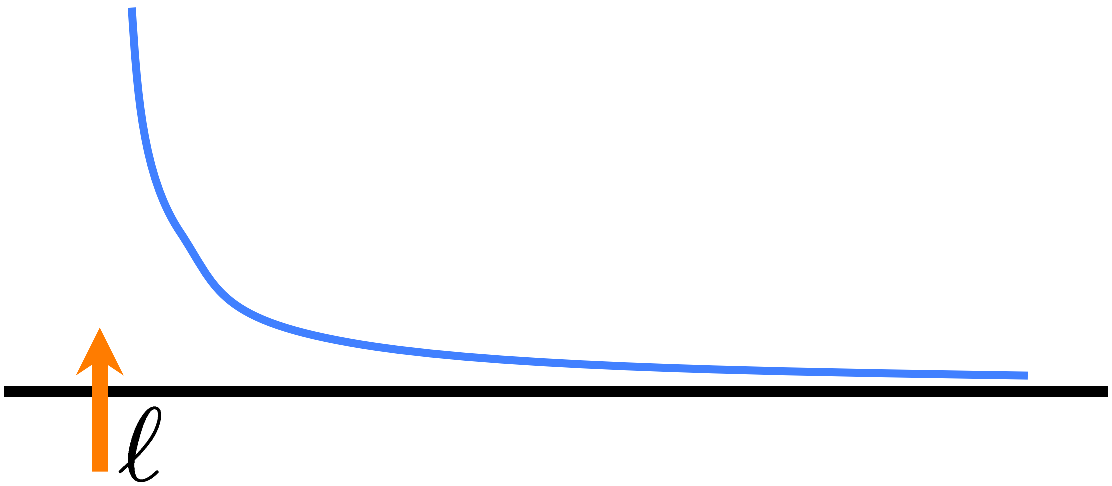
If many small eigenvalues, can only move one



Bunched large eigenvalues
repel the highest one

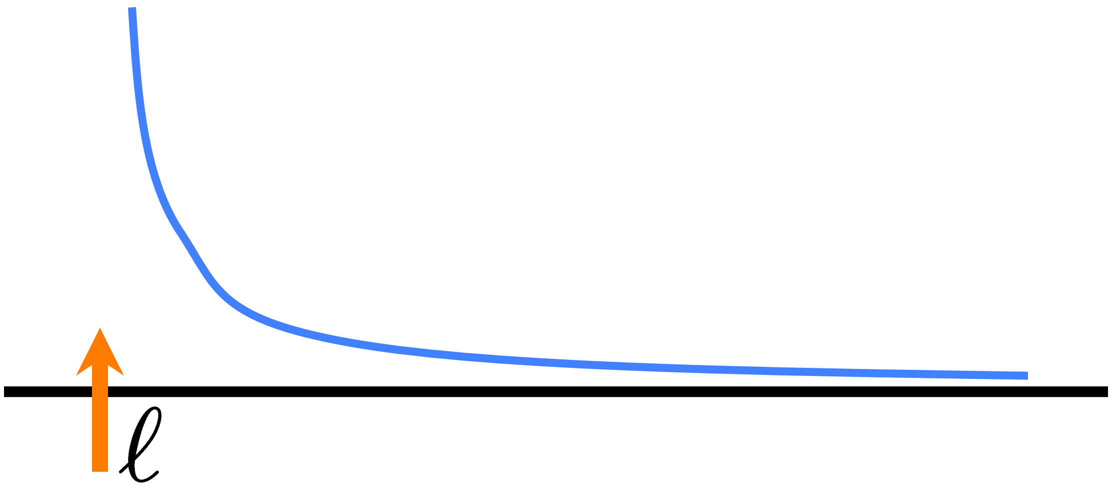
The Lower Barrier Potential Function

$$\Phi_\ell(A) = \sum_i \frac{1}{\lambda_i - \ell} = \mathrm{Tr}\left((A - \ell I)^{-1}\right)$$



The Lower Barrier Potential Function

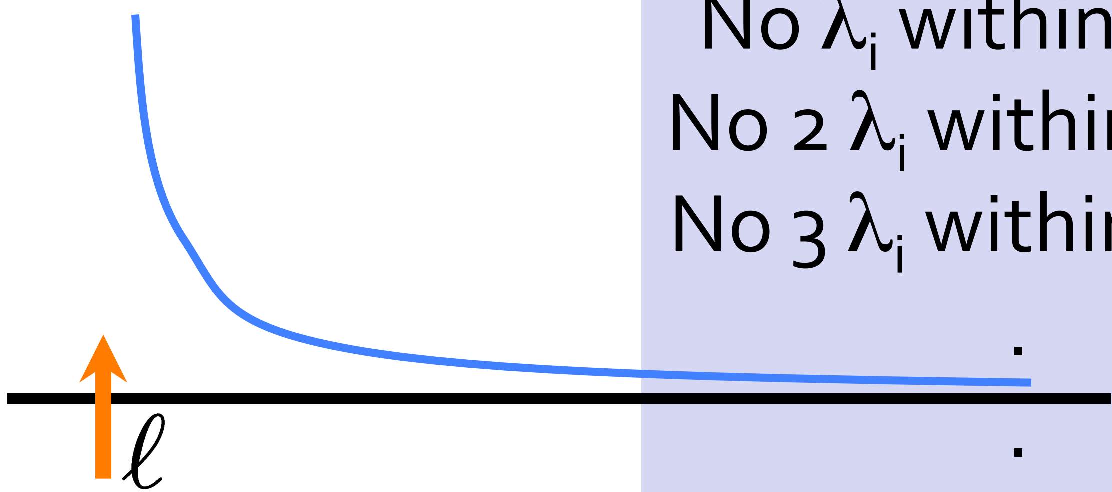
$$\Phi_\ell(A) = \sum_i \frac{1}{\lambda_i - \ell} = \mathrm{Tr}\left((A - \ell I)^{-1}\right)$$



$$\Phi_\ell(A) \leq 1 \implies \lambda_{min}(A) \geq \ell + 1$$

The Lower Barrier Potential Function

$$\Phi_\ell(A) = \sum_i \frac{1}{\lambda_i - \ell} = \mathrm{Tr}\left((A - \ell I)^{-1}\right)$$

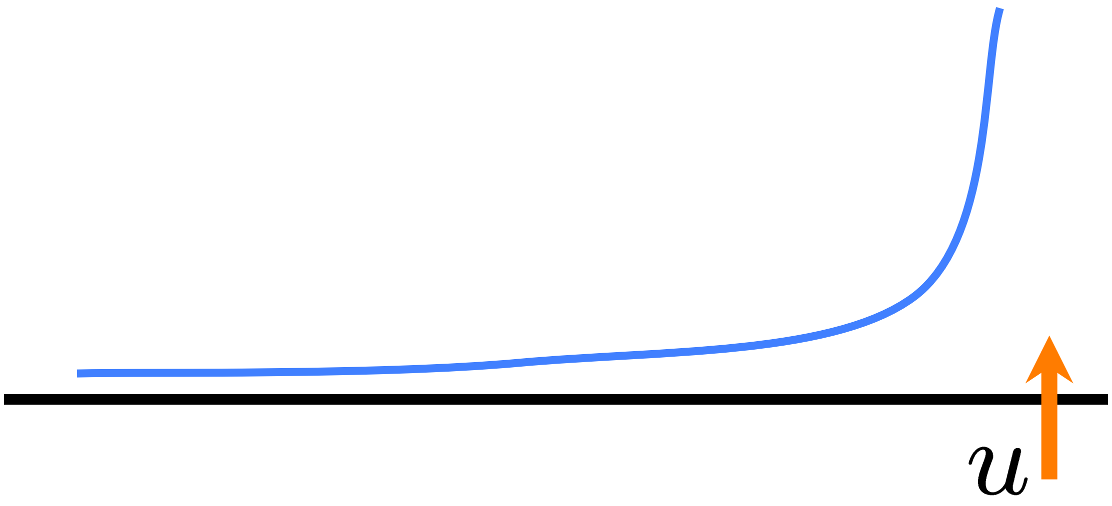


No λ_i within dist. 1
No 2 λ_i within dist. 2
No 3 λ_i within dist. 3
.
.
No k λ_i within dist. k

$$\Phi_\ell(A) \leq 1 \implies \lambda_{min}(A) \geq \ell + 1$$

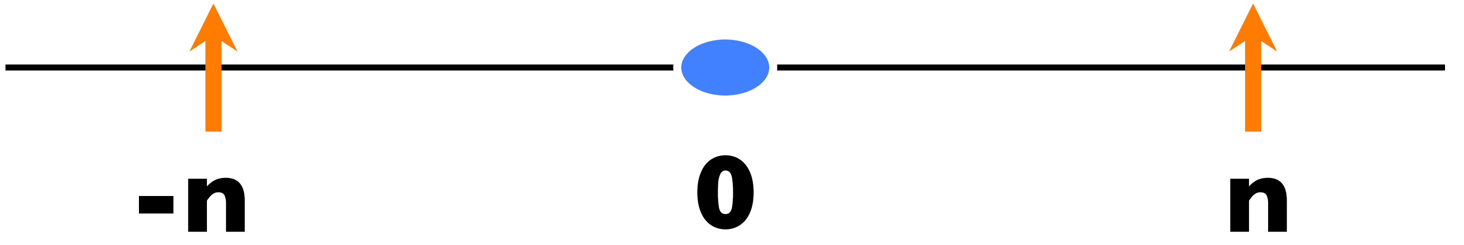
The Upper Barrier Potential Function

$$\Phi^u(A) = \sum_i \frac{1}{u - \lambda_i} = \mathrm{Tr}\left((uI - A)^{-1}\right)$$



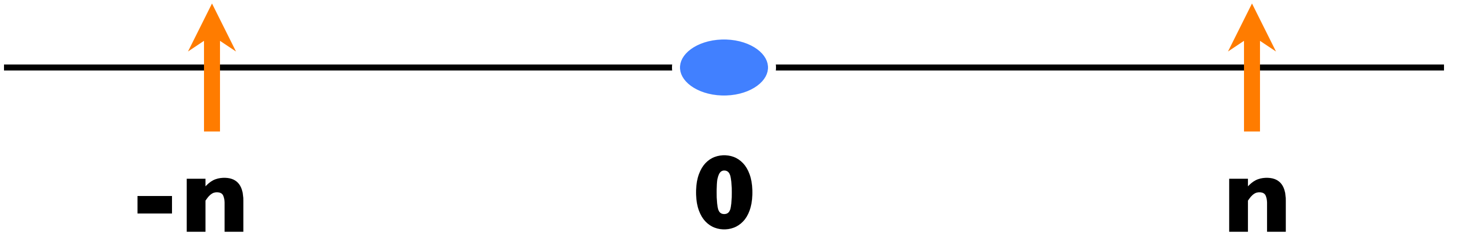
The Beginning

$$A = \varnothing$$



The Beginning

$$A = \varnothing$$

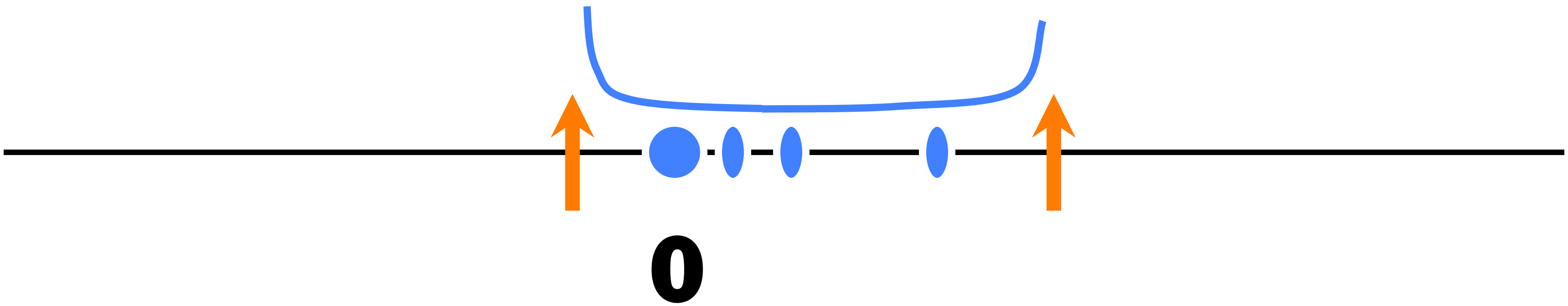


$$\Phi^n(\varnothing) = \mathrm{Tr}(nI)^{-1} = 1$$

$$\Phi_{-n}(\varnothing) = \mathrm{Tr}(nI)^{-1} = 1.$$

Step i+1

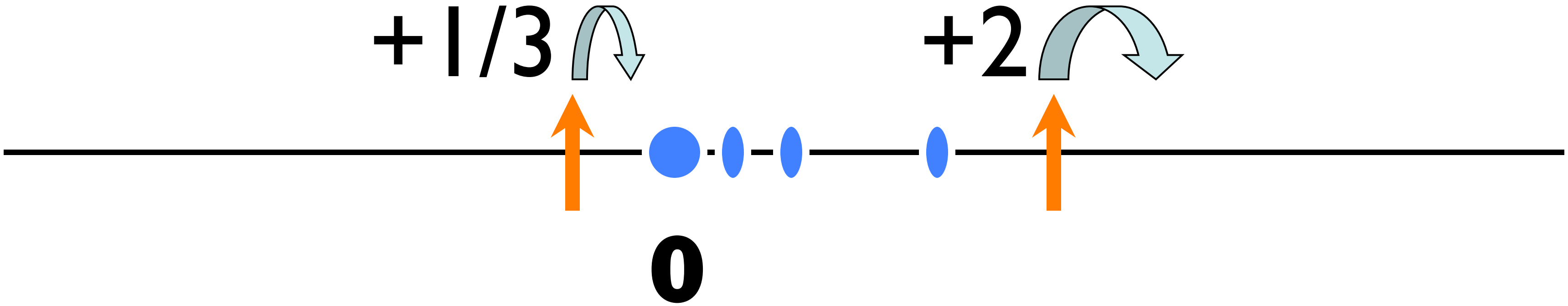
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}$$



$$\Phi^u(A) \le 1$$
$$\Phi_\ell(A) \le 1.$$

Step i+1

$A^{(i)}, A^{(i+1)}, A^{(i+2)}$

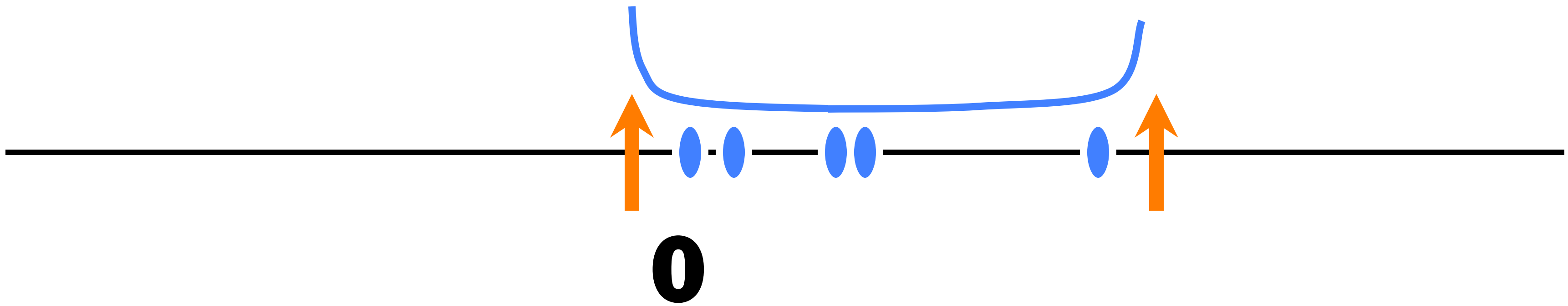


Lemma.
can always choose $+s_e v_e v_e^T$ so that potentials do not increase

$\Phi^u(A) \leq 1$
$\Phi_\ell(A) \leq 1.$

Step i+1

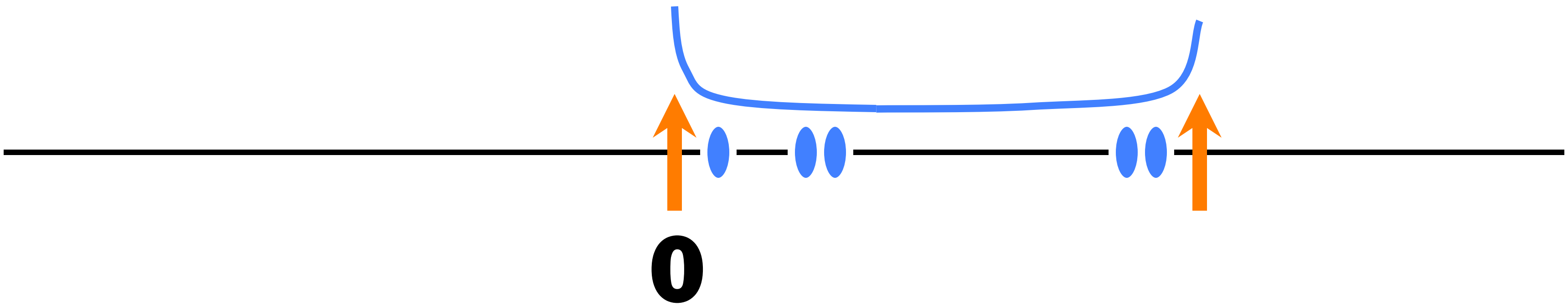
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}$$



$$\Phi^u(A) \leq 1$$
$$\Phi_\ell(A) \leq 1.$$

Step i+1

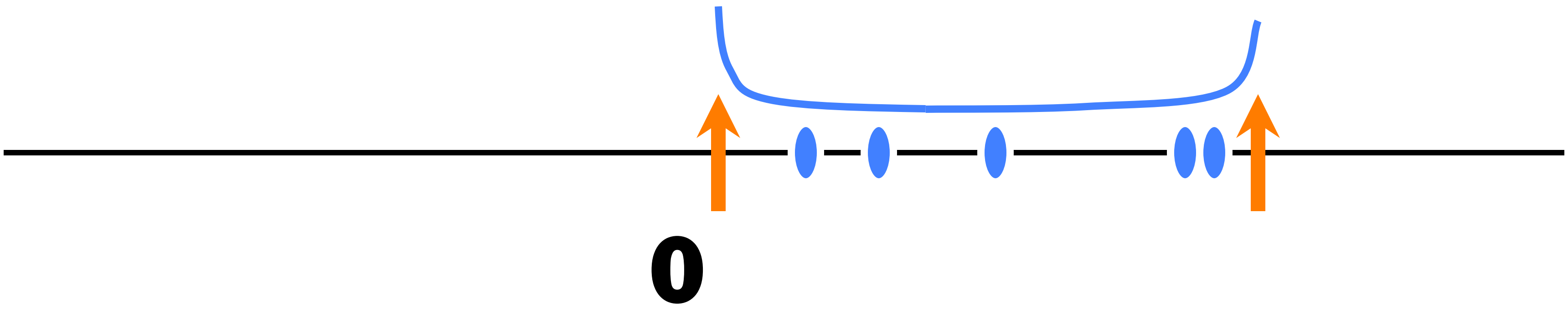
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$



$$\Phi^u(A) \leq 1$$
$$\Phi_\ell(A) \leq 1.$$

Step i+1

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$

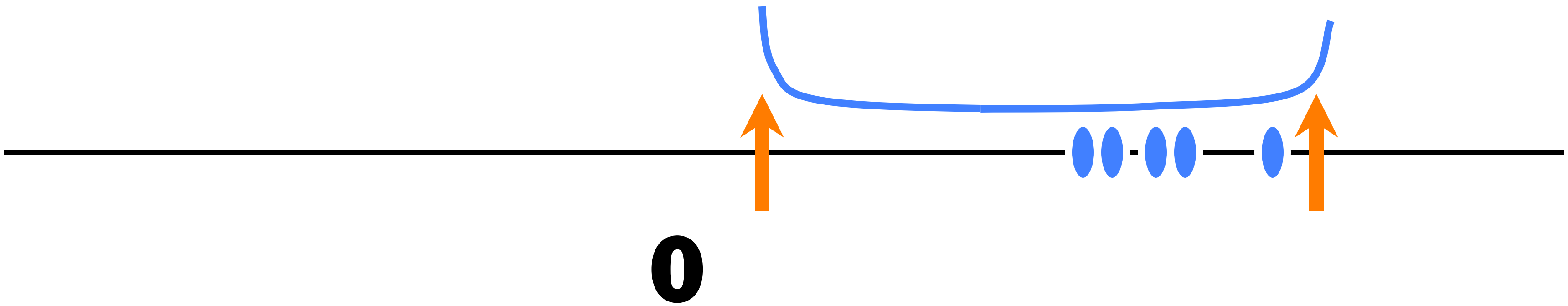


$$\Phi^u(A) \leq 1$$

$$\Phi_\ell(A) \leq 1.$$

Step i+1

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$



$$\Phi^u(A) \leq 1$$

$$\Phi_\ell(A) \leq 1.$$

Step 6n

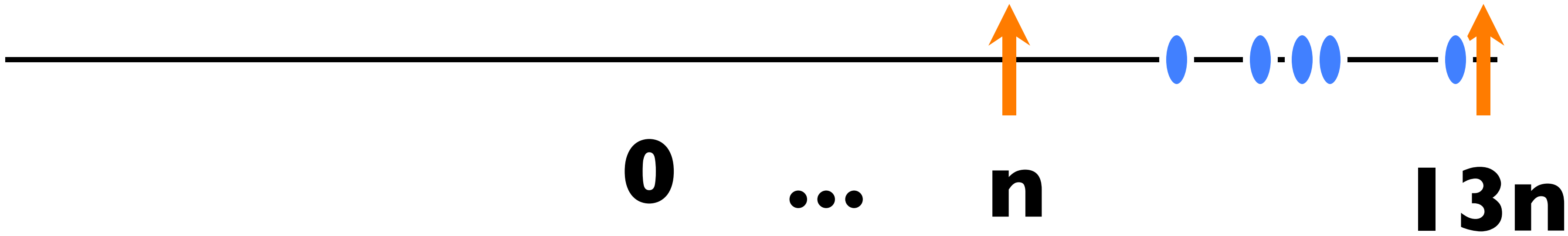
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$

$$\Phi^u(A) \leq 1$$

$$\Phi_\ell(A) \leq 1.$$

Step 6n

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$



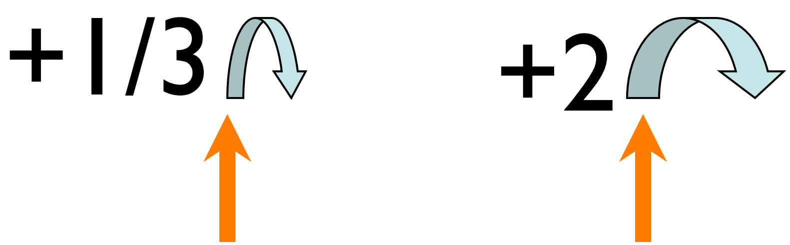
2.6-approximation with 6n vectors.

□

Goal

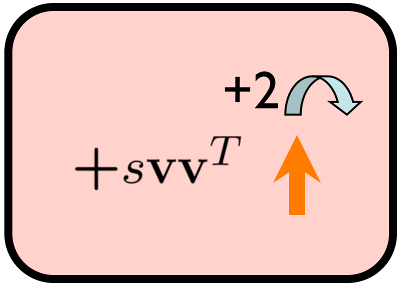
Lemma.
can always choose $+s_e v_e v_e^T$ so that potentials do not increase

$$\Phi^u(A) \le 1$$
$$\Phi_\ell(A) \le 1.$$



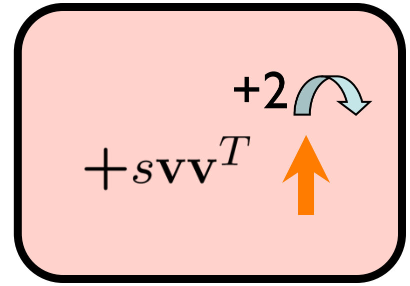
$$+s_e v_e v_e^T$$

Upper Barrier Update



Add $s\boldsymbol{v}\boldsymbol{v}^T$ and set $u' \leftarrow u + 2$.

Upper Barrier Update

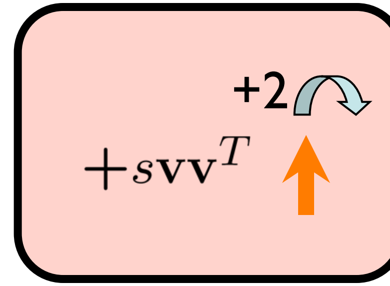


Add $s\boldsymbol{v}\boldsymbol{v}^T$ and set $u' \leftarrow u + 2$.

$$\Phi^{u'}(A + s\boldsymbol{v}\boldsymbol{v}^T)$$

$$= \mathrm{Tr}\left((u'I - A - s\boldsymbol{v}\boldsymbol{v}^T)^{-1}\right)$$

Upper Barrier Update



Add $s\boldsymbol{v}\boldsymbol{v}^T$ and set $u' \leftarrow u + 2$.

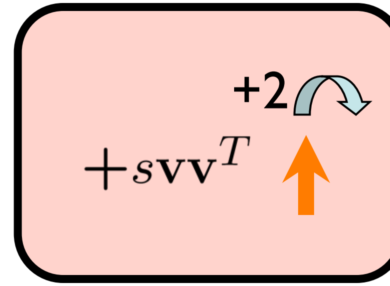
$$\Phi^{u'}(A + s\boldsymbol{v}\boldsymbol{v}^T)$$

$$= \mathrm{Tr}\left((u'I - A - s\boldsymbol{v}\boldsymbol{v}^T)^{-1}\right)$$

$$= \Phi^{u'}(A) + \frac{s\boldsymbol{v}^T(u'I - A)^{-2}\boldsymbol{v}}{1 - s\boldsymbol{v}^T(u'I - A)^{-1}\boldsymbol{v}}$$

By Sherman-Morrison Formula

Upper Barrier Update



Add $s\boldsymbol{v}\boldsymbol{v}^T$ and set $u' \leftarrow u + 2$.

$$\Phi^{u'}(A + s\boldsymbol{v}\boldsymbol{v}^T)$$

$$= \mathrm{Tr}\left((u'I - A - s\boldsymbol{v}\boldsymbol{v}^T)^{-1}\right)$$

$$= \Phi^{u'}(A) + \frac{s\boldsymbol{v}^T(u'I - A)^{-2}\boldsymbol{v}}{1 - s\boldsymbol{v}^T(u'I - A)^{-1}\boldsymbol{v}}$$

Need $\leq \Phi^{u}(A)$

How much of vv^T can we add?

Rearranging:

$$\Phi^{u'}(A + s\boldsymbol{v}\boldsymbol{v}^T) \leq \Phi^u(A)$$

iff

$$1 \geq s\boldsymbol{v}^T \left(\frac{(u'I - A)^{-2}}{\Phi^u(A) - \Phi^{u'}(A)} + (u'I - A)^{-1} \right) \boldsymbol{v}$$

How much of vv^T can we add?

Rearranging:

$$\Phi^{u'}(A + s\boldsymbol{v}\boldsymbol{v}^T) \leq \Phi^u(A)$$

iff

$$1 \geq s\boldsymbol{v}^T \left(\frac{(u'I - A)^{-2}}{\Phi^u(A) - \Phi^{u'}(A)} + (u'I - A)^{-1} \right) \boldsymbol{v}$$

Write as $1 \geq s\boldsymbol{v}^T U_A \boldsymbol{v}$

Lower Barrier

Similarly:

$$\Phi_{l'}(A + s\boldsymbol{v}\boldsymbol{v}^T) \leq \Phi_l(A)$$

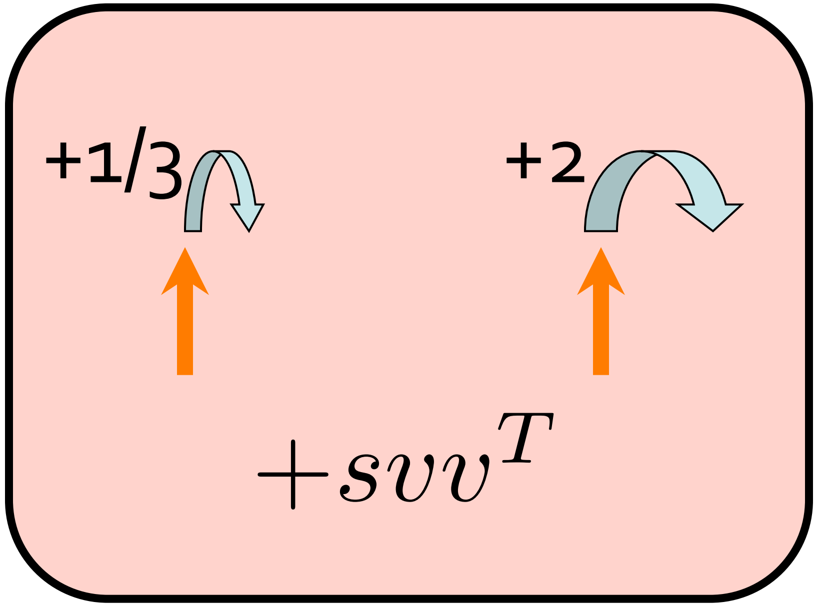
iff

$$1 \leq s\boldsymbol{v}^T \left(\frac{(A - l'I)^{-2}}{\Phi_{l'}(A) - \Phi_l(A)} - (A - l'I)^{-1} \right) \boldsymbol{v}$$

Write as $1 \leq s\boldsymbol{v}^T L_A \boldsymbol{v}$

Goal

Show that we can always add some vector while respecting *both* barriers.



Need: $s\boldsymbol{v}^T U_A \boldsymbol{v} \leq 1 \leq s\boldsymbol{v}^T L_A \boldsymbol{v}$

Two expectations

Need: $$s\boldsymbol{v}^T U_A \boldsymbol{v} \leq 1 \leq s\boldsymbol{v}^T L_A \boldsymbol{v}$$

Can show: $$\mathop{\mathbb{E}}_{e}\left[\, v_e^T U_A v_e \,\right] \leq 3/2m$$

$$\mathop{\mathbb{E}}_{e}\left[\, v_e^T L_A v_e \,\right] \geq 2/m$$

Two expectations

Need: $s\boldsymbol{v}^T U_A \boldsymbol{v} \leq 1 \leq s\boldsymbol{v}^T L_A \boldsymbol{v}$

Can show: $\mathbb{E}_e \left[\; v_e^T U_A v_e \; \right] \leq 3/2m$

$$\mathbb{E}_e \left[\; v_e^T L_A v_e \; \right] \geq 2/m$$

So: $\mathbb{E}_e \left[\; v_e^T U_A v_e \; \right] \leq \mathbb{E}_e \left[\; v_e^T L_A v_e \; \right]$

And, exists $e : v_e^T U_A v_e \leq v_e^T L_A v_e$

Two expectations

Need: $s\boldsymbol{v}^T U_A \boldsymbol{v} \leq 1 \leq s\boldsymbol{v}^T L_A \boldsymbol{v}$

Can show: $\mathop{\mathbb{E}}_{e}\left[\ v_e^T U_A v_e\ \right] \leq 3/2m$

$\mathop{\mathbb{E}}_{e}\left[\ v_e^T L_A v_e\ \right] \geq 2/m$

So: $\mathop{\mathbb{E}}_{e}\left[\ v_e^T U_A v_e\ \right] \leq \mathop{\mathbb{E}}_{e}\left[\ v_e^T L_A v_e\ \right]$

And, exists $e : v_e^T U_A v_e \leq v_e^T L_A v_e$

And s that puts 1 between them

Two expectations

Need: $\quad s\boldsymbol{v}^T U_A \boldsymbol{v} \leq 1 \leq s\boldsymbol{v}^T L_A \boldsymbol{v}$

Can show:

$$\mathop{\mathbb{E}}_{e}\left[\ v_e^T U_A v_e\ \right] \leq 3/2m$$

$$\mathop{\mathbb{E}}_{e}\left[\ v_e^T L_A v_e\ \right] \geq 2/m$$

So: $\quad \mathop{\mathbb{E}}_{e}\left[\ v_e^T U_A v_e\ \right] \leq \mathop{\mathbb{E}}_{e}\left[\ v_e^T L_A v_e\ \right]$

And, exists $e : v_e^T U_A v_e \leq v_e^T L_A v_e$

And s that puts 1 between them

Bounding expectations

$$v^T U_A v = \mathrm{Tr}\left(U_A v v^T\right)$$

$$\mathop{\mathbf{E}}_{e}\left[\ \mathrm{Tr}\left(U_A v_e v_e^T\right)\ \right] = \mathrm{Tr}\left(U_A \mathop{\mathbf{E}}_{e}\left[\ v_e v_e^T\ \right]\right)$$

$$= \mathrm{Tr}\left(U_A I\right)/m$$

$$= \mathrm{Tr}\left(U_A\right)/m$$

Bounding expectations

$$\mathrm{Tr}\left(U_A\right) = \frac{\mathrm{Tr}\left(\left(u'I - A\right)^{-2}\right)}{\Phi^{u}(A) - \Phi^{u'}(A)} + \mathrm{Tr}\left(\left(u'I - A\right)^{-1}\right)$$

$$= \Phi^{u'}(A)$$

$$\leq \Phi^{u}(A)$$

$$\leq 1$$

As barrier function is monotone decreasing

Bounding expectations

$$\mathrm{Tr}\left(U_A\right) = \frac{\mathrm{Tr}\left(\left(u'I - A\right)^{-2}\right)}{\Phi^{u}(A) - \Phi^{u'}(A)} + \mathrm{Tr}\left(\left(u'I - A\right)^{-1}\right)$$

Numerator is derivative of barrier function.

$$= \Phi^{u'}(A)$$
$$\leq \Phi^{u}(A)$$
$$\leq 1$$

As barrier function is convex,

$$\leq \frac{1}{u' - u}$$

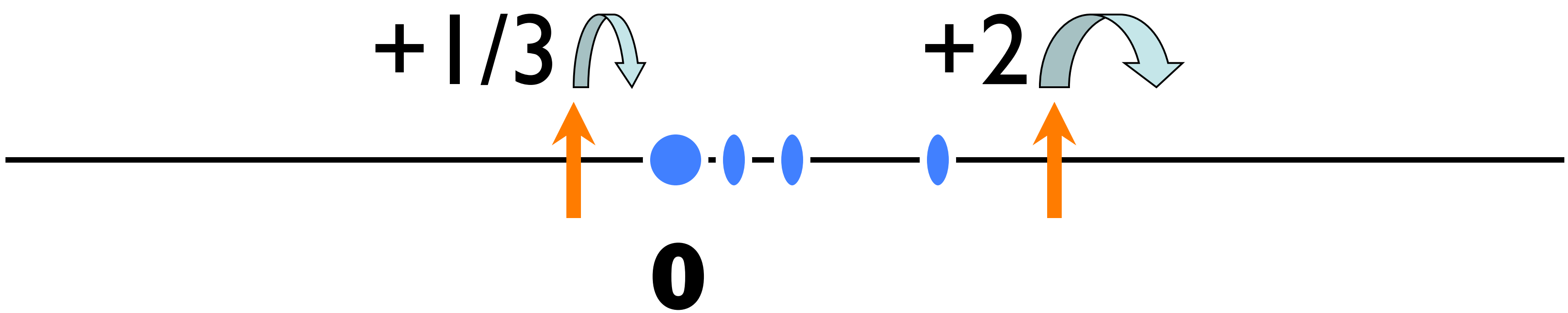
$$\mathrm{Tr}(U_A) \leq 1/2 + 1 = 3/2$$

Bounding expectations

Similarly,
$$\mathrm{Tr}\left(L_A\right) \geq \frac{1}{l' - l} - 1$$
$$= \frac{1}{1/3} - 1$$
$$= 2$$

Step i+1

$A^{(i)}, A^{(i+1)}, A^{(i+2)}$



Lemma.
can always choose $+s\mathbf{v}\mathbf{v}^T$
so that potentials do not increase

$\Phi^u(A) \leq 1$

$\Phi_\ell(A) \leq 1.$

Twice-Ramanujan Sparsifiers

Fixing dn steps and tightening parameters gives ratio

$$\frac{\lambda_{max}(A)}{\lambda_{min}(A)} \leq \frac{d+1+2\sqrt{d}}{d+1-2\sqrt{d}}$$

Less than twice as many edges as used by Ramanujan Expander of same quality

Open Questions

The Ramanujan bound

Properties of vectors from graphs?

Faster algorithm

union of random Hamiltonian cycles?