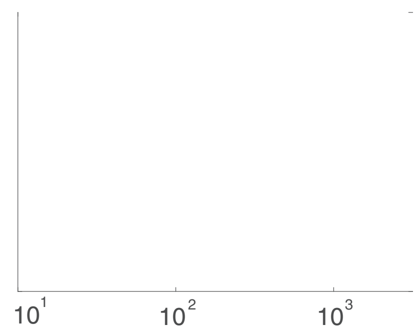
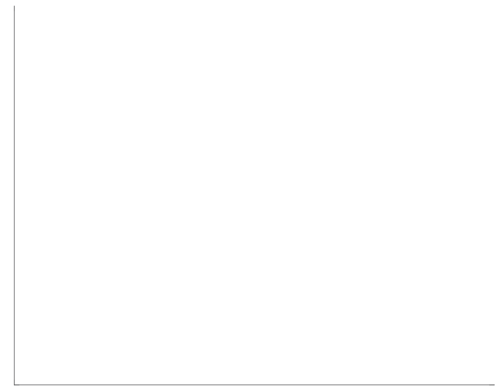
F t t ; w t

Samantha Linn ,[1,*] Sean D. Lawlesssl(e57c5ibn/2 1.72. w51e)15.8124.917 681.573 Tm0 0 1 rg(1)Tj0 g.5704 0 T†73 Tm0 0 8*g0 TD1.955


10^1
10^2
10^3



how our results would be affected by such correlations, although we expect that under certain conditions they remain qualitatively true.

Our analysis of the fastest and slowest decisions joins several recent works which highlight the importance of extreme statistics in diverse biophysical systems. For example, the earliest receptor bindings may enable a single cell to locate a source [37,38] much more accurately than later receptor bindings [39]. The fastest receptor activations may also contribute to the effectiveness of kinetic proofreading for antigen discrimination by T cells [40], while the slowest primordial follicle growth activations determine menopause timing [41] and their extreme statistics shed light on the apparent "wasteful" follicle oversupply [42].

Ramping activity of individual neurons during decision making has been observed across the brain [43,44] (although see Ref. [45]). Such dynamics may reflect the underlying evidence accumulation process preceding a decision and is often modeled by a drift-diffusion process. Decisions are thought to be triggered by the elevated activity of sufficiently many choice-related neurons [46]. These results combined with our previous work on the impact of correlations [11] suggest that early decisions tend to exhibit lower accuracy. However, a key feature of neural circuits is their recurrent connectivity, which could help neural circuits reduce or even prevent the negative effects of extreme events [47].

Our theory also applies more generally to independently evolving drift-diffusion processes on possibly unbounded domains [48]: In large populations early threshold crossings reflect only the initial states, agnostic to other system attributes, while late crossings are independent of initial states and reflect the quasistationary distribution. Hence, early crossings reflect initial biases providing fast reactions needed for time-sensitive biophysical processes [49]. If time allows, then quorum sensing processes that weight passages by order could be used [50], managing population level trade-offs between speed and accuracy.

ACKNOWLEDGMENTS

Z.P.K. and K.J. were supported by NSF DMS-2207700. Z.P.K. was supported by NIH BRAIN 1R01EB029847-01. S.D.L. was supported by NSF DMS-2325258 and NSF DMS-1944574. K.J. was supported by NSF-DBI-1707400 and NIH RF1MH130416. S.L. was supported by NSF Grant No. 2139322.

APPENDIX A: MATHEMATICAL PRELIMINARIES

Suppose $\{(\tau_n, Z_n)\}_{n\geqslant 1}$ is an independent and identically distributed (iid) sequence of realizations of the pair of (possibly correlated) random variables (τ, Z). We have in mind that τ is the decision time [or first passage time (FPT)] of some decider whose stochastic evolution of beliefs is denoted by $\{X(t)\}_{t\geqslant 0}$ and Z is a vector containing information about this decider, such as their random initial position, drift, diffusivity, and decision made. Define the cumulative distribution function (CDF) of τ,

$$F(t) := P(\tau \leqslant t).$$

Further, for any event E that is in the σ algebra generated by Z, define

$$F_E(t) := P(\tau \leqslant t \cap E).$$

In words, E is any event for which we can know whether or not it occurred by knowing Z. For example, we are interested in events E like $E = \{X(0) = \theta/2\}$, $E = \{X(0) \leqslant 0\}$, $E = \{X(\tau) = \theta\}$, etc.

For a given $N \geqslant 1$, let $n(j) \in \{1, \dots, N\}$ denote the (random) index of the jth fastest decider out of the first N deciders to make a decision. That is, suppose we order the first N FPTs (or first decision times),

$$T_{1,N} \leqslant T_{2,N} \leqslant \cdots \leqslant T_{N-1,N} \leqslant T_{N,N},$$

where $T_{j,N}$ denotes the jth fastest FPT,

$$T_{j,N} := \min\left(\{\tau_1, \dots, \tau_N\} \setminus \bigcup_{i=1}^{j-1} \{T_{i,N}\}\right), \quad j \in \{1, \dots, N\}. \tag{A1}$$

Then $n(j)$ is such that

$$\tau_{n(j)} = T_{j,N}. \tag{A2}$$

In the examples of interest, the FPTs, τ, have continuous probability distributions (i.e., $F(t)$ is a continuous function) so that the event $\tau_{n'} = \tau_{n''} < \infty$ for $n' \neq n''$ has probability zero so there is no ambiguity in Eq. (A2).

Since we have the sequence $\{(\tau_n, Z_n)\}_{n\geqslant 1}$, we denote as E_n the event E as it pertains to the nth element in the sequence $\{(\tau_n, Z_n)\}_{n\geqslant 1}$. For example, if $E = \{X(0) = \theta/2\}$, then $E_n = \{X_n(0) = \theta/2\}$. Similarly, $E_{n(j)}$ is the event E as it pertains to $Z_{n(j)}$.

Throughout the Appendix, we use the notation $\int f(t)\, dg(t)$ to denote the Riemann-Stieltjes integral of a function f with respect to a function g.

Proposition 1. For any $j \in \{1, 2, \dots, N\}$ (denoting an agent by the order j of their decision), we have that

$$P(E_{n(j)}) = j\binom{N}{j}\int_0^\infty [F(t)]^{j-1}[1 - F(t)]^{N-j}\, dF_E(t). \tag{A3}$$

In the case $j = 1$ (i.e., the fastest decider), Proposition 1 implies

$$P(E_{n(1)}) = N\int_0^\infty [1 - F(t)]^{N-1}\, dF_E(t). \tag{A4}$$

Since $1 - F$ is a decreasing function, Eq. (A4) implies that the short-time behavior of F and F_E determine the large N behavior of $P(E_{n(1)})$. More generally, Proposition 1 implies that the short-time behavior of F and F_E determine the large N behavior of $P(E_{n(j)})$ for $1 \leqslant j \ll N$.

In the case $j = N$ (i.e., the slowest decider), Proposition 1 implies

$$P(E_{n(N)}) = N\int_0^\infty [F(t)]^{N-1}\, dF_E(t). \tag{A5}$$

Since F is an increasing function, Eq. (A5) implies that the large-time behavior of F and F_E determine the large N behavior of $P(E_{n(N)})$. More generally, Proposition 1 implies that the large-time behavior of F and F_E determine the large N behavior of $P(E_{n(N-j)})$ for $1 \ll N - j$.

APPENDIX B: SOME INTEGRAL ASYMPTOTICS

The following proposition is useful for estimating the large-N behavior of some integrals of the form in Eq. (A3) and was proved in Ref. [48] (see Proposition 2 in Ref. [48]). Throughout the Appendix, "f □ g" denotes $f/g \to 1$ (e.g., as $N \to \infty$ or as $t \to 0$).

 PHYSICAL REVIEW E **110**, 024305 (2024)

and

$$\eta_i(1)$$

With these estimates, we can apply Theorem 1 to obtain estimates that the fastest decider(s) have extreme initial beliefs. In particular, suppose we want to estimate

$$P(a+\varepsilon < X_{n(1)}(0) < b-\varepsilon) \quad \text{for some small } 0 < \varepsilon \ll 1,$$

which is the probability that the fastest decider does not have extreme initial beliefs. If we define the event

$$E = \{a+\varepsilon < X(0) < b-\varepsilon\},$$

then using the notation of Appendix A, we have that

$$F_E(t) := P(\tau \leqslant t \cap E) = \int_{a+\varepsilon}^{b-\varepsilon} P(\tau \leqslant t \,|\, X(0)=x)\nu(x)\,dx \sim t^p \int_{a+\varepsilon}^{b-\varepsilon} A(x)\nu(x)e^{-C(x)/t}\,dx \quad \text{as } t\to 0^+,$$

which can be estimated as above using Laplace's method [53]. In particular, if $b > |a|$, then

$$F_E(t) \sim A(b-\varepsilon)\nu(b-\varepsilon)t^{p+1}e^{-C(b-\varepsilon)/t} \quad \text{as } t\to 0^+,$$

assuming $\nu(b-\varepsilon) > 0$, and similarly if $|a| > b$ or $|a| = b$. With this short-time behavior of $F_E(t)$, we can then plug this into Theorem 1 to show that the first deciders have the most extreme initial beliefs.

APPENDIX F: HETEROGENEOUS POPULATION WITH MULTIPLE ALTERNATIVES

We next consider the generalized case where the beliefs of the agents in the population evolve as processes with (possibly space-dependent) drift, diffusion coefficient, initial position, and even domain (in their own arbitrary space dimension $d \geqslant 1$). Suppose the belief of the ith decider evolves according to the following d-dimensional SDE,

$$dX_i = \mu_i(X_i)\,dt + \sqrt{2D_i}\,dW_i, \tag{F1}$$

where $\mu_i : \mathbb{R}^d \to \mathbb{R}^d$ is a possibly space-dependent drift, $D_i > 0$ is the diffusion coefficient, and $W(t) \in \mathbb{R}^d$ is a standard Brownian motion in d-dimensional space.

Let $L > 0$ denote an agent's (random) shortest distance they must travel to hit the closest target and let $D > 0$ denote the agent's diffusion coefficient. Define the random timescale

$$S = \frac{L^2}{4D} > 0.$$

Suppose that S has a discrete distribution on a finite set

$$0 < s_0 < s_1 < s_2 < s_3 \cdots < s_I,$$

where

$$P(S = s_i) = q_i > 0, \quad \sum_{i=0}^{I} q_i = 1.$$

Since we have $N \geqslant 1$ iid agents indexed from $n = 1$ to $n = N$, we let S_n denote the value of S for the nth agent and $S_{n(j)}$ the value of S for the jth fastest to decide.

We have that [52]

$$\lim_{t\to 0^+} t \ln P(\tau \leqslant t) = -s_0 < 0, \quad \lim_{t\to 0^+} t \ln P(\tau \leqslant t \cap S = s_i) = -s_i < 0.$$

Hence, Proposition 1 and Theorem 2 imply that for any fixed $j \geqslant 1$ and $i \in \{1, \ldots, I\}$ and any $\varepsilon > 0$,

$$N^{1-s_i/s_0-\varepsilon} \ll P(S_{n(j)} = s_i) \ll N^{1-s_i/s_0^{\varepsilon}} \quad \text{as } N\to\infty, \tag{F2}$$

where we use the notation $f \ll g$ to mean $\lim f/g = 0$. That is, in more traditional notation,

$$N^{1-s_i/s_0-\varepsilon} = o(P(S_{n(j)} = s_i)) \quad \text{as } N\to\infty,$$

$$P(S_{n(j)} = s_i) = o(N^{1-s_i/s_0+\varepsilon}) \quad \text{as } N\to\infty.$$

In the special case that the agents all move in one space dimension and the drifts are spatially constant (but may differ between agents), we can get the constant and logarithmic prefactors on the decay of $P(S_{n(j)} = s_i)$ as $N\to\infty$.

The result in Eq. (F2) says that in a large population if all the agents have the same diffusion coefficient, then the fastest deciders started closest to their decision thresholds (targets). If we allow the diffusion coefficients to vary between agents, then (F2) implies that the fastest deciders started close to their decision thresholds and/or they had big diffusion coefficients.

APPENDIX G: SLOWEST DECIDERS

Suppose the beliefs of the iid agents diffuse in some d-dimensional spatial domain $U \subset \mathbb{R}^d$ and can be absorbed at one of $m \geqslant 2$ targets $V_0, \dots, V_{m-1}$ and let $\kappa \in \{0, \dots, m-1\}$ indicate which target the decider eventually hits. Here, we will think of the m targets as parts of the $d-1$ dimensional boundary of the domain, and assume that hitting one of the targets triggers a decision. Following Refs. [54,55], suppose the beliefs of the deciders evolve as stochastic process $\{X(t)\}_{t\geqslant 0}$ that diffuse according to the SDE

$$dX(t) = -\nabla V[X(t)]\,dt + \sqrt{2D}\,dW(t), \tag{G1}$$

with reflecting boundary conditions. In Eq. (G1), the drift term is the gradient of a given potential, $V(x)$, and the noise term depends on the diffusion coefficient $D > 0$ and a standard d-dimensional Brownian motion (Wiener process) $\{W(t)\}_{t\geqslant 0}$. The survival probability conditioned on the initial position,

$$\mathbf{S}(x,t) := P(\tau > t \,|\, X(0) = x),$$

satisfies the backward Kolmogorov (also called backward Fokker-Planck) equation,

$$\begin{aligned} \frac{\partial}{\partial t}\mathbf{S} &= \mathcal{L}\mathbf{S}, \quad x \in U,\\ \mathbf{S} &= 0, \quad x \in \text{targets},\\ \frac{\partial}{\partial \mathbf{n}}\mathbf{S} &= 0, \quad x \in \text{reflecting boundary (if there is one)},\\ S &= 1, \quad t = 0. \end{aligned} \tag{G2}$$

In Eq. (G2), the differential operator $\mathcal{L}$ is the generator (i.e., the backward operator) of Eq. (G1),

$$\mathcal{L} = -\nabla V(x)\cdot\nabla + D\Delta,$$

and $\frac{\partial}{\partial \mathbf{n}}$ is the derivative with respect to the inward unit normal $\mathbf{n} : \partial U \to \mathbb{R}^d$.

Using the following weight function of Boltzmann form,

$$\rho(x) := \frac{e^{-V(x)/D}}{\int_U e^{-V(y)/D}\,dy}, \tag{G3}$$

one can check that the differential operator $\mathcal{L}$ is formally self-adjoint on the weighted space of square integrable functions (see, for example, Ref. [55]),

$$L^2_\rho(U) := \left\{ f : \int_U |f(x)|^2 \rho(x)\,dx < \infty \right\},$$

using the boundary conditions in (G2) and the following weighted inner product,

$$(f,g)_\rho := (f, g\rho) = \int_U f(x)g(x)\rho(x)\,dx,$$

where $(f,g) = \int_U f(x)g(x)\,dx$ denotes the standard L^2-inner product (i.e., with no weight function). Expanding the solution to (G2) yields,

$$\mathbf{S}(x,t) = \sum_{n\geqslant 1} (u_n, 1)_\rho e^{-\lambda_n t} u_n(x) = \sum_{n\geqslant 1} (u_n, \rho) e^{-\lambda_n t} u_n(x), \tag{G4}$$

where

$$0 < \lambda_1 < \lambda_2 \leqslant \dots, \tag{G5}$$

denote the (necessarily positive) eigenvalues of $-\mathcal{L}$. The corresponding eigenfunctions $\{u_n(x)\}_{n\geqslant 1}$ satisfy the following time-independent equation:

$$-\mathcal{L}u_n = \lambda_n u_n, \quad x \in U, \tag{G6}$$

and identical boundary conditions as $\mathbf{S}$. Further, the eigenfunctions are orthogonal and are taken to be orthonormal, which means that

$$(u_n, u_m)_\rho = \delta_{nm} \in \{0, 1\}, \tag{G7}$$

where δ_{nm} denotes the Kronecker delta function (i.e., $\delta_{nn} = 1$ and $\delta_{mn} = 0$ if $n \neq m$).

Further, it is straightforward to show that the probability that a decider reaches $+\theta$ before $-\theta$ conditioned on the initial belief $x \in [-\theta, \theta]$ is

$$p_1(x) := P(X(\tau) = +\theta) = \frac{1}{2}\left(\coth\left(\frac{\theta\mu}{D}\right) - 1\right) e^{\frac{\mu(\theta-x)}{D}}\left(e^{\frac{\mu(\theta+x)}{D}} - 1\right)$$

Therefore, applying (G11) and explicitly computing the integral yields

$$P(\kappa_{n(N-j)} = 1) \to \int_{-\theta}^{\theta} p_1(x)q(x)\,dx = \frac{1}{1 + e^{-\frac{\theta\mu}{D}}} = p_1(0) \quad \text{as } N \to \infty.$$

Hence, the slowest deciders out of $N \gg 1$ deciders make a decision as if they were initially unbiased [i.e., as if $X(0) = 0$].

APPENDIX H: PROOFS

Proof of Proposition 1. Since $\{(\tau_n, Z_n)\}_{n\geqslant 1}$ are identically distributed, we have that

$$P(A_{n(j)}) = \sum_{\substack{\text{distinct indices}\\ n_1,\ldots,n_N \in \{1,\ldots,N\}}} P(\max\{\tau_{n_1}, \ldots, \tau_{n_{j-1}}\} < \tau_{n_j} < \min\{\tau_{n_{j+1}}, \ldots, \tau$$

[18] J. I. Gold and M. N. Shadlen, The neural basis of decision making, Annu. Rev. Neurosci. **30**, 535 (2007).
[19] M. J. Mulder, E.-J. Wagenmakers, R. Ratcliff, W. Boekel, and B. U. Forstmann, Bias in the brain: A diffusion model analysis of prior probability and potential payoff, J. Neurosci. **32**, 2335 (2012).
[20] I. Rentschler, M. Jüttner, A. Unzicker, and T. Landis, Innate and learned components of human visual preference, Curr. Biol.