Optimal Embedding into Star Metrics

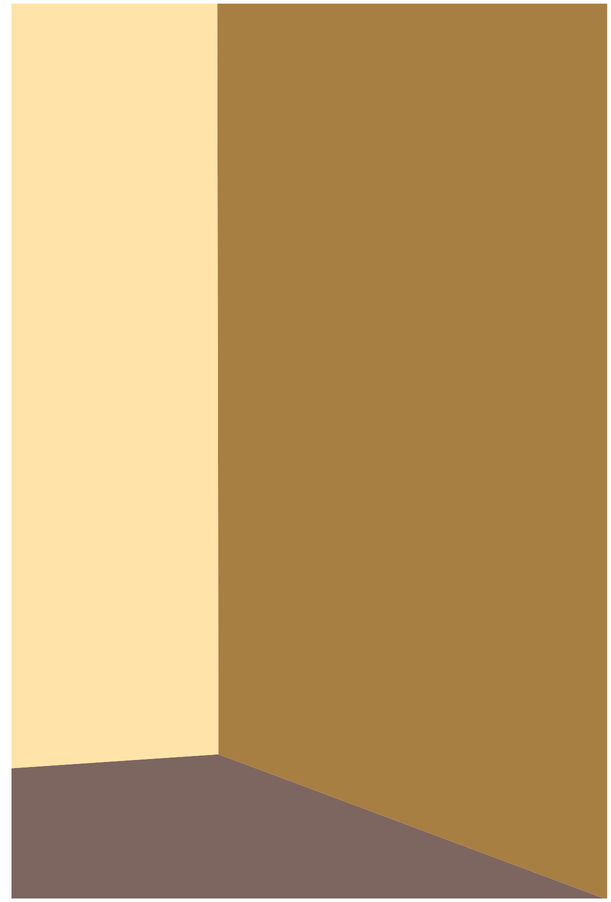
David Eppstein
Univ. of California, Irvine
Computer Science Department

Kevin Wortman
California State University, Fullerton

Metric embedding problems

Input: complicated metric space

Quadratic # degrees of freedom
e.g. distance matrix
shortest paths in weighted graph
n-dimensional L-infinity space



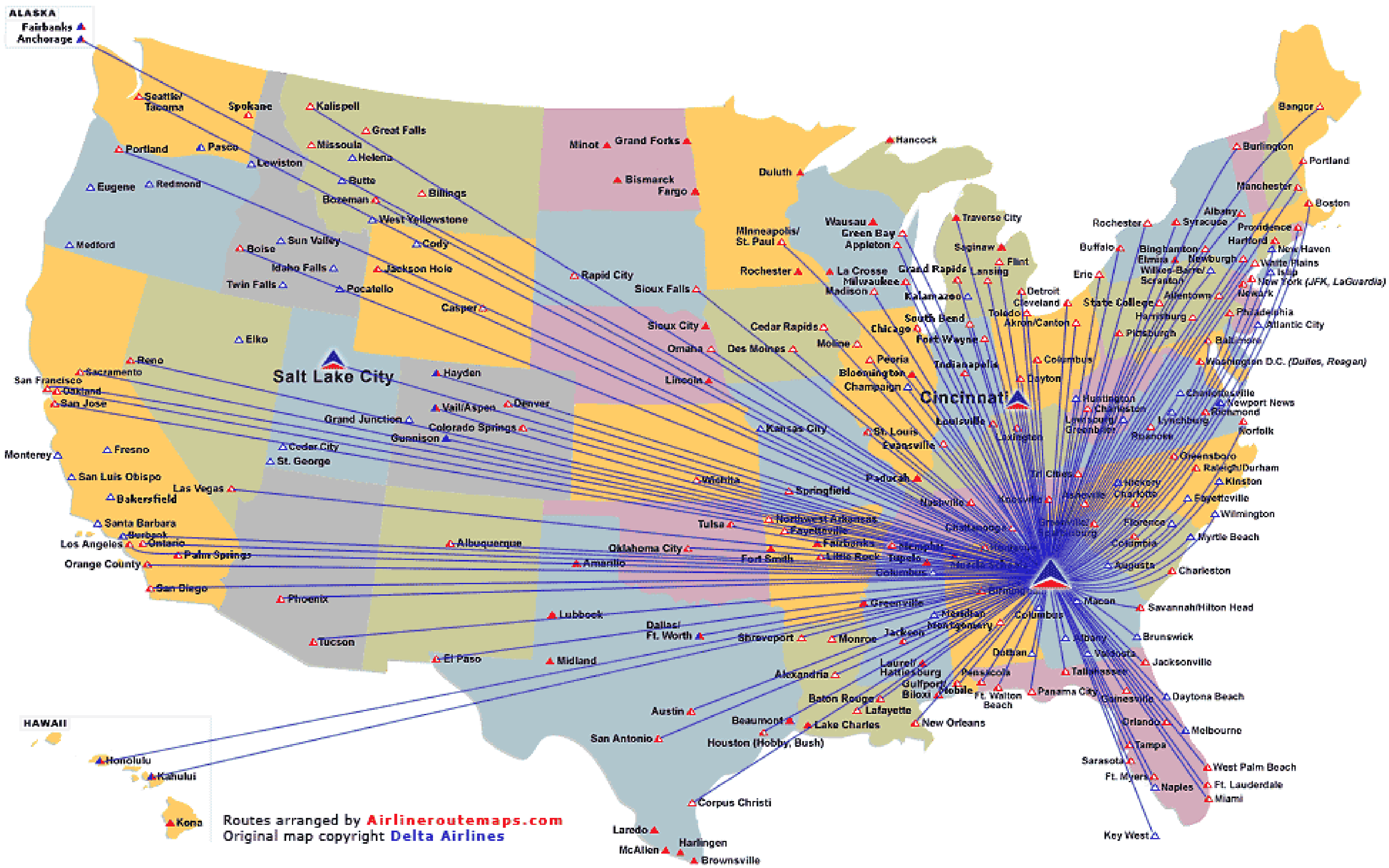
Output: simple metric space

As close as possible to
the input metric space

Linear # degrees of freedom

Star metrics

Defining property: there exists a hub h that is between every other pair of points



For all s and t, d(s,t) = d(s,h) + d(h,t)

Once distances from the hub are known, all other distances are determined

How to measure the quality of a metric embedding?

Dilation: how much farther you would have to travel in the new metric

Scale the new metric so that all distances are at least as large as they were in the old metric

Dilation = maximum ratio
new distance / old distance
among all pairs of input points

(factor by which some distances grow)

Minimum dilation star problem

Given an n by n matrix D[s,t] of distances in a metric space

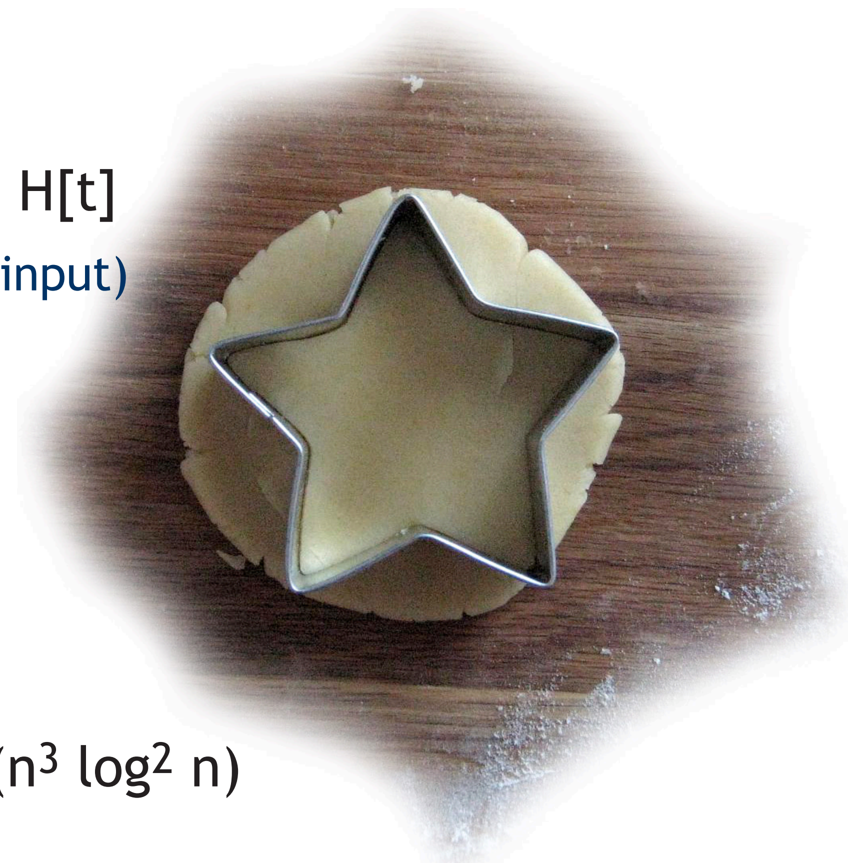
Produce a vector H[s] of distances
from each input point to a new hub

Satisfying, for all s and t, $D[s,t] \leq H[s] + H[t]$

(scale output distances to be at least as big as input)

Minimize $\max_{s,t} (H[s] + H[t])/D[s,t]$

(find the star with optimal dilation)



We solve this in **strongly polynomial time** $O(n^3 \log^2 n)$

Related Work

Metric embedding into L1, L2, etc. with guaranteed dilation bounds
[Large literature]

Does not find optimal embedding

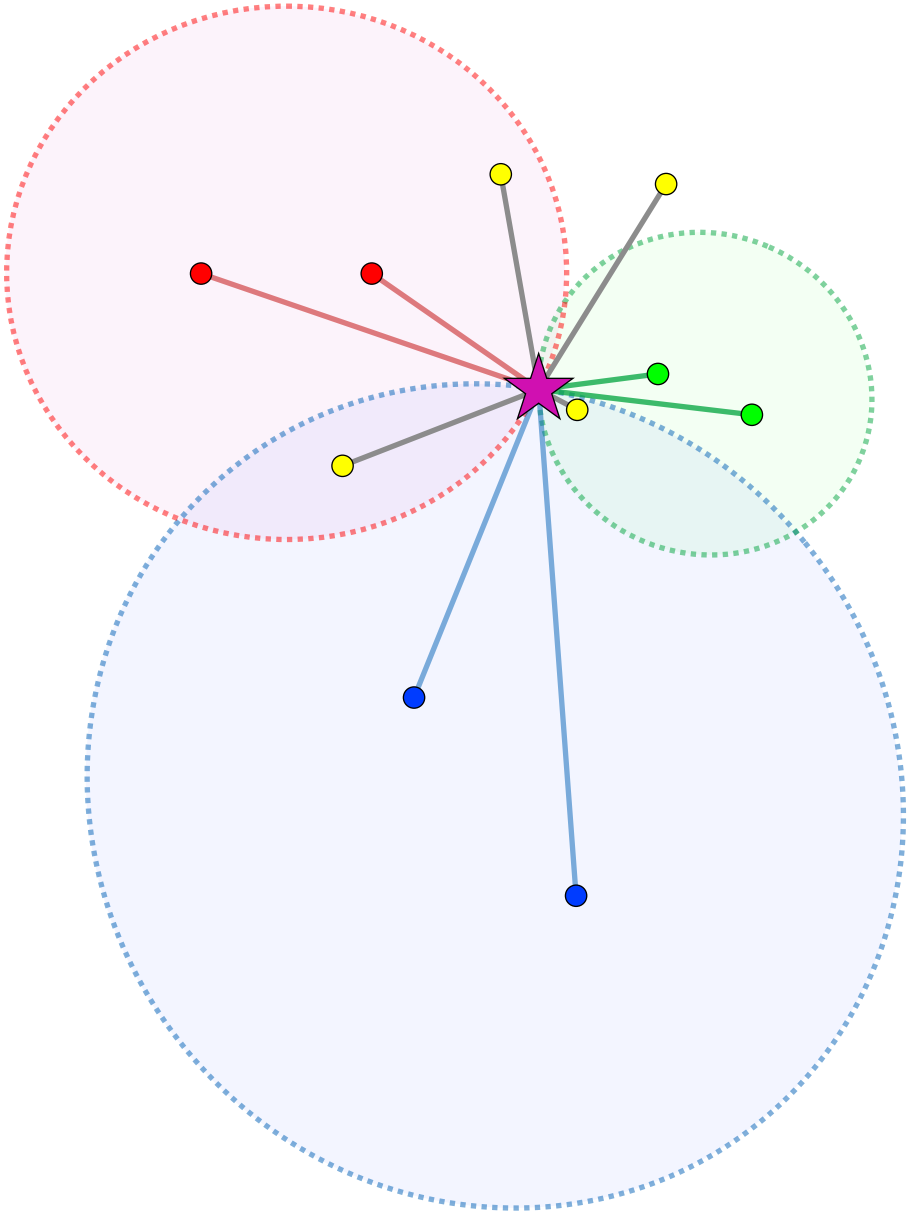
Optimal embedding into ultrametric
minimizing maximum difference (new distance - input distance)
[Farach, Kannan, Warnow, Algorithmica 1995]

Insensitive to distortion of small distances

Optimal embedding of unweighted graph to line
[Fomin, Lokshtanov, Saurabh, WG 2009]

Singly exponential time (at least not factorial)

Related work: Euclidean min dilation star [E & W, SoCG'05]



Given n points in Euclidean plane or higher dimensional space

Find a hub within that space minimizing dilation of star network with Euclidean distances as lengths

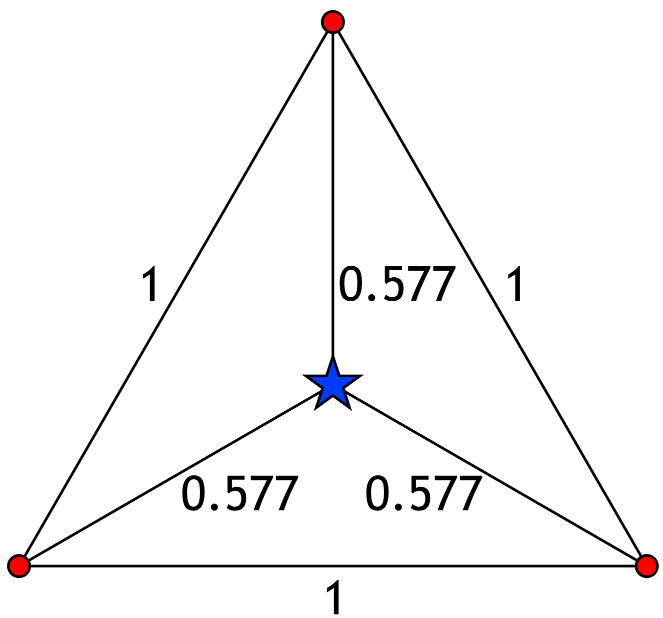
Equivalently, maximize min eccentricity of ellipses passing through the hub having pairs of input points as foci

$O(n \log n)$ in any fixed dimension if hub can be any point of the space

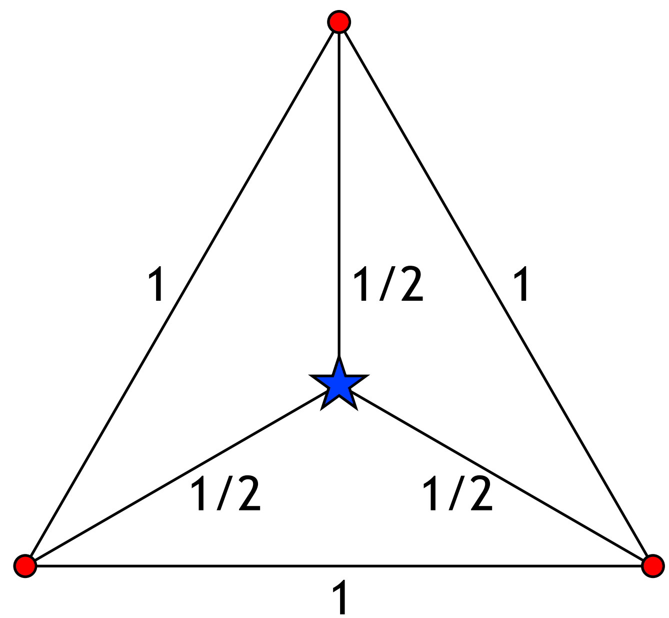
$O(n\, 2^{\alpha(n)} \log^2 n)$ if hub must be one of the input points (2d only)

The difference between geometric and metric stars

Example: input is an equilateral triangle
(three points with equal distances)

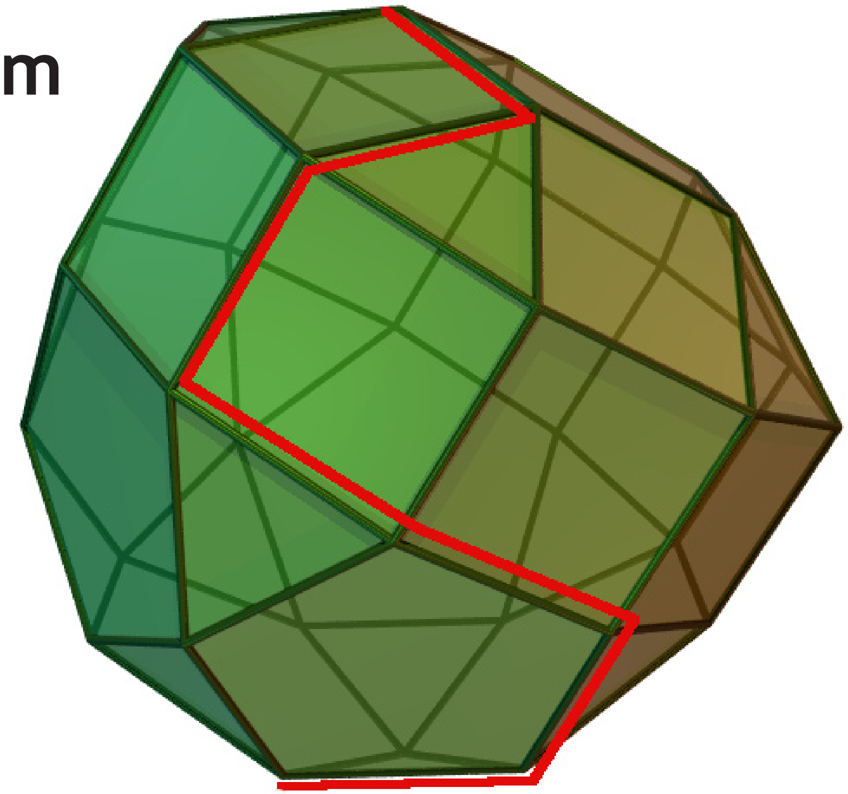


Euclidean min dilation star
dilation = 1.155



Metric min dilation star
dilation = 1

Solution ideas, I:
Express MDS as a linear program



Find H[x] and Δ

Satisfying

H[x] ≥ 0

H[x] + H[y] ≥ distance(x,y)

H[x] + H[y] ≤ Δ distance(x,y)

for all x and y

Minimizing Δ

Not in form for known strongly-polynomial LP algorithms

more than O(1) variables total

more than two variables in some inequalities

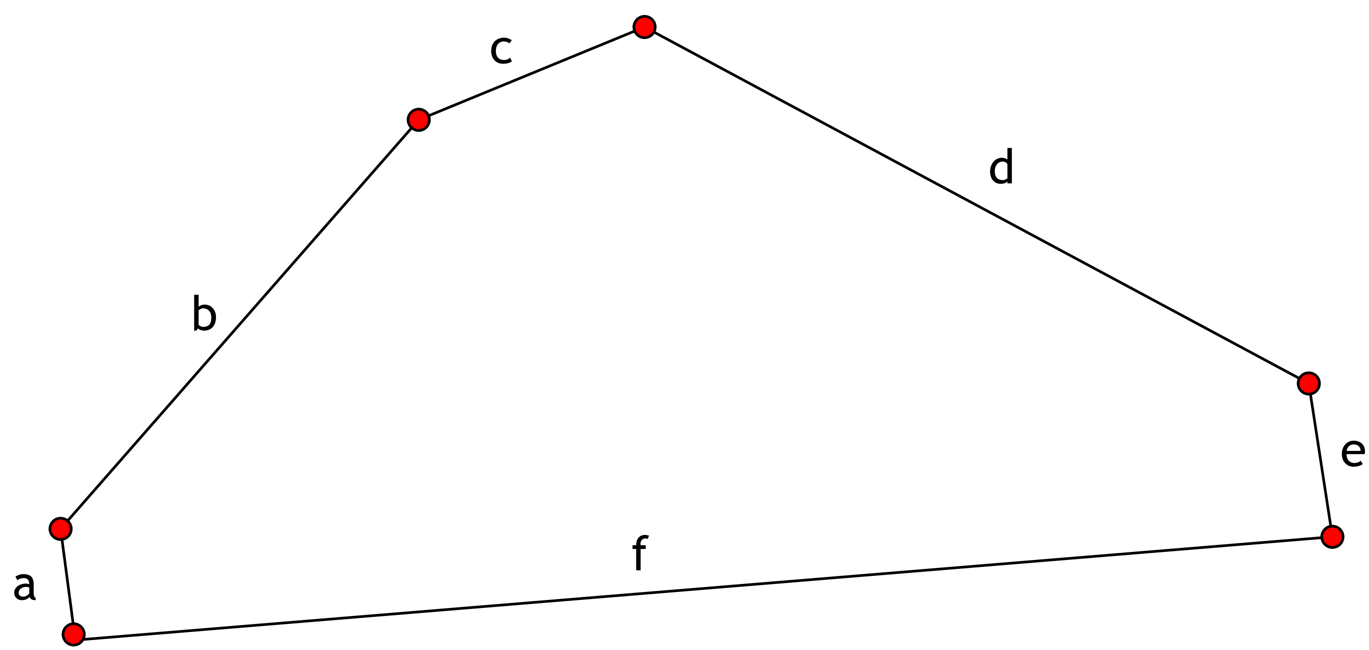
Solution ideas, II: Characterize LP basis

Any even cycle in original metric space lower-bounds the dilation:

dilation ≥ (sum of even edge lengths) / (sum of odd edge lengths)

(because in star, both sums are forced to equal each other)

Optimal dilation turns out to equal worst cycle of this type

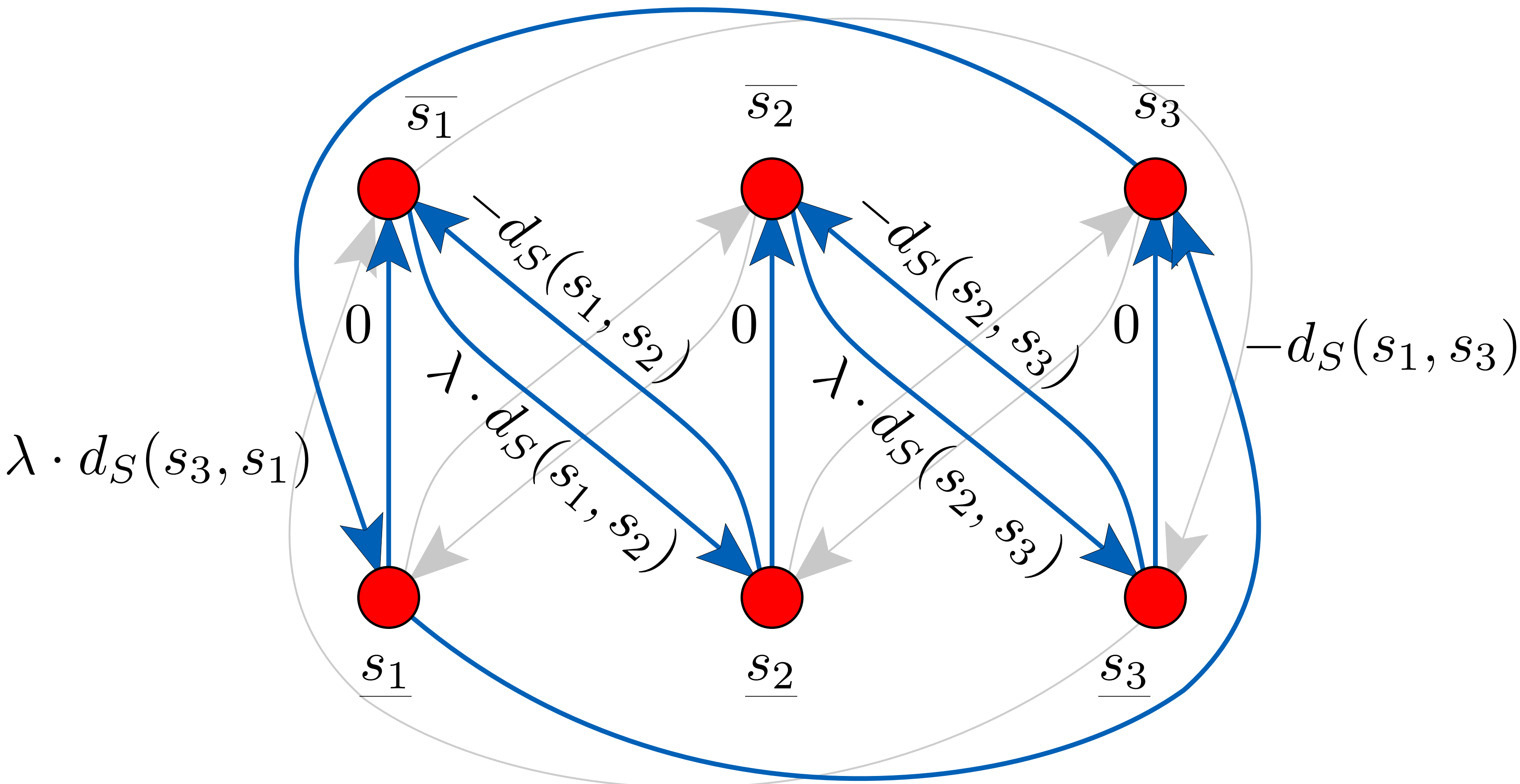


$$\text{dilation} \geq (b + d + f) \;/\; (a + c + e)$$

Solution ideas, III: transform to a graph problem

Parametric negative cycle detection:
digraph where edge weights are linear functions of a parameter λ,
find smallest value of λ such that all cycle lengths are non-negative

Optimal λ = optimal dilation Δ of original metric embedding problem



Solution ideas, IV: strongly polynomial solution for parametric negative cycle detection problem

Megiddo's parametric search [Megiddo, JACM 1983]:

Simulate parallel algorithm for optimization problem
as if it were given the optimal parameter value as input
using a decision algorithm to help simulate each branch step

Simulated algorithm [Savage, Ph.D. thesis 1977]:

Compute all pairs shortest paths by repeated matrix squaring

Decision algorithm for comparing parameter value to optimum:

Bellman–Ford, detect negative cycles in non-parametric graph

Solution ideas, V: cycle detection details

Store a matrix of piecewise linear parametric functions

Represents lengths of paths with at most 2^i hops

Initially: $i = 0$, matrix stores the edge length functions

Repeat log n times, until $2^i > n$:

Square in (min,+) matrix arithmetic to increment i

$O(n^3 \log n)$ time to combine piecewise linear functions

Binary search for the optimal parameter value
among the breakpoints of the path-length functions
simplifying matrix entries back to non-piecewise functions

$O(\log n)$ calls to Bellman-Ford

Solution ideas VI: finding the actual embedding

Now that we know the correct dilation...

Plug it into the same parametric graph

Add an extra "source" vertex to the graph

Compute distances to all other vertices (Bellman-Ford)

H[x] = (half of) difference between two distances
from source to two vertices representing x

Some algebra + triangle inequality shows this is a valid embedding

Lower bound shows this is the optimal embedding

Conclusions and open problems

First known strongly-polynomial minimum-dilation metric embedding

But...
The dilation might not be very good

Can we find optimal dilation
(or within a constant factor of optimal)
for embeddings into metrics such as convex combinations of trees
for which we also know good guarantees on dilation?

Some known hardness results, much more to be done