Randomized averaging algorithms:
when can errors & unreliability just be ignored?

Paolo Frasca & ***Julien Hendrickx***

Lund – LCCC workshop Oct 2014

(Linear) Averaging

Agents have values $x_i(t)$

Averaging iteration:

$$x_i(t+1) = \sum a_{ij}(t)x_j(t)$$

$$a_{ij}(t) \geq 0$$

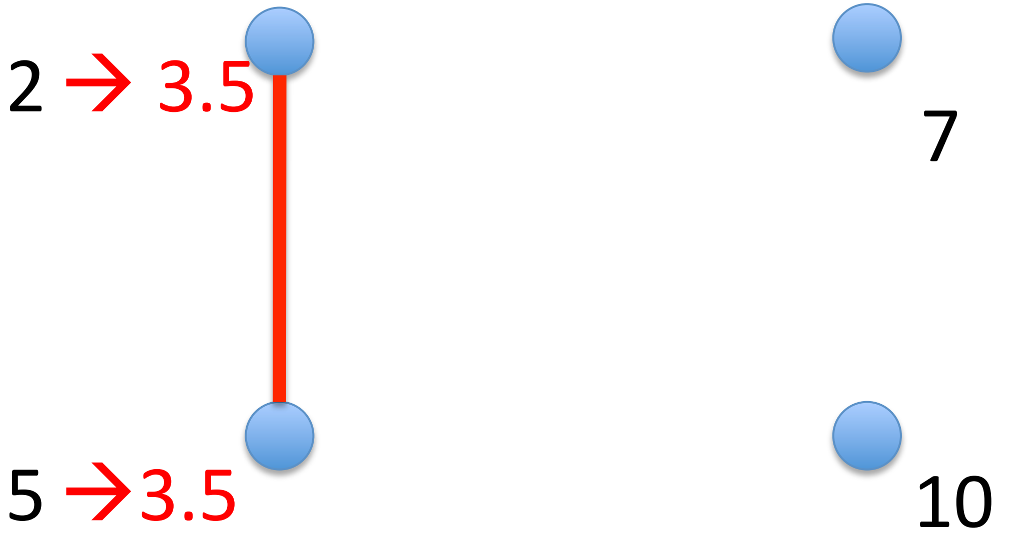
$$\sum_j a_{ij}(t) = 1$$

If average preserving (ex: if $a_{ij} = a_{ji}$) and enough interactions

$$x_i \to \bar{x} = \frac{1}{n}\sum x_i(0)$$

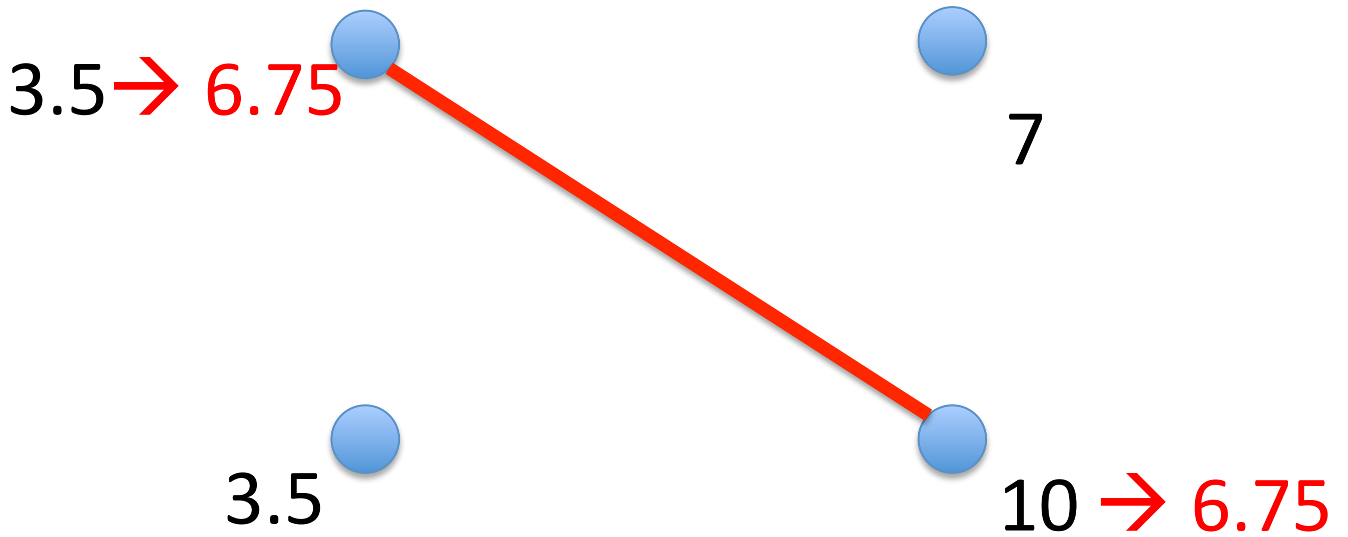
Basis for many decentralized algorithms

Ex: Gossip



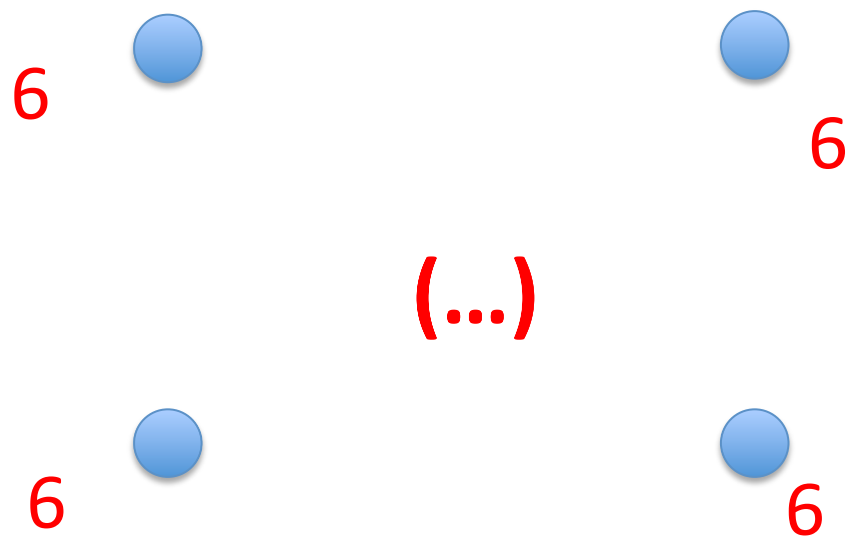
$$\bar{x} = 6$$

Ex: Gossip



$\bar{x} = 6$

Ex: Gossip



$\bar{x} = 6$

Convergence to the average

What if transmissions are asymmetric/uncertain?

Communication requirements for averaging

- Simple consensus:

$$x(t+1) = A(t)x(t)$$ $A(t)$ stochastic

Average preserved if *A(t)* doubly stochastic. $\mathbf{1}^T A(t) = \mathbf{1}^T$. Needs

- ***symmetric / balanced*** interactions
- therefore ***synchronous*** interactions

- Push sum *[Kempe et al 03]*, based on mass preservation

Directed asynchronous communications OK

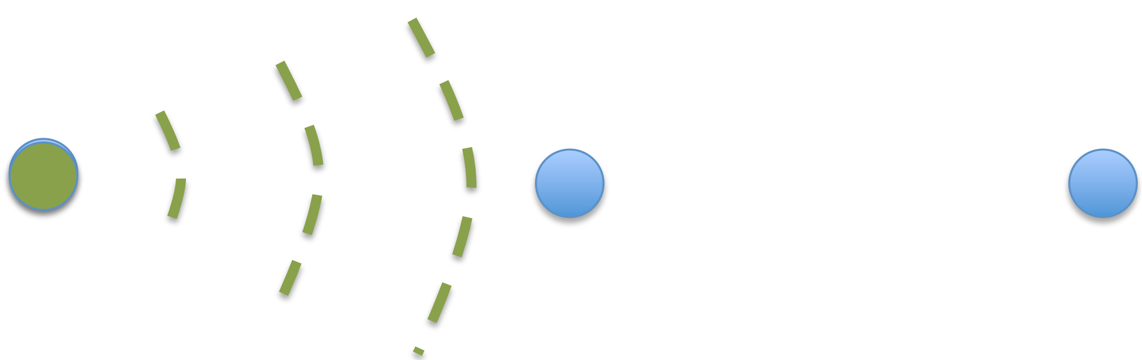
But needs ***reliability***: agents must know

- if message received,
- how many agents receive message

Unpredictability and Unreliability

- Directed & uncertain communication

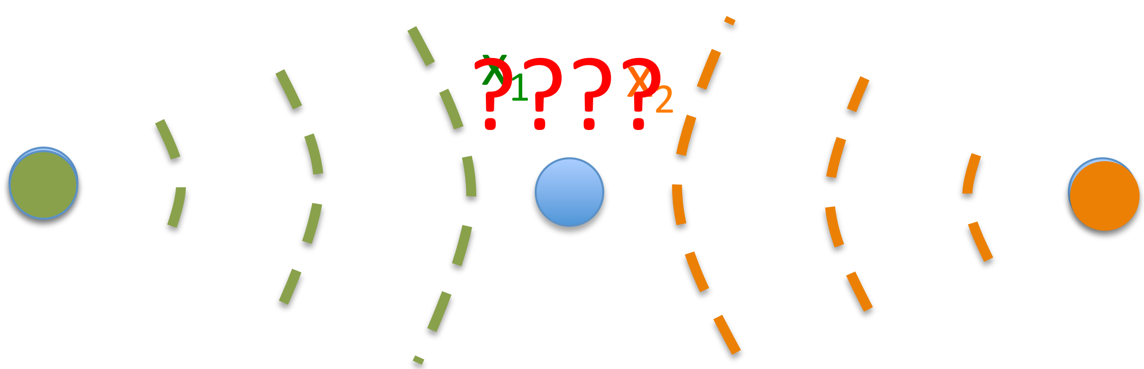
- Packet loss
- Broadcasts at random times



Unpredictability and Unreliability

- Directed & uncertain communication

- Packet loss
- Broadcasts at random times
- Collisions due to interferences



Unpredictability and Unreliability: 2 approaches

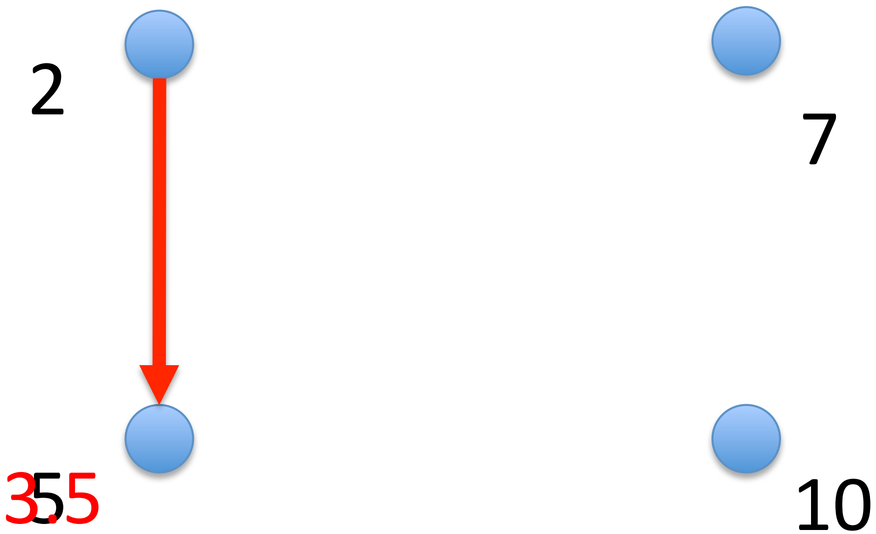
1) Verification and correction mechanisms

idea: *errors may create big trouble*

2) Ignore issues, assume average behavior and hope for the best

Idea: *errors will cancel out*

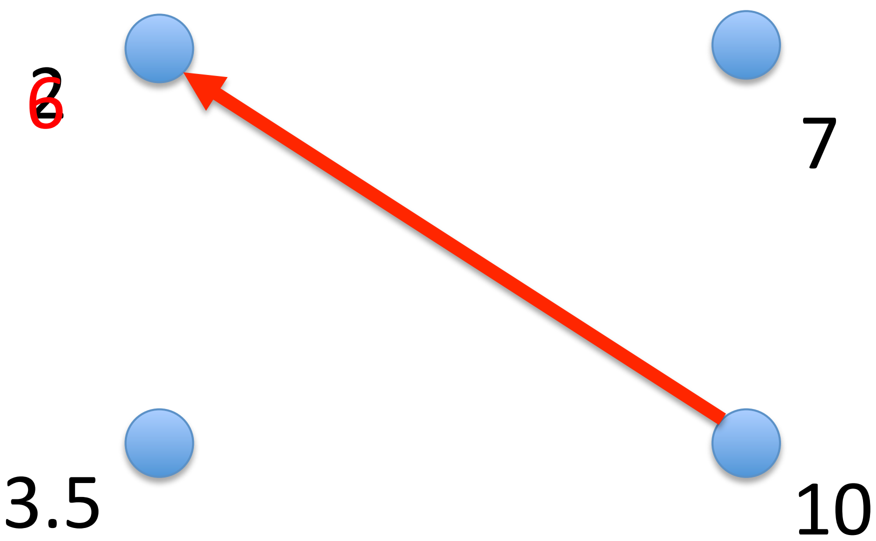
Naïve extension: Asymmetric Gossip



$$\bar{x}(0) = 6$$

$$\bar{x} = 5.625$$

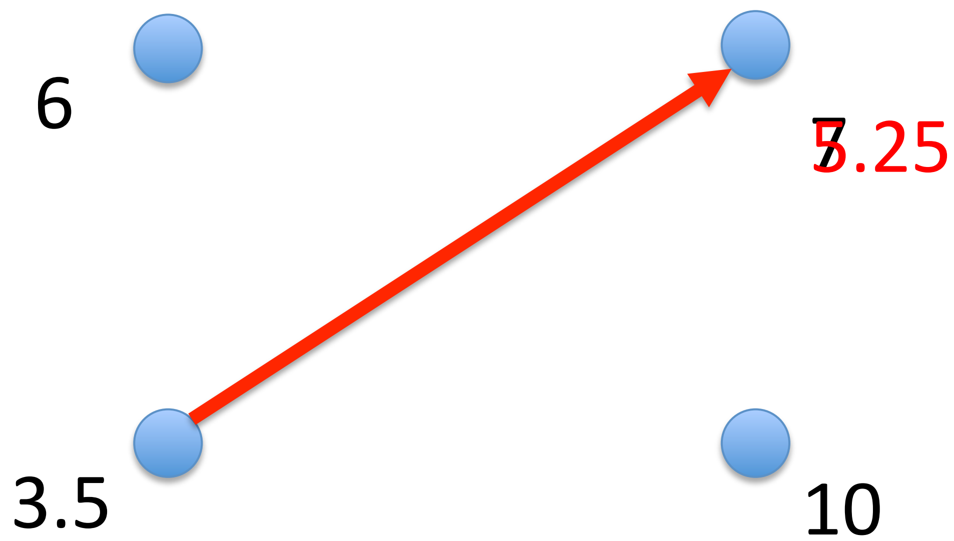
Naïve extension: Asymmetric Gossip



$$\bar{x}(0) = 6$$

$$\bar{x} = 5\ \ 6.625$$

Naïve extension: Asymmetric Gossip



$$\bar{x}(0) = 6$$

$$\bar{x} = 6.625 \quad 6.1875$$

Does not converge to average, but seems close…

Guarantee?

Outline

- Introduction
- ***Formulation and previous results***
  - ***Formulation***
  - ***Exact approach***
  - ***Convergence-based approach***
- Our approach
- Particularization and Applications
- Conclusions

Problem formulation

$$x(t+1) = A(t)x(t)$$

where $A(t) \in \Re_{+}^{n \times n}$

- Averaging matrices ($A(t)\mathbf{1} = \mathbf{1}$)
- ***i.i.d. random variables***

Any distribution, but
events at different times independent

Problem formulation

$$x(t+1) = A(t)x(t)$$

where $A(t) \in \Re_{+}^{n \times n}$

- Averaging matrices ($A(t)\mathbf{1} = \mathbf{1}$)
- ***i.i.d. random variables***
- preserve average on expectation ($\mathbf{1}^T E A(t) = \mathbf{1}^T$)

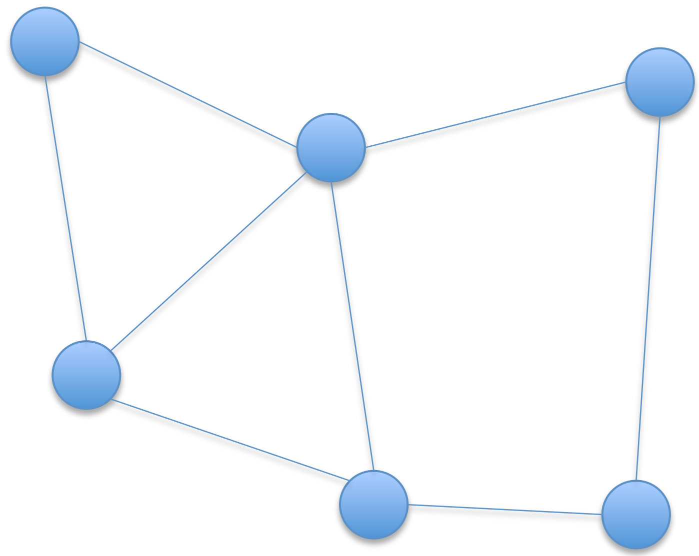
➔ $E\bar{x}(t) = \bar{x}(0)$

Convergence to "consensus" w.p. 1 under conditions

$$x(t) \to \mathbf{1}x_{\infty} \quad \text{with} \quad Ex_{\infty} = \bar{x}(0)$$

Goal: bound ***standard asymptotic error*** $E\left(x_{\infty} - \bar{x}(0)\right)^2$

Examples: nominal network



1. Synchronous symmetric updates,

$$x_i(t+1) = x_i(t) + \sum_j a_{ij}(x_j(t) - x_i(t))$$

   but some messages are lost

2. Nodes wake up and update at random times

3. Nodes wake up, send messages to neighbor(s) at random time

Exact approach

Standard asymptotic error computed exactly using x(0) and V:

$$E(A^T V A) = V \qquad \mathbf{1}^T V \mathbf{1} = 1$$

But, very hard to obtain expression for V

→ Theoretical computation only for specific cases (involved!)

Ex: *[Fagnani & Zampieri, SICON 2009]*
Fixed averaging algorithm with packet losses
deployed on Cayley graphs of Abelian groups

Exact approach

Standard asymptotic error computed exactly using x(0) and V:

$$E(A^T V A) = V \qquad \mathbf{1}^T V \mathbf{1} = 1$$

But, very hard to obtain expression for V

→ Theoretical computation only for specific cases (involved!)
→ Numerical verification for given system (hard)

- Numerical issues
- May need enumeration of all A

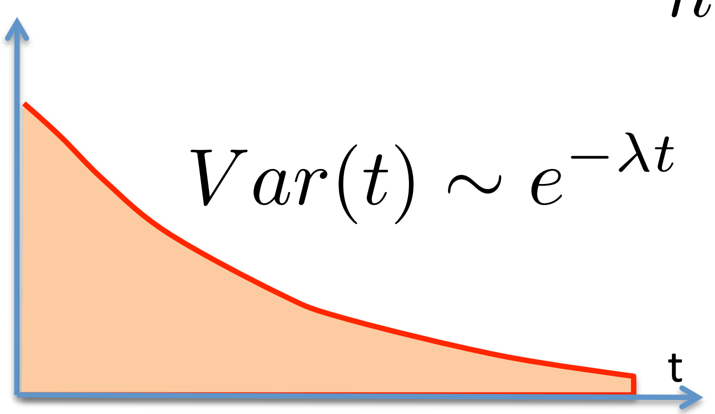
often many (one for each possible combination of updates)

Convergence based approach

$Var(t)$ Individual variance $\mathsf{E}\,\frac{1}{n}\sum\left(x_i(t)-\bar{x}(t)\right)^2$

$err(t)$ Expected group error $\mathsf{E}\left(\bar{x}(t)-\bar{x}(0)\right)^2$

Idea $\Delta err(t) \leq \frac{K}{n} Var(t)$

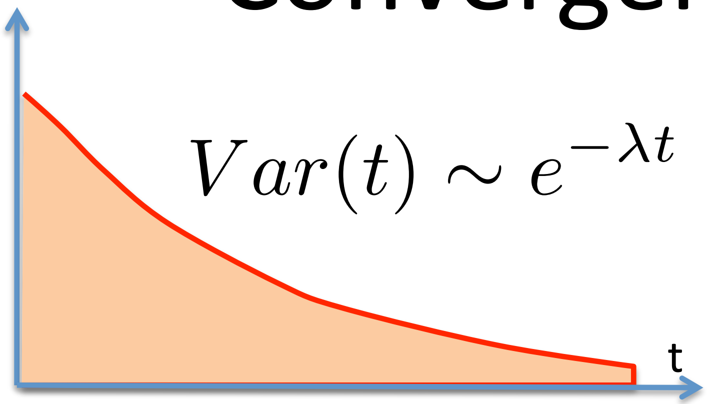


$$err(\infty) \leq O\left(\frac{K}{n\lambda}\right) Var(0)$$

λ linked to algebraic connectivity of "network"

Ex: [Fagnani & Frasca, 2011]

Convergence based approach



$$err(\infty) \leq O\left(\frac{K}{n\lambda}\right) Var(0)$$

λ linked to algebraic connectivity of "network"

Conclusion: keep λ large when n grows ➔ expanders, etc.

But, expanders not easy to build physically (range constraints, etc.)

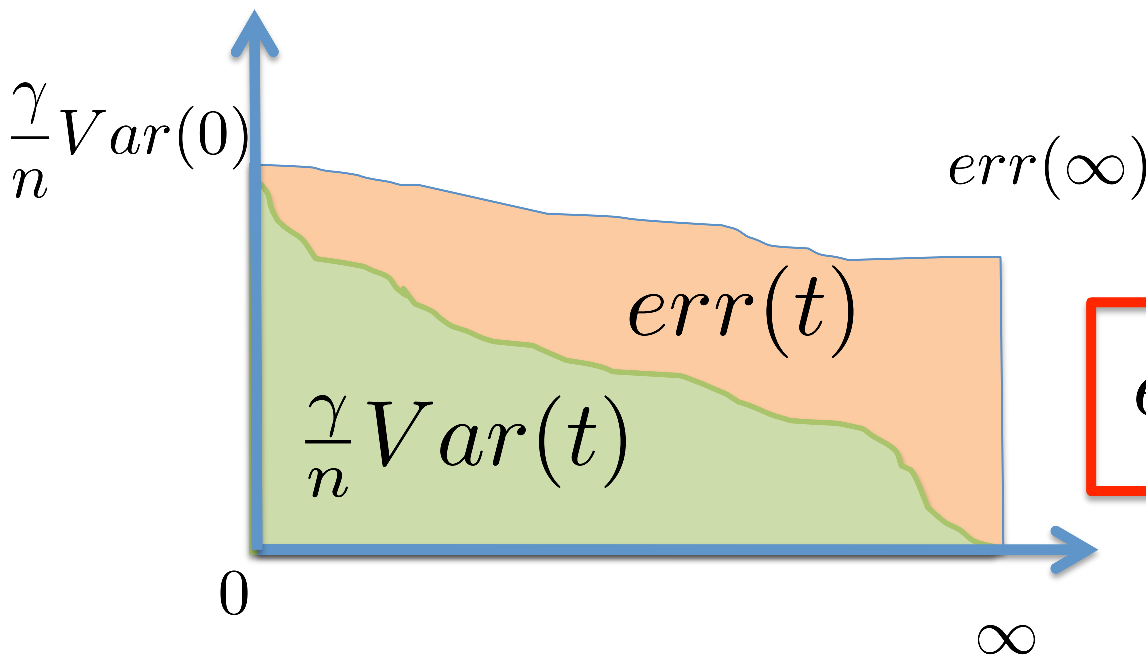
and λ often not critical in experimental results

Outline

- Introduction
- Formulation and previous results
- ***Our approach***
- Particularization and Applications
- Conclusion and further works

Our approach: conservation

Find γ s.t. $err(t) + \frac{\gamma}{n} Var(t)$ non-increasing



$$err(\infty) \leq \frac{\gamma}{n} Var(0)$$

- If γ independent of n (or o(n)), "***accuracy***" for large n
- Not always the case, exist systems with large errors

How to pick γ ?

$$err(t) + \frac{\gamma}{n} Var(t)$$ non-increasing

γ valid if $E\left(L^T \mathbf{1}\mathbf{1}^T L\right) \leq \gamma\, E\left(L + L^T - L^T L\right)$

with $L(t) = I - A(t)$

Appears involved, but

- Generic γ can be found for large classes of systems
- Results very easy to use
- Often leads to strong bounds

Outline

- Introduction
- Formulation and previous results
- Our approach
- ***Particularization and Applications***
   1. ***Limited updates***
   2. ***Uncorrelated updates***
   3. ***Uncorrelated broadcasts***
   4. ***Limited correlations***
- Conclusions

1. Limited updates

$$A_{\max} \geq \sum_i \sum_{j \neq i} a_{ij}(t)$$ Max total simultaneous updates

$$a_{ii}^{\min} \leq a_{ii}(t)$$ Min. self confidence

Then $$\gamma = \frac{A_{\max}}{a_{ii}^{\min}}$$

$$err(\infty) \leq \frac{1}{n} \left(\frac{A_{\max}}{a_{ii}^{\min}} \right) V(0)$$

"Accuracy" if self confidence + amount of updates << n

Asymmetric gossip

$$A_{\max} \geq \sum_i \sum_{j \neq i} a_{ij}(t) \quad \Rightarrow \quad err(\infty) \leq \frac{1}{n} \left(\frac{A_{\max}}{a_{ii}^{\min}} \right) V(0)$$

Iteration: At each t, ***1 random node*** i chooses* j and updates to

$$x_i(t) + q\left(x_j(t) - x_i(t)\right)$$

$$A_{\max} = q$$

$$a_{ii}^{\min} = 1 - q$$

$$\Rightarrow \quad err(\infty) \leq \frac{1}{n} \left(\frac{q}{1-q} \right) V(0)$$

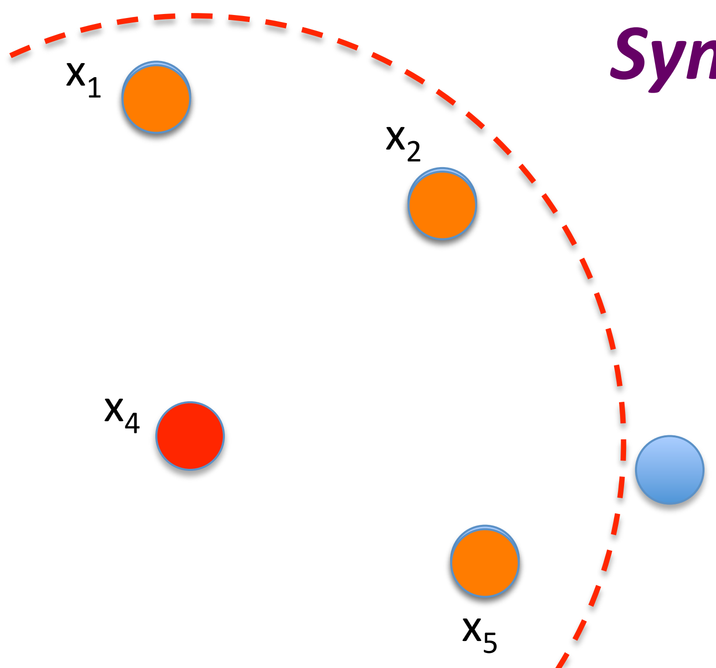
Accuracy for large n. Bound similar to *[Fagnani & Zampieri 08]*

* With symmetric probabilities

Broadcast gossip

- At each t, a node j wakes up with uniform probability
- j transmits x_j to all within listening range (assumed symmetric)
- update $$x_i(t+1) = x_i(t) + q(x_j(t) - x_i(t))$$

$q \in (0,1)$



Symmetry → $E(\bar{x}(t)) = \bar{x}(0)$

Broadcast gossip

- At each t, a node j wakes up with uniform probability
- j transmits x_j to all within listening range (assumed symmetric)
- update $x_i(t+1) = x_i(t) + q(x_j(t) - x_i(t))$

$q \in (0,1)$

<u>Previous results</u>

- Involved, rely on spectral quantities related to network defined by listening ranges
- No general guarantee of accuracy for large n

Ex: $$err(\infty) \leq O\left(\frac{d_{\max}^2}{\lambda_1 n}\right) V(0)$$

$d_{\max}$ Largest degree

λ_1 Smallest nonzero eigenvalue laplacian

[Aysal et al. 2009], [Fagnani & Delvenne 2010], [F. Fagnani and P. Frasca. 2011a,b]

Broadcast gossip: our approach

- At each t, a node j wakes up with uniform probability
- j transmits x_j to all within listening range (assumed symmetric)
- update $$x_i(t+1) = x_i(t) + q(x_j(t) - x_i(t))$$

$q \in (0,1)$

Limited updates rule:

$$A_{\max} = q d_{\max}$$
$$a_{ii}^{\min} = 1 - q$$

➔ $$err(\infty) \leq \frac{d_{\max}}{n} \left(\frac{q}{1-q} \right) V(0)$$

- Accuracy for large n if $d_{\max} < o(n)$
- Independent of graph spectrum

2. Independent updates

$$err(\infty) \leq \frac{\gamma}{n} Var(0)$$

If nodes make independent decisions about updates and weights

a_{ij}, a_{kl} Uncorrelated if $i \neq k$ Then

$$\gamma = \frac{1 - a_{ii}^{\min}}{a_{ii}^{\min}}$$

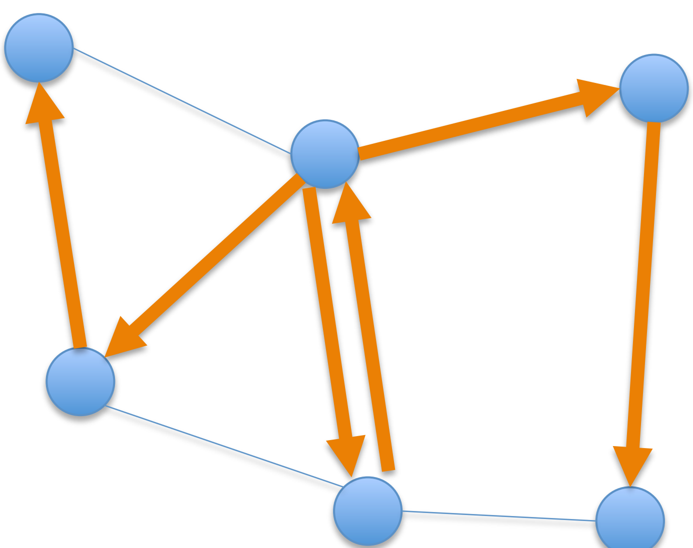
➔ Always accurate when $n \to \infty$ if $a_{ii}^{\min}$ bounded

Synchronous Asymmetric Gossip

Iteration: Every node i chooses neighbor j_i and updates to

$$x_i(t+1) = (1-q)x_i(t) + q x_{j_i}(t)$$

Probabilities s.t $E(\bar{x}(t)) = \bar{x}(0)$



Synchronous Asymmetric Gossip

Iteration: Every node i chooses neighbor j_i and updates to

$$x_i(t+1) = (1-q)x_i(t) + qx_{j_i}(t)$$

Probabilities s.t $E(\bar{x}(t)) = \bar{x}(0)$

Previous results

- Related to eigenvalues of graph Laplacian
- Accuracy for particular cases

[Fagnani & Zampieri, 2008]

Synchronous Asymmetric Gossip

Iteration: Every node i chooses neighbor j_i and updates to

$$x_i(t+1) = (1-q)x_i(t) + qx_{j_i}(t)$$

Probabilities s.t $E(\bar{x}(t)) = \bar{x}(0)$

Indep updates: $err(\infty) \leq \frac{\gamma}{n} Var(0)$ $\gamma = \frac{1 - a_{ii}^{\min}}{a_{ii}^{\min}}$

$a_{ii}^{\min} = 1 - q$

$$err(\infty) \leq \frac{1}{n}\left(\frac{q}{1-q}\right)V(0)$$

For all graphs, independently of spectral quantities

3. Independent broadcasts

$$err(\infty) \leq \frac{\gamma}{n} Var(0)$$

- 2. Independent decisions about updates and weights

$$\gamma = \frac{1 - a_{ii}^{\min}}{a_{ii}^{\min}}$$

- 3. Independent broadcast decisions (i.e. columns in A)

$$\gamma = \frac{a_{col}^{\max}}{a_{ii}^{\min}}$$

Max importance of a node

Minimal self confidence

$$a_{col}^{\max} \geq \sum_{i:i \neq j} a_{ij}(t) \qquad \forall j, t$$

4. General correlation rule

$$err(\infty) \leq \frac{\gamma}{n} Var(0)$$

If ***no*** a_{ij} ***correlated to more than 4C*** other coefficients

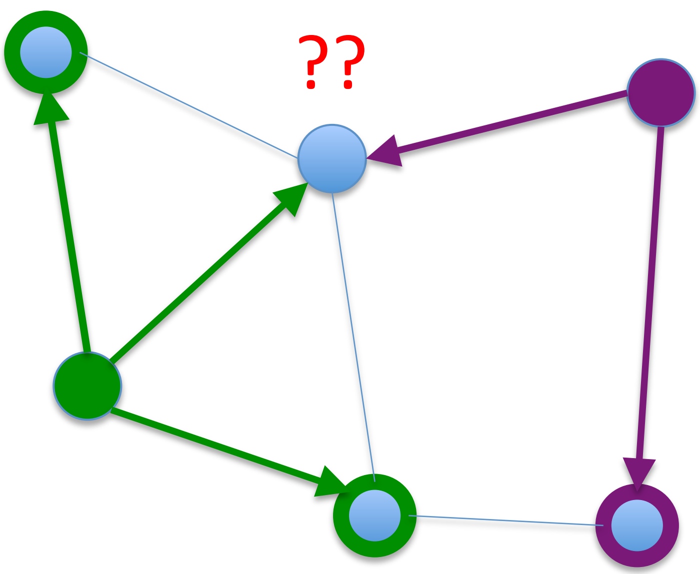
$$\gamma = \frac{C}{a_{ii}^{\min}}$$

(conservative)

Broadcast with collisions

At every time:

- Some nodes awake and broadcast their values
- Nodes receiving one value update
- Nodes receiving two or more values do not



Weights and probabilities
s. t. $E(\bar{x}(t)) = \bar{x}(0)$

Challenges:

- Multiple updates
- Multiple correlations

Broadcast with collisions

At every time:

- Some nodes awake and broadcast their values
- Nodes receiving one value update
- Nodes receiving two or more values do not

<u>*Previous results*</u>

- Results on Cayley graphs of Abelian Groups

- Accuracy for certain sparse networks when n grows

[Fagnani & Frasca, 2011]

Broadcast with collisions

$$\gamma = \frac{C}{a_{ii}^{\min}}$$

If ***no*** a_{ij} ***correlated to more than 4C*** other coefficients

How many correlations?

$a_{12} > 0$ If 1 transmits

Unless 3 transmits

Which also affects a_{34}

Correlations up to distance 3

➔ $4C \leq d_{max}^3$

Broadcast with collisions

$$\gamma = \frac{C}{a_{ii}^{\min}}$$

If ***no*** a_{ij} ***correlated to more than 4C*** other coefficients

Correlations up to distance 3 $\quad 4C \leq d_{max}^3$

$$err(\infty) \leq \frac{d_{\max}^3}{n} \frac{1}{a_{ii}^{\min}} V(0)$$

Accuracy for large n if $d_{max} \leq o(n^{1/3})$

Outline

- Introduction
- Formulation and previous results
- Our approach
- Particularization and Applications
- ***Conclusions***

Summary

- $\frac{\gamma}{n}$ group error + individual var. non-increasing*
- γ often easy to compute
- Bound independent of convergence speed and spectral properties; only local properties
- Prove asymptotical accuracy of many schemes
- Equals or outperforms (almost) all previous ad hoc results

* actually $\frac{\gamma}{n+\gamma}$

Summary

$$err(\infty) \leq \frac{\text{Largest "correlated event"}}{\text{Min "self confidence"}} \quad \frac{Var(0)}{n}$$

Ignoring problems and hoping for the best is OK if

- Limited correlations *(often OK in multi-agent systems)*
- Sufficient self-confidence for agents

Robustness of large class algorithms with respect to important fluctuations

Possible developments

- Expected average not entirely preserved
- Correlation between different times
- More detail → Less conservative bounds
- Other algorithms

Thank you for your attention!

References

Paolo Frasca and J.H, ***On the mean square error of randomized averaging algorithms,*** Automatica 2013 arXiv:1111.4572v1

Paolo Frasca and J.H., ***Large network consensus is robust to packet losses and interferences,*** Proceedings of ECC 2013, Zurich, July 2013