

Heterogeneous Preference and Local Nonlinearity in Consensus Decision Making

Andrew T. Hartnett,^{1,2,*} Emmanuel Schertzer,^{3,4} Simon A. Levin,² and Iain D. Couzin^{2,5,6}

¹*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*

²*Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544, USA*

³*UPMC Université Paris 06, Laboratoire de Probabilités et Modèles Aléatoires, CNRS UMR 7599, 75005 Paris, France*

⁴*Collège de France, Center for Interdisciplinary Research in Biology, CNRS UMR 7241, 75005 Paris, France*

⁵*Department of Collective Behaviour, Max Planck Institute for Ornithology, D-78457 Konstanz, Germany*

⁶*Chair of Biodiversity and Collective Behaviour, University of Konstanz, D-78457 Konstanz, Germany*

(Received 3 May 2015; published 22 January 2016)

In recent years, a large body of research has focused on unveiling the fundamental physical processes that living systems utilize to perform functions, such as coordinated action and collective decision making. Here, we demonstrate that important features of collective decision making among higher organisms are captured effectively by a novel formulation of well-characterized physical spin systems, where the spin state is equivalent to two opposing preferences, and a bias in the preferred state represents the strength of individual opinions. We reveal that individuals (spins) without a preference (unbiased or uninformed) play a central role in collective decision making, both in maximizing the ability of the system to achieve consensus (via enhancement of the propagation of spin states) and in minimizing the time taken to do so (via a process reminiscent of stochastic resonance). Which state (option) is selected collectively, however, is shown to depend strongly on the nonlinearity of local interactions. Relatively linear social response results in unbiased individuals reinforcing the majority preference, even in the face of a strongly biased numerical minority (thus promoting democratic outcomes). If interactions are highly nonlinear, however, unbiased individuals exert the opposite influence, promoting a strongly biased minority and inhibiting majority preference. These results enhance our understanding of physical computation in biological collectives and suggest new avenues to explore in the collective dynamics of spin systems.

DOI: [10.1103/PhysRevLett.116.038701](https://doi.org/10.1103/PhysRevLett.116.038701)

Collective decision making is ubiquitous across cellular [1,2], animal [3], human [4,5], and engineered systems [6]. Despite varying greatly in terms of the type and composition of components (agents), common dynamical features allow collectives to make rapid, accurate decisions even in complex environments or amidst conflicting needs [7]. The dynamics of populations of neurons, for example, have been precisely related to binary “spins” (spikes) communicating via pairwise interactions [8,9], a characteristic of models employed widely in statistical physics. Neural dynamics also exhibit striking parallels with collective decision making among organisms themselves, such as the process by which honeybee colonies decide among alternative nest sites [10,11]. Thus, by abstracting microscopic details and relating biological computation directly to that among physical particles, we can obtain considerable insight regarding collective behavior and possibly reveal novel aspects of physical systems.

Here, we explore the dynamics of populations making collective decisions. Despite being inspired by recent

experimental data from animal groups, we deliberately shift the focus from understanding collective decisions in a particular experimental context to a general understanding of the underlying dynamical interplay among population constitution, space, and the character of local interactions. In doing so we demonstrate that many of the complex, and possibly counterintuitive, features seen when considering collective decision making by organisms can be represented via consideration of a continuum, or “family,” of spin system models.

A main focus of our work is to consider observations in animal groups that suggest that unbiased (or uninformed) individuals can strongly influence the outcome of collective consensus decisions [11]. Previous phenomenological models, while able to capture some elements in common with experiments [11], provide relatively limited insight into the mechanism of action. Here, we exploit the well-characterized nature of physical spin systems to deepen our theoretical understanding of collective decision making, employing a two-choice decision task for a group with three distinct subpopulations: two subpopulations with conflicting preferences (informed or preference individuals), and a variable proportion of unbiased individuals (who participate in the decision making process, but have no information and/or no preferred outcome).

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

In physical terms one could think of each informed subpopulation as being sensitive to a different external field. Each field exerts an influence on the corresponding subpopulation, increasing its tendency to remain in, or return to, its preferred state. Thus, beyond biological applications, our work suggests new avenues to explore in interacting particle systems, and could have applications in a wide range of systems from magnetic materials to neural and social dynamics.

Description of the model.—We consider a family of nonlinear (discrete time) voter models on the 2D square lattice [12–16] in which individuals may, or may not, exhibit a bias regarding their preferred state; our social collective is a spin system in which each lattice point is characterized by an opinion state $\sigma_z \in \{-1, +1\}$ and a fixed bias ω_z . Opinion states change in time and are shared with neighbors during interactions. For simplicity, we assume that $\omega_z \in \{\omega_0, \omega_+, \omega_-\}$, corresponding to unbiased ($\omega_0 = 1$), biased to $+1$ ($\omega_+ > 1$), and biased to -1 [$\omega_- \in (0, 1)$] subpopulations, respectively. We assume that individuals from each biased subpopulation are randomly and uniformly distributed in the landscape according to the density $[\rho_0, (1 - \rho_0)\rho_+, (1 - \rho_0)\rho_-]$.

At each time step, an individual samples the opinions of its four neighbors, mimicking local influence or social pressure, and computes a social field that is distorted by its own personal bias ω_z :

$$h_{\omega_z} = \frac{\omega_z n_z^+ - n_z^-}{\omega_z n_z^+ + n_z^-} \in [-1, 1] \quad (1)$$

in which n_z^+ (n_z^-) is the number of neighbors with $+1$ (-1) opinions.

Bias is implemented as an amplification of the influence of neighbors who agree with an individual's internal preference. This amounts to an increased willingness to switch into, or an intransigent tendency to remain in, a preferred state given at least one supporting neighbor.

Given its social field h_{ω_z} , an individual then updates its opinion state probabilistically. Individuals in the -1 state at time t transition to the $+1$ state at time $t + 1$ with a probability $G^\beta(h_{\omega_z})$ and individuals transition from the $+1$ to the -1 state with a probability $1 - G^\beta(h_{\omega_z})$, where $G^\beta(h)$ is the following sigmoid function:

$$G^\beta(h) = \frac{1}{2} \left(1 + \frac{\tanh(\beta h)}{\tanh(\beta)} \right) \in [0, 1] \quad (2)$$

and where $\beta \in (0, \infty)$ serves to parametrize the nonlinearity of the model. When $\omega_z = 1 \forall z$ (no bias), and in the limit as $\beta \rightarrow 0$, our model reduces to the classical voter model [17,18]. By contrast, as $\beta \rightarrow \infty$, $G^\beta(h_{\omega_z})$ becomes a step function and we recover zero temperature (discrete time) Ising dynamics [15,19] (also known as the majority-rule model). Interpolating between these two extremes allows us to investigate models that exhibit a spectrum of coarsening behaviors [20].

Consensus and time to equilibrium.—The impact of unbiased individuals on the equilibrium of the system is readily apparent [Fig. 1(a)]. In the absence of a sufficient density of such individuals, the system becomes frozen in a frustrated state with pockets of each opinion [22]. A set of similarly biased individuals within a small region of space is self-reinforcing, and at low densities of unbiased individuals these self-reinforcing pockets are common, preventing the flow of information across the system. By contrast, clusters of unbiased individuals can be thought of as pooling information or allowing opinions to propagate. Thus, increasing the density of such individuals contributes to achieving consensus in two key ways—spatially diluting biased individuals, and pooling opinion states.

For each set of system parameters, we find there is a specific density of unbiased individuals (ρ_0^*) at which self-reinforcing clusters disappear and the system will reach

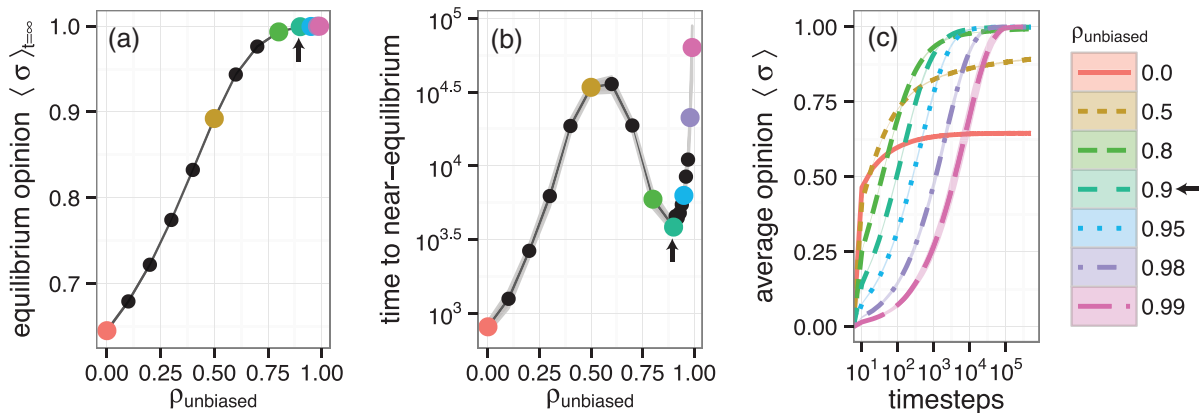


FIG. 1. The equilibrium magnetization is shown for a nonlinear system ($\beta = 0.5$) as a function of the density of unbiased individuals (a). A magnetization of $+1$ (-1) corresponds to majority (minority) consensus. In all three panels, we started with a random uniform distribution of opinion. The majority and minority have equivalent biases ($\omega_+ = 1/\omega_- = 2.96$), ensuring positive mean opinions ($\rho_+ = 1.5\rho_-$). In (b), the time steps taken to reach 98% of the equilibrium value $\langle \sigma \rangle$ are shown as a function of the density of unbiased individuals. Panel (c) shows that the existence of some biased individuals accelerates domain formation on a fast time scale; however, too many lead to multiple stable domains.

consensus ($\langle\sigma\rangle = \pm 1$). For a wide range of parameters, this specific density also corresponds to both the global minimum for the time to consensus and a local minimum in the time to equilibrium (where equilibrium may be a frustrated state) [Fig. 1(b)]. In particular, the time to equilibrium is a nonmonotonic function of the density ρ_0 , and, in the vicinity of ρ_0^* , adding unbiased individuals to the system can both increase magnetization and reduce the time to equilibrium.

In a broad sense, this behavior is analogous to (aperiodic) stochastic resonance [23–25], in which adding some noise, here the unbiased individuals, to the system can actually amplify the signal, here the magnetization. Unbiased individuals can be interpreted as noise, in the sense that they do not carry any intrinsic preference, and that in the absence of biased individuals, the population would drift to a $+1$ or -1 magnetization with equal probability. Furthermore, our system exhibits some similarities with random nonlinear dynamical systems, where noise allows the system to escape the local minima of a potential—here the self-reinforcing clusters of biased individuals.

Lastly, while in the absence of any biased individuals both absorbing states are equally probable, very little “signal” is required to drive the system to a given absorbing state. A small number of biased individuals (< 10) is enough to significantly increase the probability of reaching the preferred absorbing state (See Supplemental Material at [26] for Fig. S2).

Effect of nonlinearities.—We now turn our attention to the impact of the unbiased individuals on the equilibrium of the system for different values of the nonlinearity parameter β . As in Ref. [11], we consider a majority with a weak preference for the $+1$ state and a minority with a strong preference for the -1 state (i.e., $\rho_- < \rho_+$ but $1/\omega_- > \omega_+$), and a variable number of unbiased individuals.

Before examining the complex behavior that arises through the interaction of space and nonlinearity, let us first consider the role of nonlinearity alone in a mean-field model. This space-free model can be thought of as a null model that serves as both a convenient way to first define the notion of intrinsic preference, and second to later highlight the role of space. In this setting, the fraction v of individuals in the $+1$ state is described by the dynamics

$$v'(t) = \langle b^\beta(v(t)) \rangle - v(t), \quad (3)$$

where $\langle b^\beta(v) \rangle$ can be interpreted as the preference for the $+1$ state in a population with magnetization $\langle\sigma\rangle = 2v - 1$:

$$\langle b^\beta(v) \rangle = \rho_0 b_0^\beta(v) + (1 - \rho_0)(\rho_+ b_+^\beta(v) + \rho_- b_-^\beta(v)),$$

where $b_i^\beta(v)$ is the probability for an individual of type i to switch from -1 to $+1$ in a population where the magnetization of each individual is chosen independently according to the distribution $\mathbb{P}(1) = 1 - \mathbb{P}(-1) = v$. When social influence in the population does not favor any opinion

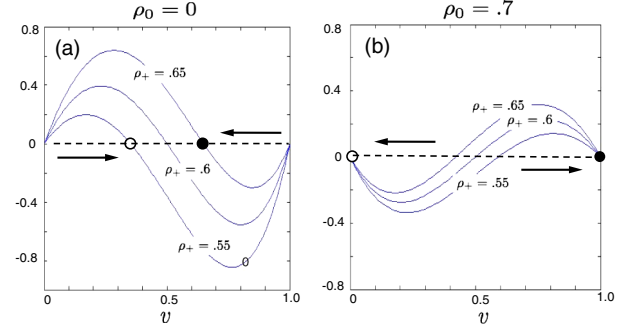


FIG. 2. The function $\langle b^\beta(v) \rangle - v$ for different values of ρ_+ and ρ_0 . [$\beta = 2$, $b_+^\beta(1/2) = 0.79$, $b_-^\beta(1/2) = 0.07$]. For both $\rho_0 = 0, 0.7$, the consensus states are stationary points for Eq. (3) and there exists a unique stationary point ρ^* on the interval $(0,1)$. In the absence of unbiased individuals (a), the consensus states are unstable, and the dynamics converge to the frustrated state ρ^* . While, for a relatively high density of unbiased individuals (b), the internal stationary point becomes unstable and the dynamics converge to one of the two consensus states. Points represent the magnetization at equilibrium for $\rho_+ = 0.55$ (white) and $\rho_+ = 0.65$ (black). Unbiased individuals do not change the sign of the magnetization at equilibrium, but they do amplify it.

($\langle\sigma\rangle = 0$ or $v = 1/2$), the coefficient $\langle b^\beta(1/2) \rangle$ can be interpreted as a measure of intrinsic preference.

If we start the dynamics at $v(0) = 1/2$, the sign of the magnetization at equilibrium is determined by the population’s intrinsic preference. In turn, straightforward analytics show that the fraction of unbiased individuals does not affect the sign of this intrinsic preference, but only affects its magnitude [Figs. 2(a) and 2(b)].

In the spatial model, the situation is quite different. Beyond altering the coarsening dynamics [27,28], the social nonlinearity can qualitatively change the equilibrium magnetization of the system.

First, in contrast with the mean-field model, the boundary between the $+1$ or -1 dominated regions of the parameter space depends on the concentration of unbiased individuals. As ρ_0 increases, the region of parameter space where the minority dominates at equilibrium expands; the expansion is more prominent for higher values of β (Fig. 3). This is consistent with the finding that unbiased individuals can promote a specific subpopulation and, in doing so, reverse the outcome of a collective decision [11].

Secondly, in the region of parameter space where the numerical advantage of the majority is roughly balanced by the increased bias of the minority (i.e., when the average intrinsic preference of the population is close to zero) the system deviates substantially from its mean-field behavior. Without unbiased individuals, the system equilibrium is close to zero magnetization [Fig. 4 (inset)]. As the proportion of unbiased individuals increases, the magnetization of the system is determined by the value of the nonlinearity parameter β [Fig. 4 (main figure)].

If social influence during interactions tends to be relatively linear, unbiased individuals reinforce the majority preference in the face of a more strongly biased numerical

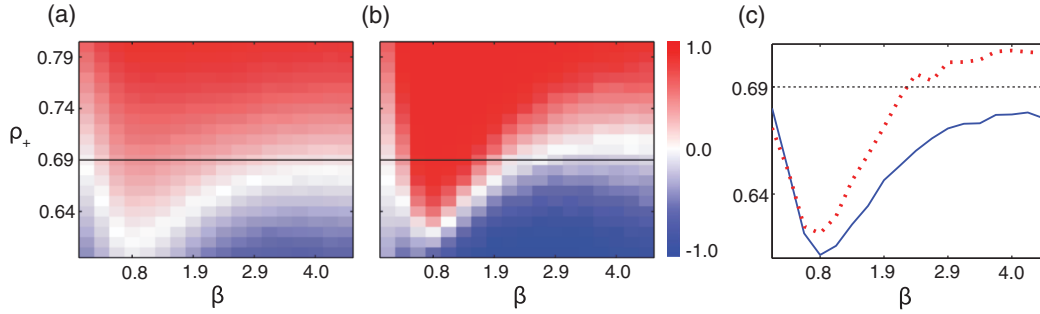


FIG. 3. Intrinsic preferences are fixed across experiments [$b_+^\beta(1/2) = 0.69$ and $b_-^\beta(1/2) = 0.08$] so that the sign of the intrinsic preference is controlled by the relative proportion of weakly opinionated individuals ρ_+ . For the mean-field model, a horizontal line ($\rho_+ = 0.69$) separates the $+1$ and -1 dominated regions, independent of the density of unbiased individuals [panel (a), $\rho_0 = 0$; panel (b), $\rho_0 = 0.8$]. Panel (c) shows the boundary separating the two regions for the spatial model for $\rho_0 = 0$ (solid blue curve) and $\rho_0 = 0.8$ (dotted red curve). The boundary is shifted upward for higher values of ρ_0 .

minority, thus promoting democratic outcomes. However, as social influence becomes increasingly nonlinear, the role of unbiased individuals is reversed; they now tend to promote a strongly biased minority, thus inhibiting the role of numerical advantage and amplifying the role of preference strength. If local interactions can be modulated by internal or external factors, this result suggests that group decisions may be readily adaptable [29,30].

Relaxing spatial constraints.—That the mean-field model exhibits qualitatively different behavior than the spatial model suggests that spatial constraints are important to both the dynamics and the final state of the system. To understand the role of such constraints, we consider an intermediate model between our original fixed-position model and a pure mean-field approach.

We employ a model of the kind studied by Durrett and Neuhauser with fast stirring [31]. Neighboring individuals can now exchange positions, and do so much more frequently than they update their opinions (update with probability $\epsilon \ll 1$). We make a further simplifying assumption that the roles of biased and unbiased

individuals are randomly redistributed in the landscape each time step [32].

In the limit as $\epsilon \rightarrow 0$, the local density of $+1$ opinions is well described by the solution to the following partial differential equation [31]:

$$\partial_t u(x, t) = \frac{1}{2} \Delta u(x, t) + \left[\langle b^\beta(u(x, t)) \rangle - u(x, t) \right], \quad (4)$$

which is obtained from the mean-field equation (3) by adding a Laplacian term.

Standard analysis of reaction-diffusion equations [33] states that when Eq. (4) has a unique stationary point on $(0,1)$, the asymptotic behavior of the equation is determined by the stability of this point. When stable, the equilibrium of the fast-stirring dynamics coincides with the mean-field equilibrium; if unstable, the equilibrium of the system is determined by $\int_0^1 \langle b^\beta(v) \rangle dv$, which can be interpreted as the mean preference of the population. When it is greater (less) than $1/2$, the system reaches a consensus at $+1$ (-1).

In Fig. 2, we started from a frustrated population in which the mean-field equilibrium is achieved at some internal stable stationary point $\rho^* \in (0, 1)$ in the absence of unbiased individuals [Fig. 2(a)]. From the previous analysis, $u(x, t) \rightarrow \rho^*$ as $t \uparrow \infty$, and the sign of the magnetization is determined by the population's intrinsic preference. By increasing the density of unbiased individuals ρ_0 , the internal stationary point becomes unstable [Fig. 2(b)] and the system reaches a consensus determined by the population's mean preference.

In summary, in the region of the parameter space where Eq. (3) admits a unique stationary point on $(0,1)$, the region of dominance for the minority opinion is either determined by the intrinsic preference of the population (low density of unbiased individuals) or by its mean preference (high density of unbiased individuals). As in the original spatial model, we then observe that the minority dominated region expands as the density of unbiased individuals increases, and that this effect is more important for higher values of β [34].

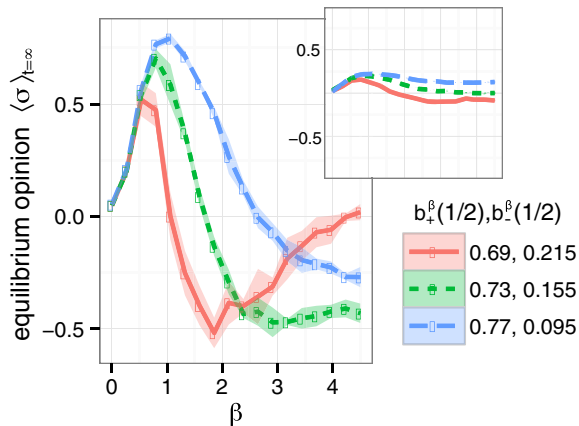


FIG. 4. The average magnetization at equilibrium is shown as a function of the local nonlinearity for different biases of both the majority and minority. In the main figure, the fraction of unbiased individuals is fixed at 80%. In the inset, the fraction of unbiased individuals is equal to 0%.

Future directions.—In this work we demonstrate that important features of collective decision making can be captured and understood in the context of spin systems. There are both natural and biologically important extensions to be pursued in future work. For example, decision making systems (e.g., neural or animal groups) in nature can be offered multiple options. Numerical simulations, and potentially limited analyses, could be conducted for a dynamical model interpolating between the multistate voter model [35] and a generalization of the Potts model [36].

Additionally, we hope to focus on more complex and naturalistic topologies. Simulating consensus dynamics on graphs constructed from real sensory networks in fish schools [37,38] may provide both better estimates of the actual density of unbiased individuals in such schools and a test bed for understanding how and when fluctuations due to finite system size play a meaningful role in decision making.

Lastly, our work suggests that for a given topology and set of biases there is an optimal density of unbiased individuals. Further analysis, particularly with respect to the different time scales of coarsening, could provide insights into a mechanism whereby agents could self-organize toward such an optimal density.

The authors would like to thank Colin Torney for several helpful conversations. This work was supported by National Science Foundation Grants No. PHY-0848755 and No. IOS-1355061, EAGER Grant No. IOS-1251585, Office of Naval Research Grants No. N00014-09-1-1074 and No. N00014-14-1-0635, Army Research Office Grants No. W911NF-11-1-0385 and No. W911NF-14-1-0431, and Human Frontiers Science Program Grant No. RGP0065/2012.

* a.t.hartnett@gmail.com

- [1] E. B. Jacob, I. Becker, Y. Shapira, and H. Levine, *Trends Microbiol.* **12**, 366 (2004).
- [2] T. S. Deisboeck and I. D. Couzin, *BioEssays* **31**, 190 (2009).
- [3] J. Krause and G. D. Ruxton, *Living in Groups* (Oxford University Press, New York, 2002).
- [4] F. Galton, *Nature (London)* **75**, 450 (1907).
- [5] J. Surowiecki, *The Wisdom of Crowds: Why the Many are Smarter Than the Few and how Collective Wisdom Shapes Business, Economies, Societies, and Nations* (Doubleday, Garden City, NY, 2004).
- [6] W. Ren and R. Beard, *IEEE Trans. Autom. Control* **50**, 655 (2005).
- [7] J. Krause, G. Ruxton, and S. Krause, *Trends in Ecology & Evolution* **25**, 28 (2010).
- [8] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek, *Nature (London)* **440**, 1007 (2006).
- [9] G. Tkačik, O. Marre, D. Amodei, E. Schneidman, W. Bialek, and M. J. Berry II, *PLoS Comput. Biol.* **10**, e1003408 (2014).
- [10] T. D. Seeley, and P. K. Visscher, *Apidologie* **35**, 101 (2004).
- [11] I. D. Couzin, C. C. Ioannou, G. Demirel, T. Gross, C. J. Torney, A. Hartnett, L. Conradt, S. A. Levin, and N. E. Leonard, *Science* **334**, 1578 (2011).
- [12] C. Castellano, S. Fortunato, and V. Loreto, *Rev. Mod. Phys.* **81**, 591 (2009).
- [13] C. Castellano, M. A. Muñoz, and R. Pastor-Satorras, *Phys. Rev. E* **80**, 041129 (2009).
- [14] M. Mobilia and S. Redner, *Phys. Rev. E* **68**, 046106 (2003).
- [15] P. L. Krapivsky and S. Redner, *Phys. Rev. Lett.* **90**, 238701 (2003).
- [16] J. Molofsky, R. Durrett, J. Dushoff, D. Griffeth, and S. Levin, *Theor. Popul. Biol.* **55**, 270 (1999).
- [17] T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Vol. 324 (Springer, New York, 1999).
- [18] J. Drouffe and C. Godreche, *J. Phys. A* **32**, 249 (1999).
- [19] S. Galam, *Eur. Phys. J. B* **25**, 403 (2002).
- [20] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.116.038701> and Ref. [21] for simulation details.
- [21] M. Weigel, *J. Comput. Phys.* **231**, 3064 (2012).
- [22] Standard arguments in Markov chain theory tend to show that the only two possible equilibria of the system are the two consensus states, and that other observed (frustrated) equilibria are only pseudoequilibria of the system. Additional experiments were conducted to confirm that these frozen states are indistinguishable from the “true” equilibria. See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.116.038701> for Fig. S1.
- [23] J. J. Collins, C. C. Chow, and T. T. Imhoff, *Phys. Rev. E* **52**, R3321 (1995).
- [24] K. Wiesenfeld and F. Moss, *Nature (London)* **373**, 33 (1995).
- [25] B. McNamara and K. Wiesenfeld, *Phys. Rev. A* **39**, 4854 (1989).
- [26] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.116.038701> for Fig. S2.
- [27] D. R. Wozny, U. R. Beierholm, and L. Shams, *PLoS Comput. Biol.* **6**, e1000871 (2010).
- [28] S. Arganda, A. Pérez-Escudero, and G. G. de Polavieja, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 20508 (2012).
- [29] C. Torney, Z. Neufeld, and I. Couzin, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22055 (2009).
- [30] A. Berdahl, C. Torney, C. Ioannou, J. Faria, and I. Couzin, *Science* **339**, 574 (2013).
- [31] R. Durrett and C. Neuhauser, *Ann. Probab.* **22**, 289 (1994).
- [32] This amounts to considering a population of identical agents either pooling social information (with probability ρ_0) or acting on an aggregate of personal (biased towards ± 1) and social information. This is analogous to the situation in Ref. [29] in which individuals change their social behavior based on local context.
- [33] P. C. Fife and J. B. McLeod, *Arch. Ration. Mech. Anal.* **65**, 335 (1977).
- [34] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.116.038701> for Fig. S3 showing the boundary between the -1 and $+1$ dominated regions across values of the nonlinearity parameter β .
- [35] M. Starnini, A. Baronchelli, and R. Pastor-Satorras, *J. Stat. Mech.* (2012) P10027.
- [36] F. Y. Wu, *Rev. Mod. Phys.* **54**, 235 (1982).
- [37] A. Strandburg-Peshkin, C. R. Twomey, N. W. Bode, A. B. Kao, Y. Katz, C. C. Ioannou, S. B. Rosenthal, C. J. Torney, H. S. Wu, S. A. Levin, and I. D. Couzin, *Curr. Biol.* **23**, R709 (2013).
- [38] S. B. Rosenthal, C. R. Twomey, A. T. Hartnett, H. S. Wu, and I. D. Couzin, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4690 (2015).