Eluding catastrophic shifts

Paula Villa Martín, Juan A. Bonachela, Simon A. Levin, and Miguel A. Muñoz

Abstract

Transitions between regimes with radically different properties are ubiquitous in nature. Such transitions can occur either smoothly or in an abrupt and catastrophic fashion. Important examples of the latter can be found in ecology, climate sciences, and economics, to name a few, where regime shifts have catastrophic consequences that are mostly irreversible (e.g., desertification, coral reef collapses, and market crashes). Predicting and preventing these abrupt transitions remains a challenging and important task. Usually, simple deterministic equations are used to model and rationalize these complex situations. However, stochastic effects might have a profound effect. Here we use 1D and 2D spatially explicit models to show that intrinsic (demographic) stochasticity can alter deterministic predictions dramatically, especially in the presence of other realistic features such as limited mobility or spatial heterogeneity. In particular, these ingredients can alter the possibility of catastrophic shifts by giving rise to much smoother and easily reversible continuous ones. The ideas presented here can help further understand catastrophic shifts and contribute to the discussion about the possibility of preventing such shifts to minimize their disruptive ecological, economic, and societal consequences.

Catastrophic shifts are natural phenomena of paramount importance (1, 2). For instance, socioeconomic and socioecological systems (3, 4) as well as ecosystems such as lakes, savannas, or oceans (with their embedded fisheries or coral reefs) can experience, as a consequence of small changes in environmental conditions, sudden collapses after which recovery can be extremely difficult (5–9). Hereinafter, we refer to these transitions as “abrupt” or “sudden” because there is a radical change in the steady-state properties once the threshold is crossed (regardless of the velocity of environmental variations).

Abrupt catastrophic transitions from vegetation-covered states to desertic ones in semiarid regions constitute an illustrative example (10–17). The latter habitats are characterized by positive feedback loops between vegetation and water: the presence of water fosters plant growth that, in turn, fosters water accumulation in plant-covered regions by, for instance, reducing evaporation rates. The fate of the ecosystem is determined by the overall precipitation rate, with a green or a desert stable state for high and low rates, respectively. There is an intermediate bistable precipitation regime compatible with either a barren or a vegetated landscape. Thus, as a response to some small environmental change, or if the feedback loop is incidentally interrupted, there can be a regime shift, implying a collapse of the vegetation cover and the mostly irreversible emergence of a deserted landscape (1, 10–12, 18).

This mechanism, in which population growth (or, more generally, activity generation) is reinforced by a positive feedback, constitutes the basic ingredient for multistability and for possible catastrophic regime shifts, also called tipping points or critical transitions. Similar facilitation mechanisms appear in a vast variety of examples in population ecology (Allee effect) (19); neuroscience (synaptic facilitation) (20); systems biology (21); and, in general, climate, biological, and social sciences (1, 22, 23).

Opposite to abrupt shifts, many other systems in nature and society exhibit much smoother transitions between active and quiescent states, with a more easily reversed progressive deterioration. Examples of the latter appear in epidemic spreading, fixation of alleles in population genetics, computer virus propagation, and autocatalytic chemical reactions, to name a few (24–28).

Predicting and anticipating catastrophic regime shifts and distinguishing them from their smoother counterparts constitutes a timely subject with a vast number of important applications, including the prevention of biodiversity collapse or radical climate changes as the result of anthropogenic pressures (21, 29–32). Indeed, early warning indicators of regime shifts (such as increasingly slower return rates from perturbation and rising variance) have been proposed, and some of them have been empirically tested (9, 14, 30, 31), even if their degree of robustness and reliability is still under debate (33, 34). Most of these approaches rely on understanding derived from simple deterministic equations in which spatial dependence is averaged out (21, 29–32). Recently, spatially explicit versions of these deterministic approaches have also been considered in the literature (35). In particular, ingredients such as spatial heterogeneity and mobility (e.g., diffusion) have been incorporated into those models, leading to interesting consequences and a much richer phenomenology that includes patchiness and pattern formation (31, 36–44). Moreover, it has been suggested that emerging spatiotemporal patterns could be potentially used as early indicators of tipping points or that transitions in these improved models can become more gradual (15–17, 45–47).

Common to most previous studies is the fact that stochastic effects such as demographic or intrinsic noise—an unavoidable feature of real systems—are typically left out of the picture (see, however, e.g., refs. 10, 37, 47, and 48). Stochasticity or noise is known to play a fundamental role in complex problems with many degrees of freedom, inducing nontrivial effects such as

Significance

Catastrophic shifts such as desertification processes, massive extinctions, or stock market collapses are ubiquitous threats in nature and society. In these events, there is a shift from one steady state to a radically different one, from which recovery is exceedingly difficult. Thus, there is a huge interest in predicting and eventually preventing catastrophic shifts. Here we explore the influence of key mechanisms such as demographic fluctuations, heterogeneity, and diffusion, which appear generically in real circumstances. The mechanisms we study could ideally be exploited to smooth abrupt shifts and to make transitions progressive and easier to revert. Thus, our findings could be of potential importance for ecosystem management and biodiversity conservation.

Author contributions: P.V.M., J.A.B., S.A.L., and M.A.M. designed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper. The authors declare no conflict of interest.

This article is a PNAS Direct Submission. G.S. is a guest editor invited by the Editorial Board.

1 Present address: Marine Population Modelling Group, Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, Scotland, United Kingdom.

2 To whom correspondence should be addressed. Email: mamunoz@onsager.ugr.es.

E1828–E1836 | PNAS | Published online March 30, 2015

www.pnas.org/cgi/doi/10.1073/pnas.1414708112
as noise-induced transitions, stochastic resonance, and stochastic amplification of fluctuations (49, 50). In particular, demographic noise can have dramatic effects in spatially explicit low-dimensional systems, profoundly influencing the features of the expected transitions (51). Thus, although external or environmental noise could also play an important role (52, 53), here we focus on analyzing explicitly the effects that demographic or internal noise may have on abrupt shifts.

Spatially explicit stochastic differential equations (such as Langevin equations) constitute the most appropriate mathematical formalism to study the role of stochasticity on phase transitions (54). Alternatively, individual-based models could be also used (see, e.g., refs. 10 and 37 and refs. therein), but Langevin equations can be usually derived from these individual-based approaches by using standard techniques (55). Furthermore, Langevin equations provide a much more generic and robust framework, highlighting universal features and therefore transcending the specificities of particular systems (54). Thus, beyond the theory of dynamical systems, the language and tools of statistical mechanics prove to be best suited for shedding light upon stochastic problems with many degrees of freedom. Within this framework, simple bifurcations exhibited by deterministic (or mean field (51)) systems are just fingerprints of true phase transitions. Catastrophic shifts stand for first-order or discontinuous (i.e., abrupt) transitions showing bistability and hysteresis (56), whereas smooth bifurcations correspond to continuous or second-order phase transitions, in which the system reaches scale-invariant (fractal) organization with diverging characteristic lengths and times and other remarkable and distinct features (51).

In this paper, we study spatially explicit stochastic systems that are susceptible a priori to exhibiting catastrophic transitions, i.e., systems for which deterministic approaches would predict alternative stable states and abrupt transitions between them. We focus on the more ecologically relevant 2D case but briefly discuss also the cases of one and three spatial dimensions. Our goal is to scrutinize how intrinsic stochasticity influences these systems, in particular, to explore whether stochasticity in combination with other realistic mechanisms such as limited diffusion and spatial (quenched) heterogeneity may alter the nature of phase transitions. By using a combination of computational and analytical techniques, together with known results from the statistical mechanical/bifurcation phase transitions, we show that these realistic ingredients can potentially round up abrupt discontinuities, giving rise to more predictable, progressive, and easy-to-reverse transitions. As a possible application of our results, we speculate that basic and widespread aspects of natural systems could be potentially exploited to prevent abrupt regime shifts and their undesired consequences.

Results

Model Building. Mathematically, smooth regime shifts into quiescent states are usually described and understood in terms of simple deterministic equations such as the logistic equation, \( \frac{\partial \rho(t)}{\partial t} = a \rho(t) - b \rho^2(t) \). In the latter, \( \partial \) stands for time derivative, \( \rho \geq 0 \) is the relevant variable (e.g., population density) that we call henceforth “activity,” \( a \) is the growth rate, and \( b > 0 \) fixes the maximal activity density (e.g., carrying capacity) (24). This equation describes a smooth (transcritical) transition between an active and a quiescent state as \( a \) is varied beyond a critical value (Fig. 1A and C). This equation can be easily modified to include a generic facilitation term, representing the positive feedback mechanisms discussed in the previous section. In its simplest form, facilitation alters linearly the growth rate \( a \), enhancing it in the presence of activity: \( a \rightarrow a + \alpha \rho \), with \( \alpha > 0 \). This variation generates an effective quadratic term \(-\alpha \rho^2\), which is fully equivalent to leaving the growth term unaltered and replacing \( b \rightarrow b + \alpha \) in the logistic equation. Thus, the coefficient of the resulting quadratic term—which, for simplicity, we continue to call \( b \)—can potentially change its sign, and the new equation takes the form

\[
\frac{\partial \rho(t)}{\partial t} = a \rho(t) - b \rho^2(t) - c \rho^3(t), \quad b < 0. \quad [1]
\]

A new higher-order (cubic) term has been added to enforce a finite carrying capacity (i.e., to prevent the population density to diverge when \( b < 0 \)). Eq. 1 is the simplest equation used to study catastrophic shifts at a deterministic level (1, 2, 22–24) (Fig. 1). This equation admits two alternative stable solutions (bistability) as well as an abrupt, discontinuous (i.e., fold) bifurcation (Fig. 1B and D) (24, 25, 57). As the figure illustrates, the right-hand side (rhs) of Eq. 1 can be interpreted as the gradient of a potential (58), and the sign of the parameter \( b \) controls the nature of the transition at the deterministic level: continuous for \( b > 0 \) or abrupt for \( b < 0 \). This observation is essential for the discussion that follows.

Aiming at capturing the relevant phenomenology of catastrophic shifts in a parsimonious yet complete way, we extend the equation above to include explicit spatial dependence and demographic or intrinsic stochasticity,

\[
\frac{\partial \rho(x, t)}{\partial t} = a \rho(x, t) - b \rho^2(x, t) - c \rho^3(x, t) + D \nabla^2 \rho(x, t) + \eta(x, t), \quad b < 0, \quad [2]
\]

where \( \rho(x, t) \) quantifies the activity at position \( x \) and time \( t \). This equation, similar to the one used to describe the strong Allee effect, is

\[\text{Fig. 1. Bifurcation diagrams and deterministic potential } V(\rho) \text{ for continuous (A and C) and abrupt (B and D) transitions. (A and B) Lines represent the steady-state solutions of } \rho(t) = a \rho(t) - b \rho^2(t) - c \rho^3(t) \text{ as } a \text{ is varied, with } b \text{ either positive (A) or negative (B), displaying continuous and abrupt transitions, respectively. Four particular values of } a \text{ are highlighted with spheres (blue for quiescent states and increasingly more reddish for active ones). (C and D) Effective potential } V(\rho) = \left( \frac{\xi}{2} - \frac{b}{6} \rho^3 - \frac{1}{4} \rho^4 \right), \text{ from which the deterministic forces above are derived, plotted as a function of the activity, } \rho, \text{ for the same values of a highlighted above. In C, } b > 0, \text{ and the transition is smooth and continuous (transcritical bifurcation), whereas in D, } b < 0, \text{ and there is an abrupt jump in the location of the potential absolute minimum, } a^* = -b^2/4c, \text{ as the control parameter } a \text{ is varied corresponding to a discontinuous transition (fold bifurcation). In summary, the sign of } b \text{ in Eq. 1 controls the nature of the transition at the deterministic or fluctuation-less level.} \]
effect (19), consists of three different contributions: (i) a local deterministic force (with $b<0$), which coincides at each location $x$ with the rhs in Eq. 1; (ii) diffusion, represented by the Laplacian term $D \nabla^2 \rho(x,t)$, with proportionality constant $D>0$ (this term accounts for dispersal of activity to neighboring locations); and (iii) demographic stochasticity, encoded in the Gaussian (white) noise $\eta(x,t)$ with zero mean and variance proportional to $\sigma^2 \rho(x,t)$ (this multiplicative noise ensures that demographic fluctuations do not exist in the bulk of regions deprived of activity).

Eq. 2, with $b<0$, could a priori—i.e., thinking in a deterministic or mean field way (51)—be expected to capture the behavior of catastrophic shifts in spatially extended systems. However, as a result of the presence of noise, its emerging phenomenology might not be straightforwardly inferable from the mean field reasoning. Therefore, we first need to provide an overview of the actual properties of systems described by Eq. 2 with $b<0$ under standard circumstances. Afterward, we shall scrutinize how these results might be altered once other ingredients such as large demographic noise, limited diffusion, and/or environmental disorder are introduced. Here we resort to extensive computational analyses as well as renormalization group calculations to discuss aspects of Eq. 2 that are relevant to our discussion. For other (fundamental) aspects of this equation, such as existence and uniqueness of solutions as well as more formal analytical approaches including small-noise calculations, we refer the reader to the existing vast mathematical literature (59–67).

**Computational Results.** Integrating numerically Eq. 2 is not a trivial task owing to the presence of multiplicative noise. However, as described in Methods, there exists to this end an exact integration scheme (68) that has already been successfully used to study spatially explicit problems in ecology (69). As a note of caution, let us remark that determining numerically the nature of a phase transition in an extended system can be a difficult enterprise; the literature is plagued with claims of discontinuous transitions in systems with quiescent states (70) that eventually were proven to be continuous ones once sufficiently large sizes and times and enough statistics were collected (26). Thus, to avoid any ambiguity in our conclusions, we performed very extensive large-scale computer simulations and different types of numerical experiments. We considered discrete square lattices of size up to $1,024 \times 1,024$ with either periodic or open boundary conditions, averaged over up to $10^6$ realizations, for each of the different types of computational experiments we performed (Methods): (i) decay experiments from an initial homogeneously active state, (ii) spreading experiments from an initially localized seed of activity in an otherwise quiescent state, and (iii) interfacial experiments in which the evolution of an initially half-empty/half-occupied lattice—with a planar interface in between—is investigated.

**Catastrophic shifts can appear in 2D noisy systems.** A summary of the main features shown by a 2D system described by Eq. 2 with typical parameter values is presented in Fig. 2. In particular, Fig. 2A shows the averaged activity in the steady states as a function of the control parameter $a$ revealing the existence of two alternative homogeneous stable solutions: an active one with $p>0$ and a quiescent one with $p=0$, with an intermediate regime of bistability and all of the characteristic signs of a discontinuous transition. In particular, as illustrated in the inset of Fig. 2A, which steady state is reached may depend upon initial conditions revealing the existence of bistability and hysteresis, i.e., trademarks of first-order phase transitions.

Regarding probabilistic aspects, which are essential here, let us remark that small systems—even in the active phase—may fall into the quiescent state owing to rare demographic fluctuations; however, the averaged time for this to happen grows exponentially with system size in the active phase (55), and it is much larger than computational times for the sizes we have considered. On the other hand, different routes to extinction—all of them of stochastic nature—exist in the regime of bistability (these have been studied computationally by some of us in ref. 71 and analytically in ref. 66 for a discrete version of our model) confirming the discontinuous nature of the transition.

As a visual illustration, Fig. 2B and C shows examples of how an island of one of the phases may propagate, invading a sea of the other, when the latter phase is close to the threshold of instability. We have found no evidence of the existence of an intermediate critical behavior with a nontrivial power law exponent—as would correspond to a continuous transition—in any of the numerical (spreading or decay) experiments we have performed.

Finally, we have also conducted interfacial experiments to determine the relative stability of both phases. If the system described by Eq. 2 shows a catastrophic shift, there should be a value of $a$ for which the interface separating two perfect halves of the system—each half in one of the two coexisting states—does not move on average; this is the so-called Maxwell point (39). As shown in Fig. 2D, there indeed exists a Maxwell point for the system described by Eq. 2.

Hence, all this evidence allows us to safely conclude that Eq. 2 experiences a true discontinuous, first-order phase transition in two dimensions, in agreement with deterministic expectations. This conclusion is quite robust against changes in parameter values, but as detailed in what follows, it may eventually break down in the presence of certain additional mechanisms, giving rise to very different scenarios.

**Factors preventing catastrophic shifts in 2D noisy systems.** The role of enhanced (demographic) noise. The noise amplitude, $\sigma^2$, is a measure of the level of demographic stochasticity present in the system. This factor is, thus, a straightforward indicator of how far the actual stochastic system is from its deterministic counterpart. To explore the consequences of high stochasticity, we have carried out the same type of computational experiments described above but now enhancing the noise amplitude (from $\sigma^2=1$ above to $\sigma^2=4$) while keeping fixed the remaining parameters. As illustrated in Fig. 3A, the situation is very different from the one in the previous section. In particular, now there is a continuous phase transition at a specific value of the control parameter, $a_c$. At this point, the relevant quantities in spreading and decay experiments exhibit power law (scale-free) behavior, characteristic of continuous transitions (see inset). Moreover, we have detected critical depinning-like behavior that coincides with the expected values in standard continuous transitions into quiescent states, i.e., those characteristic of the well-known directed percolation (DP) universality class (Methods (26–28, 72)). For completeness, we have also verified that there is no bistability, and as a consequence, a Maxwell point cannot possibly be identified. Further computational experiments confirmed that for any parametrization, it is possible to find a threshold for the noise amplitude that alters the character of the transition, therefore leading to the same conclusions. This provides strong evidence that Eq. 2 exhibits a continuous (more specifically, a DP) phase transition if demographic stochasticity is large enough. This conclusion is in contrast with the deterministic expectation and thus manifests that strong demographic fluctuations play an essential role in these low-dimensional spatially explicit systems.

**The role of limited diffusion.** Similarly, we have scrutinized the effect of reducing the diffusion constant, $D$, in Eq. 2, while keeping fixed all other parameters. Limited diffusion is very widespread in ecological systems. For example, under certain conditions, plants may restrict their range of seed dispersal as an evolutionary strategy to enhance survival (73). When limited diffusion occurs, one could expect spatial effects to be enhanced, and thus, departures from mean field behavior are more likely to occur; at the opposite extreme of very large diffusion (e.g., long-ranged seed dispersal), results are expected to be much closer to the deterministic limit. Indeed, for small $D$ (e.g., $D=0.1$ in Fig. 3B),
computational evidence reveals the existence of a continuous transition in the DP class, in contrast to the catastrophic shift reported above for $D=1$.

As an explicit illustration of the probabilistic nature of the discussed phenomena, the inset of Fig. 3 includes results for the surviving probability as a function of time $P_s(t)$ in spreading

**Fig. 2.** Computational results for Eq. 2 in a 2D lattice, showing an abrupt regime shift. Parameter values are $D=c=1$, $\sigma^2 = 1$, and $b=-2$. (A) Steady-state averaged activity as a function of the tunable parameter $a$ (different colors correspond to different type of initial conditions: light green for small initial activities and dark green for large initial activities). There are two distinct stable solutions: one with an associated large stationary activity (which disappears around $a_- \approx -0.98$) and a trivial or quiescent one with $P=0$ (which becomes unstable at some point near $a_- \approx -0.95$). These two alternative steady states are separated by a line of unstable solutions (dashed gray). In the interval between the two limits of stability, two alternative stable states compete. The existence of bistability is confirmed by results in the inset, showing that the steady state depends upon initial conditions ($a_- \approx -0.95$). B illustrates how an initially localized seed of activity expands throughout the system (spreading experiment) near the threshold of instability of the quiescent phase. Similarly, C illustrates the decay of an initially homogeneous state toward the quiescent state (decay experiment). (D) In interfacial experiments, half of the system is initially occupied and half is empty (Left); either the active phase invades the quiescent one (Center) or vice versa (Right). These two regimes are separated by a Maxwell point at which the interface does not move on average. In B-D, arrows indicate the direction of system’s advance with time.

**Fig. 3.** Realistic ingredients can alter the nature of potentially catastrophic shifts in 2D environments. Using the same parametrization as in Fig. 2, we study separately the effects induced in Eq. 2 by (A) enhanced demographic noise ($\sigma^2 = 4$, $D=1$), (B) limited diffusion ($D=0.1$, $\sigma^2 = 1$), and (C) spatial heterogeneity ($b \rightarrow b(x)$). In all three cases, the transition becomes continuous, with no sign of bistability nor discontinuous jumps. In A and B, power law behavior is observed (see insets) for all of the computed time-dependent (decay and/or spreading) quantities right at the critical point ($a_- \approx -0.708$ in A and $a_- \approx -0.5236$ in B). Moreover, the corresponding exponent values (both of them close to 0.45) coincide with the expected values for the directed percolation class (Methods). Curves in the insets correspond to values of $a$ between $-0.702$ and $-0.718$ in A and between $-0.5234$ and $-0.5238$ in B, both in equal intervals. In C, alongside the continuous transition, there appears a broad region within the quiescent phase, called a Griffiths phase, characterized by an extremely slow decay of activity, i.e., by power laws with continuously varying exponents, as illustrated in the inset (values of $a$ between $-0.250$ and $-0.300$ in equal intervals).
experiments, showing a smooth continuous change in its asymptotic behavior and a DP-like power law decay at criticality. Similarly to the enhanced stochasticity case, there always can be found a threshold for \( D \) below which Eq. 2 exhibits a continuous transition (Methods).

**The role of (quenched) spatial heterogeneity.** A relevant ingredient which is unavoidably present in real systems is spatial heterogeneity. Here we focus on cases where disorder does not change with time (i.e., quenched disorder). Local differences in environmental conditions can generate regions that are more prompt to collapse and others that are more resilient, giving rise to patchy and irregular activity patterns. A general result in statistical mechanics proves that discontinuous transitions cannot possibly occur in 2D disordered systems at thermodynamic equilibrium (74–76). This conclusion has been recently extended to more general systems, not necessarily at equilibrium, including transitions into quiescent states if stochasticity is present (71).

On the basis of this result, one could expect abrupt transitions to be smoothed once both heterogeneity and noise are considered. To study explicitly the consequences of heterogeneity in a system described by Eq. 2, we assume \( b \) to be position-dependent, i.e., \( b \rightarrow b(x) \). The value of \( b(x) \) at each location \( x \) is randomly extracted from a uniform distribution in the interval \([-2, 0)\), essentially ensuring a different \( b < 0 \) at each location. Results of extensive computer simulations summarized in Fig. 3C, show that any amount of spatial heterogeneity induces a smooth transition. In this transition, the collapse from the active phase to the quiescent one occurs in a rather gradual way, with a progressive deterioration of the less favorable regions. Furthermore—and differently for the previous cases—spatial disorder induces a broad region around the transition point in which power law scaling is generically observed. In particular, the averaged activity decays in a very slow (power law) fashion as a function of time toward the quiescent state, not just at the critical point (as usually happens) but rather for a whole range of values of the control parameter \( a \). This region, with generic scale-free behavior, is usually dubbed “Griffiths phase” and stems from the fact that unfavorable zones are emptied first and then, progressively, more and more resilient zones collapse in a step-by-step fashion (77).

Therefore, we have confirmed that spatial random heterogeneity is a sufficient ingredient to destroy abrupt regime shifts in 2D stochastic systems, giving rise to smooth transitions, in agreement with ref. 71.

**The role of spatial dimensionality.** All of the results above have been obtained for 2D systems. However, some of the reported noise-induced effects might depend profoundly on the system dimensionality. Thus, we now discuss the 1D and 3D cases.

For 1D systems, fluctuation effects are expected to be extremely severe (51). Indeed, existing analytical arguments predict that stochasticity completely washes away discontinuous transitions into absorbing states, converting them into continuous ones (26). Thus, catastrophic shifts into quiescent states cannot possibly occur in 1D systems (26, 48). We have verified computationally this prediction: our simulations show clearly a continuous phase transition in all 1D cases.

In 3D systems, the combined effect of amplified demographic noise and limited diffusion still affects the nature of the transition, even if to a lesser extent (as shown by our analytical calculations; see below). On the other hand, and contrarily to the 2D case, spatial random heterogeneity combined with demographic stochasticity does not suffice to destroy abrupt regime shifts in 3D systems: alternative stable states and abrupt shifts can survive the introduction of spatial disorder (71). In consequence, catastrophic shifts can occur more easily in 3D systems than in their 2D counterparts.

In summary, the smaller the spatial dimension, the more likely fluctuations play a fundamental role, potentially breaking deterministic predictions, preventing catastrophic shifts and generating much more gradual and smooth transitions.

**Analytical Results.** In addition to the strong numerical evidence presented so far, we now provide analytical understanding on why the transition may become continuous under the above-discussed circumstances, in particular, for low diffusion as well as for the large noise case. To this end, we rely again on statistical mechanics and use renormalization group theory (Methods) (26, 51).

In fluctuating spatially extended systems, crucial information about large-scale features—including the nature of possible phase transitions—cannot be derived from the associated deterministic potential (Fig. 1). The reason is that the true (or renormalized) effective potential includes fluctuation effects, which are lacking in such deterministic or bare potential (51, 78). Therefore, to rationalize the previous numerical conclusions, we need to think in terms of the (true) renormalized potential, \( V_R(\rho) \). In particular, fluctuations have the net effect of shifting the effective parameter values characterizing the potential, from their original deterministic or bare values to their renormalized or dressed variants.

Renormalization group techniques were devised to compute analytically \( V_R(\rho) \) as the scale of description is enlarged (26, 78). To illustrate how this works, we have computationally measured the probability distribution for the local activity, \( \text{Prob}(\rho_m) \), in the stationary steady state, where \( \rho_m \) is the activity averaged in square boxes of progressively larger linear size, \( m \) (i.e., at coarser and coarser scales). In this way, it is possible to measure the renormalized coarse-grained effective potential as \(-\log[\text{Prob}(\rho_m)] \) (51, 78). The most likely value of the activity at each coarse-grained scale lies at the minimum of the corresponding potential.

As an example, results for the limited diffusion case (\( D = 0.1, \sigma^2 = 1 \)) are shown in Fig. 4 for different values of the control parameter \( a \). For fine-grain scales such as \( m = 1 \), the effective potential is expected to coincide with the deterministic one (51). Indeed, it exhibits a discontinuous transition as its global minimum jumps abruptly from 0 to a nonvanishing value in a discontinuous way. However, as the level of coarse-graining is increased, a dramatic change of behavior is observed. For instance, for boxes of size \( m = 64 \), it can be already seen that the effective potential experiences a continuous transition from 0 to arbitrarily small activity values. This illustrates the change in the nature of the phase transition—occurring for small diffusion constants or for large demographic noise amplitudes—once fluctuations and spatial effects are taken into account.

These results can be understood using renormalization group ideas. Indeed, as explained in detail in Methods, we have also used standard procedures to perform an analytical renormalization group calculation [à la Wilson (78)]. This allows us to compute (up to first order in a perturbative expansion) how every parameter appearing in the potential changes or flows upon coarse-graining. In particular, Fig. 5 clearly illustrates how an originally negative bare value of \( b \) can become positive as the coarse-graining scale (represented here by the parameter \( J \)) is enlarged in a 2D system. The fact that \( b \) becomes positive indicates that the behavior at larger scales corresponds to a continuous DP transition.

In particular, this change in sign occurs—other parameters being fixed (e.g., \( \sigma^2 = 1 \))—for sufficiently small values of \( D \), in perfect agreement with our numerical observations, which reported a transmutation in the nature of the transition only in the low-diffusion limit. Similarly, keeping an intermediate diffusivity value \( D = 1 \), the same phenomenon occurs for sufficiently large noise amplitudes \( \sigma^2 \gg 1 \) (Methods).

This renormalization group calculation can also be used to illustrate that the system’s dimensionality plays an important role in these results. In particular, fluctuation-induced corrections vanish for spatial dimensions larger than \( d = 4 \) (where deterministic results are expected to hold) and are much more severe for low dimensions (e.g., \( d = 1 \) and \( d = 2 \)). Thus, discontinuous transitions and catastrophic shifts are predicted to be
Conclusions and Discussion

Catastrophic shifts occur when, as a consequence environmental or external changes, a system crosses abruptly from one phase to a radically different competing one, from which recovery may be exceedingly difficult due to hysteresis effects. Such abrupt regime shifts can affect the system at multiple levels, entailing, for instance, important ecological and/or socioeconomic consequences. Classical examples include studies of insect outbreaks (79), shallow lakes, savannas, socioecological systems, and markets (1–4). Therefore, there is an increasing interest in finding early warning signals that may help to predict when one of these tipping points is about to occur. Most studies and predictions concerning catastrophic shifts and their early warning indicators have been based on deterministic equations where demographic stochasticiy—a natural and unavoidable ingredient of real systems—is left out of the picture. Statistical mechanics tells us that intrinsic noise can play a fundamental role in the behavior of complex systems with many degrees of freedom, generating nontrivial effects such as stochastic resonance and noise-induced transitions among many others (49). Therefore, it is important to go beyond deterministic approaches to develop robust and reliable predictors of the occurrence of catastrophic shifts.

In this work, we have introduced and analyzed the simplest possible stochastic theory of catastrophic shifts in spatially extended systems, namely, Eq. 2. A simplistic deterministic analysis of this equation would average out the noise and would lead, in general, to the prediction of alternative stable states and an abrupt transition between them (Eq. 1). Here we have studied instead the full stochastic model, including demographic noise, by using a combination of computational and analytical tools and have explored the effects of stochasticity. First, we have verified that for low or moderate levels of demographic noise, 2D systems may truly exhibit a bona fide first-order transition, with bistability and hysteresis. Thus, catastrophic shifts can actually appear in noisy systems. However, we also show that adding any of the following ingredients, (i) enhanced demographic variability, (ii) limited dispersal/diffusivity, and/or (iii) spatial (quenched) heterogeneity, suffices to alter the nature of the phase transition, giving rise to a second-order (continuous) one.

Most of our results have been obtained for 2D systems, which have obvious applications to ecological problems such as desertification processes or vegetation dynamics in savannas. However, some of the reported noise-induced effects depend profoundly on the system dimension. For example, in one spatial dimension, relevant for the study of, e.g., the oceanic water column, or rivers, fluctuation effects prohibit the very existence of discontinuous transitions, as has already been suggested in the literature (26). In three spatial dimensions the smoothing effects of limited diffusion and amplified demographic noise are still present even if to a lesser extent (Analytical Results), and contrarily to the 2D case, discontinuous transitions can survive to the introduction of spatial (quenched) heterogeneity. Thus, as a rule of thumb, the smaller the spatial dimension, the more likely fluctuations play an important role, potentially breaking deterministic predictions and smooth abrupt transitions. However, fluctuation effects need to be carefully analyzed in each spatial dimension to reach robust conclusions.

In this work we have put the focus onto demographic or intrinsic noise, but environmental or external sources of noise can also potentially be important factors (52, 53, 80). Preliminary studies suggest that this type of variability could also alter the order of phase transitions (53), but more detailed analyses of this, as well as of the interplay between demographic and environmental noise, would be highly desirable.

This study offers obvious opportunities for ecosystem management. All of the relevant features present in Eq. 2 and its variations have straightforward counterparts in natural systems. Thus, identifying these mechanisms in specific problems may provide a reliable indicator as to when a transition is expected and of whether it is expected to be abrupt or smooth. Moreover, the conclusions here can potentially help prevent catastrophic transitions from occurring by forcing their transformation into
continuous ones. Most of the current ecosystem management strategies focus on stopping or slowing down the ongoing change before the shift occurs. For instance, the declaration of a species as protected aims to prevent species extinction. In the system’s phase diagram, this is equivalent to preventing the control parameter $a$ (e.g., poaching pressure) from reaching its transition value. Unfortunately, this goal is not always possible to achieve, e.g., when the control parameter is linked to natural resource availability, climatological factors, or hardly unavoidable human activities. In these cases, our study offers alternative strategies with which the catastrophic effects of those shifts can be reduced (that is, the discontinuous shift can be transformed into a continuous one), allowing the transition to be more predictable and even eventually reverted. Continuous transitions show a single stable state changing progressively with environmental conditions; therefore, they are easier to handle, foresee, and undo.

Some examples of ecosystem engineering that could potentially take the system in that direction include introducing or enhancing spatial disorder (e.g., grazing, watering, or burning selected zones in the vegetation example), forcing a reduction of effective diffusion (e.g., preventing seed dispersal by herbivores), or artificially enhancing demographic variability. Similarly, these ideas may be potentially useful in the design of practical programs for ecosystem restoration and management policies to avoid the collapse of natural resources. For instance, using any of the mechanisms we present here to smooth an abrupt transition to extinction could potentially open the door to the existence of low-density states of the focal species, which were not possible in the discontinuous case. These low-density states could be ideally used as early warning indicators and therefore help prevent such extinctions. On the other hand, introducing these mechanisms may enlarge the absorbing phase (i.e., shift the transition point $a_c$ toward less negative values). Therefore, the system may become more vulnerable because the same pressure will drive the population extinct. Thus, the suitability of these mechanisms for ecosystem management depends on this important trade-off between predictability and vulnerability, which needs to be carefully evaluated.

In summary, we have proposed a general framework under which specific studies of potential catastrophic shifts should be set to obtain more reliable and informative predictions. Given the growing concerns about the effect of anthropogenic pressures on climate and biodiversity, we hope that this framework will help to understand better and open new research roads to explore possible strategies to mitigate the radical and harmful effects of sudden undesirable regime shifts.

Methods

Smooth Transitions into Quiescent States. Extinctions can be seen as transitions from an active to a quiescent or absorbing phase, in which essentially all dynamics and activity cease. Most of them—in opposition to catastrophic shifts—correspond to second-order/continuous phase transitions in which the system changes in a much smoother way, moving gradually from one state to the other in a rather universal, i.e., detail-independent, fashion (26, 27, 28, 81, 82). More specifically, all these systems organize at the transition critical point in a scale-invariant fractal way, with its concomitant power law distributions and scaling behavior (51).

The essential properties of this broad family of transitions into absorbing or quiescent states—customarily called DP class—are well captured by the Langevin equation (26, 27, 28, 81, 82):

$$\partial_t \rho(x,t) = a \rho(x,t) - b \rho(x,t)^3 + \nabla \cdot \left( D \nabla \rho(x,t) \right) + \Delta \rho(x,t), \quad b, D > 0, \quad \rho \geq 0,$$

which differs from Eq. 2 just in the sign of $b$ and the fact that the cubic term is, therefore, not needed (i.e., it is irrelevant). Observe that in the absence of noise and assuming spatial homogeneity, this equation reduces to the simple deterministic equation for continuous transitions discussed in Introduction. Eq. 3 has been studied both analytically (with renormalization group techniques (81, 82) and other methods (26)) and numerically (68), and the properties of the associated continuous transition have been well established. The resulting phase transition is continuous, as it is at a deterministic level, but important quantitative aspects do not coincide with the deterministic predictions; noise and spatial dimension are crucial aspects to describe them. For instance, at criticality, the activity decay $\rho(t) \sim t^{-0.46}$ and the evolution of the total activity in spreading experiments, $N(t) \sim t^{0.23}$, are power laws with universal exponents, defining the DP universality class in 2D systems (72).

Integration Scheme. Integrating stochastic (partial) differential equations with multiplicative noise, i.e., a noise depending on the system state, is a nontrivial task. The key problem is that standard integration techniques do lead to negative values of the activity variables $\rho$ and, thus, to numerical instabilities. Nevertheless, an efficient and accurate numerical integration can be obtained through a split-step scheme introduced a few years ago (68). The scheme consists of separating, at each spatial site, the integration of some of the deterministic terms from that of the stochastic part (plus eventually linear and common deterministic terms), e.g., $d\rho = d\rho_{\text{det}} + d\rho_{\text{stoch}}$, where the Gaussian (white) noise $\eta(x,t)$ has zero mean and variance proportional to $\sigma^2 \rho(x,t)$. The scheme allows us to exactly sample, at each spatial location, the time-dependent solution of the associated Fokker-Planck equation (55):

$$P(\rho(t+\Delta t)) = \frac{\partial}{\partial t} \rho(x,t) = 2\Delta \rho(x,t)^{\nu/2} I_0(4\Delta \rho(x,t)^{\nu/2} - 1),$$

where $I_0$ is a Bessel function of order $\nu = 2a/\sqrt{a^2 - 1}$, and $\nu = 1$. The most important step is to realize that this equation can be rewritten, with the help of a Taylor series expansion, as

$$P(\rho(t+\Delta t)) = \sum_{n=0}^{\infty} \left( \frac{j_n(4\Delta \rho(x,t)^{\nu/2})^{\nu/2}}{n!} \right) \frac{\partial^n}{\partial \rho^n} \rho(x,t)^{\nu/2},$$

and noticing that $\rho(t+\Delta t)$ can be obtained by a mixture of gamma and Poisson probability distributions that will reconstitute, on average, all terms of the latter equation. Thus, the method alternates two steps: (i) the integration of deterministic terms using some standard algorithm (such as an Euler or Runge-Kutta) and then (ii) using its output, Eq. 5 is used to obtain the final updated value of $\rho(t+\Delta t)$ at each spatial location. More details and applications of this numerical scheme can be found in ref. 68.

Computational Experiments. Different types of computational experiments have been performed to ascertain the nature of phase transitions: (i) in decay experiments the system is initialized with a homogeneously active state, $\rho(x,t=0)=1$, and the evolution of $\rho(t)$ is monitored, averaging over all sites and over many different realizations. (ii) In spreading experiments, we follow the dynamics of an initially localized seed of 100 active sites forming a 10 x 10 square box at the center of an otherwise empty lattice, measuring how the averaged total (integrated) activity changes as a function of time. (iii) In interfacial experiments, an initially half-empty/half-occupied lattice is considered, and the dynamics of the interface separating the two halves is analyzed (Fig. 2). In first-order transitions this interface moves on average in one direction or the other depending on the value of the control parameter and remains stable right at the Maxwell point. Instead, in second-order
continuous transitions the interface is quickly erased rather than moving as a whole.

Renormalization Group Analysis. Renormalization group techniques (51, 78) have been applied to equations such as Eqs. 3 and 2. In particular, Eq. 3—sometimes called Reggeon field theory or Gribov process—captures the relevant features of continuous transitions into absorbing or quiescent phases, defining the so-called DP class (26, 81, 82). For a clear and concise presentation of how renormalization group techniques can be applied to Eq. 3 we refer the reader to ref. 27. The calculation consists of a perturbative expansion around the critical dimension, $d = 4$, above which standard deterministic (mean field) results hold.

Here we just follow the calculation in ref. 27 and briefly describe the modifications required to deal with Eq. 2 rather than with Eq. 3. All of the basic ingredients of the perturbative theory remain unchanged (Fig. 6) (27), but the sign of $b$ needs to be inverted, and an additional cubic nonlinearity, $-c_{3}$ (which has an associated new vertex as shown in Fig. 6) needs to be included.

Naive dimensional analysis tells us that the new cubic term is irrelevant around four spatial dimensions (at which the perturbative expansion is performed); however, if $b$ is negative, then $c$ is needed to stabilize the theory, and thus, it is a so-called dangerously irrelevant operator, which needs to be explicitly taken into consideration to obtain stable results.

The renormalization procedure consists of first rescaling coordinates and fields, $x \rightarrow \Lambda x$, $t \rightarrow t', \Lambda$, and $\rho \rightarrow \rho / \Lambda^2$ [where $\Lambda$ is an infinitesimal dilatation in momentum space, which can be expressed as $\Lambda = \exp[\lambda t]$], and then eliminating short-range fluctuations, i.e., integrating out the moments in the diagram shown in Fig. 6) and yields $3^{d-c_{3}}/\Omega = (2\pi)^{d+2}/(12\cdot 4^{d+2})$, where $\Omega$ is the surface of a $d$-dimensional hypersphere. Incorporating this additional correction to the standard DP renormalization group flow equations and fixing the spatial dimension to $d = 2$, we obtain the flow diagram shown in Fig. 5. Starting with a negative value of $b$, the flow keeps it negative for large values of $D$, whereas for small $D$ values the renormalized value crosses the line $b = 0$, thus becoming positive and remaining so. As soon as $b$ becomes positive, the standard DP theory is recovered, the term $c$ becomes irrelevant, and therefore, it starts flowing to 0 at larger coarse-graining scales. As $c$ approaches 0, the value of $b$ in the renormalization group becomes identical to that of the standard directed percolation class, and in particular, it reaches the DP line.

Similarly, keeping the same value of $D$, the same phenomenon can be observed by increasing the (demographic) noise variance $\sigma^2$. Let us remark that a tricritical point, at which the renormalized $b$ vanishes, should also appear at some value of $D$ (located at $D = 0.9$ in Fig. 5). This point separates continuous from discontinuous transitions and can be also investigated in detail using renormalization group techniques (83). Similar but milder results are obtained in three spatial dimensions, $d = 3$. Indeed, observe that as $S_{2}/(2\pi)$ increases with decreasing dimensionality, the effect becomes more pronounced for low-dimensional systems: the lower the dimension, the larger the value of $D$ at which the transition changes nature.

In summary, a renormalization group calculation enables us to determine analytically that for low diffusion constants and/or for large noise amplitudes, a change in the nature of the phase transition is to be expected from a first-order behavior at a mean field (fluctuation-less) level to a second-order one once fluctuations have been taken into consideration, especially in low-dimensional systems.

ACKNOWLEDGMENTS. We thank Paolo Moretti and Jorge Hidalgo for a careful reading of the manuscript and for useful comments. We acknowledge support from the J. de Andalucía Project of Excellence P09-FQM-4682, the Spanish Ministerio de Economía y Competitividad Project FIS2013-43201-P, scholarship PU2012/05750, the Army Research Office Grant W911NG-11-1-0385, the National Science Foundation Grant OCE-1046001, the Andrew W. Mellon Foundation (Dynamics of South African Vegetation), and Foundation Questions in Evolutionary Biology (FQEB) Grant (FQEB-RFP-12-14) from the John Templeton Foundation.