Speed-Accuracy Trade-off in Collective Decision Making

Vaibhav Srivastava

Department of Mechanical & Aerospace Engineering
Princeton University

Joint work with: Naomi Leonard

December 10, 2013

IEEE Conference on Decision and Control
Florence, Italy

Collective Decision Making in two alternative choice tasks


Birds deciding whether to migrate or not

Collective Decision Making in two alternative choice tasks


Birds deciding whether to migrate or not

Leader election in a two party system

Collective Decision Making in two alternative choice tasks

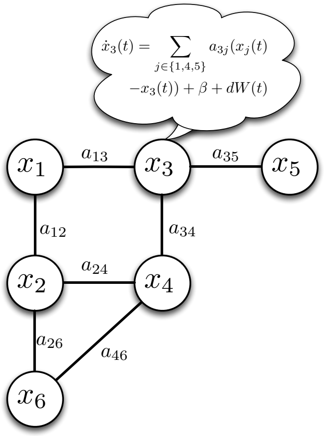

Birds deciding whether to migrate or not

Leader election in a two party system

Social information assimilation + decision-making = Socio-Cognitive Networks

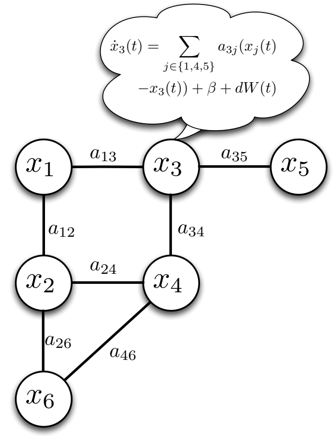
Collective Decision Making in Socio-Cognitive Networks

Alternative 1
versus
Alternative 2



Collective Decision Making in Socio-Cognitive Networks

Alternative 1
versus
Alternative 2



- Ideal group versus Condorcet group

R. D. Sorkin, C. J. Hays, and R. West. Signal-detection analysis of group decision making. *Psychological review*, 108(1):183, 2001

P. Braca, S. Marano, V. Matta, and P. Willett. Asymptotic optimality of running consensus in testing binary hypotheses. *IEEE Transactions on Signal Processing*, 58(2):814–825, 2010

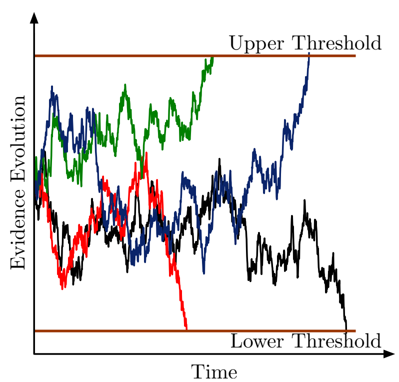
Drift Diffusion Model and the Free Response Paradigm

- Models human decision making in two alternative choice tasks
- Evidence evolution in a two alternative choice task is modeled by

$$dx(t) = \beta dt + dW(t), \quad x(t) = x_0$$

- Decision process at time τ is

$$\begin{cases} x(\tau) > \eta, & \text{choose alternative 1,} \\ x(\tau) < -\eta, & \text{choose alternative 2,} \\ \text{else,} & \text{collect more evidence.} \end{cases}$$



R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen. The physics of optimal decision making: A formal analysis of performance in two-alternative forced choice tasks. *Psychological Review*, 113(4):700–765, 2006

Social Interaction and the DeGroot Model

- $\mathbf{p}$: vector of opinions in a network
- A: row stochastic matrix
- models consensus seeking in a social network by

$$\mathbf{p}(t+1) = A\mathbf{p}(t).$$

- same as the celebrated consensus dynamics in multi-agent systems
- Continuous time consensus seeking in a social network modeled by

$$\dot{\mathbf{p}}(t) = -L\mathbf{p}(t), \quad \mathbf{p}(0) = \mathbf{p}_0 \qquad L = \text{Laplacian Matrix}$$

M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974

J. N. Tsitsiklis. *Problems in Decentralized Decision Making and Computation*. PhD thesis, Massachusetts Institute of Technology, November 1984

A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003

R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007

Coupled Drift Diffusion Model

- n decision-makers collect noisy signals and interact with each other
- the evidence aggregation process well modeled by

$$d\mathbf{x}(t) = \underbrace{-L\mathbf{x}(t)dt}_{\text{social term}} + \underbrace{\beta\mathbf{1}_n dt + \sigma dW(t)}_{\text{noisy signal}}, \quad \mathbf{x}(0) = \mathbf{0}_n. \qquad (1)$$

Quantities of interest:

- Expected decision times
- Error rates (probability of wrong decision)

Coupled Drift Diffusion Model

- n decision-makers collect noisy signals and interact with each other
- the evidence aggregation process well modeled by

$$d\mathbf{x}(t) = \underbrace{-L\mathbf{x}(t)dt}_{\text{social term}} + \underbrace{\beta\mathbf{1}_n dt + \sigma dW(t)}_{\text{noisy signal}}, \quad \mathbf{x}(0) = \mathbf{0}_n. \qquad (1)$$

Quantities of interest:

- Expected decision times
- Error rates (probability of wrong decision)

Standard approach:

- solve first passage time associated with the FP equation for (1)
- an elliptic PDE with n variables

I. Poulakakis, L. Scardovi, and N. E. Leonard. Node classification in networks of stochastic evidence accumulators. *arXiv preprint arXiv:1210.4235*, October 2012

Asymptotic Optimality of the Coupled DDM

- **Evidence vector:** $\mathbf{x}(t) = x_{\text{cen}}(t)\mathbf{1}_n + \boldsymbol{\epsilon}(t)$

$$\mathrm{d}x_{\text{cen}}(t) = \beta\mathrm{d}t + \frac{1}{n}\mathbf{1}_n^\top \mathrm{d}\mathbf{W}(t), \; x_{\text{cen}}(0) = 0$$

$$\mathrm{d}\boldsymbol{\epsilon}(t) = -L\boldsymbol{\epsilon}(t)\mathrm{d}t + (I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top)\mathrm{d}\mathbf{W}_n(t), \boldsymbol{\epsilon}(0) = \mathbf{0}_n.$$

- $\epsilon_k(t) \to \mathcal{N}(0, 1/\mu_k), \qquad \frac{1}{\mu_k} = \sum_{p=2}^n \frac{1}{2\lambda_p} {u_k^{(p)}}^2$
- μ_k is a certainty index determined purely by the interaction graph

Asymptotic Optimality of the Coupled DDM

- **Evidence vector:** $\mathbf{x}(t) = x_{\text{cen}}(t)\mathbf{1}_n + \boldsymbol{\epsilon}(t)$

$$\mathrm{d}x_{\text{cen}}(t) = \beta\mathrm{d}t + \frac{1}{n}\mathbf{1}_n^\top \mathrm{d}\mathbf{W}(t), \; x_{\text{cen}}(0) = 0$$

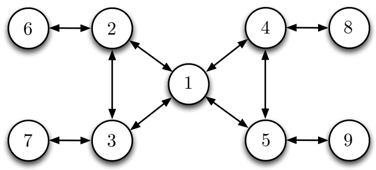
$$\mathrm{d}\boldsymbol{\epsilon}(t) = -L\boldsymbol{\epsilon}(t)\mathrm{d}t + (I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top)\mathrm{d}\mathbf{W}_n(t), \boldsymbol{\epsilon}(0) = \mathbf{0}_n.$$

- $\epsilon_k(t) \to \mathcal{N}(0, 1/\mu_k), \qquad \frac{1}{\mu_k} = \sum_{p=2}^n \frac{1}{2\lambda_p} {u_k^{(p)}}^2$
- μ_k is a certainty index determined purely by the interaction graph

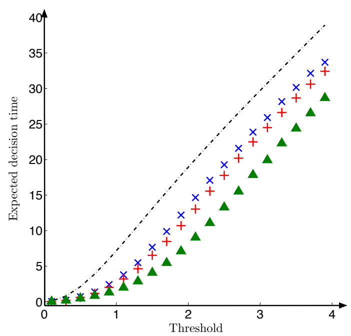
Asymptotic optimality

$$\frac{x_k(t) - \beta t}{\sqrt{t}} = \frac{x_{\text{cen}}(t) - \beta t}{\sqrt{t}} + \frac{\epsilon_k(t)}{\sqrt{t}} \implies x_k(t) = x_{\text{cen}}(t) + o(1)$$

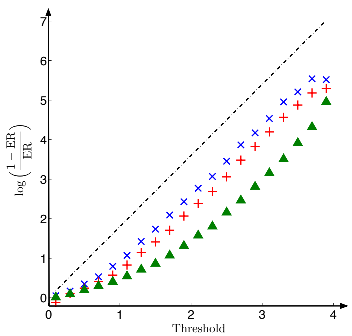
Numerical Illustration: Asymptotic Optimality



Interaction graph



Expected decision time



Log odds of error rates

Decoupled Approximation to the Coupled DDM

- decoupled approximation to $\boldsymbol{\epsilon}(t)$

$$\mathrm{d}\boldsymbol{\epsilon}(t) = -L\boldsymbol{\epsilon}(t)\mathrm{d}t + (I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top)\mathrm{d}\mathbf{W}_n(t), \boldsymbol{\epsilon}(0) = \mathbf{0}_n$$

- $\epsilon_k(t)$ is a continuous Gaussian process and converges to $\mathcal{N}(0, 1/\mu_k)$

Decoupled Approximation to the Coupled DDM

- decoupled approximation to $\boldsymbol{\epsilon}(t)$

$$\mathrm{d}\boldsymbol{\epsilon}(t) = -L\boldsymbol{\epsilon}(t)\mathrm{d}t + (I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top)\mathrm{d}\mathbf{W}_n(t), \boldsymbol{\epsilon}(0) = \mathbf{0}_n$$

- $\epsilon_k(t)$ is a continuous Gaussian process and converges to $\mathcal{N}(0, 1/\mu_k)$
- approximate $\epsilon_k(t)$ by the O-U process

$$\mathrm{d}\varepsilon_k(t) = -\frac{\mu_k}{2}\varepsilon_k(t) + \mathrm{d}W(t), \quad \varepsilon_k(0) = 0$$

Decoupled Approximation to the Coupled DDM

- decoupled approximation to $\boldsymbol{\epsilon}(t)$

$$\mathrm{d}\boldsymbol{\epsilon}(t) = -L\boldsymbol{\epsilon}(t)\mathrm{d}t + (I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top)\mathrm{d}\mathbf{W}_n(t), \boldsymbol{\epsilon}(0) = \mathbf{0}_n$$

- $\epsilon_k(t)$ is a continuous Gaussian process and converges to $\mathcal{N}(0, 1/\mu_k)$
- approximate $\epsilon_k(t)$ by the O-U process

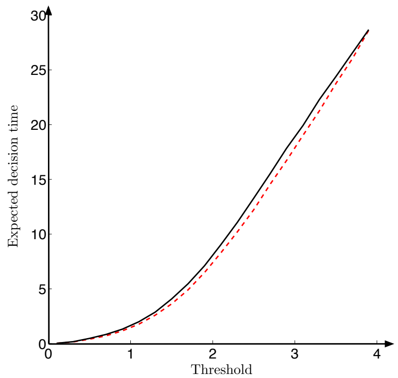
$$\mathrm{d}\varepsilon_k(t) = -\frac{\mu_k}{2}\varepsilon_k(t) + \mathrm{d}W(t), \quad \varepsilon_k(0) = 0$$

Efficiency of approximation

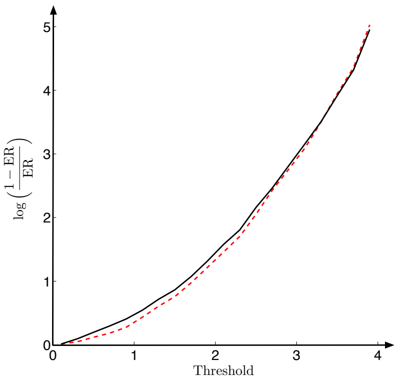
$$\lim_{t\to+\infty} \mathrm{corr}(\epsilon_k(t), \varepsilon_k(t)) = \mu_k \sum_{p=1}^{n} \frac{1}{2\mathrm{eig}_p(L + \mathrm{diag}(\boldsymbol{\mu}/2))}(\tilde{u}_k^{(p)})^2 - \frac{2}{n}$$

1. approximate evidence at node k: $x_{\mathrm{cen}}(t) + \varepsilon_k(t)$
2. **Decision time and Error Rate:** need to solve n elliptic PDEs with two variables opposed to a PDE with n variables earlier

Numerical Illustration: Decoupled Approximation



Expected decision time



Log odds of error rates

The reduced DDM approximates the coupled DDM well

Further Approximations

- bound the contribution by the O-U process $\varepsilon_k(t)$
- for sufficiently large K, with high probability

$$\max_{s\in[0,t]} |\varepsilon_k(t)| \leq \frac{K}{\sqrt{\mu_k}}$$

- effective threshold for the centralized DDM belongs to the set $(\eta - K/\sqrt{\mu_k}, \eta + K/\sqrt{\mu_k})$

Further Approximations

- bound the contribution by the O-U process $\varepsilon_k(t)$
- for sufficiently large K, with high probability

$$\max_{s\in[0,t]} |\varepsilon_k(t)| \leq \frac{K}{\sqrt{\mu_k}}$$

- effective threshold for the centralized DDM belongs to the set $(\eta - K/\sqrt{\mu_k}, \eta + K/\sqrt{\mu_k})$

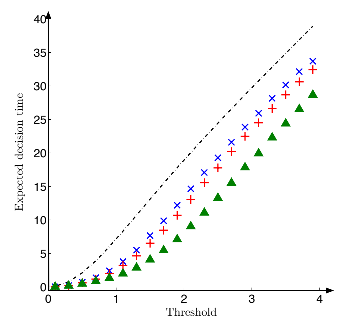
Bounds on Decision Time and Error Rates

$$\frac{\eta_k - \frac{K}{\sqrt{\mu_k}}}{\beta} \tanh\left(\beta n\left(\eta_k - \frac{K}{\sqrt{\mu_k}}\right)\right) \leq \mathrm{ET}_k \leq \frac{\eta_k + \frac{K}{\sqrt{\mu_k}}}{\beta} \tanh\left(\beta n\left(\eta_k + \frac{K}{\sqrt{\mu_k}}\right)\right)$$

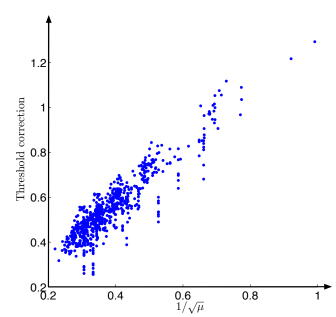
$$\frac{1}{1+\exp\left(2\beta n(\eta_k + \frac{K}{\sqrt{\mu}})\right)} \leq \mathrm{ER}_k \leq \frac{1}{1+\exp\left(2\beta n(\eta_k - \frac{K}{\sqrt{\mu}})\right)}$$

V. Srivastava (Princeton University) Collective Decision Making Dec 10, 2013 11 / 14

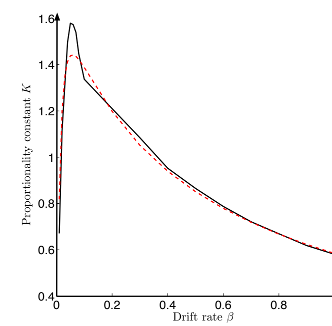
Empirical Estimates for Threshold Correction



Expected decision time



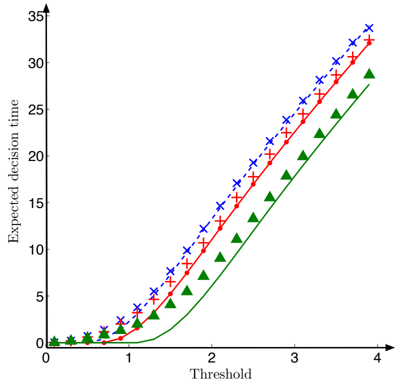
Log odds of error rates



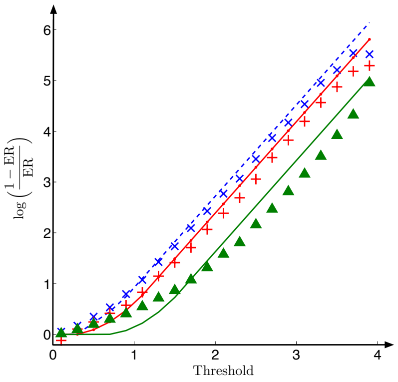
Log odds of error rates

1. coupled DDM at each node well approximated by centralized DDM with a modified threshold
2. effective threshold at node $k = \eta - \frac{K(\beta)}{\sqrt{\mu_k}}$, $\quad K(\beta) = \frac{e^{-\frac{1}{4\sqrt{\beta}}}}{\sqrt{\beta}(1+\beta/3)}$

Numerical Illustration: Threshold Corrected Centralized DDM



Expected decision time



Log odds of error rates

The centralized DDM with corrected thresholds approximates the coupled DDM well

Conclusions and Future Directions

Conclusions:

1. towards rigorous modeling and analysis of socio-cognitive networks
2. coupled DDM as model for social decision-making in 2-AC tasks
3. a computationally tractable decoupled approximation to coupled DDM
4. further approximation by the threshold corrected centralized DDM
5. ideas extend to multi-alternative choice tasks
   and 2-AC tasks with recency effect

Future Directions:

1. relaxing the continuous communication assumption
2. heterogeneous individuals
3. general decision-making tasks, e.g., multi-armed bandits