

Heterogeneity Improves Speed and Accuracy in Social Networks

Bhargav Karamched^{1,2,*} Megan Stickler³ William Ott³ Benjamin Lindner^{4,5}

Zachary P. Kilpatrick^{6,†} and Krešimir Josić^{3,7,‡}

¹Department of Mathematics, Florida State University, Tallahassee, Florida 32306, USA

²Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306, USA

³Department of Mathematics, University of Houston, Houston, Texas 77004, USA

⁴Physics Department of Humboldt University Berlin, Newtonstraße 15, 12489 Berlin, Germany

⁵Bernstein Center for Computational Neuroscience Berlin, Philippstraße 13, Haus 2, 10115 Berlin, Germany

⁶Department of Applied Mathematics, University of Colorado Boulder, Boulder, Colorado 80309, USA

⁷Department of Biology and Biochemistry, University of Houston, Houston, Texas 77004, USA



(Received 11 July 2020; revised 4 September 2020; accepted 24 September 2020; published 16 November 2020)

How does temporally structured private and social information shape collective decisions? To address this question we consider a network of rational agents who independently accumulate private evidence that triggers a decision upon reaching a threshold. When seen by the whole network, the first agent's choice initiates a wave of new decisions; later decisions have less impact. In heterogeneous networks, first decisions are made quickly by impulsive individuals who need little evidence to make a choice but, even when wrong, can reveal the correct options to nearly everyone else. We conclude that groups comprised of diverse individuals can make more efficient decisions than homogenous ones.

DOI: [10.1103/PhysRevLett.125.218302](https://doi.org/10.1103/PhysRevLett.125.218302)

A central question in biology, sociology, and economics is how the exchange of information shapes group decisions [1–7]. Various organisms observe the choices of their peers to guide their own decisions [8–11]: Argentinian ants form trails by following their peers [12], African wild dogs depart a congregation in response to their neighbor's sneezes [13], and pedestrians look to each other to decide when to cross a road [14].

How do individuals combine private evidence and social information to make decisions? To address this question we have proposed a tractable model of collective decision making and analyzed decisions in small networks [15]. Here we extend this work to large heterogeneous networks. We show that in a group of identical agents, a wrong first decision leads approximately half the network astray. However, in heterogeneous networks a wrong first choice is usually made by hasty, uninformed agents and only convinces others who are similarly quick to decide. Cautious agents can observe the decisions of early adopters and make the right choice. Thus, in diverse groups decisions by unreliable agents, even when wrong, can reveal the better option.

Previous models of collective decision making ignored temporal aspects of evidence accumulation [6,16,17] or did not describe rational agents [18,19]. Our model incorporates both aspects and allows us to understand decision makers' departure from rationality [20].

Model description.—We consider an all-to-all network, or clique, of agents, each deciding between two options (Fig. 1). Like day traders, or strangers in a market, agents

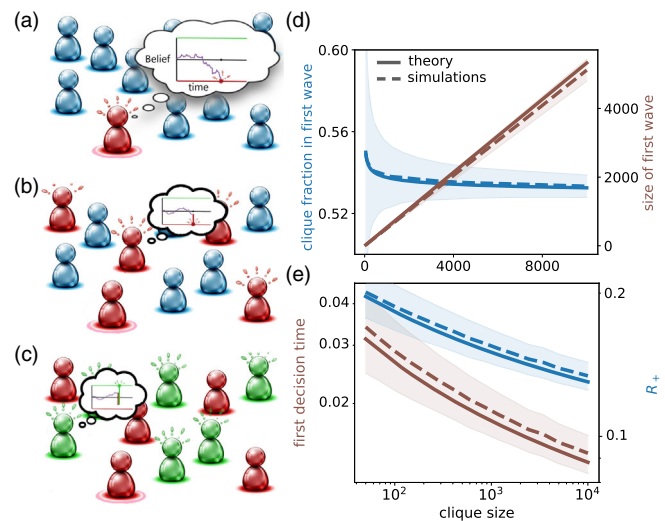


FIG. 1. Waves of collective decisions. (a) The first in a clique of identical agents gathers sufficient private evidence but decides incorrectly (red). (b) The first decision convinces a few agents to agree. Since this wave is small, it reveals to undecided (blue) agents that the first decision was likely wrong. (c) The difference between decided and undecided agents leads the remaining agents to choose correctly (green). (d) The first wave increases with N (red) but comprises a smaller fraction of the population [blue; Eq. (5)]. Here, the first decision is correct. (e) The time to the first decision decreases with network size (red), allowing each agent less time to accumulate private information [Eq. (4)]. Information provided by an individual first wave decision also decreases [R_+ , blue; Eq. (6)]. $\theta = 0.7$ in (d) and (e). Here, and below, solid and dashed lines represent simulations and theory, respectively, and shaded regions capture one standard deviation around the mean.

make private observations and gather social evidence by observing the choices of all other agents. They do not share private information but know the statistics of the observations each agent makes. A decision cannot be undone.

An example provides intuition: consider a group of people deciding between two products to buy. They study the products' specifications and read reviews, making a sequence of private observations. They also observe which product their friends choose. Each person combines private observations (product reviews) with social information (decisions of friends). They do not exchange information directly but know the type of information their friends gather, and thus how beliefs evolve [6,16]. Once purchased, the product cannot be returned.

Evolution of beliefs: We assume N agents accumulate noisy private observations and optimally combine them with information obtained from observing the decisions of their neighbors to choose between two hypotheses, H^+ or H^- . Either hypothesis is *a priori* equally likely to be correct. Each agent, i , makes decisions based on their belief, $y_i(t)$, which equals the log-likelihood ratio (LLR) between the hypotheses given all available evidence [21]. After a sequence of private observations, $\xi_{1:t}^{(i)}$, the belief is $y_i(t) = \log[\mathbb{P}(H^+|\xi_{1:t}^{(i)})/\mathbb{P}(H^-|\xi_{1:t}^{(i)})]$. If private observations are rapid and uncorrelated in time and between agents, beliefs evolve as

$$dy_i = \pm \alpha dt + \sqrt{2\alpha} dW_i, \quad (1)$$

where the sign of the drift equals that of the correct hypotheses, and $W_i(t)$ are independent, standard Wiener processes [22,23]. Each observer starts with no evidence, so $y_i(0) = 0$. We assume henceforth that H^+ is correct, and that $\alpha = 1$. When H^- is correct or $\alpha \neq 1$ the analysis is similar.

Each agent, i , sets a threshold, θ_i , and chooses H^+ (H^-) at time T_i if $y_i(T_i) \geq \theta_i$ [$y_i(T_i) \leq -\theta_i$], and $y_i(t) \in (-\theta_i, \theta_i)$ for $0 \leq t < T_i$. All other agents observe a decider's choice, but may not know their threshold. We consider omniscient agents who know each other's thresholds and the case of consensus bias where each agent assumes all others have the same threshold they do.

Belief updates from decision: Without loss of generality, we assume the belief of agent $i = 1$ is the first to reach threshold at time $t = T$ [Fig. 1(a)].

Until this decision, beliefs of all agents, $y_i(t)$ with $i = 2, \dots, N$, evolve independently according to Eq. (1). Upon observing the first decision, omniscient agents update their belief by the evidence independently accumulated by the first decider, $y_i(T) \rightarrow y_i(T) \pm \theta_1$ [15] (see the Supplemental Material, Sec. III [24]). Observing a positive (H^+) first decision causes any belief that satisfies $y_i(T^-) \in [\theta_i - \theta_1, \theta_i]$, to cross the positive threshold, θ_i , evoking a positive decision by agent i . Agents subject to consensus bias update their belief as $y_i(T) \rightarrow y_i(T) \pm \theta_i$.

Contrasting with previous work [15], we assume agents make decisions in synchronous waves, ending all social information exchange before accumulating further private evidence: A wave of a_1 agreeing agents follows the first choice [Fig. 1(b)]. Each of the remaining $N - a_1 - 1$ undecided agents then obtains information by observing who followed the first decision and who remained undecided. How do the undecided agents make use of this newly revealed information?

Homogeneous populations.—To answer this question, first suppose agents have identical thresholds, $\theta_i = \theta$, for all i , so that omniscience and consensus bias are equivalent. Observing that agent $i \neq 1$ follows a positive first decision tells other agents that $y_i(T^-) \in [0, \theta]$. Therefore observing a first wave decision of agent i leads to an increment in belief equal to [15]

$$\begin{aligned} \text{LLR}[y_i(T) \in (0, \theta)] &\stackrel{\text{def}}{=} \log \left(\frac{\mathbb{P}[y_i(T) \in [0, \theta] | H^+]}{\mathbb{P}[y_i(T) \in [0, \theta] | H^-]} \right) \\ &= \log \left(\frac{\int_0^\theta p_+(x, T) dx}{\int_0^\theta p_-(x, T) dx} \right) \equiv R_+(T). \end{aligned}$$

Here $p_\pm(x, t)\Delta x = \mathbb{P}[y_i(t) \in (x, x + \Delta x) | H^\pm] + O(\Delta x^2)$ is the conditional probability density for the belief of agent i at time t . Since thresholds are symmetric, $\int_0^\theta p_-(x, t) dx = \int_{-\theta}^0 p_+(x, t) dx$, so observing an agent j who remains undecided after the first decision reveals $y_j(T) \in (-\theta, 0]$, leading to an increment $\text{LLR}\{y_j(T) \in (-\theta, 0]\} \equiv R_-(T) = -R_+(T)$. Thus, in contrast to previous work [6,7,27,28], the first decision, but not subsequent ones, reveals the exact information gathered by the decider.

Agents know the statistics of private observations and can compute $p_+(x, t)$ and $p_-(x, t)$. Thus agents know that beliefs evolve according to Eq. (1) and that the belief distribution prior to any decision satisfies:

$$\partial_t p_\pm = \mp \partial_x p_\pm + \partial_{xx}^2 p_\pm, \quad p_\pm(\pm\theta, t) = 0, \quad (2)$$

if H^\pm is correct, with $p_\pm(x, 0) = \delta(x)$. Agents do not know which hypothesis is correct, and compute the belief update, $R_+(T)$, using only belief distributions, $p_\pm(x, t)$.

Agents undecided after the first wave combine the information from all observed decisions and indecisions. Since private measurements are independent, information obtained from agents in the first wave is additive. The resulting belief increment is

$$\begin{aligned} c_1^+ &\stackrel{\text{def}}{=} a_1 R_+(T) + (N - a_1 - 2) R_-(T) \\ &= (2a_1 - N + 2) R_+(T). \end{aligned} \quad (3)$$

If $a_1 > N/2 - 1$, the weight of new evidence favors the choice of the first agent, and $c_1^+ > 0$. Conversely, observing

that more agents remain undecided provides evidence against the first agent's decision.

All undecided agents increment their belief by c_1^+ , causing a second wave of a_2 decisions, all of equal sign [Fig. 1(c)]. Agents in the first wave agree with the first decision, while agents in the second wave agree when the sign of c_1^+ matches the first decision. Observers undecided after this second wave update their beliefs by a new increment, c_2^+ . Waves of decisions follow until either all agents make a choice or no new agent makes a decision after some belief update, c_k^+ , $k \geq 2$ [15]. Undecided agents then continue to accumulate private information ([24], Sec. V). Whether the first decision is right or wrong, we show that in large populations the first two waves encompass the entire population.

If the first agent wrongly chooses H^- , computations are similar: observing a decision in the first wave provides a belief increment $R_-(t) = -R_+(t)$, and observing an undecided agent provides an increment $R_+(t)$, resulting in a belief increment $c_1^- = (2a_1 - N + 2)R_-(T)$. Further decision waves follow equivalently.

Decisions in large groups: As N grows, $T \rightarrow 0$, and we approximate the solution to Eq. (2) using the method of images [29,30] [24], Eq. (S3)]. Extreme value theory then gives [31–35]:

$$\mathbb{E}[T] \approx \frac{\theta^2}{4 \ln N}. \quad (4)$$

The mean time decreases logarithmically with N , allowing each agent less time to gather private information [Fig. 1(e)]. When T is small the remaining beliefs are distributed almost symmetrically around the origin. We find that

$$\mathbb{E}[a_1 | y_1(T) = \pm \theta] \approx \frac{N-1}{2} \left(1 \pm \frac{\theta}{\sqrt{4\pi \ln N}} \right). \quad (5)$$

Slightly more than half of a large clique immediately follows a correct first decision [Fig. 1(d)], and slightly less than half the clique follows a wrong first choice.

The number of agents in excess of half the population following a correct first decision scales as $N(\ln N)^{-1/2}$. But as N grows agents in the first wave accumulate less private information prior to their choice. For large N , the expected social information communicated by each decision is ([24], Sec. IX)

$$\mathbb{E}[R_+(T)] \approx 2\mathbb{E}[\sqrt{T/\pi}] \approx \theta/\sqrt{\pi \ln N}. \quad (6)$$

As N increases, a_1 grows [Fig. 1(d)], but each first wave decision provides less information [Fig. 1(e)]. However, the logarithmic decrease in $R_{\pm}(T)$ is outweighed by the nearly linear growth in a_1 : using Eqs. (3), (4), and (6), we find that the expected belief update, $\hat{c}_1^{\pm} \equiv \mathbb{E}[c_1^{\pm}]$, to undecided

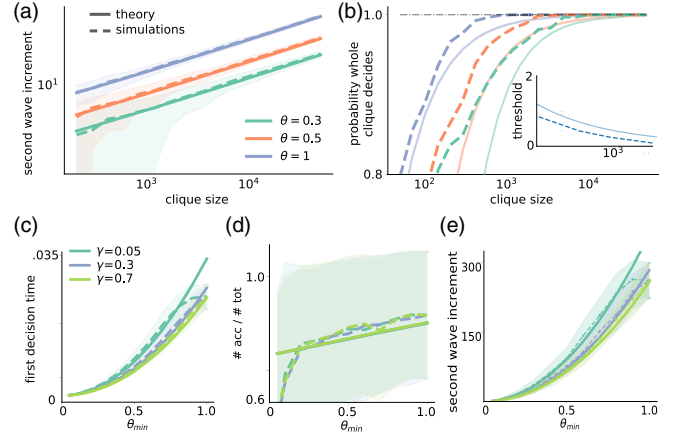


FIG. 2. Decision statistics for homogeneous and dichotomous cliques. (a) Belief increment \hat{c}_1^{\pm} . (b) Probability the full clique decides after the second wave. Chebyshev's inequality provides an upper bound on N by which the probability is reached ([24], Sec. XI). Inset: threshold θ at which $\hat{c}_1^{\pm} = 2\theta$ as N varies. (c) First decision time for dichotomous threshold cliques for various γ [Eq. (S10)]. (d) Fraction of accurate deciders in dichotomous threshold cliques under consensus bias [Eqs. (S11) and (S12)]. (e) Belief increment of agents in the second wave in dichotomous threshold cliques under consensus bias [Eq. (S12)]. Clique size $N = 15000$ in panels (c)–(e).

agents in the second wave grows nearly linearly in N ([24], Sec. IX),

$$\hat{c}_1^{\pm} \approx \frac{\theta^2 N}{2\pi \ln N}. \quad (7)$$

Here \hat{c}_1^{\pm} is positive: if the first decision is correct, then more than half the network is in the first wave, and both $(2\mathbb{E}[a_1] - N - 2)$ and $R_+(T)$ are positive in Eq. (3). Both of these terms are negative when the first decision is wrong. Thus the second wave is self-correcting when the network is sufficiently large, $\hat{c}_1^{\pm} > 2\theta$ [Fig. 2(a)]. The belief increment diverges, and its sign agrees with that of the correct choice, so that in large networks all undecided agents make the correct choice in the second wave. We use Chebyshev's inequality to show that when $N \geq 4\pi[\theta^2(1-x)]^{-1}$ the clique decides by the second wave with probability at least x [Fig 2(b) and [24], Sec. XI].

Heterogeneous populations.—A population of decision makers is rarely homogeneous. Some people decide quickly based on little evidence. Others require substantial information before choosing [36,37]. Does such diversity impact decisions of the collective?

To model such diversity we assume that decision thresholds are distributed over an interval $[\theta_{\min}, \theta_{\max}]$. Agents with a low threshold are more likely to decide first but also to make a wrong choice [38]. The ensuing exchange of social information depends on assumptions agents make about each other: while populations under consensus bias behave like homogeneous populations, omniscient agents

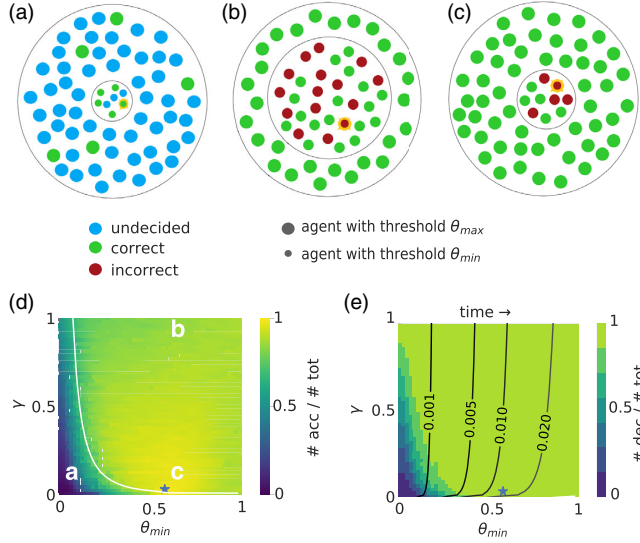


FIG. 3. Balancing hasty and deliberate decisions in dichotomous cliques. (a) With few low-threshold agents, the remaining agents receive insufficient information to decide after the first wave. (b) With many low-threshold agents, a wrong first decision sways much of the network. (c) With the right number of low-threshold agents, a few hasty agents follow an incorrect decision, but the difference between agreeing and disagreeing low-threshold agents drives the rest to choose correctly. (d) Fraction of the clique choosing accurately for a dichotomous threshold clique. White line represents Eq. (8). (e) Fraction of the clique deciding by the end of the second wave. Isoclines indicate time to first decision. $N = 15\,000$ in (b) and (c).

can leverage quick, unreliable decisions to improve the response of the population.

Dichotomous threshold distribution: The case of agents with either a high or a low threshold is tractable and sheds light on more general examples. Before a decision the belief of each agent evolves according to Eq. (1) with absorbing boundaries at $-\theta_i < 0 < \theta_i$. We assume that γN agents share threshold θ_{\min} and $(1 - \gamma)N$ share threshold θ_{\max} for $0 < \theta_{\min} < \theta_{\max}$ and $\gamma \in (0, 1)$. The first decision is then likely made by an agent with a low threshold, and is thus fast but unreliable ([24], Sec. XII). We use the approximation $\mathbb{E}[T] \approx \theta_{\min}^2 / 4 \ln(\gamma N)$ which breaks down when $0 < \gamma \ll 1$, but works well otherwise [Fig. 2(c)].

A clique under consensus bias is homogeneous from an observer's perspective and thus behaves like a homogeneous population. Indeed, the expected size of the first wave is given by an expression similar to Eq. (5), $\mathbb{E}[a_1] \approx (N - 1/2)(1 \pm (\theta_{\min}/\sqrt{4\pi \ln \gamma N}))$ ([24], Sec. XIII). The expected belief increment in the second wave is $\hat{c}_1^\pm \approx \theta_{\min}^2 \gamma N / 2\pi \ln \gamma N$ ([24], Sec. XIV) which is analogous to Eq. (7) and is governed by the timing of the first choice [Fig. 2(e)]. In large populations decisions happen quickly, before the belief distributions can interact with the boundaries. Therefore \hat{c}_1^\pm is approximately independent of the

observer's threshold: following the first wave all agents make the same update.

As in homogeneous networks, \hat{c}_1^\pm grows with N , and when $\hat{c}_1^\pm \geq 2\theta_{\max}$, we expect all agents to decide by the second wave. If the first decision is correct, the entire clique follows. A wrong first choice is followed by about half the network [Fig. 2(d)], while the second wave decides correctly. Hence under consensus bias, dichotomous cliques behave like homogenous cliques with threshold θ_{\min} : uninformed agents govern decisions, leading to fast, inaccurate choices.

In contrast, omniscient agents correctly weigh evidence revealed by a hasty first decider. We expect about half of the low-threshold agents, $\gamma N/2$, to decide in the first wave. Indeed, we find $\mathbb{E}[a_1] \approx (\gamma N - 1/2)(1 \pm (\theta_{\min}/\sqrt{4\pi \ln \gamma N}))$. The evidence revealed by a few low-threshold agents is unlikely to sway high-threshold agents [Fig. 3(a)]. However, if the subpopulation of low-threshold agents is sufficiently large, the difference between those convinced and unconvinced by the first choice triggers a correct decision in the rest of the population [Figs. 3(b) and 3(c)].

Thus, in a network of omniscient agents, hasty observers govern the speed of the first decision and comprise the first wave. The remaining agents can then observe the choices of the early adopters to make the right decision. The fraction of wrong decisions can thus be smaller than in homogeneous networks.

In finite populations this argument requires γ and θ_{\min} to be large enough for the first wave to convince the remainder of the population [Fig. 3(a)] but small enough to buffer the majority from following an incorrect first choice [Fig. 3(b)]. Hence, the population makes the best decisions at intermediate values of γ and θ_{\min} [star in Fig. 3(d)]. A balance is reached when $\hat{c}_1^- = 2\theta_{\max}$ ([24], Sec. XVII), which corresponds to a fraction of low-threshold agents given by

$$\gamma \approx \frac{4\pi\theta_{\max} \ln N}{N \theta_{\min}^2}. \quad (8)$$

Maximal accuracy is achieved when this balance holds [star, white line in Fig. 3(d)]. Almost all agents decide by the second wave [Fig. 3(e)].

Finite populations with dichotomous thresholds can sacrifice a small fraction of early adopters so the majority makes a fast, correct choice. Agents in heterogeneous networks can thus decide more quickly and outperform agents in homogeneous networks in recovering from a wrong first choice [Figs. 3(c) and 4].

Different threshold distributions: With different distributions supported on the interval $[\theta_{\min}, \theta_{\max}]$ the expected time to the first decision is again governed by θ_{\min} . Under consensus bias $\mathbb{E}[a_1] \approx N - 1/2[1 \pm (\theta_{\min}/\sqrt{4\pi \ln N})]$ ([24], Sec. XX). In either case, \hat{c}_1^\pm satisfies Eq. (7) with

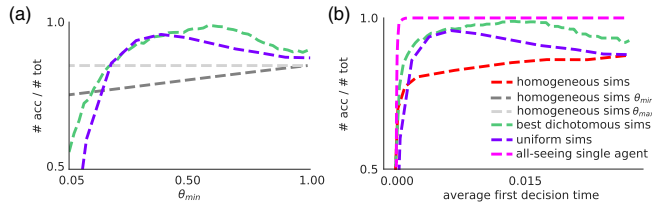


FIG. 4. Improved accuracy for heterogeneous cliques. (a) Mean fraction of clique choosing accurately after two waves for different threshold distributions in omniscient populations. For the dichotomous case, γ is chosen to maximize accuracy for each θ_{\min} value. (b) Over a range of possible first decision times, heterogeneous cliques give better accuracy than homogeneous ones with omniscient social updating ([24], Sec. XXII for simulation details). Also shown: accuracy of a single agent with access to private information of all agents. $N = 15\,000$ and $\theta_{\max} = 1$.

θ_{\min} replacing θ . For sufficiently large N , $\hat{c}_1^{\pm} > 2\theta_{\max}$. Therefore, under consensus bias the clique again behaves as a homogeneous clique with threshold θ_{\min} .

The omniscient case is more complicated, but simulations show that trends observed in the dichotomous case persist for a large class of threshold distributions. Hasty agents decide first, and deliberate agents decide based on which early adopters followed the first choice (Figs. 4 and S10), leading to faster and more accurate choices than in homogeneous networks.

Conclusion.—Our model of collective decision making is analytically tractable and shows how diverse populations can make better decisions than homogeneous ones, extending previous results [6,17–19]. Previous models often described agents forced to make decisions in sequence [27], while we assumed agents decide when faced with sufficient evidence. Decision-makers in real social networks likely combine these strategies, leading to asynchronous but clustered decisions.

Our work may describe why social organization emerges in animal groups. For example, low decision thresholds promote quick decisions based on little evidence, characteristic of “bold” individuals observed across the animal kingdom [39]. Such individuals may emerge as leaders since they often decide first. “Shy” individuals who require more evidence to make decisions may follow [40,41].

More realistic features can be included in our model: observations could be correlated [42] and agents could accumulate evidence at different rates, giving inhomogeneous drift and diffusion coefficients [22,23]. Our framework can thus be extended to understand decisions in diverse communities.

We would like to thank T. Taillefumier for helpful comments. This work was supported by NSF/NIH CRCNS Grant No. R01MH115557. B. K. and K. J. were supported by NSF Grant No. DMS-1662305. K. J. was also

supported by NSF NeuroNex Grant No. DBI-1707400. Z. P. K. was also supported by NSF Grant No. DMS-1853630. W. O. was supported by NSF Grant No. DMS-1816315 and NIH Grant No. R01GM117138.

*bkaramched@fsu.edu

†zpkilpat@colorado.edu

‡josic@math.uh.edu

- [1] I. D. Couzin, *Trends Cognit. Sci.* **13**, 36 (2009).
- [2] W. Edwards, *Psychol. Bull.* **51**, 380 (1954).
- [3] C. D. Frith and U. Frith, *Annu. Rev. Psychol.* **63**, 287 (2012).
- [4] A. Pérez-Escudero and G. G. D. Polavieja, *PLoS Comput. Biol.* **7**, e1002282 (2011).
- [5] S. Arganda, A. Pérez-Escudero, and G. G. D. Polavieja, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 20508 (2012).
- [6] R. P. Mann, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E10387 (2018).
- [7] R. P. Mann, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 10388 (2020).
- [8] A. J. W. Ward, D. J. T. Sumpter, I. D. Couzin, P. J. B. Hart, and J. Krause, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 6948 (2008).
- [9] A. Ward, J. Krause, and D. Sumpter, *PLoS One* **7**, e32411 (2012).
- [10] G. E. Gall, A. Strandburg-Peshkin, T. Clutton-Brock, and M. B. Manser, *Anim. Behav.* **132**, 91 (2017).
- [11] A. Strandburg-Peshkin, D. R. Farine, I. D. Couzin, and M. C. Crofoot, *Science* **348**, 1358 (2015).
- [12] A. Perna, B. Granovskiy, S. Garnier, S. C. Nicolis, M. Labédan, G. Theraulaz, V. Fourcassié, and D. J. T. Sumpter, *PLoS Comput. Biol.* **8**, e1002592 (2012).
- [13] R. H. Walker, A. J. King, J. W. McNutt, and N. R. Jordan, *Proc. R. Soc. B* **284**, 20170347 (2017).
- [14] J. Faria, S. Krause, and J. Krause, *Behav. Ecol.* **21**, 1236 (2010).
- [15] B. Karamched, S. Stolarczyk, Z. P. Kilpatrick, and K. Josić, *SIAM J. Appl. Dyn. Syst.* **19**, 1884 (2020).
- [16] S. B. Goyal, *Connections: An Introduction to the Economics of Networks* (Princeton University Press, Princeton, NJ, 2012), pp. 1–289.
- [17] M. Mueller-Frank, *Theor. Econ.* **8**, 1 (2013).
- [18] D. J. Watts, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5766 (2002).
- [19] R. J. Caginalp and B. Doiron, *SIAM J. Appl. Dyn. Syst.* **16**, 1543 (2017).
- [20] W. S. Geisler, *Visual neuroscience* **10**, 12 (2003).
- [21] Often the belief refers to the posterior probability of the state. The two definitions are closely related as $p(H_-|\xi) = 1/[1 + \exp(y)]$.
- [22] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, *Psychol. Rev.* **113**, 700 (2006).
- [23] A. Veliz-Cuba, Z. P. Kilpatrick, and K. Josić, *SIAM Rev.* **58**, 264 (2016).
- [24] See the Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.125.218302> for details about the simulations and calculations and Refs. [25,26].
- [25] R. A. Fisher and L. H. C. Tippet, in *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 24

- (Cambridge University Press, Cambridge, England, 1928), pp. 180–190.
- [26] A. T. Hartnett, E. Schertzer, S. A. Levin, and I. D. Couzin, *Phys. Rev. Lett.* **116**, 038701 (2016).
 - [27] A. V. Banerjee, *Q. J. Econ.* **107**, 797 (1992).
 - [28] A. Kirman, *Q. J. Econ.* **108**, 137 (1993).
 - [29] D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes* (Chapman and Hall, London, 1965).
 - [30] J. Drugowitsch, *Sci. Rep.* **6**, 20490 (2016).
 - [31] G. H. Weiss, K. E. Shuler, and K. Lindenberg, *J. Stat. Phys.* **31**, 255 (1983).
 - [32] S. B. Yuste and K. Lindenberg, *J. Stat. Phys.* **85**, 501 (1996).
 - [33] Z. Schuss, K. Basnayake, and D. Holcman, *Phys. Life Rev.* **28**, 52 (2019).
 - [34] K. Basnayake, Z. Schuss, and D. Holcman, *J. Nonlinear Sci.* **29**, 461 (2019).
 - [35] S. D. Lawley and J. B. Madrid, *J. Nonlinear Sci.* **1** (2020).
 - [36] T. Postmes, R. Spears, and S. Cihangir, *J. Pers. Soc. Psychol.* **80**, 918 (2001).
 - [37] N. R. Franks, A. Dornhaus, J. P. Fitzsimmons, and M. Stevens, *Proc. R. Soc. B* **270**, 2457 (2003).
 - [38] V. Srivastava and N. E. Leonard, *IEEE Trans. Control Network Syst.* **1**, 121 (2014).
 - [39] L. M. Aplin, D. R. Farine, R. P. Mann, and B. C. Sheldon, *Proc. R. Soc. B* **281**, 20141016 (2014).
 - [40] J. W. Jolles, N. J. Boogert, V. H. Sridhar, I. D. Couzin, and A. Manica, *Curr. Biol.* **27**, 2862 (2017).
 - [41] T. Sasaki, R. P. Mann, K. N. Warren, T. Herbert, T. Wilson, and D. Biro, *Phil. Trans. R. Soc. B* **373**, 20170038 (2018).
 - [42] R. Moreno-Bote, *Neural Comput.* **22**, 1786 (2010).