Predictability of extreme events in a branching diffusion model

Andrei Gabrielov
Departments of Mathematics and Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN, 47907-1395[*]

Vladimir Keilis-Borok
Institute of Geophysics and Planetary Physics, and Department of Earth and Space Sciences, University of California Los Angeles, 3845 Slichter Hall, Los Angeles, CA 90095-1567[†]

Ilya Zaliapin
Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557-0084, Corresponding author.[‡]

(Dated: August 10, 2007)

We propose a framework for studying predictability of extreme events in complex systems. Major conceptual elements — *direct cascading* or *fragmentation*, *spatial dynamics*, and *external driving* — are combined in a classical age-dependent multi-type branching diffusion process with immigration. A complete analytic description of the size- and space-dependent distributions of particles is derived. We then formulate an extreme event prediction problem and determine characteristic patterns of the system behavior as an extreme event approaches. In particlular, our results imply specific premonitory deviations from self-similarity, which have been heuristically observed in real-world and modeled complex systems. Our results suggest a simple universal mechanism of such premonitory patterns and natural framework for their analytic study.

PACS numbers: 89.75.Hc, 89.75.-k, 91.30.pd, 02.50.-r, 91.62.Ty, 64.60.Ht

INTRODUCTION

Extreme events (also called critical transitions, disasters, catastrophes and crises) are a most important yet least understood feature of many natural and human-made processes. Among examples are destructive earthquakes, El-Niños, economic depressions, stock-market crashes, and major terrorist acts. Extreme events are relatively rare, and at the same time they inflict a lion's share of the damage to population, economy, and environment. Accordingly, studying the extreme events is pivotal both for fundamental predictive understanding of complex systems and for disaster preparedness (see [1, 2] and references therein).

In this paper we work within a framework that emphasizes mechanisms underlying formation of extreme events. Prominent among such mechanisms is *direct cascading* or *fragmentation*. Among other applications, this mechanism is at the heart of the study of 3D turbulence [3]. A statistical model of direct cascade is conveniently given by the branching processes; they describe populations in which each individual can produce descendants (offsprings) according to some probability distribution. A branching process may incorporate *spatial* dynamics, several types of particles (multi-type processes), age-dependence (random lifetimes of particles), and immigration due to external driving forces [4].

In many real-world systems, observations are only possible within a specific domain of the phase space of a system. Accordingly, we consider here a system with an *unobservable* source of external driving ultimately responsible for extreme events. We assume that observations can only be made on a *subspace* of the phase space. The direct cascade (branching) within a system starts with injection of the largest particles into the source. These particles are divided into smaller and smaller ones, while spreading away from the source and eventually reaching the subspace of observations. An important observer's goal is to locate the external driving source. The distance between the observation subspace and the source thus becomes a natural control parameter. An extreme event in this system can be defined as emergence of a large particle in the observation subspace. Clearly, as the source approaches the subspace of observation, the total number of observed particles increases, the bigger particles become relatively more frequent, and the probability of an extreme event increases. In this paper, we give a complete quantitative description of this phenomenon for an age-dependent multi-type branching diffusion process with immigration in $\mathbb{R}^n$.

It turns out that our model closely reproduces the major premonitory patterns of extreme events observed in hierarchical complex systems. Extreme events in such systems are preceded by transformation of size distribution in the permanent background activity (see *e.g.*, [1]). In particular, general activity increases, in favor of relatively strong although sub-extreme events. That was established first by analysis of multiple fracturing and seismicity [5, 6], and later generalized to socio-economic processes [7]. Our results suggest a simple universal mechanism of such premonitory patterns.

MODEL

The system consists of particles indexed by their *generation* $k = 0, 1, \ldots$. Particles of zero generation (*immigrants*) are injected into the system by an external forcing. Particles of any generation $k > 0$ are produced as a result of splitting of particles of generation $k-1$. Immigrants ($k = 0$) are born at the origin $\mathbf{x} := (x_1, \ldots, x_n) = \mathbf{0}$ according to a homogeneous Poisson process with intensity μ. Each particle lives for some random time τ and then transforms (splits) into a random number β of particles of the next generation. The probability laws of the lifetime τ and branching β are rank-, time-, and space-independent. New particles are born at the location of their parent at the moment of splitting.

The lifetime distribution is exponential: $\mathsf{P}\{\tau < t\} = 1 - e^{-\lambda\, t}$, $\lambda > 0$. The conditional probability that a particle transforms into $n \geq 0$ new particles (0 means that it disappears) given that the transformation took place is denoted by p_n. The probability generating function for the number β of new particles is thus

$$h(s) = \sum_n p_n\, s^n. \tag{1}$$

The expected number of offsprings (also called the *branching number*) is $B := E(\beta) = h'(1)$ (see *e.g.*, [4]).

Each particle diffuses in $\mathbb{R}^n$ independently of other particles. This means that the density $p(\mathbf{x}, \mathbf{y}, t)$ of a particle that was born at instant 0 at point $\mathbf{y}$ solves the equation

$$\frac{\partial p}{\partial t} = D \left(\sum_i \frac{\partial^2}{\partial x_i^2} \right) p \equiv D \triangle_{\mathbf{x}}\, p \tag{2}$$

with the initial condition $p(\mathbf{x}, \mathbf{y}, 0) = \delta(\mathbf{x} - \mathbf{y})$. The solution of (2) is given by

$$p(\mathbf{x}, \mathbf{y}, t) = (4\,\pi\, D\, t)^{-n/2} \exp\left\{ -\frac{|\mathbf{x} - \mathbf{y}|^2}{4\, D\, t} \right\}, \tag{3}$$

where $|\mathbf{x}|^2 = \sum_i x_i^2$.

It is convenient to introduce particle *rank* $r := r_{\max} - k$ for an arbitrary integer $r_{\max}$ and thus consider particles of ranks $r \leq r_{\max}$. This reflects our focus on direct cascading, which often assumes that particles with larger size (*e.g.*, length, volume, mass, energy, momentum, *etc.*) split into smaller ones according to an appropriate conservation law. Figure 1 illustrates the model dynamics.

SPATIO-TEMPORAL PARTICLE RANK DISTRIBUTIONS

The model decsribed above is a superposition of independent branching processes generated by individual immigrants. We consider first the case of a single immigrant; then we expand these results to the case of multiple immigrants. Finally, we analyze the rank distribution of particles. Proofs of all statements will be published in a forthcoming paper.

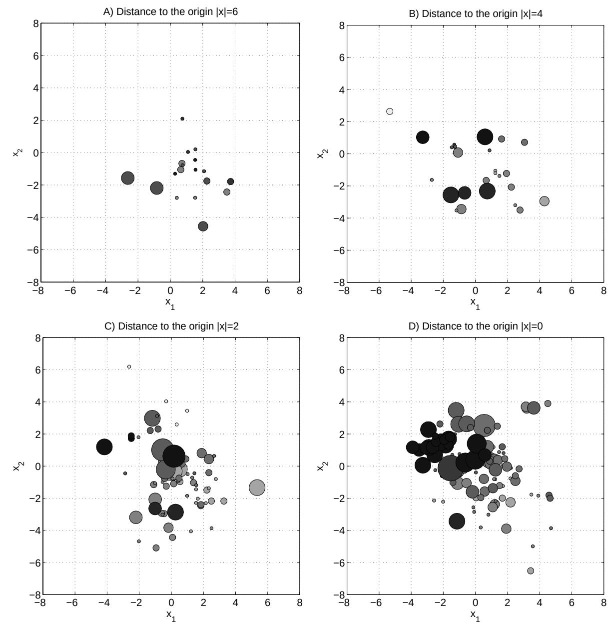

FIG. 1: Example of a 3D model population. Different panels show 2D subspaces of the model 3D space at different distances $|\mathbf{x}|$ to the origin. Model parameters are $\mu = \lambda = 1$, $D = 1$, $B = 2$. Circle size is proportional to the particle rank. Different shades correspond to populations from different immigrants, the descendants of earlier immigrants have lighter shade. The clustering of particles is explained by the splitting histories. Note that, as the origin approaches, the particle activity significantly changes, indicating the increased probability of an extreme event.

Single immigrant

Let $p_{k,i}(G, \mathbf{y}, t)$ be the conditional probability that at time $t \geq 0$ there exist $i \geq 0$ particles of generation $k \geq 0$ within spatial region $G \subset \mathbb{R}^n$ given that at time 0 a single immigrant was injected at point $\mathbf{y}$. The corresponding generating function is

$$F_k(G, \mathbf{y}, t; s) = \sum_i p_{k,i}(G, \mathbf{y}, t) s^i. \tag{4}$$

Proposition 1 *The generating functions $F_k(G, \mathbf{y}, t; s)$ solve the following recursive system of non-linear partial differential equations:*

$$\frac{\partial}{\partial t} F_k = -D \triangle_{\mathbf{y}}\, F_k - \lambda\, F_k + \lambda\, h\left(F_{k-1}\right), \quad k \geq 1, \tag{5}$$

with the initial conditions $F_k(G, \mathbf{y}, 0; s) \equiv 1$, $k \geq 1$, and

$$F_0(G, \mathbf{y}, t; s) = (1 - P) + P\, s, \tag{6}$$

where $P := e^{-\lambda\, t} \int_G p(\mathbf{x}, \mathbf{y}, t) d\mathbf{x}$.

Next, consider the expected number $\bar{A}_k(G, \mathbf{y}, t)$ of generation k particles at instant t within the region G produced by a single immigrant injected at point $\mathbf{y}$ at time $t = 0$. It is given by the following partial derivative (see *e.g.*, [4])

$$\bar{A}_k(G, \mathbf{y}, t) := \frac{\partial F_k(G, \mathbf{y}, t; s)}{\partial s} \,|_{s=1} . \tag{7}$$

Consider also the expectation density $A_k(\mathbf{x}, \mathbf{y}, t)$ that satisfies, for any $G \subset \mathbb{R}^n$,

$$\bar{A}_k(G, \mathbf{y}, t) = \int_G A_k(\mathbf{x}, \mathbf{y}, t)\, d\mathbf{x}. \tag{8}$$

Corollary 2 *The expectation densities $A_k(\mathbf{x}, \mathbf{y}, t)$ solve the following recursive system of linear partial differential equations:*

$$\frac{\partial A_k}{\partial t} = D \triangle_{\mathbf{x}} A_k - \lambda A_k + \lambda B A_{k-1}, \quad k \geq 1, \tag{9}$$

with the initial conditions $A_k(\mathbf{x}, \mathbf{y}, 0) \equiv 0,\ k \geq 1$,

$$\begin{aligned} A_0(\mathbf{x}, \mathbf{y}, 0) &= \delta(\mathbf{y} - \mathbf{x}), \\ A_0(\mathbf{x}, \mathbf{y}, t) &= e^{-\lambda t} p(\mathbf{x}, \mathbf{y}, t),\ t > 0. \end{aligned} \tag{10}$$

The solution to this system is given by

$$A_k(\mathbf{x}, \mathbf{y}, t) = \frac{(\lambda B t)^k}{k!} A_0(\mathbf{x}, \mathbf{y}, t). \tag{11}$$

The system (9) has a transparent intuitive meaning. The rate of change of the expectation density $A_k(\mathbf{x}, \mathbf{y}, t)$ is affected by the three processes: diffusion of the existing particles of generation k in $\mathbb{R}^n$ (first term in the rhs of (9)), splitting of the existing particles of generation k at the rate λ (second term), and splitting of generation $k-1$ particles that produce on average B new particles of generation k (third term).

Multiple immigrants

Here we expand the results of the previous section to the case of multiple immigrants that appear at the origin according to a homogeneous Poisson process with intensity μ. The expectation $\mathcal{A}_k$ of the number of particles of generation k is given, according to the properties of expectations, by

$$\mathcal{A}_k(\mathbf{x}, t) = \int_0^t A_k(\mathbf{x}, \mathbf{0}, s)\, \mu\, ds \tag{12}$$

The steady-state spatial distribution $\mathcal{A}_k(\mathbf{x})$ corresponds to the limit $t \to \infty$ and is given by

$$\mathcal{A}_k(z) = \frac{\mu}{\lambda k!} \left(\frac{B}{2}\right)^k \left(\frac{2 \pi D}{\lambda}\right)^{-n/2} z^{\nu} K_{\nu}(z). \tag{13}$$

Here $z := |\mathbf{x}| \sqrt{\lambda/D}$, $\nu = k - n/2 + 1$ and K_ν is the modified Bessel function of the second kind.

Rank distribution and spatial deviations

Recall that the particle rank is defined as $r = r_{\max} - k$. The spatially averaged steady-state rank distribution is a pure exponential law with index B:

$$A_k = \int_{\mathbb{R}^n} \int_0^\infty A_k(\mathbf{x}, \mathbf{0}, t) \mu \, dt\, d\mathbf{x} = \frac{\mu}{\lambda} B^k \propto B^{-r}. \tag{14}$$

To analyze deviations from the pure exponent, we consider the ratio $\gamma_k(\mathbf{x})$ between the number of particles of two consecutive generations:

$$\gamma_k(\mathbf{x}) := \frac{\mathcal{A}_k(\mathbf{x})}{\mathcal{A}_{k+1}(\mathbf{x})}. \tag{15}$$

For the purely exponential rank distribution, $A_k(\mathbf{x}) = c\, B^k$, the value of $\gamma_k(\mathbf{x}) = 1/B$ is independent of k and $\mathbf{x}$; while deviations from the pure exponent will cause γ_k to vary as a function of k and/or $\mathbf{x}$. Combining (13) and (15) we find

$$\gamma_k(\mathbf{x}) = \frac{2\,(k+1)}{B\, z} \frac{K_\nu(z)}{K_{\nu+1}(z)}, \tag{16}$$

where, as before, $z := |\mathbf{x}| \sqrt{\lambda/D}$ and $\nu = k - n/2 + 1$.

Proposition 3 *The asymptotic behavior of the function $\gamma_k(z)$ is given by*

$$\lim_{z \to 0} \gamma_k(z) = \begin{cases} \infty, & \nu \leq 0, \\ \dfrac{1}{B}\left(1 + \dfrac{n}{2\nu}\right), & \nu > 0, \end{cases} \tag{17}$$

$$\gamma_k(z) \sim \frac{2(k+1)}{B\, z},\ z \to \infty,\ \text{fixed } k, \tag{18}$$

$$\gamma_k(z) \sim \frac{1}{B}\left(1 + \frac{n}{2\nu}\right),\ k \to \infty,\ \text{fixed } z. \tag{19}$$

Proposition 3 allows one to describe all deviations of the particle rank distribution from the pure exponential law (14). Figure 2 illustrates our findings. First, Eq. (19) implies that at any spatial point, the distribution asymptotically approaches the exponential form as rank r decreases (generation k increases). Thus the deviations can only be observed at the largest ranks (small generation numbers). Analysis of the large-rank distribution is done using Eqs. (17),(18). Near the origin, where the immigrants enter the system, Eq. (17) implies that $\gamma_k(z) > \gamma_{k+1}(z) > 1/B$ for $\nu > 0$. Hence, one observes the *upward deviations* from the pure exponent: for the same number of rank r particles, the number of rank $r+1$ particles is larger than predicted by (14). The same behavior is in fact observed for $\nu \leq 0$ (the details will be published elsewhere). In addition, for $\nu \leq 0$ the ratios $\gamma_k(z)$ do not merely deviate from $1/B$, but diverge to infinity at the origin. Away from the origin, according to Eq. (18), we have $\gamma_k(z) < \gamma_{k+1}(z) < 1/B$, which implies *downward deviations* from the pure exponent: for the same number of rank r particles, the number of rank $r+1$ particles is smaller than predicted by (14).

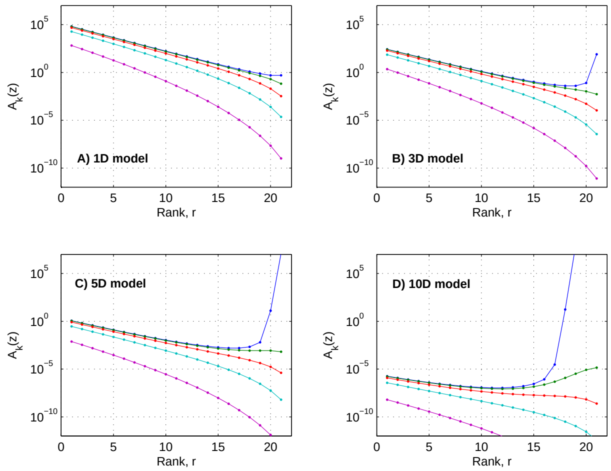

FIG. 2: Deviations from self-similarity: Expected number $A_k(z)$ of generation k particles at distance z from the origin (cf. Proposition 3). The distance z is increasing (from top to bottom line in each panel) as $z = 10^{-3}, 2, 5, 10, 20$. Model dimension is $n = 1$ (panel A), $n = 3$ (panel B), $n = 5$ (panel C), and $n = 10$ (panel D). Other model parameters: $\mu = \lambda = 1$, $D = 1$, $B = 2$, $r_{\max} = 21$. One can clearly see the transition from downward to upward deviation of the rank distributions from the pure exponential form as we approach the origin.

DISCUSSION

Motivation for this work is the problem of prediction of extreme events in complex systems. Our point of departure is a classical model of spatially distributed population of particles of different ranks governed by direct cascade of branching and external driving. In the probability theory this model is known as the age-dependent multi-type branching diffusion process with immigration [4]. We introduce here a new approach to the study of this process. We assume that observations are only possible on a subspace of the system phase space while the source of external driving remains unobservable. The natural question under this approach is the dependence of size-distributions of particles on the distance to the source. The complete analytical solution to this problem is given by the Proposition 1.

It is natural to consider rank as a logarithmic measure of the particle size. If we assume a size-conservation law in the model, the exponential rank distriburtion derived in (14) corresponds to a self-similar, power-law distribution of particle sizes, characteristic for many complex systems. Thus, the Proposition 3 describes space-dependent deviations from the self-similarity (see also Fig. 2); in particular, deviations premonitory to an extreme event. The numerical experiments (that will be published elsewhere) confirm the validity of our analytical results and asymptotics in a finite model.

The model studied here exhibits very rich and intriguing premonitory behavior. Figure 1 shows several 2D snapshots of a 3D model at different distances from the source. One can see that, as the source approaches, the following changes in the background activity emerge: a) The intensity (total number of particles) increases; b) Particles of larger size become relatively more numerous; c) Particle clustering becomes more prominent; d) The correlation radius increases; e) Coherent structures emerge. In other words, the model exhibits a broad set of premonitory phenomena previously observed heuristically in real and modeled systems: multiple fracturing [6], seismicity [5], socio-economics [7], percolation [8], hydrodynamics, hierarchical models of extreme event development [1]. These phenomena are at the heart of earthquake prediction algorithms well validated during 20 years of forward world-wide tests (see *e.g.*, [1]).

In this paper we analyse only the first-moment properties of the system; such properties can explain the premonitory intensity increase (item a above) and transformation of the particle rank distribution (item b). At the same time, the framework developed here allows one to quantitatively analyze other premonitory phenomena; this can be readily done by considering the higher-moment properties.

This study was supported in part by NSF grant No. ATM 0620838.

* Electronic address: `agabriel@math.purdue.edu`
† Electronic address: `vkb@ess.ucla.edu`
‡ Electronic address: `zal@unr.edu`

[1] V. I. Keilis-Borok and A. A. Soloviev, A. A. (eds), *Nonlinear Dynamics of the Lithosphere and Earthquake Prediction.* (Springer, Heidelberg, 2003). D. Sornette, *Critical Phenomena in Natural Sciences.* 2-nd ed. (Springer-Verlag, Heidelberg, 2004).
[2] S. Albeverio, V. Jentsch, and H. Kantz (eds), *Extreme Events in Nature and Society* (Springer, Heidelberg, 2005). P. Embrechts, C. Kluppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance* (Springer, 2004).
[3] U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge University Press, 1996).
[4] K. B. Athreya and P. E. Ney, *Branching Processes* (Dover Publications, 2004).
[5] V. I. Keilis-Borok, Proc. Natl. Ac. Sci. USA, **93**, 3748-3755 (1996). V. I. Keilis-Borok, Ann. Rev. Earth Planet. Sci., **30**, 1-33 (2002). J. Rundle, D. Turcotte, and W. Klein (eds), *Geocomplexity and the Physics of Earthquakes.* (AGU, Washington DC, 2000). D. L. Turcotte, Ann. Rev. Earth Planet. Sci., **19**, 263-281 (1991). S. C. Jaume and L. R. Sykes, Pure Appl. Geophys., **155** (2-4): 279-305 (1999).
[6] I. M. Rotwain, V. I. Keilis-Borok, and L. Botvina, Phys. Earth Planet. Inter., **101**, 61-71 (1997).
[7] V. I. Keilis-Borok, A. A. Soloviev, C. B. Allegre CB, *et al.* Pattern Recognition **38**, (3), 423-435 (2005).
[8] I. Zaliapin, H. Wong, and A. Gabrielov, Phys. Rev. E, **71**, 066118 (2004).