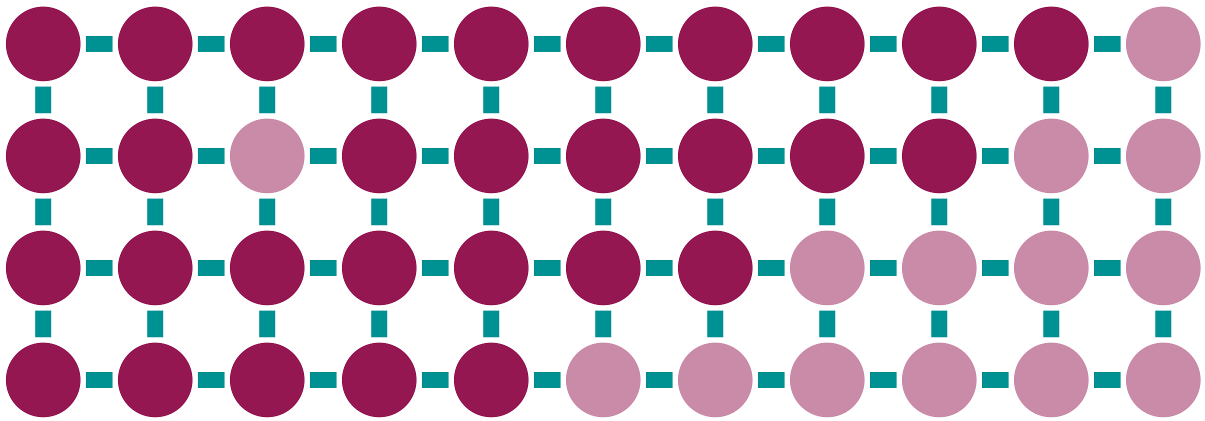
Discrete Extension and Selection Problems

Yuval Rabani - The Hebrew University of Jerusalem

motivation - image segmentation [KT '99]

Input: a raster image degraded by noise.
Output: a restored image.
Assumption: small spatial discontinuity.



Homogeneous Markov random fields with pairwise interactions.

motivation - Lipschitz extension

finite $X \subset Y$, Banach spaces Y,Z, $\varphi: X \to Z$

$\varphi': Y \to Z$ extends φ

$$e(X,Y,Z) = \sup_{\varphi} \inf_{\varphi'} \{\|\varphi'\|_{Lip} \,/\, \|\varphi\|_{Lip}\}$$

Problem: what's $e_n(Y,Z) = \sup_{|X|=n} \{e(X,Y,Z)\}$?

$e_n(Y,L_2) = O(\sqrt{\log n})$ [JL '84]

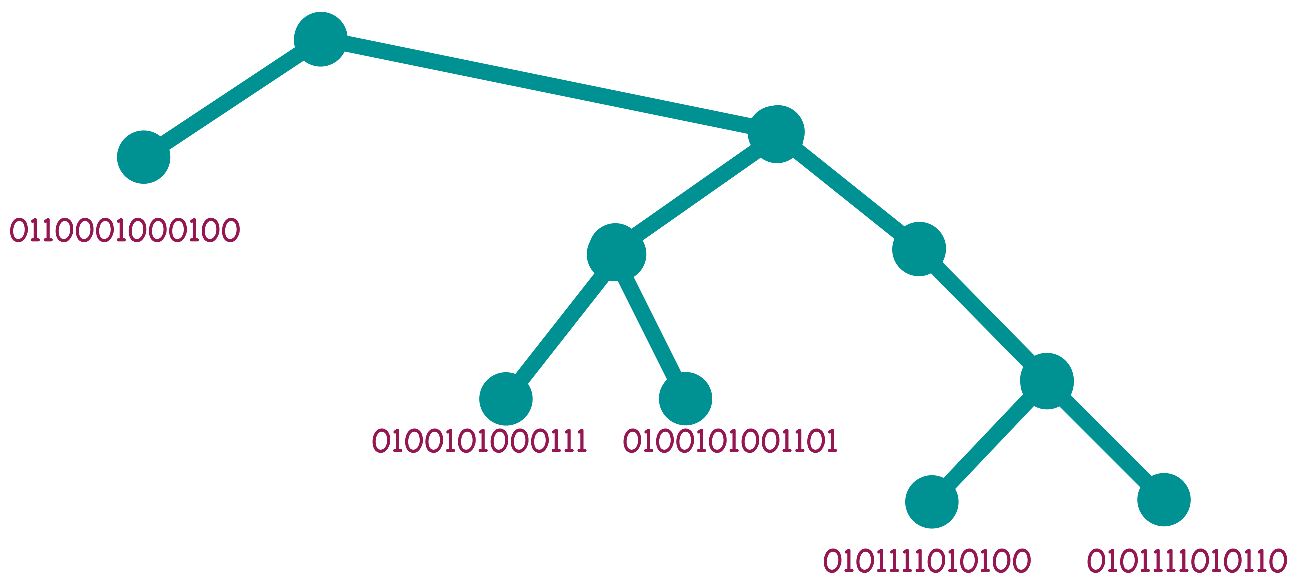
$e_n(Y,Z) = O(\log n)$

$e_n(Y,Z) = \Omega(\sqrt{\log n/\log\log n})$

(the last two: [JLS '86])

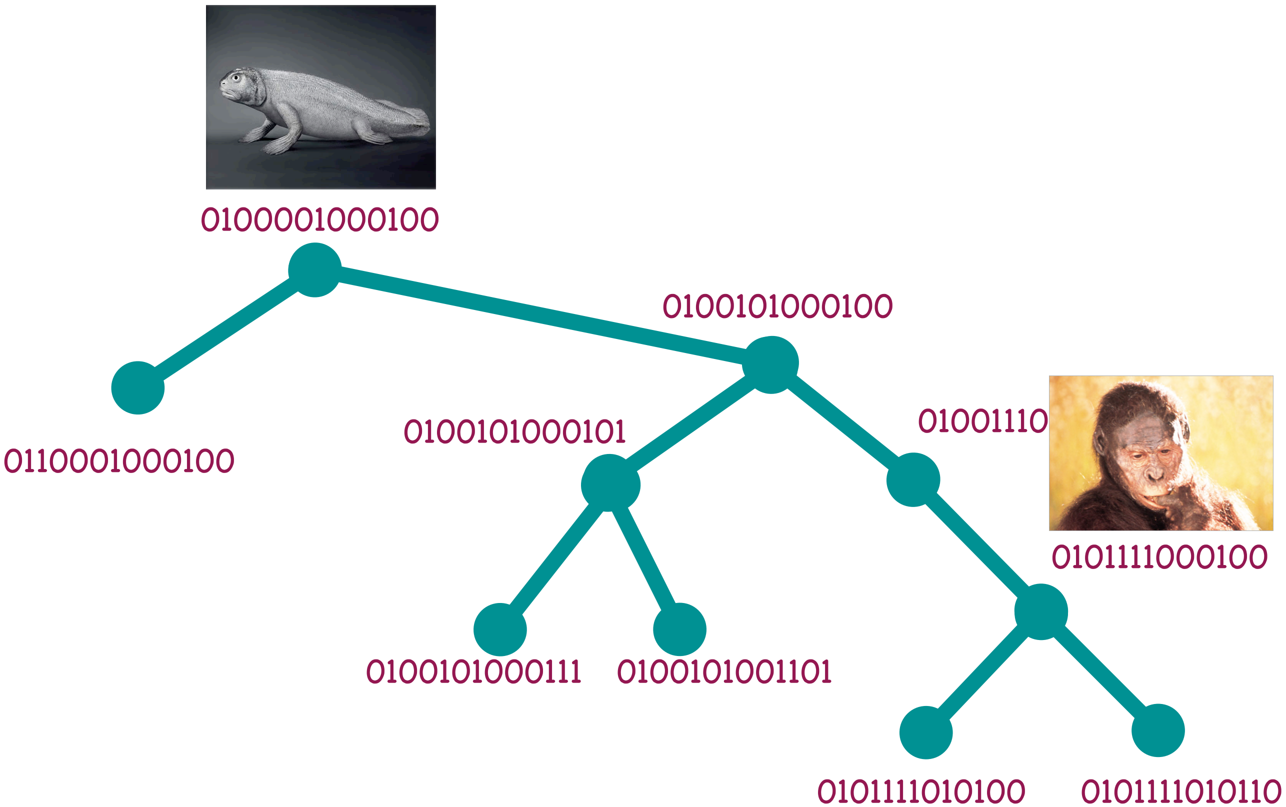
motivation - DNA multiple sequence alignment

Input: evolutionary tree, DNA seq on leaves
Output: hypothesized ancestors sequences



motivation - DNA multiple sequence alignment

Input: evolutionary tree, DNA seq on leaves
Output: hypothesized ancestors sequences



multiway cut

Input: graph $G=(V,E)$, terminal set $T \subset V$ | $k = |T|$

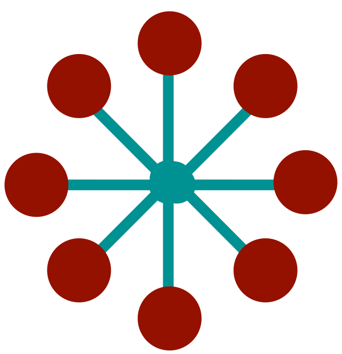
Output: $\varphi: V \to T$ extending $id: T \to T$

Objective: min #edges uv w/ $\varphi(u) \neq \varphi(v)$

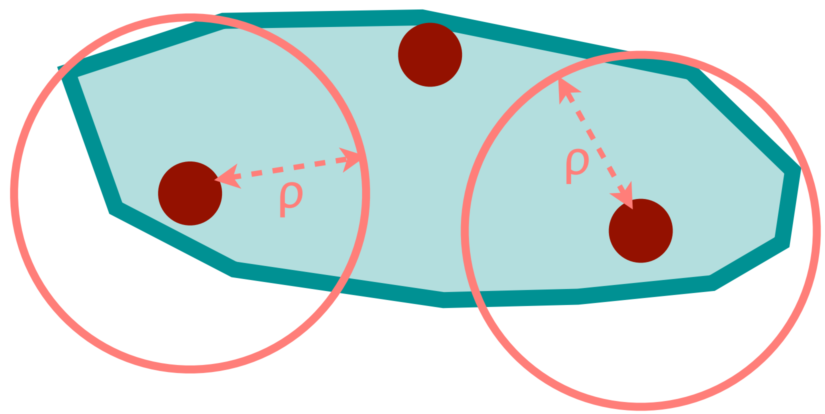
<u>metric relaxation</u>: minimize $\Sigma_{uv \in E}\, d(u,v)$ s.t.

d is a metric on V and d is uniform on T.

$\Rightarrow$ $(2-2/k)$-approx., matching [DJPSY '94].



bad example



$\rho < \frac{1}{2}$

rounding algorithm

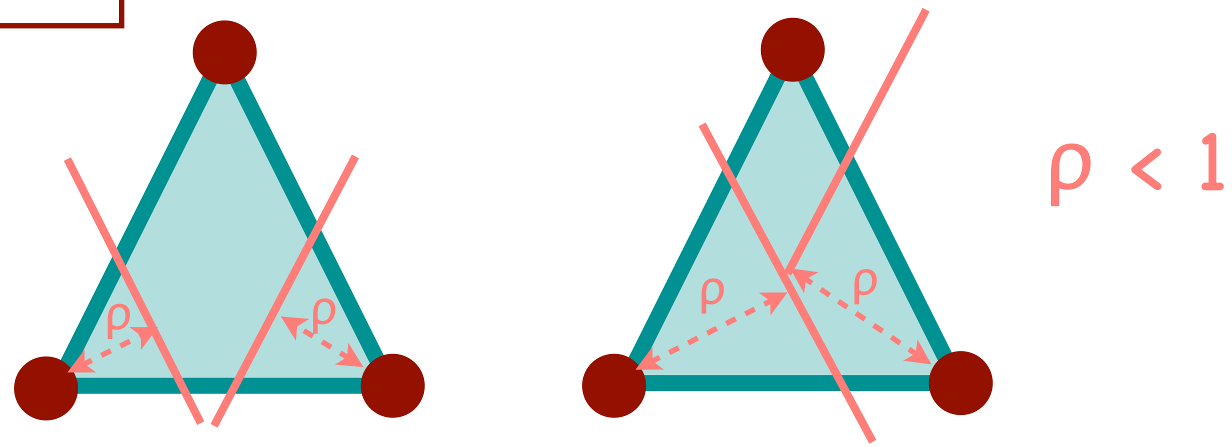
multiway cut (cont.)

transportation relaxation [CKR '98]

minimize $\sum_{uv \in E} \frac{1}{2} \cdot \|\theta(u)-\theta(v)\|_1$ ← statistical distance

s.t. $\theta: V \to \Delta^k$ maps T to the vertices

(k-1)-simplex



ratio in [8/7-o(1), 1.349) [FK '00, KKSTY '99]

for k=3 ratio is 12/11 [KKSTY '99, CT '99]

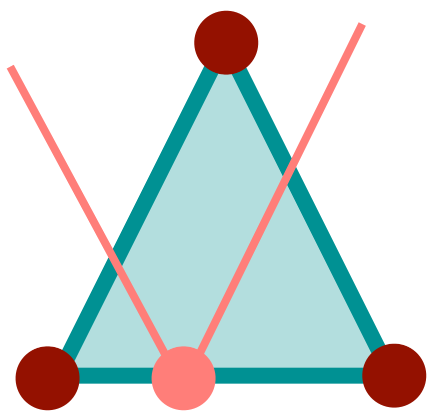
uniform labeling

Input: graph $G=(V,E)$, label set L, $a:V \to 2^L$

$k = |L|$

Output: $\varphi:V \to L$ s.t. $\varphi(u) \in a(u)$

Objective: min #edges uv w/ $\varphi(u) \neq \varphi(v)$



ratio in $[2-2/k, 2]$ [KT '99]

for $k=3$ ratio is $4/3$ [Chuzhoy '00]

0-extension [Karzanov '98]

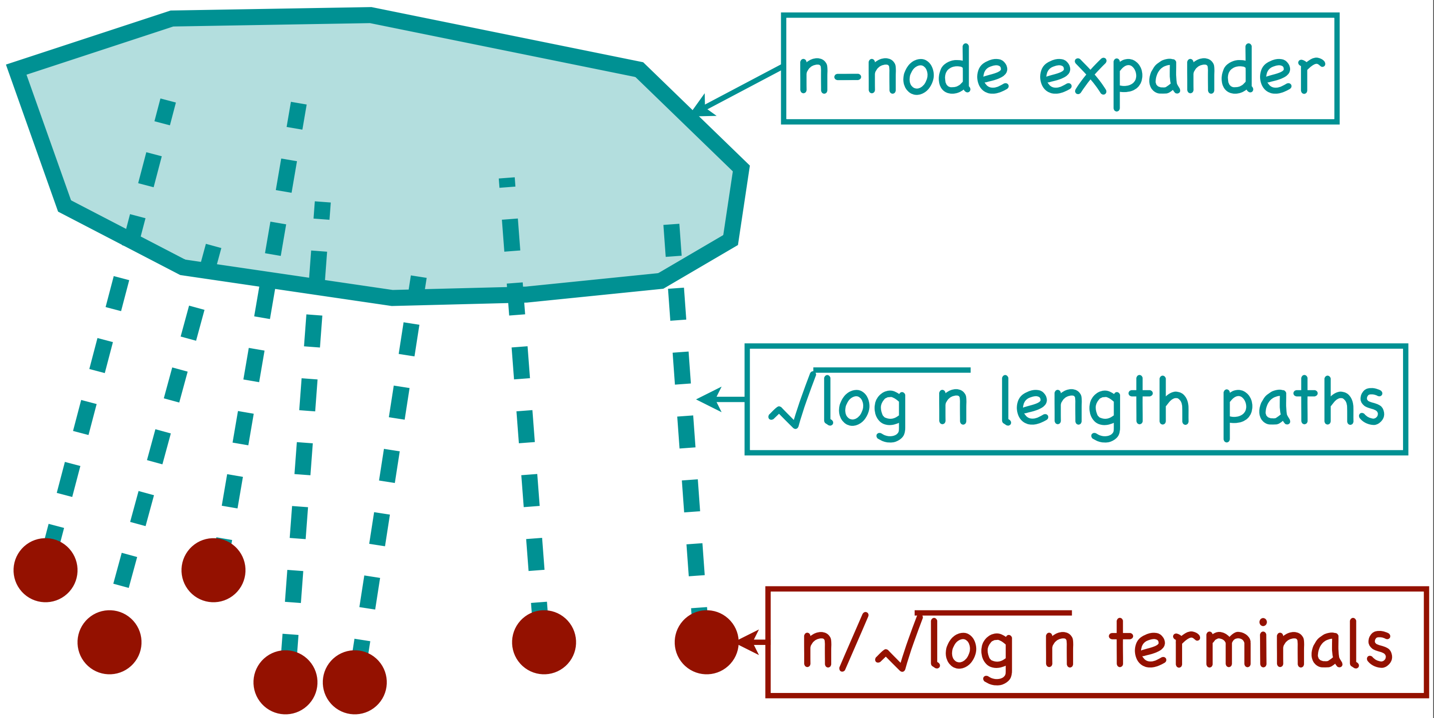
Input: $G=(V,E)$, $T \subset V$, metric d on T

Output: $\varphi: V \rightarrow T$ extending $id: T \rightarrow T$

Objective: minimize $\Sigma_{uv \in E}\, d(\varphi(u),\varphi(v))$

metric relaxation: minimize $\Sigma_{uv \in E}\, d'(u,v)$ s.t. d' is a metric on V that extends d.

[CKR '00]: $\Omega(\sqrt{\log |T|})$



padded decompositions (graph version)

graph $G=(V,E)$, shortest path metric d_G
distribution Pr over injections $\varphi:V\to V$
$\mathrm{diam}(\varphi,u) = \mathrm{diam}(\varphi^{-1}(u))$
$\mathrm{diam}(\varphi) = \max_{u\in V} \mathrm{diam}(\varphi,u)$

Objective 1: $T \subset V$, $|T| = k$, $\varphi:V\to T$ (id on T)
min $\max_{uv\in E}$ $E[\chi(\varphi(u)\neq\varphi(v))\cdot \mathrm{diam}(\varphi,u)]$

CKR decompositions [CKR '00]:
pick u.a.r perm. σ on T, factor $\rho \in [1,s]$
$\varphi(u)$ = first $t\in T$ s.t. $d_G(u,t) \leq \rho\cdot\Delta_u$

Thm [FHRT '03]: $E[\chi(\varphi(u)\neq\varphi(v))\cdot \mathrm{diam}(\varphi,u)]$
$= O(\log k/\log\log k)$ (for $s = \log k/\log\log k$)

padded decompositions (graph version)

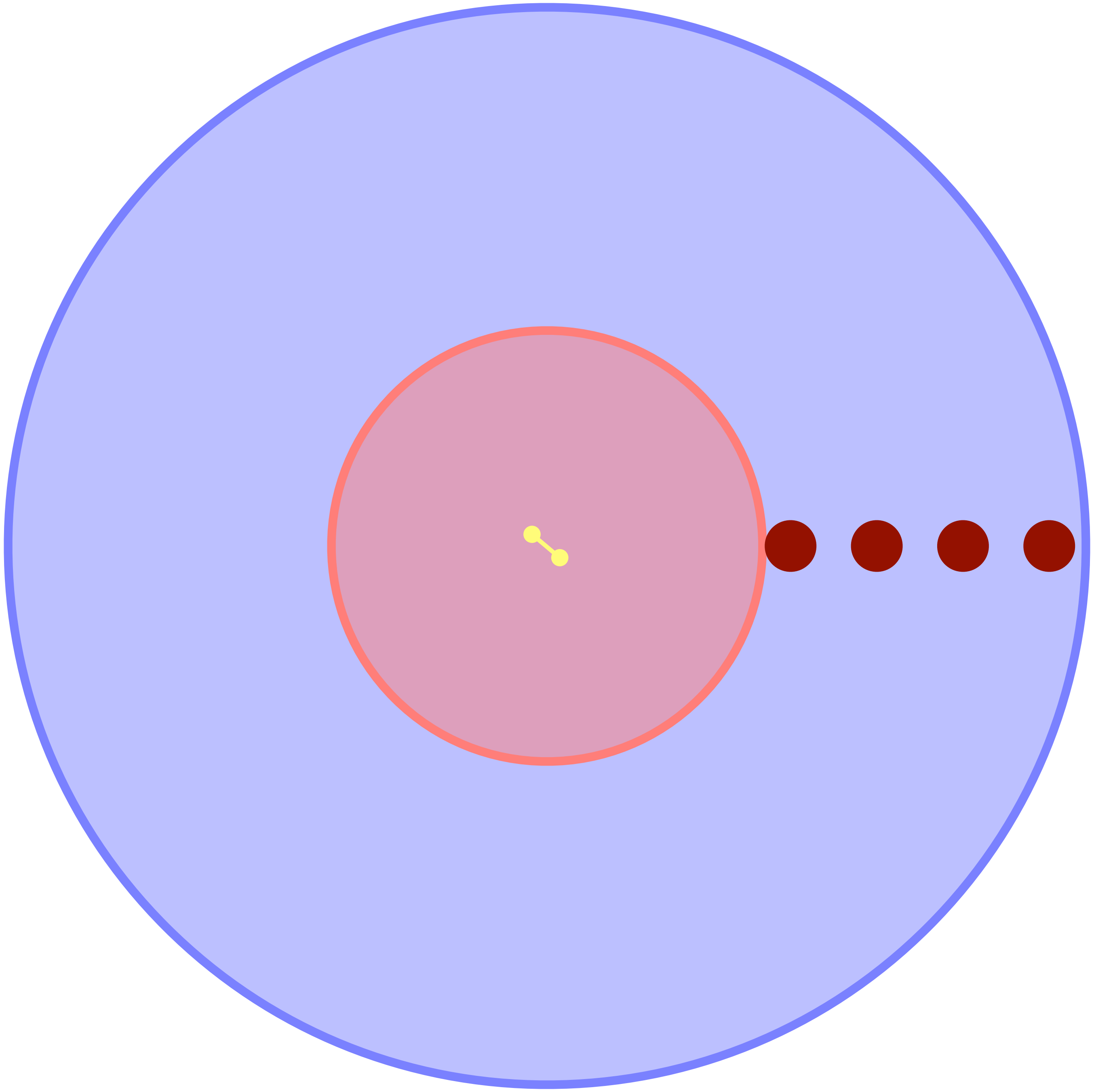
graph $G=(V,E)$, shortest path metric d_G
distribution Pr over injections $\varphi:V\to V$
$\mathrm{diam}(\varphi,u) = \mathrm{diam}(\varphi^{-1}(u))$
$\mathrm{diam}(\varphi) = \max_{u\in V} \mathrm{diam}(\varphi,u)$

Objective 1: $T \subset V$, $|T| = k$, $\varphi:V\to T$ (id on T)
$\min \max_{uv\in E} E[\chi(\varphi(u)\neq\varphi(v))\cdot \mathrm{diam}(\varphi,u)]$

CKR decompositions [CKR '00]:
pick u.a.r perm. σ on T, factor $\rho \in [1,s]$
$\varphi(u)$ = first $t\in T$ s.t. $d_G(u,t) \leq \rho \cdot \Delta_u$ ← $\Delta_u = d_G(u,T)$

Thm [FHRT '03]: $E[\chi(\varphi(u)\neq\varphi(v))\cdot \mathrm{diam}(\varphi,u)]$
$= O(\log k/\log\log k)$ (for $s = \log k/\log\log k$)

the idea in a nutshell



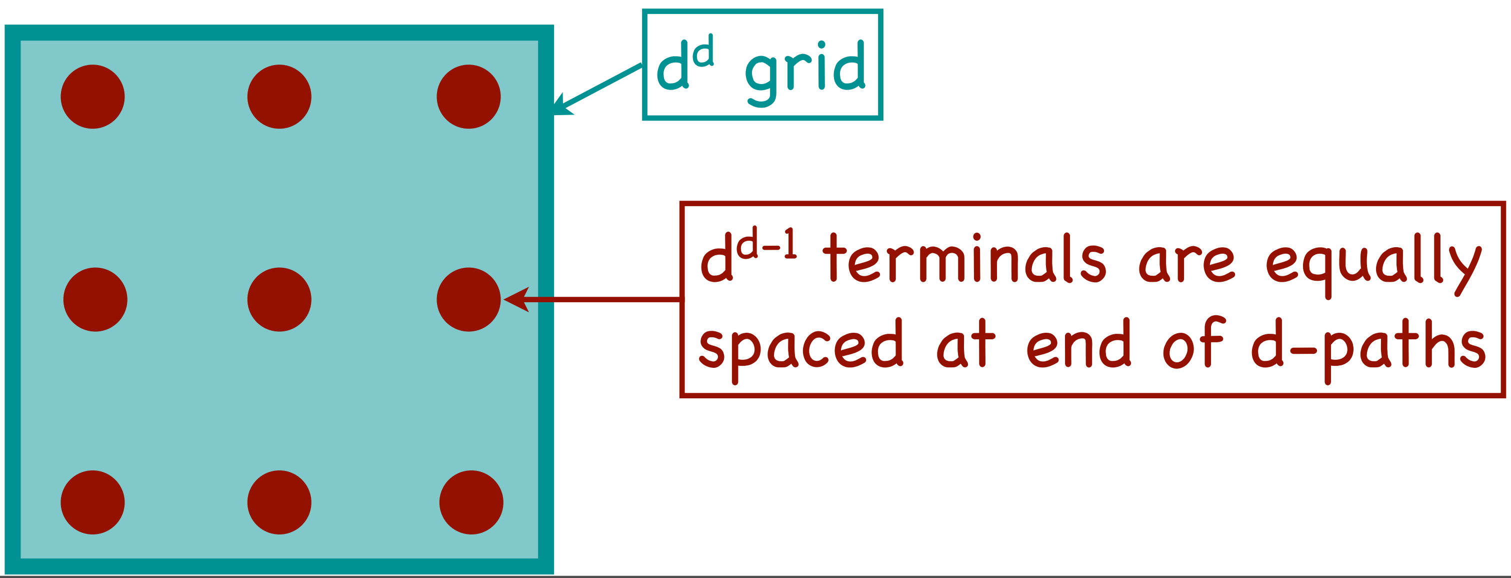
further observations on 0-extension

- $O(\sqrt{diam(T)})$-approximation [CKR '00]:

pick u.a.r $\rho \in [\sqrt{diam(T)}, 2\cdot\sqrt{diam(T)}]$

$$\varphi(u) = \begin{cases} u & \text{if } u \in T \\ \text{closest terminal} & \text{if } d_G(u,T) < \rho \\ \text{terminal 1} & \text{otherwise} \end{cases}$$

- $O(\log k/\log\log k)$ is tight for minimizing $\Sigma_{cut}\,[diam(\varphi,u) + diam(\varphi,v)]$:



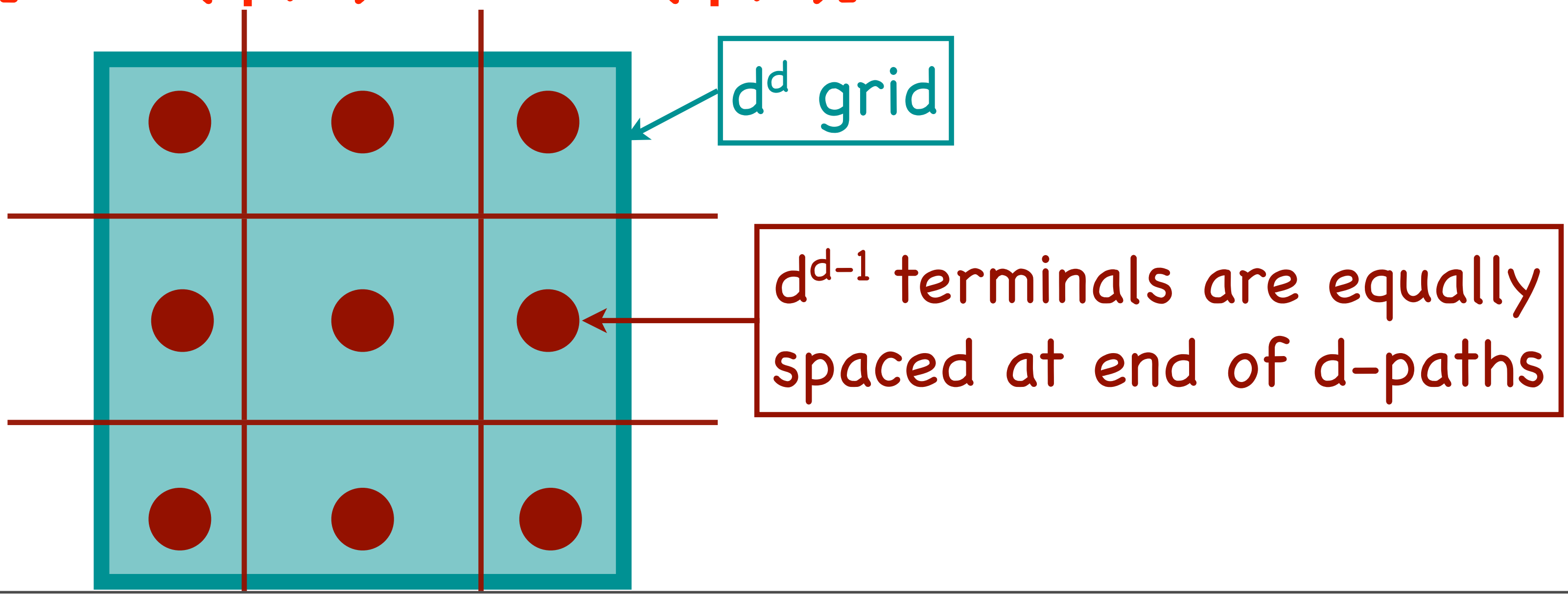
further observations on 0-extension

- $O(\sqrt{diam(T)})$-approximation [CKR '00]:

pick u.a.r $\rho \in [\sqrt{diam(T)}, 2 \cdot \sqrt{diam(T)}]$

$$\varphi(u) = \begin{cases} u & \text{if } u \in T \\ \text{closest terminal} & \text{if } d_G(u,T) < \rho \\ \text{terminal 1} & \text{otherwise} \end{cases}$$

- $O(\log k/\log\log k)$ is tight for minimizing Σ_{cut} [diam(φ,u) + diam(φ,v)]:



more on decompositions

Objective 2: $\varphi: V \to V$

min $\max_{uv \in E}$ $\Pr[\varphi(u) \neq \varphi(v)]$ s.t. $\mathrm{diam}(\varphi) \leq r$

max $\min_u$ $\Pr[N(u) \subset \varphi^{-1}(\varphi(u))]$ s.t. $\mathrm{diam}(\varphi) \leq r$

min $\mathrm{diam}(\varphi)$ s.t. $\forall u$, $\Pr[N(u) \subset \varphi^{-1}(\varphi(u))] \geq ½$

Thm [MN '06]:

$\Pr[N(u) \subset \varphi^{-1}(\varphi(u))] \geq (|B(u,r/8)| \,/\, |B(u,r)|)^{16/r}$

* Improves 1 – $(1/r) \cdot \log(|B(u,r)| \,/\, |B(u,r/8)|)$

⇒ Thm [FRT '04]: any finite metric (X,d) embeds into a convex combination of dominating HSTs with distortion $O(\log |X|)$.

metric labeling [KT '99]

Input: $G=(V,E)$, metric space (L,d), $a:V\to 2^L$ $k = |L|$

Output: $\varphi:V\to L$ s.t. $\varphi(u) \in a(u)$

Objective: minimize $\Sigma_{uv\in E}\, d(\varphi(u),\varphi(v))$

$O(\log k)$ approximation [KT '99, FRT '04]

embeds (L,d) into HSTs, then const. factor approximation for HSTs

transportation relaxations

transportation relaxation [CKNZ '01]

minimize $\sum_{uv \in E} \mathrm{Tran}_d(\theta(u),\theta(v))$ ← transportation metric

s.t. $\theta: V \to \Delta^k$ maps u to $\mathrm{span}(a(u))$

bad example [KKMR '06]:

L = nodes of a k-node expander H, $d = d_H$

$V(G) = \{uv : u,v \in L, u \neq v\}$

$E(G) = \{uv\text{-}uv' : vv' \in E(H)\}$

$\theta(uv)$ = uniform distribution on $\{u,v\}$

$\Rightarrow \Omega(\log |L|)$ integrality gap

similar construction $\Rightarrow \Omega(\sqrt{\log |L|})$

integrality gap for 0-extension

hardness results

metric labeling:

$\Omega(\sqrt{\log |L|})$ unless NP ⊂ quasi-P [CN '04]

k provers, ∀ pair {i,j}, i gets a clause C_{ij} and j gets a variable $x_{ij} \in C_{ij}$

0-extension:

$\Omega(\sqrt[4]{\log |L|})$ (similar argument) [KKMR '06]

all problems:

unique games-hard to approximate beyond the transportation relaxation integrality gap [MNRS '08]

label extension [RK '98]

Input: tree $T=(V,E)$, leaf labels in $\{0,1\}^k$
Output: $L:V\rightarrow \{0,1\}^k$ (unchanged on leaves)
Objective: minimize $\max_{uv\in E} H(L(u),L(v))$

Trivial lower bound:
$\max\{1, \max\{H(L(u),L(v))/d_T(u,v) : u,v \text{ leaves}\}\}$

Thm [Matoušek '90]: $e(d_T,Z) = O(1)$
(in particular, for $Z = [0,1]^k$)

$\Rightarrow$ $O(1)$ approximation [RSS '08]

label completion/selection [RSS '08]

Input: $G=(V,E)$, $L:V \to \{0,1,*\}^k$

Output: $L':V \to \{0,1\}^k$ consistent with L

Objective: minimize $\max_{uv \in E} H(L'(u),L'(v))$

obvious relaxation:

minimize $\max\{1, \max_{uv \in E} \|L'(u) - L'(v)\|_1\}$ s.t.

$L': V \to [0,1]^k$ is consistent with L

$O(\log^2 k)$-approx. for trees
$\Omega(\log k)$ integrality gap for trees
$O(\log |E| / \log\log |E|)$-approx.

} [RSS '08]

applications of padded decompositions

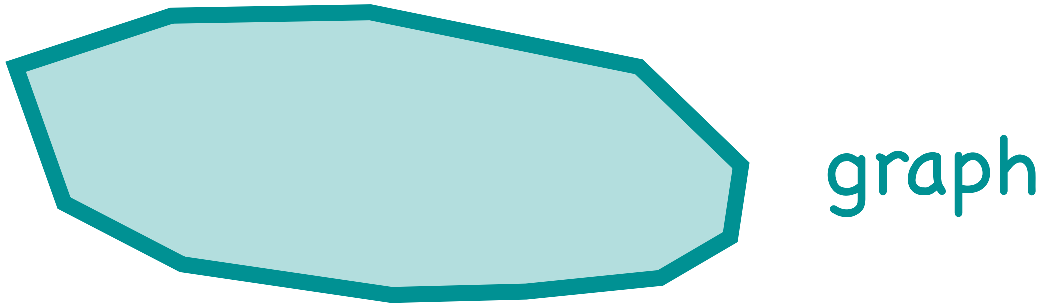
• measured descent [KLMN '04]:
an n-point λ-doubling metric embeds into ℓ_2 w/distortion $O(\sqrt{\lambda \log n})$

• Lipschitz extension [LN '05]:
$e_n(Y,Z) = O(\log n/\log\log n)$

• metric Ramsey properties [MN '06]:
an n-point metric has an $n^{1-\varepsilon}$-point subset that embeds into ℓ_2 w/distortion $O(1/\varepsilon)$

• local-global tradeoffs [CMM '07]:
every m-point subset of an n-point metric M embeds into ℓ_1 w/distortion c
$\Rightarrow$ M embeds into ℓ_1 w/dist. $O(c \cdot \log(n/m))$

vertex sparsifiers [Moitra '09, LM '10]



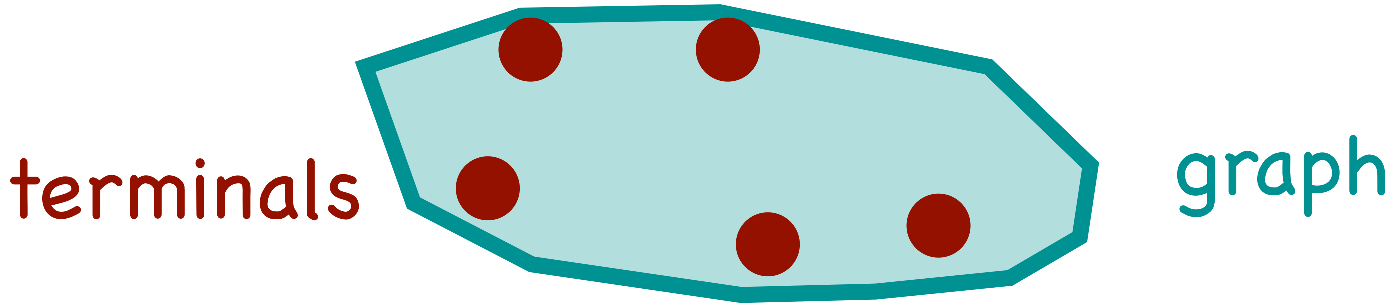
$O(\log k/\log\log k)$* cut (also flow) sparsifiers [CLLM '10, MM '10, EGKRTT '10]

$\Omega(\sqrt{\log k} / \log\log k)$ lower bound [MM '10]

does $e_n(\ell_2,\ell_1)$ = const.? [Ball '92] $\Rightarrow$
$e_n(\ell_1,\ell_1) = \tilde{O}(\sqrt{\log n}) \Rightarrow \tilde{O}(\sqrt{\log n})$ cut spars.

* integrality gap for the metric relaxation for 0-extension

vertex sparsifiers [Moitra '09, LM '10]



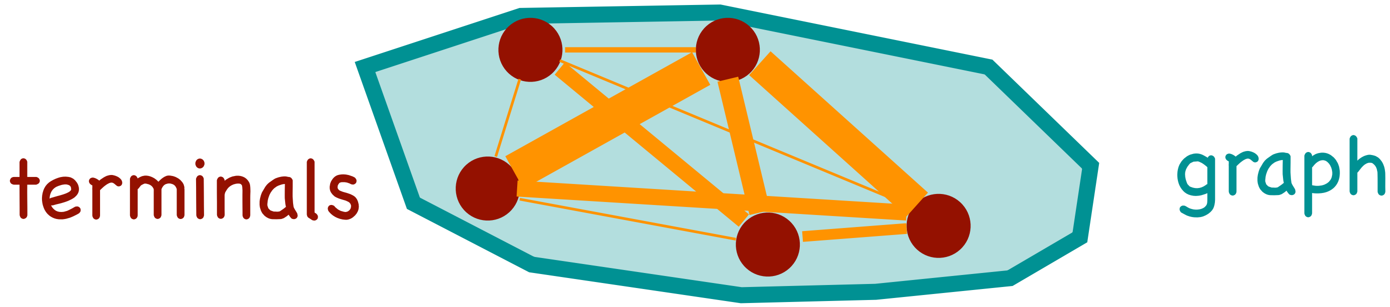
$O(\log k/\log\log k)$* cut (also flow) sparsifiers [CLLM '10, MM '10, EGKRTT '10]

$\Omega(\sqrt{\log k} / \log\log k)$ lower bound [MM '10]

does $e_n(\ell_2,\ell_1)$ = const.? [Ball '92] ⇒
$e_n(\ell_1,\ell_1) = \tilde{O}(\sqrt{\log n}) \Rightarrow \tilde{O}(\sqrt{\log n})$ cut spars.

* integrality gap for the metric relaxation for 0-extension

vertex sparsifiers [Moitra '09, LM '10]



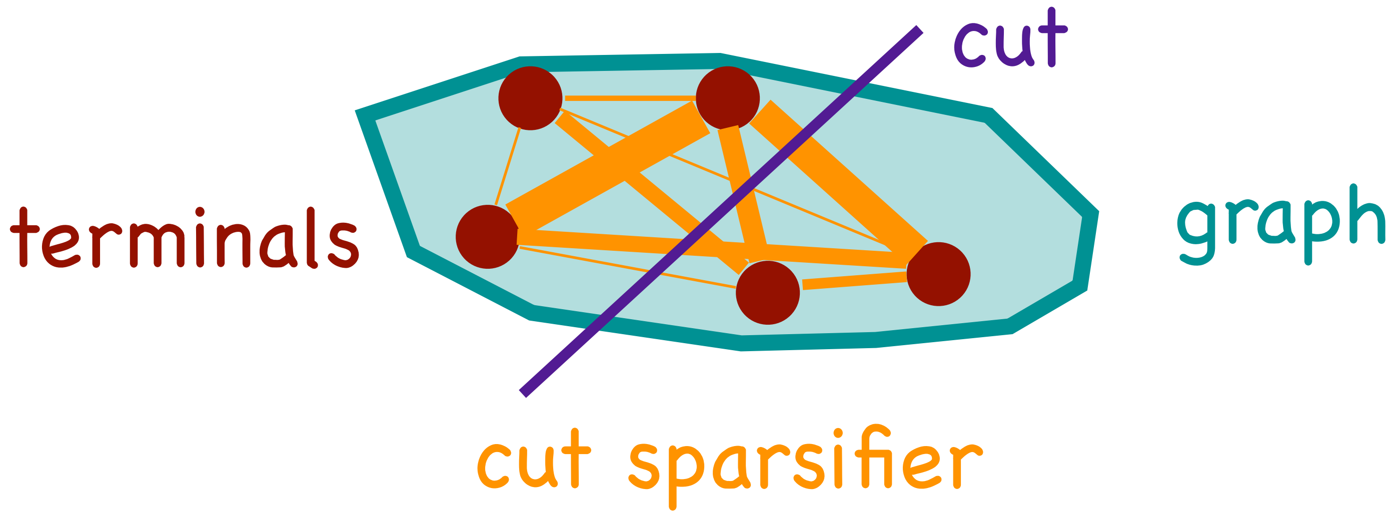
O(log k/loglog k)* cut (also flow) sparsifiers [CLLM '10, MM '10, EGKRTT '10]

$\Omega(\sqrt{\log k} / \log\log k)$ lower bound [MM '10]

does $e_n(\ell_2,\ell_1)$ = const.? [Ball '92] ⇒
$e_n(\ell_1,\ell_1) = \tilde{O}(\sqrt{\log n}) \Rightarrow \tilde{O}(\sqrt{\log n})$ cut spars.

* integrality gap for the metric relaxation for 0-extension

vertex sparsifiers [Moitra '09, LM '10]



$O(\log k/\log\log k)$* cut (also flow) sparsifiers [CLLM '10, MM '10, EGKRTT '10]

$\Omega(\sqrt{\log k} / \log\log k)$ lower bound [MM '10]

does $e_n(\ell_2,\ell_1)$ = const.? [Ball '92] ⇒ $e_n(\ell_1,\ell_1) = \tilde{O}(\sqrt{\log n}) \Rightarrow \tilde{O}(\sqrt{\log n})$ cut spars.

* integrality gap for the metric relaxation for 0-extension

Thank you!