Gossip in $CAT(\kappa)$ metric spaces

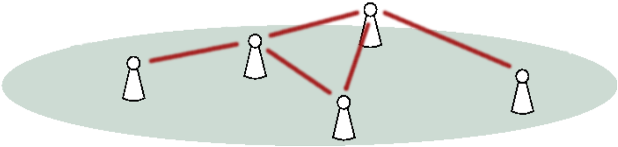
Anass Bellachehab   Jérémie Jakubowicz

Télécom SudParis, Institut Mines-Télécom & CNRS UMR 5157

GSI 2015
Palaiseau October 28

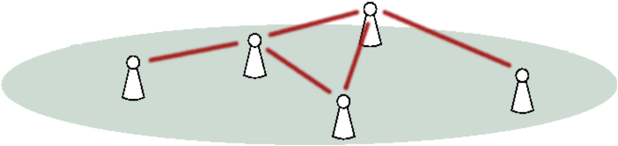
Problem

We consider a network of N agents such that:



- The network is represented by a connected, undirected graph $G = (V, E)$, where $V = \{1, \ldots, N\}$ stands for the set of agents and E denotes the set of available communication links between agents.
- At any given time t an agent v stores stores data represented as an element $x_v(t)$ of a data space $\mathcal{M}$.
- $X_t = (x_1(t), \ldots, x_N(t))$ is the tuple of data values of the whole network at instant t.

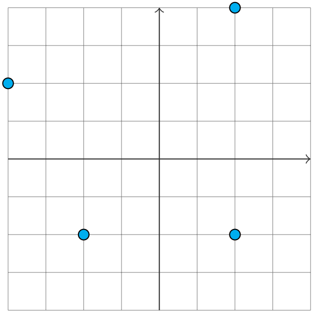
Problem (cont'd)



- Each agent has its own Poisson clock that ticks with a common intensity λ (the clocks are identically made) independently of other clocks.
- When an agent clock ticks, the agent is able to perform some computations and wake up some neighboring agents.

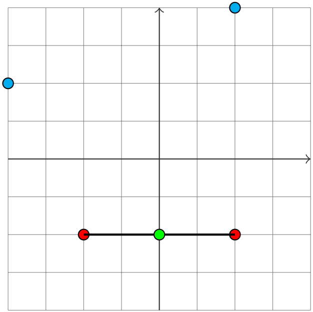
The goal is to take the system from an initial state $X(0)$ to a **consensus state**; meaning a state of the form $X_\infty = (x_\infty, \ldots, x_\infty)$ with: $x_\infty \in \mathcal{M}$.

Random Pairwise Gossip (Xiao & Boyd'04)



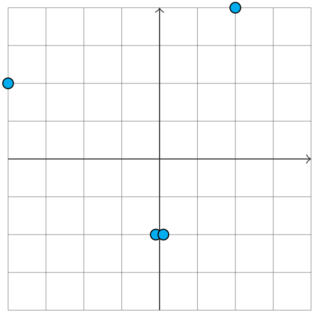
$$x_0 = \begin{pmatrix} -1 & -1 \\ 1 & -1 \\ -2 & 1 \\ 1 & 2 \end{pmatrix}$$

Random Pairwise Gossip (Xiao & Boyd'04)



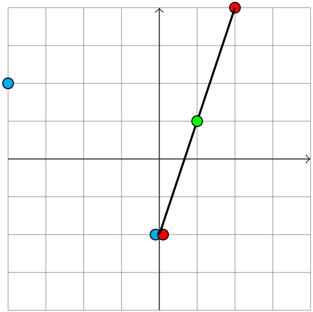
$$x_0 = \begin{pmatrix} -1 & -1 \\ 1 & -1 \\ -2 & 1 \\ 1 & 2 \end{pmatrix}$$

Random Pairwise Gossip (Xiao & Boyd'04)



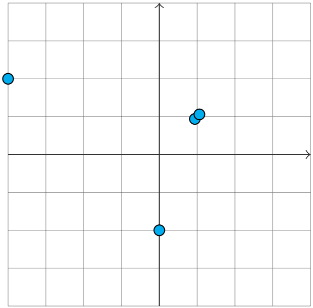
$$x_1 = \begin{pmatrix} 0 & -1 \\ 0 & -1 \\ -2 & 1 \\ 1 & 2 \end{pmatrix}$$

Random Pairwise Gossip (Xiao & Boyd'04)



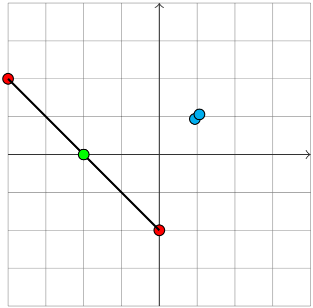
$$x_1 = \begin{pmatrix} 0 & -1 \\ 0 & -1 \\ -2 & 1 \\ 1 & 2 \end{pmatrix}$$

Random Pairwise Gossip (Xiao & Boyd'04)



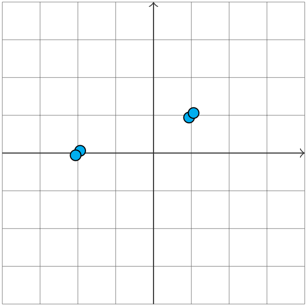
$$x_1 = \begin{pmatrix} 0.5 & 0.5 \\ 0 & -1 \\ -2 & 1 \\ 0.5 & 0.5 \end{pmatrix}$$

Random Pairwise Gossip (Xiao & Boyd'04)



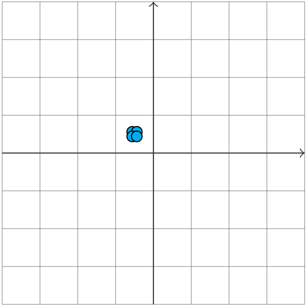
$$x_1 = \begin{pmatrix} 0.5 & 0.5 \\ 0 & -1 \\ -2 & 1 \\ 0.5 & 0.5 \end{pmatrix}$$

Random Pairwise Gossip (Xiao & Boyd'04)



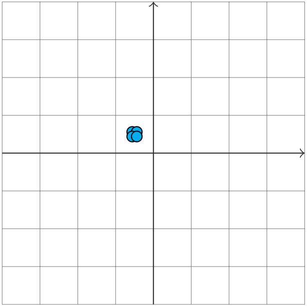
$$x_2 = \begin{pmatrix} 0.5 & 0.5 \\ -1 & 0 \\ -1 & 0 \\ 0.5 & 0.5 \end{pmatrix}$$

Random Pairwise Gossip (Xiao & Boyd'04)



$$x_\infty = \begin{pmatrix} -0.25 & 0.25 \\ -0.25 & 0.25 \\ -0.25 & 0.25 \\ -0.25 & 0.25 \end{pmatrix}$$

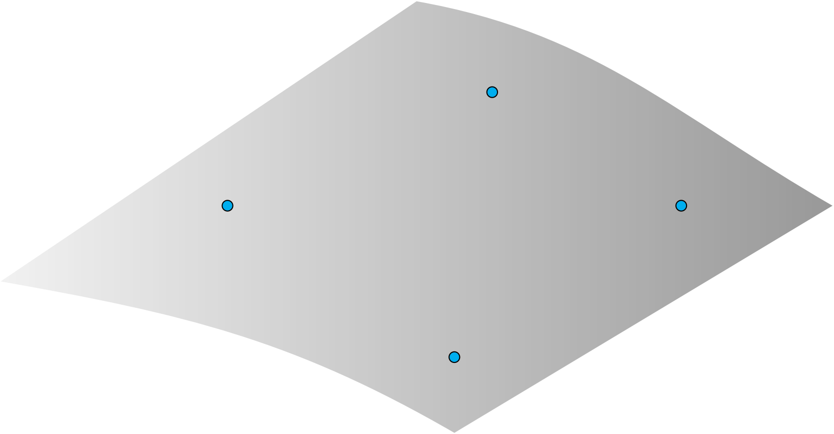
Random Pairwise Gossip (Xiao & Boyd'04)



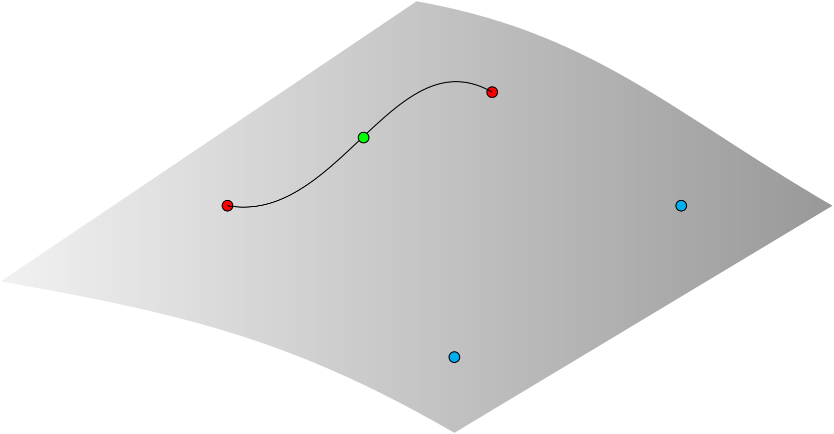
$$x_\infty = \begin{pmatrix} -0.25 & 0.25 \\ -0.25 & 0.25 \\ -0.25 & 0.25 \\ -0.25 & 0.25 \end{pmatrix}$$

$$x_n = \left(I - \frac{1}{2}(\delta_{i_n} - \delta_{j_n})(\delta_{i_n} - \delta_{j_n})^T \right) x_{n-1}$$

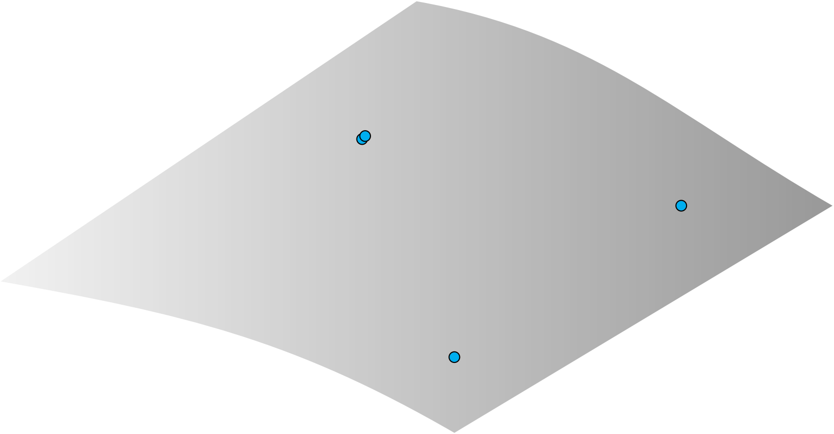
A natural extension in a metric setting



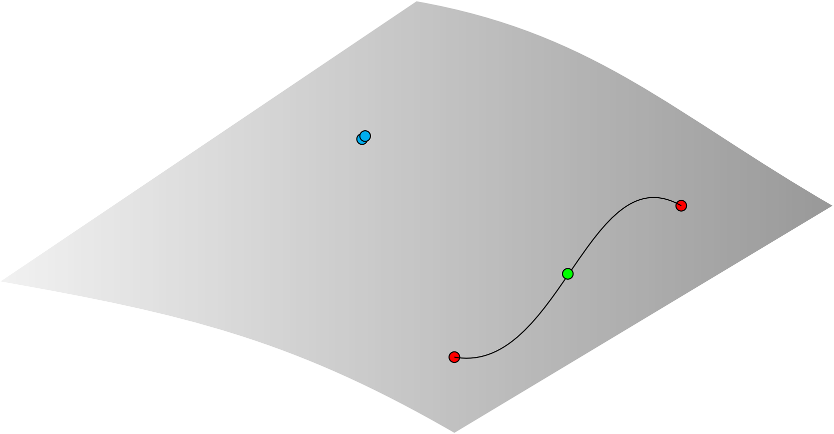
A natural extension in a metric setting



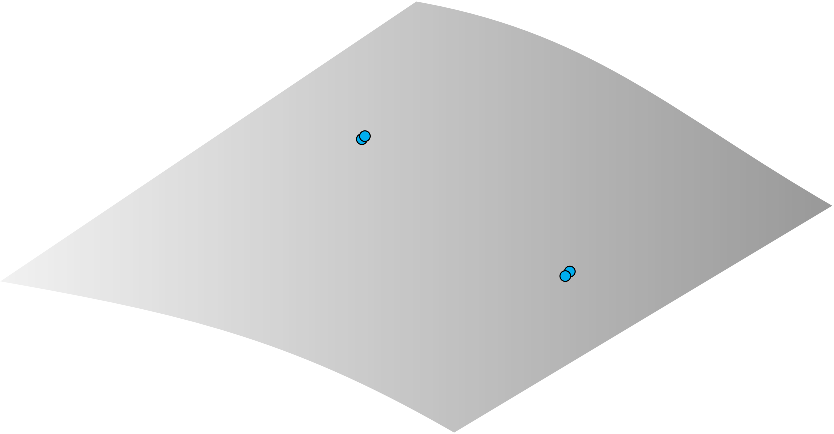
A natural extension in a metric setting



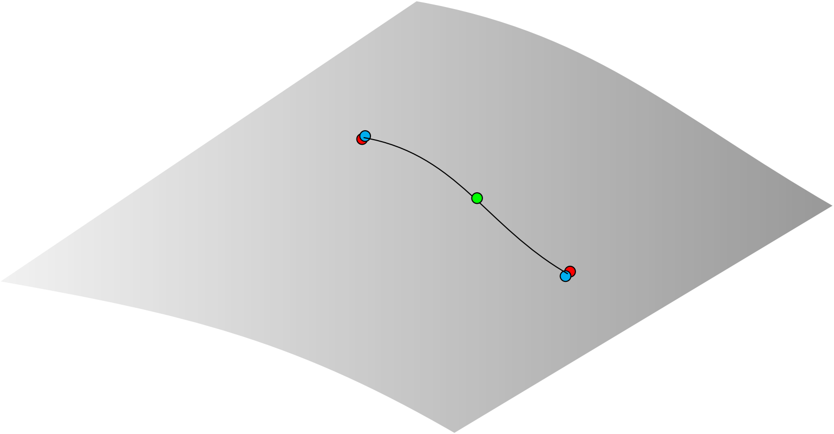
A natural extension in a metric setting



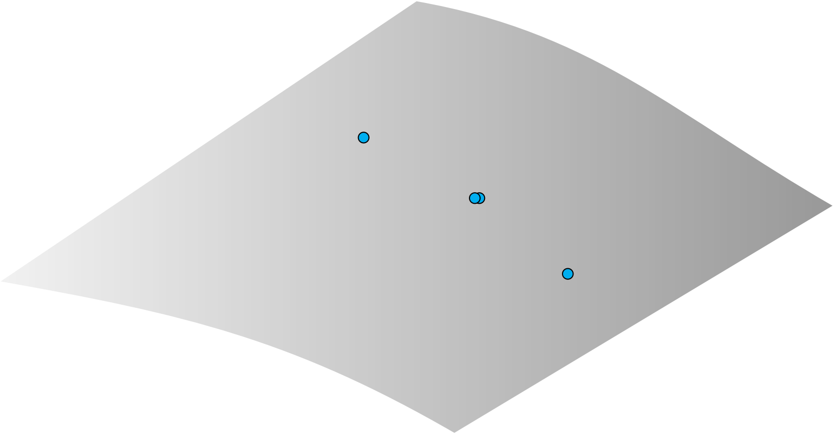
A natural extension in a metric setting



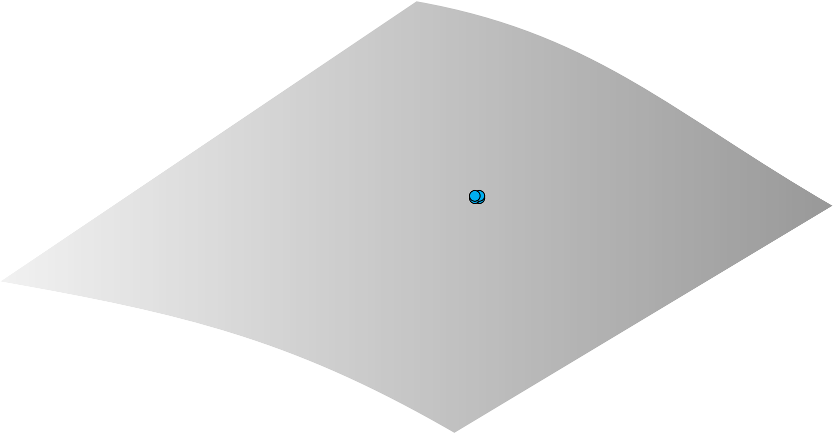
A natural extension in a metric setting



A natural extension in a metric setting



A natural extension in a metric setting



Outline

1. Motivation
2. State of the art
3. CAT(κ) spaces
4. Previous result for $\kappa = 0$
5. Why the $\kappa > 0$ case is more complex
6. Our result

Motivation

In its Euclidean setting, Random Pairwise Midpoint cannot address several useful type of data:

- Sphere positions (Sphere)
- Line orientations (Projective space)
- Solid orientations (Rotations)
- Subspaces (Grassmanians)
- Phylogenetic Trees (Metric space)
- Cayley graphs (Metric space)
- Reconfigurable systems (Metric space)

State of the art

- Consensus optimization on manifolds : [Sarlette-Sepulchre'08],[Tron *et al.*'12],[Bonnabel'13]
- Synchronization on the circle : [Sarlette *et al.*'08]
- Synchronization on SO(3) : [Tron *et al.*'12]
- Our previous work: Distibuted pairwise gossip on $CAT(0)$ spaces

Caveat: In this work, we deal the problem of synchonization, *i.e.* attaining a consensus, whatever its value; contrarily to the Euclidean case where it is known that random pairwise midpoints converges to $\bar{x}_0$.

CAT(κ) spaces

Model spaces

Consider a model surface $\mathcal{M}_\kappa$ with constant sectional curvature κ:

- $\kappa < 0$ corresponds to a *hyperbolic space*
- $\kappa = 0$ corresponds to a *Euclidean space*
- $\kappa > 0$ corresponds to a *sphere*

Geodesics

Assume $\mathcal{M}$ is a metric space equipped with metric d. A map $\gamma : [0, l] \to \mathcal{M}$ such that:

$$\forall 0 \leq t, t' \leq l, \quad d\left(\gamma(t), \gamma(t')\right) = |t - t'|$$

is called a *geodesic* in $\mathcal{M}$; $a = \gamma(0)$ and $b = \gamma(l)$ are its endpoints. If there exists one and only one geodesic linking a to b, it is denoted $[a, b]$.

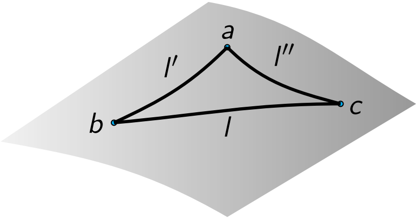
CAT(κ) spaces (cont'd)

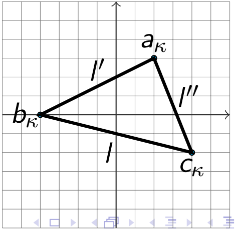
Triangles

A triple of geodesics γ, γ' and γ'' with respective endpoints a, b and c is called a *triangle* and is denoted $\triangle(\gamma, \gamma', \gamma'')$ or $\triangle(a, b, c)$ when there is no ambiguity.

Comparison triangles

When $\kappa \leq 0$, given a triangle $\triangle(\gamma, \gamma', \gamma'')$, there always exist a triangle $\triangle(a_\kappa, b_\kappa, c_\kappa)$ in $\mathcal{M}_\kappa$ such that $d(a, b) = d(a_\kappa, b_\kappa)$, $d(b, c) = d(b_\kappa, c_\kappa)$ and $d(c, a) = d(c_\kappa, a_\kappa)$ with $a = \gamma(0)$, $b = \gamma'(0)$ and $c = \gamma''(0)$.





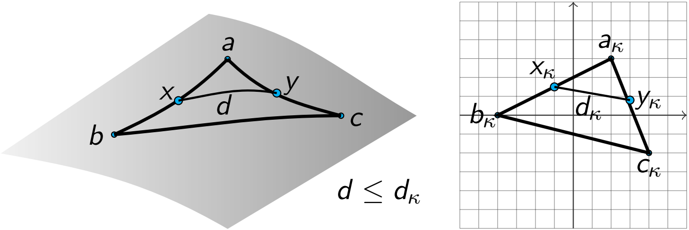
CAT(κ) spaces (cont'd)

CAT(κ) inequality

A triangle $\triangle(\gamma, \gamma', \gamma'')$ in a metric space $\mathcal{M}$ satisfies the CAT(κ) inequality if for any $x \in [a, b]$ and $y \in [a, c]$ one has:

$$d(x, y) \leq d(x_\kappa, y_\kappa)$$

where $x_\kappa \in [a_\kappa, b_\kappa]$ is such that $d(a_\kappa, x_\kappa) = d(a, x)$ and $y_\kappa \in [a_\kappa, c_\kappa]$ is such that $d(a_\kappa, y_\kappa) = d(a, y)$.



A metric space is said CAT(κ) if every pair of points can be joined by a geodesic and every triangle with perimeter less than $2D_\kappa = \frac{2\pi}{\sqrt{\kappa}}$ satisfy the CAT(κ) inequality.

Formal setting

Assumptions

1. Time is *discrete* $t = 0, 1, \ldots$
2. At each time each agent holds a "value" $x_{t,v}$ in a $CAT(\kappa)$ metric space $\mathcal{M}$
3. At each time t, an agent V_t randomly wakes up and wakes up a neighbor W_t, according to the probability distribution:

$$\mathbb{P}[\{V_k, W_k\} = \{v, w\}] = \begin{cases} P_{v,w} > 0 & \text{if } v \sim w \\ 0 & \text{otherwise} \end{cases}$$

Algorithm description

$$x_{t,v} = \begin{cases} \mathrm{Midpoint}(x_{t-1,V_t}, x_{t-1,W_t}) & \text{if } v \in \{V_t, W_t\} \\ x_{t-1,v} & \text{otherwise} \end{cases}$$

Previous result

The algorithm is sound

Because geodesics exist and are unique in CAT(0) spaces.

Convergence

The algorithm converges to a consensus with probability 1, whatever the initial state x_0.

Rate of convergence

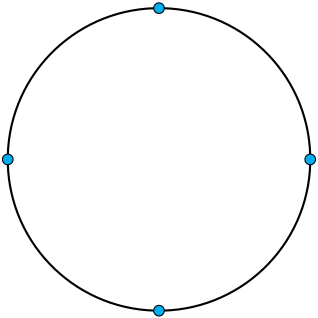
Convergence occur at a *linear rate*: define

$$\sigma^2(x) = \sum_{v \sim w} d^2(x_v, x_w) \; ;$$

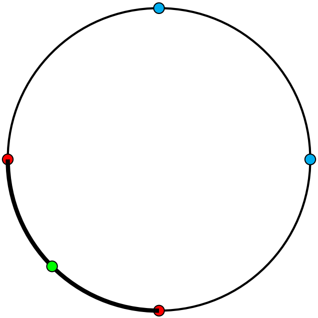
then, there exists a constant $L < 0$ such that

$$\mathbb{E}\sigma^2(X_k) \leq C_0 \exp(Lk)$$

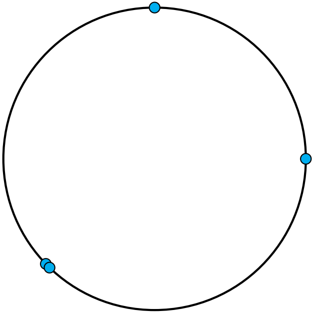
What changes for the $\kappa > 0$ (the case of the sphere)



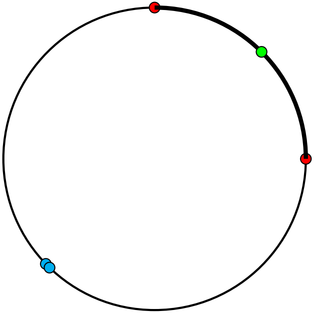
What changes for the $\kappa > 0$ (the case of the sphere)



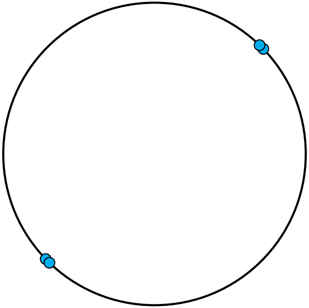
What changes for the $\kappa > 0$ (the case of the sphere)



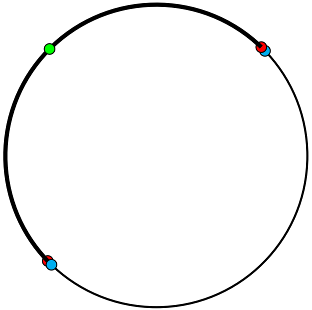
What changes for the $\kappa > 0$ (the case of the sphere)



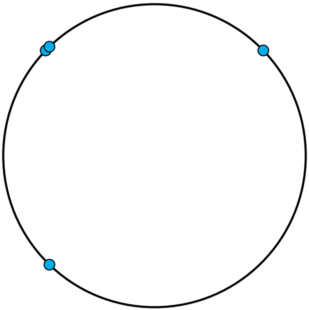
What changes for the $\kappa > 0$ (the case of the sphere)



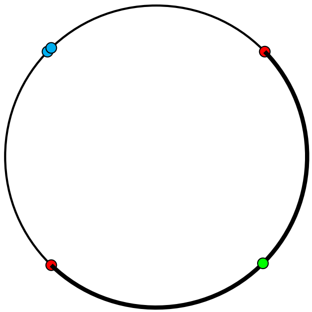
What changes for the $\kappa > 0$ (the case of the sphere)



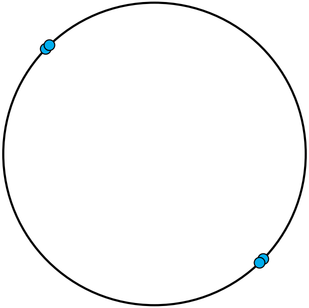
What changes for the $\kappa > 0$ (the case of the sphere)



What changes for the $\kappa > 0$ (the case of the sphere)



What changes for the $\kappa > 0$ (the case of the sphere)



Our result

Provided the diameter of the initial set of values is less than $D_\kappa/2$,

The algorithm is sound

Because geodesics exist and are unique using this restriction.

Convergence

The algorithm converges to a consensus with probability 1.

Rate of convergence

Convergence occur at a *linear rate*: define

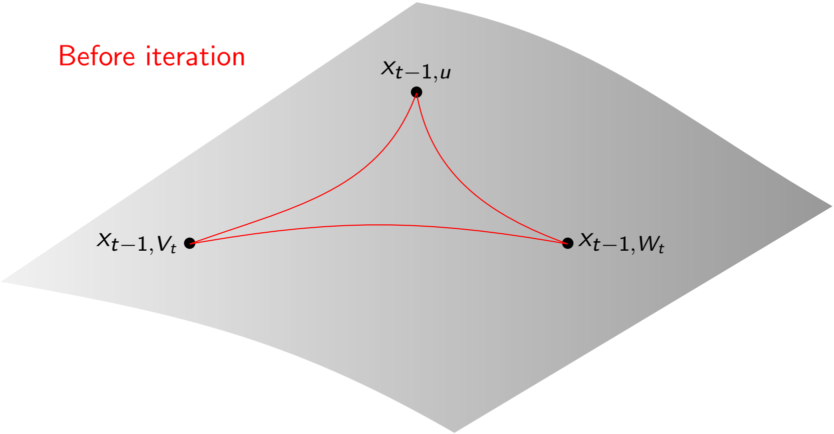
$$\sigma^2(x) = \sum_{v \sim w} \chi_\kappa \left(d(x_v, x_w) \right) \; ;$$

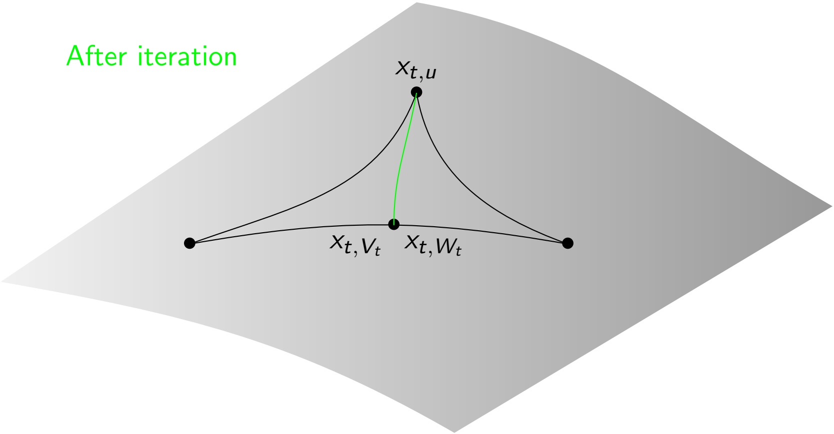
with:

$$\chi_\kappa(x) = 1 - \cos(\sqrt{\kappa} x)$$

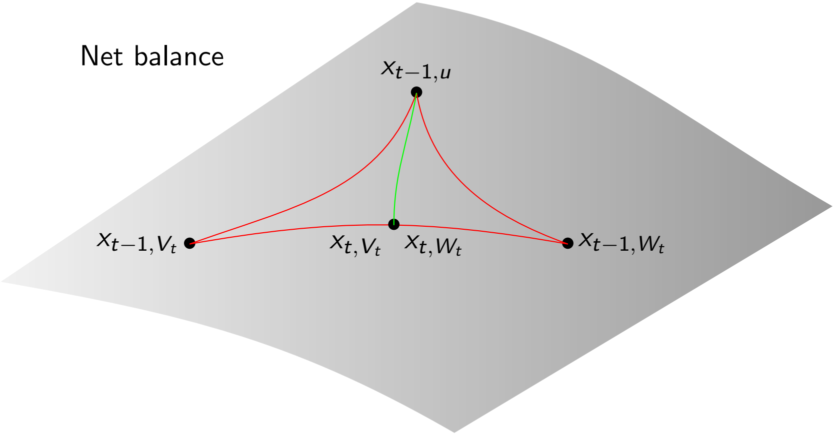
then, there exists a constant $L \in (-1, 0)$ such that:

$$\mathbb{E}\sigma^2(X_k) \leq C_0 \exp(Lk)$$


Before iteration
$x_{t-1,u}$
x_{t-1,V_t}
x_{t-1,W_t}


After iteration
$X_{t,u}$
X_{t,V_t}
X_{t,W_t}

Net balance



Sketch of proof (Net balance)

Let us look at the increments:

$$N(\sigma_\kappa^2(X_t) - \sigma_\kappa^2(X_{t-1})) = -\chi_\kappa(d(X_{V_t}(t-1), X_{W_t}(t-1)))$$

$$+ \sum_{\substack{u \in V \\ u \neq V_t, u \neq W_t}} T_\kappa(V_t, W_t, u)$$

with:

$$T_\kappa(V_t, W_t, u) = 2\chi_\kappa(d(X_u(t), M_t)) - \chi_\kappa(d(X_u(t), X_{V_t}(t-1)))$$

$$-\chi_\kappa(d(X_u(t), X_{W_t}(t-1)))$$

Using the inequality:

$$\chi_\kappa\left(d\left(\left\langle\frac{p+q}{2}\right\rangle\right), r\right) \leq \frac{\chi_\kappa(d(p,r)) + \chi_\kappa(d(q,r))}{2}$$

Sketch of proof (Two propositions)

We can prove the a first propostion:

$$\mathbb{E}[\sigma_\kappa^2(X_{k+1}) - \sigma_\kappa^2(X_k)] \leq -\frac{1}{N}\mathbb{E}\Delta_\kappa(X_k)$$

with:

$$\Delta_\kappa(x) = \frac{1}{2N} \sum_{\substack{v \sim w \\ \{v,w\} \in E}} P_{v,w}\chi_\kappa(d(x_v, x_w))$$

Using graph connectedeness we prove a second proposition:
Assume $G = (V, E)$ is an undirected connected graph, there exists a constant $C_G \geq 1$ depending on the graph only such that:

$$\forall x \in \mathcal{M}^N, \quad \frac{1}{2}\Delta_\kappa(x) \leq \sigma_\kappa^2(x) \leq C_G \Delta_\kappa(x)$$

Sketch of proof (cont'd)

The following lemma
Assume a_n is a sequence of nonnegative numbers such that $a_{n+1} - a_n \leq -\beta a_n$ with $\beta \in (0,1)$. Then,

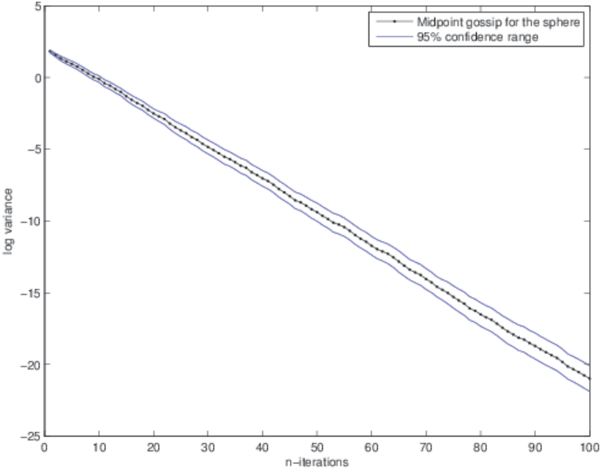
$$\forall n \geq 0, \quad a_n \leq a_0 \exp(-\beta n)$$

Combined with the two propositions, gives the desired result.

$$\mathbb{E}\sigma^2(X_k) \leq \exp(Lk)$$

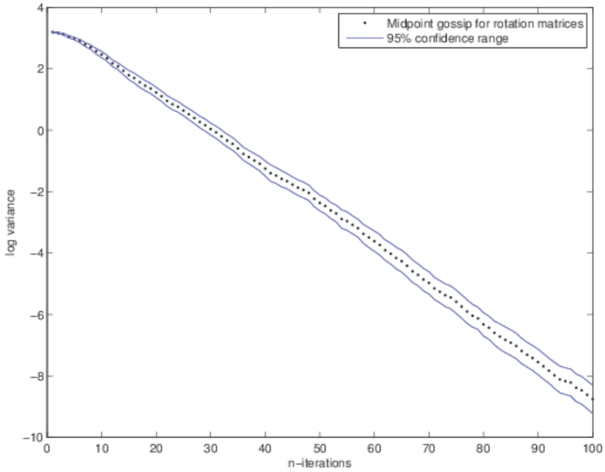
Simulation results

- Sphere



Simulation results

- Rotations



Summary

- We have proved that, when the data belong to complete CAT(κ) metric space, *provided the initial values are close enough*, the same algorithm makes sense and also converge linearly.
- We have checked that our results are consistent with simulations.