Mapping the distinct origins of bimodality in a classic model with alternative stable states

Karen Abbott, Vasilis Dakos

Mapping the distinct origins of bimodality in a classic model with alternative stable states

Karen C. Abbott · Vasilis Dakos

Abstract Stochastic ecological dynamics result from both transient and asymptotic features of the underlying system, yet explanations for observed patterns often emphasize asymptotics. For example, an ecological state (e.g. a particular population size or community composition) that occurs frequently and/or persists for a meaningful duration might be assumed to be a stable equilibrium, even though transients can also persist for a long time and may recur. In this paper, we consider one particular pattern – a bimodal distribution of states as a system is observed through time – and consider alternative causes for this pattern. First, we consider the "asymptotic" explanation that each mode corresponds to a distinct stable state. Second, we consider that one mode might correspond to a long transient. We explore the dynamics that result from each of these causes in a classic bistable model, focusing particularly on the degree of environmental stochasticity needed to generate a bimodal distribution of states in each case. Our results highlight that observations of a system's dynamics do not provide enough information to determine the number and location of stable states. We conclude that a more serious and systematic consideration for the possible role of transients in driving observed dynamics will lead to stronger insights and understanding.

K.C. Abbott ⊠
Department of Biology
Case Western Reserve University
Cleveland, OH 44106 U.S.A.
Tel: 216-368-3561
E-mail: kcabbott@case.edu

V. Dakos
Institut des Sciences de l'Evolution de Montpellier (ISEM)
Université de Montpellier
34095 Montpellier, cedex 05, France
E-mail: vasilis.dakos@umontpellier.fr

1 Introduction

Linking pattern to process is a central aspiration in ecology, but it is notoriously challenging because different processes can produce the same pattern (Cale et al. 1989, Urban and Matter 2018). When an ecological observation matches the expected behavior of a deterministic, mechanistic model, this match is often taken as evidence that the mechanisms included in the model explain the observed data. Further, such a match could be seen to suggest that although stochasticity was undoubtedly present in the real, observed system, it did not mask the signal of the underlying deterministic processes. However, our appreciation for the profound role that stochasticity plays in structuring ecological populations and communities continues to grow (Ellner and Turchin 2005, Benincà et al. 2011, Nolting and Abbott 2016, Boettiger 2018, Shoemaker et al. 2020). With this understanding comes increased interest in moving beyond thinking of stochasticity as a potential masker of patterns, but as a key part of the process that generated the pattern (Boettiger 2018). Accomplishing this change requires a better understanding of how stochasticity interacts with deterministic processes to drive particular patterns of interest.

Bimodal distributions of population sizes or other ecosystem properties are intriguing patterns, often explained through deterministic processes that give rise to alternative stable states. Indeed, identification of bimodality has been suggested as a powerful diagnostic tool of the presence of alternative stable states when direct experimental evidence is lacking or difficult to collect (Scheffer and Carpenter 2003). In particular, the increasing availability of remotely sensed data, the compilation of datasets from multiple ecosystems, and the accumulation of long-term monitoring data have recently allowed scientists to clearly document multimodal patterns in ecological variables previously hypothesized to have alternative stable states (Scheffer et al. 2015). Some of the most pronounced examples of multimodal distributions come from data on vegetation cover that reflect forest and savanna states in tropical forests (Staver et al. 2011, Hirota et al. 2011), woodland and treeless states in boreal biomes (Scheffer et al. 2012, Abis and Brovkin 2019), and states with distinct functionality in drylands (Berdugo et al. 2017).

However, it is well known that other processes besides alternative stable states can generate multimodal distributions (Scheffer and Carpenter 2003). For example, a system with a single stable state and a long transient state (associated for example with a saddle point or ghost attractor (Hastings et al. 2018)) can show bimodality that is just as strong as in systems with alternative stable states (Abbott and Nolting 2017). This bimodality occurs when stochasticity allows the system to visit a transient state from which escape is slow. When the speed of escape is similar to the rate of stochastic switching between stable states, it can be impossible to tell from observations how many stable states are present (Abbott and Nolting 2017).

In this study, we use a stochastic version of what is perhaps the best-known example of an ecological model with alternative stable states – the Lotka-Volterra competition model – to illustrate two distinct

mechanisms that can give rise to multi-modality, and to better understand the role that stochasticity plays under each mechanism. The distribution of population sizes of competing species in this model appears bimodal when there are two stable states (priority effects). In this scenario, stochasticity allows the system to switch between them. However, stochasticity can also allow multi-modality to appear even when there is only a single deterministically-stable state (competitive coexistence). This happens when stochasticity allows the system to visit the vicinity of the single-species saddle point which, though unstable, can transiently persist for ecologically meaningful amounts of time. In both cases (the bistable case and the case with a single stable state but the potential for long transients), stochasticity is crucial for the appearance of multi-modality in the distribution of population sizes. With too little noise, the system is very likely to remain near one stable state, regardless of the presence or type of other equilibria, and the distribution of population sizes will be unimodal. With too much noise, the distribution of population sizes will reflect the distribution of stochastic perturbations as the dynamics proceed mostly-randomly, revealing little about the underlying deterministic processes (noise-swamping). For both mechanisms, it is at intermediate noise levels that multi-modality can appear.

The notion that stochasticity has a 'sweet spot' – an intermediate noise level that is large enough to be qualitatively interesting but small enough not to swamp everything else out – is quite general. Classical examples are noise-sustained coexistence in a community of competing species (e.g. Chesson 1985) and population persistence and extinction risk that depend non-monotonically on the color of noise (e.g. Ripa and Lundberg 1996). More intricate cases can occur due to stochastic resonance (Wiesenfeld and Jaramillo 1998). Outbreaks in periodic populations can be amplified at a certain noise intensity (e.g. Blarer and Doebeli 1999), or noise may even provoke species oscillations in an otherwise stable system; at a certain timescale such oscillations will even be amplified (e.g. Benincà et al. 2011). More drastically, noise can even induce the appearance of underlying chaotic feedbacks (Rand and Wilson 1991), where unpredictability becomes maximal at a 'sweet spot' of noise intensity (Ellner and Turchin 2005).

What remains unexplored, however, is whether the sweet spot itself contains useful biological information. In the context of our Lotka-Volterra example, is the intermediate noise level that leads to bimodality the same regardless of the number of stable states? Or, is the sweet spot for stochastically switching between alternative stable states different from the one for shifting between a lone stable state and a long transient? If the sweet spots differ, then when we observe a bimodal distribution of states, knowledge of the noise intensity could help us to identify which underlying mechanism is acting.

Here, using stochastic versions of the Lotka-Volterra competition model, we characterize the number and locations of modes in the distribution of population densities across a range of interspecific competition coefficients (which determine the number of stable states) and noise intensities. Through simulation, we find the intermediate noise range that gives rise to multi-modal population distributions and we study the number and location of these modes. We compute quasi-potentials (Nolting and Abbott 2016) and

use them to uncover the basis for changes in the number of modes, the relative stability of each mode, and the expected transition rate between modes, regardless of whether all modes represent stable states. We find that there is a similar, intermediate noise intensity that is likely to produce bimodal distributions of states even if the underlying model is not in the bistable regime. We interpret this finding through a broader theoretical lens to conclude that in our stochastic world, ecological understanding requires us to consider both equilibrium and transient explanations for observed dynamics.

2 Methods

2.1 Models

To illustrate the origins of multi-modal population distributions, with or without alternative stable states, we use the well-known Lotka-Volterra competition model with stochasticity added in two ways. In the main text, we consider a model with additive noise,

$$dN_i(t) = \left[r_i N_i(t) \left(1 - \frac{1}{K_i} \sum_{j=1}^{2} \alpha_{ij} N_j(t) \right) \right] dt + \sigma dW_i(t) \ , \qquad N_i(t) \geq 0 \ , \quad i = 1, 2 \ , \tag{1}$$

interpreted in the Itô sense. $N_i(t)$ is the population density of species i at time t. K_i is its carrying capacity, r_i is its intrinsic population growth rate, and α_{ij} is the competitive effect of species j on species i. Stochasticity is included in the second term, in which dW_i are independent Weiner processes with mean 0 and infinitesimal variance 1. The parameter σ controls the strength of stochasticity.

To aid biological interpretation at the expense of notational precision, we rewrite Eqn. (1) in the form of an ordinary differential equation (the standard Lotka-Volterra competition model) with an additional random variable, $\varepsilon_i(t)$, representing a stochastic density independent immigration/emigration at rate,

$$\frac{dN_i(t)}{dt} = r_i N_i(t) \left(1 - \frac{1}{K_i} \sum_{j=1}^{2} \alpha_{ij} N_j(t) \right) + \varepsilon_i(t) \ , \qquad N_i(t) \geq 0 \ , \quad i = 1, 2. \tag{2}$$

We model $\varepsilon_i(t)$ as a temporally-uncorrelated Gaussian stochastic process with mean 0 and variance σ. We simulate Eqn. (2) by discretizing the ordinary differential equation and drawing $\varepsilon_i(t)$ values as described in the next section; this is equivalent to simulating the stochastic differential equation in Eqn. (1) using the Euler-Maruyama method.

Because it represents population density, $N_i(t)$ is constrained to be non-negative. If the value of $\varepsilon_i(t)$ drawn at a particular time would dictate that more individuals of species i emigrate than are currently present (e.g. a large negative value of $\varepsilon_i(t)$ when $N_i(t)$ is small), we avoid a nonsensical result by setting $N_i(t)$ to zero. Subsequent random immigration (i.e. positive $\varepsilon_i(t)$) could rescue the population from local extinction in the additive noise model.

In the supplementary materials (Online Resource 1), we also consider a model with stochastic perturbations to the net per capita population growth rate,

$$\frac{dN_i(t)}{dt} = r_i N_i(t) \left(1 - \frac{1}{K_i} \sum_{j=1}^{2} \alpha_{ij} N_j(t) \right) + \varepsilon_i(t) N_i(t) , \quad i = 1, 2. \tag{3}$$

We call this the multiplicative noise model because the random variable, ε_i, now appears in a term that multiplies population size, $N_i(t)$. In this model, a population that goes extinct remains extinct.

The benefit of the additive noise model is that the variance in the stochastic term is independent of the current population size, which allows us to manipulate stochastic variance independently from the parameters that influence the stationary distribution of population sizes. Also, because extinction is not permanent in the additive noise model (just as local extinction is typically not permanent in real populations that are not isolated), the additive noise model allows us to observe repeated transitions to and from temporary competitive exclusion. For these reasons, we focus on the additive noise model for the remainder of the main text.

2.2 Analysis

We are interested in the role of stochasticity in driving the appearance of multi-modal distributions of population size and, particularly, whether this role is different when the deterministic model is bistable versus when it is not. To study the effect of noise intensity, we varied σ, the standard deviation of the noise term, $\varepsilon_i(t)$.

We set $\alpha_{ii} = 1$ and $K_i = 1$ for $i = 1, 2$ and study the model's behavior across α_{12}, α_{21} combinations that span two bifurcations: from stable coexistence to competitive exclusion of species 1; then from competitive exclusion of species 1 to bistability (each species can stably exclude the other). Throughout, we maintain the relationship $\alpha_{12} = 1.03\alpha_{21}$ and adjust α_{21} to cross these bifurcations, which occur as α_{12} and then α_{21}, respectively, cross 1. We set $r_1 = 0.1$ and $r_2 = 0.15$, meaning that the slightly stronger competitor (species 2) is also a faster grower. Supplemental analyses (not shown) confirm that our main results are not specific to this parameter setting.

2.2.1 Stochastic simulations

We simulated the stochastic models in MATLAB by discretizing Eqn. (2) with a step size of $dt = 0.01$ (see Online Resource 2, Fig. 11 for examples). Before simulation, we drew 50 independent sequences of $\varepsilon_i(t)$ values with mean 0 and variance 1 for each $i = 1, 2$. These sequences were rescaled to have variance $\sigma^2 dt$ and then used to generate 50 replicate realizations of the stochastic dynamics for each parameter set (combination of α_{21} (with $\alpha_{12} = 1.03\alpha_{21}$) and σ). Because we reuse the same sequences of $\varepsilon_i(t)$, note that any differences we see when we compare across parameter sets are due to changes in the parameter

143 values rather than variance in the stochastic process itself. Stochastic simulations were run for 1000 time

144 units from initial conditions near the coexistence equilibrium ("near" meaning displaced from the interior

145 equilibrium by the first drawn values of ε_1 and ε_2 adjusted to have the appropriate variance).

146 *2.2.2 Modes in the distribution of population sizes*

147 We estimated the number of modes in the distribution of population sizes in the last 500 time units of the

148 simulated time series. Multi-modality was assessed by fitting a 2D Gaussian mixture distribution to the

149 time series of both species (see Online Resource 2, Fig. 12 for examples). As we had four deterministic

150 equilibria (extinction, species 1 alone, species 2 alone, coexistence), we fitted distributions of 1 to 5 com-

151 ponents (mixtures of 1 to 5 Gaussian distributions) to allow for an extra degree of freedom when selecting

152 for the best fitting component based on the minimum Bayesian Information Criterion (BIC). We did this

153 using the *fitgmdist* function in MATLAB that employs a maximum likelihood procedure by selecting a

154 full covariance type and a regularization parameter set to 0.005 to avoid poorly estimated covariances

155 (that occurred when variation in our time series was minimal, especially in the case of multiplicative

156 noise). We used the number of components returned as an estimate of the different clusters (modes) in

157 the time series. We also estimated the Euclidean distance between the means of each component as a

158 measure of difference between clusters (modes), and we approximated the size of each cluster (mode)

159 by estimating the average Mahalanobis distance of all points to the mean of each cluster. We applied

160 these estimations for each of the 50 realizations of each of 2,500 $\alpha_{21} - \sigma$ combinations (we changed

161 α_{21} from 0.5 to 1.5, and σ from 0.01 to 0.5 in log scale, both in 50 steps, maintaining the relationship

162 $\alpha_{12} = 1.03\alpha_{21}$ throughout). Lastly, for each of these $\alpha_{21} - \sigma$ combinations, we report summary statistics

163 averaged across the 50 realizations on the mean number of modes, coefficient of variation (CV) in number

164 of modes, median number of modes, modal number of modes, standard deviation in number of modes,

165 and evenness of clusters, as well as average size of clusters and average distance between clusters in state

166 space. We report the first two of these summary statistics in the main text, and the rest appear in Online

167 Resource 3.

168 *2.2.3 Quasi-potentials*

169 We computed quasi-potential surfaces (Moore et al. 2015, 2016, Nolting and Abbott 2016) for the additive

170 noise model (Eqn. (2)) under four different parameter combinations: two in the parameter region with

171 stable coexistence and two in the bistable region. The quasi-potential is a generalization of the better-

172 known concept of a potential function. The potential of a deterministic system is a function that describes

173 how the system's state will change through time, as a function of its current state. It is often interpreted

174 as a landscape, with valleys and peaks that correspond to states the system will move toward or away

175 from, respectively, like a ball rolling on the landscape. Because it describes the global dynamics – that

¹⁷⁶ is, state changes for *any* state – the potential is a much more powerful tool than local linear stability
¹⁷⁷ analysis that only characterizes dynamics close to an equilibrium state. However, most ecological models
¹⁷⁸ of two or more species lack a potential function, because their dynamics are governed by additional forces
¹⁷⁹ (like a propensity to cycle) besides just the pull "downhill" toward stable states. The quasi-potential is
¹⁸⁰ a surface that can be interpreted in the same way as the potential, and it can be computed numerically
¹⁸¹ for deterministic models that, like the Lotka-Volterra model, lack a true potential function (Moore et al.
¹⁸² 2015, 2016, Nolting and Abbott 2016). It is therefore an ideal tool for characterizing the global dynamics
¹⁸³ of our 2-dimensional system.

¹⁸⁴ For stochastic models such as Eqn. (2), we compute the quasi-potential for the deterministic part of
¹⁸⁵ the model only. Stochasticity can be thought of as random jostling of the ball that would otherwise tend
¹⁸⁶ to roll downhill on the quasi-potential surface. The difference in quasi-potential height between two states
¹⁸⁷ quantifies how readily the system will move from one to the other. In other words, a larger stochastic
¹⁸⁸ perturbation is needed to move to a state that is steeply uphill from the current state, and a deeper
¹⁸⁹ valley can be thought of as more stable than a shallow valley (Nolting and Abbott 2016). We computed
¹⁹⁰ quasi-potentials using the QPot package (Moore et al. 2015) in R (R Core Team 2017) following the
¹⁹¹ methods detailed in Moore et al. (2016) for $\alpha_{21} = 0.7, 0.9, 1.1,$ and 1.3 and other parameters as above.

¹⁹² *2.2.4 First passage times*

¹⁹³ In ecosystems with alternative stable states, a key property we often wish to understand is the expected
¹⁹⁴ frequency with which the system should shift from one stable state to another. The mean first passage
¹⁹⁵ time (MFPT) is the expected waiting time for a system at one stable state (in one valley on the potential
¹⁹⁶ or quasi-potential landscape) to cross an intervening unstable equilibrium (a peak), after which it will
¹⁹⁷ converge onto a different stable state (roll into a different valley). We calculated mean first passage times
¹⁹⁸ for the same four parameter combinations ($\alpha_{21} = 0.7, 0.9, 1.1,$ and 1.3). In all cases, we calculated the
¹⁹⁹ passage times between the coexistence equilibrium and each of the exclusion equilibria ($N_1 = 1, N_2 = 0$
²⁰⁰ and $N_1 = 0, N_2 = 1$). When the coexistence equilibrium was stable ($\alpha_{21} = 0.7, 0.9$), the passage time
²⁰¹ reflects how long it takes to get *from* this state to an unstable exclusion equilibrium. When the coexistence
²⁰² equilibrium was unstable, the passage time is how long it takes to get *to* this state from a stable exclusion
²⁰³ equilibrium. Thus, in all cases, our first passage times reflect the amount of time the system will stay
²⁰⁴ near a stable state before reaching an unstable one.

²⁰⁵ We simulated model Eqn. (2) using the procedure described above (see *2.2.1 Stochastic simulations*)
²⁰⁶ to find the mean and distribution of first passage times. For each of the four example parameter sets, and
²⁰⁷ for each of 12 values of σ between 0.001 and 1, we ran 50 replicate simulations beginning near a stable
²⁰⁸ equilibrium. We ran each simulation until the populations left the vicinity of the stable equilibrium in
²⁰⁹ the direction of the adjacent saddle (see Online Resource 4). The time at which this exit occurred was

210 recorded as the first passage time and the MFPT was the mean first passage time in the 50 realizations

211 for a given parameter set.

212 We can also compute mean first passage times analytically from the quasi-potentials (Nolting and

213 Abbott 2016), although this method relies on a small-noise approximation and is thus expected to

214 become inaccurate as σ increases. The calculation is based on the difference in the quasi-potential height

215 at the stable and unstable equilibria – a greater height difference (i.e. a deeper valley) means that

216 larger stochastic perturbations would be needed to move the system "uphill" away from the stable state,

217 resulting in a longer first passage time. If the quasi-potential height is V_s at the stable equilibrium and

218 V_u at the unstable equilibrium, the MFPT from the stable to unstable equilibrium under noise intensity

219 σ is,

$$\text{MFPT} \approx \frac{2\pi}{|\lambda_{1u}|} \sqrt{\frac{|\det \nabla^2 V_u|}{\det \nabla^2 V_s}} \exp\left(\frac{2(V_u - V_s)}{\sigma^2} \right) \tag{4}$$

220 (Bouchet and Reygner 2015, Nolting and Abbott 2016), where $|\lambda_{1u}|$ is the magnitude of the dominant

221 eigenvalue of the Jacobian matrix evaluated at the unstable equilibrium and $\nabla^2 V_x$ $(x = s, u)$ is the

222 Hessian matrix (i.e. a matrix of second-order partial derivatives) evaluated at V_x. Once we have computed

223 quasi-potentials, we have the necessary heights, V_x, at the equilibria as well as heights at nearby states

224 to estimate the derivatives.

3 Results

226 Simulated population dynamics from the additive noise model (Eqn. (2)) illustrate the two distinct ways

227 in which bimodality in the distribution of population sizes can arise. Bimodality can arise from true

228 bistability (Fig. 1a–b), or from stochastic dynamics around a single stable equilibrium with the influence

229 of long transients elsewhere in state space (Fig. 1c–d). Both of these examples are in the 'sweet spot':

230 stochasticity is strong enough to allow us to observe more than just a unimodal distribution of population

231 sizes around a single stable state, yet not so strong as to mask the biological mechanisms represented in

232 the model.

233 By analyzing many replicate simulations across α_{ij} values, we can map out that sweet spot: the

234 σ values that lead, on average, to bimodal distributions of the state variables (Fig. 2a-b), with modes

235 corresponding to biological features of the model. Moving left to right along the x-axis of panel (a) moves

236 the system from a region with stable coexistence ($\alpha_{21} < 0.97$ so that $(\alpha_{12} = 1.03\alpha_{21}) < 1$), through a

237 narrow window ($0.97 < \alpha_{21} < 1$) where species 1 is competitively excluded, then (when $\alpha_{21} > 1$) into

238 the bistable region. Thus, multi-modality (i.e. instances where the mean number of modes observed in

239 simulations meaningfully exceeds 1) to the left of the dashed line in Fig. 2a cannot be due to alternative

240 stable states, whereas multi-modality on the right-hand side may be.

On the left-hand side, where the deterministic model has a single stable state corresponding to species coexistence ($\alpha_{12} < 1$), our baseline expectation would be a unimodal distribution of population sizes around this coexistence state. For sufficiently small noise, this is what we see, but for higher intensity noise, the populations are able to visit the proximity of the two saddle points at $N_1, N_2 = (0, 1)$ and $(1,0)$, and relatively slow escape from these saddles creates additional modes in the population distribution (Fig. 2c,e). In these instances, we may see two modes (as in the Fig. 1 examples) as populations drop to lower density and are slow to recover. These modes are shifted toward the two unstable boundary equilibria ($N_1, N_2 = (0, 1)$ and $(1,0)$) relative to the stable coexistence equilibrium (Fig. 2c). We may also see three distinct modes, one near the coexistence equilibrium and the other two closer to the boundary equilibria (Fig. 2e). We are more likely to observe three modes with higher intensity noise (larger σ, Fig. 2a). As α_{21} (and thus α_{12}) increase toward the bistable region, the beginning of the sweet spot – the minimum intensity noise needed to produce, on average, two or more modes – declines (Fig. 2a).

In contrast, in the bistable region ($\alpha_{21} > 1$, right of the dashed line in Fig. 2a), the location of the sweet spot appears to have weaker dependence on the strength of interspecific competition. For low intensity noise (such as the region blown up in Fig. 2b), the amount of noise needed to cause switching between alternative stable states increases the deeper into the bistable region we go. However, for high intensity stochasticity (i.e. the top part of Fig. 2a, still right of the dashed line), the average number of modes is fairly insensitive to the strength of competition. At sufficiently high stochasticity, we again may see bi- (Fig. 2d) or trimodal distributions (Fig. 2f).

The shapes of the quasi-potentials can help us understand these observations. The quasi-potential surface transitions from single-welled to double-welled when the deterministic model undergoes bifurcation from stable coexistence to bistable competitive exclusion (Fig. 3). Less stochasticity is needed to transition over flatter parts of the quasi-potential surface. Therefore, Figs. 3a–b suggest that when interspecific competition is weaker (a relative to b), the coexistence state is more stable and so a higher intensity of noise will be needed to push the system away from stable coexistence to one of the relatively flat regions around the boundary equilibria. This change in quasi-potential shape explains the downward slope of the contour lines on the left-hand side of Fig. 2a: far from the bistable region, the quasi-potential well around the coexistence state is deep (as in Fig. 3a) and strong stochasticity is needed to see more than a single mode, but closer to the bistable region, the quasi-potential is relatively flat between the three equilibria (as in Fig. 3b) and multi-modality can appear with weaker noise.

Under parameter values for which the deterministic model has alternative stable states, the quasi-potential is double-welled. Lower intensity noise should be needed to transition between shallower wells (as in Fig. 3c) than between deeper wells (Fig. 3d). This explains the increasing slope of the contour lines on the right-hand side of Fig. 2a (and in Fig. 2b), at least for lower values of σ. We do not see

275 evidence of this effect with larger σ suggesting that at some point, the noise is sufficient to cause similar

276 transitions between wells despite differences in quasi-potential well depth.

277   The quasi-potential allows us to estimate mean first passage times, although this estimate is only

278 accurate for small σ (Fig. 4, solid curves near the y-axis). For larger values of σ, we used simulations

279 to calculate mean and the distribution of first passage times (Fig. 4, violin plots). Intuitively, mean

280 first passage times are shorter when noise intensity is higher (higher σ), regardless of the strength of

281 competition. First passage times were longer and more variable when there was a single stable equilibrium

282 (Fig. 4a,b) than when there were alternative stable states (Fig. 4c,d).

283   In aggregate, our results reveal strong similarities in the stochastic dynamics of this model in two

284 different regimes (single stable state versus alternative stable states). In both regimes, bimodal distribu-

285 tions of states appear when the strength of stochasticity is in an intermediate 'sweet spot', with modes

286 corresponding to both stable equilibria and saddles. This sweet spot can be reached with weaker noise

287 levels the closer the system is to the bifurcation that separates the regimes. In both regimes, this effect

288 can be understood by observing that the quasi-potentials necessarily flatten near this bifurcation. In

289 both regimes, mean first passage times are shorter for stronger noise, and trimodal distributions of states

290 are more likely to appear with stronger noise.

291   So, what are the major differences between the regimes? We can see that higher intensity noise is

292 generally required to produce multi-modal distributions of population sizes when there is a single stable

293 state (Fig. 2a). We also find that there is more variability among replicates in the number of modes when

294 there is a single stable state (Fig. 5, which shows more yellow shades on the left-hand side and more blue

295 on the right). In other words, the likelihood of consistently observing exactly two modes is greater when

296 those modes correspond to alternative stable states. We also saw higher variance in first passage times

297 (more elongated violin plots in Fig. 4a-b relative to c-d) with a single stable state.

298   The multiplicative noise model (Eqn. (3)) leads to similar conclusions (Online Resource 1). We again

299 see bimodality in both regimes, with a sweet spot in the absence of alternative stable states around 0.3

300 for our parameter setting, compared to about 0.45 when the model is bistable (Fig. 7 in Online Resource

301 1). The fact that the sweet spot for σ is now higher, not lower, in the bistable regime may simply

302 reflect differences in the population densities that multiply the noise term in Eqn. (3). In any case, the

303 sweet spots are again similar on the two sides of the bifurcation. Like the additive noise model, when

304 there is only one stable state, the noise intensity needed to produce an average of two modes declines

305 as competition increases toward the bistable region. Within the bistable region, the number of modes

306 is again relatively insensitive to changes in the competition coefficients. As in the additive noise model,

307 the number of modes is more variable among realizations when there is only one stable state. The only

308 major qualitative difference between models, which is a direct consequence of their formulation, is that

309 the multiplicative model is extinction prone under high intensity noise.

4 Discussion

The Lotka-Volterra model is a textbook example of alternative stable states and is often a student's first encounter with priority effects. Even though its behavior is so well known, these analyses revealed new insights. Most importantly, we found that there is a similar (roughly $\sigma = 0.2$-0.4), intermediate noise intensity that is likely to produce bimodal distributions of states *whether or not* the underlying model is in the bistable regime (Fig. 2a). This occurs because transient visits to the vicinity of unstable saddle points can have very similar effects as visits to alternative stable states (Fig. 2c-f). This in turn is due to the fact that both saddles and stable states occur in locally flat regions of the quasi-potential surface (Fig. 3).

Quasi-potentials were useful in this study for visualizing the deterministic part of the dynamics in a way that permits an intuitive understanding of the effects of adding noise (e.g. by noting flat versus steep areas). Height differences between two points on the surface reflect the amount of "work" needed to counteract the downhill pull to move between those states along the most likely path (Freidlin and Wentzell 2012). By considering the most likely path, lower probability events (such as those considered in large deviation theory; see e.g. Wainrib 2013) are not reflected. This feature explains why mean first passage times derived from the quasi-potential became less accurate as the noise intensity – and thus the opportunity to travel along paths farther from the most likely one – increased (Fig. 4). There are other ways to extract a potential-like surface from a stochastic process, and each provides a different and complementary interpretation. For example, the Helmholtz-Hodge decomposition yields a surface in which height differences reflect work needed to move along an average, rather than most likely, path between two states (Bhatia et al. 2013, Strang 2020). Whether this surface more accurately predicts quantities like mean first passage time in ecological models remains to be seen. However, this surface no longer corresponds directly to the deterministic part of the process (e.g. peaks and wells no longer occur at the equilibria of the underlying ordinary differential equation model) and thus much of the intuitive appeal of the quasi-potential is lost. It is for this reason that we opted to use the quasi-potential in the present study.

Our results make clear that bimodality alone cannot be taken as evidence of alternative stable states. In the Introduction, we asked whether the location of the 'sweet spot' – the intermediate noise level needed to see bimodality – could help us discriminate between true alternative states and look-alikes. Our anti-climactic answer is, maybe but probably not. We did observe differences in the strength of noise at the sweet spot (stronger when there is only one stable state and the system is not near the bifurcation; Fig. 2). We also found higher variability among replicates in first passage times (Fig. 4) and number of modes (Fig. 5) when there was a single stable state versus when there were alternative stable states. However, none of these differences was dramatic. To be useful as a diagnostic, we would need much larger

344 (e.g. order of magnitude) differences that would allow us to recognize noise as strong or variance as high

345 without a detailed comparative study (stronger or higher than what?).

346 By studying a different model capable of alternative stable states, and again comparing bistable

347 and uni-stable regimes, Abbott and Nolting (2017) illustrated how saddle points effectively masquerade

348 as additional stable states due to their similar influence on the stochastic dynamics. Time series of

349 stochastic transitions between two stable states were indistinguishable from transitions between a single

350 stable state and a saddle by any measure we considered: the relationship between mean and equilibrium

351 population sizes; the temporal variance in population sizes; and the results of several statistical tests for

352 the strength of bimodality (Abbott and Nolting 2017). We did not, however, test for differences in the

353 amount of stochasticity needed to see bimodality, and the present study fills that gap. Taken together,

354 we see strong evidence that because long transients can occur near saddles, the effects of saddles and

355 asymptotically stable states on stochastic dynamics are exceedingly similar.

356 The influence of saddles on transient dynamics is well-known (Hastings et al. 2018, Morozov et

357 al. 2020), and yet the tacit presumption that common and or recurrent states must correspond to

358 stable, asymptotic equilibria remains ingrained in ecological thinking. This is understandable, because

359 asymptotic equilibria are intuitively appealing and because some common states really do correspond

360 to stable states (e.g. Fig. 1a). Nevertheless, this study joins the call to consider both asymptotic and

361 transient explanations for observed dynamics (Hastings 2001, 2004, Abbott and Nolting 2017, Hastings

362 et al. 2018, Francis et al. in review).

363 We began this article with a nod to the age-old problem of linking pattern to process, and on the sur-

364 face, it may appear that our results only add to the list of reasons this is difficult. Looking more deeply,

365 though, it points to new ways forward. For instance, while bimodality should not be taken on its own as

366 evidence of alternative stable states, it could imply the existence of other non-stable states like saddles

367 and ghost attractors that could potentially become stable if conditions change. More realistic patterns of

368 stochasticity (like correlated noise) might cause the modes we identified to become quasi-stable for the

369 relevant timescales of a particular system; this warrants further investigation. Our results also highlight

370 a need for mechanistic models in the interpretation of data on population and community dynamics.

371 If a validated mechanistic model for a particular system has alternative stable states, this provides di-

372 rect evidence that bimodal state distributions are likely due to underlying bistability. Importantly, if a

373 validated mechanistic model does *not* show alternative stable states yet the data appear bimodal, our

374 analysis shows that the model may nonetheless be correct! Deeper understanding of stochastic and tran-

375 sient dynamics serves to strengthen data-model integration, ultimately leading to better understanding

376 of real ecological systems.
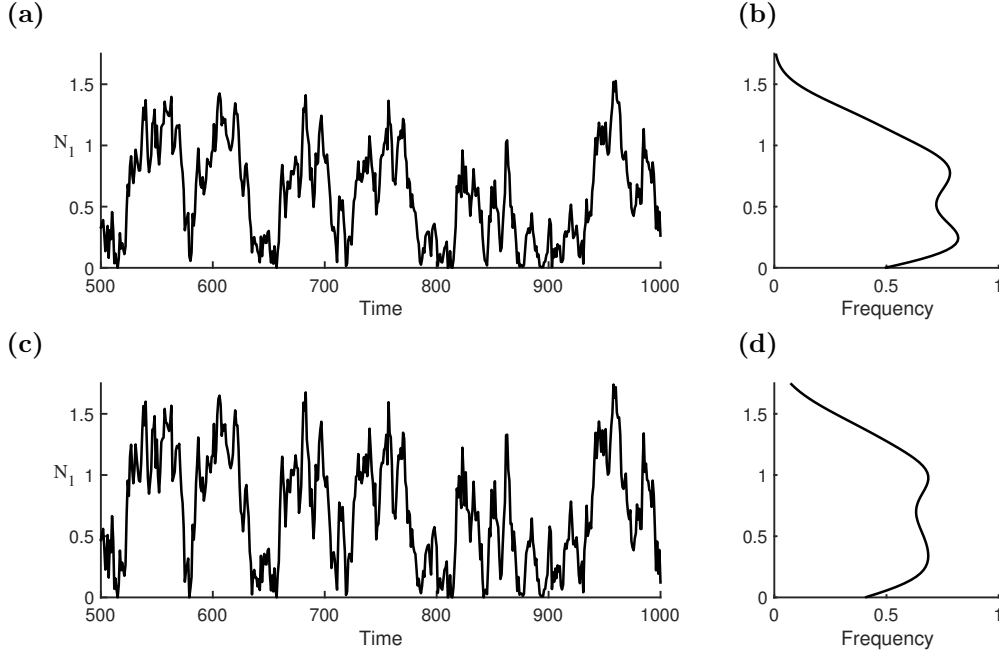
Fig. 1 Single realizations of model Eqn. (2), shown as (a,c) time series and (b,d) the corresponding distributions of abundances for species 1. The example in (a–b) shows bimodaility that arises from bistability of the two boundary equilibria. (Recall that random immigration prevents permanent extinction in this model, so the boundary equilibria are not absorbing. Because populations are prevented from becoming negative, the modes created by the two stable states are shifted inward from 0 and 1.) The example in (c–d) shows bimodality that arises from stochastic switching between the stable coexistence equilibrium at $N_1 \approx 0.56$ and the unstable boundary equilibria. The same sequences of Gaussian(0,1) values were used to construct $\varepsilon_i(t)$ in both examples, adjusting by the appropriate σ. In all panels, $r_1 = 0.1, r_2 = 0.15, K_1 = K_2 = 1$; in (a–b) $\alpha_{21} = 1.3, \alpha_{12} = 1.339, \sigma = 0.15$; in (c–d), $\alpha_{21} = 0.7, \alpha_{12} = 0.721, \sigma = 0.2$. Note that although we show just N_1 here for clarity of presentation, our analyses are based on the bivariate distribution of N_1, N_2 values, as shown in Fig. 2.
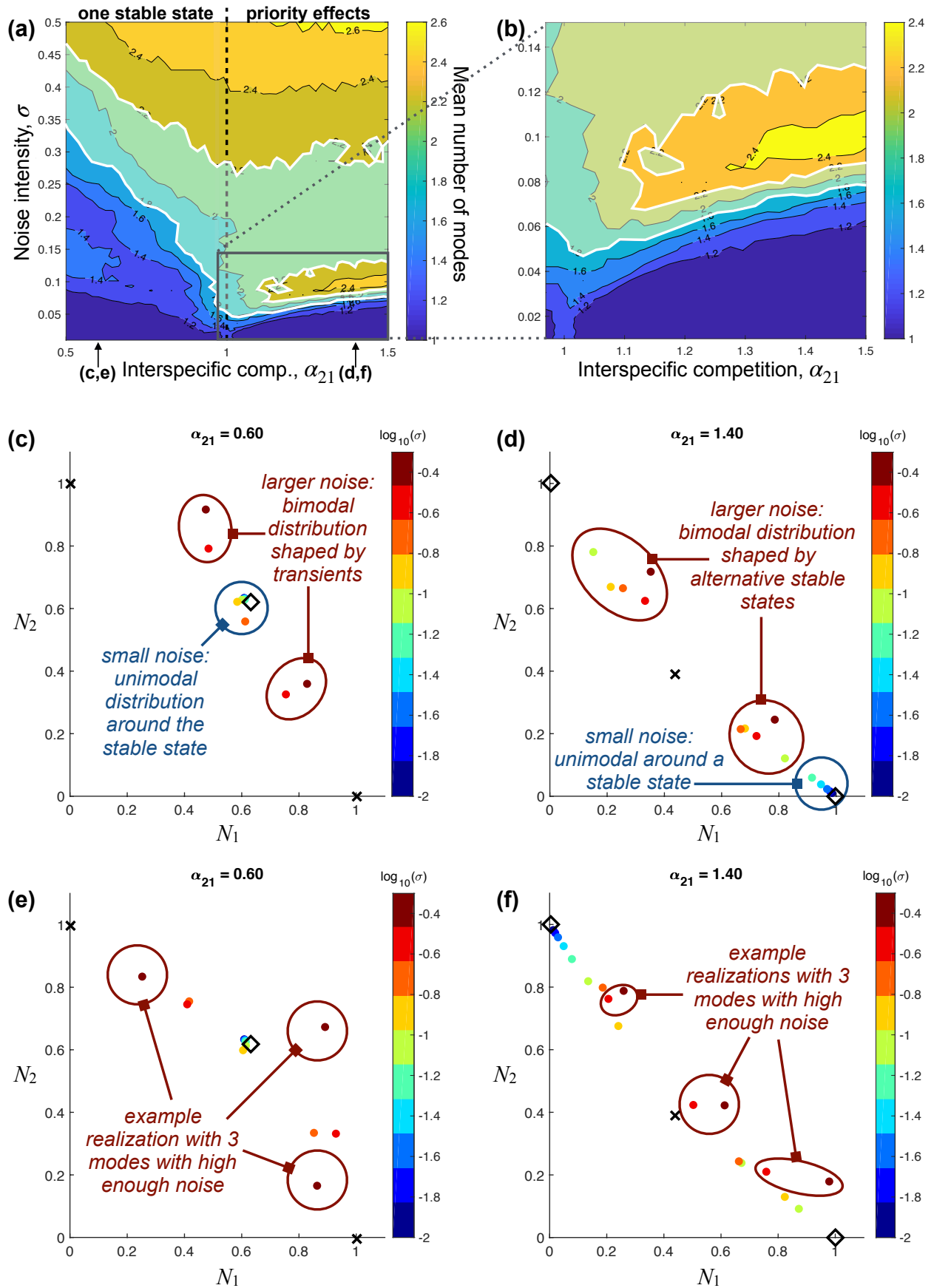
Fig. 2 See next page for caption.

Fig. 2 caption: (a-b) Mean number of modes for different combinations of α_{21} and σ, averaged across 50 realizations; (b) is a zoomed in view from (a). The 'sweet spot' for bimodality is outlined and shaded with white. (We define the sweet spot here as parameter combinations resulting in a mean number of modes between 1.8 and 2.2. Note that this definition is conservative relative to other possible definitions, like a median or mode of 2 modes (Online Resource 3).) The deterministic model is bistable to the right of the dashed line. α_{21} values used in subsequent panels are marked with arrows along the x-axis. (c-f) Locations of the modal population sizes in example realizations of the model. Colors correspond to different noise intensities (σ values), and circles mark the mode(s) observed in a single realization for each noise intensity. Multiple dots of the same color signify multi-modal distributions under that noise intensity. Coexistence and boundary equilibria are marked with a black diamond if stable and a black $\times$ if unstable. For the α_{21} value used in (c,e), the species would stably coexist in the absence of noise. For small to moderate noise intensities, the modal population sizes correspond to the stable coexistence equilibrium. However, with high noise intensity, the distribution becomes (c) bimodal or even (e) trimodal as populations have the opportunity to transiently visit the unstable single-species equilibria. (As in Fig. 1, the modes are shifted inward from the $N_1 = 0$ and $N_2 = 0$ boundaries due to the stochastic immigration/emigration term, and this effect increases with the noise intensity.) In (d,f), the noise-free model is bistable, with one species excluding the other. With small enough noise, we see a unimodal distribution of population sizes around one of these stable equilibria (which one depends on the particular sequence of noise). However, with higher noise intensity, the distribution becomes (d) bimodal due to the ability of the populations to switch between stable states, or (f) trimodal as populations transiently visit the unstable coexistence equilibrium as well.
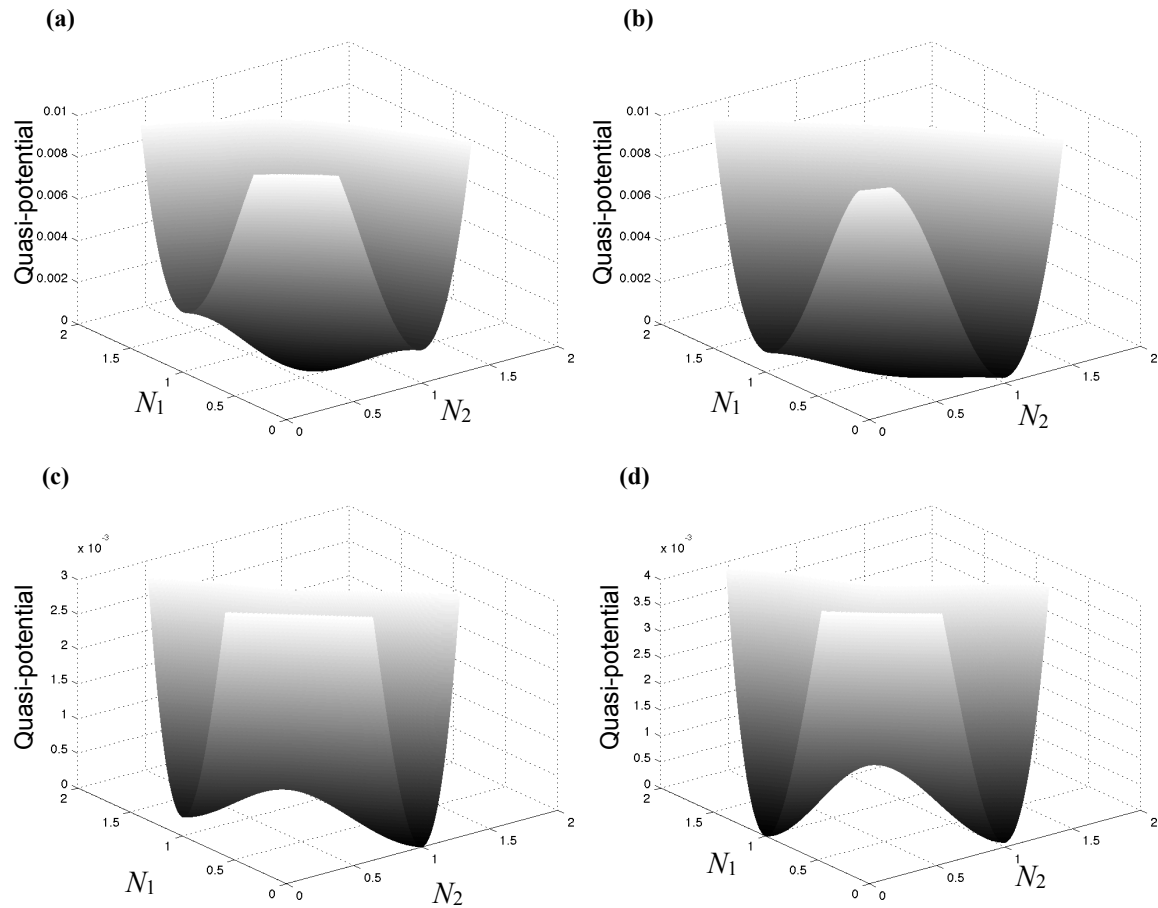
Fig. 3 Quasi-potentials for 4 levels of interspecific competition: (a) $\alpha_{21} = 0.7$, (b) $\alpha_{21} = 0.9$, (c) $\alpha_{21} = 1.1$, (d) $\alpha_{21} = 1.3$. In all panels, $\alpha_{12} = 1.03 \times \alpha_{21}$. There is one stable equilibrium (coexistence) in (a,b) and in (c,d) the model is bistable. Grayscale shading scales with height on the z-axis to aid interpretation of the images. $r_1 = 0.1, r_2 = 0.15, K_1 = K_2 = 1, \alpha_{11} = \alpha_{22} = 1$.
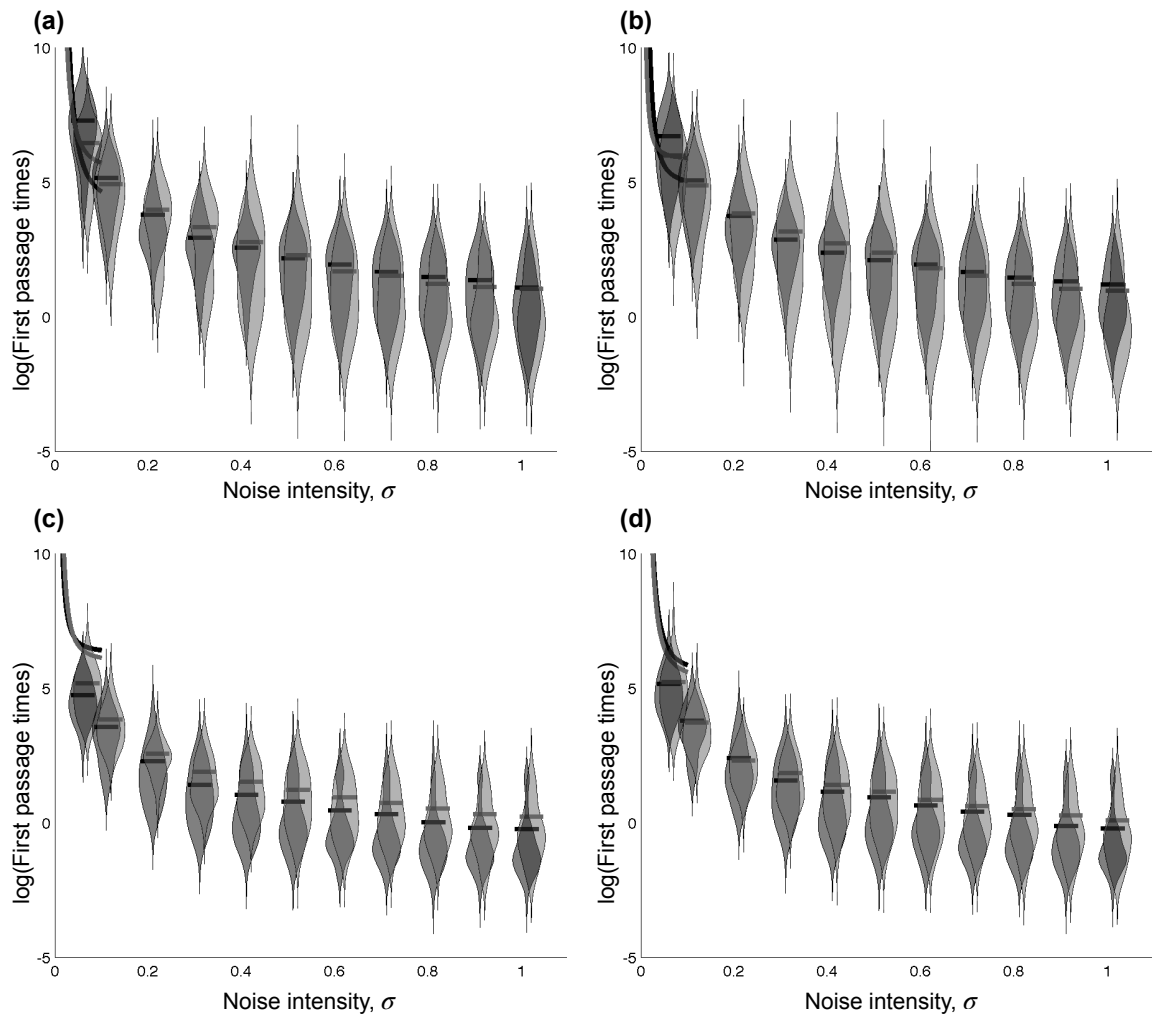
Fig. 4 Log of first passage times for (a) $\alpha_{21} = 0.7$, (b) $\alpha_{21} = 0.9$, (c) $\alpha_{21} = 1.1$, (d) $\alpha_{21} = 1.3$, for different values of noise intensity, σ. Violin plots (gray shapes, jittered to allow both shapes at each σ value to be seen) were computed from simulations of model Eqn. (2). Darker gray shapes show the distribution of log first passage times between the coexistence equilibrium and the N_1-only boundary; shape width represents the frequency with which each transition time was observed. The solid bars superimposed on each shape mark the log of the mean first passage time. Lighter, semi-transparent shapes show the same information for passage times between the coexistence equilibrium and the N_2-only boundary. Mean first passage times were also calculated from the quasi-potentials for small noise ($0 < \sigma \leq 0.1$) and are plotted as black (for transitions to and from the N_1-only equilibrium) and gray (to/from the N_2-only equilibrium) curves in the upper left part of each panel. $r_1 = 0.1, r_2 = 0.105, K_1 = K_2 = 1, \alpha_{11} = \alpha_{22} = 1$.
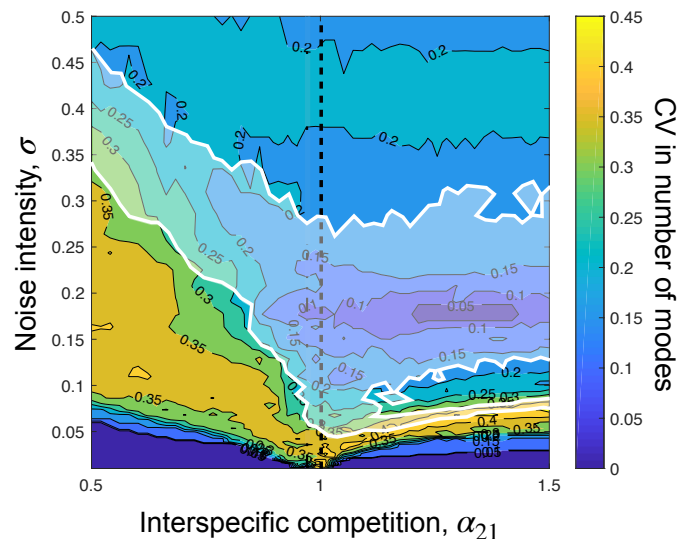
Fig. 5 Coefficient of variation in the number of modes, for the same 50 realizations as those plotted in Fig. 2a. The 1.8 and 2.2 contours from Fig. 2a are redrawn in white and the area in between shaded to delineate the bimodal region. The bistable region is to the right of the dashed line. Notice that in the bistable region on the right, CVs tend to be lower overall (more blue and fewer yellow shades) and lower in particular within the bimodal region outlined in white (CV $\approx 0 - 0.2$ on the right, versus $\approx 0.2 - 0.3$ on the left). This means that even when we expect a bimodal state distribution (mean number of modes between 1.8 and 2.2, outlined in white), the actual number of modes observed in different realizations will be less variable when there are alternative stable states.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Abbott KC, Nolting BC (2017) Alternative (un)stable states in a stochastic predator-prey model. Ecol Complex 32B:181–195.

Abis B, Brovkin V (2019) Alternative tree-cover states of the boreal ecosystem: a conceptual model. Global Ecol Biogeogr 5:612–627.

Benincà E, Dakos V, Van Nes E, Huisman J, Scheffer M (2011) Resonance of plankton communities with temperature fluctuations. Am Nat 4:E85–95.

Berdugo M, Kéfi S, Soliveres S, Maestre FT (2017) Plant spatial patterns identify alternative ecosystem multifunctionality states in global drylands. Nat Ecol Evol 1:0003.

Bhatia H, Norgard G, Pascucci V, Bremer, P (2012) The Helmholtz-Hodge Decomposition–A Survey. IEEE Trans Vis Comp Graphics 19:1386–1404.

Blarer A, Doebeli M (1999) Resonance effects and outbreaks in ecological time series. Ecol Lett 2:167–177.

Boettiger C (2018) From noise to knowledge: how randomness generates novel phenomena and reveals information. Ecol Lett 21:1255–1267.

Bouchet F, Reygner J (2015) Generalisation of the Eyring–Kramers transition rate formula to irreversible diffusion processes. Annales Henri Poincaré 17:3499–3532.

Cale WB, Nenbry GM, Yeakley JA (1989) Inferring process from pattern in natural communities. BioScience 39:600–605.

Chesson PA (1985) Coexistence of competitors in spatially and temporally varying environments. Theor Pop Biol 3:263–287.

Ellner SP, Turchin P (2005) When can noise induce chaos and why does it matter: a critique. Oikos 3:620–631.

Francis T, Abbott KC, Cuddington K, Gellner G, Hastings AM, Lai Y-C, Morozov A, Petrovskii S, Zeeman ML (in review) Management implications of long transients in ecological systems. Nat Sustain.

428 Freidlin MI, Wentzell AD (2012) Random Perturbations of Dynamical Systems , 3rd Edition. Springer,
429   Berlin.

430 Hastings A (2001) Transient dynamics and persistence of ecological systems. Ecol Lett 4:215–220.

431 Hastings A (2004) Transients: the key to long-term ecologial understanding? Trends Ecol Evol 19:39–45.

432 Hastings AM, Abbott KC, Cuddington K, Francis T, Gellner G, Lai Y-C, Morozov A, Petrovskii S,
433   Scranton K, Zeeman ML (2018) Transient phenomena in ecology. Science 361:eaat6412.

434 Hirota M, Holmgren M, Van Nes EH, Scheffer M (2011) Global resilience of tropical forest and savanna
435   to critical transitions. Science 334:232–235.

436 Moore CM, Stieha CR, Nolting BC, Cameron MK, Abbott KC (2015) QPot: An R package for calculating
437   quasi-potentials. https://cran.r-project.org/web/packages/QPot/index.html

438 Moore CM, Stieha CR, Nolting BC, Cameron MK, Abbott KC (2016) QPot: An R package for stochastic
439   differential equation quasi-potential analysis. The R Journal 8:19–38.

440 Morozov A, Abbott KC, Cuddington K, Francis T, Gellner G, Hastings AM, Lai Y-C, Petrovskii S,
441   Scranton K, Zeeman ML (2020) Long transients in ecology: theory and applications. Phys Life Rev
442   32:1–40.

443 Nolting BC, Abbott KC (2016) Balls, cups, and quasi-potentials: quantifying stability in stochastic
444   systems. Ecology 97:850–864.

445 R Core Development Team (2017) R: A Language and Environment for Statistical Computing. R Foun-
446   dation for Statistical Computing. https://www.R-project.org

447 Rand D, Wilson H (1991) Chaotic stochasticity: A ubiquitous source of unpredictability in epidemics.
448   Proc Roy Soc B 246:179–184.

449 Ripa J, Lundberg P (1996) Noise colour and the risk of population extinctions. Proc Roy Soc B 1377:1751–
450   1753.

451 Scheffer M, Carpenter SR (2003) Catastrophic regime shifts in ecosystems: linking theory to observation.
452   Trends Ecol Evol 18:648–656.

453 Scheffer M, Hirota M, Holmgren M, Van Nes EH, Chapin FS (2012) Thresholds for boreal biome tran-
454   sitions. Proc Nat Acad Sci USA 109:21384–21389.

455 Scheffer M, Carpenter SR, Dakos V, Van Nes EH (2015) Generic indicators of ecological resilience:
456   inferring the chance of a critical transition. Ann Rev Ecol Evol Syst 46:145–167.

457 Shoemaker LG, Sullivan LL, Donohue I, Cabral J, Williams RJ, Mayfield M, Chase J, Chu C, Harpole
458   WS, Huth A, HilleRisLambers J, James ARM, Kraft NJB, May F, Muthukrishnan R, Satterlee S,
459   Taubert F, Wang X, Wiegand T, Yang Q and Abbott KC(2020) Integrating the underlying structure
460   of stochasticity into community ecology. Ecology 101: e02922.

461 Staver AC, Archibald S, Levin SA (2011) The global extent and determinants of savanna and forest as
462   alternative biome states. Science 334:230–232.

463 Strang A (2020) Applications of the Helmholtz-Hodge decomposition to networks and random processes.

464 Dissertation, Case Western Reserve University.

465 Urban NA, Matter SF (2018) Metapopulation mirages: problems parsing process from pattern. Ecol

466 Modelling 375:20–29.

467 Wainrib G (2013) A brief introduction to large deviations theory. In: Bachar M, Batzel J, Ditlevsen S

468 (eds) Stochastic Biomathematical Models. Springer, Berlin, pp 57–72.

469 Wiesenfeld K, Jaramillo F (1998) Minireview of stochastic resonance. Chaos 8:539–548.

Manuscript with changes in red
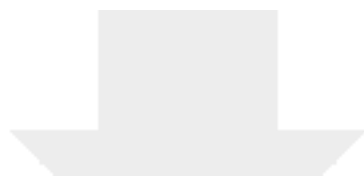
Click here to access/download
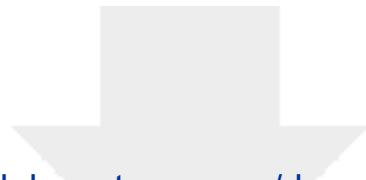Electronic Supplementary Material
main_MarkupInRed.pdf
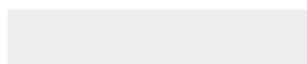
Click here to access/download
Electronic Supplementary Material
OnlineResource1.pdf

Online Resource 2

Click here to access/download
Electronic Supplementary Material
OnlineResource2.pdf
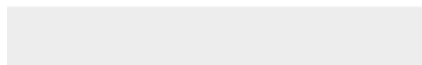
Click here to access/download
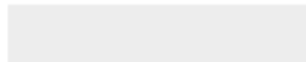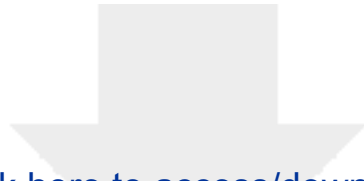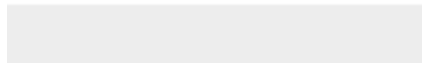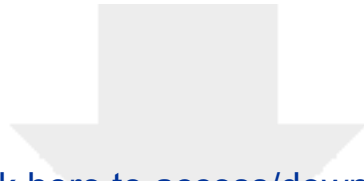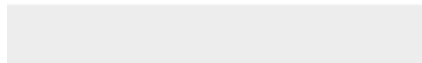**Electronic Supplementary Material**
OnlineResource3.pdf

Click here to access/download
Electronic Supplementary Material
OnlineResource4.pdf

Dear Kim and Alan,

Thank you for inviting us to respond to the reviewer's comments on our manuscript, "Mapping the distinct origins of bimodality in a classic model with alternative stable states." We followed nearly all recommendations and feel the paper has improved as a result. We are submitting the revised version of our manuscript both with and without changes marked in red, for your convenience. Our responses to all comments are in red below. Please let us know if you see the need for any further changes before final acceptance.

Sincerely,
Karen and Vasilis

Assoc Editor: I have nothing to add to these stellar reviews, accept an agreement with reviewer #2 about the use of vague pronoun references, and with reviewer #1 about the "noisiness" of fig 5

Thanks for emphasizing these points – please see below for our detailed responses.

Reviewer #1: I have carefully reviewed the manuscript, "Mapping the distinct origins of bimodality in a classic model with alternative stable states." This is a very nice treatment of an important issue that is well referenced and clearly written. Here, the authors illustrate how intermediate levels of stochasticity can give rise to bimodal distributions in state dynamics even when the underlying deterministic system has only a single stable state. The authors focus on the case of how certain phenomena which can create long transients can provide a mechanism for stochastic bimodality. While the authors acknowledge stochastic bimodality has been well documented under other mechanisms before, I think this is the first time it has been shown in these long transients. It is also a rather compelling simple example of what I think is still a generally under-appreciated point (that bimodality is not sufficient evidence to establish alternate stable states.)

I have a few minor suggestions below which I think could improve the manuscript, but no substantive objections to the treatment of the issues here.

Eq 1 & 2 are shown as ordinary differential equations with stochastic terms. While this is quite common practice, I believe it to be somewhat sloppy notation -- stochastic differential equations (SDEs) are not ODEs, and the notation is rather ambiguous as to whether the equation refers to an Ito or Statonovich integral (I believe the authors have the Ito form in mind).

This is a good point and our first inclination was to simply replace Eqns 1 & 2 (now Eqns 2 & 3) with their SDE forms. However, the random variable epsilon doesn't appear directly in the SDEs and we feel epsilon is a useful quantity for biological interpretation. Therefore, we have chosen to provide both versions of the additive noise model: the SDE in Eqn 1 (specifying that we intend the Ito form on L 97) and the sloppy-but-friendly

ODE + random variable in Eqn 2. We discuss their relation in the new text on L 101-107 before moving ahead with the Eqn 2 form. We feel this approach removes the mathematical ambiguity caused by Eqn 2 (previously Eqn 1) while maintaining readability by ecologists who are more familiar with ODEs than SDEs.

pg 5 line 4: It would be nice if the authors summarized consistencies or any salient differences of the multiplicative noise model here. The authors have provided a reasonable justification for focusing on the additive noise model (though the corresponding ecological assumptions could be detailed a bit more precisely -- particularly as noise in immigration / emigration process might seem less intuitive than environmental or demographic noise -- though near an equilibrium the additive noise model can be a good approximation of the multiplicative one too.). However, they merely note that multiplicative noise is treated in the supplement without any comment as to what that supplement shows. The supplement itself also avoids much textual discussion comparing the results, but basically leaves it up to the reader to compare supplemental figures to those of the main text.

We have remedied this with a new paragraph at the end of the Results section (L 298-309).

pg 5, line 54: As there are quite a variety of methods for determining bimodality, some citation for the method chosen would be nice.

We took a rather ad hoc approach, after getting dissatisfying results from published methods such as the dip test (Hartigan and Hartigan 1985) and the critical window size test (Silverman 1981). We are therefore unaware of an appropriate citation to use here.

pg 10, line 23: Discussion: I quite like the discussion of the 'sweet spot' in noise size, which as the authors observe, has nice resonance with other stochastic theory while being a distinct example here. It may be too speculative, but it calls to mind the prior literature on chaotic dynamics with the hypothesis that ecological systems may often be at the 'edge of chaos.' Just tossing it out there in case the authors find it any use drawing the comparison.

This is a thought-provoking connection, and we're tempted to try it. However, we feel there is an important difference that precludes us from making this connection casually. The "edge of chaos" literature argues that selection has positioned ecological systems near the chaotic regime. We are not arguing here that ecological systems *are* typically at the sweet spot in noise size, just that *if* they're at the sweet spot, here's what we can expect to see. The question of whether nature is organized in such a way as to make it more likely that we're at this sweet spot is really interesting, but not something we feel equipped to responsibly speculate on at the moment.

Fig 4: I found the illustration of the break-down of the small noise assumption in the mean first passage time from the quasi-potential to actually be quite interesting. While it's clear that the approximation should fail at 'some point', it's not so common to see

that point illustrated so clearly, and I thought the textual discussion and the figure rather glossed over this. It may also be worth mentioning in the discussion that the large deviation theory or WKB approximation may provide a better estimate than the quasi-potential MFPT for this.

We appreciate the invitation to talk more deeply about quasi-potentials and related approaches, and gladly added a paragraph to the Discussion (L 319-335) to address this point.

Fig 5. I found this figure rather harder to read than all the other figures. I think the color scale and overly fine (noisy) resolution for the contours distracts a bit from the main point of the figure -- it might be worth revisiting this to see if there is a way to make the bimodal region in intermediate noise intensities 'pop out' a bit more.

We agree that this figure is busy but are unsure how to remedy that. To aid interpretation, we have added text to the figure caption to help readers see what we are asking them to take away from this figure.

Reviewer #2: What a wonderful paper! Clearly written, well explained, and presenting profound results from a very simple model. It is a privilege to provide this review, and I'm terribly sorry it has taken me so long to get the paper read. I only have very minor editorial comments, which the authors are free to ignore, as the paper really is just fine as it is. Should the authors be interested in following my suggested revisions however, I think that they imrpove the readability of the behaviour.

Here is a summary of my corrections:
1. Years ago, a co-author taught me that the word "this" should always be followed by the object to which it refers (point of grammar). I do find that following this rule does make writing clearer.

This advice is helpful. (<- edited from "This is helpful advice" – see, I'm learning!) Edits, including all those suggested in the marked up pdf provided by the reviewer, have been made throughout.

2. A few typos, commas that need to be moved, etc.

Corrected. We thank this reviewer for her/his attention to detail!

3. I was a little dissatisfied with "movement away from the vicinity of the stable steady state" being equated with "arriving at the unstable steady state" being shunted entirely to Online resources. It would be helpful to have a few words of explanation included in the main text.

Point taken. The phrase "in the direction of the saddle" has been added (L 209) to clarify that we indeed look for movement away from the stable state in the right direction

to signify a first passage. While remaining details are still left to the supplement, we suspect this phrase would have prevented the reviewer from feeling uneasy about whether we measured what we say we measured.

Thanks again for giving me the opportunity to review this paper, and I hope to see it in print soon!