Decentralized Ranking Aggregation: Gossip Algorithms for Borda and Copeland Consensus

Anna van Elst **Kerrian Le Caillec**

Igor Colin Stephan Clémençon

LTCI, Télécom Paris, Institut Polytechnique de Paris

April 16, 2026

kerrian.lecaillec@telecom-paris.fr

Outline

1. Motivation & Decentralized Setting
2. Background and Notations on Consensus Methods
3. Proposed Algorithms and Convergence Guarantees
4. Numerical Experiments
5. Conclusion

Why Decentralized Ranking Aggregation?

Classical (Centralized)

- All n rankings sent to one server
- Server computes consensus

Problems:

- Single point of failure
- Bandwidth bottleneck: $\Omega(nm)$ bits
- Requires trusted coordinator
- Privacy concerns

Decentralized (This Work)

- Agents talk **only to neighbors**
- No central coordinator

Key idea:

- Random gossip: pairs exchange & average
- Local updates $\rightarrow$ global consensus
- $O(m)$ bits per interaction

Can we match centralized accuracy with only local communication?

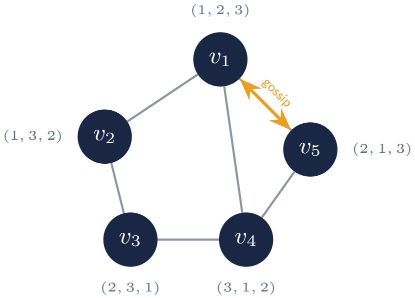
Formal Setup

Model

- $n \geq 2$ voters, $m \geq 1$ items indexed by $[\![m]\!] = \{1, \dots, m\}$
- Each voter $v \in [\![n]\!]$ holds $\sigma_v \in \mathfrak{S}_m$
- Connected communication graph $G = (V, E)$
- **Asynchronous**: each agent's clock $\sim \mathcal{P}(1)$
- Edges activated with probability $p_e > 0$; distribution $\mathbf{P} = (p_e)_{e \in E}$

Goal

Design gossip algorithms that **accurately estimate** the Borda, Copeland and Footrule consensus using **only local pairwise communication**.



Each node holds a local ranking ($m = 3$). Pairwise gossip drives convergence to the global consensus.

Randomized Gossip Averaging

Randomized averaging (Boyd et al., 2006)

Each node v maintains estimate $x_v(t)$, initialized to $x_v(0)$. At each step:

$$\text{Sample edge } e = (u, v) \sim p_e, \quad x_u(t+1) = x_v(t+1) = \frac{x_u(t) + x_v(t)}{2}.$$

All other nodes unchanged. Converges in ℓ^2 to $\bar{x} = \frac{1}{n}\sum_v x_v(0)$.

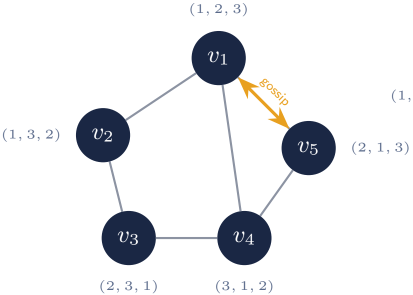
Lemma (Convergence Rate)

For each coordinate i and $t > 0$:

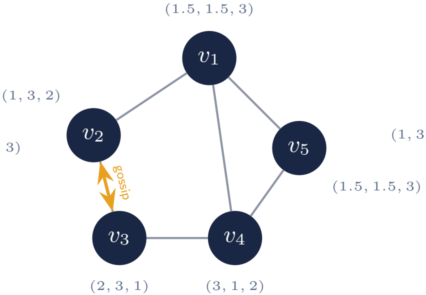
$$\mathbb{E}\left[\|X_i(t) - \bar{x}_i \mathbf{1}_n\|^2\right] \;\leq\; \left(1 - \frac{c}{2}\right)^t \|X_i(0) - \bar{x}_i \mathbf{1}_n\|^2,$$

where $c > 0$ is the **spectral gap** *of the weighted Laplacian $L(\mathbf{P}) = \sum_{e \in E} p_e L_e$.*

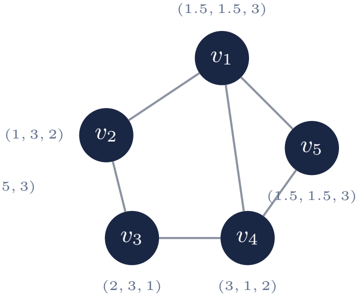
- Better-connected graph $\Leftrightarrow$ larger c $\Leftrightarrow$ faster convergence



At iteration $t = 0$.



At iteration $t = 1$.



At iteration $t = 2$.

Example of the gossip averaging process over three iterations, where each node $v \in V$ starts with an initial ranking σ_v and, under the Borda representation, communicates the corresponding vector in $\mathbb{R}^m$ associated with σ_v.

Score-based Consensus Methods

Borda (de Borda, 1781)

Score: $s_i^B = \frac{1}{n}\sum_{v=1}^{n} \sigma_v(i)$

Rank items *ascending* by s_i^B.

Copeland (Copeland, 1951)

Score: $s_i^C = \sum_{j \neq i} \left[\mathbf{1}\{\hat{p}_{ij} > \frac{1}{2}\} - \mathbf{1}\{\hat{p}_{ij} < \frac{1}{2}\} \right]$

where $\hat{p}_{ij} = \frac{1}{n}\sum_v \mathbf{1}\{\sigma_v(i) < \sigma_v(j)\}$.

Rank *descending* by s_i^C. **Condorcet-consistent.**

Metric-based methods

Kemeny (Kemeny, 1959)

$\sigma_*^K = \arg\min_{\sigma \in \mathfrak{S}_m} \sum_{v=1}^n d_\tau(\sigma, \sigma_v)$

Minimizes total Kendall-τ distance.
Satisfies Condorcet criterion.

Footrule / L^1 (Fagin et al., 2003)

Minimizes sum of $d_1(\sigma, \sigma') = \sum_{i=1}^m |\sigma(i) - \sigma'(i)|$.
Coordinate-wise *median* of ranks.

Local Kemenization Algorithm (Dwork et al., 2001)

- Scan adjacent pairs (u, v) in current ranking
- If $p_{uv} < 1/2$ (majority prefers $v \succ u$), swap them
- Repeat until no swaps occur

Output: Locally Kemeny-optimal ranking

Decentralized Borda Consensus

Pseudocode

Init: $x_v \leftarrow \sigma_v$ (own ranking vector)
for $t = 1, 2, \ldots, T$ **do**
  Select edge $(u, v) \in E$ w.p. p_e
  $x_u, x_v \leftarrow \dfrac{x_u + x_v}{2}$
end for
Sort x_v ascending $\rightarrow \hat{\sigma}_v$
return $\hat{\sigma}_v(T)$ (Borda estimate)

- Each voter $v \in V$ encodes σ_v as a full Borda score vector in $\mathbb{R}^m$
- Gossip averaging yields estimates of the mean Borda scores
- The consensus ranking is obtained by sorting the aggregated scores

Decentralized Copeland Consensus

Pseudocode

Init: $x_v \leftarrow \left(\hat{p}_{ij}^v\right)_{i \neq j}$, $\hat{p}_{ij}^v = \mathbf{1}[\sigma_v(i) < \sigma_v(j)]$
for $t = 1, 2, \ldots, T$ **do**
  Select edge (u, v); $x_u, x_v \leftarrow \frac{x_u + x_v}{2}$
end for
$s_{v,i} = \sum_{j \neq i} \left[\mathbf{1}(x_{v,(i,j)} \leq \frac{1}{2}) - \mathbf{1}(x_{v,(i,j)} > \frac{1}{2})\right]$
Sort $s_{v,\cdot}$ descending $\rightarrow \hat{\sigma}_v$
return $\hat{\sigma}_v(T)$

- Each voter $v \in [\![n]\!]$ estimates the pairwise probability matrix $(p_{ij})_{i \neq j}$
- Then, v computes locally Copeland scores.
- **Consensus σ_*^C is partial ranking:** $i \prec_{\sigma_*^C} j \iff s_i^C < s_j^C$. When $s_i^C = s_j^C$, items i and j are incomparable.

Convergence of Algorithms

Theorem (Borda Convergence)

Assume Borda scores are distinct. For any $t > 0$*:*

$$\frac{1}{n}\sum_{v=1}^{n}\mathbb{E}\big[d_\tau(\sigma_*^B, \hat{\sigma}_v(t))\big] \leq \min\big\{C_1 e^{-c_1 t},\ C_2 e^{-c_2 t}\big\}$$

with rates $c_1 = c/2$, $c_2 = c/4$, $\Delta^B = \min_{i \neq j} |s_i^B - s_j^B|$ *(min score gap),* $\gamma^B = \sum_i \|S_i(0) - s_i \mathbf{1}_n\|^2$*:*

$$C_1 = \frac{2(m-1)\gamma^B}{n(\Delta^B)^2}, \quad C_2 = \frac{(m-1)\sqrt{m}\,\gamma^B}{\sqrt{n}\,\Delta^B},$$

When scores are tied ($s_i^B = s_j^B$**):**

- Tied pairs contribute zero to d_τ (any ordering is consistent with partial consensus)
- Δ^B considers only non-tied pairs; if all scores equal, bound is vacuous

Convergence of Algorithms

Theorem (Copeland Convergence)

Under weak stochastic transitivity, for any $t > 0$:

$$\frac{1}{n}\sum_{v=1}^{n}\mathbb{E}\big[d_\tau(\sigma_*^C, \hat{\sigma}_v(t))\big] \;\leq\; \frac{m-1}{\Delta^C}\min\big\{K_1 e^{-c_1 t},\; K_2 e^{-c_2 t}\big\}$$

with $c_1 = c/2$, $c_2 = c/4$ as before, $\Delta^C = \min_{i\neq j}|s_i^C - s_j^C|$, and K_1, K_2 depending on initial pairwise estimates and margins $\delta_{ij} = |\hat{p}_{ij} - \frac{1}{2}|$.

Handling pairwise ties:

- $J_i := \{j \neq i : \hat{p}_{ij} \neq 1/2\}$ — only non-tied pairwise comparisons
- When $\hat{p}_{ij} = 1/2$, pair (i, j) excluded from score calculation
- Weak stochastic transitivity excludes cycles but permits ties

Extensions: Footrule & Decentralized Kemenization

Decentralized Footrule

Init: $x_v \leftarrow \sigma_v \in \mathbb{R}^m$
for $t = 1, 2, \ldots, T$ **do**
  Select edge (u, v)
  Perform asynchronous ADMM update on (x_u, x_v) toward coordinate-wise median
end for
$\hat{x}_v \leftarrow x_v$
Sort $\hat{x}_v$ in ascending order $\rightarrow \hat{\sigma}_v$
return $\hat{\sigma}_v(T)$

Note: Median estimation is computationally harder than averaging, leading to slower convergence.

Decentralized Local Kemenization

Init: $x_v \leftarrow (\hat{p}^v_{ij})_{i \neq j}, \quad \hat{p}^v_{ij} = \mathbf{1}[\sigma_v(i) < \sigma_v(j)]$
for $t = 1, 2, \ldots, T$ **do**
  Select edge (u, v); $x_u, x_v \leftarrow \frac{x_u + x_v}{2}$
end for
Compute preliminary ranking $\tilde{\sigma}_v$ from x_v
Apply local Kemenization using $x_{v,(i,j)}$ as estimates of p_{ij}
return $\hat{\sigma}_v(T)$

All four algorithms transmit $O(m)$ or $O(m^2)$ values per interaction, matching the information-theoretic lower bound.

Synthetic Experiments: Mallows Model

Setup

$P(\sigma) \propto \phi^{d_\tau(\sigma,\sigma_0)}$, $\quad n = 151$, $m = 8$, $\phi = 0.5$, $T = 2{,}000$, $N = 100$ trials.
Graph types: **complete**, **Watts–Strogatz**, **geometric**.

MSE of Scores (Fig. 2a,b)

- Borda & Footrule: MSE < 0.5 in $< 1{,}000$ iterations across all topologies
- Footrule slightly slower (median $>$ mean difficulty)
- Copeland pairwise scores converge fast and stably

Kendall-τ Error (Fig. 2c)

- Rapid convergence for Borda, Copeland, Footrule
- Ordering: Complete $>$ Watts–Strogatz $>$ Geometric
- Consistent with connectivity hierarchy

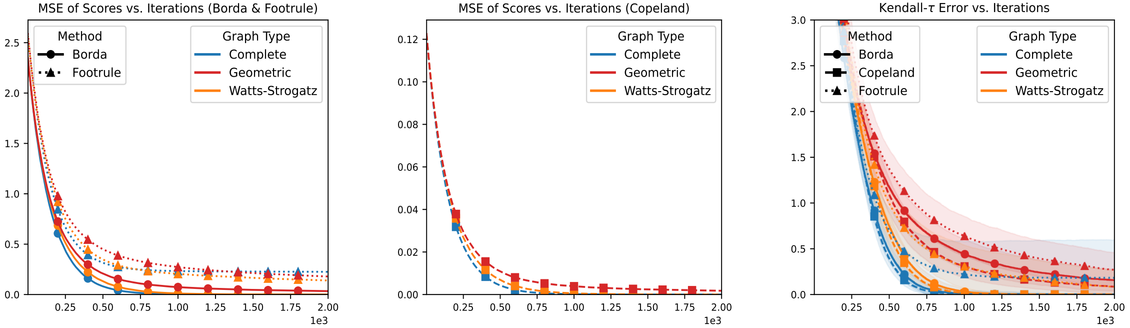
Exponential Fit

Empirical decay matches the theoretical bound:

$$\text{MSE}(t) \simeq C\, e^{-ct/2}$$

Both Borda and Copeland confirm the $O(e^{-ct/2})$ rate in practice.

Synthetic Data Results



(a) MSE of Borda/Footrule scores

(b) MSE of pairwise probabilities

(c) d_τ error to (respective) consensus

Figure: Convergence of consensus methods (Borda, Copeland, Footrule) on rankings sampled from a Mallows model ($n = 151$ agents, $m = 8$ items, $\phi = 0.5$) across different graph topologies. Results show mean $\pm$ standard deviation over $N = 100$ trials.

Real-World Datasets

Datasets

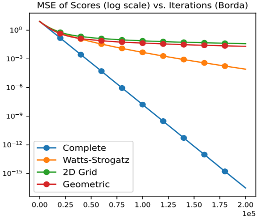
- **Sushi**: $n = 5{,}000$, $m = 10$, strict complete rankings.
  $N = 15$ trials, $T = 200{,}000$ iterations.
- **Debian**: $n = 504$, $m = 7$, *partial* rankings.
  $N = 250$ trials, $T = 50{,}000$ iterations.

Additional topology: **2D grid**.

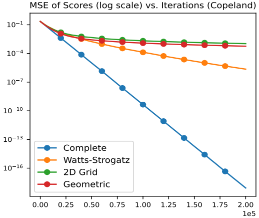
Partial Ranking Handling

- Borda: normalize ranked items, impute average rank for unranked ones
- Copeland: use weak pairwise preferences $\hat{p}_{ij} = \mathbf{1}\{\sigma(i) < \sigma(j)\} + \frac{1}{2}\mathbf{1}\{\sigma(i) = \sigma(j)\}$

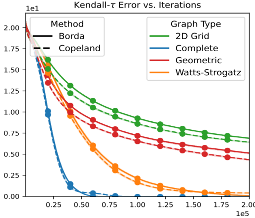
Dataset-based Results



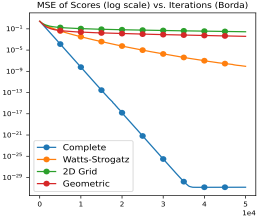
(a) MSE of Borda scores



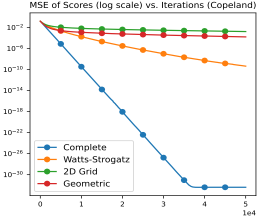
(b) MSE of Copeland scores



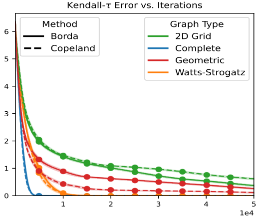
(c) d_τ error to (respective) consensus



(d) MSE of Borda scores



(e) MSE of Copeland scores



(f) d_τ error to (respective) consensus

Conclusion & Perspectives

- **Gossip algorithms** for decentralized Borda, Copeland, Footrule, and Local Kemenization
- **Convergence guarantees:** $O(e^{-ct/2})$ for Borda & Copeland; rate controlled by graph spectral gap c
- **Robustness analysis:** Copeland and Kemeny are superior; Local Kemenization recovers Kemeny in all tested settings

Research directions

- **Byzantine robustness** — agents actively sending communicating wrong information.
- **Adaptive gossip** — topology/data-aware edge sampling.
- **Privacy-preserving gossip** — differentially private pairwise scores.
- **Finite-time guarantees** for Footrule relaxation.

Thank you for your attention!

Appendix: Gossip Averaging Proof Sketch

Lemma (Gossip Matrix Properties)

At each iteration, if edge (i, j) is selected w.p. p_e, the transition matrix is $W_2(t) = I_n - L_e/2$ where $L_e = (e_i - e_j)(e_i - e_j)^\top$.

- *$W_2(t)$ is symmetric and doubly stochastic; $W_2(t)^2 = W_2(t)$.*
- *$\bar{W}_2 = \mathbb{E}[W_2(t)] = I_n - \frac{1}{2}\sum_e p_e L_e$ is doubly stochastic.*
- *$\tilde{W}_2 := \bar{W}_2 - \mathbf{1}\mathbf{1}^\top/n$ satisfies $\|\tilde{W}_2\|_{\mathrm{op}} \leq \lambda_2 = 1 - c/2$.*

Proof. Let $Y(t) = X_i(t) - \bar{x}_i \mathbf{1}_n$. Then:

$$\mathbb{E}\big[\|Y(t)\|^2 \mid Y(t-1)\big] = Y(t-1)^\top \mathbb{E}[W_2(t)]\, Y(t-1) \leq \lambda_2 \|Y(t-1)\|^2.$$

Iterating and using $\lambda_2 = 1 - c/2$:

$$\mathbb{E}[\|Y(t)\|^2] \leq \left(1 - \tfrac{c}{2}\right)^t \|Y(0)\|^2.$$

Appendix: AsylADMM Algorithm

Notation:

- d_k: degree of node k, $\mathcal{N}_k$: edges incident to node k
- $\mathrm{prox}_{\gamma f}(z) = \mathrm{argmin}_x \left\{ f(x) + \frac{1}{2\gamma}\|x - z\|^2 \right\}$

Distributed optimization problem:

$$\underset{x \in \mathbb{R}^n}{\mathrm{argmin}} \sum_{i=1}^{n} f_i(x_i) \quad \text{s.t.} \quad x_i = x_j \ \ \forall (i,j) \in E$$

ADMM reformulation:

$$\min_{x \in \mathbb{R}^n,\, z \in \mathbb{R}^{2m}} f(x) + g(z) \quad \text{s.t.} \quad Mx = z$$

where $f(x) = \sum_{i=1}^{n} f_i(x_i)$, $g(z) = \sum_{e \in E} \iota_C(z_e)$, $C = \mathrm{span}(\mathbf{1}_2)$

Example (quantile estimation via pinball loss):

$$f_k(x) = L_\alpha(a_k - x) = \big(\alpha - \mathbb{I}\{a_k - x \leq 0\}\big)(a_k - x)$$

Appendix: AsylADMM Algorithm

AsylADMM: Async & Lite ADMM

Require: Initial vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$; step size $\rho > 0$

Initialization:
for all nodes $v = 1, \ldots, n$ **do**
    $\mathbf{x}_v \leftarrow \mathbf{a}_v$
    $\hat{\mu}_k \leftarrow 0$
end for
for $t = 0, 1, \ldots$ **do**
    Draw edge $e = (u, v) \in E$ with probability p_e
    Compute average $\mathbf{z}_e \leftarrow \frac{1}{2}(\mathbf{x}_u + \mathbf{x}_v)$
    for agents $k \in \{u, v\}$ **do**
        $\hat{\mu}_k \leftarrow \hat{\mu}_k + \frac{\rho}{d_k}(\mathbf{z}_e - \mathbf{x}_k)$
        $x_k \leftarrow \operatorname{prox}_{f_k}^{\rho d_k}\left(\mathbf{z}_e + \frac{\hat{\mu}_k}{\rho}\right)$
    end for
end for

Robustness Under Data Contamination

Setup: $n = 100$, $m = 8$, $\phi = 0.5$, contamination rate $\varepsilon = 0.3$.

Table 2 — Adversarial contamination (ε samples $\sim$ Mallows(σ_0^{-1}, φ))

Method	$d_\tau(\cdot, \sigma_0)$	ΔL w/o LK	ΔL w/ LK
Borda	0.91 ± 0.85	0.03 ± 0.04	0.00 ± 0.00
Copeland	0.59 ± 0.71	0.00 ± 0.02	0.00 ± 0.01
Footrule	1.26 ± 0.99	0.13 ± 0.11	0.01 ± 0.02
Kemeny	0.58 ± 0.67	0.00 ± 0.00	0.00 ± 0.00

Key Takeaway

Kemeny $\approx$ **Copeland** $\gg$ **Borda** $\gg$ **Footrule** in contamination robustness.

Local Kemenization applied to Borda or Copeland recovers the Kemeny consensus in virtually all cases — a lightweight but highly effective post-processing step.

Summary of Theoretical Guarantees

Main Results at a Glance

Method	Rate	Comm. per step	Condorcet?
Borda	$O(e^{-ct/2})$	$O(m)$	No
Copeland	$O(e^{-ct/2})$	$O(m^2)$	Yes
Footrule	(AsylADMM rate)	$O(m)$	No
Local Kem.	$O(e^{-ct/2})$	$O(m^2)$	Yes

What drives convergence speed?

- **Spectral gap** c depends on graph topology — complete $\geq$ Watts–Strogatz $\geq$ geometric $\geq$ 2D grid