Sharp phase transition of the repeated average process

Lingfu Zhang

California Institute of Technology

Based on work in preparation with Dong Yao (Jiangsu Normal University, China)

Penn/Temple Probability Seminar **Feb 3, 2026**

The repeated average process

(Aldous-Lanoue, 11: A lecture on the averaging process)

- A (finite, connected) graph $G = (V, E)$
- A (real) number x_v at each vertex $v \in V$

At each step

- Pick an edge $e = (u, v)$ at random, uniformly from E
- Replace both x_u and x_v by $(x_u + x_v)/2$

Some motivations

Statistical physics dissipation of energy/temperature, etc.

Opinion dynamics in a social network

Economy distribution of wealth

Quantum computing concatenation of quantum circuits, as a product of random matrices

The repeated average process

(Aldous-Lanoue, 11: A lecture on the averaging process)

- A (finite, connected) graph $G = (V, E)$
- A (real) number x_v at each vertex $v \in V$

At each step

- Pick an edge $e = (u, v)$ at random, uniformly from E
- Replace both x_u and x_v by $(x_u + x_v)/2$

Many related models

Random redistribution $(x_u, x_v) \to (U(x_u+x_v), (1-U)(x_u+x_v))$

Kac's walk in thermodynamics $(x_u, x_v) \to (\cos(\theta)x_u - \sin(\theta)x_v, \sin(\theta)x_u + \cos(\theta)x_v)$

General interactions between 2 or more vertices

(Deffuant model, gossip algorithms …)

(for more background, see *Aldous-Lanoue, 11, Diaconis-Saloff-Coste, 14, Chatterjee-Diaconis-Sly-Z., 19, Movassagh-Szegedy-Wang, 22*)

The repeated average process

Let $\{x_v(t)\}_{v\in V}$ be the numbers after t steps:

- Sum $\sum_{v\in V} x_v(t)$ preserved
- As time $t \to \infty$, each $x_v(t) \to \bar{x} = |V|^{-1} \sum_{v\in V} x_v(t)$ (assume G finite and connected)

How fast does $\{x_v(t)\}_{v\in V}$ converge to $\{\bar{x}\}_{v\in V}$?

Some common/natural metrics:
(Below assume that each $x_v(0) \geq 0$, and $\sum_{v\in V} x_v(0) = 1$; then $\bar{x} = |V|^{-1}$.)

- L^2 distance $S(t) = \sum_{v\in V}(x_v(t) - \bar{x})^2$
- L^1 distance $T(t) = \sum_{v\in V} |x_v(t) - \bar{x}|$
- Entropy $H(t) = -\log(\bar{x}) + \sum_{v\in V} x_v \log(x_v)$

Each $\to 0$ as $t \to \infty$; how fast?

The repeated average process

Related to edge random walk on the same graph

Poissonize continuous time (instead of discrete steps), each edge activated at rate $|E|^{-1}$

(Aldous-Lanoue, 11) For each vertex v and $t \geq 0$,

$$\mathbb{E}x_v(t) = \sum_{u \in V} x_u(0)\mathbf{Q}^t(u,v) =: \mathbf{P}^t(v),$$

$\mathbf{Q}$: transition probability for the continuous time edge random walk on $G = (V, E)$, speed $(2|E|)^{-1}$.

$\mathbf{P}^t(v)$: probability of walkers at v at time t, with initial distribution $\mathbf{P}^t(v) = x_v(0)$

Comparison random walk vs repeated average

- L^2 distance $\mathbb{E}S(t) = \mathbb{E}\sum_{v \in V}(x_v(t) - \bar{x})^2 \geq \sum_{v \in V}(\mathbf{P}^t(v) - \bar{x})^2 =: S_{RW}(t)$
- L^1 distance $\mathbb{E}T(t) = \mathbb{E}\sum_{v \in V}|x_v(t) - \bar{x}| \geq \sum_{v \in V}|\mathbf{P}^t(v) - \bar{x}| =: T_{RW}(t)$
- Entropy $\mathbb{E}H(t) = -\log(\bar{x}) + \mathbb{E}\sum_{v \in V} x_v \log(x_v)$

$$\geq -\log(\bar{x}) + \mathbb{E}\sum_{v \in V} \mathbf{P}^t(v)\log(\mathbf{P}^t(v)) =: H_{RW}(t)$$

*Roughly, random walk mixing is **faster** than convergence of the repeated average process*

L^2 decay via spectrum gap

(Aldous-Lanoue, 11) For the graph Laplacian $D - A$ of $G = (V, E)$, let its eigenvalues be $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{|V|}$. Then for $\epsilon > 0$,

$$\mathbb{E}[S(t+\epsilon)|S(t)] \leq e^{-\epsilon\lambda_2/(2|E|)} S(t),$$

where $S(t) = \sum_{v \in V} (x_v(t) - \bar{x})^2$.

For comparison **Classical L^2 decay of random walk** $\partial_t S_{RW}(t) \leq -\lambda_2 S_{RW}(t)/(2|E|)$

So

$$S_{RW}(t+\epsilon) \leq e^{-\epsilon\lambda_2/(2|E|)} S_{RW}(t),$$

where $S_{RW}(t) = \sum_{v \in V} (\mathbf{P}^t(v) - \bar{x})^2$.

L^2 decay

Some other L^2 results (lower bound)

- For $G = K_n$ (the complete graph)

$$\mathbb{E}[S(t+\epsilon)|S(t)] = e^{-\epsilon/(n-1)}S(t),$$

(Chatterjee-Diaconis-Sly-Z. 19)

- A related lower bound (general G)

$$\sup_{\{x_v(0)\}_{v\in V}:S(0)=1} \mathbb{E}S(t) \geq e^{-\epsilon\lambda_2/(2|E|)}$$

(Movassagh-Szegedy-Wang, 22)

Entropy decay (log-Sobolev)

(Aldous-Lanoue, 11) For $\epsilon > 0$,

$$\mathbb{E}[H(t+\epsilon)|H(t)] \leq e^{-\epsilon\alpha/(2|E|)}H(t),$$

where $H(t) = -\log(\bar{x}) + \sum_{v\in V} x_v(t)\log(x_v(t))$, and α is the <u>*log-Sobolev constant*</u> of the graph Laplacian.

For comparison **Classical entropy decay of random walk** $\partial_t H_{RW}(t) \leq -\alpha H_{RW}(t)/(2|E|)$

So

$$H_{RW}(t+\epsilon) \leq e^{-\epsilon\alpha/(2|E|)}H_{RW}(t),$$

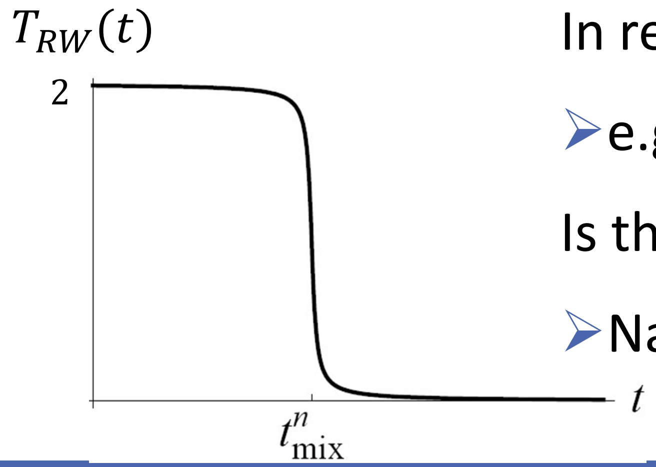
where $H_{RW}(t) = -\log(\bar{x}) + \mathbb{E}\sum_{v\in V} \mathbf{P}^t(v)\log(\mathbf{P}^t(v))$.

L^1 decay

A natural & more challenging question: decay of $T(t) = \sum_{v \in V} |x_v(t) - \bar{x}|$?

Random walk analog $T_{RW}(t) := \sum_{v \in V} |\mathbf{P}^t(v) - \bar{x}| \in [0,2]$

- Also known as the **total-variation distance** between $\mathbf{P}^t$ (distribution at time t) and uniform (stationary).
- Measure the convergence of sampling algorithms/MCMC.
- Cut-off (Card shuffling, *Diaconis-Shahshahani, 80s*; *Aldous-Diaconis, 86*)



In repeated average: $T(t)$ is the linear way of describing consensus/equity

- e.g., Gini index in wealth distribution

Is there also a **cut-off**? i.e., window of decay from ≈ 2 to $\approx 0 \ll$ total time

- Natural guess: YES whenever random walk has cut-off

Unlike L^2 / log-Sobolev, L^1 can be **different** from random walk.

Complete graph K_n

- Edge random walk (start from a given vertex, continuous time, speed $1/(2|E|) = 1/(n(n-1))$)

$T_{RW}(t) = \frac{2(n-1)}{n} e^{-t/n}$ (mixed after one move of walker)

- Repeated average: need time $\gg n$

(Chatterjee-Diaconis-Sly-Z., 19)

For $T(t) = \sum_{i=1}^{n} |x_i(t) - n^{-1}|$, and starting from $(x_1(0), \dots, x_n(0)) = (1,0, \dots, 0)$, for any $a \in \mathbb{R}$ we have

$$T\left(n(\log_2(n) + a\sqrt{\log_2(n)})/2\right) \to 2\Phi(-a),$$

in probability, where Φ is the CDF of standard Gaussian.

- Since the process is linear, $(1,0, \dots, 0)$ is the extremal initial configuration, with the slowest decay.

In words, cut-off at time $n\log_2(n)/2$, plus order $n\sqrt{\log(n)}$ window, and Gaussian profile.

Complete graph K_n

Proof Ideas using a fragmentation process

1. Start with one 1 and many 0.
2. After one (non-trivial) operation, two $\frac{1}{2}$ and many 0.
3. After another (non-trivial) operation, $\frac{1}{2}, \frac{1}{4}, \frac{1}{4},$ and many 0.
4. After another (non-trivial) operation, 3 possibilities:
   i. $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8},$ and many 0, or
   ii. $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4},$ and many 0, or
   iii. $\frac{1}{2}, \frac{3}{8}, \frac{3}{8},$ and many 0, but this is very unlikely

- Approximately, binary splitting, with $\mathrm{Exp}(2/n)$ waiting time (for each split)
- Need $\log_2(n)$ splits, to get order $1/n$
- Total time needed (for a typical number to be of order $1/n$): sum of $\log_2(n)$ many $\mathrm{Exp}(2/n)$; roughly
$$n\left(\log_2(n) + X\sqrt{\log_2(n)}\right)/2$$
  where X is standard Gaussian
- If most numbers are of order $1/n$, the L^2 distance $S(t)$ is of order $1/n$
- Use L^2 decay
$$\mathbb{E}[S(t+\epsilon)|S(t)] = e^{-\epsilon/(n-1)}\ S(t)$$
- Run for another Mn time:
$$\mathbb{E}[T(t+Mn)]^2 < \mathbb{E}[nS(t+Mn)] < e^{-M}$$
  (using Cauchy-Schwartz inequality)

Some other graphs?

(Movassagh-Szegedy-Wang, 22) L^1 decay needs $\geq n\log_2(n)/2$ time in **any** graph.

Via estimating an 'augmented entropy function'.

(Quattropani-Sau, 21) L^1 decay no cut-off in `finite-dim graph', e.g., lattice of fixed dim.

Lower bound: random walk; Upper bound: L^2 decay

(Caputo-Quattropani-Sau, 22) L^1 decay has cut-off in hypercube $\{0,1\}^d$

Same cut-off time as random walk: both $T(t)$ and $T_{RW}(t)$ decays to 0 from 2 abruptly around $t = d2^{d-2}\log(d)$, with cut-off window $O(d2^d)$

Lower bound: random walk; Upper bound: L^2 decay

A general criterion

Consider a random walk $W(t)$ coupled with the repeated average process:

- Initially, $\mathbb{P}[W(0) = v] = x_v(0)$.
- Whenever an edge $e = (u, v)$ is activated to average, and $W(t_-) = u$, then with probability $1/2$, $W(t_+) = u$; and with probability $1/2$, $W(t_+) = v$.
- Under this coupling, $W(t)$ is random on top of the repeated average process; and $\mathbb{P}[W(t) = v | \{x_u(t)\}_{u \in V}] = x_v(t)$.

❖ **We care about** $x_{W(t)}(t)$, i.e., randomly chosen from $\{x_v\}_{v \in V}$, with probability given by itself.

(Yao-Z., 26+) For a sequence of connected graphs G_n (with n vertices), suppose there are $t_n > 0$, such that as $n \to \infty$, $t_n \to \infty$, and

- (spectral gap) $\liminf_{n \to \infty} \lambda_2 n / |E| > 0$;
- (concentration in scale) $n x_{W(t_n)}(t_n) \to \infty$ and $n^{1-\epsilon} x_{W(t_n)}(t_n) \to 0$ in probability, for any $\epsilon > 0$.

Then for any $\epsilon > 0$, in probability as $n \to \infty$:

$$T(t_n) \to 2, \qquad T\big(t_n + \epsilon n \log(n)\big) \to 0.$$

By (Movassagh-Szegedy-Wang, 22), must have $t_n \geq n \log_2(n)/2$; therefore this implies cut-off.

A general criterion

(Yao-Z., 26+) For a sequence of connected graphs G_n (with n vertices), suppose there are $t_n > 0$, such that as $n \to \infty$, $t_n \to \infty$, and

- (spectral gap) $\liminf_{n\to\infty} \lambda_2 n/|E| > 0$;
- (concentration in scale) $nx_{W(t_n)}(t_n) \to \infty$ and $n^{1-\epsilon}x_{W(t_n)}(t_n) \to 0$ in probability, for any $\epsilon > 0$.

Then for any $\epsilon > 0$, in probability as $n \to \infty$: $T(t_n) \to 2$, $T\big(t_n + \epsilon n\log(n)\big) \to 0$.

Proof In G_n at time t_n, the above condition implies:

- $\sum_v x_v(t_n)\mathbf{1}[x_v(t_n) > C_n n^{-1}] \to 1$ for some $C_n \to \infty$. Then since
  $T(t_n) = \sum_v |x_v(t_n) - n^{-1}| = 2\sum_v (x_v(t_n) - n^{-1})\mathbf{1}[x_v(t_n) > n^{-1}] \geq 2\sum_v x_v(t_n)\mathbf{1}[x_v(t_n) > C_n n^{-1}] - 2/C_n$
  we have $T(t_n) \to 2$ in probability, as $n \to \infty$.
- L^2: $S(t_n) = \sum_v (x_v(t_n) - n^{-1})^2 = \sum_v \big(x_v(t_n)\big)^2 - n^{-1} = \mathbb{E}\big[x_{W(t_n)}(t_n)\big] - n^{-1} < n^{\epsilon-1}, \forall \epsilon > 0$.
- By spectral gap, $S(t_n + \epsilon' n\log(n)) = o(n^{-1})$, so $T\big(t_n + \epsilon' n\log(n)\big) \to 0$. (By Cauchy-Schwartz)

More precise understanding of $x_{W(t)}(t)$ would give better estimate of L^1 decay.

Random d regular graph

- Fix $d \geq 3$, take a uniformly random connected d regular graph with n (labeled) vertices.
- Fix any $v \in V$. Start with $x_v(0) = 1, x_w(0) = 0$ for $w \neq v$.
- Decay of $T(t) = \sum_{v \in V} |x_v(t) - n^{-1}|$?

Random walk *(Lubetzky-Sly, 08)* Cut-off at time $\frac{d}{2(d-2)} n \log_{d-1}(n)$, with Gaussian profile:

$$T_{RW}\left(\frac{d}{2(d-2)} n \log_{d-1}(n) + an\sqrt{\frac{d(d-1)}{(d-2)^3}\log_{d-1}(n)}\right) \to 2\Phi(-a), \quad \text{as } n \to \infty$$

Local tree-like *(Lubetzky-Sly, 08)* For a random d regular graph, with high probability (i.e., probability $\to 1$ as $n \to \infty$) the following is true: for any vertex v, there is at most one cycle in its $\log_{d-1}(n)/5$ neighborhood.

- *Note that diameter of a random d regular graph is* $\approx \log_{d-1}(n)$
- *Can be proved using the random pairing construction of random d regular graphs*

Random d regular graph

(Yao-Z., 26+) There are constants $C_d, D_d > 0$, such that for any a, as $n \to \infty$,

$$T\left(C_d n\log(n) + D_d a n\sqrt{\log(n)}\right) \to 2\Phi(-a),$$

in probability.

Moreover, $C_d > \frac{2d}{(d-2)\log(d-1)}$.

(The last inequality is obvious for large d, since $C_d > 1/(2\log(2))$ by *Movassagh-Szegedy-Wang, 22*).

In words, cut-off at time $C_d n\log(n)$, plus order $n\sqrt{\log(n)}$ window, and Gaussian profile.
*And cut-off time is **different** from random walk cut-off time!*

Difficulties

- Different (from random walk) cut-off time: no obvious lower bound
- Fragmentation as in complete graph? Not binary splitting

Random d regular graph

Proof Ideas Try using the general criterion

- Spectral gap: $\lambda_2 \approx d - 2\sqrt{d-1}$ (*Friedman, 04*).
- Decay of $x_{W(t)}(t)$?

Consider infinite d regular tree instead:
(start with 1 at root, and a coupled walker at the root; each edge activated at rate 1)

- Sub-additivity: $-\log(x_{W(t+s)}(t+s)) \leq -\log(x_{W(t)}(t)) - \log(x_{W(s)}(s))$ *(independent copies)*
- $\liminf_{t\to\infty} \frac{-\log(x_{W(t)}(t))}{t} > 0$, (by *Movassagh-Szegedy-Wang, 22*)

Then $\frac{-\log(x_{W(t)}(t))}{t} \to L_d > 0$ in probability as $n \to \infty$, i.e., concentration in scale. *(Let $C_d = d/(2L_d)$.)*

This sub-additivity argument should apply to any graphs with 'transitive structure' and spectral gap.

Next

- Show that $t^{-1/2}(-\log(x_{W(t)}(t)) - L_d t)$ converges to Gaussian; prove `decay of correlation in time'
- Compare d regular tree and random d regular graph
- For $C_d > \frac{2d}{(d-2)\log(d-1)}$: compare random walk to repeated average, show `extra-unbalancing'

Further questions/directions

- Other random/deterministic graphs with spectral gap (expanders)

  e.g., (connected component of) Erdos-Renyi graphs, configuration model, random geometric graph, …

  Overall, the techniques **should work**, if the graph is 'homogeneous'

- Other general pair-wise interacting processes:

  Challenge for some interactions, e.g.,

  - Random redistribution $(x_u, x_v) \to (U(x_u+x_v), (1-U)(x_u+x_v))$
  - Kac's walk in thermodynamics $(x_u, x_v) \to (\cos(\theta)x_u - \sin(\theta)x_v, \sin(\theta)x_u + \cos(\theta)x_v)$

  They do **not** converge to all being equal; need to use some different distance, such as Wasserstein-1 from optimal transport. But the techniques **should** still get decay of typical x_v.

  (See also *Caputo-Quattropani-Sau, 25* for general interactions on complete graphs)

- Beyond being 'pair-wise'? e.g., hyper-graph? (See *Caputo-Quattropani-Sau, 24* for complete hyper-graphs)