Clusters in an Epidemic Model with Long-Range Dispersal

Xiangyu Cao, Pierre Le Doussal, Alberto Rosso

Clusters in an epidemic model with long-range dispersal

Xiangyu Cao,[1] Pierre Le Doussal,[1] and Alberto Rosso[2]

[1] *Laboratoire de Physique de l'École normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université Paris Cité, F-75005 Paris, France*
[2] *Université Paris-Saclay, CNRS, LPTMS, 91405, Orsay, France*

In presence of long range dispersal, epidemics spread in spatially disconnected regions known as clusters. Here, we characterize exactly their statistical properties in a solvable model, in both the supercritical (outbreak) and critical regimes. We identify two diverging length scales, corresponding to the bulk and the outskirt of the epidemic. We reveal a nontrivial critical exponent that governs the cluster number, the distribution of their sizes and of the distances between them. We also discuss applications to depinning avalanches with long range elasticity.

Catastrophic events such as avalanches, material failure, and initial-stage epidemic outbreaks, often occur as a chain reaction. Their simplest model was that of Bienaymé and Galton-Watson (BGW)[1, 2], originally conceived for genealogy. In a continuous time version one starts with a single infected individual. During a short time lapse dt each infected individual recovers with probability $\gamma \mathrm{d}t$, and causes a new infection with probability $\beta \mathrm{d}t$. On average, each infection generates $R_0 = \beta/\gamma$ new ones: R_0 determines the fate of the epidemic. When $R_0 < 1$, it goes to extinction rapidly. When $R_0 > 1$, the size of the population that has been infected up to time t grows exponentially, $S \sim e^{(\beta-\gamma)t}$, as in the initial outbreak stage of an epidemic. At the critical point, $R_0 = 1$, the probability that the epidemic has survived up to time t decreases as $\sim 1/t$, and in that case it will have infected $\sim t^2$ individuals. As a result, S has strong fluctuations and has a power law distribution $P(S) \sim S^{-3/2}$ with a cutoff at $S_{\max} \sim t^2$. The critical case mimics the scale free behaviour displayed by avalanches in disordered materials, i.e. the propagation of an instability which triggers further instabilities via elastic interaction [3].

The BGW model ignores the spatial spreading of the epidemic. Branching diffusion models consider that infected individuals also perform some random walk in a d dimensional space, independently of recovery and infection. Often, one specifies the random walk to be a short-range Brownian motion. Then the region affected by the epidemic is a connected set, whose geometric properties have been characterized [4–10]. For instance, at criticality, the radius ξ of this set grows as $\xi \sim S^{1/4}$.

However, Brownian diffusion models cannot capture the *long-range* dispersal that ubiquitously occurs in nature, due to e.g. wind, ocean currents, and air traffic [11–17], spreading an epidemic far from its origin. A similar situation is observed in disordered materials where long-range interactions can trigger disconnected avalanches, e.g., in the propagation of crack fronts [18–21], wetting lines [22–24] or plasticity [25, 26]. In this work we model the long-range dispersal of the infected individuals as follows: during dt, an individual jumps from x to x' with probability $p_\alpha(x - x')\mathrm{d}^d x' \mathrm{d}t$, where $p_\alpha(x)$ decays as a power law at large distances:

$$p_\alpha(x) = \frac{\theta(|x| - \epsilon)}{|x|^{\alpha+d}}\,, \ \alpha > 0\,. \tag{1}$$

Here $|x|$ is the Euclidean norm, θ is the Heaviside step function and $\epsilon \ll 1$ is a short-distance cutoff. Similar long-range models have been studied on a lattice, where the outbreak always displays a sub-exponential growth [27–33]. Here, we assume an infinite pool of susceptible individuals everywhere, which ensures an exponential outbreak when $R_0 > 1$.

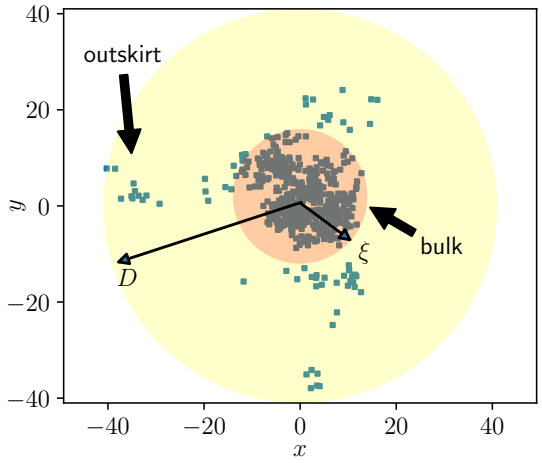

FIG. 1. Spatial distribution of an critical epidemic started at the origin, totaling 1000 infections. Due to the long-range dispersal [(1), $\alpha = 1.5$] of infected individuals, the points visited form disconnected clusters. The bulk of radius ξ, concentrating a majority of infections, is surrounded by a sparse outskirt, containing *all* infections.

A typical epidemic obtained from a numerical simulation of our model is shown in Fig. 1. One may distinguish two regions characterized by distinct length scales. The *bulk*, of radius ξ, contains most of the infections. Farther away, a sparser *outskirt* of radius D contains all the remaining infections. The existence of the outskirt is a consequence of the long-range jumps. One aim of this work is to obtain how ξ and D scale with the infected population S. Another fundamental consequence of long-range dispersal, is the presence of *clusters*, i.e.

spatially disconnected regions affected by the epidemic. As is apparent from Fig. 1, the clusters vary in sizes and their spatial distribution is not uniform. The second goal of this Letter is to introduce a method to properly define the clusters. We then characterize their random geometry: how the number of clusters grows with S, how their sizes are distributed, what are the distances separating them, etc. Our exact results are obtained by the analysis of a non-linear "instanton" equation. We stress that our methods are applicable to real-world data. As a proof of principle, we tested our theory against the Covid-19 outbreak data in the US. Remarkably, a prediction of our model, (12) below, describes well the spatial distribution of the clusters during the first week of March 2020 [34].

The epidemic model introduced above provides a discrete realization, equivalent, near criticality [35], to the mean-field theory [36, 37] describing the spatial structure of the avalanches of slowly driven elastic interfaces in a disordered medium. In crack experiments, clusters have been directly observed [38], and their number and size distribution have been characterized [39, 40]. These works proposed that these properties are fully encoded in the global properties of the crack front, e.g., in its roughness exponent [41–43]. Here, we make a first step at examining this issue analytically; our results indicate that the cluster statistics probably involve a new independent exponent. In what follows, we report our main results, and sketch the main points of their derivation, see [34] for details.

Bulk and outskirt. We first determine the length scales of the bulk, ξ, and outskirt, D, by simple arguments. We consider our model with a single infected individual at the origin initially ($t = 0$). At criticality ($R_0 = 1$), the bulk length ξ can be estimated as the *typical* displacement of a random walk with jump distribution (1). When $\alpha < 2$, we have a Lévy flight, and thus

$$\xi \sim t^{\frac{1}{\alpha}} \sim S^{\frac{1}{2\alpha}} \,, \; \alpha < 2 \,, \tag{2}$$

where the last estimate comes from the scaling $S \sim t^2$. When $\alpha > 2$, we recover the short-range behavior $\xi \sim \sqrt{t} \sim S^{\frac{1}{4}}$ [37, 44]. On the other hand, the outskirt's diameter D is estimated as the *farthest* jump among $\sim S$ independent attempts:

$$D \sim S^{\frac{1}{\alpha}} \,. \tag{3}$$

Hence, the outskirt is much larger than the bulk if $\alpha < 4$: only for $\alpha > 4$ do we completely recover a short-range behavior, with $D \sim \xi \sim S^{1/4}$. This is already a surprise, as naively one would expect a short-range takeover at $\alpha = 2$.

In the supercritical regime, the argument for the outskirt diameter D and the result (3) still hold. The scaling of the bulk size ξ is different. Indeed, the infected population grows exponentially, $S \sim e^{(\beta-\gamma)t}$. As a consequence, the density of infected individuals is exponentially large at the epicenter, $x = 0$, and decays as $\sim S|x|^{-\alpha-d}$ (1). The bulk extent is then determined by the distance, $|x| = \xi$, at which the density reaches unity:

$$\xi \sim S^{\frac{1}{\alpha+d}} \sim e^{\frac{\beta-\gamma}{\alpha+d}t} \tag{4}$$

Note that, when $R_0 > 1$, the separation of scales $D \gg \xi$ remains for any α and the short-range behavior with a linear growth $\xi \propto t$ is never recovered. This is in contrast with lattice models [27, 28], where a reduction to short-range does happen at $\alpha = d + 1$.

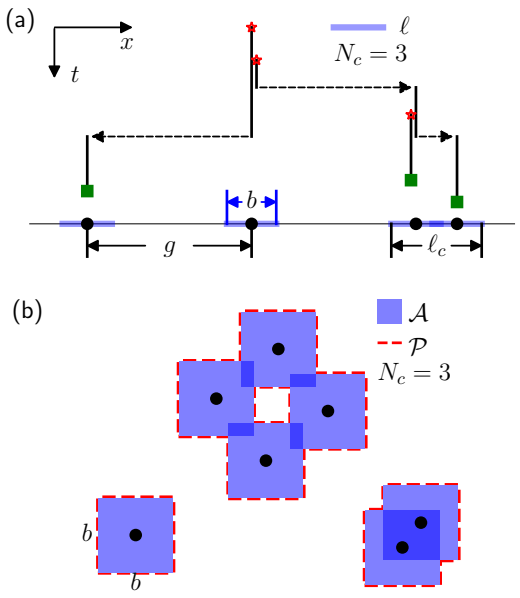

FIG. 2. (a) Illustration of an epidemic in 1D. An infection (recovery, jump) is indicated by a red star (green square, dashed line, resp.). The points visited are coarse-grained by an interval of length b. They form $N_c = 3$ clusters, with total extension $\ell = \sum \ell_c$. The gaps g are defined independently of b. (b) In 2D, a point is coarse-grained by a square of side b, to define the cluster number $N_c(= 3)$, the area $\mathcal{A}$, and the perimeter $\mathcal{P}$. Note that in 2D, a cluster can be non-convex and have holes.

Defining clusters. In our model, the ensemble of positions ever occupied by an infected individual up to time t is a finite set, as only a finite number of jumps have occurred. How do we define its clusters? For simplicity we focus on one and two dimensions. We introduce a coarse-graining scale $b \gg \epsilon$, and thicken each point by a patch of size b — an interval of length b in 1D, and a square of size b in 2D — centered at that point, see Fig. 2. The patches attached to different points can then overlap and form clusters. To characterize their spatial distribution, we introduce the following observables: (i) The number of cluster N_c; (ii) the length/area of individual clusters, ℓ_c in 1D and $\mathcal{A}_c$ in 2D. The sum of all ℓ_c ($\mathcal{A}_c$) is the epidemic's *extension*, ℓ (area, $\mathcal{A}$, respectively). (iii) We also characterize the distances between clusters. In 1D, a natural choice is the distribution of gaps (Fig. 2). It is not hard to see that, the number of gaps larger than g is related to the cluster number with $b = g$:

$$N_c(b = g) = (\text{number of gaps} > g) + 1 \,. \tag{5}$$

In 2D, the notion of gaps is not obvious, and we take $N_c(b=g)$ as a probe of the distances between clusters. We obtained the b dependence of all the quantities; for conciseness, we report results with $b=1$ unless otherwise stated.

Clusters at criticality. When $R_0=1$, statistical fluctuations are strong, and there are various ways of averaging. Here, we focus on averages conditioned on a large infected population S (assuming non-extinction), denoted as $\langle\mathcal{O}\rangle_S$ for an observable $\mathcal{O}$. From the S-conditioned averages, we can obtain the asymptotics of the average over all realizations up to time t, using

$$\langle\mathcal{O}(t)\rangle \sim \int^{S_{\max}} P(S)\,\langle\mathcal{O}\rangle_S\,\mathrm{d}S, \quad S_{\max}\sim t^2\,. \tag{6}$$

Thus, if $\langle\mathcal{O}\rangle_S \sim S^a$, $\langle\mathcal{O}(t)\rangle \sim t^{\max(2a-1,0)}$ (See Table I of [34] for results).

We have seen that when $\alpha<4$, the outskirt is much larger than the bulk, and we expect many clusters. Interestingly, the interval $\alpha\in(0,4)$ is divided into several regimes, with qualitatively different behaviors of $\langle\ell\rangle_S, \langle\mathcal{A}\rangle_S$ and $\langle N_c\rangle_S$. Let us start with the most nontrivial one, $\alpha\in(d/2,d)$. There, we find that the average extension and area are related to the bulk extent in a rather expected way:

$$\langle\ell\rangle_S\sim\xi\,,\ \langle\mathcal{A}\rangle_S\sim\xi^2\,. \tag{7}$$

It is worth noting that the above quantities are independent of b for a large range of b, see (15) below. Now, the average number of clusters scales with ξ via a new and nontrivial exponent

$$\langle N_c\rangle_S\sim\xi^\chi\,,\ \alpha<\chi<d\,. \tag{8}$$

The exponent χ is a function of α and d, and determined by a transcendental equation given in [34] together with a plot. It satisfies $\alpha<\chi<d$, which means that the number of clusters grows with S but remains much lower than the area or extension. Thus, the cluster areas $\mathcal{A}_c$ and extensions ℓ_c must have broad distributions (with divergent mean as $S\to\infty$). Computing them is beyond the reach of the present techniques. However, assuming that they follow a single power law in the interval $[1,\xi^d]$, we can surmise their exponent [40]:

$$P(\ell_c)\sim\ell_c^{-\chi-1}\,,\ P(\mathcal{A}_c)\sim\mathcal{A}_c^{-\chi/2-1}\,. \tag{9}$$

Concerning the gaps between clusters, we found that $\langle N_c(b=g)\rangle_S$ has two regimes with distinct power laws:

$$\frac{\langle N_c(b=g)\rangle_S}{\sqrt{S}}\sim\begin{cases}(g/g_c)^{-\frac{d(\chi-\alpha)}{d-\alpha}} & 1\ll g\ll g_c\\ (g/g_c)^{-\frac{\alpha d}{d+\alpha}} & g_c\ll g\ll D\end{cases}, \tag{10}$$

where $g_c=\xi^{1-\alpha/d}$ is the crossover gap length. The two gap regimes $g\ll g_c$ and $g\gg g_c$ correspond to gaps in the bulk and in the outskirt, respectively. To better understand this result, let us consider one dimension [45].

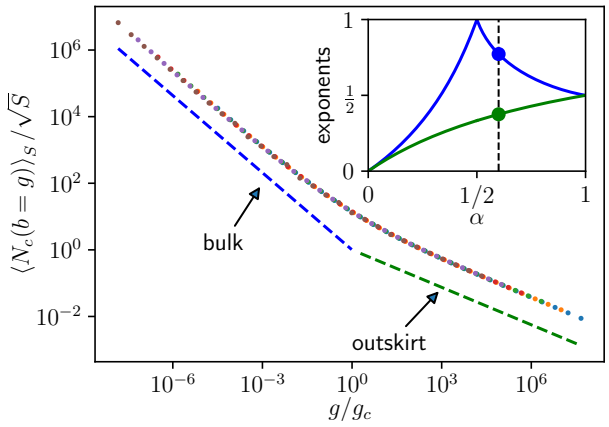

FIG. 3. Gap distribution in 1D with $\alpha=0.6$, obtained by numerical solution of (13)[34]. Data points with various sizes $S=10^{20},\ldots,10^{28}$ are collapsed using (10). The dashed lines indicate the predicted exponents in two regimes. Inset: The dependence of the two exponents on α.

Observe that, the total length of the gaps no greater than g_c is exactly the bulk size:

$$\sum(\text{gaps}\le g_c)\sim g_c\,\langle N_c(b=g_c)\rangle_S\sim\xi\,. \tag{11}$$

Now, if we consider all the bulk gaps up to a size $g\ll g_c$, their number is almost $N_c(b=1)$, but their total size is a negligible fraction of the bulk. On the other hand, the outskirt gaps of are a minority in number, but their total size is much greater than the bulk size. Of course, there is no sharp transition between bulk and outskirt, but rather a smooth crossover. Indeed, in Fig. 3, we show that to demarcate the two power laws requires several orders of g/g_c. Otherwise, one may observe a "compromise" of the theoretical predictions.

So far we focused on the regime $\alpha\in(d/2,d)$. The other ones are simpler. In a nutshell, for strong long-range dispersal ($\alpha<d/2$), the clusters become atomic and have a finite size in average. Therefore we have $\langle N_c\rangle_S\sim S$, and $\langle\ell\rangle_S, \langle\mathcal{A}\rangle_S\sim S$ as well. For weak long-range dispersal ($\alpha>d$), the bulk becomes more compact, and gaps of size $\gtrsim 1$ exist only in the outskirt. See [34] for a detailed discussion.

Clusters of an outbreak. In the super-critical ($R_0>1$) regime, the statistical fluctuations are weak. We can thus consider the averages up to t, which are dominated by realizations with an infected population $S\sim e^{(\beta-\gamma)t}$. Recall that the bulk and outskirt diameter grow exponentially as (3) and (4), for any $\alpha>0$. Now, the cluster structure of an outbreak is also simpler, and we found the same qualitative picture for any α. The bulk is compact and has no large gaps. Its extension/are is $\langle\ell\rangle\sim\xi, \langle\mathcal{A}\rangle\sim\xi^2$, where $\xi\sim S^{1/(d+\alpha)}\sim e^{(\beta-\gamma)t/(d+\alpha)}$ (4). The outskirt is sparse, and has an exponential number of clusters, $\langle N_c\rangle\sim\xi^d$. Notably, their spatial structure is time-independent: the gap distribution is stationary up

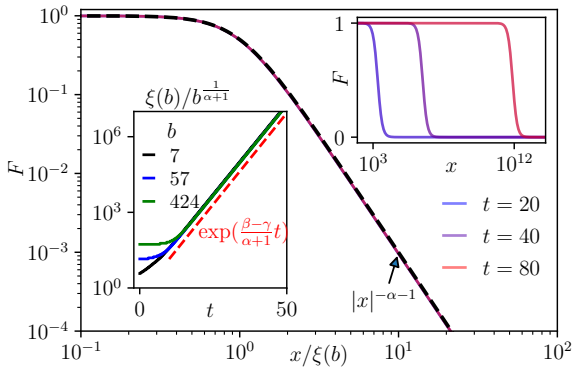

FIG. 4. Traveling wave solution to (13) in the supercritical regime, ($\alpha = 2, \beta = 1, \gamma = 0, b = 7$). The solution at $t = 20, 40, 80$ (upper inset) collapsed onto the front profile $F(x,t) = f(x/\xi(b))$, $f(y) = 1/(1+y^{1+\alpha})$ (black dashed curve). The front position $\xi(b)$ is defined by $F(\xi(b)) = 1/2$. Its time dependence is plotted in the lower inset for 3 values of b. The collapse confirms the b dependence of the front position [34].

to a normalization and a cutoff,

$$\langle N_c(b=g)\rangle \sim \xi^d g^{-\frac{d\alpha}{\alpha+d}}\,,\ g \ll D\,. \tag{12}$$

In [34], we tested this prediction against Covid-19 data, finding an encouraging agreement.

Method. We highlight some key points of our analytical approach. The main object is a function $F(x,t|b)$, which is the probability that x belongs to the patch of some point visited before t. A standard backward recursion argument shows that F satisfies a semi-linear "instanton" equation [37, 46–49]:

$$\partial_t F = \mathcal{D}^\alpha F + (\beta-\gamma)F - \beta F^2\,,\ F|_{t=0} = 0\,, \tag{13}$$

for any x outside the patch of the origin; inside that, $F = 1$. Here $(\mathcal{D}^\alpha f)(x) := \int p_\alpha(x-y)(f(y)-f(x))\mathrm{d}^d y$ is the "fractional diffusion" term. From the solution F, we can obtain the area (extension) by integrating it over the plane (line). The cluster number is obtained by differentiating with respect to b. In 1D, we have

$$N_c(b) = \partial_b \ell(b)\,. \tag{14}$$

A similar trick exists in 2D [34].

Therefore, the problem boils down to the asymptotic analysis of (13). In the super-critical regime, the exponential spreading of its traveling wave solution follows from existing rigorous results [50]; for a self-contained derivation and our results on clusters, see [34]. In Fig. 4, we plot the front profile. Note that it decays as a power law, and does not have a characteristic width. In contrast, in traveling wave equations with short-range diffusion, the wavefront position has linear growth in time and its width is of order unity.

The results at criticality follows from the stationary solution of (13). The solution in the regime $\alpha \in (d/2, d)$ involves a noteworthy feature. To discuss that without going into technical details, consider the following puzzle, say in 1D. Recall that the cluster number and the extension are related by a b-derivative (14). Then, how can they scale differently: $\ell \sim \xi, N_c \sim \xi^\chi \ll \xi$? The crux is that, the leading asymptotics of ℓ is b-independent, while N_c derives from a subleading term:

$$\langle \ell(b)\rangle_S = c_0\xi + c_1(b)\xi^\chi\,,\ (b \ll g_c) \tag{15}$$

where c_0 is b-independent. To extract the cluster statistics from the solution of (13), it is necessary to identify its *subleading* asymptotics, in addition to the previously known leading one [51]. This mathematical detail has a physical interpretation: cluster statistics are associated with irrelevant perturbations in the sense of the renormalization group. During the coarse-graining process, the clusters merge and information about them is gradually erased.

Conclusion. We have characterized the clusters of an epidemic model with long-range dispersal, which is equivalent near criticality to the mean-field theory of depinning avalanches with long-range elasticity. We found that two diverging length scales — the bulk and the outskirt — emerge in both super-critical and critical regimes. In the latter, the bulk can have a rich structure with broadly distributed cluster sizes as well as gap sizes. Our analytical approach based on the instanton equation can be extended to study the effect of inhomogeneous networks [14], realistic mixing patterns [14, 52], superspreading events [53], or the regions where the epidemic is still active at time t [54, 55]. It will be also interesting to see how much the qualitative features revealed here appear in other epidemic models, e.g., contact point processes [56, 57]. Finally, concerning depinning avalanches, our model provides a mean-field description which should be quantitatively correct for realistic long-range systems when $d \geq 2\alpha$. To describe these systems for $d < 2\alpha$, loop corrections to mean field theory should be taken into account. In particular, our results imply a cluster number distribution $P(N_c) \sim N_c^{-\mu}$ where $\mu = \alpha/\chi + 1$ for $\alpha \in (d/2, d)$. At the critical dimension $d = 2\alpha$, we recover the BGW value $\mu = 3/2$, but in our model $\mu > 3/2$ is a new exponent when $d < 2\alpha$. Meanwhile, numerical studies [40] of realistic models suggest that $\mu \approx 3/2$ for all $d < 2\alpha$. It will be interesting to see how to retain the "dangerously irrelevant" cluster statistics in the field theory and whether the loop corrections can account for this numerical observation.

ACKNOWLEDGMENTS

We thank Jean-Philippe Bouchaud for pointing out the literature on human mobility. We thank William Terrot for preliminary work on the project, and Grégory Schehr, Vincenzo Schimmenti for valuable comments on the manuscript. PLD acknowledges support from ANR

under the grant ANR-17-CE30-0027-01 RaMaTraF. XC and PLD thank LPTMS for hospitality.

SUPPLEMENTAL MATERIAL

A. Derivation of the instanton equation

We recall the standard backward recursion argument used to derive the instanton equation. Let the positions of the infected individuals at time t be $x_1, \dots, x_{I(t)}$ where $I(t)$ is the number of infected. Note that at $t = 0$ we have one infected individual at $x_1 = 0$. Consider the probability that the b-neighborhood (called a patch in the main text) around x has not been infected until t:

$$E(x, t|b) = 1 - F(x, t|b) = \text{Prob}(\|x_i(s) - x\| > b/2, \forall i = 1, \dots, I(s), s < t) \tag{16}$$

where $\|x\| = \|x\|_\infty$ is the infinite-norm of x, defined as $\|(x^1, \dots, x^d)\| = \max\{x^1, \dots, x^d\}$ (x^a is the a-th component of a point x). In particular, $\{\|x\| < b/2\}$ is the box of linear size b centered at the origin.

Since the spatial diffusion is symmetric (the probability of going from $x \to y$ and $y \to x$ are the same), it is not hard to see that E is equal to the probability that the b-neighborhood of the origin has not been visited, if the epidemic starts at x:

$$E(x, t|b) = \text{Prob}(\|x_i(s)\| > b/2 | x_0(0) = x)\,. \tag{17}$$

Note that if $\|x\| \le b/2$, $E = 0$ by definition. For $\|x\| > b/2$, we can apply a backward recursion of E by considering what can happen during $t \in (0, \mathrm{d}t)$.

1. Another individual is infected, with probability $\beta \mathrm{d}t$. In that case $E(x) \to E(x)^2$ (because from that moment, the two individuals act independently from now on with the same law).

2. The patient 0 recovers with probability $\gamma \mathrm{d}t$. Then $E(x) \to 1$ (note that we assumed $\|x\| > b/2$).

3. The patient performs a jump to y with probability $p_\alpha(y - x)\mathrm{d}^d y \mathrm{d}t$. In that case $E(x) \to E(y)$. Note that $p_\alpha(x)$ is a probability rate, and thus not normalized. With probability $1 - \mathrm{d}t \times \int p_\alpha(x)\mathrm{d}^d x$, the individual makes no jump.

Gathering all the possibilities, we have

$$E(x, t + \mathrm{d}t) - E(x, t) = \beta(E(x)^2 - E(x))\mathrm{d}t + \gamma(1 - E(x))\mathrm{d}t + \int p_\alpha(x - y)(E(y) - E(x))\mathrm{d}^d y \mathrm{d}t\,,\ \|x\| > b/2\,. \tag{18}$$

Now noting that $F = 1 - E$ and the definition of the fractional diffusion operator

$$(\mathcal{D}^\alpha f)(x) = \int p_\alpha(x - y)(f(y) - f(x))\mathrm{d}^d y\,,\ p_\alpha(x) = |x|^{-d-\alpha}\theta(|x| - \epsilon)\,, \tag{19}$$

we obtain the instanton equation, (22) below.

In the main text we mentioned that in 1D, the cluster number can be obtained by deriving the extension with with respect to b, namely $N_c(b) = \partial_b \ell(b)$. In 2D, deriving the area once gives the *perimeter* $\mathcal{P}$; deriving twice, we obtain the difference between cluster and hole numbers:

$$\mathcal{P}(b) = 2\partial_b \mathcal{A}(b)\,,\ \partial_b \mathcal{P}(b) = 4(N_c - N_h)\,. \tag{20}$$

Yet, we are able to constrain the asymptotics of N_c using the bounds

$$\mathcal{P}/(4b) \ge N_c \ge N_c - N_h\,. \tag{21}$$

These geometric formulas are not hard to derive, upon observing Figure 2 of the main text. We also note that similar formulas (with different prefactors) hold if we replace squares by disks in 2D. So our asymptotic results are independent of this choice.

B. Super-critical regime: traveling wave solution

We consider the instanton equation for $F(x,t|b)$, which is the probability that the b-neighborhood of x has been visited by an infected individual by time t (the epidemic starts with a single infected individual at $x=0, t=0$)

$$\partial_t F = \mathcal{D}^\alpha F + (\beta-\gamma)F - \beta F^2\,,\ \|x\| > b/2 \tag{22}$$

$$F(\|x\| < b/2) = 1\,,\ F|_{t=0} = 0\,, \tag{23}$$

in the super-critical regime ($\beta > \gamma$). Note that the initial condition $F|_{t=0} = 0$ is consequence of the strict inequality $s < t$ in the definition (16): for $t = 0$, $F = 0$ because no infected individual exists for $t < 0$.

Here, we provide a simple self-contained derivation of the traveling wavefront position (including b dependence) and of front profile. The result applies to any d and α. Very initially, we can neglect the last two terms and solve the approximate equation $\partial_t F = \mathcal{D}^\alpha F|_{t=0}$. We get

$$F \approx t\mathcal{D}^\alpha\left[\theta(b/2 - \|x\|)\right]\,, \tag{24}$$

$$\mathcal{D}^\alpha\left[\theta(b/2 - \|x\|)\right] \approx \begin{cases} (\|x\| - b/2)^{-\alpha} & 0 < \|x\| - b/2 \ll b \\ b^d|x|^{-\alpha-d} & |x| \gg b\,. \end{cases} \tag{25}$$

The first regime corresponds to points outside the b-neighborhood but very close to it (such that the neighborhood appears semi-infinite). In that case there can be prefactors in the above formula depending on x/b, but they are unimportant for what follows. The second regime corresponds to points far away from the neighborhood; the formula we gave is asymptotically exact.

At $t = \mathcal{O}(1)$ (uniformly for all x), the linear in t growth is be overtaken by the exponential growth generated by the $(\beta-\gamma)F$ term. In fact that term dominates the RHS of (22), so that

$$F \sim S\mathcal{D}^\alpha\left[\theta(b/2 - \|x\|)\right]\,,\ \text{where } S = e^{(\beta-\gamma)t}\,, \tag{26}$$

until $F \sim 1$ and the nonlinear term $-\beta F^2$ stops the growth. If $S \gg b^\alpha$, e.g. if the scale b is smaller than outskirt scale $D = S^{1/\alpha}$, we can obtain an equation for the wave front position $\xi(b)$

$$S \sim b^{-d}\xi(b)^{\alpha+d} \implies \xi(b) = S^{\frac{1}{\alpha+d}} b^{\frac{d}{\alpha+d}}\,,\ \text{if } D \gg b\,. \tag{27}$$

We remark that $\xi(b)$ is related to the bulk extent ξ by the relation $\xi = \xi(b=1)$, but they are different quantities. Recall that in the super-critical regime, we define the bulk extent as the distance from origin at which the density of the infected population becomes of order one. ξ is thus independent of b, and can be done by a simple argument, as given in the main text, and does not require analyzing the instanton equation. In contrast, $\xi(b)$ is the b-dependent wavefront position of the instanton equation.

Now, we can plug the traveling wave ansatz

$$F(x,t) = f(|x|/\xi(b)) \tag{28}$$

into (22) to find the front profile. As a result, at large ξ, we find

$$-\frac{(\beta-\gamma)yf'(y)}{\alpha+d} = (\beta-\gamma)f(y) - \beta f(y)^2\,. \tag{29}$$

Note that the $\mathcal{D}^\alpha F$ term gives a negligible contribution. We can explicitly solve for $f(y)$:

$$f(y) = \frac{\beta-\gamma}{\beta + (y/y_0)^{d+\alpha}} \tag{30}$$

where y_0 is an unknown constant. Note that $f(y\to 0) \to 1 - \gamma/\beta = 1 - R_0^{-1}$ and $f(y\to\infty) \sim 1/y^{d+\alpha}$. These predictions are verified in Fig. 4 of the main text.

In summary we have shown that for any b fixed, as $t \to \infty$,

$$F(x,t) \to f(|x|/\xi(b))\,,\ \xi(b) = e^{\frac{\beta-\gamma}{\alpha+d}t}\, b^{\frac{d}{\alpha+d}} \tag{31}$$

with f given by (30). The bulk extent is given by ξ with $b = 1$. Integrating over x in 1D and 2D, we have

$$\langle \ell \rangle = \int F\mathrm{d}x \sim \xi(b)\,,\ \langle \mathcal{A} \rangle = \int F\mathrm{d}^2x \sim \xi(b)^2\,. \tag{32}$$

Both terms have a nontrivial b dependence. So in 1D, the cluster number is

$$\langle N_c \rangle = \partial_b \langle \ell \rangle \sim \xi(b)/b \,, \, (d=1) \,. \tag{33}$$

In 2D, we have

$$\langle \mathcal{P} \rangle = 2\partial_b \langle \mathcal{A} \rangle \sim \xi(b)^2/b \,, \, \langle N_c - N_h \rangle = \frac{1}{4}\partial_b \langle \mathcal{P} \rangle \sim \xi(b)^2/b^2 \,, \, (d=2) \,. \tag{34}$$

Now since $\mathcal{P}/(4b) \geq N_c \geq N_c - N_h$, the above results sandwich the asymptotics of $\langle N_c \rangle$:

$$\langle N_c \rangle \sim \xi(b)^2/b^2 \,, \, (d=2) \,. \tag{35}$$

This sandwiching argument will be systematically repeated below to obtain the cluster number asymptotics in 2D, see (70), (71), (77), (83) below.

C. Application to the Covid-19 outbreak in the United States

As a proof of principle of our method, we test our approach against the real-world on the Covid-19 outbreak in the United States in March 2020. The long-range dispersal is important for describing the epidemic spreading in human society. Indeed, it has been shown [14, 15] that human mobility is well described by a Lévy flight with $\alpha \approx 0.6$, with a cutoff of $\sim 10^3$km.

We test the prediction of our model in the supercritical regime, on the distribution of gaps between clusters. We recall that this is defined as the number of clusters N_c as a function of the coarse-grain distance b. The prediction is that its average value is time-independent, up to a normalisation that depends on the number of infections S. More precisely,

$$\langle N_c(b) \rangle \sim S^{\frac{d}{\alpha+d}} b^{-\frac{\alpha d}{\alpha+d}} \,. \tag{36}$$

We test this prediction against the data on the initial outbreak of Covid-19 in the United States (continental states) in March 2020. County-level daily infection numbers are made available by The New York Times (https://github.com/nytimes/covid-19-data). For any given day, we obtain a set of points which are the geographical center of the counties where infections have been reported, see Fig. 5 (top). Then, for any distance b, we form a graph by connecting all pairs of infected counties with geodesic distance $\leq b$, and compute $\langle N_c(b) \rangle$ as the number of connected components of the graph. We can extract the total infection number S for each day.

The results are shown in Fig. 5 (bottom). We found that in a time window of roughly a week (March 1 - March 5), and for 10km $\leq b \leq 10^3$km, $\langle N_c(b) \rangle$ is time-independent up to a global pre-factor that increases with time. Upon dividing by $S^{-\frac{d}{\alpha+d}}$, see (36), the data for different days are collapsed. Moreover, the b dependence is consistent with a power law $b^{-\frac{\alpha d}{\alpha+d}}$, where $\alpha \approx 0.6$ as previously found by independent studies [14, 15] (one of the works measured α by tracking the displacement of dollar bills). After the first week of March, the epidemic covers almost all the counties of the United States, see Fig. 5 (top). Thus, the behavior of $\langle N_c(b) \rangle$ changes qualitatively and is no longer described by our theory.

It may seem concerning that the prediction works only for a few days. However, this is consistent with the rapid growth of the outskirt diameter D predicted by our model. Indeed, the daily infection data we used indicate that the total number of infections doubles every two days in the beginning of March 2020, $S \sim 2^{n_{\text{days}}/2}$. By the scaling law $D \sim S^{1/\alpha}$, the diameter of the epidemic doubles every day. The distance resolution in our analysis is about 10km, and the Lévy nature of human mobility is valid up to $\sim 10^3$km [14, 15], giving us a window of two orders of magnitude. Therefore the number of days where our theory is expected to work is

$$n_{\text{days}} = \frac{\ln(10^2)}{\ln(2)/(2\alpha)} \approx 5.5 \,,$$

which is approximately a week.

To summarize, we tested the prediction on the gap distribution against real-world epidemic data and obtained an encouraging agreement in the initial stage of the outbreak. The quantitative prediction (36) appears to be robust despite many real-world factors that are not taken into account. This is a demonstration of universality in statistical physics.

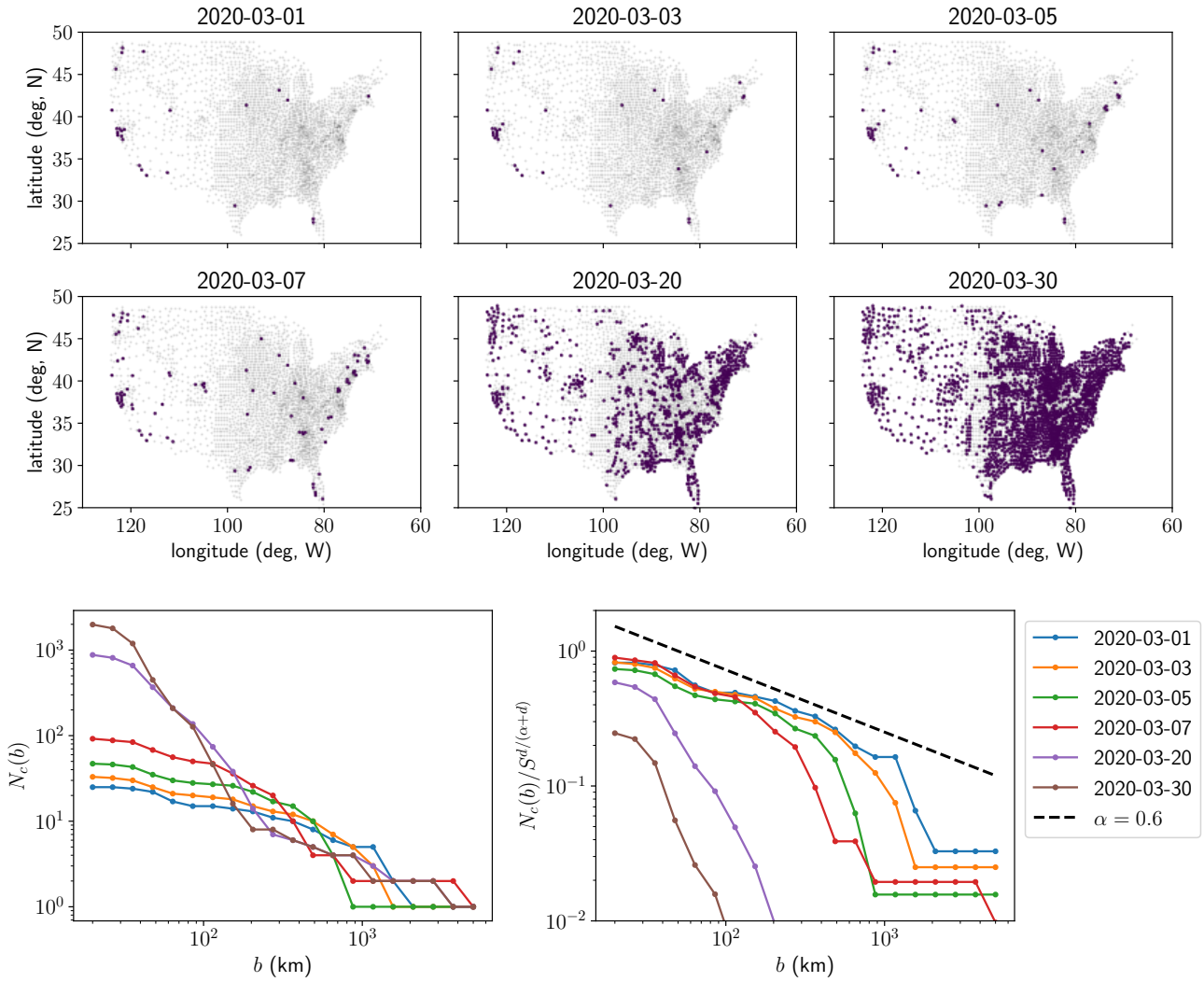

FIG. 5. *Top.* The purpose dots are the geographical center of the counties of the continental US where Covid-19 infections have been reported. The grey dots are all other counties. *Bottom.* The number of connected components of the graph obtained by connecting infected counties with distance $\leq b$, as a function of b, for several days in March 2020. The data in early March is compared to the prediction (36), with $d = 2$ and $\alpha = 0.6$ (dashed line).

D. Critical regime

1. *General strategy*

We approach criticality from the sub-critical side, where the epidemic always goes to extinction. Thus we shall always consider the $t \to \infty$ limit. For simplicity, we may let

$$\beta = 1\,,\ \gamma = 1 + m^2\,, \tag{37}$$

where m^2 is a small positive number ("mass squared" in the avalanche context) that controls the distance to criticality.

We can obtain averages conditioned on the total infection number S by deriving with respect to m^2. Indeed, this is because the distribution of the total infection number, S, follows a power law $P(S) \sim S^{-3/2}$ with a cutoff at $S_m = m^{-4}$. Therefore, for any observable $\mathcal{O}$, we have

$$\langle \mathcal{O} \rangle \sim \int^{S_m} S^{-3/2} \langle \mathcal{O} \rangle_S \, \mathrm{d}S \tag{38}$$

$d=1$	$\alpha<1/2$	$1/2<\alpha<1$	$1<\alpha<3$	$3<\alpha<4$	$\alpha>4$
$\langle \ell \rangle_S \sim S^?$	1	$1/(2\alpha)$	$1/(1+\alpha)$	$1/4$	
$\langle N_c \rangle_S \sim S^?$	1	$\chi/(2\alpha)$	$1/(1+\alpha)$	$1-\alpha/4$	0
$\langle \ell(t) \rangle \sim t^?$	1	$(1-\alpha)/\alpha$	0		
$\langle N_c(t) \rangle \sim t^?$	1	$(\chi-\alpha)/\alpha$	0		

$d=2$	$\alpha<1$	$1<\alpha<2$	$2<\alpha<4$	$\alpha>4$
$\langle \mathcal{A} \rangle_S \sim S^?$	1	$1/\alpha$	$1/2$	
$\langle N_c \rangle_S \sim S^?$	1	$\chi/(2\alpha)$	$1-\alpha/4$	1
$\langle \mathcal{A}(t) \rangle \sim t^?$	1	$(2-\alpha)/\alpha$	0	
$\langle N_c(t) \rangle \sim t^?$	1	$(\chi-\alpha)/\alpha$	0	

TABLE I. Summary of the main asymptotic results in the critical regime. Scaling exponents are displayed for both the time average and conditioned average on a large infection number S. In two dimensions, the perimeter's scaling is identical to the cluster number.

where $\langle \mathcal{O} \rangle_S$ is the average at criticality and conditioned on S, while $\langle \mathcal{O} \rangle$ is the (non-conditioned) average over the near-critical ensemble (37). Differentiating both sides with respect to m^2, we obtain

$$\langle \mathcal{O} \rangle_S \sim - \left. \frac{\partial \langle \mathcal{O} \rangle}{\partial m^2} \right|_{m^2 \to S^{-\frac{1}{2}}} . \tag{39}$$

Most often, $\langle \mathcal{O} \rangle$ contains a term that is proportional to a power of m^2. Then, $\langle \mathcal{O} \rangle_S$ is proportional to that term, multiplied by $m^{-2} \sim \sqrt{S}$.

We can also directly obtain finite time averages at criticality from the $t = \infty$, sub-critical one, by simply substituting $m^2 \to 1/t$. This is because the critical and near-critical dynamics are indistinguishable until $t \sim m^{-2}$, after which the sub-critical one saturates.

In view of the above considerations, we shall concentrate on the the stationary instanton equation in d dimensions, $d = 1, 2$:

$$\mathcal{D}^\alpha F = F^2 + m^2 F \, , \, \|x\| > b/2 \, , \, F(\|x\| < b/2) = 1 \, . \tag{40}$$

Indeed, we verified numerically that the time-dependent instanton equation always converges (point-wise) to a stationary solution as $t \to \infty$ in the subcritical regime. The integral of F, and its b-derivative provides the sub-critical averages. To analyze the asymptotic behavior of the solution, we shall consider a few approximate solutions to it. Each of them is dominant in some range of parameters. Then we show how to assembly them in various regimes. In what follows, we shall assume $\alpha < 2$, until Section D 8, where the competition with short-range physics is discussed.

The main results of the analysis below are summarized in Table I.

Note. Unless otherwise stated, we assume that $b \geq 1$, and restrain from considering smaller values of b. Considering $b \geq 1$ is enough for deriving the results of the main text. We will comment on situations where considering $b \ll 1$ might be useful (see Section D 8 below).

2. *Extent of the bulk and the outskirt*

Before proceeding with the analysis of the instanton equation, we recall the simpler calculation, of the average density $S(x)$ of infected individuals at distance x from the origin [the size of the infected population is is the integral of $S(x)$]. By a similar backward recursion argument as above, we can show that it satisfies a linear equation

$$\mathcal{D}^\alpha S(x) = m^2 S(x) - \delta(x) \, . \tag{41}$$

Therefore, $S(x)$ is nothing but the Green function of a fractional Gaussian free field with mass m^2 and a kinetic term $\propto |k|^\alpha$ when $\alpha < 2$ (and $|k|^2$ when $\alpha > 2$, see (79) below) in the momentum space. The bulk extent is the correlation length of this field:

$$\begin{cases} \xi = m^{-\frac{2}{\alpha}} & \alpha < 2 \\ \xi_{\rm SR} = m^{-1} & \alpha > 2 \end{cases} . \tag{42}$$

(Here, in the Supplemental Material, to avoid confusion, we shall use the subscript SR to denote the short-range bulk extent. The symbol ξ without subscript is always equal to $m^{-2/\alpha}$ even when $\alpha > 2$.) Beyond the correlation length, $S(x) \sim |x|^{-\alpha-d}$ has a fast-decaying tail. Therefore most of the infections happen inside the bulk. Identifying $m^2 = S^{-1/2}$ according to (39) gives the expressions in the main text: $\xi = S^{1/(2\alpha)}$ for $\alpha < 2$ and $\xi_{\rm SR} = S^{1/4}$ for $\alpha > 2$.

We also recall the simple argument leading to the scaling law of the outskirt radius $D \sim S^{1/\alpha}$, Eq. (3) of the main text. Indeed, consider S independent jump distances $r_1, \ldots, r_S$, each distributed according to Eq (1) of the main text.

It is not hard to see that $\mathrm{Prob}(|r_i| < x) \sim 1 - Cx^{-\alpha}$ for any i. By independence, $\mathrm{Prob}(\max(|r_1|, \dots |r_S|) < x) \sim (1 - Cx^{-\alpha})^S \sim \exp(-Cx^{-\alpha}S)$. Thus, the typical value of the maximal jump size is $r_{\max} \max(|r_1|, \dots |r_S|) \sim S^{1/\alpha}$. Now, assuming that the radius of the outskirt is dominated by $r_{\max}$, we obtain Eq. (3) of the main text. Although this argument seems heuristic, Eq. (3) is exact, as confirmed by the systematic analysis below, see remarks around Eq. (61).

3. *Scale invariant solution and subleading term*

We now come back to the analysis of the time-independent instanton equation (40) and consider its first approximate solution. The "scale invariant" approximation ignores the mass term and the boundary condition at $\|x\| < b/2$, and focuses on power law type solutions of the equation

$$F^2 = \mathcal{D}^\alpha F \,. \tag{43}$$

The mass introduces a cutoff of this equation when $F \sim m^2$ (since that is where $m^2 F \sim F^2$ in (40)).

To solve this equation we recall the classic formula on the Fourier transform of power laws:

$$\mathcal{F}[|x|^{-a}] := \int |x|^{-a} e^{\mathbf{i}x.k} \mathrm{d}^d x = B(a)|k|^{a-d} \,, \text{where } B(a) = \frac{\pi^{d/2} 2^{d-a} \Gamma\left(\frac{d-a}{2}\right)}{\Gamma\left(\frac{a}{2}\right)} \,. \tag{44}$$

Note that B also depends on d but we omitted this argument to keep notations concise. Then, for $\alpha < 2$, the fractional diffusion term $\mathcal{D}^\alpha$ acts on a power law in the following way:

$$\mathcal{D}^\alpha(|x|^{-h}) = |x|^{-h-\alpha} D(h, \alpha) \,, \tag{45}$$

$$D(h, \alpha) := \frac{1}{(2\pi)^d} B(h) B(d - h - \alpha) B(d + \alpha) \,. \tag{46}$$

This can be shown by performing the convolution by Fourier transform: $\mathcal{D}^\alpha f = \mathcal{F}^{-1}\left(\mathcal{F}[p]\mathcal{F}[f]\right)$.

From (45), it follows immediately that (43) admits a power law solution, which we shall call the scale-invariant approximation:

$$F_{\mathrm{sc}} = D(\alpha, \alpha)|x|^{-\alpha} \,. \tag{47}$$

Comparing this to m^2, we find the mass cutoff

$$\xi = m^{-2/\alpha} \,, \tag{48}$$

which is the same as the bulk extent. When $|x| \gg \xi$, $F_{\mathrm{sc}} \sim |x|^{-\alpha-d}$ decays fast and its contribution can be ignored for all purposes. Plotting $D(\alpha, \alpha)$, one may find that it is positive when $\alpha \in (d/2, d)$. It has a zero at $\alpha = d/2$ and a pole at $\alpha = d$. Therefore the scale invariant solution is valid as a dominant asymptotic behavior *only* in the interval $\alpha \in (d/2, d)$. A similar solution was found in 1D in [51].

Remark. We have swept some (well-known) technical details under the rug. Indeed, $\mathcal{F}[p]$ is the regularized Fourier transform, with an ϵ-dependent constant removed, and to which the formula (44) applies. Also, the RHS of (45) also misses δ terms at the origin, which are unimportant since our analysis concerns large x.

4. *Correction to scale invariant solution*

It will be important to consider admissible perturbations of the scale invariant solution F_{sc}. That is, we consider

$$F = F_{\mathrm{sc}} + \delta F \,,\ \delta F \ll F_{\mathrm{sc}} \,. \tag{49}$$

Such an F satisfies (43) if δF satisfies its linearized version

$$\mathcal{D}^\alpha(\delta F) = 2F_{\mathrm{sc}} \delta F \,. \tag{50}$$

This also admits a power-law solution at large distances

$$\delta F \propto |x|^{-\eta} \,, \text{ where } \eta \text{ satisfies } 2D(\alpha, \alpha) = D(\eta, \alpha) \,. \tag{51}$$

Anticipating what follows, we note that the exponent η is related to χ in the main text by

$$\chi = d + \alpha - \eta \,. \tag{52}$$

See Fig. 6 for a plot of χ. The explicit form of the transcendental equation $2D(\alpha,\alpha) = D(\eta,\alpha)$ (51), in terms of χ, is the following:

$$\frac{D(\eta,\alpha)}{2D(\alpha,\alpha)} = \frac{\Gamma\left(\frac{\alpha}{2}\right)\Gamma\left(\frac{\chi-\alpha}{2}\right)\Gamma\left(\frac{1}{2}(d-2\alpha)\right)\Gamma\left(\frac{1}{2}(d+2\alpha-\chi)\right)}{2\Gamma(\alpha)\Gamma\left(\frac{1}{2}(\chi-2\alpha)\right)\Gamma\left(\frac{d-\alpha}{2}\right)\Gamma\left(\frac{1}{2}(d+\alpha-\chi)\right)} = 1\,. \tag{53}$$

This equation determines χ as mentioned in the main text. Note that for each d, α there are *several branches of solutions* to this equation (generically an infinity, but only two for $\alpha = 2, 4, ..$). The branch relevant to this study is the unique one such that $\chi \in (\alpha, d)$ (or $\eta \in (\alpha, d)$) as $\alpha \in (d/2, d)$. We justify this choice by observing that $\chi \to d$ as $\alpha \to d/2$, which is expected from the continuity to the $\alpha < d/2$ regime (see below).

Remark. A similar analysis of the linear perturbation of the instanton equation around the self-consistent solution $F_{\rm SR}(x) = 2(4-d)/x^2$ was performed for the short-range Brownian model (in the continuum setting of the Brownian force model) [58] (Section VIII) and led to two possible values for the exponent $\eta = 3, -4$ in $d = 1$, $\eta = \pm 2\sqrt{2}$ in $d = 2$ and $\eta = \frac{1}{2} \pm \frac{\sqrt{17}}{2}$ in $d = 3$. Similar exponents also appeared in calculations of the fractal dimension of the boundary of the super-Brownian motion [59].

One can ask how the equation (53) recovers the SR case. This happens by setting $\alpha = 2$, while keeping d, η generic. In that case (53) simplifies and one obtains two branches of solutions:

$$\frac{D(\eta,2)}{2D(2,2)} = \frac{\eta(d-\eta-2)}{4(d-4)} \quad \Rightarrow \quad \eta = \frac{1}{2}\left(d \pm \sqrt{(d-20)d+68} - 2\right)\,, \tag{54}$$

which recover the above cited values.

It is interesting to note however that the $\alpha \to 2$ limit of the transcendental equation (53) is quite subtle in 2D. First of all, the $\alpha \to 2$ and $d \to 2$ limits do not commute. We have the following series expansion of the LHS of (53):

$$\left.\frac{D(\eta,\alpha)}{2D(\alpha,\alpha)}\right|_{\substack{d=2-\epsilon \\ \alpha=2-r\epsilon}} = \frac{\eta^2(r-1)}{16r-8} + \mathrm{O}(1)\,,\ \epsilon \to 0. \tag{55}$$

The short-range model is obtained by taking $r = 0$, i.e., sending $\alpha \to 2$ *before* $d \to 2$ as noted above (indeed, equating the above equation to 1, we obtain $\eta = \pm 2\sqrt{2}$ in agreement with above). On the other hand, taking $r = \infty$, i.e., sending $d \to 2$ before $\alpha \to 2$, we obtain $\eta \to \pm 4$ as $\alpha \to 2$. By tuning r, i.e. the angle of approach to $(\alpha, d) = (2, 2)$, we can obtain any other value of η as a limit. However, the solution that is useful for the cluster statistics in this work is a different branch, and invisible from the series expansion (55). This is because at $d = 2$, the limits $\eta \to 2$ and $\alpha \to 2$ do not commute either. We have

$$\left.\frac{D(\eta,\alpha)}{2D(\alpha,\alpha)}\right|_{\substack{d=2 \\ \alpha=2-\epsilon \\ \eta=2-r\epsilon}} = \frac{1+r}{4r} + \mathrm{O}(1)\,,\ \epsilon \to 0. \tag{56}$$

Equating the RHS to 1 gives $r = 1/3$, or $2 - \eta = (2-\alpha)/3 + \mathcal{O}((2-\alpha)^2)$ at $d = 2$. We note in passing that in 1D, $1 - \eta = (1-\alpha)/2 + \mathcal{O}((1-\alpha)^2)$.

5. *The plateau approximation*

We now consider another approximate solution: the "plateau approximation". It consists in replacing the LHS of (40) by $\mathcal{D}^\alpha F \approx \mathcal{D}^\alpha\left[\theta(b/2 - \|x\|)\right]$, which is calculated in (25). Equating that to F^2, we obtain

$$F_{\rm pl} = b^{\frac{d}{2}} |x|^{-\frac{\alpha+d}{2}}\,,\ b \ll |x| \ll X_m \tag{57}$$

where the mass cutoff scale can be again determined by $F_{\rm pl}(X_m) = m^2$:

$$X_m = b^{\frac{d}{\alpha+d}} m^{-\frac{4}{d+\alpha}}\,. \tag{58}$$

Beyond that $F_{\rm pl} \sim |x|^{-\alpha-d}$ decays fast and its contribution can be neglected.

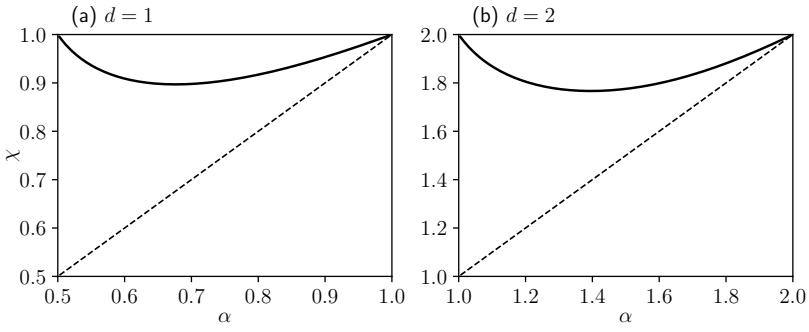

FIG. 6. Solid curve: The exponent χ as a function of α for $d = 1$ and $d = 2$. The dashed line is α.

Under the plateau approximation, the area or the extension is given by

$$\int F_{\rm pl}(x)\mathrm{d}^d x = \begin{cases} c_0(b) - c_1 X_m^d m^2 & \alpha > d \\ c_1 X_m^d m^2 & \alpha < d \end{cases} \tag{59}$$

where c_1 is some positive constant independent of b, while $c_0(b)$ depends on b but not on m^2. When $\alpha < d$, the above formula leads to the large gap power law of the gap distribution, see (71) below. When $\alpha > d$, the integral of $F_{\rm pl}$ remains of order one as $m^2 \to 0$. Yet, the subleading term $c_1 X_m^d m^2$ is m^2 dependent, and can be used to calculate the conditioned average on a large S, via (39):

$$\partial_{m^2} \int F_{\rm pl}(x)\mathrm{d}^d x \sim X_m^d \,,\ \alpha > d \,. \tag{60}$$

Remark. When the point is near the b-neighborhood, $\|x\| - b/2 \ll b$, the plateau approximation has a slower decay:

$$F_{\rm pl} \sim |\|x\| - b/2|^{-\frac{\alpha}{2}} \,,\ 0 < \|x\| - b/2 \ll b \tag{61}$$

However, this regime is important only if $X_m \lesssim b$, which is equivalent to $b \gtrsim D = m^{-4/\alpha}$. To make the following discussion less cumbersome, we shall always assume $b \ll D$ and ignore the near-plateau regime (61), unless otherwise stated. (Larger values of b no longer probe the gap distribution but rare instances of gaps greater than D.)

6. *Solution for $d/2 < \alpha < d$*

We are now ready to build the solution for the most interesting regime $d/2 < \alpha < d$, using the above pieces. Since $\alpha < d$, $(\alpha + d)/2 > \alpha$, we expect that the solution is dominated by the plateau approximation (57) at small distances and by the scale invariant one (47) at large distances. Comparing them we obtain a crossover scale

$$F_{\rm pl}(x_b) = F_{\rm sc}(x_b) \Rightarrow x_b = b^{\frac{d}{d-\alpha}} \,. \tag{62}$$

However, recall that the scale-invariant solution has a cutoff at $|x| \sim \xi$, and can exist if and only if $x_b \ll \xi$. Otherwise, the plateau approximation dominates everywhere, up to a larger mass cutoff $X_m \gtrsim \xi$. The crossover value of b is the crossover gap scale

$$x_b \sim \xi \Rightarrow b \sim g_c = \xi^{1-\alpha/d} \tag{63}$$

we referred to in the main text. In summary we have:

$$F = \begin{cases} F_{\rm pl} & |x| \ll x_b \\ F_{\rm sc} + \delta F & x_b \ll |x| \ll \xi \end{cases} \,,\ \text{if } b \ll g_c \tag{64}$$

$$F = F_{\rm pl} \,,\ |x| \ll X_m \,,\ \text{if } b \gg g_c \,. \tag{65}$$

Above, δF is the subleading correction to the scale invariant solution. We have seen that it must be proportional to $|x|^{-\eta}$ (51). It remains to fix the prefactor, which will turn out to be b-dependent. To do this, we argue that a separation of the scales $b \ll x_b \ll \xi$ imposes a single-parameter scaling of F near x_b:

$$F(x|b) = x_b^{-\alpha} \tilde{F}(x/x_b)\,, \tag{66}$$

for some scaling function $\tilde{F}(y)$ such that $\tilde{F}(y \to \infty) \sim y^{-\alpha}$ and $\tilde{F}(y \to 0) \sim y^{-(\alpha+d)/2}$. This scaling form is fixed by the leading terms in (64). We verified this ansatz with extensive numerical solution of the instanton equation, see Fig. 7. Imposing this scaling form to the subleading term fixes its prefactor

$$\delta F \sim |x/x_b|^{-\eta} x_b^{-\alpha} \sim b^{\frac{d(\eta-\alpha)}{d-\alpha}} |x|^{-\eta}\,. \tag{67}$$

It indeed has a nontrivial b-dependence, which will allow us to obtain the number of clusters and the bulk gaps distribution, see below.

We now have the complete solution in the regime $\alpha \in (d/2, d)$ (we keep all the dependence on x, m^2 and b, but drop out all other prefactors):

$$\boxed{\begin{aligned} &F \sim \begin{cases} F_{\rm pl} = b^{\frac{d}{2}} x^{-\frac{\alpha+d}{2}} & |x| \ll x_b = b^{\frac{d}{d-\alpha}} \\ F_{\rm sc} + \delta F = |x|^{-\alpha} + x_b^{\eta-\alpha}|x|^{-\eta} & x_b \ll |x| \ll \xi = m^{-2/\alpha} \end{cases}, \text{ if } b \ll g_c = \xi^{1-\alpha/d} \\ &F \sim F_{\rm pl} = b^{\frac{d}{2}} x^{-\frac{\alpha+d}{2}}\,,\ |x| \ll X_m = b^{\frac{d}{\alpha+d}} m^{-\frac{4}{d+\alpha}}\,, \text{ if } b \gg g_c\,. \end{aligned}} \tag{68}$$

When $|x| \gg \xi$ or $|x| \gg X_m$, F decays as $|x|^{-\alpha-d}$ and can be ignored.

To find the mean area and extension, we take $b \ll g_c$, and integrate over F. From (68) we can see that the integral is dominated by $F_{\rm sc} + \delta F$ at its mass cutoff scale $|x| \sim \xi$, so that:

$$\langle \ell(b) \rangle\,, \langle \mathcal{A}(b) \rangle \sim \int_{|x|<\xi} F(x) \mathrm{d}^d x \sim \xi^{d-\alpha} + x_b^{\eta-\alpha} \xi^{d-\eta}\,. \tag{69}$$

(Here and below, the RHS with $d = 1$ applies to ℓ, and $d = 2$ applies to $\mathcal{A}$.) Note that the first term in the RHS is dominant but b-independent; only the subdominant one is b-dependent. Applying (39), setting $b = 1$ and neglecting the subdominant term, we obtain the S-conditioned mean extension/area of the main text.

The mean cluster number and the bulk gap ($b \ll g_c$) distribution is dominated by the subleading term of F involving η, $|x| \sim \xi$, since the leading term is b-independent. The result can be written in a nice way using g_c:

$$\langle N_c(b = g) \rangle \sim \partial_b^d \int_{|x|<\xi} F \mathrm{d}^d x \sim (g/g_c)^{-\frac{d(d-\eta)}{d-\alpha}} = (g/g_c)^{-\frac{d(\chi-\alpha)}{d-\alpha}}\,,\ g \ll g_c\,. \tag{70}$$

Here we recall that $\chi = \alpha + d - \eta$ by definition. Upon applying (39) we find the small-gap result and the cluster number (with $b = 1$) result of the main text. We also recall that in 2D, we need to use the sandwiching argument $\langle \mathcal{P} \rangle \gtrsim \langle N_c \rangle \geq \langle N_c - N_h \rangle$, as discussed below (34).

Similarly, the outskirt gap distribution is dominated by the $b \gg g_c$ case of (68). We have

$$\langle N_c(b = g) \rangle \sim \partial_b^d \int_{|x|<X_m} F \mathrm{d}^d x \sim (g/g_c)^{\frac{\alpha d}{\alpha+d}}\,,\ g \gg g_c\,. \tag{71}$$

The S-conditioned result in the main text is then found by applying (39). Note that, we can see clearly from the calculation above that these gaps are in the outskirt, by noticing that the integral of F is dominated by the scale X_m, which is larger than ξ.

In Fig. 3 of the main text, we verified the gap distribution prediction by extensive numerical solution of the instanton equation. The derivative with respect to b an m^2 are evaluated numerically as finite differences.

7. $\alpha < d/2$: *The linear approximation*

When $\alpha < d/2$, the scale invariant solution is no longer viable in the long distance because its prefactor would be negative. In fact, at large distances, the nonlinearity becomes irrelevant and the linear approximation applies, $F \approx F_{\rm lin}$ where $F_{\rm lin}$ is such that

$$(\mathcal{D}^\alpha - m^2) F_{\rm lin} = \text{fast decaying term} \implies F_{\rm lin} \sim c(b)|x|^{-(d-\alpha)}\,,\ |x| \ll \xi = m^{-\frac{2}{\alpha}} \tag{72}$$

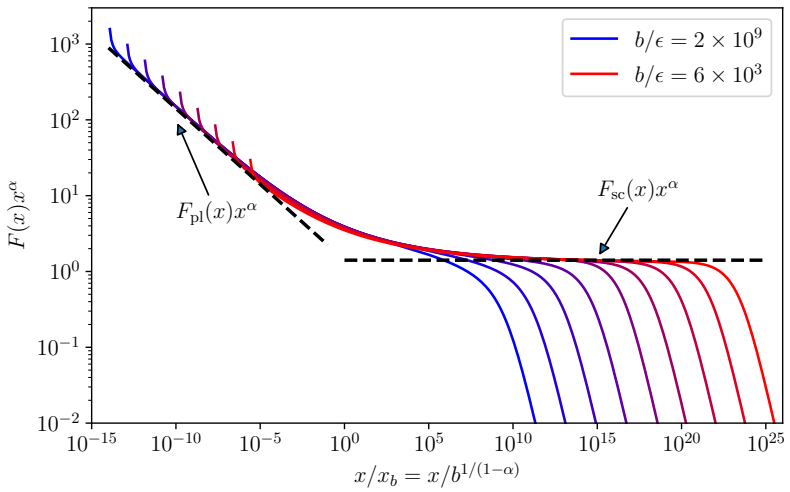

FIG. 7. Verification of the scaling ansatz (66), which is equivalent to $F(x|b) = \hat{F}(x/x_b)x^{-\alpha}$ such that $\hat{F}(y \to \infty) = D(\alpha, \alpha)$ and $\hat{F}(y \to 0) \sim y^{(\alpha-d)/2}$. We solved the instanton equation numerically at $m^2 = 10^{-19}$ and for $b/\epsilon = 6 \times 10^3, \ldots, 2 \times 10^9$ (from red to blue), and plotted Fx^α against x/x_b (for $x > b$); the data collapse near $x \sim x_b$ confirms the scaling ansatz. For $x \gg x_b$ and $x \ll x_b$ the solutions are in good agreement with the scale-invariant ($F_{\rm sc}$) and plateau ($F_{\rm pl}$) approximations, respectively. Note that the dashed lines have the exact prefactors; no fit is performed. The large distance and small distance deviations from the collapse are due to the mass cutoff and the near plateau behavior (61), respectively.

(Recall that we assume $\alpha < 2$ here). Beyond ξ, the solution again decays as $|x|^{-d-\alpha}$ and can be neglected. To fix the b-dependent prefactor, we can exploit the crossover to the plateau solution at the mass cutoff scale. A smooth crossover requires that $F_{\rm lin}(X_m) \sim F_{\rm pl}(X_m)$ when $X_m \sim \xi$. This imposes:

$$c(b) \sim b^{\frac{d(d-2\alpha)}{d-\alpha}} , \tag{73}$$

and in turn fixes the crossover scale between the linear and plateau regimes: it is still $x_b = b^{\frac{d}{d-\alpha}}$ as in the regime $\alpha \in (d/2, d)$. So the crossover with the plateau approximation works exactly as in the regime $\alpha \in (d/2, d)$. In summary, the instanton solution for the $\alpha < d/2$ is as follows:

$$\boxed{\begin{aligned} &F \sim \begin{cases} F_{\rm pl} = b^{\frac{d}{2}} x^{-\frac{\alpha+d}{2}} & |x| \ll x_b = b^{\frac{d}{d-\alpha}} \\ F_{\rm lin} = x_b^{-\alpha} |x/x_b|^{\alpha-d} & x_b \ll |x| \ll \xi \end{cases} , \text{ if } b \ll g_c = \xi^{1-\alpha/d} \\ &F \sim F_{\rm pl} = b^{\frac{d}{2}} x^{-\frac{\alpha+d}{2}} , \; |x| \ll X_m = b^{\frac{d}{\alpha+d}} m^{-\frac{4}{d+\alpha}} , \text{ if } b \gg g_c . \end{aligned}} \tag{74}$$

The extension and area (for $b \ll g_c$) are obtained by integrating F, and are dominated by $F_{\rm lin}$:

$$\langle \ell(b) \rangle , \langle \mathcal{A}(b) \rangle = \int F(x) \mathrm{d}^d x \sim \xi^\alpha x_b^{d-2\alpha} , \; b \ll g_c . \tag{75}$$

Note that these quantities have a nontrivial b-dependence already at leading order, at variance with the case of $\alpha \in (d/2, d)$. Applying (39) we obtain

$$\langle \ell(b) \rangle_S , \langle \mathcal{A}(b) \rangle_S \sim S x_b^{d-2\alpha} . \tag{76}$$

Applying the ∂_b^d derivative we obtain the gap distribution and the cluster number; applying (39) gives us the S-conditioned average:

$$\langle N_c(b = g) \rangle \sim \frac{1}{\sqrt{S}} \langle N_c(b = g) \rangle_S \sim \begin{cases} (g/g_c)^{\frac{\alpha d}{d-\alpha}} , \; g \ll g_c \\ (g/g_c)^{\frac{\alpha d}{d+\alpha}} , \; g \gg g_c \end{cases} \tag{77}$$

In particular we find $\langle N_c \rangle_S \sim \xi^{2\alpha} \sim S$. In other words, the number of clusters is proportional to the total infected population. So the size of each cluster is of order unity in average: the clusters are atomic.

Remark. We may understand the $\alpha < d/2$ regime as a degeneration of the $\alpha \in (d/2, d)$ one, in the following sense. The linear approximation is the continuation of the subleading term in the scale-invariant solution, whereas the leading term vanishes. Thereby, the gap distribution and cluster number in the regime $\alpha < d/2$ can be expressed using the same formulas of the regime $\alpha \in (d/2, d)$, upon the following replacements of the exponents:

$$\eta \to d - \alpha \,,\, \chi \to 2\alpha \,,\, (\alpha < d/2) \,. \tag{78}$$

8. *Regimes with $\alpha > d$*

When $\alpha > d$, among the above approximate solutions, only the plateau one can survive. However, when $\alpha > 2$, the fractional diffusion operator contains a normal diffusion term: more precisely, in the Fourier space

$$\mathcal{F}[p](k) = A_\epsilon |k|^2 + \cdots + B(\alpha + d)|k|^\alpha \,, \tag{79}$$

where A_ϵ is a constant depending on the short-distance cutoff of $p_\alpha(x)$, and $\dots$ denotes further even powers ($\leq \alpha$) of $|k|$ that may appear. This implies a new approximate solution: the short-range (SR) scale invariant solution,

$$F_{\rm SR} = 2(4-d)A_\epsilon |x|^{-2} \,,\, |x| \ll \xi_{\rm SR} = m^{-1} \,. \tag{80}$$

Comparing this with the plateau approximation both near the plateau $\sim |x|^{-\alpha/2}$ (61), and faraway $\sim |x|^{-(\alpha+d)/2}$ (57), we identify two threshold values: $\alpha = 4 - d$, and $\alpha = 4$. Hence, there are three cases:

1. $\alpha \in (1, 3), d = 1$. Here, the plateau approximation is valid everywhere up to the cutoff X_m:

$$\boxed{F \sim F_{\rm pl} = b^{\frac{d}{2}} x^{-\frac{\alpha+d}{2}} \,,\, |x| \ll X_m = b^{\frac{d}{\alpha+d}} m^{-\frac{4}{d+\alpha}}} \tag{81}$$

   Then the extension for any b has a nontrivial b dependence, as follows:

$$\langle \ell(b) \rangle_S \sim \partial_{m^2} \int F(x) {\rm d}x = b^{\frac{1}{1+\alpha}} S^{\frac{1}{1+\alpha}} \,. \tag{82}$$

   The gap distribution has only the large-gap regime (again given by the plateau approximation):

$$\langle N_c(b = g) \rangle_S \sim S^{\frac{1}{1+\alpha}} g^{-\frac{\alpha}{1+\alpha}} \sim \sqrt{S} (g/g_c)^{-\frac{\alpha}{\alpha+1}} \,,\, g_c = \xi^{1-\alpha} \ll 1 \,. \tag{83}$$

   The above results may appear simple formally, but their physical interpretation is rather subtle. Indeed, the average extension is always much greater than the bulk extent, which is $\xi \sim S^{1/(2\alpha)}$ for $\alpha < 2$, and $\xi_{\rm SR} \sim S^{1/4}$ for $\alpha > 2$. This transition from a long-range bulk to a short-range one is *invisible* from the asymptotic behavior of $\langle \ell \rangle$. This is because the extension is not dominated by the bulk (as is the case for all the other regimes $\alpha < d$), but by the many clusters of the outskirt, at a distance $X_m \gg \xi$ from the origin. Now, concerning the bulk itself, the above results provide only indirect information. For example, we may infer that the bulk should be compact, and devoid of gaps of size ≥ 1, because the gap distribution does not have a small-gap regime: indeed, in (83) $g_c \ll 1$ for $\alpha > 1$. If is possible for very small b, $b \lesssim g_c \ll 1$, the asymptotic behavior of the solution might be different from (81) above, and reveal further information about the bulk in the regime $\alpha \in (1, 3), d = 1$. We leave this to future study.

2. $\alpha \in (4 - d, 4)$. This is a more tricky case, since there is again a crossover from the plateau to the SR scale invariant approximation. The crossover scales can be worked out in a similar way as in the regime $\alpha \in (d/2, d)$, but their values are different:

$$\boxed{\begin{aligned} F &= \begin{cases} F_{\rm pl} = b^{\frac{d}{2}} x^{-\frac{\alpha+d}{2}} & |x| \ll b^{\frac{d}{\alpha+d-4}} \\ F_{\rm SR} \sim |x|^{-2} & b^{\frac{d}{\alpha+d-4}} \ll |x| \ll \xi_{\rm SR} = m^{-1} \end{cases} \,,\ \text{if } b \ll g_c' = \xi_{\rm SR}^{\frac{\alpha-4+d}{d}} \qquad (84) \\ F &= F_{\rm pl} = b^{\frac{d}{2}} x^{-\frac{\alpha+d}{2}} \,,\, |x| \ll X_m = b^{\frac{d}{\alpha+d}} m^{-\frac{4}{d+\alpha}} \,,\ \text{if } b \gg g_c' \,. \qquad (85) \end{aligned}}$$

   (We do not need to calculate any correction to $F_{\rm SR}$, as we explain below.) As a result, the extension and area are dominated by the short-range bulk: (89) still holds, as in the regime $\alpha > 4$. However, in difference with the

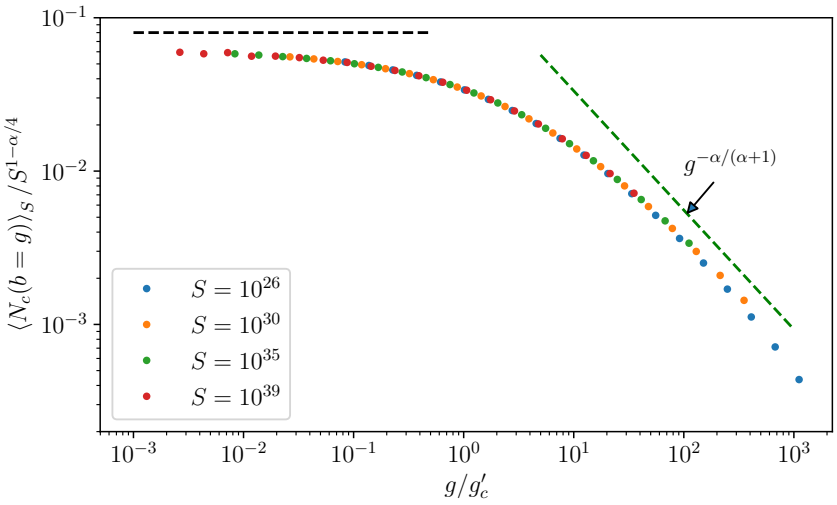

FIG. 8. Cluster number (gap distribution) with $\alpha = 3.5, d = 1$. We solved the instanton equation for $b/\epsilon \in [60, 2 \times 10^6]$, and $m^2 = 10^{-13}, 10^{-15}, 10^{-17}, 10^{-19}$ as well as $1.5m^2$, in order to extract S-conditioned averages at $S = m^{-4}$, using (39). The data collapse confirms the prediction (87) as well as that of the gap distribution for $g > g_c$.

latter regime, there are many clusters and gaps in the outskirt. Indeed, for $b \gg g_c'$, we have the gap distribution given by the plateau approximation:

$$\langle N_c(b=g)\rangle_S \sim \sqrt{S}(g/g_c)^{-\frac{\alpha d}{\alpha+d}}\,,\ g_c = \xi^{1-\alpha/d}\,,\ b \gg g_c'\,. \tag{86}$$

(Note that $\xi = m^{-2/\alpha} \sim S^{1/(2\alpha)}$ by definition.) As b further decreases below g_c', the short-range solution dominates the mass cutoff scale. So we expect that there are no more gaps in the bulk, and therefore the cluster number no longer grows:

$$\langle N_c(b \ll g_c')\rangle_S \sim \langle N_c(b=g_c')\rangle_S \sim \xi_{\rm SR}^{4-\alpha} \sim S^{1-\alpha/4}\,, \tag{87}$$

in both dimensions. In particular, the cluster number scales as $\langle N_c(b=1)\rangle_S \sim S^{1-\alpha/4}$. The above predictions are tested numerically in 1D, see (8).

We note that, in the numerical solution of the instanton equation in 1D, we did *not* observe the short-range subleading exponents (54), which is $\propto |x|^{-3}$ in 1D. Indeed a $c(b)|x|^{-3}$ correction to $F_{\rm SR}$ would imply that $\langle N_c(b=1)\rangle_S \sim \mathcal{O}(1)$, which is inconsistent with the results at large gaps $b \gg g_c'$. This does not contradict our analysis of admissible perturbations of the scale invariant solutions. An admissible perturbation may not necessarily appear.

3. $\alpha > 4$. The short-range solution decays more slowly than the near-plateau approximation $|x|^{-\alpha/2}$. This implies that the plateau contribution cannot dominate at large distances for any value of b. The solution is always the short range scale invariant one up to the cutoff $\xi_{\rm SR}$, with no significant dependence on b at long distances:

$$\boxed{F = F_{\rm SR} \sim |x|^{-2}\,,\ |x| \ll \xi_{\rm SR}} \tag{88}$$

Therefore we have a short-range behavior for the extension and area

$$\langle \ell(b)\rangle_S\,,\ \langle \mathcal{A}(b)\rangle_S \sim \partial_{m^2} \int F(x) \mathrm{d}^d x \sim \xi_{\rm SR}^d\,. \tag{89}$$

Because there is no longer a b dependence, there are no longer a large number of clusters. In other words, the short-range physics completely takes over when $\alpha > 4$, confirming the simple argument of the main text.

E. Relation and application to avalanches

In this section we discuss the above results from the perspective of depinning avalanches with long-range elasticity. We assume $\alpha < 2$ unless otherwise stated.

We recall that an avalanche of an elastic interface is described by the following equation of motion governing the interface position $u(x,t)$:

$$\partial_t u = (\mathcal{D}^\alpha - m^2)u + f(x,u) + \delta(t-t_0)\delta(x-x_0)m^2 w\,. \tag{90}$$

Here $\mathcal{D}^\alpha$ is the elastic interaction (plus a mass term), m^2 is a mass (it introduces a cut off to the avalanche size, in a similar way as does $1-R_0$), f is the quenched random force from the disordered medium, which for realistic models is usually short range in space. The last term is a localized kick in order to trigger an avalanche. The mean-field approximation, which can be justified in dimension $d > 2\alpha$, consists in replacing f by an independent Brownian motion in u for each x, such that:

$$\left\langle (f(x,u') - f(x,u))^2 \right\rangle = |u-u'|\,. \tag{91}$$

Then we obtain the Brownian force model, which is exactly solvable [36, 37]. If we place the interface at an equilibrium before the kick, the interface will move forward and eventually stop at a further equilibrium. The displacement is the avalanche. Its total size $S = \int \mathrm{d}^d x \mathrm{d}t \partial_t u$ is defined as the integrated displacement.

It is known that the Brownian force model is equivalent to a continuum limit of the epidemic model discussed in this work [35]. We will comment on this below. Now, let us take this mapping for granted and discuss the results of this work from the perspective of mean-field avalanches.

1. Applications

Roughness. In the mapping to the Brownian force model the total number of infections S corresponds to the total size of the avalanche. Its distribution is fixed by the BGW law: $P(S) \sim S^{-\tau}$, with $\tau = 3/2$. Now, the Narayan-Fisher [38, 60, 61] scaling relation (see also extensions in [62]) relates τ to the roughness exponent ζ ($u \sim x^\zeta$):

$$\tau = 2 - \frac{\alpha}{d+\zeta}\,,\ \alpha > d/2\,. \tag{92}$$

Plugging in $\tau = 3/2$, we find $\zeta = 2\alpha - d$. When $\alpha < d/2$, the above relation does not apply, and $\zeta = 0$ instead. How does this relate to our results?

- When $\alpha < d/2$, $\zeta = 0$ corresponds to our finding that $\langle \ell \rangle_S\,, \langle \mathcal{A} \rangle_S \sim S$: The avalanche is flat instead of rough, so that its extension/area is proportional to its size.
- When $\alpha > d/2$, the roughness exponent relates the bulk extent with the size in a standard way:

  $$S \sim \xi^{2\alpha} \sim \xi^{d+\zeta}\,. \tag{93}$$

  This relation is usually interpreted as attesting the self-affinity of the rough avalanche. However, the scaling relation between the extension/area and S is given by (93) (as one expects) only if $d/2 < \alpha < d$. In 1D, and when $\alpha \in (1,2)$, the mean extension $\langle \ell \rangle \sim S^{1/(1+\alpha)}$ is much larger than the bulk extent $S^{1/(2\alpha)}$, and is dominated by many small clusters in the outskirt ($\langle N_c \rangle \sim S^{1/(1+\alpha)}$ as well). These clusters are not expected to be self-affine. Yet, we expect the largest clusters in the bulk to have size $\sim \xi$ and be self-affine with roughness ζ. Probing these clusters directly is beyond the reach of the present approach.

A few more exponents. The nontrivial cluster number exponent χ gives rise to a few other predictions (conjectures), for $\alpha \in (d/2, d)$.

- By assuming self-affinity of the bulk clusters, we can relate the clusters' size S_c with their extension/area via $S_c \sim \ell_c^{2\alpha}$, $S_c \sim \mathcal{A}^\alpha$. Combining this with the conjecture on the distribution of ℓ_c and $\mathcal{A}_c$ (given in the main text), we can conjecture the cluster size distribution:

  $$P(S_c) \sim S_c^{-\frac{\chi}{2\alpha}-1}\,. \tag{94}$$

- Using the scaling relation $N_c \sim \xi^\chi \sim S^{\frac{\chi}{2\alpha}}$ and the BGW law for S, we obtain a conjecture on the distribution of cluster number:

  $$P(N_c) \sim N_c^{-\frac{\alpha}{\chi}-1}\,. \tag{95}$$

  The exponent goes to $3/2$ as $\alpha \to d/2$ (since $\chi \to d$). Away from that limit, the exponent is larger than $3/2$. Of course this is not in contradiction with the numerical observation of Ref. [40] that $P(N_c) \sim N_c^{-3/2}$ for all α in *non-mean-field* avalanches. The robustness of the exponent $3/2$ is probably a result of loop corrections.

2. Remarks on the mapping and the continuum limit

We now briefly discuss the mapping between the epidemic model and the Brownian force model, focusing on the continuum limit involved [35]. From the point of view of the epidemic model, the first step of this continuum limit is to take the infection and recovery rate to be large:

$$\beta = N \gg 1\,,\ \gamma = N + m^2\,. \tag{96}$$

Thereby, the instanton equation $\partial_t F = \mathcal{D}^\alpha F - NF^2 - m^2 F$ acquires a large parameter. The second step is to remove the large parameter from the instanton equation, by changing the observable. Recall that in our approach, $E = 1 - F$ is the probability that the b-neighborhood of the origin is *not* visited by an epidemic starting from x (16). We can soften this "hardcore repulsion" observable as follows:

$$G(x,t) := \left\langle \exp\left(-\int_{-t}^{0} \mathrm{d}s \sum_{i=1}^{I(s)} \frac{1}{N}\lambda(x_i(s),s) \right) \right\rangle\,, \tag{97}$$

where $\lambda(x,t)$ is a function of space, and the epidemic starts at $-t$ with a single infected individual at x (The choice of time coordinate looks awkward but necessary for a backward recursion to work.). In particular, taking $\lambda(x,t) = CN\theta(b/2 - \|x\|)$, with $C \to \infty$, we recover E (16). It is not hard to see that G satisfies a similar backward recursion as E:

$$\partial_t G = \mathcal{D}^\alpha G + N(1-G)^2 + m^2(1-G) - \frac{1}{N}\lambda G\,. \tag{98}$$

Finally, setting $G = \exp(-\tilde{u}/N)$, and keeping the leading order in N, we obtain

$$\partial_t \tilde{u} = \mathcal{D}^\alpha \tilde{u} - \tilde{u}^2 - m^2\tilde{u} + \lambda\,. \tag{99}$$

This is the instanton equation in the Brownian force model [37] (often a different sign convention is used where $\tilde{u} \to -\tilde{u}$ and $\lambda \to -\lambda$). Note that it is very similar to our instanton equation for F. However the differences are important, in particular the physical interpretation is distinct. In the Brownian force model, the velocity of the interface corresponds to the (scaled) density of infected individuals

$$\partial_t u(x,t) = \sum_{i=1}^{I(t)} \frac{1}{N}\delta(x_i(t) - x)\,. \tag{100}$$

In particular, if we set $\lambda(x,t) \equiv \lambda$,

$$e^{-\tilde{u}m^2 w} = G^{I(0)} = \left\langle \exp\left(-\lambda \int \partial_t u \mathrm{d}^d x \mathrm{d}t \right) \right\rangle\,,\ I(0) = m^2 wN \tag{101}$$

is the generating function of the total avalanche size, usually called S in the avalanche literature. It corresponds to th e (scaled) total lifetime of all infected individuals (until recovery), if we start with $I(0)$ infected individuals. This is in turn proportional to the total number of infections. So it is reasonable to call the latter S as well, as we did in the main text.

We remark that, once the continuum limit is taken, it becomes not obvious to speak about clusters. Indeed, putting a source term in (99), for example $\lambda = \lambda_0\theta(b/2 - \|x\|)$, amounts to imposing a finite penalty/fugacity for the time spent by infected individuals (or avalanche activities) in the b-neighborhood. To define the emptyness probability one needs to send $\lambda_0 \to \infty$. This has been done in the case of short-range elasticity to calculate, for example, the extension of the avalanche [37, 44, 58]. But such a limit would be ill-behaved with long-range elasticity, since $\tilde{u}$ would diverge everywhere. By contrast in our model, the instanton equation for F does not have a source term but rather a "boundary condition", $F|_{\|x\|<b/2} = 1$. Similar boundary conditions for $\tilde{u}$ (99) also appeared in the continuum limit, but the interpretation is different, see [35] Appendix D2, and references therein.

Remark: reduction in dimension. The 1D model can be obtained from the 2D one by a projection. Writing $(x_i(t), y_i(t))$ the coordinates of the infected individuals in the 2D model, then the process $x_i(t)$ is the 1D model with the same infection/recovery rates and the jump rate $p_\alpha(x - x')dx'dt$ where $p_\alpha(x-x') = \int dy'/((x-x')^2 + (y-y')^2)^{\frac{2+\alpha}{2}} \sim |x-x'|^{-(1+\alpha)}$, up to an irrelevant change in the precise cutoff function at scale ϵ. A similar reduction in dimension holds for the Brownian force model, see e.g. [58] Sec. IA. This reduction implies a number of bounds between 2D and 1D quantities with the same α. For example, the number of clusters increase with the dimension: $\langle N_c(b)\rangle_{\alpha,2\mathrm{D}} \geq \langle N_c(b)\rangle_{\alpha,1\mathrm{D}}$. We can check that our asymptotic results satisfy this bound. This bound is not tight: for instance when $\alpha \in (1/2, 1)$, $\langle N_c(b=1)\rangle \sim S$ in 2D but $\langle N_c(b=1)\rangle \sim S^{\chi/(2\alpha)} \ll S$ in 1D.

F. Numerical methods

In this section we describe the numerical techniques used in the extensive solution of the instanton equation. We restrict to $d = 1$.

1. Logarithmic discretisation

To solve the instanton equation numerically, we need to discretise space. The standard uniform mesh is not sufficient to attain the large length scale where the asymptotic behaviors predicted above can be clearly observed. We take a logarithmic mesh

$$x_n = \epsilon \exp(n\delta)\,,\ n = 0, 1, 2, \ldots, N. \tag{102}$$

Usually, we choose $\epsilon = 1$, the mesh size $\delta = 0.1$ or $\delta = 0.2$ is sufficient fine, and $N = 200 \sim 400$ should be chosen appropriately to avoid finite size effects. For example, for the critical regime, one should see a clear mass cutoff. The value of b is chosen to be $2x_n$ for some n.

Then, we approximate F to be piece-wise linear in the intervals $|x| \in [-x_0, x_0]$ and $|x| \in [x_n, x_{n+1}]$. We also assume $F(x) = F(-x)$. When calculating $(\mathcal{D}^\alpha F)(x)$ in the interval $|x| \in (x_n, x_{n+1})$, we approximate x to be the middle point $z_n = (x_n + x_{n+1})/2$ for the whole interval:

$$(\mathcal{D}^\alpha F)(x \in [x_n, x_{n+1}]) = \int p_\alpha(z_n - y)(F(y) - F(x))\mathrm{d}y = \sum_{m=-1}^{N} (F(z_m) - F(z_n))K_{nm}\,, \tag{103}$$

$$K_{nm} = \int_{|y|\in[x_m, x_{m+1}]} p_\alpha(z_n - y)\mathrm{d}y \tag{104}$$

where we have set $z_{-1} = 0$ (it corresponds to the interval $[-x_0, x_0]$. The matrix elements can be explicitly calculated using $p_\alpha(x) = |x|^{-1-\alpha}$, for $n \neq m$. We do not need the ones for $n = m$.

It should be noted that the logarithmic mesh amounts to different way of cutting off $p_\alpha(x)$ at short distances: the cutoff effectively depends on the position. This has an undesired effect for $\alpha > 2$: we do not get the short distance term $\propto |k|^2$ automatically, and have to add it by hand: $K \to K + K_{\mathrm{SR}}$, where K_{SR} is the above matrix with $\alpha = 2$, from which we remove all the elements with distance > 1 away from the diagonal (so that K_{SR} is tri-diagonal).

2. Iteration scheme for stationary solution

In critical regime, we need to find the stationary ($t \to \infty$) solution to the instanton equation. We find this by iteration. For this, we can write the instanton equation as

$$\sum_m K_{nm}F(z_m) = (m^2 + \sum_m K_{nm})F(z_n) + F(z_n)^2\,, |z_n| > b/2\,. \tag{105}$$

Then the iteration scheme is as follows:

$$F_0(z_n) = \theta(b/2 - z_n)\,, \tag{106}$$

$$(m^2 + \sum_m K_{nm})F_{j+1}(z_n) + F_{j+1}(z_n)^2 = \sum_m K_{nm}F_j(z_m)\,,\ z_n > b/2\,. \tag{107}$$

that is, for each iteration, we compute the matrix multiplication of the RHS, and then solve the quadratic equation for F_{j+1} (we pick the positive solution). It is not hard to show that $F_j(z)$ increases with j. Since $F \leq 1$ is also bounded from above, the (point-wise) convergence of this procedure is guaranteed. In practice, a few hundred iterations provide a sufficient convergence for all the tests we presented. This corresponds to no more than to a couple of minutes of calculation on a consumer laptop in order to generate each plot of this paper from scratch.

In the supercritical regime, we solve the time-dependent instanton equation by the Euler scheme with $\delta t = 0.01$.

3. Code availability

The code used to generate all the plots of this paper is available by following this link: https://github.com/xcao-phys/cluster.

[1] I.J. Bienaymé, *De la loi de multiplication et de la durée des familles*, Soc. Philomat. Paris Extraits, Sér 5, 37-39 (1845). L'institut 589, 13:131-132. Reprinted in D.G. Kendall, *The genealogy of genealogy: Branching processes before (and after) 1873*, Bull. London. Math. Soc. 7:225-253 (1975).

[2] H. Watson and F. Galton, *On the probability of the extinction of families*, J. Anthropol. Inst. G. B. Irel. 4, 138 (1875).

[3] Bruno Alessandro, Cinzia Beatrice, Giorgio Bertotti, and Arianna Montorsi, "Domain-wall dynamics and barkhausen effect in metallic ferromagnetic materials. i. theory," Journal of Applied Physics **68**, 2901–2907 (1990).

[4] Maury D. Bramson, "Maximal displacement of branching brownian motion," Communications on Pure and Applied Mathematics **31**, 531–581 (1978).

[5] Gordon Slade, "Scaling limits and super-brownian motion," Notices AMS **49**, 1056–1067 (2002).

[6] É. Brunet and B. Derrida, "Statistics at the tip of a branching random walk and the delay of traveling waves," EPL (Europhysics Letters) **87**, 60010 (2009).

[7] Louis-Pierre Arguin, Anton Bovier, and Nicola Kistler, "The extremal process of branching brownian motion," Probability Theory and Related Fields **157**, 535–574 (2013).

[8] Eric Dumonteil, Satya N. Majumdar, Alberto Rosso, and Andrea Zoia, "Spatial extent of an outbreak in animal epidemics," Proceedings of the National Academy of Sciences **110**, 4239–4244 (2013).

[9] Kabir Ramola, Satya N. Majumdar, and Grégory Schehr, "Spatial extent of branching brownian motion," Phys. Rev. E **91**, 042131 (2015).

[10] Kabir Ramola, Satya N. Majumdar, and Grégory Schehr, "Universal order and gap statistics of critical branching brownian motion," Phys. Rev. Lett. **112**, 210602 (2014).

[11] Andrew V. Suarez, David A. Holway, and Ted J. Case, "Patterns of spread in biological invasions dominated by long-distance jump dispersal: Insights from argentine ants," Proceedings of the National Academy of Sciences **98**, 1095–1100 (2001).

[12] James K. M. Brown and Mogens S. Hovmøller, "Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease," Science **297**, 537–541 (2002).

[13] Ran Nathan, "Long-distance dispersal of plants," Science **313**, 786–788 (2006).

[14] Dirk Brockmann, "Human mobility and spatial disease dynamics," in *Reviews of Nonlinear Dynamics and Complexity* (John Wiley & Sons, Ltd, 2009) Chap. 1, pp. 1–24.

[15] Marta C. González, César A. Hidalgo, and Albert-László Barabási, "Understanding individual human mobility patterns," Nature **453**, 779–782 (2008).

[16] Prasad Perlekar, Roberto Benzi, David R. Nelson, and Federico Toschi, "Population dynamics at high reynolds number," Phys. Rev. Lett. **105**, 144501 (2010).

[17] Vittoria Colizza, Alain Barrat, Marc Barthélemy, and Alessandro Vespignani, "The role of the airline transportation network in the prediction and predictability of global epidemics," Proceedings of the National Academy of Sciences **103**, 2015–2020 (2006), https://www.pnas.org/doi/pdf/10.1073/pnas.0510525103.

[18] J. R. Rice, "First-Order Variation in Elastic Fields Due to Variation in Location of a Planar Crack Front," Journal of Applied Mechanics **52**, 571–579 (1985).

[19] Huajian Gao and James R. Rice, "A First-Order Perturbation Analysis of Crack Trapping by Arrays of Obstacles," Journal of Applied Mechanics **56**, 828–836 (1989).

[20] Anne Tanguy, Matthieu Gounelle, and Stéphane Roux, "From individual to collective pinning: Effect of long-range elastic interactions," Phys. Rev. E **58**, 1577–1590 (1998).

[21] D. Bonamy, S. Santucci, and L. Ponson, "Crackling dynamics in material failure as the signature of a self-organized dynamic phase transition," Phys. Rev. Lett. **101**, 045501 (2008).

[22] J. F. Joanny and P. G. de Gennes, "A model for contact angle hysteresis," The Journal of Chemical Physics **81**, 552–562 (1984).

[23] Sébastien Moulinet, Alberto Rosso, Werner Krauth, and Etienne Rolley, "Width distribution of contact lines on a disordered substrate," Phys. Rev. E **69**, 035103 (2004).

[24] P. Le Doussal, K. J. Wiese, S. Moulinet, and E. Rolley, "Height fluctuations of a contact line: A direct measurement of the renormalized disorder correlator," EPL (Europhysics Letters) **87**, 56001 (2009).

[25] Jean-Christophe Baret, Damien Vandembroucq, and Stéphane Roux, "Extremal model for amorphous media plasticity," Phys. Rev. Lett. **89**, 195506 (2002).

[26] Jie Lin, Edan Lerner, Alberto Rosso, and Matthieu Wyart, "Scaling description of the yielding transition in soft amorphous solids at zero temperature," Proceedings of the National Academy of Sciences **111**, 14382–14387 (2014).

[27] Oskar Hallatschek and Daniel S. Fisher, "Acceleration of evolutionary spread by long-range dispersal," Proceedings of the National Academy of Sciences **111**, E4911–E4919 (2014).

[28] Shirshendu Chatterjee and Partha S. Dey, "Multiple phase transitions in long-range first-passage percolation on square lattices," Communications on Pure and Applied Mathematics **69**, 203–256 (2016).

[29] Xiangyu Cao, Alberto Rosso, Jean-Philippe Bouchaud, and Pierre Le Doussal, "Genuine localization transition

in a long-range hopping model," Phys. Rev. E **95**, 062118 (2017).
[30] Haye Hinrichsen, "Non-equilibrium critical phenomena and phase transitions into absorbing states," Advances in Physics **49**, 815–958 (2000).
[31] Hans-Karl Janssen and Olaf Stenull, "Field theory of directed percolation with long-range spreading," Phys. Rev. E **78**, 061117 (2008).
[32] Peter Grassberger, "SIR epidemics with long-range infection in one dimension," Journal of Statistical Mechanics: Theory and Experiment **2013**, P04004 (2013).
[33] Peter Grassberger, "Two-dimensional sir epidemics with long range infection," Journal of Statistical Physics **153**, 289–311 (2013).
[34] See Supplemental Material.
[35] For a recent review see Pierre Le Doussal, *Equivalence of mean-field avalanches and branching diffusions: From the Brownian force model to the super-Brownian motion*, ArXiv:2203.10512 (2022).
[36] P. Le Doussal and K. J. Wiese, "Distribution of velocities in an avalanche," EPL (Europhysics Letters) **97**, 46004 (2012).
[37] Pierre Le Doussal and Kay Jörg Wiese, "Avalanche dynamics of elastic interfaces," Phys. Rev. E **88**, 022106 (2013).
[38] Knut Jørgen Måløy, Stéphane Santucci, Jean Schmittbuhl, and Renaud Toussaint, "Local waiting time fluctuations along a randomly pinned crack front," Phys. Rev. Lett. **96**, 045501 (2006).
[39] Lasse Laurson, Stephane Santucci, and Stefano Zapperi, "Avalanches and clusters in planar crack front propagation," Phys. Rev. E **81**, 046116 (2010).
[40] Clément Le Priol, Pierre Le Doussal, and Alberto Rosso, "Spatial clustering of depinning avalanches in presence of long-range interactions," Phys. Rev. Lett. **126**, 025702 (2021).
[41] Deniz Ertaş and Mehran Kardar, "Critical dynamics of contact line depinning," Phys. Rev. E **49**, R2532–R2535 (1994).
[42] Pierre Le Doussal, Kay Jörg Wiese, and Pascal Chauve, "Two-loop functional renormalization group theory of the depinning transition," Phys. Rev. B **66**, 174201 (2002).
[43] Alberto Rosso and Werner Krauth, "Roughness at the depinning threshold for a long-range elastic string," Phys. Rev. E **65**, 025101 (2002).
[44] Thimothée Thiery, Pierre Le Doussal, and Kay Jörg Wiese, "Spatial shape of avalanches in the brownian force model," Journal of Statistical Mechanics: Theory and Experiment **2015**, P08019 (2015).
[45] A similar but less precise description applies to 2D if we replace a gap of length g by an "empty space" of area g^2.
[46] R. A. Fisher, "The wave of advance of advantageous genes," Annals of Eugenics **7**, 355–369 (1937).
[47] A. Kolmogorov, I. Petrovsky, and N. Piscounov, "Etude de l'équation de la diffusion avec croissance de la quantit é de matière et son application à un problème biologique," Bull. Univ. Etat Moscou A **1**, 1 (1937).
[48] D. A. Dawson, "Stochastic evolution equations and related measure processes," Journal of Multivariate Analysis **5**, 1–52 (1975).
[49] Shinzo Watanabe, "A limit theorem of branching processes and continuous state branching processes," Journal of Mathematics of Kyoto University **8**, 141 – 167 (1968).
[50] Xavier Cabré and Jean-Michel Roquejoffre, "The Influence of Fractional Diffusion in Fisher-KPP Equations," Communications in Mathematical Physics **320**, 679–722 (2013).
[51] Clément Le Priol, *Long-range interactions in the avalanches of elastic interfaces*, Theses, Université Paris sciences et lettres (2020).
[52] Dina Mistry, Maria Litvinova, Ana Pastore y Piontti, Matteo Chinazzi, Laura Fumanelli, Marcelo F. C. Gomes, Syed A. Haque, Quan-Hui Liu, Kunpeng Mu, Xinyue Xiong, M. Elizabeth Halloran, Ira M. Longini, Stefano Merler, Marco Ajelli, and Alessandro Vespignani, "Inferring high-resolution human mixing patterns for disease modeling," Nature Communications **12**, 323 (2021).
[53] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz, "Superspreading and the effect of individual variation on disease emergence," Nature **438**, 355–359 (2005).
[54] Martin Meyer, Shlomo Havlin, and Armin Bunde, "Clustering of independently diffusing individuals by birth and death processes," Phys. Rev. E **54**, 5567–5570 (1996).
[55] B. Houchmandzadeh, "Neutral clustering in a simple experimental ecological community," Phys. Rev. Lett. **101**, 078103 (2008).
[56] Joaquin Marro and Ronald Dickman, *Nonequilibrium Phase Transitions in Lattice Models*, Collection Alea-Saclay: Monographs and Texts in Statistical Physics (Cambridge University Press, 1999).
[57] Pavel L Krapivsky, Sidney Redner, and Eli Ben-Naim, *A kinetic view of statistical physics* (Cambridge University Press, 2010).
[58] Pierre Le Doussal, *More on the Brownian force model: avalanche shapes, tip driven, higher d*, ArXiv:2203.10544 (2022).
[59] Leonid Mytnik and Edwin Perkins, "The dimension of the boundary of super-brownian motion," Probability Theory and Related Fields **174**, 821–885 (2019).
[60] Onuttom Narayan and Daniel S. Fisher, "Threshold critical dynamics of driven interfaces in random media," Phys. Rev. B **48**, 7030–7042 (1993).
[61] Stefano Zapperi, Pierre Cizeau, Gianfranco Durin, and H. Eugene Stanley, "Dynamics of a ferromagnetic domain wall: Avalanches, depinning transition, and the barkhausen effect," Phys. Rev. B **58**, 6353–6366 (1998).
[62] Alexander Dobrinevski, Pierre Le Doussal, and Kay Jörg Wiese, "Avalanche shape and exponents beyond mean-field theory," EPL (Europhysics Letters) **108**, 66002 (2014).