



HAL
open science

Optimizing Noisy Complex Systems Liable to Failure

Davin Lunz

► **To cite this version:**

Davin Lunz. Optimizing Noisy Complex Systems Liable to Failure. *SIAM Journal on Applied Mathematics*, 2022, 82 (1), pp.25-48. 10.1137/21M1416126 . hal-03537040

HAL Id: hal-03537040

<https://inria.hal.science/hal-03537040v1>

Submitted on 20 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OPTIMISING NOISY COMPLEX SYSTEMS LIABLE TO FAILURE

DAVIN LUNZ*

Abstract. Inspired by complex systems in social and industrial contexts, we consider a family of coupled diffusion processes modeling system components, and an associated system objective. Each process is inherently noisy, driven by a controllable drift, and fails upon reaching a critical state. Interdependence is captured via the global objective and the governing dynamics (correlated noise, cascading failures). Analytical and numerical calculations reveal that the optimal strategies to steer such systems so as to maximise the objective are highly coupled, depending strongly on the state of the entire system. Strikingly, they exhibit a rich set of bifurcations, describing qualitatively different strategies throughout the parameter space.

Key words. Stochastic processes with reset, optimal control of PDEs, asymptotic analysis, discrete adjoint approach

AMS subject classifications. 35Q84, 49K20, 49N90, 90B25, 35B40

1. Introduction. Complex systems encompass collective dynamics born out of the interactions between constituent components, and exhibit fascinating emergent features. The theory of complex systems — with broad application in power transmission, information cascade, disease outbreak, and biochemical processes to name a few — seeks to understand emergent behaviour as a function of the individual building blocks and their interdependence, with the aim of controlling such systems to guarantee robustness [54].

Network structure and dynamics are a prominent focus of complex systems theory [8, 15, 24, 49, 53]. In particular, cascading failures are widely studied as flow through a capacity-limited network. When a component fails the load redistribution can cause neighbouring components to fail, propagating an avalanche of failure [45, 46, 48, 51]. Characterising criticality by studying what network features allow this instability to manifest provides insight for failure-resilient design [1, 14, 32] and response strategies [47]. While previous work revealed the conditions for a single network to fail, less attention has been devoted to interacting networks, each liable to failure [10].

In this work, we address the question of interacting networks liable to failure from a slightly different perspective. We consider a network of interconnected, controllable diffusion processes. Each process represents the performance of a single network. Crucially, the diffusion processes undergo ‘reset’ upon reaching a critical state, modeling a single network failure. By modeling the single network by a diffusion process, we abstract away the intra-network failure mechanism, and instead focus on properties of the inter-network system, including the case of inter-network cascading failures.

Failures are costly but may not be catastrophic in the long run. This motivates us to consider a long-term objective function that rewards high performance over long time horizons while also accounting for costly failure. This proves a vexing challenge for systems where high performance comes hand-in-hand with high risk of failure, begging the question of how to best balance the risks and rewards. Our aim is to determine the optimal control drift (that which maximises the objective) subject to the complex system dynamics, thereby balancing the inherent trade-off

*Inria Saclay – Île de France, 91120 Palaiseau, France (davin.lunz@inria.fr)
École Polytechnique, CMAP 91128 Palaiseau, France
Institut Pasteur, 75015 Paris, France.

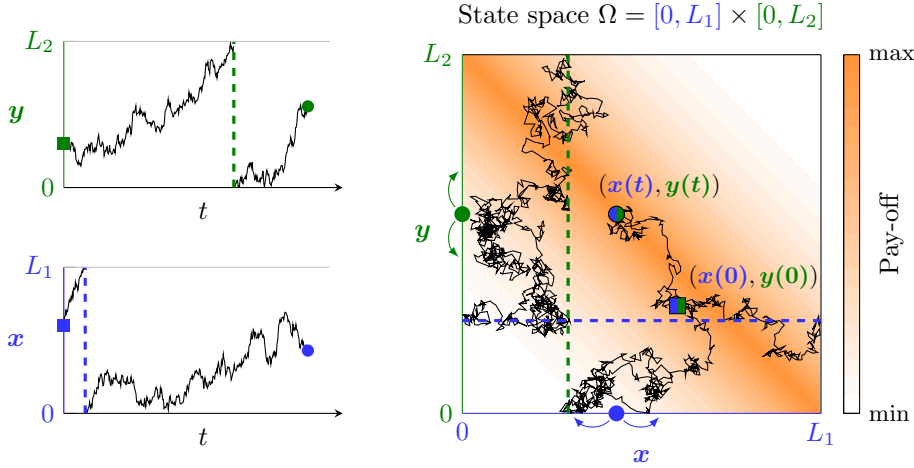


Fig. 1: Sample paths of two processes subject to reset, $x(t)$ and $y(t)$, each evolving on a domain $[0, L_i]$, resetting to zero upon hitting L_i (dashed lines). Note that the reset events act to set the state from L_i to 0, therefore, dashed lines show state changes from top to bottom in the one-dimensional plots on the left, where the corresponding coordinate jumps to zero in the two-dimensional state space $(x(t), y(t))$ on the right. The initial state is marked with a square, and the state at some later time t is marked with a circle. The processes may be *coupled* via their governing dynamics or by a pay-off function (here we depict a coupling pay-off function of $x + y$).

44 while accounting for the complexity of process interdependence.

45 The elementary processes in our model are one-dimensional diffusion processes
 46 subject to reset. Each process state evolves via a tunable drift under the influence of
 47 noise. A reset law governs jumps back to an origin. Aptly, resetting stochastic pro-
 48 cesses are versatile models, finding application from stochastic search algorithms [17,
 49 18, 22, 25, 26, 40], through biological systems of DNA polymerase binding [7, 16, 50]
 50 and active matter [20, 35, 52], to ecological populations [5, 9, 36, 37, 38], communica-
 51 tions queues [11, 33, 34], social [42] and mechanical systems in the study of reliabil-
 52 ity [12, 13, 41]. Resetting holds the process away from equilibrium [19], making it a
 53 useful model for non-equilibrium components of a complex system. Refs. [20, 21, 43]
 54 provide more extensive references.

55 We couple the individual processes via the global objective function (e.g. in power
 56 grids, where aggregate supply must match demand and an excess of either supply or
 57 demand is undesirable, the ideal output of any single power station depends on the
 58 other stations; see Figure 1). In the first instance we assume that each process observes
 59 the entire system state, and later we consider partial observability. Furthermore, we
 60 allow noise sources to be correlated among different processes, capturing inter-network
 61 noise covariance (e.g. fluctuations in wind conditions affecting wind turbines may be
 62 correlated depending on location).

63 The article is structured as follows. In section 2 we formulate the optimal control
 64 problem directly in terms of the underlying stochastic process. In section 3 we
 65 reformulate the problem in terms of its associated law by appealing to renewal theo-
 66 ry. In section 4 we outline the numerical discretisation and the adjoint optimisation

67 schemes we employ to compute solutions to the reformulated problem. The adjoint
 68 setting motivates an accompanying continuum adjoint analysis. In [section 5](#) we il-
 69 lustrate concrete solutions, showcasing some of the bifurcations solutions undergo as
 70 problem parameters vary. Finally, in [section 6](#) we discuss our findings and provide an
 71 outlook for future work.

72 **2. Problem statement.** Consider a family of d diffusion processes with state
 73 $\mathbf{x}(t) = (x_1(t), \dots, x_d(t))$ at time t , evolving within the domain $\Omega = [0, L_1] \times \dots \times$
 74 $[0, L_d] \subset \mathbb{R}^d$ governed by the Itô SDE

$$75 \quad (2.1) \quad d\mathbf{x}(t) = \mathbf{v}(\mathbf{x}(t)) dt + \boldsymbol{\sigma}(\mathbf{x}(t)) d\mathbf{W}(t), \quad \mathbf{x}(0) = \mathbf{x}_0,$$

77 for a drift coefficient $\mathbf{v} : \Omega \rightarrow \mathbb{R}^d$, a diffusion coefficient $\boldsymbol{\sigma} : \Omega \rightarrow \mathbb{R}^{d \times m}$, where $\mathbf{W}(t) \in$
 78 \mathbb{R}^m is a standard m -dimensional Wiener process. The initial condition $\mathbf{x}_0 \in \Omega$ may
 79 be chosen from a distribution over the state space $P^0(\mathbf{x})$ for $\mathbf{x} \in \Omega$ or a distribution
 80 over the 0-boundary $\Phi^0(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega_0$, or a combination of the two, such that

$$81 \quad (2.2) \quad \int_{\Omega} P^0(\mathbf{x}) d\mathbf{x} + \int_{\partial\Omega_0} \Phi^0(\mathbf{x}) d\mathbf{x} = 1.$$

83 The domain Ω is enclosed by $\partial\Omega$ comprising the hyperplanes $x_i = 0$ and $x_i = L_i$
 84 for each $i = 1, \dots, d$. We distinguish between these boundaries, denoting $\partial\Omega =$
 85 $\partial\Omega_0 \cup \partial\Omega_L$ for

$$86 \quad \partial\Omega_0 = \{x_i = 0 \text{ for any } i\}, \quad \partial\Omega_L = \{x_i = L_i \text{ for any } i\},$$

88 which we call the 0-boundary and L -boundary, respectively. The intersection $\partial\Omega_0 \cap$
 89 $\partial\Omega_L$ is nonempty but is of zero measure, and so while $\partial\Omega_0$ and $\partial\Omega_L$ are not strictly
 90 distinct, we nevertheless treat them as such. For points on the boundary we denote
 91 by \mathbf{x}^c the point opposite \mathbf{x} : if the i th coordinate of \mathbf{x} is on the boundary, $x_i \in \{0, L_i\}$,
 92 then \mathbf{x}^c is the vector sharing all coordinates $x_j^c = x_j$ for all $j \neq i$ and $x_i^c = L_i - x_i$.
 93 While this is not well-defined for points $\mathbf{x} \in \partial\Omega_0 \cap \partial\Omega_L$, these are neglected since the
 94 intersection is of zero measure.

95 Upon hitting an L -boundary, $\mathbf{x} \in \partial\Omega_L$, a process has reached criticality and
 96 ‘resets’, that is, it is transported to the 0-boundary, $\mathbf{x}^c \in \partial\Omega_0$. This accounts exclu-
 97 sively for the reset of the process that reached criticality. In [section 3.1](#) we discuss
 98 cascading failures, where the failure of one component results in the knock-on failure
 99 of other components, captured by setting multiple components to zero upon hitting
 100 an L -boundary. A particle in the vicinity of a 0-boundary experiences a reflecting
 101 boundary (that is, the resetting jump is only one way).

102 We impose bounds on the drift $\mathbf{v}(\mathbf{x}) \in [u, U]^d$, to capture the fact that our control
 103 of the processes is limited. We consider long-time pay-off functions $R(\mathbf{v})$ of the form

$$104 \quad (2.3) \quad R(\mathbf{v}) = \lim_{T \rightarrow \infty} \frac{1}{T} \left[\int_0^T f(\mathbf{x}(t)) dt - CN(T) \right],$$

106 which reward some regions of state space through f (that is, $f : \Omega \rightarrow \mathbb{R}$ is path-
 107 independent), and penalise resets via the cost C , with $N(T)$ counting the number of
 108 resets until time $t = T$.

109 We thus arrive at the problem formulation: we seek an optimal drift, maximising
 110 the objective

$$111 \quad (2.4) \quad \max_{\mathbf{v}(\mathbf{x}) \in [u, U]^d} R(\mathbf{v}),$$

112

113 subject to the dynamics (2.1) and the aforementioned reset rules.

114 Typically, the objective function to optimise requires an expectation to be taken
115 over the stochastic process. However, since we are interested exclusively in the ergodic
116 limit, there is no expectation in this formulation.

117 Our approach is to study the law of the process (2.1) subject to reset. We analyse
118 the Fokker–Planck equation governing the law in the large-time limit, and leverage
119 renewal theory to represent (2.4) as a PDE-constrained optimisation problem. We
120 then pursue both analytical and numerical results that shed light on the solution
121 structure.

122 **3. The law and its analysis.** Consider the law of the stochastic process (2.1)
123 subject to reset, which is the probability density $P(\mathbf{x}, t)$ of the particle being at
124 location \mathbf{x} at time t , and is governed by the Fokker–Planck equation

$$125 \quad (3.1a) \quad \frac{\partial P}{\partial t}(\mathbf{x}, t) + \nabla \cdot \Phi(\mathbf{x}, t) = 0,$$

126 for $\mathbf{x} \in \mathring{\Omega}$ and $t > 0$, where the probability flux Φ is given by

$$128 \quad (3.1b) \quad \Phi(\mathbf{x}, t) = \mathbf{v}(\mathbf{x})P(\mathbf{x}, t) - \nabla \cdot (\mathbf{D}(\mathbf{x}, \mathbf{v}(\mathbf{x}))P(\mathbf{x}, t)),$$

130 for the diffusivity tensor $\mathbf{D} = \boldsymbol{\sigma}\boldsymbol{\sigma}^\top/2$. The reset mechanism and reflecting boundaries
131 are represented by the boundary conditions:

$$132 \quad (3.1c) \quad P(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial\Omega_L,$$

$$133 \quad (3.1d) \quad \Phi(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) = -\Phi^0(\mathbf{x})\delta(t) + \Phi(\mathbf{x}^c, t) \cdot \mathbf{n}(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega_0,$$

135 where $\mathbf{n}(\mathbf{x})$ is the outward-facing normal at \mathbf{x} in both terms. The first term on the
136 right-hand side of (3.1d) captures an initial injection of probability mass density Φ^0
137 along $\partial\Omega_0$. The minus sign is because the normal is outward-facing, thus Φ^0 is positive.
138 We specify an initial probability density P^0 in the domain interior $\mathring{\Omega}$ via

$$139 \quad (3.1e) \quad P(\mathbf{x}, 0) = P^0(\mathbf{x}), \quad \mathbf{x} \in \mathring{\Omega}.$$

141 We also impose the unit-mass requirement (2.2) for the initial probability within the
142 domain and injected into it.

143 The system is ergodic when the diffusivity is non-degenerate [2, 3, 6, 28, 29, 39],
144 therefore, the law will converge to a unique stationary distribution. Since we are
145 interested in long time horizons during which systems recover from failure, as reflected
146 in the long-term pay-offs (2.3), we may neglect the initial transient (and the initial
147 conditions) and focus exclusively on the steady-state solution of (3.1). With this
148 restriction to the large-time limit in mind, we now apply the renewal approach.

149 We introduce the simpler first-passage process corresponding to (3.1) without the
150 resetting mechanism. We denote the corresponding density $p(\mathbf{x}, t)$ and flux $\phi(\mathbf{x}, t)$,
151 satisfying

$$152 \quad (3.2a) \quad \frac{\partial p}{\partial t}(\mathbf{x}, t) + \nabla \cdot \phi(\mathbf{x}, t) = 0,$$

154 for $\mathbf{x} \in \mathring{\Omega}$ and $t > 0$, where the flux is given by

$$155 \quad (3.2b) \quad \phi(\mathbf{x}, t) = \mathbf{v}(\mathbf{x})p(\mathbf{x}, t) - \nabla \cdot (\mathbf{D}(\mathbf{x}, \mathbf{v}(\mathbf{x}))p(\mathbf{x}, t)).$$

157 We impose the absorbing and reflecting boundary conditions

$$158 \quad (3.2c) \quad p(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial\Omega_L,$$

$$159 \quad (3.2d) \quad \phi(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) = -\hat{\Phi}^0(\mathbf{x})\delta(t), \quad \mathbf{x} \in \partial\Omega_0,$$

161 with the initial injection contribution, and impose the initial conditions

$$162 \quad (3.2e) \quad p(\mathbf{x}, 0) = \hat{P}^0(\mathbf{x}), \quad \mathbf{x} \in \hat{\Omega}.$$

164 The only difference between (3.1) and (3.2) is the flux boundary condition on $\partial\Omega_0$.

165 We introduce $Q(t)$ the probability density of the first-passage time to the L -
166 boundary being t , which is given by the probability flux at the L -boundary

$$167 \quad (3.3) \quad Q(t) = \int_{\partial\Omega_L} \phi(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{x}.$$

169 We may now express the large-time probability density P recursively, with the
170 help of p and Q via

$$171 \quad (3.4) \quad P(\mathbf{x}, t) = p(\mathbf{x}, t) + \int_0^t Q(\tau)P(\mathbf{x}, t - \tau) \, d\tau,$$

173 representing the probability of a particle to be located at \mathbf{x} if it is yet to hit an L -
174 boundary, or, if it first hit the boundary at a previous time and arrived at \mathbf{x} after
175 reset. We introduce $N(t)$ the average number of resets up until time t . This may also
176 be expressed recursively with the aid of Q via

$$177 \quad (3.5) \quad N(t) = \int_0^t Q(\tau)(1 + N(t - \tau)) \, d\tau,$$

179 representing the density of the first reset occurring at any intermediate time $0 \leq \tau \leq t$
180 followed by (one more than) the average number of resets occurring in the remaining
181 time until t .

182 The renewal-theory approach, where the resetting process is expressible recur-
183 sively, relies on the assumption that, after each reset, the process is described by
184 p in (3.2) until the subsequent reset. This observation allows us to determine the
185 initial distribution \hat{P}^0 and initial injection density $\hat{\Phi}^0$ as follows. Since the process
186 resets to the 0-boundary but not the domain interior $\hat{\Omega}$ it follows that p has no initial
187 distribution within the domain, $\hat{P}^0 = 0$. Similarly, from the periodic boundary condi-
188 tions (3.1d) we see that the initial injection density $\hat{\Phi}^0$ must coincide with the steady-
189 state normal flux at the L -boundary. In other words, the underlying assumption of
190 the renewal approach is satisfied for a zero initial distribution and the stationary flux
191 exiting the L -boundary injected at the 0-boundary. Of course, the law (3.1) of the
192 original process (2.1) may have an arbitrary initial condition P^0 and injection density
193 Φ^0 imposed; these restrictions are only required for p in order that the renewal theory
194 be applicable in the large-time limit. Despite having deduced the form of $\hat{\Phi}^0$, we
195 retain this term as is for the moment, since we do not know the steady-state normal
196 flux at the L -boundary.

197 A Laplace transform, which we denote by a tilde, diagonalises the convolutions
198 in the recursive forms (3.4) and (3.5) to give \tilde{P} and \tilde{N} in terms of \tilde{p} and \tilde{Q} , namely

$$199 \quad (3.6) \quad \tilde{P}(\mathbf{x}, s) = \frac{\tilde{p}(\mathbf{x}, s)}{1 - \tilde{Q}(s)}, \quad \tilde{N}(s) = \frac{\tilde{Q}(s)}{s(1 - \tilde{Q}(s))},$$

200

with s being the Laplace frequency. We see from (3.2b) and (3.3) that \tilde{Q} is a function only of \tilde{p} . Therefore, it follows from (3.6) that \tilde{P} and \tilde{N} are also functions only of \tilde{p} . We thus seek to solve (3.2) in Laplace space, namely

$$(3.7a) \quad s\tilde{p}(\mathbf{x}, s) + \nabla \cdot \tilde{\phi}(\mathbf{x}, s) = 0,$$

for

$$(3.7b) \quad \tilde{\phi}(\mathbf{x}, s) = \mathbf{v}(\mathbf{x})\tilde{p}(\mathbf{x}, s) - \nabla \cdot (\mathbf{D}(\mathbf{x}, \mathbf{v}(\mathbf{x}))\tilde{p}(\mathbf{x}, s)),$$

subject to the boundary conditions

$$(3.7c) \quad \tilde{p}(\mathbf{x}, s) = 0, \quad \mathbf{x} \in \partial\Omega_L,$$

$$(3.7d) \quad \tilde{\phi}(\mathbf{x}, s) \cdot \mathbf{n}(\mathbf{x}) = -\hat{\Phi}^0(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega_0.$$

Solving the transformed (3.7) analytically in the general case appears to be intractable. Our aim is to solve in the large-time asymptotic limit $s \rightarrow 0$. We pose the expansion

$$(3.8) \quad \tilde{p}(\mathbf{x}, s) \sim \tilde{p}_0(\mathbf{x}) + s\tilde{p}_1(\mathbf{x}) + \dots,$$

and adopt an analogous notation for the expansion of the corresponding flux terms $\tilde{\phi}$. The leading-order equation takes the form,

$$(3.9a) \quad \nabla \cdot \tilde{\phi}_0(\mathbf{x}) = 0,$$

for $\mathbf{x} \in \mathring{\Omega}$, subject to the boundary conditions

$$(3.9b) \quad \tilde{p}_0(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\Omega_L,$$

$$(3.9c) \quad \tilde{\phi}_0(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = -\hat{\Phi}^0(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega_0.$$

The first-order equation takes the form

$$(3.10a) \quad \nabla \cdot \tilde{\phi}_1(\mathbf{x}) = -\tilde{p}_0(\mathbf{x}),$$

for $\mathbf{x} \in \mathring{\Omega}$, subject to the boundary conditions

$$(3.10b) \quad \tilde{p}_1(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\Omega_L,$$

$$(3.10c) \quad \tilde{\phi}_1(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\Omega_0.$$

Using the divergence theorem, it follows from (2.2) and (3.9) that the leading-order normal flux through $\partial\Omega_L$ is given by

$$(3.11) \quad \begin{aligned} \int_{\partial\Omega_L} \tilde{\phi}_0(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{x} &= \int_{\Omega} \nabla \cdot \tilde{\phi}_0(\mathbf{x}) \, d\mathbf{x} - \int_{\partial\Omega_0} \tilde{\phi}_0(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{x} \\ &= \int_{\partial\Omega_0} \hat{\Phi}^0(\mathbf{x}) \, d\mathbf{x} \\ &= 1. \end{aligned}$$

Similarly at first order, it follows from (3.10) that

$$(3.12) \quad \begin{aligned} \int_{\partial\Omega_L} \tilde{\phi}_1(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{x} &= \int_{\Omega} \nabla \cdot \tilde{\phi}_1(\mathbf{x}) \, d\mathbf{x} - \int_{\partial\Omega_0} \tilde{\phi}_1(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{x} \\ &= - \int_{\Omega} \tilde{p}_0(\mathbf{x}) \, d\mathbf{x}. \end{aligned}$$

240 Combining (3.3), (3.11), and (3.12), we find that, up to first-order,

$$\begin{aligned}
 241 \quad (3.13) \quad \tilde{Q}(s) &\sim \int_{\partial\Omega_L} \tilde{\phi}_0(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{x} + s \int_{\partial\Omega_L} \tilde{\phi}_1(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{x} \\
 &= 1 - s \int_{\Omega} \tilde{p}_0(\mathbf{x}) \, d\mathbf{x}.
 \end{aligned}$$

243 From (3.6) and (3.13), after inverting the Laplace transform, we deduce that the
 244 large-time distribution and the average number of resets may be expressed as

$$\begin{aligned}
 245 \quad (3.14) \quad P(\mathbf{x}) &\sim \frac{\tilde{p}_0(\mathbf{x})}{\int_{\Omega} \tilde{p}_0(\mathbf{z}) \, d\mathbf{z}}, & \frac{N(T)}{T} &\sim \frac{1}{\int_{\Omega} \tilde{p}_0(\mathbf{x}) \, d\mathbf{x}},
 \end{aligned}$$

247 that is, P is a normalisation of the \tilde{p}_0 density.

248 Equipped with the large-time distribution (3.14) and large-time reset count (3.14),
 249 we are able to find expressions for objective functions of the form (2.3), namely

$$\begin{aligned}
 250 \quad (3.15) \quad R &= \lim_{T \rightarrow \infty} \frac{1}{T} \left[\int_0^T f(\mathbf{x}(t)) \, dt - CN(T) \right] = \frac{\int_{\Omega} f(\mathbf{x}) \tilde{p}_0(\mathbf{x}) \, d\mathbf{x} - C}{\int_{\Omega} \tilde{p}_0(\mathbf{x}) \, d\mathbf{x}}.
 \end{aligned}$$

252 Crucially, the objective (3.15) is given exclusively in terms of \tilde{p}_0 , the solution of (3.9).
 253 Ultimately, the ergodicity of the process has allowed us to replace the time averages
 254 in (2.3) with ensemble averages in (3.15).

255 The ergodicity ensures that the stationary distribution P is independent of the
 256 initial conditions P^0 and Φ^0 . Nevertheless, we find that \tilde{p}_0 depends on $\hat{\Phi}^0$ in (3.9),
 257 which we retained since we did not know the stationary boundary-normal flux of
 258 P . This may now be addressed by noting from (3.14) that the steady-state P is
 259 proportional to \tilde{p}_0 . Therefore, imposing the periodic flux boundary condition (3.1d)
 260 on \tilde{p}_0 itself guarantees the correct boundary-normal up to a multiplicative constant.
 261 The exact value of the constant is unimportant since \tilde{p}_0 is normalised in (3.14), and
 262 thus it suffices to ensure it is nonzero. Ultimately, we arrive at the equation governing
 263 \tilde{p}_0 , namely

$$\begin{aligned}
 264 \quad (3.16a) \quad \nabla \cdot \tilde{\phi}_0(\mathbf{x}) &= 0,
 \end{aligned}$$

266 for $\mathbf{x} \in \mathring{\Omega}$, subject to the boundary conditions

$$\begin{aligned}
 267 \quad (3.16b) \quad \tilde{p}_0(\mathbf{x}) &= 0, & \mathbf{x} &\in \partial\Omega_L,
 \end{aligned}$$

$$\begin{aligned}
 268 \quad (3.16c) \quad \tilde{\phi}_0(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) &= \tilde{\phi}_0(\mathbf{x}^c) \cdot \mathbf{n}(\mathbf{x}), & \mathbf{x} &\in \partial\Omega_0,
 \end{aligned}$$

270 where $\mathbf{n}(\mathbf{x})$ is the outward-facing normal at \mathbf{x} in both terms and the unit probability
 271 mass condition (2.2) takes the form

$$\begin{aligned}
 272 \quad (3.16d) \quad \int_{\partial\Omega_0} \tilde{\phi}_0(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{x} &= -1.
 \end{aligned}$$

274 It is worth noting that this subtlety involving the initial and boundary condi-
 275 tions of the p problem has not been previously addressed in the literature. This is
 276 because the previous studies have restricted the geometry to one-dimension [12, 41]
 277 (or imposed symmetry [13] with the same effect), whereas in this study the high di-
 278 mensionality and lack of symmetry necessitates a more careful analysis. When we
 279 consider cascading failures in section 3.1, the particular case where all processes fail
 280 upon any reset guarantees that the process resets to a single point whereby this prob-
 281 lem is again eliminated (but not the case of partial cascades, that is, if there is any
 282 process where, upon reaching criticality, not all processes reset).

283 **3.1. Cascading failures.** A cascading failure is any dependency in the system
 284 such that the failure of one component (say the i th component for some $i \in \{1, \dots, d\}$)
 285 results not only in its own reset, but also the reset of at least one other component
 286 (say j for any $i \neq j \in \{1, \dots, d\}$). Such dependencies may be asymmetrical. That is,
 287 the failure of process i might always cause the reset of process j , but not the converse:
 288 the failure of j need not cause the reset of i . A cascade may cause the reset of all
 289 other processes, but may also be partial; causing the reset of some other but not all
 290 other components.

291 Cascading failures are incorporated in the model by changing the boundary condi-
 292 tions imposed on the law to account for the fact that multiple processes may be reset
 293 when a single process reaches criticality. This raises a minor technical issue regarding
 294 the normal flux boundary condition. A cascading failure results in a particle reset
 295 via absorption at an L -boundary and injection at a 0 -boundary where $x_i = x_j = 0$
 296 for some $i \neq j$. The domain geometry at such a point is singular since the normal is
 297 not defined, making a normal flux boundary condition indeterminate. This was not
 298 previously a problem because, in the periodic case considered until now, the probabil-
 299 ity mass injected at the singular regions of the domain was of zero measure, however,
 300 this is no longer the case. To circumvent this issue, we may consider injection at a
 301 location vanishingly near this point. To preserve symmetry, we choose the internal
 302 point $x_i = x_j = \epsilon$ for some $\epsilon \ll 1$ which is in $\mathring{\Omega}$. A boundary point could also be
 303 chosen at the cost of a small loss of symmetry. In practice, we choose ϵ to simply be
 304 one discrete grid unit.

305 The prior analysis remains analogous. For the sake of concreteness, we demon-
 306 strate the case where reaching any critical state results in a cascading failure of all
 307 components. The governing Fokker–Planck equation (3.1) takes the form

$$308 \quad (3.17a) \quad \frac{\partial P}{\partial t}(\mathbf{x}, t) + \nabla \cdot \mathbf{\Phi}(\mathbf{x}, t) = \delta(\mathbf{x} - \boldsymbol{\epsilon}) \int_{\partial\Omega_L} \mathbf{\Phi}(\mathbf{z}, t) \cdot \mathbf{n}(\mathbf{z}) \, d\mathbf{z},$$

310 for $\mathbf{x} \in \mathring{\Omega}$ and $t > 0$, where $\boldsymbol{\epsilon}$ is the vector with ϵ in each entry, $\mathbf{n}(\mathbf{z})$ is the outward-
 311 facing normal at $\mathbf{z} \in \partial\Omega_L$, and the probability flux $\mathbf{\Phi}$ remains as in (3.1b). The reset
 312 mechanism and reflecting boundaries are represented by the boundary conditions:

$$313 \quad (3.17b) \quad P(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial\Omega_L,$$

$$314 \quad (3.17c) \quad \mathbf{\Phi}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) = -\Phi^0(\mathbf{x})\delta(t), \quad \mathbf{x} \in \partial\Omega_0,$$

316 where Φ^0 is the probability mass density injected initially. The initial probability
 317 density P^0 in the domain interior $\mathring{\Omega}$ is specified

$$318 \quad (3.17d) \quad P(\mathbf{x}, 0) = P^0(\mathbf{x}), \quad \mathbf{x} \in \mathring{\Omega}.$$

320 We again require that the initial probability has unit mass, satisfying (2.2).

321 With the same objective (2.3), we may apply the renewal theory if the first-passage
 322 process p describes the law of P between consecutive resets. For P governed by (3.17),
 323 this condition amounts to there being no injection probability for p on the 0 -boundary
 324 (since the reset occurs into the domain interior), and the initial distribution of p is to
 325 match the right-hand side of (3.17a) for the steady-state solution $\mathbf{\Phi}(\mathbf{x})$. As before, we
 326 do not know the total boundary flux of the steady-state solution, however, this is some
 327 nonzero constant that scales \tilde{p}_0 but is normalised in P via (3.14) and thus may be
 328 taken as unity without loss of generality. Seeking the large-time asymptotic solution

329 leads us analogously to the system governing the leading-order Laplace-transformed
 330 first-passage density \tilde{p}_0 , namely

$$331 \quad (3.18a) \quad \nabla \cdot \tilde{\phi}_0(\mathbf{x}) = \delta(\mathbf{x} - \epsilon),$$

333 for $\mathbf{x} \in \tilde{\Omega}$ and the usual flux form (3.7a), subject to the boundary conditions

$$334 \quad (3.18b) \quad \tilde{p}_0(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\Omega_L,$$

$$335 \quad (3.18c) \quad \tilde{\phi}_0(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\Omega_0.$$

337 In this case, application of the divergence theorem proves that the inhomogeneous
 338 condition

$$339 \quad (3.19) \quad \int_{\partial\Omega_L} \tilde{\phi}_0(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \, d\mathbf{x} = 1,$$

341 is satisfied by the solution of (3.18) and need not to be imposed to ensure a non-
 342 zero solution. For partial cascading failures, where a reset causes only some other
 343 components to fail, the form becomes hybrid; mixing periodic normal flux boundary
 344 conditions with delta functions of magnitude equal to the integrated flux over the
 345 corresponding critical boundary, while imposing (3.19). We do not further discuss
 346 partial cascades in this work.

347 To recap, by analysing the law of the process, we have reformulated the original
 348 stochastic optimal control problem (2.4) as a PDE-constrained optimisation problem
 349 of maximising (3.15) subject to the independent-failure model (3.16). This formula-
 350 tion may be extended to cascading failures, such as in (3.17). We now proceed to
 351 detail a numerical approach to solving the PDE-constrained optimisation problems.

352 4. The adjoint perspective.

353 **4.1. Numerical implementation.** Our aim in this section is to develop a nu-
 354 merical scheme to solve the PDE-constrained optimisation problem of maximising
 355 the objective (3.15) subject to the independent-failure model (3.16) or the cascading-
 356 failure model (3.18).

357 The first step is to discretise the associated PDEs (3.16) and (3.18). In [Appen-
 358 dix A](#) we provide complete details of a consistent finite-difference scheme tailored for
 359 anisotropic diffusion. Here, we assemble these discrete schemes (including the dis-
 360 cretisation of the differential operator, the associated boundary conditions, and any
 361 associated inhomogeneity constraints) in generic matrix form:

$$362 \quad (4.1) \quad \mathbf{A}\mathbf{p} = \mathbf{b},$$

364 where \mathbf{A} is the finite-difference matrix operator, and \mathbf{b} is the inhomogeneous right-
 365 hand side of the system. The vector \mathbf{p} comprises the discretised solution and perhaps
 366 a Lagrange multiplier for an inhomogeneity constraint.

367 The next step is to discretise the objective function R in (3.15), which we write
 368 as

$$369 \quad (4.2) \quad R \approx \frac{\left(\sum_{\mathbf{j} \in \Gamma} f(\mathbf{x}_{\mathbf{j}}) p_{\mathbf{j}}\right) H - C}{\left(\sum_{\mathbf{j} \in \Gamma} p_{\mathbf{j}}\right) H} =: \frac{(\mathbf{f}^\top \mathbf{p})H - C}{(\mathbf{1}^\top \mathbf{p})H},$$

371 where H is the discretisation of $d\mathbf{x}$ the differential element (that is, the product of
 372 the state spacing in each dimension, see (A.9)), $\mathbf{j} \in \Gamma$ indexes discrete points in the

373 discretised domain, so $\mathbf{p} = (p_j)_j$ is the discrete solution at states $\mathbf{x}_j \in \Omega$, $\mathbf{1}$ is a vector
 374 of ones with dimension matching \mathbf{p} , and \mathbf{f} is the vector of values $f(\mathbf{x}_j)$. In the case
 375 where \mathbf{p} also holds a Lagrange multiplier, we assume that the corresponding entries
 376 in the \mathbf{f} and $\mathbf{1}$ vectors are zero.

377 Note that the scheme (4.1) depends on the (discretised) control \mathbf{v} via the oper-
 378 ator $\mathbf{A} = \mathbf{A}(\mathbf{v})$. Thus we are equipped with a mapping from the discrete control
 379 \mathbf{v} , via the system (4.1) providing the solution \mathbf{p} , to the discretised objective func-
 380 tion $R(\mathbf{v}) := R(\mathbf{p}(\mathbf{v}))$. The final step, of optimising $R(\mathbf{v})$ with respect to \mathbf{v} , may be
 381 achieved by several approaches. Gradient-free methods require only the aforemen-
 382 tioned ingredients to optimise the objective, however, these typically require a large
 383 number of iterations which proves computationally demanding since the solve can
 384 be very costly with a highly resolved discretisation. For this reason, we choose the
 385 L-BFGS-B (limited-memory control-constrained BFGS) algorithm [44], a gradient-
 386 based quasi-Newton method, designed to handle large-scale optimisation problems.
 387 This approach provides significantly faster convergence by exploiting the gradient in-
 388 formation, however, it requires the gradient of R with respect to the control \mathbf{v} : $\nabla_{\mathbf{v}}R$.
 389 Note that we use the terms ‘drift’ and ‘control’ interchangeably; the former describes
 390 its physical function while the latter describes its conceptual role in the optimisation
 391 setting.

392 The most straightforward approach to compute $\nabla_{\mathbf{v}}R$ is via finite-differences, how-
 393 ever, this suffers from both numerical inaccuracy [4] and requires prohibitively many
 394 solves: one for each component of \mathbf{v} . This second limitation is particularly challenging
 395 as increasing the discretisation resolution increases the dimension of \mathbf{v} . To remedy
 396 both of these drawbacks, we calculate an exact expression for the derivative via the
 397 adjoint approach [23]. This requires the equivalent of a single extra solve, irrespective
 398 of the dimension of \mathbf{v} (an improvement by a factor on the order of $\mathcal{O}(2500)$ for the
 399 simulations in this study). We now demonstrate how the adjoint approach is applied
 400 to our discrete scheme (4.1) to yield the gradient $\nabla_{\mathbf{v}}R$ for use in the optimisation
 401 algorithm.

402 First, we introduce the discrete Lagrangian

$$403 \quad (4.3) \quad \mathcal{L} = \frac{\mathbf{f}^\top \mathbf{p} - C/H}{\mathbf{1}^\top \mathbf{p}} - \boldsymbol{\lambda}^\top (\mathbf{A}\mathbf{p} - \mathbf{b}),$$

405 where $\boldsymbol{\lambda}$ is an adjoint vector of dimension matching \mathbf{p} . The dependence of \mathcal{L} on the
 406 drift \mathbf{v} is explicit via \mathbf{A} and implicit via the dependence of \mathbf{p} on \mathbf{v} (also stemming
 407 from \mathbf{A}) through (4.1). Therefore, by the chain rule, we see that

$$408 \quad (4.4) \quad \nabla_{\mathbf{v}}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \mathbf{p}} \frac{\partial \mathbf{p}}{\partial \mathbf{v}} + \frac{\partial \mathcal{L}}{\partial \mathbf{v}} = \left[\frac{\mathbf{f}^\top}{\mathbf{1}^\top \mathbf{p}} - \frac{\mathbf{f}^\top \mathbf{p} - C/H}{\mathbf{1}^\top \mathbf{p}} \frac{\mathbf{1}^\top}{\mathbf{1}^\top \mathbf{p}} - \boldsymbol{\lambda}^\top \mathbf{A} \right] \frac{\partial \mathbf{p}}{\partial \mathbf{v}} - \boldsymbol{\lambda}^\top \frac{\partial \mathbf{A}}{\partial \mathbf{v}} \mathbf{p}.$$

410 Then, if $\boldsymbol{\lambda}$ satisfies the costate equation, $\partial \mathcal{L} / \partial \mathbf{p} = 0$, which takes the form

$$411 \quad (4.5) \quad \mathbf{A}^\top \boldsymbol{\lambda} = \frac{\mathbf{f}}{\mathbf{1}^\top \mathbf{p}} - \frac{\mathbf{f}^\top \mathbf{p} - C/H}{\mathbf{1}^\top \mathbf{p}} \frac{\mathbf{1}}{\mathbf{1}^\top \mathbf{p}},$$

413 we obtain the gradient of the reward, namely

$$414 \quad (4.6) \quad \nabla_{\mathbf{v}}R = \nabla_{\mathbf{v}}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \mathbf{v}} = -\boldsymbol{\lambda}^\top \frac{\partial \mathbf{A}}{\partial \mathbf{v}} \mathbf{p}.$$

416 The first equality in (4.6) follows from the fact that the Lagrangian coincides with
 417 the reward for all choices of $\boldsymbol{\lambda}$, since \mathbf{p} satisfies (4.1).

418 It is worth taking stock of the complete numerical optimisation routine. At each
 419 iteration i , we begin with the current control iterate \mathbf{v}^i and solve the forward prob-
 420 lem (4.1) to give \mathbf{p} , and thus R from (4.2). Then the adjoint system (4.5) is solved
 421 to give $\boldsymbol{\lambda}$. Note that the adjoint problem is of the same dimensionality as the for-
 422 ward problem, and thus of similar computational complexity. Finally, the gradient
 423 is calculated using the state and costate via (4.6). Thus, we have an efficient recipe
 424 for computing $R(\mathbf{v}^i)$ and $\nabla_{\mathbf{v}}R(\mathbf{v}^i)$ at each iteration. Given an initial guess \mathbf{v}_0 , the
 425 control drift is iteratively improved by the L-BFGS-B quasi-Newton method based
 426 on the values $R(\mathbf{v}^i)$ and $\nabla_{\mathbf{v}}R(\mathbf{v}^i)$ until some convergence criterion is reached. Fur-
 427 ther details and an extended discussion on the numerical convergence is provided in
 428 [Appendix B](#).

429 Before we delve into numerical results, we note that all of our simulations con-
 430 verged towards optimal drifts of type bang–bang: taking either their upper bound
 431 or lower bound but not intermediate values. This observation motivates the analysis
 432 in [section 4.2](#), where we turn to the calculus of variations to argue that, under mild
 433 assumptions, this observation is expected. The adjoint analysis undertaken in the
 434 continuum setting complements the discrete approach outlined above, demonstrating
 435 that the discrete approach is consistent with the continuum variational calculus.

436 **4.2. Bang–bang control.** Numerical results suggest that, when \mathbf{D} has no \mathbf{v} -
 437 dependence, the optimal controls are bang–bang, taking either their upper bound U
 438 or lower bound u but not intermediate values (recall from the problem statement (2.4)
 439 that we only consider bounded controls $\mathbf{v}(\mathbf{x}) \in [u, U]^d$). We offer a formal justification
 440 for this observation based on the calculus of variations. Consider an optimal drift $\mathbf{v}(\mathbf{x})$
 441 with respect to some objective function $R(\mathbf{v})$ of the form (3.15), with reward function
 442 $f(\mathbf{x})$. Importantly, we assume that f is not constant in any open ball in $\Omega \subset \mathbb{R}^d$. By
 443 way of contradiction, we assume that the optimal drift $\mathbf{v}(\mathbf{x}) \in (u, U)^d$ in some open
 444 ball $\Lambda \subset \Omega$, that is, the optimal drift lies strictly within its bounds in some small
 445 region.

446 Due to the optimality of \mathbf{v} , we expect from the first-order optimality conditions
 447 that the first variation of R with respect to the control drift, δR , vanishes (for some
 448 set of admissible perturbations we will describe). To help calculate the first variation,
 449 we introduce the (continuum) Lagrangian

$$450 \quad (4.7) \quad \mathcal{L} = R(\mathbf{v}) - \int_{\Omega} \lambda(\mathbf{x}) \left\{ \nabla \cdot [\mathbf{v}(\mathbf{x})\tilde{p}_0(\mathbf{x}) - \nabla \cdot (\mathbf{D}(\mathbf{x})\tilde{p}_0(\mathbf{x}))] \right\} d\mathbf{x},$$

452 where $\lambda(\mathbf{x})$ is an adjoint variable. Integrating by parts (and momentarily suppressing
 453 the dependence on \mathbf{x} , where clear, for brevity), we find that

$$454 \quad (4.8) \quad \int_{\Omega} \lambda \left\{ \nabla \cdot [\mathbf{v}\tilde{p}_0 - \nabla \cdot (\mathbf{D}\tilde{p}_0)] \right\} d\mathbf{x} = \int_{\Omega} (-\nabla\lambda \cdot \mathbf{v} - \nabla\nabla\lambda : \mathbf{D}) \tilde{p}_0 d\mathbf{x} \\ 455 \quad + \int_{\partial\Omega_0} \nabla\lambda \cdot (\mathbf{D}\mathbf{n})\tilde{p}_0 d\mathbf{x} + \int_{\partial\Omega_L} [\lambda(\mathbf{x}) - \lambda(\mathbf{x}^c)] \tilde{\phi}_0 \cdot \mathbf{n} d\mathbf{x}.$$

456 It follows, by taking the first variation of \mathcal{L} with respect to \tilde{p}_0 , that the costate
 457 equation is given by

$$458 \quad (4.9a) \quad \nabla\lambda(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x}) + \nabla\nabla\lambda(\mathbf{x}) : \mathbf{D}(\mathbf{x}) = -\frac{f(\mathbf{x})}{\int_{\Omega} \tilde{p}_0(\mathbf{z}) d\mathbf{z}} - \frac{\int_{\Omega} f(\mathbf{z})\tilde{p}_0(\mathbf{z}) d\mathbf{z} - C}{\left(\int_{\Omega} \tilde{p}_0(\mathbf{z}) d\mathbf{z}\right)^2},$$

459

460 subject to the boundary conditions

$$461 \quad (4.9b) \quad \nabla \lambda(\mathbf{x}) \cdot (\mathbf{D}(\mathbf{x})\mathbf{n}) = 0, \quad \mathbf{x} \in \partial\Omega_0,$$

$$462 \quad (4.9c) \quad \lambda(\mathbf{x}) = \lambda(\mathbf{x}^c), \quad \mathbf{x} \in \partial\Omega_L.$$

464 We identify the right-hand side of (4.9) as the continuum analog of the right-hand
 465 side of (4.5). Armed with the costate that satisfies (4.9) we may now analyse the
 466 first-order optimality conditions: the first variation of the objective R with respect to
 467 the control, in all admissible perturbation directions \mathbf{u} , vanishes, namely

$$468 \quad (4.10) \quad \delta R = \delta \mathcal{L} = \int_{\Omega} [\nabla \lambda(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x})] \tilde{p}_0(\mathbf{x}) \, d\mathbf{x} = 0.$$

470 An admissible perturbation \mathbf{u} is one with support in Λ such that, for small enough ϵ ,
 471 the perturbed control is admissible: $\mathbf{v}(\mathbf{x}) + \epsilon \mathbf{u}(\mathbf{x}) \in [u, U]$ for all $\mathbf{x} \in \Lambda$. It suffices
 472 to take bounded functions supported on any closed subset of Λ . Then, since \tilde{p}_0 is
 473 strictly positive on Λ (for non-degenerate diffusion the stationary distribution must
 474 be strictly positive on Ω since every location is reachable), by the fundamental lemma
 475 of the calculus of variations we deduce that $\nabla \lambda \equiv 0$ on Λ . Therefore, from (4.9a), we
 476 see that f must be constant on Λ , in contradiction to our non-stationarity assumption.
 477 The preceding argument is independent of the boundary conditions or inhomogeneities
 478 forcing \tilde{p}_0 , and therefore it remains valid in the case of cascading failures.

479 We have shown that the optimal drift must be bang–bang on the interior of the
 480 domain (assuming f is not stationary, similar to the condition in [41]). Note that
 481 when \mathbf{D} depends on the drift \mathbf{v} this result does not hold [41].

482 We now turn our attention to studying concrete numerical simulations.

483 **5. A tale of two processes.** The optimal strategy, a function in $\Omega \rightarrow [u, U]^d$,
 484 may be easily visualised for symmetric systems (i.e. dynamics and objectives not
 485 distinguishing between different processes) where $d = 2$. In this case, the symmetry
 486 ensures that the optimal drift satisfies $v_1(x, y) = v_2(y, x)$, and thus we may illustrate
 487 just $v_1 : \Omega \rightarrow [u, U]$, say, for each set of parameters, which is simply a surface over the
 488 two-dimensional domain $\Omega \subset \mathbb{R}^2$. Symmetry also allows for significant computational
 489 efficiency via dimension reduction.

490 Remarkably, numerical calculations of the optimal drift turn out to be bang–bang
 491 controls, that is, $v_i(\mathbf{x})$ takes either its upper bound U or lower bound u , but not in-
 492 termediate values. We demonstrated analytically in section 4.2 that this observation
 493 is a general characterisation of solutions for a large class of objective functions, noise
 494 couplings, and cascading-failure interdependence. This allows us to present optimal
 495 drift strategies $v_1 : \Omega \rightarrow \{u, U\}$ as contour plots over the domain Ω with the contours
 496 delineating between just two function values, U and u , representing regions of maxi-
 497 mum and minimum drift, which we shade and leave unshaded, respectively (further
 498 details in Appendix C). Since all contours on such a plot represent this single delin-
 499 eation, several such plots, for different parameter values, may be overlayed to capture
 500 how the contours (and the optimal drifts they illustrate) are affected by changing
 501 parameters. Unless specified otherwise, we use the forms

$$502 \quad (5.1) \quad \begin{aligned} d = 2, \quad L_1 = L_2 = U = 1, \quad u = C = 0, \\ \sigma = \sqrt{2}\mathbf{I}, \quad f(x, y) = -(1 - x - y)^2, \end{aligned}$$

504 where \mathbf{I} denotes the identity tensor and the associated diffusivity tensor is given by
 505 $\mathbf{D} = \mathbf{I}$. The reward function f in (5.1) encodes a pay-off for systems where the total

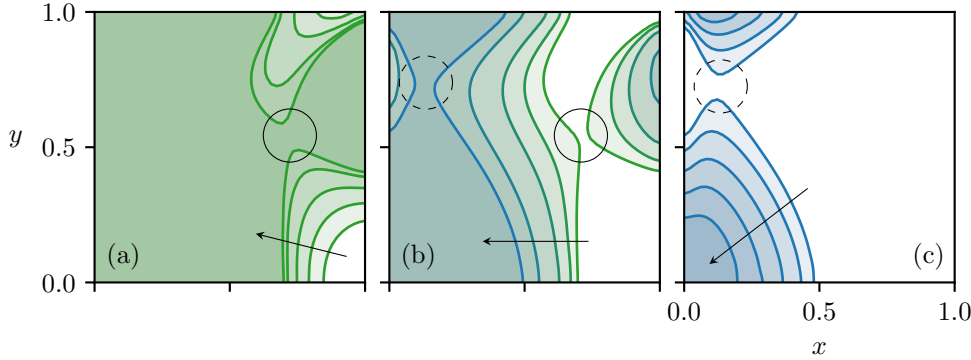


Fig. 2: Optimal drift $v_1(x, y)$ for various resetting costs: (a) $C \in \{-0.04, -0.03, -0.022, -0.016, -0.0135\}$, (b) $C \in \{-0.013, -0.008, 0, 0.015, 0.029\}$, (c) $C \in \{0.032, 0.045, 0.065, 0.09, 0.12\}$. Maximal drift ($v_1 = U$) is shaded, minimum drift ($v_1 = u$) is unshaded. Arrows show directions of increasing C . All plots share the same (x, y) axes. Circles of matching style highlight the topology-changing bifurcations. Changes in colour are for visual aid but of no quantitative relevance.

506 ‘performance’, that is, the sum of the states $x(t) + y(t)$, is near unity. We can think
 507 of this as a simple model rewarding a desired total system output, with both excess
 508 supply or deficient supply (i.e. excess demand) being undesirable. This induces a
 509 natural coupling between the processes.

510 First, we study the influence of the reset cost C . In Figure 2 we illustrate optimal
 511 strategies for various values of C , and observe how for negative reset costs (i.e. reset
 512 rewards) the state space is almost entirely shaded, as it is optimal to impose maximum
 513 drift except for corner regions (Figure 2a). The lower right is a region of high pay-
 514 off and resetting from $x = 1$ to $x = 0$ for small y results in jumping towards the
 515 low-pay-off origin. The ideal strategy is thus to drift only in the y direction (recall
 516 that $v_1(x, y) = v_2(y, x)$, therefore the y -drift is maximal in the lower right while the
 517 x -drift is minimal). This extends the duration in the high pay-off zone while driving
 518 the state to larger y such that a jump from $x = 1$ to $x = 0$ is further removed from
 519 the low pay-off zone. Similarly in the low-pay-off upper right, it is best to reset while
 520 avoiding proximity to $(x, y) = (1, 1)$ from which reset to the origin is likely.

521 As C is increased beyond some critical value near $C \approx -0.013$, the strategy
 522 bifurcates as the contours undergo a topological change: the regions of maximum drift
 523 are no longer connected (Figures 2a and 2b, solid circles). As C increases further,
 524 the region of maximum drift in the upper right vanishes, and at some critical value
 525 near $C \approx 0.03$ the remaining region of maximum drift on the left bifurcates, ‘pinching
 526 off’ into two (Figures 2b and 2c, dashed circles). Ultimately, for large values of C
 527 the optimal strategy is to impose minimum drift throughout most of the state space
 528 except for some corner regions (Figure 2c). In the lower left the pay-off is low and
 529 maximum drift still outweighs the increased risk of costly resets. In the upper left,
 530 it is better to drift in the x direction at the cost of slightly lower pay-off so that if a
 531 reset occurs the system is not near the origin.

532 The simplest (two-dimensional, symmetric) setting exemplifies how the decep-
 533 tively simple-looking problem belies a rich set of highly non-trivial optimal strategies.

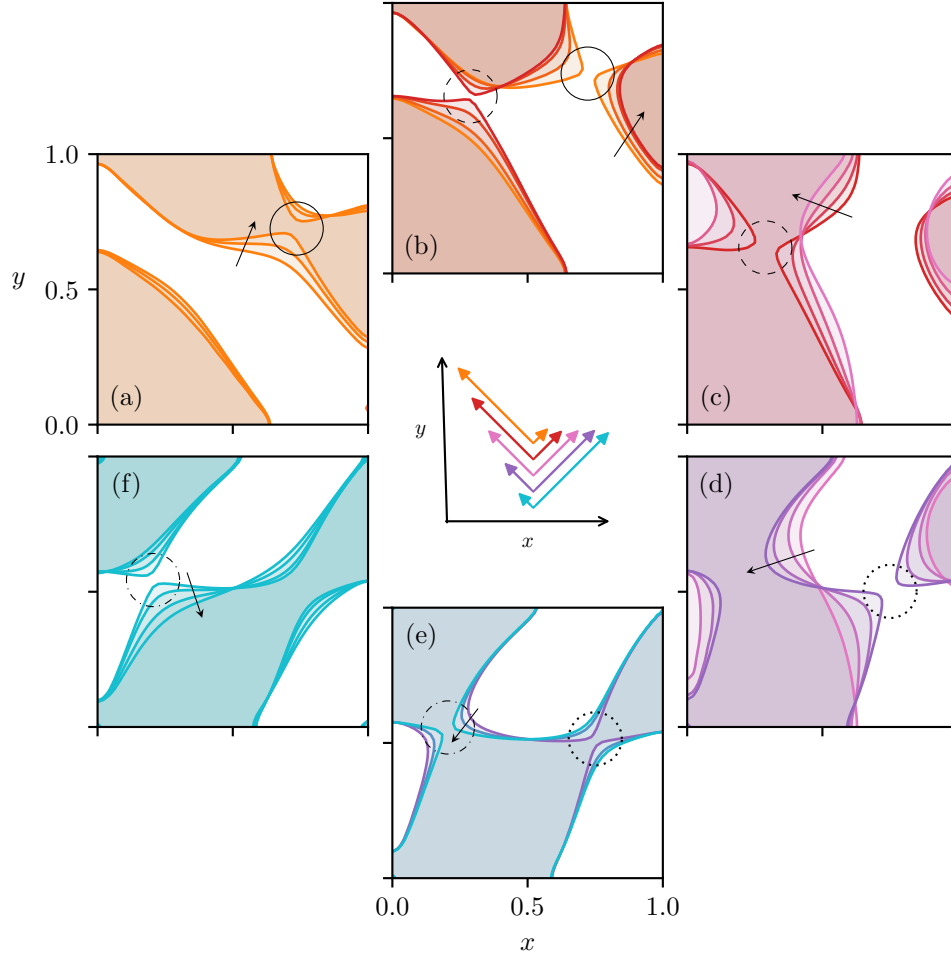


Fig. 3: Optimal drift $v_1(x, y)$ for the correlated noise (5.2) and various correlation parameters: (a) $s \in \{-0.2, -0.18, -0.165\}$, (b) $s \in \{-0.16, -0.14, -0.115, -0.105\}$, (c) $s \in \{-0.1, -0.07, -0.03, 0\}$, (d) $s \in \{0, 0.04, 0.09, 0.11\}$, (e) $s \in \{0.115, 0.125, 0.132\}$, (f) $s \in \{0.135, 0.145, 0.17, 0.2\}$. Maximal drift ($v_1 = U$) is shaded, minimum drift ($v_1 = u$) is unshaded. Black arrows show directions of increasing s . All plots share the same (x, y) axes. Circles of matching style highlight the topology-changing bifurcations. The axis in the figure centre shows the decomposition of the diffusion tensor (5.3) for $s \in \{-0.2, -0.05, 0, 0.05, 0.2\}$. Changes in colour are for visual aid but of no quantitative relevance.

534 Multiple bifurcations distinguish between different contour topologies, corresponding
 535 to the emergence of qualitatively different strategies.

536 Next, we study the influence of correlated noise via the one-dimensional family
 537 of diffusion coefficients:

$$538 \quad (5.2) \quad \sigma(s) = \sqrt{2} \begin{pmatrix} \sqrt{1-|s|} & \text{sgn}(s)\sqrt{|s|} \\ \text{sgn}(s)\sqrt{|s|} & \sqrt{1-|s|} \end{pmatrix},$$

539

540 for $s \in (-1/2, 1/2)$. The sign of s allows us to model both positively and negatively
 541 correlated noise, with $s = 0$ reducing to the case of $\mathbf{D} = \mathbf{I}$. The magnitude of
 542 noise (the expected square-deviation due to fluctuations) is preserved in each process:
 543 $\sum_j \sigma_{ij}^2 = 2$ for each row i , irrespective of s . Therefore, the ‘amount’ of noise in the
 544 system is preserved while only the correlations vary. The associated diffusivity tensor
 545 \mathbf{D} admits the spectral decomposition:

$$\begin{aligned}
 \mathbf{D} &= \begin{pmatrix} 1 & 2 \operatorname{sgn}(s) \sqrt{|s|(1-|s|)} \\ 2 \operatorname{sgn}(s) \sqrt{|s|(1-|s|)} & 1 \end{pmatrix} \\
 (5.3) \quad &= \mathbf{Z} \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \mathbf{Z}^\top, \quad \text{for } \mathbf{Z} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},
 \end{aligned}$$

548 and $\lambda_\pm = 1 \pm 2 \operatorname{sgn}(s) \sqrt{|s|(1-|s|)}$. We identify from \mathbf{Z} that the factorised diffusion
 549 directions are independent of s . The parameter s merely fixes the diffusivity in each
 550 direction, under the constraints that $\lambda_\pm > 0$ and $\lambda_+ + \lambda_- = 2$ (Figure 3 center).

551 In Figure 3 we plot the optimal drifts as s ranges from -0.2 to 0.2 . This range cap-
 552 tures all of the qualitative behaviour of the optimal drift, including four bifurcations.
 553 The key observation is that there are three persistent regions of maximum drift (in
 554 fact, these regions may also be identified in Figure 2): the lower left, upper right, and
 555 upper left. See Figure 3b where these three region are distinct. As s varies through the
 556 critical bifurcation points, these regions merge and divide. At the extremes, $s \approx \pm 0.2$,
 557 the regions’ boundaries align with the dominant direction of diffusion (see Figure 3
 558 center). For example, for $s \lesssim -0.2$, the upper regions merge to form a strip approxi-
 559 mately parallel to the boundary of the lower left region which is in turn aligned with
 560 the $(-1, 1)$ diagonal (Figure 3a). This alignment makes the y -drift strategy v_2 almost
 561 identical to the x -drift strategy v_1 (i.e. $v_2(x, y) = v_1(y, x) \approx v_1(x, y)$). Analogously
 562 for $s \gtrsim 0.2$ but in the transverse direction: the y -drift is approximately the comple-
 563 ment of the x -drift (one is maximised approximately where the other is minimised, see
 564 Figure 3f). These tightly coupled strategies are highly non-trivial, arising from the
 565 intricate interplay between the pay-off, the resetting, and the correlated dynamics.

566 Additional control constraints allow us to capture more realistically constrained
 567 scenarios. For example, if the first process was constructed to operate based only on
 568 its own state without knowledge of the full system, then we must relax the assump-
 569 tion of complete observability (algebraically, $v_1(x)$ has no y -dependence). How does
 570 that influence the optimal strategy of the second process $v_2(x, y)$? Alternatively, one
 571 process might depend on the full system state but nonetheless tailor its drift to favour
 572 certain regions of system states (for example, perhaps it greedily maximises profit in
 573 some subregion of state space which it favours despite this strategy being sub-optimal
 574 with respect to the global objective). What influence does this have on the optimal
 575 control of the second process?

576 Surprisingly, for all the cases we explored in two dimensions, the optimal strategy
 577 of the second process was nearly unchanged regardless of how we constrained the first
 578 process. This suggests that the objective landscape is effectively separable, rendering
 579 the optimal strategy of one process robust to sub-optimal strategies of the other
 580 process. This observation could be exploited for dimension reduction.

581 We have hitherto considered independent resets. However, when one system compo-
 582 nent fails it may induce the failure of other components. Such cascading failures
 583 may be accounted for in the general setting by resetting multiple processes upon the
 584 failure of some components, as detailed in section 3.1.

585 In Figure 4, we depict the optimal drifts for various reset costs C , in analogy with

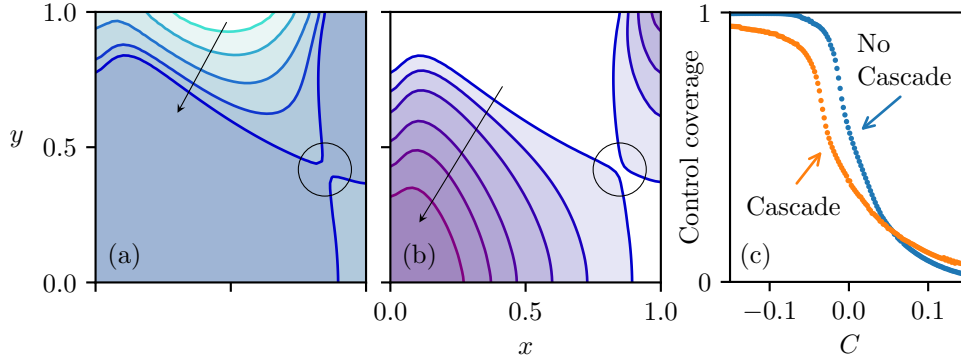


Fig. 4: (a,b) Optimal drift $v_1(x,y)$ for cascading failures and various resetting costs: (a) $C \in \{-0.1, -0.08, -0.06, -0.045, -0.0356\}$, (b) $C \in \{-0.0354, -0.02, 0, 0.03, 0.06, 0.1\}$. Maximal drift ($v_1 = U$) is shaded, minimum drift ($v_1 = u$) is unshaded. Black arrows show directions of increasing C . Both plots share the same (x,y) axes. Circles highlight the topology-changing bifurcations. Changes in colour are for visual aid but of no quantitative relevance. (c) Control coverage, the fraction of the state space on which there is maximum drift, defined in (5.4), versus reset cost C for the independent- and cascading-failure models.

586 [Figure 2](#) and using the same parameters but with cascading failures. The optimal
 587 control is qualitatively similar: two regions of maximum drift (lower left and upper
 588 right) dominate the strategies, whose extents increases for decreasing resetting costs.
 589 A third region in the upper left, as in [Figure 2](#), is recoverable in a different region of
 590 parameter space.

591 It is interesting to study the fraction of the state space on which there is maximum
 592 drift, which we call the control coverage. Since the optimal solutions are bang–bang
 593 and symmetric, this is captured by the quantity

$$594 \quad (5.4) \quad \frac{\int_{\Omega} v_1(\mathbf{x}) - u \, d\mathbf{x}}{(U - u)|\Omega|}.$$

596 When comparing the control coverage for the independent-failure model ([Figure 2](#))
 597 and the cascading-failure model ([Figure 4](#)), one notable difference is that, for $C \lesssim 0.06$,
 598 the control coverage is less extensive with cascading failures ([Figure 4c](#)). This aligns
 599 with the intuition that failure is more costly when it unleashes a cascade of component
 600 resets, and thus an optimal strategy might be less willing to incur reset. Curiously,
 601 above this critical reset cost $C \gtrsim 0.06$ the phenomenon appears to invert: the control
 602 coverage associated with cascading failures is (slightly) more extensive. Despite the
 603 discrepancy being somewhat small, we conjecture that it is not a numerical artefact,
 604 but instead reflects the fact that, with cascading failures, reset sends all trajectories
 605 to the low-pay-off origin making a more extensive maximum drift in this vicinity more
 606 optimal.

607 These simulations provide a taste of the mosaic complexity of drift strategies
 608 that optimally steer noisy complex systems. This richness is born out of the delicate
 609 balance needed in the trade-off between high performance and system crash while
 610 accounting for the interdependence of system components, in terms of their dynamics,

611 knock-on failure, and coupling in the objective.

612 **6. Conclusions.** We introduce a complex system of resetting stochastic processes, seeking to optimally control the system with respect to an objective. We
 613 study the interdependence of the processes due to the objective function, as well as
 614 interaction via correlated noise and cascading failures. Formulating in terms of the
 615 PDE governing the process law, we obtain a PDE-constrained optimisation problem
 616 via the application of renewal theory and asymptotic analysis. An adjoint analysis
 617 shows that the optimal controls are of type bang–bang. Numerical simulations,
 618 also leveraging the adjoint perspective, reveal bifurcation in the solution structure as
 619 system parameters are varied.
 620

621 The present formulation may be readily extended. More general reset laws (e.g.
 622 nonrectangular state spaces) could capture more sophisticated failure models. Asym-
 623 metries abound in reset costs, reset boundary, control constraints, and diffusion coeffi-
 624 cients. Various other objective function forms may be useful in different applications.
 625

626 We explore only the infinite time horizon. The large-time perspective confers
 627 an enormous advantage as it allows us to consider an elliptic problem rather than
 628 a parabolic problem, and thereby achieves dimension reduction by eliminating time,
 629 while remaining applicable over large time horizons. Moreover, the renewal approach
 630 makes analysing the reset count tractable, which would otherwise be a highly non-
 631 trivial complication. The analysis for an objective on a finite-time horizon cannot rely
 632 on the large-time limit, and therefore the formulation remains in the parabolic setting
 633 and the renewal theory is not directly applicable. Nevertheless, studying the temporal
 634 convergence towards the steady state could justify the use of the infinite-time case
 635 as a relevant model even over finite time horizons. This problem requires care in the
 636 numerical scheme to guarantee conservation and stability.

637 The finite-difference schemes implemented are subject to the curse of dimension-
 638 ality, however, machine-learning-based approaches make high dimensional systems
 639 tractable [31], where emergent structure and long-range properties typical of complex
 640 systems can be expected.

640 Appendix A. Discretisation scheme.

641 In this appendix we detail the discretisation employed for the PDEs describing
 642 the independent-failure model (3.16) and the cascading-failure model (3.18). The
 643 dependent variable is \tilde{p}_0 , but throughout this section when we refer to discretised
 644 quantities, we replace this with p to simplify the notation. This is emphasised so as
 645 not to be confused with the original use of p in the section 3.

646 Since the domain Ω is rectangular, the problem lends itself to a consistent finite-
 647 difference discretisation. On a grid of equal spacing h_i in each dimension $i = 1, \dots, d$,
 648 we consider the discrete domain indexed by $\mathbf{j} \in \Gamma \subset \mathbb{Z}^d$ where

$$649 \quad (A.1) \quad \Gamma = \{(j_1, \dots, j_d)^\top \mid 0 \leq j_i \leq N_i \text{ for all } i = 1, \dots, d\},$$

651 corresponding to the points $\mathbf{x}_j = (j_1 h_1, \dots, j_d h_d) \in \Omega$ where $N_i h_i \approx L_i$. We denote
 652 spatial evaluation by subscripting: $p_j \approx p(\mathbf{x}_j) = p(j_1 h_1, \dots, j_d h_d)$. All numerical
 653 results in this work use $h_i = 0.02$ and $N_i = 51$.

654 It is useful to introduce the boundary sets

$$655 \quad (A.2) \quad \begin{aligned} \partial\Gamma_0 &= \{\mathbf{j} \in \Gamma \mid j_i = 0 \text{ for some } i\}, \\ \partial\Gamma_N &= \{\mathbf{j} \in \Gamma \mid j_i = N_i \text{ for some } i\}, \\ \overset{\circ}{\Gamma} &= \Gamma \setminus (\partial\Gamma_N \cup \partial\Gamma_0). \end{aligned}$$

656

657 The sets $\partial\Gamma_0$ and $\partial\Gamma_N$ are the discrete counterparts of the continuum $\partial\Omega_0$ and $\partial\Omega_L$,
 658 respectively. Whereas in the continuum setting the intersection $\partial\Omega_0 \cap \partial\Omega_L$ is a null
 659 set, this is no longer the case in the discrete setting: different boundary conditions
 660 are imposed at these two boundaries (see (3.16b) and (3.16c)), thus, in the discrete
 661 setting, we need to decide which of the two boundary conditions are applied to points
 662 in $\partial\Gamma_0 \cap \partial\Gamma_N$. We choose to assign points in the intersection to the N -boundary, $\partial\Gamma_N$,
 663 by defining $\widehat{\partial\Gamma}_0 := \partial\Gamma_0 \setminus \partial\Gamma_N$, and calling $\widehat{\partial\Gamma}_0$ the 0-boundary.

664 Another issue concerns the appropriate boundary conditions for points on two
 665 (or more) 0-boundaries, that is, $\mathbf{j} \in \widehat{\partial\Gamma}_0$ such that $j_i = j_k = 0$ for some $i \neq k$.
 666 To deal with this, we distinguish between \mathbf{j} for which a single index is zero, and \mathbf{j}
 667 for which more than one index is zero, which we denote $\widehat{\partial\Gamma}_0^1$ and $\widehat{\partial\Gamma}_0^{>1}$, respectively:
 668 $\widehat{\partial\Gamma}_0 = \widehat{\partial\Gamma}_0^1 \cup \widehat{\partial\Gamma}_0^{>1}$. We may set the solution value in $\widehat{\partial\Gamma}_0^{>1}$ arbitrarily, as we will
 669 design a consistent scheme whose support does not enter $\widehat{\partial\Gamma}_0^{>1}$.

670 We implemented the schemes detailed in Refs. [27, 30, 55], and obtained the most
 671 stable convergence with the asymmetric scheme from Ref. [30], which we adopted in
 672 the interior of the domain, $\mathbf{j} \in \overset{\circ}{\Gamma}$, via

$$673 \quad (\text{A.3a}) \quad \sum_{k=1}^d \frac{(v_k p)_{\mathbf{j}} - (v_k p)_{\mathbf{j}-\mathbf{e}_k}}{h_k} - \frac{(\phi_k)_{\mathbf{j}+\mathbf{e}_k/2} - (\phi_k)_{\mathbf{j}-\mathbf{e}_k/2}}{h_k} = 0,$$

674 where

$$675 \quad (\text{A.3b}) \quad (\phi_k)_{\mathbf{j}+\mathbf{e}_k/2} = \frac{(D_{kk}p)_{\mathbf{j}+\mathbf{e}_k} - (D_{kk}p)_{\mathbf{j}}}{h_k}$$

$$676 \quad + \sum_{k \neq i=1}^d \frac{(D_{ki}p)_{\mathbf{j}+\mathbf{e}_i} + (D_{ki}p)_{\mathbf{j}+\mathbf{e}_k+\mathbf{e}_i} - (D_{ki}p)_{\mathbf{j}-\mathbf{e}_i} - (D_{ki}p)_{\mathbf{j}+\mathbf{e}_k-\mathbf{e}_i}}{4h_i},$$

677 with \mathbf{e}_i denoting the i th basis vector (of zero entries except the i th entry of one)
 678 and denoting the diffusivity tensor $\mathbf{D} = (D_{ik})$ in the Cartesian frame, for all \mathbf{j} such
 679 that the scheme is supported on $\overset{\circ}{\Omega} \setminus \widehat{\partial\Gamma}_0^{>1}$. For $\mathbf{j} \in \overset{\circ}{\Omega}$ where scheme (A.3) relies of
 680 evaluation in $\widehat{\partial\Gamma}_0^{>1}$, we use the scheme

$$681 \quad (\text{A.4}) \quad \sum_{k=1}^d \frac{(v_k p)_{\mathbf{j}} - (v_k p)_{\mathbf{j}-\mathbf{e}_k}}{h_k}$$

$$682 \quad - \sum_{i=1}^d \frac{1}{h_k} \left[\frac{(D_{ik}p)_{\mathbf{j}+\mathbf{e}_k} - (D_{ik}p)_{\mathbf{j}-\mathbf{e}_i+\mathbf{e}_k}}{h_i} - \frac{(D_{ik}p)_{\mathbf{j}} - (D_{ik}p)_{\mathbf{j}-\mathbf{e}_i}}{h_i} \right] = 0.$$

683 Since we only *subtract* from at most one index in the stencil (A.4), and since $j_k \geq 1$
 684 for all $\mathbf{j} \in \overset{\circ}{\Gamma}$, no point in the interior includes any indices from $\widehat{\partial\Gamma}_0^{>1}$ in its stencil.

685 Both schemes (A.3) and (A.4) are consistent. Moreover, since we will con-
 686 sider non-negative drifts in the optimisation problem, the first-order terms in both
 687 schemes (A.3) and (A.4) are of upwinded form.

688 We now detail the discrete boundary conditions. For the N -boundaries, we set

$$689 \quad (\text{A.5}) \quad p_{\mathbf{j}} = 0, \quad \mathbf{j} \in \partial\Gamma_N.$$

692 For $\mathbf{j} \in \widehat{\partial\Gamma}_0^1$, we set

$$\begin{aligned}
693 \quad (v_k p)_j &- \frac{(D_{kk}p)_{j+e_k} - (D_{kk}p)_j}{h_k} - \sum_{k \neq i=1}^d \frac{(D_{ik}p)_{j+e_i} - (D_{ik}p)_{j-e_i}}{2h_i} \\
694 \quad &= (v_k)_{j^k} p_{j^k} - \frac{(D_{kk}p)_{j^k} - (D_{kk}p)_{j^k - e_k}}{h_k} - \sum_{k \neq i=1}^d \frac{(D_{ik}p)_{j^k+e_i} - (D_{ik}p)_{j^k-e_i}}{2h_i} \\
\text{(A.6)} \quad &= \frac{(D_{kk}p)_{j^k - e_k}}{h_k},
\end{aligned}$$

697 for the single k such that $j_k = 0$, with $\mathbf{j}^k = (j_1, \dots, j_{k-1}, N_k, j_{k+1}, \dots, j_d)$ denoting
698 the point opposite \mathbf{j} , and the last equality following from (A.5) since $\mathbf{j}^k \in \partial\Gamma_N$ and
699 $\mathbf{j}^k \pm \mathbf{e}_i \in \partial\Gamma_N$ for all $i \neq k$.

700 At this stage, we have $\prod_{i=1}^d (N_i + 1)$ unknowns and equations, which we write in
701 matrix form via

$$703 \quad \text{(A.7)} \quad \mathbf{M}\mathbf{p} = \mathbf{0},$$

704 where $\mathbf{p} = (p_j)$ is an $\prod_{i=1}^d (N_i + 1)$ -dimensional vector and \mathbf{M} is the finite-difference
705 matrix. Since the system (A.7) is linear and homogeneous, any non-trivial solution
706 may be multiplied by any nonzero scalar to yield another solution. This indeterminacy
707 is resolved when the constraint (3.16d) is imposed, which may be discretised via

$$\begin{aligned}
\text{(A.8)} \quad &\sum_{\mathbf{j} \in \widehat{\partial\Gamma}_0^1} \left[(v_k p)_j - \frac{(D_{kk}p)_{j+e_k} - (D_{kk}p)_j}{h_k} - \sum_{k \neq i=1}^d \frac{(D_{ik}p)_{j+e_i} - (D_{ik}p)_{j-e_i}}{2h_i} \right] \frac{H}{h_k} = 1,
\end{aligned}$$

710 where $k = k(\mathbf{j})$ denotes the index for which $j_k = 0$ and

$$711 \quad \text{(A.9)} \quad H = \prod_{i=1}^d h_i.$$

713 The right-hand side of (A.8) is positive since we have taken the *inward*-facing normal
714 flux on the left-hand side. Denoting the discrete condition (A.8) by $\mathbf{L}^\top \mathbf{p} = 1$, we
715 impose the condition by solving the forward system

$$716 \quad \text{(A.10)} \quad \begin{pmatrix} \mathbf{M} & \mathbf{L} \\ \mathbf{L}^\top & 0 \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ q \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix},$$

718 where $\mathbf{0}$ denotes vectors of zeros and ones, respectively, and q plays the role of a
719 Lagrange multiplier.

720 The cascading failure problem (3.18) is solved using the same scheme, without
721 the inhomogeneous condition (A.8), but instead setting the entry, corresponding to
722 the discrete location of the delta function, equal to $\prod_{i=1}^d 1/h_i$ on the right-hand side
723 of (A.7). Results using both of these schemes were compared with the original scheme
724 in dimension one (where the independent and cascading failure problems coincide, but
725 the schemes differ), and both agreed with previous results from [41].

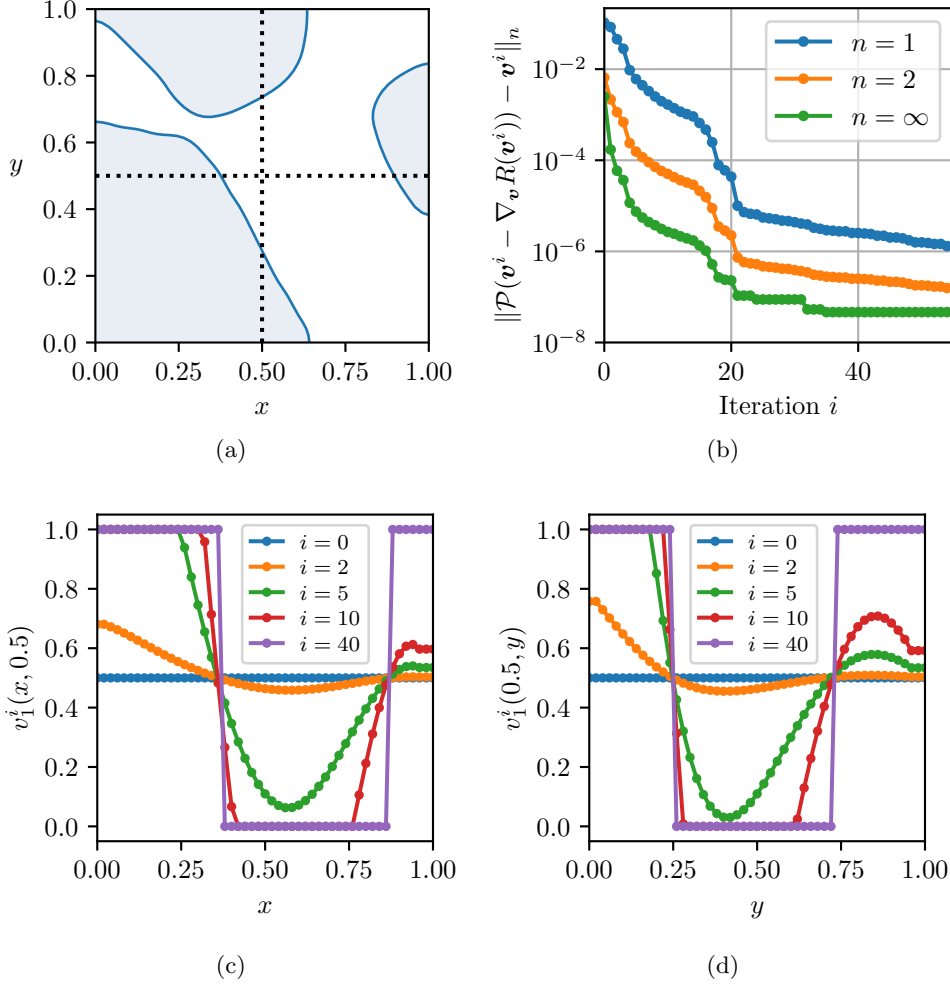


Fig. 5: Results of the numerical optimisation scheme for the independent-failure model and correlated noise (5.2) with $s = -0.11$. At iteration i , the controllable drift is denoted by $v^i(x, y)$ (and its discrete counterpart by \mathbf{v}_i). (a) Optimal drift $v_1^{55}(x, y)$ for the final iterate $i = 55$, where the maximum drift ($v_1^{55} = U$) is shaded and minimum drift ($v_1^{55} = u$) is unshaded. The dotted black lines show slices of the (x, y) -domain on which we plot the drifts v_1^i in panels (c) and (d). (b) Convergence of the ℓ^n norm of the projected gradient. (c,d) The drift $v_1^i(x, y)$ on the slices depicted in panel (a) at different iterations i of the optimisation algorithm.

726 This completes the discretisation of the governing PDEs (3.16) and (3.18).

727 Appendix B. Numerical convergence.

728 In this appendix, we demonstrate the numerical convergence achieved by the L-
 729 BFGS-B algorithm based on the adjoint approach described in section 4.1 and the
 730 discrete scheme detailed in Appendix A. We choose the case of independent failures
 731 (maximising R in (3.15) subject to (3.16)) and correlated noise, corresponding to the

732 results depicted in [Figure 3](#).

733 We start with the drift \mathbf{v}^i after $i = 55$ iterations in [Figure 5a](#). Our aim is to
 734 demonstrate that the objective evaluated with this drift, $R(\mathbf{v}^i)$, is (approximately) a
 735 local maximum, and thus this drift is (approximately) optimal. Unconstrained first-
 736 order optimality conditions require that the gradient vanishes at a local optimum.
 737 In our case, the control drift is bounded $\mathbf{v}(\mathbf{x}) \in [u, U]^d$, in which case the projected
 738 gradient plays the role of the gradient as we now describe.

739 The projection \mathcal{P} maps each component of the drift vector \mathbf{v}^i onto the bounds
 740 $[u, U]$, that is, for each component z in \mathbf{v}^i

$$741 \quad (\text{B.1}) \quad z \xrightarrow{\mathcal{P}} \begin{cases} 0, & z < u, \\ z, & z \in (u, U), \\ 1, & z > U. \end{cases}$$

742

743 The projected gradient of $R(\mathbf{v}^i)$ is given by the difference between the projection of
 744 the control perturbed by the negative gradient, $\mathcal{P}(\mathbf{v}^i - \nabla_{\mathbf{v}}R(\mathbf{v}^i))$, and the control, \mathbf{v}^i .
 745 To see why the first-order optimality conditions are simply the vanishing projected
 746 gradient, consider a local optimum at \mathbf{v}^* and separate those components strictly
 747 within the bounds from those at the bounds. The gradient $\nabla_{\mathbf{v}}R(\mathbf{v}^*)$ must vanish
 748 in the components where \mathbf{v}^* is within the bounds (otherwise the objective could be
 749 increased by a small admissible perturbation in the direction of the negative gradient,
 750 thus the point is not an optimum):

$$751 \quad (\text{B.2}) \quad v_k^* \in (u, U) \implies -[\nabla_{\mathbf{v}}R(\mathbf{v}^*)]_k = 0.$$

753 Therefore, these components must also vanish in the projected gradient. For compo-
 754 nents of \mathbf{v}^* at a bound, $v_k^* \in \{u, U\}$, the negative gradient does not point inward to
 755 the interval $[u, U]$ (otherwise the objective could be increased by a small admissible
 756 perturbation):

$$757 \quad (\text{B.3}) \quad v_k^* = u \implies -[\nabla_{\mathbf{v}}R(\mathbf{v}^*)]_k \leq 0, \quad v_k^* = U \implies -[\nabla_{\mathbf{v}}R(\mathbf{v}^*)]_k \geq 0.$$

759 Therefore, the projection of the perturbed control remains at the same boundary for
 760 these components, thus vanishing in the projected gradient. Ultimately, we find that

$$761 \quad (\text{B.4}) \quad \mathcal{P}(\mathbf{v}^* - \nabla_{\mathbf{v}}R(\mathbf{v}^*)) - \mathbf{v}^* = \mathbf{0},$$

762

763 as we set out to demonstrate.

764 In [Figure 5b](#) we show the ℓ^n norm of the projected gradient at each iteration.
 765 There projected gradient rapidly decays until iteration $i \approx 20$ after which time the
 766 convergence is slower. In [Figures 5c](#) and [5d](#) we show the control $v_1^i(x, y)$, that is,
 767 the control of the first process at iteration i , along one-dimensional slices of the
 768 domain corresponding to the black dotted lines in [Figure 5a](#). The rapid convergence
 769 is evident: by iteration $i = 10$ the drift closely matches its final configuration. The
 770 notable exceptions are regions near the L -boundary, where convergence is slower due
 771 to the relative insensitivity of the objective function to the drift. Nonetheless, from
 772 the very earliest iterations we see that the $(u + U)/2$ level set is clearly established
 773 and preserved, motivating its choice as the contour delineating regions of maximum
 774 and minimum optimal drift, as discussed in [Appendix C](#).

775 In this example, despite an intricate optimal drift topology close to a bifurcation
 776 and strong anisotropy in the diffusion operator, we observe good convergence to a

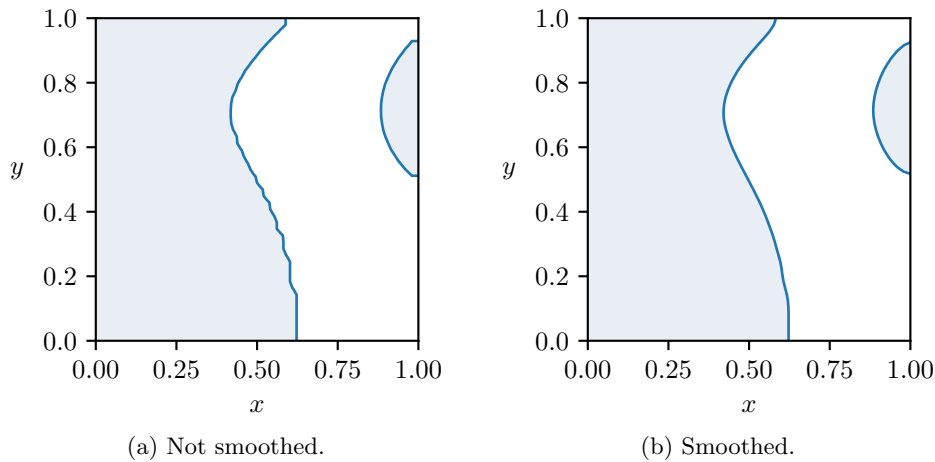


Fig. 6: Optimal drift strategy $v_1(x, y)$ (shaded regions show maximum drift, $v_1 = U$, unshaded regions show minimum drift, $v_1 = u$) for defaults parameters given in (5.1) with and without smoothing.

777 local optimum. The convergence criterion used to produce all the other figures was a
 778 threshold of 2×10^{-6} for the ℓ^∞ -norm of the projected gradient. The initial control
 779 guess used in the optimisation was the uniform $\mathbf{v}^0(\mathbf{x}) \equiv (u + U)/2$. All calculations
 780 were performed on a standard laptop computer with simulation times on the order of
 781 seconds per optimal drift.

782 Appendix C. Contour plots.

783 As detailed in section 4.2, the optimal drifts are of bang–bang type, taking either
 784 their upper bound or lower bound but not intermediate values. Therefore, to plot an
 785 optimal drift field $v_i(\mathbf{x})$ it suffices to draw the region(s) of maximum and minimum
 786 drift. We choose to draw the contour(s) of the $(u + U)/2$ level set, obtained by
 787 interpolation of the drift field $v_i(\mathbf{x})$, which we consider to separate these two regions
 788 and shade only the region of maximum drift. Since the drift field is given only at grid
 789 points, the resulting contour tends to be unnaturally rectangular (Figure 6a). We
 790 thus choose to smooth the drift field with a Gaussian filter, resulting in a smoother
 791 contour (Figure 6b). In Figures 2 to 4 we use various colours to plot the contours
 792 and shade the regions of maximal drift. These are merely for visual aid to show the
 793 continuity of the optimal drift with respect to the changing parameter values away
 794 from a bifurcation, but have no quantitative relevance.

795

REFERENCES

- 796 [1] R. ALBERT, H. JEONG, AND A.-L. BARABÁSI, *Error and attack tolerance of complex networks*,
 797 Nature, 406 (2000), pp. 378–382, <https://doi.org/10.1038/35019019>.
 798 [2] S. V. ANULOVA, H. MAI, AND A. Y. VERETENNIKOV, *On iteration improvement for averaged*
 799 *expected cost control for one-dimensional ergodic diffusions*, SIAM J. Control Optim., 58
 800 (2020), pp. 2312–2331, <https://doi.org/10.1137/19M1271944>.
 801 [3] A. ARAPOSTATHIS, V. S. BORKAR, AND M. K. GHOSH, *Ergodic Control of Diffusion Processes*,
 802 Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2011,
 803 <https://doi.org/10.1017/CBO9781139003605>.

- 804 [4] R. R. BARTON, *Computing forward difference derivatives in engineering optimization*, Eng.
805 Optimiz., 20 (1992), pp. 205–224, <https://doi.org/10.1080/03052159208941281>.
- 806 [5] R. BARTOSZYNSKI, W. J. BÜHLER, W. CHAN, AND D. K. PEARL, *Population processes under*
807 *the influence of disasters occurring independently of population size*, J. Math. Biol., 27
808 (1989), pp. 167–178, <https://doi.org/10.1007/BF00276101>.
- 809 [6] I. BEN-ARI AND R. G. PINSKY, *Ergodic behavior of diffusions with random jumps from the*
810 *boundary*, Stoch. Proc. Appl., 119 (2009), pp. 864–881, [https://doi.org/10.1016/j.spa.2008.](https://doi.org/10.1016/j.spa.2008.05.002)
811 [05.002](https://doi.org/10.1016/j.spa.2008.05.002).
- 812 [7] O. BÉNICHOU, M. MOREAU, P.-H. SUET, AND R. VOITURIEZ, *Intermittent search process and*
813 *teleportation*, J. Chem. Phys., 126 (2007), p. 234109, <https://doi.org/10.1063/1.2741516>.
- 814 [8] S. BOCCALETTI, V. LATORA, Y. MORENO, M. CHAVEZ, AND D.-U. HWANG, *Complex networks:*
815 *Structure and dynamics*, Phys. Rep., 424 (2006), pp. 175–308, [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.physrep.2005.10.009)
816 [physrep.2005.10.009](https://doi.org/10.1016/j.physrep.2005.10.009).
- 817 [9] P. J. BROCKWELL, J. GANI, AND S. I. RESNICK, *Birth, immigration and catastrophe processes*,
818 Adv. Appl. Probab., 14 (1982), pp. 709–731, <https://doi.org/10.2307/1427020>.
- 819 [10] S. V. BULDYREV, R. PARSHANI, G. PAUL, H. E. STANLEY, AND S. HAVLIN, *Catastrophic cascade*
820 *of failures in interdependent networks*, Nature, 464 (2010), pp. 1025–1028, [https://doi.org/](https://doi.org/10.1038/nature08932)
821 [10.1038/nature08932](https://doi.org/10.1038/nature08932).
- 822 [11] X. CHAO AND Y. ZHENG, *Transient analysis of immigration birth–death processes with total*
823 *catastrophes*, Probab. Eng. Inform. Sc., 17 (2003), pp. 83–106, [https://doi.org/10.1017/](https://doi.org/10.1017/S0269964803171057)
824 [S0269964803171057](https://doi.org/10.1017/S0269964803171057).
- 825 [12] B. DE BRUYNE, J. RANDON-FURLING, AND S. REDNER, *Optimization in first-passage resetting*,
826 Phys. Rev. Lett., 125 (2020), p. 050602, <https://doi.org/10.1103/PhysRevLett.125.050602>.
- 827 [13] B. DE BRUYNE, J. RANDON-FURLING, AND S. REDNER, *Optimization and growth in first-*
828 *passage resetting*, J. Stat. Mech.-Theory E., 2021 (2021), p. 013203, [https://doi.org/10.](https://doi.org/10.1088/1742-5468/abcd33)
829 [1088/1742-5468/abcd33](https://doi.org/10.1088/1742-5468/abcd33).
- 830 [14] I. DOBSON, B. A. CARRERAS, V. E. LYNCH, AND D. E. NEWMAN, *Complex systems analysis*
831 *of series of blackouts: Cascading failure, critical points, and self-organization*, Chaos, 17
832 (2007), p. 026103, <https://doi.org/10.1063/1.2737822>.
- 833 [15] S. N. DOROGOVTSSEV AND J. F. F. MENDES, *Evolution of networks*, Adv. Phys., 51 (2002),
834 pp. 1079–1187, <https://doi.org/10.1080/00018730110112519>.
- 835 [16] I. ELIAZAR, T. KOREN, AND J. KLAFTER, *Searching circular DNA strands*, J. Phys.-Condens.
836 Mat., 19 (2007), p. 065140, <https://doi.org/10.1088/0953-8984/19/6/065140>.
- 837 [17] M. R. EVANS AND S. N. MAJUMDAR, *Diffusion with optimal resetting*, J. Phys. A-Math. Theor.,
838 44 (2011), p. 435001, <https://doi.org/10.1088/1751-8113/44/43/435001>.
- 839 [18] M. R. EVANS AND S. N. MAJUMDAR, *Diffusion with stochastic resetting*, Phys. Rev. Lett., 106
840 (2011), p. 160601, <https://doi.org/10.1103/PhysRevLett.106.160601>.
- 841 [19] M. R. EVANS AND S. N. MAJUMDAR, *Diffusion with resetting in arbitrary spatial dimension*,
842 J. Phys. A-Math. Theor., 47 (2014), p. 285001, [https://doi.org/10.1088/1751-8113/47/28/](https://doi.org/10.1088/1751-8113/47/28/285001)
843 [285001](https://doi.org/10.1088/1751-8113/47/28/285001).
- 844 [20] M. R. EVANS AND S. N. MAJUMDAR, *Run and tumble particle under resetting: a renewal ap-*
845 *proach*, J. Phys. A-Math. Theor., 51 (2018), p. 475003, [https://doi.org/10.1088/1751-8121/](https://doi.org/10.1088/1751-8121/aae74e)
846 [aae74e](https://doi.org/10.1088/1751-8121/aae74e).
- 847 [21] M. R. EVANS, S. N. MAJUMDAR, AND G. SCHEHR, *Stochastic resetting and applications*, J.
848 Phys. A-Math. Theor., 53 (2020), p. 193001, <https://doi.org/10.1088/1751-8121/ab7cfe>.
- 849 [22] E. GELENBE, *Search in unknown random environments*, Phys. Rev. E, 82 (2010), p. 061112,
850 <https://doi.org/10.1103/PhysRevE.82.061112>.
- 851 [23] M. B. GILES AND N. A. PIERCE, *An introduction to the adjoint approach to design*, Flow
852 Turbul. Combust., 65 (2000), pp. 393–415, <https://doi.org/10.1023/A:1011430410075>.
- 853 [24] M. GIRVAN AND M. E. J. NEWMAN, *Community structure in social and biological networks*, P.
854 Natl. Acad. Sci. USA, 99 (2002), pp. 7821–7826, <https://doi.org/10.1073/pnas.122653799>.
- 855 [25] C. P. GOMES AND B. SELMAN, *Algorithm portfolios*, Artif. Intell., 126 (2001), pp. 43–62, [https://doi.org/10.1016/S0004-3702\(00\)00081-3](https://doi.org/10.1016/S0004-3702(00)00081-3).
- 856 [26] C. P. GOMES, B. SELMAN, N. CRATO, AND H. KAUTZ, *Heavy-tailed phenomena in satisfiability*
857 *and constraint satisfaction problems*, J. Autom. Reasoning, 24 (2000), pp. 67–100, <https://doi.org/10.1023/A:1006314320276>.
- 858 [27] D. GREENSPAN AND P. C. JAIN, *On non negative difference analogues of elliptic differential*
859 *equations*, J. Frankl. Inst., 279 (1965), pp. 360–365, [https://doi.org/10.1016/0016-0032\(65\)](https://doi.org/10.1016/0016-0032(65)90423-0)
860 [90423-0](https://doi.org/10.1016/0016-0032(65)90423-0).
- 861 [28] I. GRIGORESCU AND M. KANG, *Brownian motion on the figure eight*, J. Theor. Probab., 15
862 (2002), pp. 817–844, <https://doi.org/10.1023/A:1016232201962>.
- 863 [29] I. GRIGORESCU AND M. KANG, *Ergodic properties of multidimensional brownian motion with*
864
865

- 866 *rebirth*, Electron. J. Probab., 12 (2007), pp. 1299–1322, <https://doi.org/10.1214/EJP.v12-450>.
- 867
- 868 [30] S. GÜNTER, Q. YU, J. KRÜGER, AND K. LACKNER, *Modelling of heat transport in magnetised*
- 869 *plasmas using non-aligned coordinates*, J. Comput. Phys., 209 (2005), pp. 354–370, <https://doi.org/10.1016/j.jcp.2005.03.021>.
- 870
- 871 [31] J. HAN, A. JENTZEN, AND E. WEINAN, *Solving high-dimensional partial differential equations*
- 872 *using deep learning*, P. Natl. Acad. Sci. USA, 115 (2018), pp. 8505–8510, <https://doi.org/10.1073/pnas.1718942115>.
- 873
- 874 [32] P. HOLME, B. J. KIM, C. N. YOON, AND S. K. HAN, *Attack vulnerability of complex networks*,
- 875 Phys. Rev. E, 65 (2002), p. 056109, <https://doi.org/10.1103/PhysRevE.65.056109>.
- 876 [33] B. KRISHNA KUMAR AND D. ARIVUDAINAMBI, *Transient solution of an M/M/1 queue with*
- 877 *catastrophes*, Comput. Math. Appl., 40 (2000), pp. 1233–1240, [https://doi.org/10.1016/S0898-1221\(00\)00234-0](https://doi.org/10.1016/S0898-1221(00)00234-0).
- 878
- 879 [34] B. KRISHNA KUMAR AND S. PAVAI MADHESWARI, *Transient behaviour of the M/M/2 queue with*
- 880 *catastrophes*, Statistica, 62 (2002), pp. 129–136, <https://doi.org/10.6092/issn.1973-2201/1287>.
- 881
- 882 [35] V. KUMAR, O. SADEKAR, AND U. BASU, *Active brownian motion in two dimensions under sto-*
- 883 *chastic resetting*, Phys. Rev. E, 102 (2020), p. 052129, <https://doi.org/10.1103/PhysRevE.102.052129>.
- 884
- 885 [36] L. KUSMIERZ, S. N. MAJUMDAR, S. SABHAPANDIT, AND G. SCHEHR, *First order transition*
- 886 *for the optimal search time of Lévy flights with resetting*, Phys. Rev. Lett., 113 (2014),
- 887 p. 220602, <https://doi.org/10.1103/PhysRevLett.113.220602>.
- 888 [37] E. G. KYRIAKIDIS, *Stationary probabilities for a simple immigration-birth-death process under*
- 889 *the influence of total catastrophes*, Stat. Probabil. Lett., 20 (1994), pp. 239–240, [https://doi.org/10.1016/0167-7152\(94\)90048-5](https://doi.org/10.1016/0167-7152(94)90048-5).
- 890
- 891 [38] E. G. KYRIAKIDIS AND A. ABAKUKS, *Optimal pest control through catastrophes*, J. Appl.
- 892 Probab., 26 (1989), pp. 873–879, <https://doi.org/10.2307/3214392>.
- 893 [39] Y. J. LEUNG, W. V. LI, AND RAKESH, *Spectral analysis of brownian motion with jump boundary*,
- 894 P. Am. Math. Soc., 136 (2008), pp. 4427–4436.
- 895 [40] M. LUBY, A. SINCLAIR, AND D. ZUCKERMAN, *Optimal speedup of Las Vegas algorithms*, Inform.
- 896 Process. Lett., 47 (1993), pp. 173–180, [https://doi.org/10.1016/0020-0190\(93\)90029-9](https://doi.org/10.1016/0020-0190(93)90029-9).
- 897 [41] D. LUNZ, *Optimal feedback control in first-passage resetting*, J. Phys. A-Math. Theor., 53
- 898 (2020), p. 44LT01, <https://doi.org/10.1088/1751-8121/abc7c>.
- 899 [42] S. C. MANRUBIA AND D. H. ZANETTE, *Stochastic multiplicative processes with reset events*,
- 900 Phys. Rev. E, 59 (1999), pp. 4945–4948, <https://doi.org/10.1103/PhysRevE.59.4945>.
- 901 [43] M. MONTERO, A. MASÓ-PUIGDELLOSAS, AND J. VILLARROEL, *Continuous-time random walks*
- 902 *with reset events*, Eur. Phys. J. B, 90 (2017), p. 176, <https://doi.org/10.1140/epj/b/e2017-80348-4>.
- 903
- 904 [44] J. L. MORALES AND J. NOCEDAL, *Remark on “Algorithm 778: L-BFGS-B: Fortran subroutines*
- 905 *for large-scale bound constrained optimization”*, ACM T. Math. Software, 38 (2011), pp. 1–
- 906 4, <https://doi.org/10.1145/2049662.2049669>.
- 907 [45] Y. MORENO, J. B. GÓMEZ, AND A. F. PACHECO, *Instability of scale-free networks under*
- 908 *node-breaking avalanches*, Europhys. Lett., 58 (2002), p. 630, <https://doi.org/10.1209/epl/i2002-00442-2>.
- 909
- 910 [46] Y. MORENO, R. PASTOR-SATORRAS, A. VÁZQUEZ, AND A. VESPIGNANI, *Critical load and*
- 911 *congestion instabilities in scale-free networks*, Europhys. Lett., 62 (2003), p. 292, <https://doi.org/10.1209/epl/i2003-00140-7>.
- 912
- 913 [47] A. E. MOTTER, *Cascade control and defense in complex networks*, Phys. Rev. Lett., 93 (2004),
- 914 p. 098701, <https://doi.org/10.1103/PhysRevLett.93.098701>.
- 915 [48] A. E. MOTTER AND Y.-C. LAI, *Cascade-based attacks on complex networks*, Phys. Rev. E, 66
- 916 (2002), p. 065102, <https://doi.org/10.1103/PhysRevE.66.065102>.
- 917 [49] M. E. J. NEWMAN, *The structure and function of complex networks*, SIAM Rev., 45 (2003),
- 918 pp. 167–256, <https://doi.org/10.1137/S003614450342480>.
- 919 [50] E. ROLDÁN, A. LISICA, D. SÁNCHEZ-TALTAVULL, AND S. W. GRILL, *Stochastic resetting in*
- 920 *backtrack recovery by RNA polymerases*, Phys. Rev. E, 93 (2016), p. 062411, <https://doi.org/10.1103/PhysRevE.93.062411>.
- 921
- 922 [51] M. L. SACHTJEN, B. A. CARRERAS, AND V. E. LYNCH, *Disturbances in a power transmission*
- 923 *system*, Phys. Rev. E, 61 (2000), pp. 4877–4882, <https://doi.org/10.1103/PhysRevE.61.4877>.
- 924
- 925 [52] I. SANTRA, U. BASU, AND S. SABHAPANDIT, *Run-and-tumble particles in two dimensions under*
- 926 *stochastic resetting conditions*, J. Stat. Mech.-Theory E., 2020 (2020), p. 113206, <https://doi.org/10.1088/1742-5468/abc7b7>.
- 927

- 928 [53] S. H. STROGATZ, *Exploring complex networks*, Nature, 410 (2001), pp. 268–276, <https://doi.org/10.1038/35065725>.
929
- 930 [54] S. THURNER, R. HANEL, AND P. KLIMEK, *Introduction to the theory of complex systems*, Oxford
931 University Press, 2018, <https://doi.org/10.1093/oso/9780198821939.001.0001>.
- 932 [55] B. VAN ES, B. KOREN, AND H. J. DE BLANK, *Finite-difference methods for extremely anisotropic*
933 *diffusion*, in Advanced numerical methods for complex environmental models: needs and
934 availability, I. Faragó, Á. Havasi, and Z. Zlatev, eds., Bentham Science Publishers, 2013,
935 pp. 29–52, <https://doi.org/10.2174/9781608057788113010006>.