Supporting Information

Hallatschek and Fisher 10.1073/pnas.1404663111

SI Text

SI1. Simulation Details

A. Simulation Algorithm. For one-dimensional simulations, the state of the population is described by a linear array of N sites with periodic boundary conditions. N is chosen large enough so that end effects can be ignored (typically between 10^7 and 10^8 sites). Each site has the identity of either mutant or wild type. Initially, the whole population is wild type except for the central site, which is occupied by mutants.

In each computational time step, a source site A and target site B are chosen randomly such that their distance r is sampled from a probability density function with a tail $\mu r^{-(1+\mu)}$ at large r (see below). If A is mutant and B a wild type, then B turns into mutants and seeds a new mutant cluster. We use the convention that N time steps—i.e., an average of one jump attempt per site—comprise 1 unit of time or effective "generation." The rate of long-range jumps should be thought of as representing the product of the probability to establish a new cluster per jump and the jump rate per generation per site.

B. Jump Size Distribution. In our simulations, the distance X of a long-range jump was generated as follows. First, draw a random number Y within $(0, 1)$ and calculate the variable

$$X = [Y(L^{-\mu} - C^{-\mu}) + C^{-\mu}]^{-1/\mu}, \qquad \textbf{[S1]}$$

where C is a cutoff (see below) and L is the system size. This generates a continuous probability density function

$$\Pr(X = x) = x^{-(\mu+1)} \frac{\mu(CL)^{\mu}}{L^{\mu} - C^{\mu}} \qquad \textbf{[S2]}$$

with x values in (C, L). The actual jump distance is obtained from X by rounding down to the next integer. [Note that, because the distribution has a tail $\mu r^{-(1+\mu)}$, we have to choose $\epsilon = \mu$ in Eq. **2** of the main text.]

For the one-dimensional data in all graphs of this paper, we used $C = 1$ and system sizes ranging from $L = 10^9$ for $\mu = 0.6$ to $L = 10^8$ for $\mu = 1.4$. For such large systems, the tail of the distribution is well approximated by $p(x) \sim \mu x^{-(\mu+1)}$, as stated in the main text. We also tested variations in the cutoff C. Using $C = 10$ or $C = 100$ affected only the short-time dynamics and had very little influence on the intermediate asymptotic or long-distance behavior of the system.

For our 2D simulations, we draw jump sizes from the same distribution as the one described above and round down to the next possible site. We set the lower cutoff to $C = 1.5 > \sqrt{2}$ to make sure that jumps reach out of the source lattice point. After the jump size is drawn, the jump direction is chosen at random.

SI2. Iterative Scaling Approximation: Details and Extensions

In the main text, *Crossovers and Beyond Asymptopia*, we used a saddle-point approximation (Laplace method) to obtain a recurrence relation, Eq. **7**, for the core radius of the mutant population as a function of time. For this purpose, we assumed that length and time were measured in units of elementary crossover scales at which the growth law changes from linear to superlinear in time.

Here, we first discuss how the crossover scales can be determined, solve the recurrence relation explicitly, and work out corrections to the saddle-point approximation. Finally, we present an improved version of our geometrical argument of the source and target funnels (Fig. 4), which also allows us to estimate the probability of occupancy outside the core region.

A. Crossover Scales. The crossover scales $\ell_\times$ and $t_\times$ from linear to superlinear growth can be obtained explicitly when we assume that the total rate of long jumps is small compared with that of the short-range jumps: i.e., when ε, the coefficient of $G(r) \approx \epsilon / r^{d+\mu}$, is small. Short jumps result in diffusive motion and linear growth $\ell(t) \approx v_0 t$ with v_0 determined by the details of the selective and diffusive dynamics. Long jumps start to become important after enough time has elapsed that there have been at least some jumps of lengths of order $\ell(t)$: i.e., when $\epsilon t \ell(t)^d / \ell(t)^{\mu} \gg 1$; this occurs after a crossover time $t_\times \sim [v_0^{\mu-d} / \epsilon]^{1/(d+1-\mu)}$ at which point $\ell \sim \ell_\times \sim [v_0/\epsilon]^{1/(d+1-\mu)}$. At longer times and distances, we can measure lengths and times in units of these crossover scales, defining $\lambda \equiv \ell/\ell_\times$ and time $\theta = t/t_\times$, and expect that the behavior in these units will not depend on the underlying parameters. Note that this separation in short-time linear growth and long-time regimes can also be done for more general $G(r)$ although then the behavior will depend on the whole function—the crossover on distances of order $\ell_\times$ and the superlinear behavior on the longer distance form.

B. Solving the Recurrence Relation for the Core Radius. The recurrence relation Eq. **7** for the rescaled core radius $\lambda(t) \equiv \ell(t)/\ell_\times$ of the mutant population in terms of the rescaled time $\theta \equiv t/t_\times$,

$$\lambda^{2d+\delta}(\theta) \sim \theta \lambda(\theta/2)^{2d}, \qquad \textbf{[S3]}$$

is valid in the vicinity of the phase boundary $\delta \equiv \mu - d = 0$. We now show how the solution, quoted in [**8**], and the associated prefactors of the asymptotic growth laws can be obtained.

Defining $\varphi \equiv \log_2(\lambda)$ and $z \equiv \log_2(\theta)$ and taking the binary logarithm, $\log_2$, of Eq. **S3**, we obtain the linear recurrence relation

$$(2d + \delta)\varphi(z) \approx 2d\ \varphi(z-1) + z. \qquad \textbf{[S4]}$$

Note that in our conventions we use $\sim$ (as in Eq. **S3**) for asymptotically goes as, with unknown coefficient [i.e., loosely similar to $O(\ldots)$ notation], and $\approx$ (as in Eq. **S4**) if we know the coefficient; i.e., the ratio goes to unity.

Now, it is straightforward to see that

$$Q(z) \equiv \frac{z}{\delta} - \frac{2d}{\delta^2} \qquad \textbf{[S5]}$$

is a special solution of [**S4**]. Substituting $\varphi(z) = \tilde{\varphi}(z) + Q(z)$, we are left with the homogeneous problem

$$(2d + \delta)\tilde{\varphi}(z) \approx 2d\ \tilde{\varphi}(z-1), \qquad \textbf{[S6]}$$

which is easily solved by

$$\tilde{\varphi}(z) \approx \left(1 + \frac{\delta}{2d}\right)^{-z} \tilde{\varphi}(0). \qquad \textbf{[S7]}$$

Reinserting $\varphi(z)$ and imposing the initial condition $\varphi(0) = 0$ finally yields Eq. **8** of the main text,

$$\frac{\delta^2}{2d}\varphi(z) \approx \frac{\delta z}{2d} + \left(1+\frac{\delta}{2d}\right)^{-z} - 1. \quad \textbf{[S8]}$$

In the limit $z \to \infty$, the second and first terms dominate for $\delta > 0$ and $\delta < 0$, respectively. The resulting asymptotics

$$\log[\lambda(t)] \approx \begin{cases} B_\mu t^\eta, & \eta = \log\frac{2d/(d+\mu)}{\log 2}, \quad \delta > 0 \\ \log\left[A_\mu t^\beta\right], & \beta = \frac{1}{\mu - d}, \quad \delta < 0 \end{cases} \quad \textbf{[S9]}$$

reveal the prefactors $B_\mu = 2d\log(2)\delta^{-2}$ and $\log A_\mu = -2d\log(2)\delta^{-2}$ quoted in the main text. Note that we use the variable β throughout the *SI Text* as a shorthand for the power-law exponent $\beta = 1/(\mu - d)$.

C. Subdominant Corrections from Time Integrals. In analyzing the iterative scaling approximation in the main text, we have effectively replaced the integrand in the time integral of Eq. **4** by its value at the half-time peak that dominates the probabilities of occupation at time T. (This procedure resulted in the simple recursion relation **[8]**, which we have explicitly solved in the previous section *SI2.B*.) In doing so, we have ignored effects associated with the parameter-dependent width of the peak around $T/2$ that contributes substantially to the integral. This is valid for obtaining the leading behaviors of $\log \ell(t)$ in the large time limit, but there are corrections to these that can be larger than those that arise from the short-time small-length-scale crossovers that we discussed in the main text and we discuss them further below. For μ not much smaller than $d+1$, when the growth of ℓ is a modest power of time, the factor from the range of the time integral is only of order unity and hence no worse than other factors—including from the stochasticity—that we have neglected. However, when $\ell(t)$ grows very rapidly, the range of $(1/2)T - t$ that dominates is much smaller than T and the corrections are larger.

With rapid growth of $\ell(t)$, a saddle-point approximation to the time integral is valid: $\int_0^t dt \ell(t)\ell(T-t) \approx T c_{T/2} \ell^2(T/2)$ with the prefactor given by

$$c_t \approx \sqrt{\frac{2\pi}{-2t^2\partial_t^2 \log \ell(t)}}, \quad \textbf{[S10]}$$

from which, with the second derivative absorbing the t^2 factor, the derivative part can be rewritten as $t^2\partial_t^2 \log \ell(t) = -\partial_{\log t}\log \ell + \partial_{\log t}^2 \log \ell$. With the asymptotic growth laws we have derived, this gives $c \sim \sqrt{\mu - d}$ for $\mu - d$ small and positive, $c_t \sim 1/\sqrt{\log t}$ for $\mu = d$, and $c_t \sim t^{-\eta/2}$ for $\mu < d$. Integrating up the effects of this over the scales yields the following corrections in the various regimes: For $\mu > d$, the coefficient, A_μ, of t^β is changed by a multiplicative factor that is much less singular for $\mu \searrow d$ than that already obtained. For $\mu = d$,

$$\log[\ell(t)] \approx C\left[\log^2 t - \log t \log \log t + \mathcal{O}(\log t)\right] \quad \textbf{[S11]}$$

with $C = 1/4d \log 2$, the second term being new and the smaller correction term including the effects of the small time crossover. For $0 < \mu < d$,

$$\log[\ell(t)] \approx B_\mu t^\eta - \frac{1-\eta/2}{d-\mu}\log t \quad \textbf{[S12]}$$

with the second term for $\mu \nearrow d$ just what occurs in the crossover regime analyzed in *Crossovers and Beyond Asymptopia* (main text), where we showed that the coefficient B_μ diverges proportional to $1/(d-\mu)^2$.

D. Prefactors in Power-Law Regime. In the main text, we mainly focused on regimes in which the mutant growth is very much faster than linear; i.e., $\mu \lesssim d + 1/2$. This allowed us to approximate the integrals in the iterative scaling approximation of Eq. **1** (main text) by the use of Laplace's method. This saddle-point approximation yields the correct scaling for all exponents $\mu < d$, but (as we see below) incorrect prefactors in regimes where the actual growth is close to linear, i.e., in the power-law growth regime with $d + 1 > \mu \gtrsim d$.

To obtain a better estimate of the prefactors in this power-law regime, it is helpful to directly solve for the asymptotics of the iterative scaling argument in Eq. **1** (main text). Here, we demonstrate how this can be done for $d = 1$: Assume that most of the weight in the integral comes from regions where the jump kernel is well approximated by its power-law tail described in Eq. **2** (main text). Given $\mu < 2$ (for $d = 1$), this always holds at sufficiently long times. Then, we have

$$\epsilon \int_0^t dt' H(t') \sim 1, \quad \textbf{[S13]}$$

where

$$\begin{aligned} \mu(\mu-1)H(t') \equiv & \left(\ell(t) - \ell(t') - \ell(t-t')\right)^{1-\mu} \\ & - \left(\ell(t) - \ell(t') + \ell(t-t')\right)^{1-\mu} \\ & + \left(\ell(t) + \ell(t') + \ell(t-t')\right)^{1-\mu} \\ & - \left(\ell(t) + \ell(t') - \ell(t-t')\right)^{1-\mu}. \end{aligned} \quad \textbf{[S14]}$$

For $1 < \mu < 2$, Eq. **S13** exhibits an asymptotic power-law solution

$$\ell(t) = A_\mu (\epsilon t)^{1/(\mu-1)}. \quad \textbf{[S15]}$$

By inserting this ansatz into Eqs. **S13** and **S14**, we obtain the following result for the numerical prefactor,

$$A_\mu^{\mu-1} \approx \int_0^1 dz \tilde{R}(z), \quad \textbf{[S16]}$$

with $\tilde{R}(z)$ being equal to $H(t)$ in Eq. **S14** with $\ell(t)$ replaced by $z^{1/(\mu-1)}$.

The resulting prefactor is plotted as a function of $\mu - 1$ in Fig. S3. Note that A_μ strongly depends on the exponent μ. It sharply drops for μ approaching 1, where it follows the asymptotics $2^{-2(\mu-1)^{-2}}$. On the other hand, as μ approaches the other marginal case at $\mu = 2$, the prefactor diverges as $A_\mu \sim (2-\mu)^{-1}$, indicating the importance of intermediate asymptotic regimes, as discussed in *Crossovers and Beyond Asymptopia* (main text).

Although we have focused on the marginal case near $\mu = d$ in this article, it is clear that another case of marginality controls the crossovers near $\mu = d + 1$. Simulation results reported in Fig. S4 indicate that $\ell(t)/t \sim \log(t)$ for $\mu = 2$ in one dimension. This is consistent with our funnel argument: With nearly constant speed, the gap between the funnels remains roughly constant for a time of order t. To ensure that the source emits about one jump to the target funnel, we must have that, per unit of time, the probability of a seed jumping over the gap is of order $1/t$. Thus, the gap size ΔE should be such that $\Delta E^{-\mu+1} \sim 1/t$. For $\mu = 2$, we thus have $\Delta E \sim t$; i.e., the key jumps span distances of order t. This is ensured when $(\ell(t) - 2\ell(t/2))/t \sim$ const., i.e., if $\ell(t)/t \sim \log t$. Note that a rigorous upper bound of this form follows from the arguments presented in *SI Text*, section *SI3.A.4* for the regime $\mu > d$. The jumps of order $O(t)$ that drive the logarithmic increase in spreading velocity might be the "leaps forward" (1) recognized by Mollison in one of the earliest studies on spreading with long-range jumps.

E. Occupancy Profiles and Relevance of Secondary Seeds. In the main text, we introduced the notion of a nearly occupied core of the population [of size $\ell(t)$] as the source of most of the relevant seeds in the target funnel. However, it is clear that outside of this core there is a region of partial occupancy. This region is potentially broad, in particular for $\mu \to 0$, and may therefore lead to a significant fraction of relevant seeds. An improved theory should account for those secondary seeds and should also be able to determine the profile of mean occupancy or, equivalently, the probability that a site is occupied. Whereas we give rigorous bounds on how the total population grows in *SI Text*, section *SI3*, we first use an improved version of our funnel argument to describe the occupancy profiles.

We focus on the probability, $q(r,t)$, that at time t after a mutant establishes, it will have taken over the population a distance r away. We expect that $q(r,t)$ will be close to unity out to some core radius $\ell(t)$ and then decrease for larger r, with the average total mutant population proportional to $\ell(t)^d$. With only short-range dispersal, $\ell(t) \approx vt$ and the core is clearly delineated but when long jumps are important, the crossover from mostly occupied core to sparsely occupied halo will not be sharp. The more important quantity is the average of the total area (in two dimensions or linear extent or volume in one or three dimensions) occupied by the mutant population; we denote this $M(t) = \int d^d r q(r,t)$.

To find out when long jumps could be important, we first ask whether there are likely to be any jumps longer than $\ell(t)$ that occur up to time t. The average number of such long jumps is of order $t\ell(t)^d \int_{\ell(t)}^{\infty} r^{d-1} dr G(r)$. If $G(r)$ decreases more rapidly than $1/r^{2d+1}$, this is much less than $t/\ell(t)$ for large t. As $\ell(t)$ increases at least linearly in time, the probability that there have been any jumps longer than $\ell(t)$ is very small. The guess that ℓ indeed grows as vt, and consideration of jumps that could advance the front fast enough to contribute substantially to v, leads, similarly, to the conclusion that there is a maximum t-independent jump length beyond which the effects of jumps are negligible; indeed, their effect decreases more rapidly than $G(r)$. This reinforces the conclusion that there is only linear growth with $\mu > d+1$: a very strong breakdown of the deterministic approximation that yielded exponential growth for any power law.

When G is longer range, in particular if $G(r) \sim 1/r^{d+\mu}$ with $\mu < d+1$, many jumps longer than $\ell(t)$ will have occurred by time t. We now study the effects of such long jumps on the density profile. To do so, we investigate the behavior of $q(R,T)$ for large R and T in terms of the $\{q(r,t)\}$ at shorter times and—primarily—corresponding distances $r \sim \ell(t)$ that can be much less than R. Mutants can get to a chosen point, $\mathbf{R}$, by one making a long jump at time t from a starting point $\mathbf{x}$ to an end point, $\mathbf{y}$, and subsequently spreading from there to $\mathbf{R}$ during the remaining time interval of duration $T-t$. The rate (per volume elements) of this occurring is $q(x,t)G(|\mathbf{x}-\mathbf{y}|)q(|\mathbf{R}-\mathbf{y}|, T-t)$. In the approximation that these are independent, the probability that this does not occur at any $t < T$ from any $\mathbf{x}$ to any $\mathbf{y}$ is simply Poisson so that

$$q(R,T) \approx 1 - e^{-Q(R,T)} \qquad \text{[S17]}$$

with

$$Q(R,T) \approx \int_0^T dt \int_{z>\ell(T/2)} d^d z \int d^d x \, q(x,t) G(z) q(|\mathbf{R}-\mathbf{x}-\mathbf{z}|, T-t). \qquad \text{[S18]}$$

Here, we substituted the final point $\mathbf{y} = \mathbf{x} + \mathbf{z}$ by the sum of the jump start site and a jump vector $\mathbf{z}$, over which we integrate. Note that a lower cutoff in the z integral is necessary to exclude the many very short jumps that lead to strongly correlated establishments. The cutoff is also necessary to not count mutants that result from growth in the target area, rather than seeding from the source funnel. Our main assumption here is that if a single seed is sufficiently far from other seeds or occupied regions, then the growth from the seed is independent of the rest of the system as long as collisions are unlikely.

When the jump integral is strongly peaked at $z = R$, as is the case in or close to the stretched exponential regime, the final results will be independent of this cutoff to leading order. Then, we can approximate $Q(R,t)$ as

$$Q(R,t) \approx G(R) \int_0^t dt' M(t') M(t-t'), \qquad \text{[S19]}$$

where $M(t)$ is the expected total size of a population at a time t,

$$M(t) = \int d^d x q(x,t). \qquad \text{[S20]}$$

For a power-law kernel $G(R) = G(1) R^{-(1+\mu)}$, we make the ansatz that Eqs. **S19**, **S20**, and **S17** can be approximately solved by a scaling form

$$q(R,t) = \Xi\left(\frac{R}{\lambda(t)}\right) \qquad \text{[S21]}$$

with

$$\Xi(\xi) = 1 - \exp\left(-\xi^{-(d+\mu)}\right), \qquad \text{[S22]}$$

which leads to the condition

$$Q(\xi\lambda(t), t) = \xi^{-(d+\mu)} \kappa^2 G(1) \lambda^{-(d+\mu)} \int_0^t dt' \lambda^d(t') \lambda^d(t-t'), \qquad \text{[S23]}$$

where $M(t) = \kappa\lambda(t)^d$ and κ_μ is given by

$$\kappa = \int d\xi^d \, \Xi(\xi). \qquad \text{[S24]}$$

Thus, the above scaling form is a valid solution if the characteristic scale $\lambda(t)$ satisfies

$$\kappa^2 G(1) \lambda^{-(1+\mu)} \int_0^t dt' \lambda(t') \lambda(t-t') = 1. \qquad \text{[S25]}$$

The resulting condition is similar to our condition in the main text but differs by the numerical factor $G(1)\kappa_\mu^2$. In one dimension,

$$\kappa_\mu = 2\Gamma\left(\frac{\mu}{1+\mu}\right). \qquad \text{[S26]}$$

The divergence $\kappa_\mu^2 \sim \mu^{-2}$ as $\mu \to 0$ indicates the importance of seeds from the tail regions for small μ.

Note that at long distances, $R \gg \ell(T)$, the lengths of the jumps that dominate are close to R so that our approximation for $Q(R,T)$ should be correct even if it is a poor approximation for $R \approx \ell(T)$. Thus, the decrease in q at large distances is simply proportional to $G(R)$; more specifically,

$$q(R,T) \approx \frac{G(R)}{G(\ell(T))} \quad \textbf{[S27]}$$

(as predicted by the scaling form) so that it is of order unity at $R \approx \ell(T)$. Note that this implies that, because $G(r)$ is integrable, $\int d^d r q(r,t)$ is indeed dominated by $r \sim \ell(t)$ as we have assumed. This long-distance form for the density profile is also found in the analyses of upper and lower bounds in the next sections: Thus it can be readily proved along the same lines.

SI3. Rigorous Bounds and Outlines of Routes to Proofs

As most of our results are based on approximate analyses and heuristic arguments, it is useful to supplement these by some rigorous results. We focus on the one-dimensional case: Extensions to higher dimensions can be done similarly, although with a few complications that will require some care. We sketch here the arguments that can lead to proofs without all of the details filled in.

As via the heuristic arguments, we want to obtain the behavior at longer times in terms of the behavior at shorter times, in particular times around half as long.

We want to prove that there exist time-dependent length scales, $\ell_<(t)$ and $\ell_>(t)$, and functions, $F_<(r,t)$ and $F_>(r,t)$, such that the probability, $q(r,t)$, that a site at r from the origin is occupied at time t, is bounded above and below by

$$F_<(r,t) < q(r,t) < F_>(r,t) \quad \text{with} \quad M_< \equiv \int dr F_<(r,t) \propto \ell_<(t)$$
$$\text{and} \quad M_> \equiv \int dr F_>(r,t) \propto \ell_>(t) \quad \textbf{[S28]}$$

for all times. Then $\ell_<$ and $\ell_>$ are lower and upper bounds for $\ell(t)$—with some appropriately chosen definitions of $\ell(t)$ that differ somewhat, although not significantly, for the upper and lower bounds. Although we want the upper and lower bounds on $\ell(t)$ to be as close as possible to each other, in practice, we have obtained bounds that are good on a logarithmic scale: i.e., for $\log \ell(t)$, rather than on a linear scale. Similarly, we want to have the bounds be close to the actual expected form of q, with $F_<$ very close to unity for $r \ll \ell_<$ and proportional to $[\ell_</r]^{\mu+1}$ for $r \gg \ell_<$ and similar for the upper bounds.

It is often more convenient to consider the typical time to occupation as a function of the distance, $\tau(r)$, and derive upper and lower bounds for this, $\tau_>(r)$ and $\tau_<(r)$, respectively, such that

$$\ell_<(t=\tau_>(r)) = r \quad \text{and} \quad \ell_>(t=\tau_<(r)) = r \quad \textbf{[S29]}$$

with

$$\tau_>(r) > \tau(r) > \tau_<(r). \quad \textbf{[S30]}$$

Because of the faster than linear growth, bounds on $\tau(r)$ are generally much closer than those on $\ell(t)$.

As we want to justify the use of the heuristic iterative scaling arguments more generally, it is especially useful to obtain iterative bounds directly of the form used in those heuristic arguments: $\ell(T)$ in terms of $\{\ell(t)\}$ for t in a range near $T/2$. As the heuristic arguments do, in any case, give $\ell(T)$ only up to a multiplicative coefficient of order unity, we will generally ignore such order-unity coefficients in length scales except for coefficients that diverge or vanish exponentially rapidly as $\mu \to d$, in particular in the intermediate-range regime the coefficient, A_μ in $\ell(t) \sim A_\mu t^{1/(\mu-d)}$, which vanishes as $\log(A_\mu) \approx -\log 4/(\mu-d)^2$ as $\mu \searrow d$.

A. Upper Bounds.

1. Simple power-law bound. The simplest bound to obtain is an upper bound for $\ell(t)$ in the short- and intermediate-range regimes: i.e., in one dimension, $\mu > 1$. Define $E(t)$ to be the rightmost edge of the occupied region at time t; i.e., $c(x,t)=0$ for $x > E(t)$. The probability of a jump that fills a position $y > E(t)$ in $(t, t+dt)$ is less than $\int_{-\infty}^{E(t)} dx G(y-x) \sim 1/(y-E(t))^\mu$. For $\mu > 1$, the lower extent of the integral can be taken to $-\infty$ as the jumps arise, predominantly, from points that are not too far from the edge. [In contrast, for $\mu < 1$ jumps from the whole occupied region are important and this bound would yield a total jump probability to long distances that diverged when integrated over y, and we would have to instead use a lower extent of the integral of $-E(t)$ for the left edge.]

The advancement of the edge is bounded by a translationally and temporally invariant process of jumps of the position of the edge by distances, ΔE, whose distribution has a power-law tail. For $\mu > 2$, the mean $\langle \Delta E \rangle < \infty$, implying that the edge, and hence $\ell(t)$, cannot advance faster than linearly in time. However, for the intermediate regime, $\langle \Delta E \rangle > \infty$ so that $E(t)$ could advance as fast as a one-sided Levy flight with $E(t)$ dominated by the largest advance. As this process would yield $E \sim t^{1/(\mu-1)}$, this implies that $\ell(t)$ is bounded above by the same form as the heuristic result.

Although the simple bound captures some relevant features, in particular the dominance of jumps of length of order r to fill up a point at distance r, it is otherwise rather unsatisfactory. First, the coefficient does not vanish rapidly as $\mu \searrow 1$. And second, it suggests that the probability that an anomalously distant point, $r \gg \ell(t)$, is occupied, is, in this crude approximation of full occupancy out to the edge, simply the probability that $E(t) > r$, which falls off only as $1/r^{\mu-1}$—much more slowly than the actual $q(r,t) \sim r^{-1-\mu}$.

Nevertheless, for proving better upper bounds, the Levy-flight approximation for the dynamics of the edge is quite useful.

2. Upper bounds from source–jump–target picture. As discussed earlier, we want to make the heuristic argument of a single long jump from a source region to a target funnel region include also—or provide solid reasons to ignore—the effects of jumps from the partially filled region outside the core of the source. Very loosely, we want to write the probability that a point, R, is not occupied at time T, as

$$1 - q(R,T) \approx \exp\left[-\int_0^T dt \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy q(x,t) G(|y-x|) q(R-y, T-t) \right] \quad \textbf{[S31]}$$

with x in the source region and y in the funnel of R. However, for any positive μ, the spatial integral is dominated by $y-x$ small, so that this does not properly represent the process: There is a drastic overcounting of short jumps.

We can do much better by trying to separate the long jumps from the short ones and the source region from the funnel (in the crude approximation these overlap). To do this, we choose, for the R and T of interest, a spatiotemporal source region, S, around the origin that has a boundary at distance $B_S(t)$ that loosely reflects the growing source: $dB_S/dt \geq 0$. We then separate the process of the set of jumps that lead to R into three parts: first, jumps solely inside S, which lead to a spatiotemporal configuration of occupied sites, $\{c_S(x,t)\}$; second, bridging jumps from these out of S, say at time t from x in S to a point y in the rest of space–time, $\overline{S}$; and third, all of the subsequent dynamics from such seeds in $\overline{S}$, including inside and outside S and between these. This overcounts the possible spatiotemporal routes to R, T—especially as returns to inside S from outside are included—and thus provides an upper bound for $q(R,T)$. The probability, p_a, that a single seed, a, to y_a

at t_a leads to R being filled by T is $q(R-y_a, T-t_a)$. However, the probability that a second seed, b, leads to R filled by T is not independent as the fate of these seeds involves overlapping sets of jumps: Indeed, they are positively correlated so that

$$\begin{aligned}\mathcal{P}[c(R,T)=0|\text{seeds } a \text{ and } b] \\ \geq \mathcal{P}[c(R,T) \\ =0|\text{seed } a] \times \mathcal{P}[c(R,T)=0|\text{seed } b].\end{aligned} \quad \textbf{[S32]}$$

Because for a given occupancy profile $\{c_S(r,t)\}$, the probability density of a seed at y,t is $dydt\int_{|x|<B_S(t)} c_S(x,t)G(y-x)$, and using the generalization of the above bound to many seeds, we have

$$q(R,T) \leq 1-\exp\left[-\int_0^T dt \int_{|x|<B_S(t)} dx\, q_S(x,t) \int_{|y|>B_S(t)} dy\, G(y-x) q(R-y, T-t))\right], \quad \textbf{[S33]}$$

where $q_S(x,t) \equiv \langle c_S(x,t)\rangle$ and we have used $\langle \exp(X)\rangle \geq \exp(\langle X\rangle)$ for any random variable.

To derive a useful upper bound on $q(R,T)$ we need to choose appropriately the boundary, $B_S(t)$, of the source region and put a sufficiently stringent upper bound on $q_S(x,t)$.

3. Long-range case. For the long-range case, $\mu<1$, the integrals over $x<B_S$ and $y>B_S$ of $G(y-x)$ are dominated by long distances. Thus, the short jumps from inside to outside S do not contribute significantly. We can then simply replace q_S by the larger q to obtain a slightly weaker bound that is of exactly the form of the naive estimate except for the strict delineation of the source region, which prevents the most problematic overcounting of the effects of short jumps. A particularly simple choice is $B_S=(1/2)R$ independent of t.

We now proceed by induction. Take the bound on the scaling function to have the form $F_>(r,t)=1$ for $r<\ell_>(t)$ whereas $F_>(t)=[\ell_>/r]^{\mu+1}$ for $r \gg \ell_>(t)$ and assume that for some appropriate $\ell_>(t)$, this is indeed an upper bound for all $t<T$; i.e., $q(r,t) \leq F_>[r/\ell_>(t)]$. We can now use **[S33]** with q_S and q both replaced by $F_>$.

When $\ell_>(t)$ and $\ell_>(T-t)$ are both much less than R, the integrals over x and y will be dominated by the regions near the origin and R, respectively, yielding the spatial convolution $F_> \circ G \circ F_> \sim \ell_>(t)\ell_>(T-t)/R^{\mu+1}$. There are small positive corrections to this from two sources: first, from the regions near 0 and R, which, by expanding $y-x$ in x and $R-y$, are seen to be of order $[\ell_>(t)^2+\ell_>(T-t)^2]\ell_>(t)\ell_>(T-t)/R^{\mu+3}$; and second, from $y-x \ll R$, the regions near the source boundary, which are of order $[\ell_>(t)/R]^{\mu+1}[\ell_>(T-t)/R]^{\mu+1}R^{1-\mu}$ with the last part from the integrals over x and y. As the dominant part is exactly of the form in the heuristic treatment, integrating it over time is strongly peaked at $t \approx T/2$ (note that for either t or $T-t$ much smaller than T, one of the $F_>$ factors will be close to unity near the boundary, but these ranges of time contribute only weakly). If we use $1-e^{-Q} \leq \min(Q,1)$, then $\ell_>(T)$ can be chosen as the value of R for which $Q=1$, and for $R \gg \ell_>(T)$ the desired $F_> \sim [\ell_>/R]^{\mu+1}$ is obtained. Including the small correction factors in the convolutions necessitates slight modifications of the recursion relations for $\ell_>$ but these are negligible at long times.

4. Intermediate-range case. Obtaining an upper bound in the intermediate-range case is somewhat trickier. If we again replaced q_S by q, then the integrals over x and y would have a part dominated by both points being near the boundary: With $B_S \sim R$, this contribution to the convolution would be of order $[\ell_>(t)/R]^{\mu+1}[\ell_>(T-t)/R]^{\mu+1}$. With $t \sim T/2$ and $R \sim \ell(T)$, all of the lengths should be of order t^β with $\beta=1/(\mu-d)$, so that this boundary piece is larger by a factor of $[\ell(T)^\beta]^{\mu-1} \sim T$ than what should be the dominant part from x and y near 0 and R, respectively. Thus, we need a better upper bound on the restricted-source $q_S(r,t)$, which vanishes as $r \nearrow B_S(t)$.

To bound q_S, we can make use of the simple bound for the edge of the occupied region derived above, combined with the restrictive effects of the boundary, $B_S(t)$. Instead of choosing B_S to be constant, we choose it to have constant slope, $U \equiv dB_S/dt$, of order $\ell(T)/T$. As jumps that contribute to q_S are not allowed to cross the boundary, the distribution of jumps of the edge $E(t)$ is cut off at $\zeta(t) \equiv B_S(t)-E(t)$. Because $U>v_0$, the speed of spread in the absence of jumps beyond nearest neighboring sites, typically the gap, $\zeta(t)$, will increase with time, decreasing only by jumps. The sum of all of the jumps of E in a time interval Δt is dominated by the largest, which is of order $(\Delta t)^{1/(\mu-1)}$. This would result in the edge moving faster than U except for the cutoff. The typical gap, $\tilde{\zeta}$, is then obtained by balancing its steady decrease against the dominant jump: $U\Delta t \sim (\Delta t)^{1/(\mu-1)}$, yielding $\Delta t \sim U^{(\mu-1/)(2-\mu)}$ and hence

$$\tilde{\zeta} \sim U^{1/(2-\mu)} \sim \left[\frac{\ell(T)}{T}\right]^{1/(2-\mu)} \sim \ell(T) A_\mu^{(\mu-1)/(2-\mu)}, \quad \textbf{[S34]}$$

using $\ell(t) \sim A_\mu t^{1/(\mu-1)}$. In the limit of $\mu \searrow 1$, $\tilde{\zeta}/\ell \sim 4^{-1/(\mu-1)}$, vanishing rapidly—a reflection of the strong failure of the simple edge bound in this limit but sufficient for our present purposes. In a time $\Delta t \ll T$, the distribution of ζ in this approximation will reach a steady state. The probability that $\zeta \ll \tilde{\zeta}$ is controlled by the balance between jumps of E to near the boundary, and the steady increase in ζ from the boundary motion: Its probability density is hence of order $\zeta/\tilde{\zeta}$, which, because in this approximation all sites are occupied up to E, implies that q_S vanishes at least quadratically for small gap ζ. Combining this with the trivial bound of $q_S<q$ and choosing a convenient normalization of $\tilde{\zeta}$, we thus have

$$q_S(r,t) \leq \min\left[F_>(r,t), \frac{(B_S(t)-r)^2}{\tilde{\zeta}^2}\right]. \quad \textbf{[S35]}$$

It remains to choose $B_S(t)$ so that the bound on q_S remains sufficiently good for t small enough that the steady-state distribution of $\zeta(t)$ has not yet been reached. To keep $E(t)$ typically of order $\tilde{\zeta}$ from $B_S(t)$, we can simply chose $B_S(0)=\tilde{\zeta}$ and $U=(R-2\tilde{\zeta})/T$.

With our improved bound on q_S, for the convolution $q_S \circ G \circ q$, the small B_S-x parts are no longer dominated by B_S-x of order unity, but by B_S-x near the crossover point between the two bounds on q_S. This yields a contribution to the convolution of order $\ell_>(t)^{\alpha_S}\ell_>(T-t)^{\alpha_F}/R^{\alpha_S+\alpha_F+\mu-1}$ with $\alpha_F=(\mu+1)(2-\mu)$ and $\alpha_S=\alpha_F(3-\mu)/2$ and a multiplicative coefficient that does not depend exponentially on $1/(\mu-1)$ because the integral over x scales as $1/\tilde{\zeta}^{\mu-1}$. As $\mu \searrow 1$, $\alpha_F \to \alpha_S \to 1$, and the boundary contribution is less than the dominant part uniformly in t. Note that for $\mu>\mu_B \cong 1.5$, the bound on the near-boundary contribution can be somewhat larger for $t<T/2$ than the dominant parts, but it scales in the same way with T and thus only weakens the upper bound on the coefficient, A_μ, which is in any case of order unity in this regime.

Once the overcounting of short jumps has been sufficiently reduced, as we have now done, the rest of the analysis, in particular the large $R/\ell_>$ form of $F_>$, follows as in the long-range case.

The marginal case $\mu=1$ can be analyzed similarly to the intermediate-range case, resulting in an additional logarithmic dependence on R of the near-boundary contribution, which is, nevertheless, still much smaller than the dominant part.

The upper bounds that we have obtained are, except for modifications at small scales and for μ not much smaller than 2, essentially the same as given by the heuristic arguments, thus differing at long scales only by order-unity coefficients that, in any case, we did not expect to get correctly. All of the crossover behavior near $\mu=1$ is in the upper bounds, although that near $\mu=2$ is not.

B. Lower Bounds. To obtain lower bounds on the growth of the characteristic length scale $\ell(t)$ and the occupation probability, $q(r,t)$, a different strategy needs to be used. One of the difficulties is the dependence on the behavior at each timescale on all of the earlier timescales: For the filled region to grow typically between times T and $2T$, the stochastic processes that lead to the configuration $c(x,T)$ must not have been atypically slow or ineffective. As this applies iteratively scale by scale, we must allow for some uncertainty in whether the smaller-scale regions are typical, leading to some uncertainty at all scales, which, nevertheless, we need to bound. Because of the stochastic heterogeneity of $c(x,t)$, it is better to focus on a coarse-grained version of the occupation profile rather than on $c(x,t)$ itself, as integrations over c at time t are what act as the sources of future occupation at larger distances.

1. Mostly filled in: Marginal and long-range regimes. We consider the probability that a region is almost full; in particular, with a seed at the origin, we consider the region to one side of the origin and define

$$P_F(r,t;\Phi)\equiv\mathcal{P}\left[\frac{1}{r}\int_0^r dx\,c(x,t)>\Phi\right] \qquad \textbf{[S36]}$$

with Φ close to or equal to unity being of particular interest. To keep events sufficiently independent, we consider, as for the upper bounds, the probability of events that do not involve any jumps out of some region. In particular, we define $P_S(r,t,\Phi)$ similarly to P_F, but with the restriction that jumps do not go out of the interval $(0,r)$. For the long-range and marginal cases, we focus on partial filling, but for the intermediate range case the scale invariance mandates different treatment so we instead analyze full filling—i.e., $\Phi=1$.

The basic strategy is to start with a particular deterministic approximation to $\ell(t)$, $\tilde{\ell}(t)$ with corresponding times $\tilde{\tau}(r)$, and then show that at time not too large a multiple of $\tilde{\tau}(r)$, the region out to r will be nearly filled with high probability: i.e., that $P_S(r,\tau_>(r);\Phi)$ is close to unity for $\tau_>(r)/\tilde{\tau}(r)$ sufficiently large. We are interested in large scales as, in any case, fluctuations at the small scales can change coefficients only by order unity. We can thus be sloppy with some of the bounding inequalities: These could be improved to include the ignored corrections to the large-scale effects to make fully rigorous bounds.

As the range of time over which the typical $\ell(t)$ expands significantly plays an important role, it is useful to define

$$\tilde{D}(t)\equiv\left[\frac{d\log\tilde{\ell}(t)}{d\log t}\right]^{-1}, \qquad \textbf{[S37]}$$

which is small except for μ substantially larger than one. The dominant jumps from source to funnel involve an integral over time of $\tilde{\ell}(t)\tilde{\ell}(T-t)$, which is primarily from a range of order $T\sqrt{\tilde{D}}$ around T as discussed above. The deterministic-iterative approximation that we use as a base for the lower bounds is the solution to the iterative relation [ambiguous up to an $\mathcal{O}(1)$ multiplicative factor, which we ignore throughout]

$$\left[\tilde{\ell}(2T)\right]^{\mu+1}=T\left[\tilde{\ell}(T)\right]^2\sqrt{\tilde{D}(T)}, \qquad \textbf{[S38]}$$

corresponding to roughly one seed into a funnel of width $\tilde{\ell}(T)$ from a jump of distance $\tilde{\ell}(2T)$ from the source up to time T. For convenience, we use only half the source—x from 0 to $\tilde{\ell}(T)$. The results, $\tilde{\ell}(t)$, of this iterative approximation are, up to numerical factors that arise from these modifications and from other from nonasymptotic effects at small scales, equivalent to the upper bounds, $\ell_>(t)$ from the above. In particular, we expect the ratio between the corresponding times, $\tilde{\tau}(r)$ and $\tau_<(r)$, to approach constants that are not singular near the marginal case $\mu=d=1$.

For the lower bounds it is convenient to work with a specific set of length scales, $\tilde{\ell}_n=\tilde{\ell}(\tilde{\tau}_n)$, corresponding to a series of timescales, $\tilde{\tau}_n=2^n$ (dropping a prefactor). To mostly fill out to $\tilde{\ell}_{n+1}$ without jumps going out of $(0,\tilde{\ell}_{n+1})$ from the source of size $\tilde{\ell}_n$, most of the $K_n=\tilde{\ell}_{n+1}/\tilde{\ell}_n$ bins of size $\tilde{\ell}_n$ must be mostly filled. To get a lower bound on how long this takes and how likely it is, we make several simplifications, each of which leads to underestimates of the probability that the desired filling has occurred. First, consider only jumps into each bin that come directly from the source (rather than from other bins as can occur later). Second, ignore all but the first seed jump from the source into the bin (the effects of later jumps are not independent of those of the first). And third, include only jumps that lead from the seed in a bin that do not go outside that bin during the time during which the probability of it being mostly filled is considered. The last two conditions mean that the probability that the bin is filled to a fraction Φ by a given time, t, after the seeding jump, is at least as large as $P_S(\tilde{\ell}_n,t;\Phi)$ because a seed at the edge of the bin, which corresponds to the definition at the source, is less likely to mostly fill the bin than a seed away from the edge.

At large scales for $\mu\le 1$, the number of bins, K_n, grows with scale $K_n\sim\sqrt{\tilde{\tau}_n}$ for the marginal case and larger for the long-range case. Thus, if the probability that the furthest bin from the source is mostly filled is f_n, with the filling of the others being more probable as they are closer, it is likely that the number that are similarly mostly filled is close to K_nf_n, with significant deviations from this being very unlikely at large scales. To iterate while not losing too much in the filling fraction, we chose a series of partial filling fractions, $\{\phi_n\}$, such that $\Phi_N\equiv\prod_{n=1,N-1}\phi_n$ converges to the desired overall filling fraction, Φ, at large N, and chose conditions such that f_n is sufficiently large that the fraction of the K_n bins filled to Φ_n is greater than ϕ_n with high probability: This then implies that the region from the origin to $\tilde{\ell}_{n+1}$ will be filled to greater than Φ_{n+1} with high probability. A convenient choice is $\phi_n=1-\Delta/n^{1+\alpha}$ with any positive α and $\Delta\sum_n n^{-1-\alpha}<1-\Phi$. For convenience in dropping $\log\Delta$ factors that otherwise appear in many places, we restrict consideration to Δ not very small and do not keep careful track of α factors that also appear as we can take $\alpha\to 0$ at the expense of corrections that are down by one extra logarithm.

The filling probability of a bin is at least as large as that obtained from the requirement of the occurrence of both of two independent events: a jump into the bin from the source that occurs before some chosen initial time, T_I, and the bin being filled from that single seed by a time, T_B+T_I. The probability of a jump into a bin is at least $1-e^{-W_n}$ in terms of a conveniently chosen lower bound, W_n, on the expected number of jumps from the source into the farthest away bin, and the probability of the bin being filled from the single seed is at least $P_S(\tilde{\ell}_n,T_B;\Phi_n)$. We find iterative bounds on P_S that are convenient to write in the form

$$P_S(r,t;\Phi)\ge 1-e^{-\Lambda(r,t;\Phi)} \qquad \textbf{[S39]}$$

so that

$$1-f_n\le e^{-W_n}+e^{-\Lambda_B}\quad\text{with}\quad\Lambda_B\equiv\Lambda\left(\tilde{\ell}_n,T_B;\Phi_n\right). \qquad \textbf{[S40]}$$

For convenience we chose conditions so that $\Lambda_B\ge W_n$ and $1-f_n\le(1/2)(1-\phi_n)$, which, for K_n large, makes the probability that a fraction ϕ_n of the bins are not filled exponentially small.

We henceforth ignore this factor in the probability as it does not matter except on small scales: Adjustments to take it into account are straightforward. We thus require that

$$\Lambda_B \geq W_n \geq \log\left[\frac{(1-\phi_n)}{4}\right] = (1+\alpha)\log n + \mathcal{O}(1). \qquad \textbf{[S41]}$$

To obtain a bound on $P_S(\tilde{\ell}_{n+1}, T; \Phi_{n+1})$, we must show that a source that can give rise to an average number at least W_n of jumps into the farthest bin by time T_I occurs with probability that is somewhat larger than the desired bound at the next scale. The expected effective number of jumps out of the source of size $\tilde{\ell}_n$ into a bin of the same size a distance up to $\tilde{\ell}_{n+1}$ away before time $\tilde{\tau}_n$ was assumed in the deterministic iterative approximation to be of order $\sqrt{\tilde{D}_n}\tilde{\tau}_n$. To ensure that the average number from the actual source is sufficiently large, we can require that it be almost filled by some time T_S and include only jumps that occur between T_S and T_I as the rate of these is bounded below by the filling at T_S. The required range is

$$T_I - T_S = \frac{W_n \tilde{\tau}_n \sqrt{\tilde{D}_n}}{\Phi_n}. \qquad \textbf{[S42]}$$

We now proceed by induction and show that if

$$\Lambda\left(\tilde{\ell}_n, t; \Phi_n\right) \geq \gamma_n (t - U_n \tilde{\tau}_n) \qquad \textbf{[S43]}$$

for t in a range such that Λ is relatively large—the precise range is not crucial but minor modifications are needed to extend out to arbitrary large t—then a similar bound holds at the next scale with coefficients γ_{n+1} and U_{n+1} with both these varying slowly with n at large scales. Note that at the smallest scale the probability that a site is filled by a jump directly from the origin by time t converges exponentially to unity for long t, and thus at the smallest scales there is a trivial bound of this form. As the scale is increased, γ_n will initially change, but once the scale becomes large enough that the width of the distribution of the fraction of the bins mostly filled is small, then γ_n saturates and becomes weakly dependent on n. In the analysis below, it can be replaced by a constant.

Consider a total time T to mostly fill out to $\tilde{\ell}_{n+1}$. The time for the bins to fill with sufficiently high probability once they have been seeded is $T_B \leq U_n \tilde{\tau}_n + W_n/\gamma_n$. With $T_I - T_S$ as above, we have a time for the source to fill

$$T_S \geq T - \tilde{\tau}_n \left[U_n + \frac{W_n \sqrt{\tilde{D}_n}}{\Phi_n}\right] - \frac{W_n}{\gamma_n}. \qquad \textbf{[S44]}$$

Plugging in the probability that the source is filled in this time gives a bound on $\Lambda(T, \tilde{\ell}_{n+1}, \Phi_{n+1})$ of the same form but with, dividing out $\tilde{\tau}_{n+1} = 2\tilde{\tau}_n$,

$$U_{n+1} \leq U_n + \frac{W_n \sqrt{\tilde{D}_n}}{2\Phi_n} + \frac{W_n}{2\gamma \tilde{\tau}_n}. \qquad \textbf{[S45]}$$

As $\tilde{\tau}_n$ increases rapidly and W_n only slowly, the last term contributes only at small scales.

For the long-range regime, $\tilde{D}_n \sim e^{-\eta \log 2n}$ so the second term in **[S45]** is also small except at small scales and we conclude that U is bounded above by a μ-dependent constant. Thus, the lower bound for $\ell(t)$ and the upper bound for $\tau(r)$ have exactly the same form as the opposite bounds, except with the scale of t—i.e., $B_\mu^{-1/\eta}$—different.

For the marginal case, $\tilde{D}_n \approx 2/n$ so that U_n changes slowly at large scales. Integrating up, we see that

$$U_n < C\sqrt{n}\log n \sim \sqrt{\log \tilde{\tau}_n}\, \log\log \tilde{\tau}_n \qquad \textbf{[S46]}$$

with a coefficient independent of n (but depending on Φ and α). We can now solve for the timescale above which mostly filled is likely, $\tau_>(r) = U(\tilde{\tau}(r))\tilde{\tau}(r)$, to find a lower bound, $\ell_<(t)$, on $\ell(t)$,

$$\log(\ell(t)) \geq \log(\ell_<(t)) = \frac{\log(t)}{4\log 2}\left[\log t - 2\log\log t - \mathcal{O}(\log\log\log t)\right], \qquad \textbf{[S47]}$$

which is very close to the upper bound derived above,

$$\log(\ell(t)) \leq \log(\ell_>(t)) = \frac{\log(t)}{4\log 2}\left[\log t - \log\log t - \mathcal{O}(1)\right], \qquad \textbf{[S48]}$$

differing only in the coefficient of the correction term.

One of the advantages of this iterative approach is that the crossover regime can be handled similarly by integrating up **[S45]**. The lower bound will be similar to the upper bound throughout this crossover regime and into the asymptotic regimes for the marginal and long-range cases.

2. Fluctuations and intermediate-range regime. The reason that the fluctuation effects are relatively small for the marginal and long-range regimes is that at each successive timescale, more and more roughly independent long jumps are involved in filling up to the next length scale; i.e., K_n continues to grow. For the long-range regime, it grows so rapidly that almost all of the fluctuations come from early times: This is like what occurs for the fully mixed model. For the marginal case, the fluctuations are dominated by the smallest scales but the cumulative effects of them over the longer scales do make a difference as found in obtaining the lower bounds.

For the intermediate-range case, the ratio of length scales for each factor of 2 in the timescale saturates (when out of the crossover regime) at $K \approx 2^\beta$. This means that whatever fraction, ϕ_n, of the bins are to be filled at each scale, the probability that this occurs either decreases with scale if the product of the ϕ_n does not go to zero or saturates to a constant if the ϕ_n s do also, in which case the overall filling fraction Φ_n tends to zero as a power of time. At each scale there are a comparable number of long jumps that are needed, and thus we should expect that fluctuation effects will be scale invariant and not decrease with scale.

To get useful lower bounds on $\ell(t)$ via an upper bound on $\tau(r)$, $\tau_>(r)$, the easiest way is to require that the source be completely full and that jumps from this completely fill all of the bins at the next scale: This avoids the problems with the ϕ_n. As the probability that all of the bins are filled is (readily) bounded only by $(1 - e^{-W} - e^{-\Lambda_B})^K \approx 1 - 2Ke^{-W}$ if we again chose $\Lambda_B \geq W$, W must be larger by $\log K \approx \beta \log 2$ than for the partially filled analysis above. Carrying through to a similar analysis gives for large β a coefficient $\gamma_n \approx 2\sqrt{\beta}/(n\tilde{\tau}_n)$, which means that the filling probability decays for large times as roughly the inverse of the typical time—natural as the needed long jumps that occur at rate $\sim 1/\tilde{\tau}_n$ have a distribution of when they occur on the same timescale. Note the contrast to the rapid decay of the not mostly filled probability on a timescale of order unity from the partially filled bound derived above. The time beyond which the full filling is likely is bounded only by, in this analysis, $\tilde{\tau}_n U_n \sim \tilde{\tau}_n \sqrt{\beta} n^2$. This gives a lower bound on $\ell(t)$ proportional to $t^\beta/\log^{2\beta} t$. Although on a logarithmic scale the additional factor is smaller, we wish to do better.

The bound can be improved by considering a source that is somewhat smaller—by a factor of 2 is sufficient—than $\tilde{\ell}_n$, which increases the probability that it is filled, but means that more extra time, $T_I - T_S$, is needed to produce a mean number of jumps to the farthest bin of at least W. Using that U and $\gamma\tilde{\tau}$ vary slowly with scale, one can expand $\Lambda(r,t)$ around $r = \tilde{\ell}_n$ and analyze the changes on the bounds at the next scale. This improves the bounds to $\gamma \sim 1/(\sqrt{\beta}\tilde{\tau})$ and $U \sim \beta^{5/2}$ with $\beta = 1/(\mu - 1)$. The resulting lower bound on $\ell(t)$ is

$$\ell(t) > \widetilde{A_\mu}\beta^{-\frac{5\beta}{2}}t^\beta \quad \textbf{[S49]}$$

with the coefficient $\widetilde{A_\mu} \sim 4^{\beta^2}$ that from the deterministic iteration, which is the same, up to an order unity prefactor, as the upper bound. It is not clear where between the lower and upper bounds on the coefficient will be the typical behavior or how broad the fluctuations will be—even on a log scale.

The behavior as $\mu \nearrow 2$ we have not analyzed explicitly, instead focusing on the rapidly growing regime for $\mu \searrow 1$, but the bounds will be of similar form although more care is needed to get upper and lower bounds reasonably close to one another due to the important jumps being only a modest fraction of the size of the already occupied region.

For the marginal case, $\mu = 1$, one can find an upper bound on the time at which the region out to r is likely to be fully filled by similar methods to that for the intermediate-range power-law regime. This yields a bound $\ell_<^{\Phi=1}(t)$ of the same form as that above for partial filling ($\Phi < 1$), except with the coefficient of the $\log\log t$ term in **[S47]** equal to 6 instead of 2. The convergence of the probability of being fully filled is, however, much slower for this bound on compete filling than for the bound on being mostly filled. Whereas the latter converges for $t > \tau_>^\Phi(r) \sim \tilde{\tau}(r)\sqrt{\log\tilde{\tau}(r)}\log\log\tilde{\tau}(r)$ with a rate of order unity—dominated by the small scales—the former converges as the time increases above $\tau_>^{\Phi=1}(r) \sim \tilde{\tau}(r)\log^{5/2}\tilde{\tau}(r)$ on a timescale, $1/\gamma$, of order $\tilde{\tau}(r)\sqrt{\log\tilde{\tau}(r)}$—faster than $\tau_>^{\Phi=1}(r)$ but not much so.

Note that the convergence of the probability for being mostly filled to more than a fixed filling fraction, Φ, is a hybrid property: The probability of a fixed site being filled by time t, $q(r,t)$, is bounded below by (roughly) the product of Φ and the probability that the region out to r is filled to above Φ. To get the convergence of this to unity, Φ needs to be adjusted and the thus-far ignored $\log(1-\Phi)$ factors kept track of. This also necessitates treating intermediate scales differently as the number of bins that do not need to be filled, $(1-f_n)K_n$, is not large. However, a different approach would provide a better bound: Focusing on a specific site being filled with high probability can be done by a method more analogous to the funnel picture in the main text. For the site to be filled, it needs to be in a small-scale bin that is mostly but not necessarily fully filled with high probability, which needs itself to be in a larger bin similarly, etc. However, these can be filled from source regions that are not fully filled: Being partly filled with high enough probability is sufficient. We have not carried out such analysis in detail in part because the actual mechanism by which sites that are empty for an anomalously long time will be filled is more complicated as it will involve filling from nearby regions on a hierarchy of scales that were filled at more typical times.

The analyses here can be immediately extended to give lower bounds on the average density profile at long distances: These will be of the same form as the lower bounds, thus demonstrating that the predicted $\ell(t)^{-(1+\mu)}/r$ is essentially correct.

3. Comparisons with results of Chatterjee and Dey. As noted in the main text, when this work was essentially complete, a preprint by Chatterjee and Dey (CD) appeared, which derives and proves some results closely related to ours in the context of long-range first passage percolation, which is essentially equivalent to the lattice dispersal model, with the jump kernel $G(r) \sim r^{-\alpha}$ equivalent to the $1/r^{-(d+\mu)}$ that we use (2). Although some of the quantities CD focus on are different, the leading asymptotic scaling behaviors they obtain are essentially the same, and their proofs apply in all dimensions. Our $q(r,t)$ corresponds to the probability that the first passage time $T^F(r)$ is less than t and their diameter, $D(t)$—the maximum distance between any pair of occupied points at time t—is, with high probability that decays as a power of ℓ/D—not many times $\ell(t)$, as we both obtain.

CD give heuristic arguments for the extent of the linear regime, the behavior in the power-law regime, and where this breaks down (at $\mu = d$), which are related to ours. Some of their bounds also make use of inequalities related to the simplified form of our self-consistency condition, Eq. **4**.

However, CD's results are suboptimal. In particular, for the coefficient, C, of $\log\ell(t)/\log^2 t$ in the marginal case, they obtain only upper and lower bounds instead of our exact result $C = 1/4d\log 2$. Indeed, in one dimension we obtain rather tight upper and lower bounds on the errors,

$$1 - c_-\frac{\log\log t}{\log t} < \frac{4\log 2\log\ell(t)}{\log^2 t} < 1 - \frac{\log\log t}{\log t} \quad \textbf{[S50]}$$

with high probability—in senses that can be made precise from our analysis—with the coefficient c_- either 2 or 6, depending on the definition of $\ell(t)$ used. In the intermediate-range power-law growth regime, CD's theorems do not appear to exclude $\log t$ prefactors in $\ell(t)$, although their analysis might well do so. However, the main difference is our analysis of the whole crossover regime for μ near d, including the divergences and vanishings of coefficients of the asymptotic forms, which they do not consider. These are crucial for comparisons with simulations because of the very long length scales of the crossovers.

To turn our upper and lower bounds into formal proofs in one dimension requires primarily filling in some details associated with the small-scale regime. For higher dimensions substantial additional work may be needed, although we believe the strategies we developed here should work without major modifications.

1. Mollison D (1972) The rate of spatial propagation of simple epidemics. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, eds Le Cam LM, Neyman J, Scott EL (Univ of California Press, Berkeley, CA), Vol 3, pp 579–614.
2. Chatterjee S, Dey PS (2013) Multiple phase transitions in long-range first-passage percolation on square lattices. *arXiv*:1309.5757.

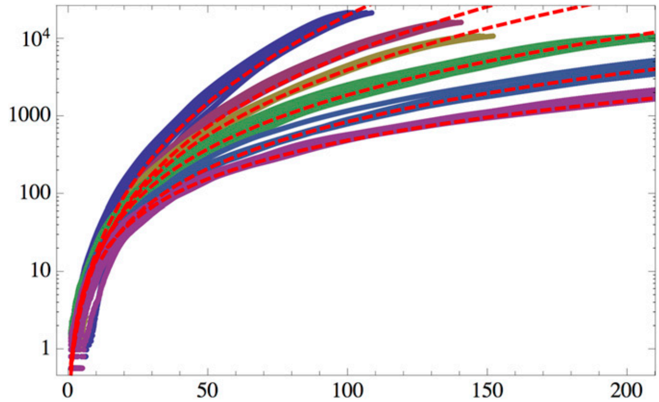

Fig. S1. Summary of the spreading dynamics in two spatial dimensions. The effective radius $\sqrt{M(t)/\pi}$ of the region occupied by the mutant population is plotted as a function of time *t*, for various long-range jump kernels. Each colored cloud represents data obtained from 10 runs for a given jump kernel with tail exponent μ as indicated. Red dashed lines represent predictions, obtained from Eq. **8** (main text) with fitted crossover scales. The jump exponents are $\mu \in \{1.6, 1.8, 1.9, 2.1, 2.3, 2.5\}$.

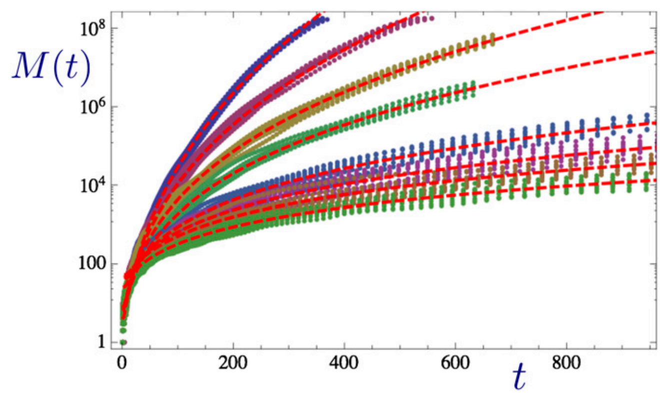

Fig. S2. Summary of the spreading dynamics in one spatial dimension with short-range as well as long-range dispersal. The total number of occupied sites $M(t)$ (the "mass") is shown as a function of time. For these simulations each cluster expands at a linear speed even in the absence of long-range jumps. In each time step, a long-range jump is performed only with probability $\tilde{\epsilon} = 0.1$. For the short-range part, a pair of neighboring sites is chosen at random. If this pair happens to fall on a boundary of a mutant cluster, i.e., the identity of both sites is mixed, then the wild-type site is switched to a mutant site. This leads to expansion of mutant clusters at average speed of $v_0 = 2$ sites per generation. The jump exponents are $\mu \in \{0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4\}$ from the fastest- to the slowest-growing case. The theoretical predictions indicated by the red dashed lines fit the data after choosing appropriate crossover timescales and length scales between linear growth and superlinear growth.

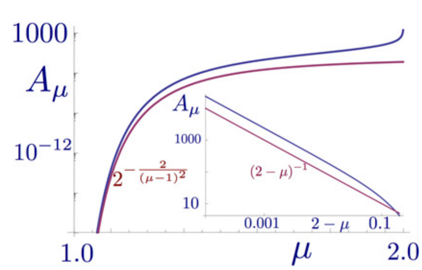

Fig. S3. The prefactor A_μ of the predicted power law growth in Eq. **3** (main text) in one dimension: $\ell(t) \approx A_\mu t^\beta$ with $\beta = 1/(\mu - 1)$ for $1 < \mu < 2$. The blue curve is obtained numerically from solving Eq. **1** (main text) with the power-law ansatz; the red curve represents an analytic approximation derived in *Crossovers and Beyond Asymptopia* (main text). Note the sharp (nonanalytic) drop of the prefactor as μ approaches 1. The reason is very slow crossover to the power law from an intermediate asymptotic regime controlled by the dynamics of the marginal case. As μ approaches 2, the prefactor diverges as $A_\mu \sim (2-\mu)^{-1}$, indicative of another slow crossover at $\mu = 2$ (Fig. S4).

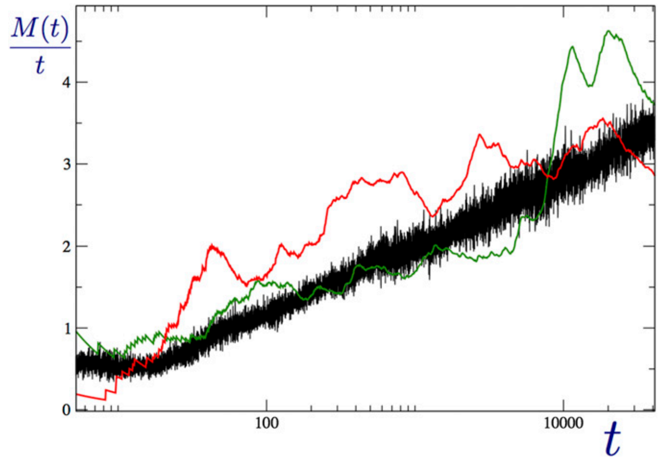

Fig. S4. Dynamics of growth in one dimension at the marginal point between superlinear and linear growth, $\mu = 2$. The number of mutant sites, $M(t)$, scaled by time, $M(t)/t$, is plotted as a function of time, averaged over 10 realizations (black) and for two individual realizations (red and green). Whereas the averaged data suggest $M(t) \sim t\ln(t)$, the individual realizations indicate strong fluctuations caused by occasional rare jumps, which are of order t. These leaps forward (1) are driving the logarithmic increase of the spreading velocity.