Neither pulled nor pushed: Genetic drift and front wandering uncover a new class of reaction-diffusion waves

Gabriel Birzu
Department of Physics, Boston University, Boston, MA 02215, USA
Oskar Hallatschek
Departments of Physics and Integrative Biology,
University of California, Berkeley, California 94720, USA
and
Kirill S. Korolev
Department of Physics and Graduate Program in Bioinformatics,
Boston University, Boston, MA 02215, USA
korolev@bu.edu

August 30, 2017

Abstract

Short Abstract: Traveling waves describe diverse natural phenomena from crystal growth in physics to range expansions in biology. Two classes of waves exist with very different properties: pulled and pushed. Pulled waves are noisy because they are driven by high growth rates at the expansion edge, where the number of organisms is small. In contrast, fluctuations are suppressed in pushed waves because the region of maximal growth is shifted towards the population bulk. Although it is commonly believed that expansions are either pulled or pushed, we found an intermediate class of waves with bulk-driven growth, but exceedingly large fluctuations. Such waves have many unusual properties because their foci of growth, ancestry, and diversity are spatially separated from each other.

Long Abstract: Epidemics, flame propagation, and cardiac rhythms are classic examples of reaction-diffusion waves that describe a switch from one alternative state to another. Only two types of waves are known: pulled, driven by the leading edge, and pushed, driven by the bulk of the wave. Here, we report a distinct class of semi-pushed waves for which both the bulk and the leading edge contribute to the dynamics. These hybrid waves have the kinetics of pushed waves, but exhibit giant fluctuations similar to pulled waves. The transitions between pulled, semi-pushed, and fully-pushed waves occur at universal ratios of the wave velocity to the Fisher velocity (linear spreading velocity). We derive these results in the context of a species invading a new habitat by examining front diffusion, rate of genetic drift, and fluctuation-induced corrections to the expansion velocity. All three quantities decrease as a power law of the population density with the same exponent. For fully-pushed waves, the exponent is -1 consistent with the central limit theorem. In semi-pushed waves, however, the fluctuations average out much more slowly, and the exponent approaches 0 towards the transition to pulled waves. As a result, a rapid loss of genetic diversity and large fluctuations in front position occur even for populations with cooperative growth and other forms of an Allee effect. The outcome of spatial spreading in such populations could therefore be less predictable than previously thought.

Introduction

Wave-like phenomena are ubiquitous in nature and have been extensively studied across many disciplines. In physics, traveling waves describe chemical reactions, kinetics of phase transitions, and fluid flow [1–8]. In biology, traveling waves describe invasions, disease outbreaks, and spatial processes during development and cytoskeleton remodeling [9–22]. Even non-spatial phenomena such as Darwinian evolution and dynamics on networks can be successfully modeled by waves propagating in more abstract spaces such as fitness [22–27].

The wide range of applications stimulated substantial effort to develop a general theory of traveling waves that is now commonly used to understand, predict, and control spreading phenomena [1, 9, 15, 17, 18, 22, 28, 29]. A major achievement of this theory was the division of traveling waves into two classes with very different properties [1, 6, 9, 28, 30–35]. The first class contains waves that are "pulled" forward by the dynamics at the leading edge. Kinetics of pulled waves are independent from the nonlinearities behind the front, but extremely sensitive to noise and external perturbations [1, 31, 36]. In contrast, the waves in the second class are resilient to fluctuations and are "pushed" forward by the nonlinear dynamics throughout the wave front. Although it is commonly assumed that all traveling waves are either pulled or pushed, the transition between the two classes has not been systematically investigated [1, 27, 31–33, 36–38].

Here, we show that the simple pulled *vs.* pushed classification does not hold. Instead, there are three distinct classes of traveling waves. The class of a wave is completely determined by the dimensionless ratio of the expansion velocity to the Fisher velocity, which is the spreading rate obtained by linearizing the reaction-diffusion equation for small population densities. High and low ratios correspond to the standard classes of pulled and pushed waves respectively, but the intermediate ratios belong to a new class with very unusual properties. These intermediate waves have the kinetics of pushed waves because they are driven by growth behind the front, yet they show extreme sensitivity to fluctuations similar to pulled waves. We term such waves semi-pushed and show that their hybrid behavior is driven by the spatial separation between the foci of growth, ancestry, and diversity within the wave front. The spreading and evolutionary dynamics in semi-pushed waves is therefore likely to be quite different from those in pulled and pushed waves and may require new theoretical developments [27].

Model

Traveling waves occur when a transport mechanism couples dynamics at different spatial locations. The nature of these wave-generating processes could be very different and ranges from reactions and diffusion in chemistry to growth and dispersal in ecology. The simplest and most widely-used model of a reaction-diffusion wave is the generalized Fisher-Kolmogorov equation:

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + r(n)n + \sqrt{\gamma_n(n)n}\eta(t,x), \tag{1}$$

which, in the context of ecology, describes how a species colonizes a new habitat [1, 9, 39–41]. Here, $n(t,x)$ is the population density of the species, D is the dispersal rate, and $r(n)$ is the density-dependent per capita growth rate. The last term accounts for demographic fluctuations: $\eta(t,x)$ is a Gaussian white noise, and $\gamma_n(n)$ quantifies the strength of the fluctuations. In simple

birth-death models, γ_n is a constant, but we allow for an arbitrary dependence on n provided that $\gamma_n(0) > 0$. The origin of the noise term and its effects on the wave dynamics are further discussed in the SI.

Pulled waves occur when $r(n)$ is maximal at small n; for example, when the growth is logistic: $r(n) = r_0(1 - n/N)$ [1, 9]. Here, r_0 is the growth rate at low densities, and N is the carrying capacity that sets the population density behind the front. For pulled waves, the expansion dynamics are controlled by the very tip of the front, where the organisms not only grow at the fastest rate, but also have an unhindered access to the uncolonized territories. As a result, the expansion velocity is independent of the functional form of $r(n)$ and is given by the celebrated result due to Fisher, Kolmogorov, and Skellam [39–41]:

$$v_{\text{F}} = 2\sqrt{Dr(0)}. \tag{2}$$

Pushed waves occur when a species grows best at intermediate population densities [1, 9]. Such non-monotonic behavior of $r(n)$ arises through a diverse set of mechanisms and is known as an Allee effect in ecology [42, 43]. Most common causes of an Allee effect are cooperative feeding, collective defense against predators, and the difficulty in finding mates at low population densities [44–47]. The velocity of pushed waves is always greater than Fisher's prediction ($v > v_{\text{F}}$) and depends on all aspects of the functional form of $r(n)$ [1, 9].

Allee effects are typically described by adding a cooperative term to the logistic equation:

$$r(n) = r_0 \left(1 - \frac{n}{N}\right) \left(1 + B\frac{n}{N}\right), \tag{3}$$

where B is the strength of cooperativity. For this model, the exact solutions are known for the expansion velocity and the population density profile; see SI and Ref. [9, 48, 49]. For $B \leq 2$, expansions are pulled, and the expansion velocity equals v_{F}, which is independent of B. That is cooperativity does not always increase the expansion velocity even though it always increases the growth rates at high densities. For $B > 2$, expansions are pushed, and v increases with B. Figure 1A illustrates this transition from pulled to pushed waves as cooperativity is increased. In Methods and SI, we also present several alternative models of an Allee effect and show that our conclusions do not depend on a particular choice of $r(n)$.

Increasing the value of cooperativity beyond $B = 2$ not only makes the expansion faster, but also shifts the region of high growth from the tip to the interior of the expansion front (Fig. 1BC). This shift is the most fundamental difference between pulled and pushed waves because it indicates the transition from a wave being "pulled" by its leading edge to a wave being "pushed" by its bulk growth.

The transition from "pulled" to "pushed" dynamics is also evident in the number of organisms that trace their ancestry to the leading edge *vs.* the bulk of the front. The expected number of descendants has been determined for any spatial position along the front for both pulled and pushed waves [37, 38, 50, 51]. For pulled waves, only the very tip of the expansion contributes to future generations. On the contrary, the organisms at the leading edge leave few progeny in pushed waves,

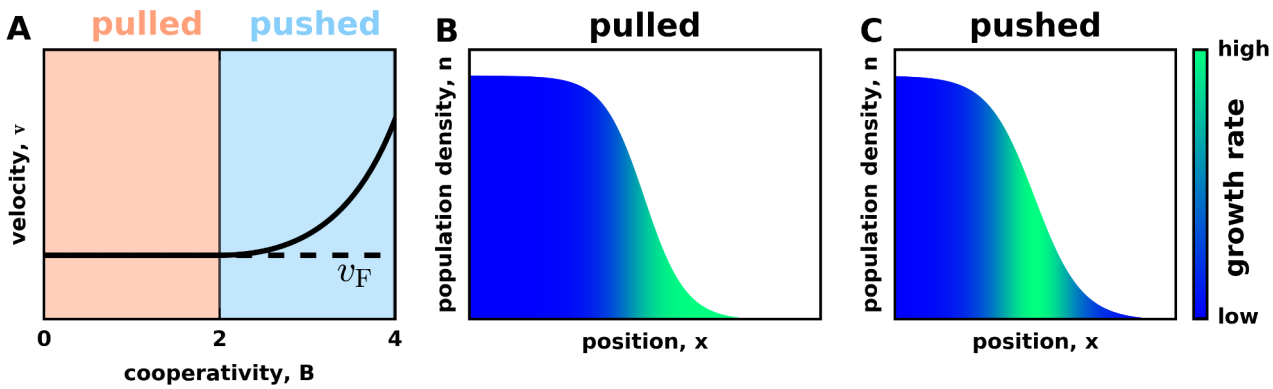

Figure 1. Waves transition from pulled to pushed as growth becomes more cooperative. **(A)** shows the expansion velocity as a function of cooperativity for the growth rate specified by Eq. (3). For low cooperativity, expansions are pulled and their velocity equals the Fisher velocity. Beyond the critical value of $B = 2$, expansions become pushed and their velocity exceeds v_F. Note that the region of high growth is at the leading edge of the front in pulled waves **(B)**, but in the interior of the front in pushed waves **(C)**. This difference is due to the dependence of the growth rate on the population density. For low cooperativity, the growth rate is maximal at low population densities, but, for high cooperativity, the growth rate is maximal at intermediate population densities. In all panels, the exact solution of Eq. (1) is plotted; $B = 0$ in panel B, and $B = 4$ in panel C.

and the population descends primarily from the organisms in the region of high growth. This shift in the spatial focus of ancestry has a profound effect on species evolution. In pulled waves, only mutations near the very edge of the expansion have an appreciable fixation probability, but the entire expansion front contributes to evolution in pushed waves (Fig. 2).

The edge-dominated dynamics make pulled waves extremely sensitive to the vagaries of reproduction, death, and dispersal [1, 31, 35, 36]. Indeed, the number of organisms at the leading edge is always small, so strong number fluctuations are expected even in populations with a large carrying capacity, N. These fluctuation affect both ecological properties, such as the shape and position of the wave front, and evolutionary properties, such as the genetic diversity of the expanding population. Consistent with these expectations, experiments with pulled waves reported an unusual roughness of the expansion front [6] and a rapid loss of genetic diversity [52, 53].

While the difference between pulled and pushed waves is well-established, little is known about the transition between the two types of behavior. In particular, it is not clear how increasing the nonlinearity of $r(n)$ transforms the patterns of fluctuations and other properties of a traveling wave. To answer this question, we developed a computer-based model of range expansions that can simulate both ecological and evolutionary processes. Specifically, our simulations described the dynamics of both the population density and the relative abundance of two neutral genotypes. The former was used to estimate the fluctuations in the position of the front, and the latter was used to quantify the decay rate of genetic diversity. In simulations, the species expanded in a one-dimensional array of habitable patches connected by dispersal between the nearest neighbors. Each time step consisted of a deterministic dispersal and growth followed by random sampling to simulate demographic fluctuations and genetic drift (see Methods and SI). By increasing the cooperativity

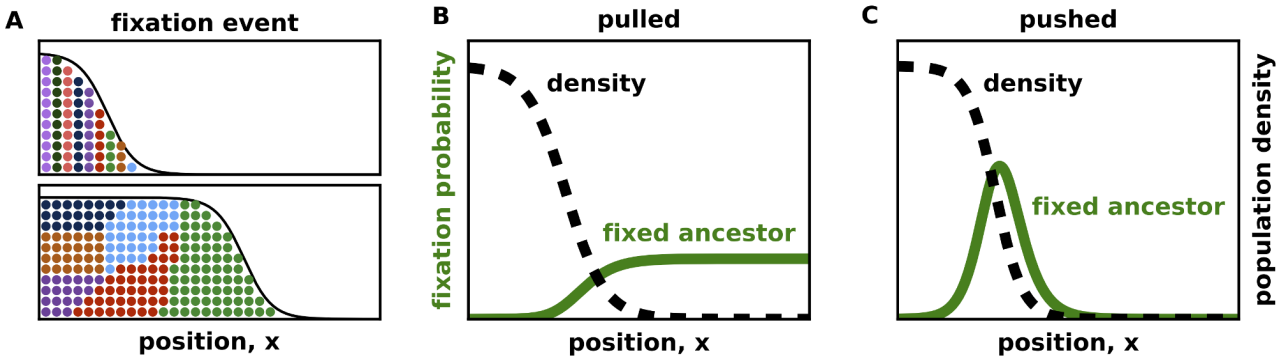

Figure 2. Ancestral lineages occupy distinct locations in pulled and pushed waves. (A) illustrates the fixation of a particular genotype. Initially, a unique and heritable color was assigned to every organism to visualize its ancestral lineage. There are no fitness differences in the population, so fixations are caused by genetic drift. **(B)** and **(C)** show the probability that the fixed genotype was initially present at a specific position in the reference frame comoving with the expansion. The transition from pulled to pushed waves is marked by a shift in the fixation probability from the tip to the interior of the expansion front. This shift indicates that most ancestral lineages are focused at the leading edge in pulled waves, but near the middle of the front in pushed waves. The fixation probabilities were computed analytically, following Refs. [37, 38], as described in the SI. We used $B = 0$ in panel B and $B = 4$ in panel C.

of the growth rate, we observed a clear transition from pulled ($v = v_{\rm F}$) to pushed ($v > v_{\rm F}$) waves accompanied by a dramatic reduction in fluctuations. This transition is illustrated in Fig. 3, where a small increase in cooperativity results in more than an order of magnitude drop in the rates of front wandering and genetic drift.

Results

Fluctuations provide an easy readout of the internal dynamics in a traveling wave, so we decided to determine how they change as a function of cooperativity. Since the magnitude of the fluctuations also depends on the population density, we looked for a qualitative change in this dependence while varying B. In particular, we aimed to determine whether population dynamics change discontinuously at the transition between pulled and pushed waves.

Spatial wandering of the front

We first examined the fluctuations of the front position in the comoving reference frame. The position of the front $X_{\rm f}$ was defined as the total population size in the colonized space normalized by the carrying capacity $X_{\rm f} = \frac{1}{N}\int_0^{+\infty} n(t,x)dx$. As expected [1, 35, 36, 54], $X_{\rm f}$ performed a random walk due to demographic fluctuations in addition to the average motion with a constant velocity (Fig. 3AB). For both pulled and pushed waves, the variance of $X_{\rm f}$ grew linearly in time (Fig. 4A), i.e. the front wandering was diffusive and could be quantified by an effective diffusion constant $D_{\rm f}$.

The magnitude of the front wandering is expected to depend strongly on the type of the expansion [1, 35, 36, 54]. For pulled waves, Ref. [35] found that $D_{\rm f} \sim \ln^{-3} N$, but a very different scaling $D_{\rm f} \sim$

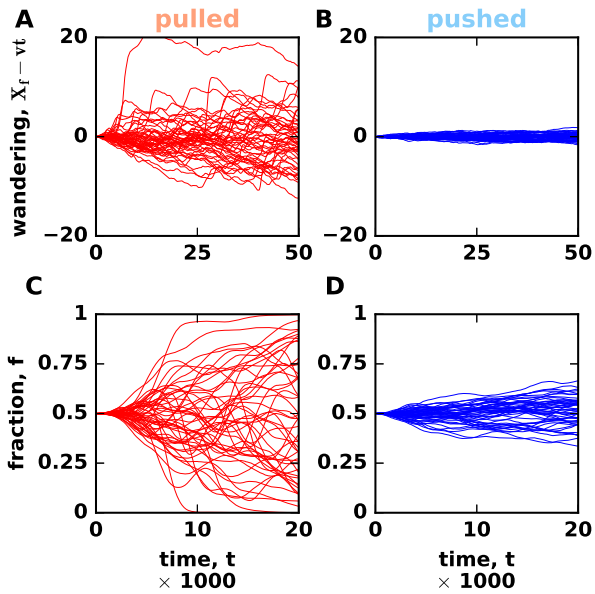

Figure 3. Fluctuations are much stronger in pulled than in pushed waves. The top row compares front wandering between pulled **(A)** and pushed **(B)** expansions. Each line shows the position of the front $X_f(t)$ in a single simulation relative to the mean over all simulations in the plot. The bottom row compares the strength of genetic drift between pulled **(C)** and pushed **(D)** expansions. We started the simulations with two neutral genotypes equally distributed throughout the front and then tracked how the fraction of one of the genotypes changes with time. This fraction was computed from 300 patches centered on X_f to exclude the fluctuations well behind the expansion front.

N^{-1} was predicted for certain pushed waves [54]; see Fig. 4B. Given that pulled and pushed waves belong to distinct universality classes, it is easy to assume that the transition between the two scaling regimens should be discontinuous [1, 35–38, 54]. This assumption, however, has not been carefully investigated, and we hypothesized that there could be an intermediate regime with $D_f \sim N^{\alpha_D}$. From simulations, we computed how α_D changes with B and indeed found that pushed waves have intermediate values of α_D between 0 and -1 when $B \in (2, 4)$ (Fig. 8).

The dependence of the scaling exponent on the value of cooperativity is shown in Fig. 4C. For large B, we found that α_D is constant and equal to -1, which is consistent with the previous work [54]. Below a critical value of cooperativity, however, the exponent α_D continually changes with B towards 0. The critical cooperativity is much larger than the transition point between pulled and pushed waves, so the change in the scaling occurs within the class of pushed waves. This transition divides pushed waves into two subclasses, which we termed fully-pushed and semi-pushed waves. For pulled waves, we found that α_D is independent of B, but our estimate of α_D deviated slightly from the expected value due to the finite range of N in the simulations (compare N^{α_D} and $\ln^{-3} N$ fits in Fig. 4B).

Loss of genetic diversity

Our analysis of the front wandering showed that pushed waves consist of two classes with a very

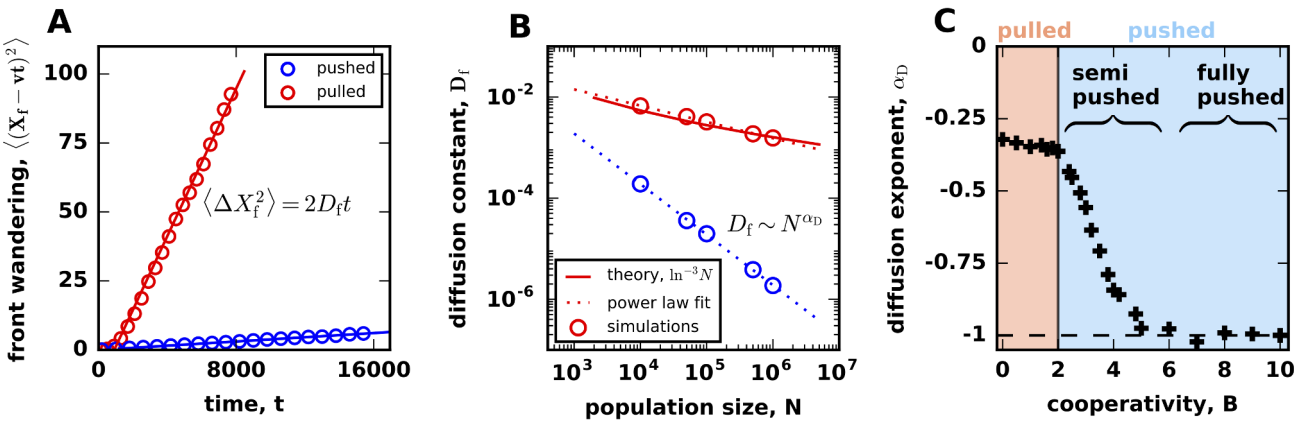

Figure 4. Front wandering identifies a new class of pushed waves. (A) Fluctuations in the front position can be described by simple diffusion for both pulled and pushed waves. **(B)** The front diffusion is caused by the number fluctuations, so the effective diffusion constant, D_f, decreases with the carrying capacity, N. For pulled waves, $D_f \sim \ln^{-3} N$ [35], while, for pushed waves, D_f can decrease much faster as N^{-1} [54]. We quantify the scaling of D_f with N by the exponent α_D equal to the slope on the log-log plot shown. The overlap of the two red lines highlight the fact that, even though α_D should equal 0 for pulled waves, the limited range of N results in a different value of $\alpha_D \approx -0.33$. **(C)** The dependence of the scaling exponent on cooperativity identifies two distinct classes of pushed waves.

different response to demographic fluctuations. To determine whether this difference extends to other properties of expansions, we turned to genetic drift, a completely different process that describes fluctuations in the genetic composition of the front. Genetic drift occurs even in the absence of front wandering (see SI and Ref. [55]), so these two properties are largely independent from each other and capture complementary aspects related to physical and evolutionary dynamics in traveling waves.

We quantified genetic fluctuations by the rate at which genetic diversity is lost during an expansion. The simulations were started in a diverse state with each habitable patch containing an equal number of two neutral genotypes. As the expansion proceeded, the relative fractions of the genotypes fluctuated and eventually one of them was lost from the expansion front (Fig. 3C). To capture the loss of diversity, we computed the average heterozygosity H, defined as the probability to sample two different genotypes at the front. Mathematically, H equals the average of $2f(1-f)$, where f is the fraction of one of the genotypes in an array of patches comoving with the front, and the averaging is done over independent realizations. Consistent with the previous work [35, 37], we found that the heterozygosity decays exponentially at long times: $H \sim e^{-\Lambda t}$ for both pulled and pushed waves (Fig. 5A). Therefore, the rate Λ was used to measure the strength of genetic drift across all values of cooperativity.

By analogy with the front wandering, we reasoned that Λ would scales as N^{α_H} for large N, and α_H would serve as an effective "order parameter" that distinguishes different classes of traveling waves. Indeed, Ref. [35] showed that $\Lambda \sim \ln^{-3} N$ for pulled waves, i.e. the expected α_H is zero. Although no conclusive results have been reported for pushed waves, the work on adaptation waves in fitness space suggests $\alpha_H = -1$ for fully-pushed waves [27]. Our simulations confirmed both of these

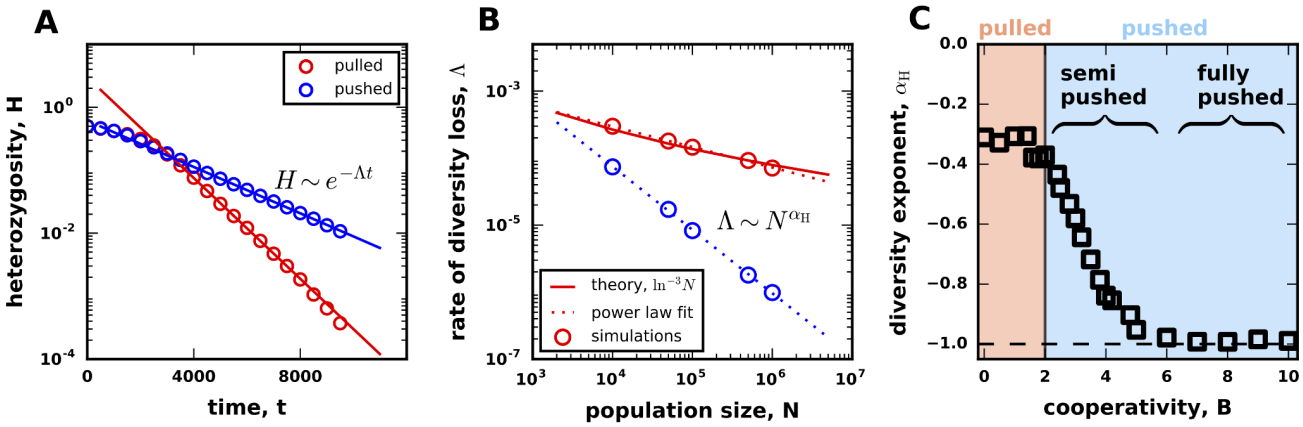

Figure 5. Genetic diversity is lost at different rates in pulled, semi-pushed waves, and fully-pushed waves. (A) The average heterozygosity, H, is a measure of diversity equal to the probability to sample two distinct genotypes in the population. For both pulled and pushed expansions, the decay of genetic diversity is exponential in time: $H \sim e^{-\Lambda t}$, so we used Λ to measure the strength of genetic drift. **(B)** Genetic drift decrease with N. For pulled waves, $\Lambda \sim \ln^{-3} N$ [35], while, for fully-pushed waves, we predict that $\Lambda \sim N^{-1}$; see Eq. (4). To quantify the dependence of Λ on N, we fit $\Lambda \sim N^{\alpha_H}$. The dashed red line shows that even though α_H should equal 0 for pulled waves, the limited range of N results in a different value of $\alpha_H \approx -0.33$. **(C)** The dependence of the scaling exponent on cooperativity identifies the same three classes of waves as in Fig. 4C; the transitions between the classes occur at the same values of B.

predictions (Fig. 5B) and showed $\Lambda \sim N^{\alpha_H}$ scaling for all values of cooperativity.

The dependence of α_H on B shows that genetic fluctuations follow exactly the same pattern as the front wandering (Fig. 5C). In particular, both exponents undergo a simultaneous transition from $\alpha_H = \alpha_D = -1$ to a continual dependence on B as cooperativity is decreased. Thus, genetic fluctuations also become large as waves switch from fully-pushed to semi-pushed. In the region of pulled waves, α_D and α_H are independent of B, but their values deviate from the theoretical expectation due to the finite range of N explored in the simulations[1]. Overall, the consistent behavior of the fluctuations in the position and composition of the front strongly suggests the existence of two classes of pushed waves, each with a distinct set of properties.

The origin of semi-pushed waves

We next sought an analytical argument that can explain the origin of the giant fluctuations in semi-pushed waves. In the SI, we explain and extend the approaches from Refs. [37] and [54] to compute D_f and Λ using a perturbation expansion in $1/N$. The main results are

[1]This difficulty in distinguishing between $\ln^{-3} N$ and $N^{-0.34}$ scaling from simulations alone has led some authors to suggest a power-law dependence of Λ on N for pulled waves [37] contrary to the theoretical prediction [35].

$$
\begin{aligned}
D_{\mathrm{f}} &= \frac{1}{N}\frac{\int_{-\infty}^{+\infty}\gamma_n(\rho)[\rho'(\zeta)]^2\rho(\zeta)e^{\frac{2v\zeta}{D}}d\zeta}{2\left(\int_{-\infty}^{+\infty}[\rho'(\zeta)]^2 e^{\frac{v\zeta}{D}}d\zeta\right)^2}, \\
\Lambda &= \frac{1}{N}\frac{\int_{-\infty}^{+\infty}\gamma_f(\rho)\rho^3(\zeta)e^{\frac{2v\zeta}{D}}d\zeta}{\left(\int_{-\infty}^{+\infty}\rho^2(\zeta)e^{\frac{v\zeta}{D}}d\zeta\right)^2}.
\end{aligned}
\tag{4}
$$

Here, primes denote derivatives; $\zeta = x - vt$ is the coordinate in the reference frame comoving with the expansion; $\rho(\zeta) = n(\zeta)/N$ is the normalized population density profile in the steady state; v is the expansion velocity; and D is the dispersal rate as in Eq. (1). For the sake of generality, we also introduced the strength of the genetic drift γ_f, which could be different from the strength of demographic fluctuations γ_n (see SI); note that $\gamma_f(0) > 0$.

The N^{-1} scaling that we observed for fully-pushed waves is readily apparent from Eqs. (4). Moreover, we calculated the integrals analytically for the model specified by Eq. (3); see SI. These exact results show excellent agreement with the simulations (Fig. 9) and thus confirm the validity of the perturbation approach for fully-pushed waves.

Why does the N^{-1} scaling break down in semi-pushed waves? We found that the integrals in the numerators in Eqs. (4) become more and more dominated by large ζ as cooperativity decreases, and, at a critical value of B, they diverge. To pinpoint this transition, we determined the behavior of $\rho(\zeta)$ for large ζ by linearizing Eq. (1) for small population densities:

$$
D\frac{d^2\rho}{d\zeta^2} + v\frac{d\rho}{d\zeta} + r(0)\rho = 0, \tag{5}
$$

where we replaced n by ρ and shifted into the reference frame comoving with the front. Equation (5) is linear, so the population density decreases exponentially at the front as $\rho \sim e^{-k\zeta}$. The value of k is obtained by substituting this exponential form into Eq. (5) and is given by $k = \frac{v}{2D}\left(1 + \sqrt{1 - v_{\mathrm{F}}^2/v^2}\right)$ with v_{F} as in Eq. (2) (see SI). From the asymptotic behavior of ρ, it is clear that the integrands in the numerators in Eqs. (4) scale as $e^{(2v/D-3k)\zeta}$, and the integrals diverge when $v/D = 3k/2$. The integrals in the denominators converge for all pushed waves.

The divergence condition can be stated more clearly by expressing k in terms of v and then solving for the critical velocity v_{critical}. From this calculation, we found that the transition from fully-pushed to semi-pushed waves occurs at a universal ratio of the expansion velocity v to the linear spreading velocity v_{F}:

$$
v_{\mathrm{critical}} = \frac{3}{2\sqrt{2}}v_{\mathrm{F}}. \tag{6}
$$

This result does not rely on Eq. (3) and holds for any model of cooperative growth.

The ratio v/v_{F} increases with cooperativity and serves as a model-independent metric of the extent

to which a wave is pushed. Equation (6) and the results below further show that this metric is universal, i.e. different models with the same $v/v_{\rm F}$ have the same patterns of fluctuations. We can then classify all reaction-diffusion waves using this metric. Pulled waves correspond to the special point of $v/v_{\rm F} = 1$. When $1 < v/v_{\rm F} < \frac{3}{2\sqrt{2}}$, waves are semi-pushed, and fully-pushed waves occur when $v/v_{\rm F} \geq \frac{3}{2\sqrt{2}}$. Fully-pushed waves also occur when $r(0) < 0$; see SI. Such situations are called propagation into metastable state in physics [1] and strong Allee effect in ecology [44]. Because the growth rate at the front is negative, $v_{\rm F}$ does not exist, and the expansion proceeds only due to the growth in the bulk, where the fluctuations are small.

Although the perturbation theory breaks down for $v < v_{\rm critical}$, we can nevertheless estimate the scaling exponents $\alpha_{\rm D}$ and $\alpha_{\rm H}$ by imposing an appropriate cutoff in the integrals in Eqs. (4). One reasonable choice of the cutoff is $\rho_c \sim 1/N$, which ensures that there is no growth in patches that have fewer than one organism. In the SI, we show that this cutoff is appropriate for deterministic fronts with $\gamma_n = 0$, but a different cutoff is needed for fluctuating fronts with $\gamma_n > 0$. The need for a different cutoff had been recognized for a long time both from simulations [33] and theoretical considerations [35]. However, a method to compute the cutoff has been developed only recently. For pulled waves, the correct value of the cutoff was obtained in Ref. [56] using a nonstandard moment-closure approximation for Eq. (1). We extended this method to pushed waves and found that the integrals should be cut off for ζ greater than $\zeta_c \sim \ln(N)/(v/D - k)$; see SI.

Upon applying this cutoff to Eqs. (4), we find that the fluctuations in semi-pushed waves have a power-law dependence on N with a nontrivial exponent between 0 and -1. The exponent is the same for both Λ and $D_{\rm f}$ and depends only on $v/v_{\rm F}$. Our theoretical results can then be summarized as follows

$$\alpha_{\rm D} = \alpha_{\rm H} = \begin{cases} 0, & v/v_{\rm F} = 1, \\ -2\dfrac{\sqrt{1 - v_{\rm F}^2/v^2}}{1 - \sqrt{1 - v_{\rm F}^2/v^2}}, & v/v_{\rm F} \in (1, \dfrac{3}{2\sqrt{2}}), \\ -1, & v/v_{\rm F} \geq \dfrac{3}{2\sqrt{2}}. \end{cases} \tag{7}$$

In the case of pulled waves, our cutoff-based calculation not only predicts the correct values of $\alpha_{\rm D} = \alpha_{\rm H} = 0$, but also reproduces the expected $\ln^{-3} N$ scaling (SI).

To test the validity of the cutoff approach, we compared its predictions to the simulations of Eq. (3) and two other models of cooperative growth; see Fig. 6, Methods, and SI. The simulations confirm that the values of $\alpha_{\rm D}$ and $\alpha_{\rm H}$ are equal to each other and depend only on $v/v_{\rm F}$. Moreover, there is an excellent quantitative agreement between the theory and the data, given the errors in $\alpha_{\rm D}$ and $\alpha_{\rm H}$ due to the finite range of N in our simulations.

The success of the cutoff-based calculation leads to the following conclusion about the dynamics in semi-pushed waves: The fluctuations are controlled only by the very tip of the front while the growth and ancestry are controlled by the front bulk (see Figs. 1C and 2C). Thus, the counter-intuitive behavior of semi-pushed waves originates from the spatial segregation of different processes within a wave front. This segregation is not present in either pulled or fully-pushed waves and signifies a new state of the internal dynamics in a traveling wave.

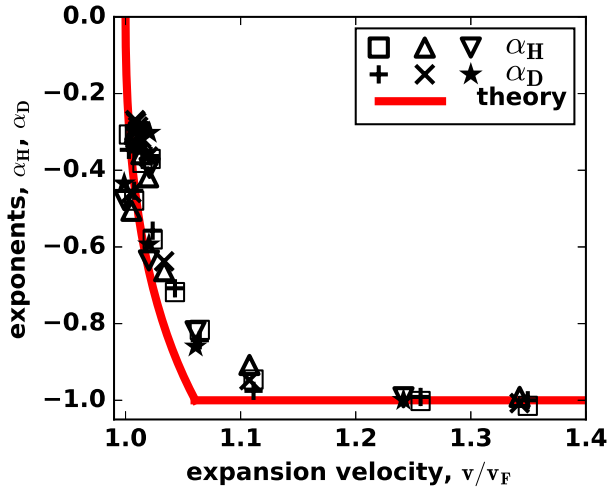

Figure 6. The universal transition from semi-pushed to fully-pushed waves. For three different models on an Allee effect, the scaling exponents for the heterozygosity and front diffusion collapse on the same curve when plotted as a function of v/v_F. Thus, v/v_F serves as a universal metric that quantifies the effects of cooperativity and separates semi-pushed from fully-pushed waves. We used used squares and plus signs for the model specified by Eq. (3), triangles and crosses for the model specified by Eq. (8), inverted triangles and stars for the model specified by Eq. (9), and the red line for the theoretical prediction from Eq. (4).

Corrections to the expansion velocity due to demographic fluctuations

Finally, we examined how the expansion velocity depends on the strength of demographic fluctuations. For waves traveling in fitness space, deterministic limit results in an infinite adaptation rate, so the value of v is largely determined by demographic fluctuations [23, 24, 56]. For range expansions, the dependence of v on N is weaker, and demographic fluctuations typically result only in a correction to the deterministic prediction for the wave velocity v_d. Historically, corrections to v_d have been used to test the theories of fluctuating fronts [1, 33, 36], so we computed $v - v_d$ using a perturbation theory in $1/N$ (see SI). This calculation showed that $v - v_d \sim N^{\alpha_V}$ with α_V equal to $\alpha_D = \alpha_H$. Thus, we predicted $1/N$ scaling for fully-pushed waves and a weaker power-law dependence for semi-pushed waves with the exponent given by Eq. (7); the scaling for pulled waves is logarithmic in the carrying capacity[2]. Our simulations showed good agreement with these theoretical predictions (Fig. 11) and, therefore, provided further support for the existence of three distinct classes of traveling waves.

Discussion

Spatially extended systems often change through a wave-like process. In reaction-diffusion systems, two types of waves have been known for a long time: pulled and pushed. Pulled waves are driven by the dynamics at the leading edge, and all their properties can be obtained by linearizing the equations of motion. In contrast, the kinetics of pushed waves are determined by nonlinear reaction processes. The distinction between pulled and pushed waves has been further supported by the recent work on the evolutionary dynamics during range expansions [37, 38]. In pulled waves,

[2]Note that, for pulled waves, $v - v_d \sim \ln^{-2} N$, which is slightly different from the $\ln^{-3} N$ scaling of D_f and Λ [35, 56].

mutations spread only if they occur at the expansion edge, but the entire front contributes to adaptation in pushed waves.

A natural conclusion from the previous work is that all aspects of the wave behavior are determined by whether the wave is pulled or pushed. Here, we challenged this view by reporting how fluctuation patterns change as the growth of a species becomes more nonlinear. Our main finding is that both front wandering (a physical property) and genetic drift (an evolutionary property) show identical behavior with increasing nonlinearity and undergo two phase transitions. The first phase transition is the classic transition between pulled and pushed waves. The second phase transition is novel and separates pushed waves into two distinct subclasses, which we termed fully-pushed and semi-pushed waves.

In fully-pushed waves, fluctuations obey the central limit theorem and decrease with the carrying capacity as N^{-1}. This simple behavior arises because all processes are localized in a region behind the front. The number of organisms in this region grows linearly with N, so the variance of the fluctuations scales as N^{-1}. The central limit theorem seems not to apply to semi-pushed waves, for which we observed a nontrivial power-law scaling with variable exponents. The new scaling reflects the balance between the large fluctuations at the leading edge and the localization of the growth and ancestry processes behind the front. The departure from the N^{-1} scaling is the strongest in pulled waves, where all processes localize at the tip of the front. Since the number of organisms at the leading edge is always close to 1, the fluctuations are very large and decrease only logarithmically with the population size.

The unusual behavior of semi-pushed waves arises due to the spatial segregation between the foci of growth, ancestry, and diversity (Fig. 7). At the transition from pulled to semi-pushed waves, the locations of maximal growth and fixation probabilities shift from the edge to the bulk of the front. However, there is no corresponding shift in the focus of diversity defined as the location where ancestral lineage are most likely to coalesce (i.e. the point where two randomly sampled individuals are expected to have their most recent common ancestor). This focus, which determines the strength of genetic drift and other fluctuations, moves from the edge to the bulk of the front at a separate phase transition separating semi-pushed from fully-pushed waves. Thus, in pulled an fully-pushed waves, all foci are colocalized, but they spatially separate in semi-pushed waves. As a result, semi-pushed waves possess characteristics of both pulled and pushed expansions. The segregation of processes within the wave front is relatively unexplored and could underlie other transitions in expansion dynamics.

The analysis of fluctuations also revealed surprising universality in pushed waves. Because pushed waves are nonlinear, their velocity and front shape depend on all aspects of the growth rate, and it is natural to assume that there are as many types of pushed waves as there are nonlinear growth functions. Contrary to this expectation, we showed that many consequences of nonlinearities can be captured by a single dimensionless parameter v/v_{F}. This ratio was first used to distinguish pulled and pushed waves, but we found that v/v_{F} also determines the transition from semi-pushed to fully-pushed waves and the magnitude of the fluctuations. We therefore suggest that v/v_{F} could be a very useful and possibly universal metric of the extent to which an expansion is pushed. Such a metric is needed to compare dynamics across different ecosystems and could also play an important role in connecting the theory to experiments and field studies.

The important role of v/v_{F} is further supported by the recent analysis of genealogies in populations

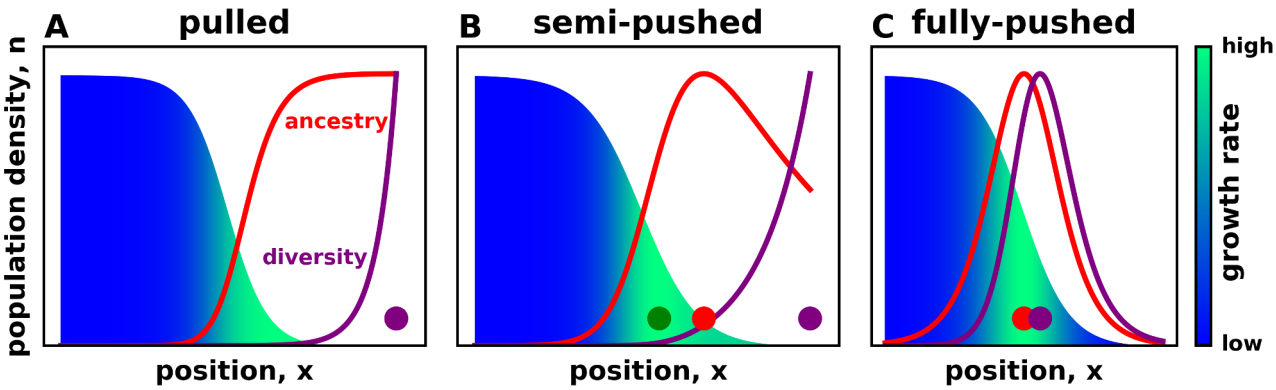

Figure 7. Foci of growth, ancestry, and diversity spatially segregate in semi-pushed waves. The three panels compare the spatial distribution of growth, ancestry, and diversity among pulled, semi-pushed, and fully-pushed waves. The color gradient shows how the per-capita growth rate changes along the wave front. The ancestry curve shows the distribution of the most recent common ancestor of the entire population at the front, which is the same as the probability of a neutral mutation arising at a particular location and then reaching fixation. The diversity curve shows the spatial distribution of the most recent common ancestor of two individuals sampled randomly from the front. In other words, this curve shows the probability that two ancestral lineages coalesce at a specific location at the front. Thus, the maximum of the diversity curve corresponds to the location that contributes most to the rate of diversity loss. The color gradient and the curves are theoretical predictions from Eqs. (24), (153), and (154). The colored dots show the positions of foci of growth (green), ancestry (red), and diversity (purple). In both pulled and pushed waves these foci are colocalized, but they are spatially separated in semi-pushed waves. As a result of this, semi-pushed waves posses characteristics of both pulled and pushed waves.

whose evolution can be described by traveling waves in fitness space [27]. In this context, pulled waves describe populations that adapt rapidly with many beneficial mutations present at the same time, while pushed waves are similar to populations that evolve slowly, e.g. when all mutations are neutral. These two extremes of evolutionary dynamics result in very different genealogies. Genealogies under neutral evolution are described by the Kingman coalescent [57–59], in which ancestral lineages merge at a constant rate and only in pairs. Genealogies in rapidly adapting populations are described by the Bolthausen-Sznitman coalescent, for which multiple mergers are common [26, 60, 61]. Brunet and Derrida [27] observed a gradual transition between these two coalescents by comparing populations that by chance adapted at different rates. This observation suggests the genealogies in expanding populations could undergo a similar transition: from Bolthausen-Sznitman coalescent in pulled waves, to generic Beta coalescent in semi-pushed waves, and to Kingman coalescent in fully-pushed waves.

Despite the critical role that evolution plays in biological invasions [10, 19, 29, 62–71], only a handful of studies examined the link between genetic diversity and species ecology in this context [37, 38, 51, 55]. The main result of the previous work is that Allee effects reduce genetic drift and preserve diversity. This conclusion, however, was reached without systematically varying the strength on the Allee effect in simulations and was often motivated by the behavior of the fixation probabilities rather than the diversity itself. Our findings not only provide firm analytical and numerical support for the previous results, but also demonstrate that the simple picture of reduced fluctuations in

pushed waves does not accurately reflect the entire complexity of the eco-evolutionary feedback in traveling waves. In particular, we showed that the strength and location of genetic drift varies greatly between semi-pushed and fully-pushed waves. As a result, even a large Allee effect that makes the expansion pushed could be insufficient to substantially slow down the rate of diversity loss.

Methods

The simulations in Figs. 3-5 were carried out for the growth model defined by Eq. (3). In Fig. 6 and certain supplemental figures, we also used two other growth models to demonstrate that our results do not depend on the choice of $r(n)$. These growth models are specified by the following equations:

$$r(n) = g_0 \left(1 - \frac{n}{N}\right) \left(\frac{n}{N} - \frac{n^*}{N}\right), \tag{8}$$

and

$$r(n) = g_0 \left[1 - \left(\frac{n}{N}\right)^3\right] \left[\left(\frac{n}{N}\right)^3 - \left(\frac{n^*}{N}\right)^3\right], \tag{9}$$

where $N > 0$ is the carrying capacity, $g_0 > 0$ sets the time scale of growth, and c^* is the Allee threshold, which could assume both positive and negative values; see SI.

We simulated range expansions of two neutral genotypes in a one-dimensional habitat modeled by an array of patches separated by distance a; the time was discretized in steps of duration τ. Thus, the abundance of each genotype was represented as $n_i(t, x)$, where $i \in \{1, 2\}$ is the index of the genotype, and t and x are integer multiples of τ and a. Each time step, we updated the abundance of both genotypes simultaneously by drawing from a multinomial distribution with N trials and probability p_i to sample genotype i. The values of p_i reflected the expected abundances of the genotypes following dispersal and growth:

$$p_i = \frac{\frac{m}{2} n_i(t, x-a) + (1-m) n_i(t, x) + \frac{m}{2} n_i(t, x+a)}{N(1 - r(\tilde{n})\tau)}, \tag{10}$$

where $\tilde{n} = \frac{m}{2} n_1(t, x-a) + (1-m) n_1(t, x) + \frac{m}{2} n_1(t, x+a) + \frac{m}{2} n_2(t, x-a) + (1-m) n_2(t, x) + \frac{m}{2} n_2(t, x+a)$ is the total population density after dispersal. Note that $p_1 + p_2 < 1$ in patches, where the population density is less than the carrying capacity.

In the continuum limit, when $r(n)\tau \ll 1$ and $ka \ll 1$, our model becomes equivalent to Eq. (1) for the population density and to Eq. (58) for the relative fraction of the two genotypes with $D = ma^2/2$, $\gamma_n = (1 - n/N)/\tau$, and $\gamma_f = 1/(a\tau)$. For simplicity, we set both a and τ to 1 in all of our simulations.

Acknowledgements We thank Jeff Gore and Saurabh Gandhi for useful discussions. This work was supported by a grant from the Simons Foundation (#409704, Kirill S. Korolev; #327934, Oskar Hallatschek), by the startup fund from Boston University to Kirill S. Korolev, and by a National Science Foundation Career Award (#1555330, Oskar Hallatschek). Simulations were carried out on the Boston University Shared Computing Cluster.

Supplemental Information

Supplemental Information (SI) provides additional results and explanations that further support the classification of reaction-diffusion waves into pulled, semi-pushed, and fully-pushed waves. SI can be roughly divided into two parts. The first part (up to "Cutoffs for deterministic and fluctuating fronts") mostly reviews previous findings while the second part contains mostly new results. We describe the content of each part in more detail below.

The goal of the first half is to introduce common notation, clarify terminology, and state the results in a way that makes it easy to compare theoretical predictions to simulations, experiments, and field studies. The first two sections summarize the standard theory of deterministic reaction-diffusion waves and explain the terms that physicists and ecologists use for cooperative growth. The third section discusses the patterns of ancestry in reaction-diffusion waves. The fourth section introduces demographic fluctuations and genetic drift paying special attention to distinguishing fluctuations in population density from fluctuations in the genetic composition of the population. This distinction is not always drawn in the literature, but is important for applying the theory to specific populations. Sections V-VIII develop a perturbation theory in $1/N$ to compute Δv, D_{f}, and Λ. The only new results here are the second order perturbation theory for Δv and the expressions for D_{f} and Λ for exactly solvable models.

The second half of the SI begins with section IX, which shows how to regularize the perturbation theory by introducing an effective cutoff at low population densities. The value of the cutoff differs between deterministic and fluctuating fronts and is a nontrivial result from our work. The following section contains our main arguments for the existence of the three distinct classes of reaction-diffusion waves. This section combines the results of the perturbation theory with an appropriate cutoff and provides the derivation of the scaling exponents α_{V}, α_{D}, and α_{H}. The separation between foci of growth, ancestry, and diversity is discussed in Section XI, and the details of computer simulations and data analysis are given in Section XII. The final section of the SI contains additional results from simulations. In particular, we show that (i) the perturbation theory agrees with simulations for fully-pushed waves without any adjustable parameters; (ii) the scaling of Δv with N is the same as for D_{f} and Λ; (iii) the predicted scaling exponents match simulation results for both deterministic and fluctuating fronts. In this section, we also show that waves propagating into a metastable state are fully-pushed.

Table of Contents

I. Classification of the growth rate $r(n)$ in physics and ecology: metastability and Allee effects 18

II. Deterministic theory and classification of reaction-diffusion waves 18

Reduced equation for traveling wave solutions

Behavior near the boundaries

Waves propagating into a metastable state are pushed

Infinite number of solutions for propagation into an unstable state

Fisher waves - a simple case of pulled waves

Transition between pulled and pushed regimes of propagation into an unstable state

Summary

Example of pulled and pushed waves in exactly solvable models
Connection to the model of cooperative growth in the main text
Other exactly solvable models
Comments on notation
III. Dynamics of neutral markers and fixation probabilities 25
Forward-in-time dynamics
Fixation probabilities
Contribution to neutral evolution by different regions of the front
Spatial distribution of ancestors
Backward-in-time dynamics and the patterns of ancestry
Fixation probabilities and ancestry in pulled vs. pushed waves
Evaluation of integrals
IV. Demographic fluctuations and genetic drift 32
Fluctuations in population size
Fluctuations in population composition
Relationship between demographic fluctuations and genetic drift
Fluctuations in spatial models
V. Correction to the wave velocity, v, due to a cutoff 35
VI. Diffusion constant of the front, D_{f} 38
Perturbation theory for demographic fluctuations
Perturbation theory for migration fluctuations
Results for exactly solvable models
VII. Correction to velocity due to demographic fluctuations 43
VIII. Rate of diversity loss, Λ 47
Forward-in-time analysis of the decay of heterozygosity
Backward-in-time analysis of lineage coalescence
Explicit results for Λ in exactly solvable models and connection
IX. Cutoffs for deterministic and fluctuating fronts 52
Cutoff for deterministic fronts
Cutoff for fluctuating fronts
Cutoff for pushed waves expanding into a metastable state
Cutoff for pushed waves expanding into an unstable state
Cutoff for pulled waves
X. Scaling of Δv, D_{f}, and Λ in pulled, semi-pushed, and fully-pushed waves . 56
$1/N$ scaling in fully-pushed waves
N^{α} scaling in semi-pushed waves
Logarithmic scaling in pulled waves
Scaling with N in deterministic fronts
Comparison of Δv due to a cutoff in growth and due to demographic fluctuations

XI. Precise definition of the foci of growth, ancestry, and diversity 59
XII. Computer simulations 60
Alternative representation of the simulation algorithm
Initial conditions and simulation details
Front velocity measurements
Front diffusion measurements
Heterozygosity measurements
XIII. Supplemental results and figures 63

I. Classification of the growth rate $r(n)$ in physics and ecology: metastability and Allee effects

Here, we introduce the terminology used to characterize the growth term in Eq. (1) in physics and ecology.

The physics literature typically distinguishes between propagation into an unstable state when $r(0) > 0$ and propagation into a metastable state when $r(0) < 0$. The main difference between these two cases is their response to small perturbations. For $r(0) > 0$, the introduction of any number of organisms into an uncolonized habitat results in a successful invasion, so $n = 0$ is an unstable state. In contrast, for $r(0) < 0$, invasions fail when the number of introduced organisms is sufficiently small. Large introductions, however, do result in an invasion, so $n = 0$ is a metastable state. Since $n = N$ is stable against any perturbation, it is referred to as a stable state. When $n = 0$ and $n = N$ are the only states stable against small perturbations, $r(n)$ is often termed bistable.

In ecology, populations with a metastable state at $n = 0$ ($r(0) < 0$) are said to exhibit a strong Allee effect. When $r(0) > 0$, the growth dynamics is further classified as exhibiting either a weak Allee effect or no Allee effect. A population exhibits no Allee effect if $r(0) \geq r(n)$ for all n; otherwise, it exhibits a weak Allee effect. A common example of $r(n)$ without an Allee effect is the logistic growth, for which $r(n)$ decays monotonically from r_0 at $n = 0$ to 0 at $n = N$.

II. Deterministic theory and classification of reaction-diffusion waves

The goal of this section is to explain the difference between pulled and pushed waves. Despite considerable work on the topic, it is easy to conflate related but distinct properties of traveling waves, and few concise and self-contained accounts are available in the literature. Here, we only provide minimal and mostly intuitive discussion following Ref. [1], which is one of the most lucid and comprehensive reviews on propagating reaction-diffusion fronts. This section contains no new results.

The distinction between pulled and pushed waves arises already at the level of deterministic reaction-diffusion equations, so stochastic effects are not considered in this section. Specifically, we are interested in the asymptotic behavior as $t \to +\infty$ of the solutions of the following one-dimensional problem:

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + r(n)n, \tag{11}$$

where $r(n)$ is the per capita growth rate that is negative for large n, but could be either positive or negative at small n. Except for possibly $n = 0$, we assume that there is only one other stable fixed point of $dn/dt = r(n)n$ at $n = N$, where N is the carrying capacity. We further assume that the initial conditions are sufficiently localized, e.g. $n(0, x)$ is strictly zero outside a finite domain. Under these assumptions, the long time behavior of $n(t, x)$ is that of a traveling wave:

$$\begin{aligned} &n(t,x) = n(\zeta), \\ &\zeta = x - vt, \\ &n(-\infty) = N, \\ &n(+\infty) = 0. \end{aligned} \tag{12}$$

and our main task is to determine the wave velocity v and the shape of the front in the comoving reference frame $n(\zeta)$. Throughout this paper, we focus on the right-moving part of the expansion; the behavior of the left-moving expansion is completely analogous.

Reduced equation for traveling wave solutions
By substituting Eq. (12) into Eq. (11), we obtain the necessary condition on v and $n(\zeta)$:

$$Dn'' + vn' + r(n)n = 0, \tag{13}$$

where primes denote derivatives with respect to ζ. The solutions of Eq. (13) clearly have a translational degree of freedom, i.e., if $n(\zeta)$ is a solution, then $n(\zeta + \text{const})$ is a solution. We can eliminate this degree of freedom by choosing the references frame such that $n(0) = N/2$. The general solution of Eq. (13) then has only one remaining degree of freedom, which corresponds to the value of $n'(0)$. For a given v, the existence of the solution depends on whether the value of $n'(0)$ can be adjusted to match the boundary conditions specified by Eq. (12).

Depending on the behavior of the solution at $\zeta \to \pm\infty$, each boundary condition may or may not provide a constraint on the solution and, therefore, remove either one or zero degrees of freedom. To determine the number constraints, we linearize Eq. (13) near each of the boundaries.

Behavior near the boundaries
For $\zeta \to -\infty$, we let $u = N - n$ and obtain:

$$Du'' + vu' + \left.\frac{dr}{dn}\right|_{n=N} Nu = 0, \tag{14}$$

which has the following solution:

$$u = A_b e^{k_b x} + C_b e^{q_b x}, \tag{15}$$

where

$$\begin{aligned} k_b &= \frac{\sqrt{v^2 - 4D \left.\frac{dr}{dn}\right|_{n=N} N} - v}{2D}, \\ q_b &= \frac{-\sqrt{v^2 - 4D \left.\frac{dr}{dn}\right|_{n=N} N} - v}{2D}. \end{aligned} \tag{16}$$

Here, we used subscript b to indicate that we refer to the behavior of the population bulk. Because the carrying capacity is an attractive fixed point, dr/dn is negative at $n = N$, and therefore $k > 0$ and $q < 0$. We then conclude that the boundary condition at $\zeta \to -\infty$ requires that $C_b = 0$ and thus selects a unique value of $n'(0)$.

Next, we analyze the behavior at the front and linearize Eq. (13) for small n:

$$Dn'' + vn' + r(0)n = 0, \tag{17}$$

which has the following solution:

$$n = Ae^{-kx} + Ce^{-qx}, \tag{18}$$

where

$$\begin{aligned} k &= \frac{v + \sqrt{v^2 - 4Dr(0)}}{2D}, \\ q &= \frac{v - \sqrt{v^2 - 4Dr(0)}}{2D}. \end{aligned} \tag{19}$$

Waves propagating into a metastable state are pushed
The implications of Eq. (19) depend on the sign of $r(0)$. For a strong Allee effect ($r(0) < 0$), that is when the wave propagates into a metastable state, we find that $k > 0$ and $q < 0$. In consequence, the boundary condition requires that $C = 0$ and imposes an additional constraint. For an arbitrary value of v this constraint cannot be satisfied because the value of $n'(0)$ is already determined by the behavior in the bulk. However, there could be a value of v for which the constraints in the bulk and at the front are satisfied simultaneously. This special v is then the velocity of the wave. Because the value of v depends population dynamics throughout the wave front, the propagation into a metastable state is classified as a pushed wave.

Infinite number of solutions for propagation into an unstable state
For $r(0) > 0$, i.e. when the wave propagates into an unstable state, the analysis is more subtle

because the behavior at the front does not fully constrain the velocity of the wave. To demonstrate this, we draw the following two conclusions from Eq. (19). First, v must be greater or equal to $v_{\text{F}} = 2\sqrt{Dr(0)}$; otherwise, the solutions are oscillating around zero as $\zeta \to +\infty$ and violate the biological constraint that $n \geq 0$. Second, both k and q are positive for $v \geq v_{\text{F}}$, so the solution decays to 0 for arbitrary A and C. Thus, the boundary condition at the front does not impose an additional constraint, and one can find a solution of Eq. (13) satisfying Eq. (12) for arbitrary $v \geq v_{\text{F}}$.

Fisher waves—a simple case of pulled waves

The multiplicity of solutions for $r(0) > 0$ posed a great challenge for applied mathematics, statistical physics, and chemistry, and its resolution greatly stimulated the development of the theory of front propagation. The simplest context in which one can show that the wave velocity is unique is when there is no Allee effect, i.e. $r(0) \geq r(n)$ for all $n \in (0, N)$. This condition guarantees that the expansion velocity for Eq. (11) is less or equal than the expansion velocity for the dynamical equation linearized around $n = 0$:

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + r(0)n. \tag{20}$$

Moreover, we show below that the upper bound on v from the linearized dynamics coincides with the lower bound imposed by Eq. (19). Thus, the velocity of the wave is $2\sqrt{Dr(0)}$. This analysis was first carried out in Ref. [40] for an equation, which is now commonly known as the Fisher, Fisher-Kolmogorov, or Fisher-Kolmogorov-Petrovskii-Piskunov equation [39, 40]. Therefore, we refer to waves with $r(0) \geq r(n)$ as Fisher or Fisher-like waves. Since the expansion of the population is determined by the linearized dynamics, Fisher waves are pulled.

The upper bound on the velocity from the linearized equation can be obtained for an arbitrary initial condition via the standard technique of Fourier and Laplace transforms [1]. However, it is much simpler to use an initial condition for which a closed form solution is available. The results of such an analysis are completely general because the velocity of the wave should not depend on the exact shape of $n(0, x)$ at least when the population density is zero outside a finite domain. So, we make a convenient choice of $n(0, x) = \delta(x)$, where $\delta(x)$ is the Dirac delta function. Then, the solution of Eq. (20) is the product of an exponential growth term and a diffusively widening Gaussian:

$$n(t, x) = \frac{1}{\sqrt{4\pi Dt}} e^{r(0)t} e^{-\frac{x^2}{4Dt}}. \tag{21}$$

To determine the long-time behavior, it is convenient to shift into a comoving reference frame (see Eq. (12)):

$$n(t, \zeta) = \frac{1}{\sqrt{4\pi Dt}} e^{\left(r(0) - \frac{v^2}{4D}\right)t} e^{-\frac{\zeta^2}{4Dt}} e^{-\frac{v}{4D}\zeta}. \tag{22}$$

From the first exponential term, it immediately clear that the wave velocity must be equal to $2\sqrt{Dr(0)}$; otherwise the population size will either exponentially grow or decline as $t \to +\infty$. As we said

earlier, this result demonstrates that the upper and the lower bounds on v coincide with each other and uniquely specify the expansion velocity.

The last exponential term in Eq. (22) further shows that the wave profile decays exponentially with ζ as $e^{-k_{\rm F}\zeta}$, where

$$k_{\rm F} = \sqrt{\frac{r(0)}{D}} \tag{23}$$

consistent with Eq. (19) for $v = v_{\rm F}$.

Finally, the middle exponential term in Eq. (22) describes the transition from a sharp density profile at $t = 0$ to the asymptotic exponential decay of $n(\zeta)$. For $\zeta < 2\sqrt{Dt}$, this term is order one, and the front is well approximated by the asymptotic exponential profile. For $\zeta > 2\sqrt{Dt}$, the Gaussian term becomes important, and the shape of the front is primarily dictated by the diffusive spreading from the initial conditions. Thus, the linear spreading dynamics builds up a gradual decay of the population density starting from a much sharper population front due to localized initial conditions.

The steepness of the front created by the linearized dynamics is very important for the transition from pulled to pushed waves, so we emphasize that Eq. (23) specifies the lower bound on the rate of exponential decay of the population density. Indeed, Eq. (22) indicates that n decays no slower than $e^{-k_{\rm F}\zeta}$ at all times.

Transition between pulled and pushed regimes of propagation into an unstable state
Up to here, we have shown that waves are pushed when the Allee effect is strong, but the waves are pulled when there is no Allee effect. Now, we shift the focus to the remaining case of the weak Allee effect and show that the transition from pulled to pushed waves occurs when the growth at intermediate n is sufficiently high to support an expansion velocity greater than the linear spreading velocity $v_{\rm F}$.

To understand the origin of pushed waves, we need to consider the behavior of the population density near the front for $v > v_{\rm F}$. Equations (18) and (19) predict that $n(\zeta)$ is a sum of two exponentially decaying terms with different decay rates: One term decays with the rate $k > k_{\rm F}$, but the other with the rate $q < k_{\rm F}$. This behavior is in general inconsistent with the solution of the linearized dynamical equation (20). Indeed, our analysis of a localized initial condition suggests and more rigorous analysis proves [1] that $n(t, x)$ decays to zero at least as fast as $k_{\rm F}$. Therefore, we must require that $C = 0$ just as in the case of propagation into a metastable state.

The extra requirement makes the problem of satisfying the boundary conditions overdetermined. As a result, there are two alternatives. First, there could be no solutions for any $v > v_{\rm F}$. In this case, the expansion must be pulled because it is the only feasible solution. Second, a special value of v exists for which the boundary conditions at $\zeta \to \pm\infty$ can be satisfied simultaneously. In this case, the wave is pushed because the pulled expansion at low n is quickly overtaken by the faster expansion from the bulk.

For completeness, we also mention that initial conditions that are not localized and decay asymptotically as $e^{-\tilde{k}x}$ with $\tilde{k} < k_{\rm F}$ lead to pulled expansions with $v = 1/2 v_{\rm F}(\tilde{k}/k_{\rm F} + k_{\rm F}/\tilde{k}) > v_{\rm F}$. This

result immediately follows from the solution of Eq. (20) using either Laplace transforms or a substitution of an exponential ansatz $n = e^{-\tilde{k}(x-vt)}$. Discreteness of molecules or individuals, however, make such initial conditions impossible, so the wave propagation with $v > v_{\rm F}$ can only describe transient behavior for initial conditions with a slow decay at large x.

Summary

To summarize, we have shown that all waves propagating into a metastable state are pushed, and their profile decays as $e^{-k\zeta}$ at the front with k given by Eq. (19). Propagation into an unstable state could be either pulled or pushed depending on the relative strength of the growth behind vs. at the front. Pushed waves expand with $v > v_{\rm F}$, and their profile decays as $e^{-k\zeta}$. In contrast, pulled waves expand with $v = v_{\rm F}$, and their profile decays as $e^{-k_{\rm F}\zeta}$. In general, the class of a wave propagating into an unstable state cannot be determined without solving the full nonlinear problem (Eq. (11) or Eq. (13)). However, there are a few rigorous results that determine the class of the wave from simple properties of $r(n)$. To the best of our knowledge, the most general result of this type is that waves are always pulled when there is no Allee effect.

Example of pulled and pushed waves in exactly solvable models

We conclude this section by illustrating the transition from pulled to pushed waves in an exactly solvable model. Our main goal is to provide an concrete example for the abstract concepts discussed so far. In addition, the models presented below are used in simulations and to explicitly calculate the diffusion constant of the front and the rate of diversity loss in the regime of fully-pushed waves.

Consider $r(n)$ specified by

$$r(n) = g_0 \left(1 - \frac{n}{N}\right) \left(\frac{n}{N} - \frac{n^*}{N}\right), \tag{24}$$

where g_0 sets the time scale of growth, N is the carrying capacity, and n^* is a parameter that controls the strength of an Allee effect. For every value of n^*, the growth function $r(n)$ can be classified in one of three types: no Allee effect, weak Allee effect, or strong Allee effect. The growth function does not exhibit an Allee effect when $r(0) \geq r(n)$ for all $n \in (0, N)$. For the model defined above, the region of no Allee effect corresponds to $n^*/N \leq -1$. When an Allee effect is present, one distinguished between a strong Allee effect, $r(0) < 0$, and a weak Allee effect, $r(0) \geq 0$. Thus, the Allee effect is weak for $n^*/N \in (-1, 0]$ and strong for $n^* > 0$. In the latter case, n^* represents the minimal population density required for net growth and is known as the Allee threshold. In the following, we will refer to n^* as the Allee threshold regardless of its sign.

With $r(n)$ specified by Eq. (24), Eq. (13) admits an exact solution [9, 48, 49, 72]:

$$\begin{aligned} n(\zeta) &= \frac{N}{1 + e^{\sqrt{\frac{g_0}{2D}}\zeta}} \\ v &= \sqrt{\frac{g_0 D}{2}} \left(1 - 2\frac{n^*}{N}\right). \end{aligned} \tag{25}$$

For $n^* > 0$, we expect a unique pushed wave, so Eq. (25) must provide the desired solution.

Allee threshold, n^*/N	$(-\infty, -1]$	$(-1, -0.5]$	(-0.5, 0]	(0, 0.5)
Allee effect	none	weak		strong
Stability of invaded state	unstable			metastable
Type of expansion	pulled		pushed	

Table 1. Comparison of wave type, state stability, and Allee effect for an exactly solvable model of range expansions given by Eq. (24). Note that the transition from pulled to pushed waves does not coincide with a change in the type of growth. In particular, the pulled-pushed transition is distinct from the transition between no Allee effect and a weak Allee effect; it is also distinct from the transition between propagation into unstable and metastable state. For $c^*/N > 1/2$, the relative stability of the populated and unpopulated states changes, and the expansion wave propagates from $n = 0$ state into $n = N$ state.

For $n^*/N < -1$, we know that the wave must be pulled with $v = v_{\mathrm{F}} = 2\sqrt{-Dg_0 n^*/N}$ and $k = k_{\mathrm{F}} = \sqrt{-g_0 n^*/(ND)}$. To identify the transition between pulled and pushed waves, we equate the two expressions for the velocity and obtain that the critical value of the Allee threshold is given by $n^*/N = -1/2$. Thus, the transition between pulled and pushed waves occurs within the region of a weak Allee effect in agreement with the general theory developed above. The behavior of this exactly solvable model is summarized in Table 1.

Connection to the model of cooperative growth in the main text
The model of cooperative growth that we defined in the main text (Eq. (3)) is a simple reparameterization of Eq. (24) with $r_0 = -g_0 n^*/N$ and $B = -N/n^*$. In consequence, $r(n)$ exhibits no Allee effect for $B \leq 1$ and a weak Allee effect for $B > 1$. A strong Allee effect is not possible for any B because $r(0) > 0$. Hence, the wave always propagates into an unstable state. This model choice was convenient for us because it ensures that the transition between semi-pushed and fully-pushed waves is unambiguously distinct from the transition between propagation into unstable and metastable states.

The transition between pulled and pushed waves occurs at $B = 2$. For pushed waves ($B > 2$), the velocity and front shape are given by

$$\begin{aligned} n(\zeta) &= \frac{N}{1 + e^{\sqrt{\frac{r_0 B}{2D}}\zeta}} \\ v &= \sqrt{\frac{r_0 D B}{2}}\left(1 + \frac{2}{B}\right), \end{aligned} \tag{26}$$

while for pulled waves ($B \leq 2$) the corresponding results are

$$\begin{aligned} n(\zeta) &\sim e^{-\sqrt{\frac{r_0}{D}}\zeta}, \quad \zeta \to +\infty, \\ v &= 2\sqrt{D r_0}. \end{aligned} \tag{27}$$

Although Eq. (3) is equivalent to Eq. (24), our computer simulations of these models reveal complementary information because they explore different cuts through the parameter space. In particular, when we vary B in one model, we change both g_0 and n^*/N in the other model. Similarly, changes

in n^*/N modify both r_0 and B. Concordant results for the two parameterizations indicate that the transitions from pulled to semi-pushed and from semi-pushed to fully-pushed waves are universal and do not depend on the precise definition of cooperativity or the strength of an Allee effect.

Other exactly solvable models
For completeness, we also mention that exact solutions for pushed waves are known for a slightly more general class of $r(n)$ than the quadratic growth function discussed so far. The following results are from Ref. [73], which is an excellent resource for exactly solvable models of traveling waves.

For $r(n)$ defined by

$$r(n) = g_0 \left[1 - \left(\frac{n}{N}\right)^b\right] \left[\left(\frac{n}{N}\right)^b - \left(\frac{n^*}{N}\right)^b\right], \tag{28}$$

pushed waves occurs for $n^*/N > -(b+1)^{-1/b}$, and their velocity and profile shape are given by

$$\begin{aligned} v &= \sqrt{(b+1)g_0 D}\left[\frac{1}{b+1} - \left(\frac{n^*}{N}\right)^b\right], \\ n(\zeta) &= \frac{N}{\left(1 + e^{b\sqrt{\frac{g_0}{(b+1)D}}\zeta}\right)^{1/b}}. \end{aligned} \tag{29}$$

The transition point from semi-pushed to fully-pushed waves follows from these results and the condition that $v = \sqrt{9/8}v_{\mathrm{F}}$. The value of this critical Allee threshold is given by $n^*/N = -[2(b+1)]^{-1/b}$. Finally, the transition from weak to no Allee effect occurs at $n^* = -N$ and from weak to strong Allee effect at $n^* = 0$.

Comments on notation
For all models of $r(n)$, it is sometimes convenient to use the normalized population density $\rho = n/N$. For the exactly solvable models introduced above, it is also convenient to define $\rho^* = n^*/N$. This notation is used in the main text and in the following sections.

We can also now be more precise about the definitions of population bulk, front, interior regions of the front, and the leading edge, which we use throughout the paper. The population bulk is defined the region where n is close to N and Eq. (15) holds. Similarly, the leading edge, the tip of the front, front edge, *etc.* refer to the region of $n \ll N$, where Eq. (18) holds. The region with intermediate n is termed as the front or more precisely as the interior region of the front or the bulk of the front. We tried to avoid using the generic term front whenever that can cause confusion between the leading edge and the interior region of the front.

III. Dynamics of neutral markers and fixation probabilities

Heritable neutral markers provide a window in the internal dynamics of an expanding population. These dynamics can be studied either forward in time or backward in time. The former approach

describes how the spatial distribution of neutral markers changes over time and provides an easy way to compute the fixation probabilities of neutral mutations. The latter approach describes the patterns of ancestry that emerge during a range expansion and provides a natural way to infer population parameters from genetic data. The main goal of this section is to demonstrate that both forward-in-time and backward-in-time dynamics fundamentally change at the transition from pulled to pushed waves. In pulled waves, all organisms trace their ancestry to the very tip of the expansion front, which is also the only source of successful mutations. In contrast, the entire expansion front contributes to the evolutionary dynamics in pushed waves. This section contains no new results except for the analytical calculation of the fixation probabilities in the exactly solvable models. The discussion largely follows that in Refs. [37] and [38]. Our main goal here is to introduce the notation to be used in the following sections and explain how the patterns of ancestry depend on cooperativity.

Forward-in-time dynamics

Let us consider a subpopulation carrying a neutral marker i and describe how its population density $n_i(t,x)$ changes forward in time. Since the growth and migration rates are the same for all markers, n_i obey the following equation:

$$\frac{\partial n_i}{\partial t} = D\frac{\partial^2 n_i}{\partial x^2} + r(n)n_i, \tag{30}$$

where, as before, n is the total population density, which is given by $n = \sum_i n_i$ if all individuals are labeled by some marker.

To isolate the behavior of neutral markers from the overall population growth, it is convenient to define their relative frequency in the population: $f_i = n_i/n$. From Eqs. (11) and (30), it follows that [37, 50]

$$\frac{\partial f_i}{\partial t} = D\frac{\partial^2 f_i}{\partial x^2} + 2D\frac{\partial \ln n}{\partial x}\frac{\partial f_i}{\partial x}. \tag{31}$$

The new advection-like term arises from the nonlinear change of variables and accounts for a larger change in f_i due to immigration from regions with high population density compared to regions with low population density. The main effect of the new term is to establish a "flow" of f_i from the posterior to the anterior of the front.

It is also convenient to shift into the comoving reference frame (see Eq. (12)) in order to focus on the dynamics that occurs at a fixed position within the front region rather than at a fixed position in the stationary reference frame. The result reads

$$\frac{\partial f_i}{\partial t} = D\frac{\partial^2 f_i}{\partial \zeta^2} + v\frac{\partial f_i}{\partial \zeta} + 2D\frac{\partial \ln n}{\partial \zeta}\frac{\partial f_i}{\partial \zeta}. \tag{32}$$

We now drop the index i because, for the rest of this section, we focus on the frequency of a single marker, which we denote simply by $f(t,\zeta)$. In the following, we also assume that $n(t,\zeta)$ has

reached the steady-state given by Eq. (13). The evolutionary dynamics are typically much slower than ecological dynamics, so the initial transient in the dynamics of n could be neglected.

The analysis of Eq. (32) is greatly simplified by the existence of a time-invariant quantity:

$$\pi = \frac{\int_{-\infty}^{+\infty} f(t,\zeta) n^2(\zeta) e^{v\zeta/D} d\zeta}{\int_{-\infty}^{+\infty} n^2(\zeta) e^{v\zeta/D} d\zeta}, \tag{33}$$

which was first demonstrated in Refs. [37] and [38].

To show that π is conserved, we evaluate its time derivative:

$$\frac{d\pi}{dt} = \frac{\int_{-\infty}^{+\infty} \frac{\partial f}{\partial t} n^2 e^{v\zeta/D} d\zeta}{\int_{-\infty}^{+\infty} n^2 e^{v\zeta/D} d\zeta}, \tag{34}$$

and replace $\partial f/\partial t$ by the right hand side of Eq. (32). The numerator can then be simplified via the integration by parts to eliminate the derivatives of f with respect to ζ in favor of f. This leads to the cancellation of all the terms and thereby proves that π does not depend on time.

The conservation on π makes it quite straightforward to determine the fixation probabilities, the spatial distribution of ancestors, and the contribution of different parts of the front to the neutral evolution. We now discuss each of these results separately.

Fixation probabilities
Since only spatial derivatives of f enter Eq. (32), we conclude that $f = \text{const}$ is a solution that describes the steady state after migration has smoothed out the spatial variations in the initial conditions. The value of the constant is given by π because

$$\pi = \lim_{t\to+\infty} \frac{\int_{-\infty}^{+\infty} f(t,\zeta) n^2(\zeta) e^{v\zeta/D} d\zeta}{\int_{-\infty}^{+\infty} n^2(\zeta) e^{v\zeta/D} d\zeta} = \frac{\int_{-\infty}^{+\infty} \left(\lim_{t\to+\infty} f\right) n^2 e^{v\zeta/D} d\zeta}{\int_{-\infty}^{+\infty} n^2 e^{v\zeta/D} d\zeta} = \lim_{t\to+\infty} f(t,\zeta). \tag{35}$$

At the level of deterministic dynamics, this result captures how the final fraction of a genotype depends on its initial distribution in the population. Genetic drift, however, leads to the extinction of all but one genotype, so $f(t,\zeta)$ should be interpreted as the average over the stochastic dynamics (one can take the expectation value of both sides of Eq. (58) below). Since the expected value of f is 1 times the probability of fixation plus 0 times the probability of extinction, we immediately conclude that the fixation probability equals π. Thus, given the initial distribution of a neutral marker $f(0,\zeta)$, we can obtain its fixation probability by evaluating the integrals in Eq. (33) at $t = 0$.

We can also express this result in terms of the absolute abundance of the neutral marker n_i (here we keep the subscript to distinguish n_i from the total population density). Since $f_i = n_i/n$, the fixation probability is given by

$$\pi = \frac{\int_{-\infty}^{+\infty} n_i(t,\zeta) n(\zeta) e^{v\zeta/D} d\zeta}{\int_{-\infty}^{+\infty} n^2(\zeta) e^{v\zeta/D} d\zeta}. \tag{36}$$

For a single organism present at location ζ_0, we can approximate n_i as $\delta(\zeta - \zeta_0)$ and thus obtain the fixation probability of a single mutant:

$$u(\zeta_0) = \frac{n(\zeta_0) e^{v\zeta_0/D}}{\int_{-\infty}^{+\infty} n^2(\zeta) e^{v\zeta/D} d\zeta}. \tag{37}$$

Contribution to neutral evolution by different regions of the front
From Eq. (37), we can determine how different regions contribute to the neutral evolution during a range expansion. Neutral evolution proceeds through two steps: first a random mutations appears somewhere in the population and second the frequency of the mutation fluctuations until the mutation either reaches fixation or becomes extinct. For a given mutation, the probability that it first occurred at location ζ is proportional to $n(\zeta)$, and its fixation probability is given by $u(\zeta)$. Thus, the fraction of fixed mutations that first originated at ζ is given by $n^2(\zeta)e^{v\zeta/D} / \int_{-\infty}^{+\infty} n^2(\zeta)e^{v\zeta/D} d\zeta$.

Spatial distribution of ancestors
Finally, we note that the last result also represents the probability that the ancestor of a randomly sampled individual from the population used to live at location ζ sufficiently long ago. To demonstrate this, we label all individuals between ζ and $\zeta + d\zeta$ at a long time in the past and note that the probability of a random individual to have its ancestor at ζ equals the expected number of labeled descendants. Since the long time limit of f is given by π, we immediately conclude that the probability distribution of ancestor locations is given by

$$S(\zeta) = \frac{n^2(\zeta) e^{v\zeta/D}}{\int_{-\infty}^{+\infty} n^2(\zeta) e^{v\zeta/D} d\zeta}. \tag{38}$$

Backward-in-time dynamics and the patterns of ancestry
To characterise the patterns of ancestry in a population, it is convenient to describe the dynamics of ancestral lineages backward in time. Following the approach of Ref. [37], we show below that the probability $S(\tau, \zeta)$ that an ancestor of a given individual lived at position ζ time τ ago is governed by the following equation [37]:

$$\frac{\partial S}{\partial \tau} = D \frac{\partial^2 S}{\partial \zeta^2} - v \frac{\partial S}{\partial \zeta} - 2D \frac{\partial}{\partial \zeta} \left(\frac{\partial \ln n}{\partial \zeta} S \right). \tag{39}$$

Here, the diffusion term randomizes the position of the ancestor; the term proportional to v pushes the ancestor towards the tip of the wave and reflects the change into the comoving reference frame; the last term pushes the ancestor towards the population bulk and reflects the fact that the ancestor is more likely to have emigrated from the region where the population density is higher. For large τ, it is easy to check that these forces balance and result in a stationary distribution $S(\zeta)$ given by

Eq. (38). Note that, unlike in Eq. (32) for the dynamics of f, the right hand side of Eq. (39) contains a divergence of a flux and, therefore, preserves the normalization of S, i.e. $\int S(\tau,\zeta)d\zeta = \text{const}$. Indeed, the probability that an ancestor was present somewhere in the population must always equal to one.

The derivation of Eq. (39) from Eq. (32) follows the standard procedure for changing from the forward-in-time to the backward-in-time description [74, 75] and consists of three steps. The first step is to define the "propagator" function that can describe both forward-in-time and backward-in-time processes. We denote this function as $G(t_d,\zeta_d;t_a,\zeta_a)$ and define it as the probability that a descendant located at ζ_d at time t_d originated from an ancestor who lived at time t_a at position ζ_a. On the one hand, with t_d and ζ_d fixed, G can be viewed as a function of t_a and ζ_a that specifies the probability distribution of ancestor location at a specific time. On the other hand, with t_a and ζ_a fixed, G can be viewed as a function of t_d and ζ_d that describes the spatial and temporal dynamics of the expected frequency of the descendants from all organisms that were present at ζ_a at time t_a. This two-way interpretation follows from the labeling thought-experiment that we used to derive $S(\zeta)$ using the forward-in-time formulation.

In the second step, we claim that G obeys the same equation as $f(t,\zeta)$, i.e. Eq. (32). This statement immediately follows from the forward-in-time interpretation of G. We formally state this result as

$$\frac{\partial G}{\partial t_d} = \mathcal{L}_{\zeta_d} G, \tag{40}$$

where $\mathcal{L}_{\zeta_d}$ is the linear operator from the right hand side of Eq. (32), and we used the subscript ζ_d to indicate variable on which the operator acts.

The third step is to derive an equation for G that involves only the ancestor-related variables. To this purpose, we consider an infinitesimal change in t_a,

$$\begin{aligned} G(t_d,\zeta_d;t_a-dt,\zeta_a) &= \int_{-\infty}^{+\infty} G(t_d,\zeta_d;t_a,\zeta')G(t_a,\zeta';t_a-dt,\zeta_a)d\zeta' \\ &= \int_{-\infty}^{+\infty} G(t_d,\zeta_d;t_a,\zeta')[G(t_a-dt,\zeta';t_a-dt,\zeta_a)+dt\mathcal{L}_{\zeta'}G(t_a-dt,\zeta';t_a-dt,\zeta_a)]d\zeta' \\ &= \int_{-\infty}^{+\infty} G(t_d,\zeta_d;t_a,\zeta')[\delta(\zeta'-\zeta_a)+dt\mathcal{L}_{\zeta'}\delta(\zeta'-\zeta_a)]d\zeta' \end{aligned} \tag{41}$$

where we used the Markov property of the forward-in-time dynamics, then Eq. (40), and finally the fact that $G(t,\zeta_a;t,\zeta_d)=\delta(\zeta_a-\zeta_d)$, which immediately follows from the definition of G. Equation (41) can be further simplified by expanding the left hand side in dt:

$$-\frac{\partial G}{\partial t_a} = \int_{-\infty}^{+\infty} G(t_d,\zeta_d;t_a,\zeta')\mathcal{L}_{\zeta'}\delta(\zeta'-\zeta_a)d\zeta'. \tag{42}$$

Finally, we integrate by parts to transfer the derivatives in $\mathcal{L}_{\zeta_a}$ from the delta function to $G(t_d, \zeta_d; t_a, \zeta')$:

$$-\frac{\partial G}{\partial t_a} = \int_{-\infty}^{+\infty} \delta(\zeta' - \zeta_a)\mathcal{L}^+_{\zeta'} G(t_d, \zeta_d; t_a, \zeta')d\zeta' = \mathcal{L}^+_{\zeta_a} G(t_d, \zeta_d; t_a, \zeta_a), \tag{43}$$

where, $\mathcal{L}^+_{\zeta_a}$ is the operator that results from the integration by parts and is known as the adjoint operator of $\mathcal{L}_{\zeta_a}$.

Equation (43) is the desired backward-in-time formulation that involves only the ancestor-related variables. To see that it is equivalent to Eq. (39), one needs to explicitly compute $\mathcal{L}^+_{\zeta_a}$, substitute the definition of $\tau = t_d - t_a$, and change the notation from G to S.

Fixation probabilities and ancestry in pulled vs. pushed waves
We conclude this section by comparing the fixation probabilities and patterns of ancestry in pulled and pushed waves. We focus on $S(\zeta)$ as a typical example; other quantities, e.g., $u(\zeta)$ can be analyzed in the same fashion.

Up to a constant factor, $S(\zeta)$ is given by $n^2(\zeta)e^{v\zeta/D}$. For large negative ζ, the exponential factor tends rapidly to zero indicating that the bulk of the wave contributes little to the neutral evolution and is unlikely to contain the ancestor of future generations. This conclusion applies to both pulled and pushed waves. The behavior at the front is more subtle because $n(\zeta) \to 0$ and $e^{v\zeta/D} \to +\infty$ as $\zeta \to +\infty$. To determine the scaling of $S(\zeta)$ at the front, we replace n by its asymptotic form $e^{-k\zeta}$ and obtain that

$$S(\zeta) \propto e^{-\zeta(2k - v/D)} \propto e^{-\zeta(k-q)}, \tag{44}$$

where the last expression follows from the fact that $v = D(k + q)$; see Eq. (19).

For pushed waves, $k > q$ so the tip of the front makes a vanishing contribution to the neutral evolution. Therefore, the main contribution to $S(\zeta)$ must come from the interior regions of the front. In fact, for the exactly solvable model specified by Eq. (24), one can express ζ in terms of n and show that

$$S(\zeta) = \frac{2\pi\rho^*}{\sin(2\pi\rho^*)}\rho^{1+2\rho^*}(1-\rho)^{1-2\rho^*}, \tag{45}$$

where $\rho = n/N$, and $\rho^* = n^*/N$. This result clearly demonstrates that $S(\zeta)$ is peaked at intermediate population densities (specifically at $\rho = (1 + 2\rho^*)/2$).

For pulled waves, $k = q$, and $n^2(\zeta)e^{v\zeta/D} \to$ const as $\zeta \to +\infty$. Thus, every point arbitrarily far ahead of the front contributes equally to the neutral evolution. Since the region ahead of the front is infinite, the relative contribution of the bulk and the interior of the front must be negligible compared to that of the leading edge. A more careful analysis requires one to impose a cutoff on ζ

at sufficiently low densities so that $S(\zeta)$ can be normalized. We show how to introduce such a cutoff in section IX.

In summary, we determined fixation probabilities and patterns of ancestry, which are plotted in Fig. 2 of the main text. We also compared the dynamics in pulled and pushed waves and found that they are driven by distinct spatial regions of the front. In pulled waves, the very tip of the front not only "pulls" the wave forward, but also acts as the focus of ancestry and the sole source of successful mutations. In pushed waves, however, the entire front contributes to both the expansion dynamics and evolutionary processes. We refer the readers to Ref. [37, 38, 56] for the original derivations of these results and further discussion.

Evaluation of integrals

Let us briefly explain how one can evaluate the integrals that appear in Eq. (38) and in similar equations for the diffusion constant of the front and the rate of diversity loss. The main insight is to change the independent variable from ζ to ρ using equation Eq. (25). The following formulas are useful for this purpose:

$$\begin{aligned}
\zeta &= \sqrt{\frac{2D}{g_0}} \ln\left(\frac{1-\rho}{\rho}\right), \\
d\zeta &= -\sqrt{\frac{2D}{g_0}} \frac{d\rho}{\rho(1-\rho)}, \\
e^{\zeta v/D} &= \rho^{-(1-2\rho^*)}(1-\rho)^{1-2\rho^*}, \\
\frac{d\rho}{d\zeta} &= -\sqrt{\frac{g_0}{2D}}\rho(1-\rho).
\end{aligned} \tag{46}$$

After this change of variable, all integrals become beta functions [76]. For example,

$$\int_{-\infty}^{+\infty} n^2(\zeta) e^{\frac{v\zeta}{D}} d\zeta = \sqrt{\frac{2D}{g_0}} \int_0^1 \rho^{2\rho^*}(1-\rho)^{-2\rho^*} = \mathrm{Be}(1+2\rho^*, 1-2\rho^*). \tag{47}$$

The integrals of this type can be evaluated in the complex plane. Specifically, one can equate the integral around the branch cut from 0 to 1 to the residue at $\zeta = \infty$. For the integral above, this results in

$$\sqrt{\frac{2D}{g_0}} \int_0^1 \rho^{2\rho^*}(1-\rho)^{-2\rho^*} = \frac{2\pi\rho^*}{\sin(2\pi\rho^*)}. \tag{48}$$

An alternative method to evaluate the integrals is to use the following properties of the gamma

and beta functions [76]:

$$\begin{aligned} \mathrm{Be}(x,y) &= \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \\ \Gamma(z)\Gamma(1-z) &= \frac{\pi}{\sin \pi z}. \end{aligned} \tag{49}$$

Using these formulas, we can evaluate all beta functions of the type $\mathrm{Be}(n+z, m-z)$, where n and m are positive integers and $z \in (0,1)$. The general expressions are

$$\begin{aligned} \mathrm{Be}(n+z, 1-z) &= \frac{\prod_{k=0}^{n-1}(z+k)}{n!} \frac{\pi}{\sin \pi z}, \\ \mathrm{Be}(n+z, m-z) &= \frac{\prod_{k=0}^{n-1}(z+k) \prod_{l=1}^{m-1}(l-z)}{(n+m-1)!} \frac{\pi}{\sin \pi z}, \quad m > 1. \end{aligned} \tag{50}$$

IV. Demographic fluctuations and genetic drift

In this section, we describe how to move beyond the deterministic approximation in Eqs. (11) and (31) and account for the effects of demographic fluctuations and genetic drift. Because the magnitude of the fluctuations depends on the details of the reproductive process, we need to introduce two additional functions of the population density $\gamma_n(n)$ and $\gamma_f(n)$ that describe the strength of fluctuations in the population size and composition respectively. This section contains no new results, and its main purpose is to specify the relevant notation and carefully discuss how stochastic dynamics should be added to deterministic reaction-diffusion equations.

For simplicity, we will consider well-mixed populations first and limit the discussion to neutral markers, e.g. genotypes that do not differ in fitness.

Fluctuations in population size

Demographic fluctuations in the size $n(t)$ of a well-mixed population arise due to the randomness of births and deaths. The simplest and most commonly used assumption is that an independent decision is made for each organism on whether it dies or to reproduces [77, 78]. In a short time interval Δt, the number of births and deaths are therefore two independent Poisson random variables with mean and variance equal to $\mu_b(n) n \Delta t$ for births and to $\mu_d(n) n \Delta t$ for deaths. Here, $\mu_b(n)$ and $\mu_d(n)$ are the per capita rates of birth and death respectively. Since the change in the population size is the difference between these two independent random variables, we conclude that the mean change of n is $[\mu_b(n) - \mu_d(n)] n \Delta t$ and the variance is $[\mu_b(n) + \mu_d(n)] n \Delta t$. In the continuum limit, the dynamics of the population size is then described by the following stochastic differential equation

$$\frac{dn}{dt} = [b(n) - d(n)]n + \sqrt{[b(n) + d(n)]n}\,\eta(t), \tag{51}$$

where $\eta(t)$ is the Itô white noise, i.e. $\langle \eta(t_1)\eta(t_2) \rangle = \delta(t_1 - t_2)$. In the following, we denote ensemble averages by angular brackets, and use $\delta(t)$ for the Dirac delta function.

The assumption that births and death events are independent random variables is however too restrictive. For example, the number of birth always equals the number of deaths in the classic

Wright-Fisher model, which exhibits no fluctuations in n as a result. In addition, the number of births or deaths could deviate from the Poisson distribution and therefore have the variance not equal to the mean. To account for such scenarios, we need to generalize Eq. (51) as follows

$$\frac{dn}{dt} = r(n)n + \sqrt{\gamma_n(n)n}\eta(t), \tag{52}$$

where $r(n)$ is the difference between the birth and death rates, and γ_n characterises the strength of the demographic fluctuations. The value of γ_n can be easily determined from model parameters because $\gamma_n n\Delta t$ is the sum of the variances of births and deaths during Δt minus twice their covariance.

Fluctuations in population composition
Genetic drift arises because the choice of the genotype that is affected by a specific birth or death event is random. This randomness does not lead to a change in the average abundance of neutral genotypes, but induces a random walk in the space of population compositions described by the species fractions $\{f_i\}$. It is easy to show that both births and deaths contribute equally to the increase in the variance of f_i in a short time interval Δt [77–79], so the strength of the genetic drift depends only on the total number of updates due to both births and deaths: $\gamma_f(n)n\Delta t$, where $\gamma_f = \mu_b(n) + \mu_d(n)$. Since probability to choose genotype i for an update is proportional to its current fraction in the population $f_i(t)$, the number of updates for each genotype will be given by a multinomial distribution with $\gamma_f(n)n\Delta t$ trials and outcome probabilities given by $\{f_i\}$. This leads to the following continuum limit [77–80]:

$$\frac{df_i}{dt} = \sqrt{\frac{\gamma_f(n)}{n} f_i(1-f_i)}\eta_i(t), \tag{53}$$

with the covariance structure of the noises η_i specified by

$$\langle \eta_i(t_1)\eta_j(t_2)\rangle = \delta(t_1 - t_2)\begin{cases} 1, & i = j, \\ -\sqrt{\dfrac{f_i f_j}{(1-f_i)(1-f_j)}}, & i \neq j. \end{cases} \tag{54}$$

The factor of $1/n$ under the square root in Eq. (53) arises because a single birth or death event changes the frequency of the genotype at most by $1/n$. The dependence on f_i reflects the properties of the multinomial distribution and ensures that the sum of f_i does not fluctuate and remains equal to 1.

Since genetic drift and demographic fluctuations are independent from each other, i.e. $\langle \eta(t_1)\eta_i(t_2)\rangle = 0$, and Eqs. (52), (53), and (54) completely specify population dynamics. In particular, one can easily obtain the dynamical equations for the genotype abundances n_i by differentiating $n_i = f_i n$.

An alternative, but completely equivalent, formulation of Eqs. (53) and (54) reads

$$\frac{df_i}{dt} = \sqrt{\frac{\gamma_f(n)}{n} f_i} \left(\tilde{\eta}_i(t) - \sqrt{f_i} \sum_j \tilde{\eta}_j \right), \tag{55}$$

where $\langle \tilde{\eta}_i(t_1)\tilde{\eta}_j(t_2)\rangle = \delta_{ij}\delta(t_1 - t_2)$; we use δ_{ij} to denote the Kronecker delta, i.e. the identity matrix. This alternative definition arises naturally when one derives the equations for f_i starting from the dynamical equations for species abundances n_i and shows that $\sum_j f_j$ is constant more clearly. We provide this formulation only for completeness and do not use in the following.

Relationships between demographic fluctuations and genetic drift
For the simple processes of uncorrelated births and deaths described by Eq. (51), one can show that $\gamma_f(n) = \gamma_n(n)$ [77, 78], but this is equality does not hold in general. For example, in the simulations that we describe below γ_f is independent of n, but γ_n monotonically decreases to zero as the population size approaches the carrying capacity. Nevertheless, in a wide set of models, $\gamma_f(0) = \gamma_n(0)$ because the dynamics of different genotypes becomes uncorrelated at low population densities and their fluctuations are determined by $\gamma_0 = \mu_b(0) + \mu_d(0)$. As a result, the fluctuations in pulled and semi-pushed waves depends only on γ_0 when the carrying capacity is large enough to justify the asymptotic limit.

Exceptions to $\gamma_f(0) = \gamma_n(0)$ are in principle possible, for example, when many cycles of birth and death occur without an appreciable change in the total population. Such dynamics could arise when a slow and quasi-deterministic niche construction is required to increase the current limit on the population size.

For completeness, we also mention that the dynamical equations for n_i take a particularly simple form when $\gamma_n(n) = \gamma_f(n) = \gamma(n)$:

$$\frac{dn_i}{dt} = g(n)n_i + \sqrt{\gamma(n)n_i}\tilde{\eta}_i(t), \tag{56}$$

where $\langle \tilde{\eta}_i(t_1)\tilde{\eta}_j(t_2)\rangle = \delta_{ij}\delta(t_1 - t_2)$. Thus, for simple birth-death models, the fluctuations in genotype abundances are independent from each other as expected. Although Eq. (56) is often used as a starting point for the analysis [81], it does not capture the full complexity of possible eco-evolutionary dynamics.

Fluctuations in spatial models
It is straightforward to extend the above discussion to spatial populations where n and f_i depend on both t and x. The net result is that Eqs. (11) and (31) acquire stochastic terms specified by Eqs. (52) and (53). The results read

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + r(n)n + \sqrt{\gamma_n(n)n}\eta(t), \tag{57}$$

and

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} + 2\frac{\partial \ln n}{\partial x}\frac{\partial f}{\partial x} + \sqrt{\frac{\gamma_f(n)}{n} f_i(1-f_i)}\eta_i(t). \tag{58}$$

The noise-noise correlations are specified by the following equations

$$\langle \eta(t_1, x_1)\eta_i(t_2, x_2)\rangle = 0, \tag{59}$$

$$\langle \eta(t_1, x_1)\eta(t_2, x_2)\rangle = \delta(t_1 - t_2)\delta(x_1 - x_2), \tag{60}$$

and

$$\langle \eta_i(t_1, x_1)\eta_j(t_2, x_2)\rangle = \delta(t_1 - t_2)\delta(x_1 - x_2)\left[\delta_{ij} - (1-\delta_{ij})\sqrt{\frac{f_i f_j}{(1-f_i)(1-f_j)}}\right]. \tag{61}$$

Note that, in Eqs. (57) and (58), we omitted a noise term that accounts for the randomness of migration (or diffusion in the context of chemical reactions). Such noise inevitably arises when each organisms makes an independent decision on whether to migrate to a particular nearby site. Because this noise conserves the number of individual it appears as a derivative of the flux in the dynamical equation for n_i. The general form of this noise is $\partial_x[\sqrt{\gamma_m n_i}\chi_i]$ with $\langle \chi_i(t_1, x_1)\chi_j(t_2, x_2)\rangle = \delta_{ij}\delta(t_1 - t_2)\delta(x_1 - x_2)$, and χ_i are uncorrelated with η and η_i [54].

Migration noise does not typically lead to any new qualitative dynamics, and we will show below it leads to the same scaling of the fluctuations with the bulk population density N. Moreover, migration noise is often negligible compared to genetic drift. For example, it can be neglected when the migration rate is small or the number of organisms at the dispersal stage is much larger then the number of reproducing adults (compare the number of seeds vs. the number of trees). We do not consider migration noise further because it is absent in our computer simulation. For the sake of simplicity and greater computational speed, we chose to perform the migration update deterministically.

V. Correction to the wave velocity, v, due to a cutoff

How do demographic fluctuations modify the dynamics of wave propagation? This question is central to our paper and has generated significant interest in nonequilibrium statistical physics. Most early studies explored how demographic stochasticity modifies the expansion velocity [23, 31, 33, 36, 82]. While velocity corrections are small and likely negligible in the context of range expansions, they are essential for the description of evolving populations, which are often modeled as traveling waves in fitness space [23, 24, 56]. More importantly, wave velocity serves a salient and easy to measure observable that has been frequently used to test the theories of fluctuating fronts. This section shows how to compute the corrections to wave velocity using perturbation theory. All results in this section have been derived previously in Refs. [1, 31, 54, 83–85]. Our main goal here is to introduce the relevant notation and to explain the perturbation theory in the simplest context.

Because non-linear stochastic equations are notoriously difficult to analyze, a direct calculation of v is challenging, and several approximate approaches were developed instead [1]. In this section, we describe the simplest of these approaches that imposes a cutoff on the growth rate below a certain population density n_c:

$$r_{\text{cutoff}}(n) = r(n)\theta(n - n_c), \tag{62}$$

where $\theta(n)$ is the Heaviside step function, which equals one for positive arguments and zero for negative arguments.

Although the value of n_c must reflect the strength of the demographic fluctuations, it is not entirely clear how to determine n_c a priori. A natural guess is to set n_c to one over the size of the patch size in simulations so that no growth occurs in regions where the expected number of individuals is less than one. However, this choice does not capture the full complexity of demographic fluctuations as shown in section IX. For now, we keep n_c as an unspecified parameter and focus on the corrections to v due to the change in the growth rate specified by Eq. (62). The position of the cuttoff where the deterministic profile reaches n_c is denoted as ζ_c, i.e. $n(\zeta_c) = n_c$.

The corrections to v can be computed using a perturbation expansion in $\Delta r(n) = r_{\text{cutoff}}(n) - r(n)$. This approach has been developed by different groups either for computing the corrections due to a cutoff or for computing the diffusion constant of the front [1, 31, 54, 83–85].

Let us first introduce a convenient notation for the perturbation expansion that is also used in the following sections, where the perturbation is a stochastic variable rather than a deterministic cutoff. All quantities that solve the deterministic, unperturbed problem (Eq. (11) or Eq. (13)) are denoted with subscript d. All quantities that solve the full, perturbed problem are denoted without a subscript. And, the differences between the two types of quantities are denotes with Δ.

With this notation, the perturbed equation reads

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + n r_{\text{cutoff}}(n), \tag{63}$$

or equivalently

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + n r(n) + n\Delta r(n). \tag{64}$$

We seek the solution correct to the first order in Δr via the following ansatz

$$n(t, x) = n_d(x - v_d t - \Delta v t) + \Delta n(x - v_d t - \Delta v t), \tag{65}$$

where Δn is the correction to the shape of the stationary density profile, and Δv is the correction to the expansion velocity.

The zeroth order in perturbation theory yields the unperturbed equation:

$$Dn_d'' + v_d n_d' + r n_d = 0, \tag{66}$$

which is automatically satisfied by our choice of n_d.

To obtain the equations for the next order, we expand $\partial n/\partial t$ as

$$\frac{\partial n}{\partial t} \approx -v_d n_d' - n_d' \Delta v - v_d \Delta n', \tag{67}$$

the diffusion term as

$$D\frac{\partial^2 n}{\partial x^2} \approx D n_d'' + \Delta n'', \tag{68}$$

and the growth term as

$$r_{\text{cutoff}}(n) n \approx n_d r(n_d) + r(n_d)\Delta n + r'(n_d) n_d \Delta n + n_d \Delta r(n_d). \tag{69}$$

As before, we use primes to denote derivatives of functions of a single argument.

The resulting equation for the first order in perturbation theory reads

$$\mathfrak{L}_p \Delta n = -n_d \Delta v - n_d \Delta r, \tag{70}$$

where

$$\mathfrak{L}_p = D\frac{d^2}{d\zeta^2} + v_d \frac{d}{d\zeta} + r(n_d) + r'(n_d) n_d, \tag{71}$$

is the linear operator that acts on the comoving spatial variable $\zeta = x - vt = x - (v_d + \Delta v)t$.

Although Eq. (70) has two unknowns Δn and Δv, both quantities can be determined simultaneously because the solution for Δn exists only for a specific value of Δv. The constraint on Δv comes from the fact that $\mathfrak{L}_p$ has an eigenvalue equal to zero and, therefore, its image does not span the entire space of functions possible on the right hand side of Eq. (70). As a result, Δv must be chosen to make $-n_d \Delta v - n_d \Delta r$ lie in the image of $\mathfrak{L}_p$.

The zero mode of $\mathfrak{L}_p$ originates from the translational invariance of the unperturbed problem, for which both $n_d(\zeta)$ and $n_d(\zeta + \text{const})$ are solutions. Therefore, there should be no restoring

force from the dynamical equation for Δn that effectively translates the front by an infinitesimal distance $\delta\zeta$. Since $n_d(\zeta+\delta\zeta) \approx n_d(\zeta) + n_d'(\zeta)\delta\zeta$, we expect that $\Delta n \propto n_d'(\zeta)$ should not alter the left hand side of Eq. (70). Consistent with reasoning, the differentiation of Eq. (66) with respect to ζ shows that $\mathfrak{L}_p n_d' = 0$. Thus, $\mathfrak{L}_p$ indeed has a zero mode with n_d' being the right eigenvector. The corresponding left eigenvector can be obtained by solving $\mathfrak{L}_p^+ L(\zeta) = 0$ and is given by

$$L(\zeta) = n_d'(\zeta) e^{v_d \zeta/D}. \tag{72}$$

To compute Δv, we multiply both sides of Eq. (70) by $L(\zeta)$ and integrate over ζ. Since $L\mathfrak{L}_p$ is equivalent to zero, the terms on the left hand side cancel, and we obtain that

$$\Delta v = -\frac{\int_{-\infty}^{+\infty} e^{v_d\zeta/D} n_d'(\zeta) n_d(\zeta) \Delta r(n_d(\zeta)) d\zeta}{\int_{-\infty}^{+\infty} e^{v_d\zeta/D} [n_d'(\zeta)]^2 d\zeta}, \tag{73}$$

which is the same result as in Refs. [31, 86]. For the specific form of Δr due to a cutoff, this formula simplifies to

$$\Delta v = \frac{\int_{\zeta_c}^{+\infty} e^{v_d\zeta/D} n_d'(\zeta) n_d(\zeta) r(n_d(\zeta)) d\zeta}{\int_{-\infty}^{+\infty} e^{v_d\zeta/D} [n_d'(\zeta)]^2 d\zeta}, \tag{74}$$

which is the main result of this section.

The solvability condition that we used to compute Δv has a simple interpretation: All perturbations that act along the zero eigenmode of $\mathfrak{L}_p$ accumulate unattenuated and contribute to the translation of the front, i.e. to Δv rather than to Δn. This fact can be seen more clearly from the time-dependent perturbation theory that we use in sections VI and VII to compute the diffusion constant of the front and the corrections to the wave velocity due to demographic noise rather than a cutoff.

VI. Diffusion constant of the front, D_f

While a cutoff can account for changes in the velocity due to demographic fluctuations, it cannot capture the fluctuations in the front shape and position. In this section, we describe the stochastic properties of the front using an extension of the perturbation theory developed above. Originally developed in Refs. [83] and [54, 85], this approach shows that the position of the front performs a random walk that can be described by an effective diffusion constant. Following Ref. [54], we derive the general formula for D_f given by Eq. (4) of the main text and evaluate it explicitly for the exactly solvable models introduced in the beginning of the SI. The calculations for the exactly solvable models are the only new results in this section.

Perturbation theory for demographic fluctuations
The calculation follows exactly the same steps as in section V. We begin by restating Eq. (57) in a more convenient form:

$$\frac{\partial \rho}{\partial t} = D\frac{\partial^2 \rho}{\partial x^2} + r(\rho)\rho + \frac{1}{\sqrt{N}}\Gamma(\rho)\eta \tag{75}$$

where Γ denotes the strength of the noise term

$$\Gamma(\rho) = \sqrt{\gamma_n(\rho)\rho}. \tag{76}$$

In this section, we use the normalized population density $\rho = n/N$ instead of n to indicate that the stochastic term is small and scales as $1/\sqrt{N}$. We also introduce a more compact notation for the noise strength Γ to avoid taking explicit derivatives of $\sqrt{\gamma_n(\rho)\rho}$.

We seek the solution of Eq. (75) in the following form

$$\rho(t, x) = \rho_d(x - v_d t - \xi(t)) + \Delta\rho(t, x - v_d t - \xi(t)), \tag{77}$$

where ρ_d is the deterministic stationary solution satisfying Eq. (13), $\xi(t)$ is the shift in the front position due to fluctuations, and $\Delta\rho(t, \zeta)$ accounts for the effect of the perturbation on the front shape. Because the perturbation, $\Gamma\eta$, is time dependent, $\Delta\rho$ explicitly depends on time in addition to the dependence on t through the comoving coordinate $\zeta = x - v_d t - \xi(t)$.

To the first order in perturbation theory, there are no terms due to the special rules of Itô calculus, and we obtain the following expansions for the deterministic terms in Eq. (75)

$$\frac{\partial \rho}{\partial t} \approx -v_d \rho_d' - \rho_d' \xi' - v_d \frac{\partial \Delta\rho}{\partial \zeta} + \frac{\partial \Delta\rho}{\partial t}, \tag{78}$$

the diffusion term as

$$D\frac{\partial^2 \rho}{\partial x^2} \approx D\rho_d'' + D\frac{\partial^2 \Delta\rho}{\partial \zeta^2}, \tag{79}$$

and the growth term as

$$r(\rho)\rho \approx \rho_d r(\rho_d) + r(\rho_d)\Delta\rho + r'(\rho_d)\rho_d \Delta\rho. \tag{80}$$

For functions with a single argument, primes denote derivatives with respect to that argument.

As before, the zeroth order of the perturbation theory is automatically satisfied, and the first non-trivial equation arises at the first order:

$$\frac{\partial \Delta\rho}{\partial t} - \mathfrak{L}_p \Delta\rho = \rho_d' \xi' + \frac{1}{\sqrt{N}} \Gamma(\rho_d)\eta, \tag{81}$$

where $\mathfrak{L}_p$ is the same as in Eq. (71).

To obtain the equation for ξ, we multiply both sides by $L(\zeta)$, the left eigenvector of $\mathfrak{L}_p$ with zero eigenvalue, and integrating over ζ. The result reads

$$\frac{\partial}{\partial t} \int_{-\infty}^{+\infty} L\Delta\rho d\zeta = \int_{-\infty}^{+\infty} L\rho_d' \xi' d\zeta + \frac{1}{\sqrt{N}} \int_{-\infty}^{+\infty} L\Gamma(\rho_d)\eta d\zeta. \tag{82}$$

We now use the fact that $\int_{-\infty}^{+\infty} L(\zeta)\Delta\rho(\zeta)d\zeta = 0$. The projection of $\Delta\rho$ on L vanishes because translations of ρ_d are excluded from the fluctuations of the front shape and are instead included through $\xi(t)$. Imposing this condition is also necessary for the perturbation theory to be self-consistent. Otherwise, according to Eq. (82), $\int_{-\infty}^{\infty} L\Delta\rho d\zeta$ would perform an unconstrained random walk and grow arbitrarily large, which would violate the assumption that $\Delta\rho$ is small. After imposing $\int_{-\infty}^{+\infty} L\Delta\rho d\zeta = 0$, we obtain

$$\xi'(t) = -\frac{1}{\sqrt{N}} \frac{\int_{-\infty}^{+\infty} L(\zeta)\Gamma(\rho_d(\zeta))\eta(t,\zeta)d\zeta}{\int_{-\infty}^{+\infty} L(\zeta)\rho_d'(\zeta)d\zeta}, \tag{83}$$

From Eq. (83), it immediately follows that

$$\langle \xi'(t) \rangle = 0. \tag{84}$$

Thus, there are no corrections to the wave velocity at this order in the perturbation theory, and the motion of the front position is a random walk.

The deviation between the position of the front relative to the deterministic expectation is given by ξ, which we obtain by integrating Eq. (83):

$$X_{\mathrm{f}} - v_d t = \xi = -\frac{1}{\sqrt{N}} \int_0^t \frac{\int_{-\infty}^{+\infty} L(\zeta)\Gamma(\rho_d(\zeta))\eta(t,\zeta)d\zeta}{\int_{-\infty}^{+\infty} L(\zeta)\rho_d'(\zeta)d\zeta} dt \tag{85}$$

To determine the diffusion constant of front wandering, we evaluate the mean square displacement of the front position:

$$
\begin{aligned}
D_{\rm f} &= \frac{{\rm Var}\{X_{\rm f}^2\}}{2t} = \frac{\langle \xi^2 \rangle}{2t} \\
&= \frac{1}{2tN} \left\langle \int_0^t \frac{\int_{-\infty}^{+\infty} L(\zeta_1)\Gamma(\rho_d(\zeta_1))\eta(t_1,\zeta_1)d\zeta_1}{\int_{-\infty}^{+\infty} L(\zeta)\rho_d'(\zeta)d\zeta} dt_1 \int_0^t \frac{\int_{-\infty}^{+\infty} L(\zeta_2)\Gamma(\rho_d(\zeta_2))\eta(t_2,\zeta_2)d\zeta_2}{\int_{-\infty}^{+\infty} L(\zeta)\rho_d'(\zeta)d\zeta} dt_2 \right\rangle \\
&= \frac{1}{2tN} \frac{\int_0^t dt_1 \int_0^t dt_2 \int_{-\infty}^{+\infty} d\zeta_1 \int_{-\infty}^{+\infty} d\zeta_2 L(\zeta_1)\Gamma(\rho_d(\zeta_1))L(\zeta_2)\Gamma(\rho_d(\zeta_2)) \langle \eta(t_1,\zeta_1)\eta(t_2,\zeta_2)\rangle}{\left(\int_{-\infty}^{+\infty} L(\zeta)\rho_d'(\zeta)d\zeta\right)^2} \\
&= \frac{1}{2N} \frac{\int_{-\infty}^{+\infty} d\zeta L^2(\zeta)\Gamma^2(\rho_d(\zeta))}{\left(\int_{-\infty}^{+\infty} L(\zeta)\rho_d'(\zeta)d\zeta\right)^2},
\end{aligned}
\tag{86}
$$

where we used Eq. (85) to express $\xi(t)$ and Eq. (60) to average over the noise.

Finally, we substitute the expression for $L(\zeta)$ from Eq. (72) and use the explicit form of Γ from Eq. (89) to obtain Eq. (4) from the main text:

$$
D_{\rm f} = \frac{1}{2N} \frac{\int_{-\infty}^{+\infty} [\rho_d'(\zeta)]^2 \rho_d(\zeta)\gamma_n(\rho_d(\zeta)) e^{\frac{2v_d\zeta}{D}} d\zeta}{\left(\int_{-\infty}^{+\infty} [\rho_d'(\zeta)]^2 e^{\frac{v_d\zeta}{D}} d\zeta\right)^2},
\tag{87}
$$

which was originally derived in Refs. [83] and [54, 85].

Perturbation theory for migration fluctuations
The analysis that we performed to compute $D_{\rm f}$ due to demographic noise can be easily generalized to account for the noise in migration; see the discussion below Eq. (61). This was fist done in Ref. [54] that extend the perturbation theory to the following equation

$$
\frac{\partial \rho}{\partial t} = D\frac{\partial^2 \rho}{\partial x^2} + r(\rho)\rho + \frac{1}{\sqrt{N}}\Gamma(\rho)\eta + \frac{1}{\sqrt{N}}\frac{\partial(\Gamma_m(\rho)\chi)}{\partial \zeta},
\tag{88}
$$

where Γ_m denotes the strength of the migration fluctuations

$$
\Gamma_m(\rho) = \sqrt{\gamma_m(\rho)\rho},
\tag{89}
$$

and $\chi = \sum_i \sqrt{n_i}\chi_i/\sqrt{n}$ is a unit-strength, delta-correlated, Gaussian noise that enters the equation for the total population density $n = \sum_i n_i$ of all neutral genotypes. Note that, for the standard diffusion, $\gamma_m(\rho) = \text{const}$, but we allow the dependence on ρ because it does not affect the calculation below.

The solution for ξ acquires an additional term due to migration fluctuations:

$$\xi = -\frac{1}{\sqrt{N}}\int_0^t \frac{\int_{-\infty}^{+\infty} L(\zeta)\Gamma(\rho_d(\zeta))\eta(t,\zeta)d\zeta}{\int_{-\infty}^{+\infty} L(\zeta)\rho_d'(\zeta)d\zeta}dt - \frac{1}{\sqrt{N}}\int_0^t \frac{\int_{-\infty}^{+\infty} L(\zeta)\frac{\partial\Gamma_m(\rho_d(\zeta))\chi(t,\zeta)}{\partial\zeta}d\zeta}{\int_{-\infty}^{+\infty} L(\zeta)\rho_d'(\zeta)d\zeta}dt. \tag{90}$$

Because χ and η are uncorrelated, their contributions to D_{f} simply add:

$$D_{\mathrm{f}} = \frac{1}{2N}\frac{\int_{-\infty}^{+\infty}[\rho_d'(\zeta)]^2\rho_d(\zeta)\gamma_n(\rho_d(\zeta))e^{\frac{2v_d\zeta}{D}}d\zeta}{\left(\int_{-\infty}^{+\infty}[\rho_d'(\zeta)]^2 e^{\frac{v_d\zeta}{D}}d\zeta\right)^2} + \frac{1}{2N}\frac{\int_{-\infty}^{+\infty}\gamma_m(\rho_d(\zeta))\rho_d(\zeta)(\rho_d''(\zeta)+\rho_d'(\zeta)v_d/D)^2 e^{\frac{2v_d\zeta}{D}}d\zeta}{\left(\int_{-\infty}^{+\infty}[\rho_d'(\zeta)]^2 e^{\frac{v_d\zeta}{D}}d\zeta\right)^2}. \tag{91}$$

The higher order derivatives of ρ appear in the second term due to the integration by parts that is necessary to remove derivatives from the delta function due to $\langle\chi(t_1,\zeta_1)\chi(t_2,\zeta_2)\rangle$. See Ref. [54] for the original derivation and further details.

It is now clear that the qualitative behavior of the two terms in Eq. (91) is the same. Indeed, the denominators are identical, and the integrands in the numerators have the same scaling behavior at the front, where divergences could occur. To see this, one can substitute the asymptotic behavior of the population density, $\rho \sim e^{-k\zeta}$, and confirm that both numerators scale as $e^{-\zeta(3k-2v_d/D)}$. Thus, the transition from fully-pushed to semi-pushed waves leads to the divergence of both integrals, and the scaling exponent α_{D} is the same for both migration and demographic fluctuations. For simplicity, only demographic fluctuations are considered in all other sections of this paper.

Results for exactly solvable models

In the regime of fully-pushed waves, we can evaluate D_{f} explicitly for the exactly solvable models introduced in section I. The details of these calculations are summarized in the subsection on integral evaluation at the end of section III.

For the model specified by Eq. (24), we find that

$$D_{\mathrm{f}} = \frac{3}{16\pi}\frac{\gamma_n^0}{N}\sqrt{\frac{2D}{g_0}}\frac{(1-4\rho^*)(3-4\rho^*)}{\rho^*(1-2\rho^*)(1-\rho^*)^2}\tan 2\pi\rho^*, \tag{92}$$

when $\gamma_m = 0$ and $\gamma_n(n) = \gamma_n^0 = \text{const}$.

Note that the choice of $\gamma_n(n)$ is not specified by the deterministic model of population growth and needs to be determined either from the microscopic dynamics or from empirical observations. For models formulated in terms of independent birth and death rates, $\gamma_n(n)$ is a constant on the order of $1/\tau$, where τ is the generation time. However, different γ_n are possible. For example, our simulations that are based on the Wright-Fisher model have $\gamma_f = \gamma_n^0(1-n/N)$, and the corresponding theoretical prediction for D_{f} reads

$$D_{\mathrm{f}} = \frac{3}{20\pi}\frac{\gamma_n^0}{N}\sqrt{\frac{2D}{g_0}}\frac{(1-4\rho^*)(3-4\rho^*)}{\rho^*(1-\rho^*)(1-2\rho^*)}\tan 2\pi\rho^*. \tag{93}$$

For the same model of an Allee effect, the contribution of the noise due to migration with $\gamma_m = \gamma_m^0 = \text{const}$ is given by

$$D_{\mathrm{f}} = \frac{\sqrt{2}}{40\pi}\frac{\gamma_m^0}{N}\sqrt{\frac{g_0}{D}}\frac{(1+\rho^*)(1+7\rho^*)(1-4\rho^*)(3-4\rho^*)}{\rho^*(1-2\rho^*)(1-\rho^*)^2}\tan 2\pi\rho^* \tag{94}$$

assuming $\gamma_n = 0$.

For completeness, we also provide the results for other models and different choices of γ_n and γ_m.

$$D_{\mathrm{f}} = \frac{3}{16\pi}\frac{\gamma_n^0}{N}\sqrt{\frac{2D}{r_0 B}}\frac{B^2(B+4)(3B+4)}{(B+2)(B+1)^2}\tan\frac{2\pi}{B} \tag{95}$$

for the model of cooperative growth defined in the main text with $\gamma_n = \gamma_n^0$ and $\gamma_m = 0$.

$$D_{\mathrm{f}} = \frac{3}{20\pi}\frac{\gamma_n^0}{N}\sqrt{\frac{2D}{r_0 B}}\frac{B(B+4)(3B+4)}{(B+2)(B+1)}\tan\frac{2\pi}{B} \tag{96}$$

for the model of cooperative growth defined in the main text with $\gamma_n = \gamma_n^0(1-n/N)$ and $\gamma_m = 0$.

$$D_{\mathrm{f}} = \frac{\sqrt{2}}{40\pi}\frac{\gamma_m^0}{N}\sqrt{\frac{r_0 B}{D}}\frac{(B-1)(B-7)(B+4)(3B+4)}{(B+2)(B+1)^2}\tan\frac{2\pi}{B} \tag{97}$$

for the model of cooperative growth defined in the main text with $\gamma_m = \gamma_m^0 = \text{const}$ and $\gamma_n = 0$.

VII. Correction to velocity due to demographic fluctuations

In this section, we compute the correction to the wave velocity directly from the stochastic formulation in Eq. (75) instead of relying on a growth-rate cutoff at low densities. Our main finding is that, for pushed waves, the scaling of Δv with N coincides[3] with that of front diffusion constant D_{f} and the rate of diversity loss Λ. Note that this result cannot be obtain from the cutoff-based calculation of Δv without knowing the correct dependence of n_c on v/v_{F}. Thus, the calculation of Δv in the stochastic model provides an additional insight in the dynamics of fluctuating fronts. To the best of our knowledge, the results presented in this section are new.

[3]The scaling behavior of Δv is slightly different for pulled waves because $\Delta v \sim \ln^{-2} N$; see Ref. [1, 35, 36, 82].

Because the first order correction to v is zero (See Eq. (83)), we proceed to the second order in perturbation theory. In this calculation, it is convenient to distinguish the contributions to ζ and $\Delta\rho$ that come from the different orders of the perturbative expansion:

$$\begin{aligned}\zeta &= x - v_d t - \xi_{(1)}(t) - \xi_{(2)}(t),\\ \rho(t,\zeta) &= \rho_d(\zeta) + \Delta\rho_{(1)}(t,\zeta) + \Delta\rho_{(2)}(t,\zeta),\end{aligned} \tag{98}$$

where the order is indicated by a subscript in brackets. For fully-pushed waves, we expect that the first order corrections $\xi_{(1)}(t)$ and $\Delta\rho_{(1)}(t,\zeta)$ scale as $1/\sqrt{N}$, and the second order corrections $\xi_{(2)}(t)$ and $\Delta\rho_{(2)}(t,\zeta)$ scale as $1/N$. Therefore, we expand all terms in Eq. (75) up to order $1/N$.

For $\partial\rho/\partial t$, we obtain

$$\frac{\partial\rho}{\partial t} \approx \rho_d'(-v_d - \xi_{(1)}' - \xi_{(2)}') + \frac{1}{2}\rho_d''\left\langle\frac{(d\xi_{(1)})^2}{dt}\right\rangle + \frac{\partial\Delta\rho_{(1)}}{\partial\zeta}(-v_d - \xi_{(1)}') + \frac{\partial\Delta\rho_{(1)}}{\partial t} + \frac{\partial\Delta\rho_{(2)}}{\partial\zeta}(-v_d) + \frac{\partial\Delta\rho_{(2)}}{\partial t}, \tag{99}$$

where $\rho_d''/2\left\langle(d\xi_{(1)})^2/dt\right\rangle$ arises due to the Itô formula of stochastic calculus, which prescribes how to compute derivatives of nonlinear functions; see Refs. [74, 75, 80, 87]. The unusual derivative $\left\langle(d\xi_{(1)})^2/dt\right\rangle$ is non-zero because the displacement of a random walk grows as $\sqrt{dt}$. Using Eqs. (83) and (86), we express this derivative in terms of D_{f}, which we know to the order $1/N$ from the first order of the perturbation theory:

$$\left\langle(d\xi_{(1)})^2/dt\right\rangle = 2D_{\mathrm{f}}. \tag{100}$$

The expansion of other terms is more straightforward and does not involve any additional terms due to the special rules of Itô calculus:

$$\begin{aligned}&D\frac{\partial^2\rho}{\partial x^2} \approx D\rho_d'' + D\frac{\partial^2\Delta\rho_{(1)}}{\partial\zeta^2} + D\frac{\partial^2\Delta\rho_{(2)}}{\partial\zeta^2},\\ &r(\rho)\rho \approx [r(\rho)\rho]|_{\rho=\rho_d} + [r(\rho)\rho]'|_{\rho=\rho_d}\Delta\rho_{(1)} + \frac{1}{2}[r(\rho)\rho]''|_{\rho=\rho_d}(\Delta\rho_{(1)})^2 + [r(\rho)\rho]'|_{\rho=\rho_d}\Delta\rho_{(2)},\\ &\Gamma(\rho) \approx \Gamma(\rho_d) + \Gamma'(\rho_d)\Delta\rho_{(1)},\end{aligned} \tag{101}$$

where we kept only the terms that scale at most as $1/N$ and used $|_{\rho=\rho_d}$ to indicate that the expression to the left is evaluated at $\rho = \rho_d$. Upon choosing $\Delta\rho_{(1)}$ and $\xi_{(1)}$ that satisfy the first order equation, i.e. Eq. (81), we obtain the following equation for $\Delta\rho_{(2)}$ and $\xi_{(2)}$:

$$\frac{\partial\Delta\rho_{(2)}}{\partial t} - \mathfrak{L}_p\Delta\rho_{(2)} = -D_{\mathrm{f}}\rho_d'' + \frac{\partial\Delta\rho_{(1)}}{\partial\zeta}\xi_{(1)}' + \frac{1}{2}[r''(\rho_d)\rho_d + 2r'(\rho_d)](\Delta\rho_{(1)})^2 + \frac{1}{\sqrt{N}}\Gamma'(\rho_d)\Delta\rho_{(1)}\eta + \rho_d'\xi_{(2)}'. \tag{102}$$

The value of $\xi'_{(2)}$ needs to be chosen to satisfy the solvability condition, which we obtain by multiplying both sides of Eq. (102) by $L(\zeta)$, integrating over ζ, and requiring that $\Delta\rho_{(2)}$ has zero projection on $L(\zeta)$. The result reads

$$\xi'_{(2)}(t) = -\frac{\int_{-\infty}^{+\infty} L(\zeta)\left[-D_{\rm f}\rho''_d + \frac{\partial\Delta\rho_{(1)}}{\partial\zeta}\xi'_{(1)} + \frac{1}{2}[r''(\rho_d)\rho_d + 2r'(\rho_d)](\Delta\rho_{(1)})^2 + \frac{1}{\sqrt{N}}\Gamma'(\rho_d)\Delta\rho_{(1)}\eta\right]d\zeta}{\int_{-\infty}^{+\infty} L(\zeta)\rho'_d(\zeta)d\zeta}. \tag{103}$$

The correction to the velocity, Δv, can now be obtained by averaging Eq. (103) over η and substituting the explicit expression for $L(\zeta)$ from Eq. (72):

$$\Delta v = \langle\xi'_{(2)}\rangle = D_{\rm f}\frac{\int_{-\infty}^{+\infty} e^{v_d\zeta/D}\rho'_d(\zeta)\rho''_d(\zeta)d\zeta}{\int_{-\infty}^{+\infty} e^{v_d\zeta/D}[\rho'_d(\zeta)]^2 d\zeta} - \frac{\int_{-\infty}^{+\infty} e^{v_d\zeta/D}\rho'_d(\zeta)\frac{1}{2}[r''(\rho_d)\rho_d + 2r'(\rho_d)]\langle[\Delta\rho_{(1)}(t,\zeta)]^2\rangle d\zeta}{\int_{-\infty}^{+\infty} e^{v_d\zeta/D}[\rho'_d(\zeta)]^2 d\zeta}. \tag{104}$$

Note that $\langle\Delta\rho_{(1)}\eta\rangle = 0$ and, therefore, $\langle\Delta\rho_{(1)}\xi'\rangle = 0$ because $\Delta\rho_{(1)}(t,\zeta)$ depends only on $\eta(\tilde{t},\zeta)$ with $\tilde{t} < t$ and $\langle\eta(\tilde{t},\zeta)\eta(t,\zeta)\rangle = 0$.[4] The first term could be further simplified through integration by parts in the numerator, assuming that the integrals converge:

$$\Delta v = -v_d\frac{D_{\rm f}}{2D} - \frac{\int_{-\infty}^{+\infty} e^{v_d\zeta/D}\rho'_d(\zeta)\frac{1}{2}[r''(\rho_d)\rho_d + 2r'(\rho_d)]\langle[\Delta\rho_{(1)}(t,\zeta)]^2\rangle d\zeta}{\int_{-\infty}^{+\infty} e^{v_d\zeta/D}[\rho'_d(\zeta)]^2 d\zeta}. \tag{105}$$

To complete the calculation of Δv, we need to obtain $\Delta\rho_{(1)}$ by solving Eq. (81). Before performing this calculation, let us state the main findings and discuss their implications. For fully-pushed waves, we find that $\Delta\rho_{(1)} \sim 1/\sqrt{N}$, and all integrals in Eq. (105) converge. Thus, $\Delta v \sim 1/N$ in this regime. For semi-pushed waves, one needs to apply a cutoff at large ζ to ensure convergence. We show that the divergence of the last term in Eq. (105) does not exceed that of $D_{\rm f}$. Thus, the leading behavior is controlled by the first term, and the scaling of Δv coincides with that of $D_{\rm f}$. The scaling behavior of Δv and $D_{\rm f}$ is slightly different for pulled waves: $\Delta v \sim \ln^{-2} N$ and $D_{\rm f} \sim \ln^{-3} N$; see Ref. [1, 35, 36, 82].

The calculation of $\Delta\rho_{(1)}$ can be simplified by transforming Eq. (81) into a Hermitian form. This is accomplished by the following change of variables that eliminates the term linear in $\partial/\partial\zeta$ from $\mathfrak{L}_p$:

$$\Delta\rho_{(1)}(t,\zeta) = e^{-\frac{v_d\zeta}{2D}}\Psi(t,\zeta). \tag{106}$$

Equation (81) then takes the following form

[4] This simplification is specific to the Itô calculus and does not occur in Stratonovich's formulation. The results of course do not depend on the type of calculus used as long as all calculations are carried using the same calculus and the initial problem statement is correctly formulated. In population dynamics, demographic fluctuations affect only future generations, so Itô's formulation appears naturally.

$$\frac{\partial \Psi}{\partial t} - \mathfrak{H}_p \Psi = e^{\frac{v_d \zeta}{2D}} \left[\rho_d' \xi_{(1)}' + \frac{1}{\sqrt{N}} \Gamma(\rho_d) \eta \right], \tag{107}$$

where $\mathfrak{H}_p$ is a Hermitian operator:

$$\mathfrak{H}_p = D \frac{\partial^2}{\partial \zeta^2} - \frac{v_d^2}{4D} + r(\rho_d) + r'(\rho_d)\rho_d. \tag{108}$$

We solve Eq. (107) using the method of separation of variables. Let us denote the eigenvalues and normalized eigenvectors of $\mathfrak{H}_p$ by λ_l and $\mathfrak{h}_l$ respectively. The index l labels both discrete and continuous parts of the spectrum of $\mathfrak{H}_p$ such that λ_l are in the decreasing order; $l = 0$ corresponds to the zero mode. In the basis of $\mathfrak{h}_l$, we express Ψ as:

$$\Psi(t, \zeta) = \sum_l a_l(t) \mathfrak{h}_l(\zeta). \tag{109}$$

The unknown coefficients $a_l(t)$ are determined by projecting Eq. (107) on $\mathfrak{h}_l$:

$$\frac{da_l}{dt} - \lambda_l a_l = \int_{-\infty}^{+\infty} e^{\frac{v_d \zeta}{2D}} \left[\rho_d' \xi_{(1)}' + \frac{1}{\sqrt{N}} \Gamma(\rho_d) \eta \right] \mathfrak{h}_l d\zeta. \tag{110}$$

and then solving these linear equations:

$$a_l(t) = \int_{-\infty}^{t} e^{\lambda_l (t - \tilde{t})} \left\{ \int_{-\infty}^{+\infty} e^{\frac{v_d \zeta}{2D}} \left[\rho_d'(\zeta) \xi_{(1)}'(\tilde{t}) + \frac{1}{\sqrt{N}} \Gamma(\rho_d(\zeta)) \eta(\tilde{t}, \zeta) \right] \mathfrak{h}_l(\zeta) d\zeta \right\} d\tilde{t}. \tag{111}$$

Here, we assumed that the front has been propagating for a very long time and, therefore, set the lower integration limit of the integral over $\tilde{t}$ to $-\infty$. The next step is to substitute the solution for $\xi_{(1)}'$ from Eq. (83):

$$a_l(t) = \frac{1}{\sqrt{N}} \int_{-\infty}^{t} e^{\lambda_l (t - \tilde{t})} \left\{ \int_{-\infty}^{+\infty} \left[e^{\frac{v_d \zeta}{2D}} \Gamma(\rho_d(\zeta)) \eta(\tilde{t}, \zeta) - \mathfrak{h}_0(\zeta) \int_{-\infty}^{+\infty} e^{\frac{v_d \tilde{\zeta}}{2D}} \Gamma(\rho_d(\tilde{\zeta})) \eta(\tilde{t}, \tilde{\zeta}) \mathfrak{h}_0(\tilde{\zeta}) d\tilde{\zeta} \right] \mathfrak{h}_l(\zeta) d\zeta \right\} d\tilde{t}, \tag{112}$$

where we used the fact that

$$\mathfrak{h}_0(\zeta) = \frac{e^{\frac{v_d \zeta}{2D}} \rho_d'(\zeta)}{\sqrt{\int_{-\infty}^{+\infty} e^{\frac{v_d \zeta}{D}} [\rho_d'(\zeta)]^2 d\zeta}}, \tag{113}$$

and, therefore,

$$\xi'_{(1)}(t) = -\frac{1}{\sqrt{N}} \frac{\int_{-\infty}^{+\infty} e^{\frac{v_d \zeta}{2D}} \Gamma(\rho_d(\zeta)) \eta(t,\zeta) \mathfrak{h}_0(\zeta) d\zeta}{\sqrt{\int_{-\infty}^{+\infty} e^{\frac{v_d \zeta}{D}} [\rho'_d(\zeta)]^2 d\zeta}}. \tag{114}$$

Equation (112) is further simplified by carrying out the integration over ζ in the last term and using the orthogonality of $\mathfrak{h}_0$ and $\mathfrak{h}_l$ for $l > 0$:

$$a_l(t) = \frac{1-\delta_{0l}}{\sqrt{N}} \int_{-\infty}^{t} e^{\lambda_l (t-\tilde{t})} \left\{ \int_{-\infty}^{+\infty} e^{\frac{v_d \zeta}{2D}} \Gamma(\rho_d(\zeta)) \eta(\tilde{t},\zeta) \mathfrak{h}_l(\zeta) d\zeta \right\} d\tilde{t}, \tag{115}$$

Note that, $a_0 = 0$ consistent with the solvability condition that $\Delta\rho_{(1)}$ has a vanishing projection on the zero mode.

With the solution for $\Delta\rho_{(1)}$ at hand, we proceed to calculate the average $[\Delta\rho_{(1)}]^2$ that enters Eq. (105):

$$\begin{aligned} \langle [\Delta\rho_{(1)}(t,\zeta)]^2 \rangle &= \sum_{l_1>0} \sum_{l_2>0} \mathfrak{h}_{l_1}(\zeta) \mathfrak{h}_{l_2}(\zeta) \langle a_{l_1}(t) a_{l_2}(t) \rangle \\ &= \frac{1}{N} \sum_{l_1>0} \sum_{l_2>0} \mathfrak{h}_{l_1}(\zeta) \mathfrak{h}_{l_2}(\zeta) \frac{-1}{\lambda_{l_1} + \lambda_{l_2}} \int_{-\infty}^{+\infty} e^{\frac{v_d \zeta}{D}} \Gamma^2(\rho_d(\zeta)) \mathfrak{h}_{l_1}(\zeta) \mathfrak{h}_{l_2}(\zeta) d\zeta. \end{aligned} \tag{116}$$

Upon substituting this result into Eq. (105), we obtain

$$\begin{aligned} \Delta v = &-v_d \frac{D_\mathrm{f}}{2D} - \frac{1}{N} \frac{1}{\int_{-\infty}^{+\infty} e^{v_d \zeta/D} [\rho'_d(\zeta)]^2 d\zeta} \times \\ &\sum_{l_1>0} \sum_{l_2>0} \frac{\left[\int_{-\infty}^{+\infty} e^{\frac{v_d \zeta}{D}} \gamma_n(\rho_d(\zeta)) \rho_d(\zeta) \mathfrak{h}_{l_1}(\zeta) \mathfrak{h}_{l_2}(\zeta) d\zeta\right] \left[\int_{-\infty}^{+\infty} e^{v_d \zeta/D} \rho'_d(\zeta) \frac{1}{2} [r''(\rho_d)\rho_d + 2r'(\rho_d)] \mathfrak{h}_{l_1}(\zeta) \mathfrak{h}_{l_2}(\zeta) d\zeta\right]}{-(\lambda_{l_1} + \lambda_{l_2})}. \end{aligned} \tag{117}$$

Since the eigenvectors $\mathfrak{h}_l$ decay at least as fast as $\mathfrak{h}_0 \sim e^{\frac{v_d \zeta}{2D}} \rho'_d(\zeta)$ as $\zeta \to +\infty$, all the integrands in Eq. (117) decay faster than $e^{-(3k-2v_d/D)\zeta}$. For fully-pushed waves, all the integrals converge, and the correction to the velocity scales as $1/N$. For semi-pushed waves, the term with D_f shows the fastest divergence with the cutoff and, therefore, determines the scaling of Δv with N.

VIII. Rate of diversity loss, Λ

In this section, we describe how genetic diversity is lost during a range expansion and provide the derivation of Eq. (4) from the main text, which was first derived in Ref. [37]. For simplicity, we

consider an expansion that started with two neutral genotypes present throughout the population and determine how the probability to sample two different genotypes at the front decreases with time. The calculation of Λ is based on the perturbation theory in $1/N$ and relies on a mean-field assumption that $n(t,x)$ can be approximated by $\langle n(t,x)\rangle$. This analysis is asymptotically exact for fully-pushed waves and could be extended to semi-pushed and pulled waves by applying a cutoff at large ζ as we show in section IX. The current section contain no new results except for the calculation of Λ in exactly solvable models of fully-pushed waves.

Forward-in-time analysis of the decay of heterozygosity
We quantify the genetic diversity in the population by the average heterozygosity:

$$H(t,\zeta_1,\zeta_2) = \langle f(t,\zeta_1)[1-f(t,\zeta_2)] + [1-f(t,\zeta_1)]f(t,\zeta_2)\rangle, \tag{118}$$

which is the probability to sample two different genotypes at positions ζ_1 and ζ_2 in the comoving reference frame at time t. Here, f denotes the frequency of one the two genotypes; the frequency of the other genotype is $1-f$.

To obtain a closed equation for the dynamics of H, we assume that $n(t,\zeta)$ is given by its non-fluctuating stationary limit, $n(\zeta)$, from Eq. (13). Then, we differentiate Eq. (118) with respect to time and use Eq. (58) to eliminate the time derivatives of f. The result reads

$$\frac{\partial H}{\partial t} = \left(\mathcal{L}_{\zeta_1} + \mathcal{L}_{\zeta_2}\right) H - \delta(\zeta_1-\zeta_2)\frac{\gamma_f(n)}{n} H, \tag{119}$$

where

$$\mathcal{L}_\zeta = D\frac{\partial^2}{\partial\zeta^2} + v\frac{\partial}{\partial\zeta} + 2D\frac{\partial \ln n}{\partial\zeta}\frac{\partial}{\partial\zeta}. \tag{120}$$

we note that the first term in Eq. (119) follows from the rules of regular calculus, but the last term arises due to the Itô formula of stochastic calculus, which prescribes how to compute derivatives of nonlinear functions; see Refs. [74, 75, 80, 87]. This last term encapsulates the effect of genetic drift and ensures that genetic diversity decays to zero due to the fixation of one of the genotypes.

Since H obeys a linear equation, it will decay to zero exponentially in time with the decay rate given by the solution of the following eigen-value problem:

$$-\Lambda H = \left(\mathcal{L}_{\zeta_1} + \mathcal{L}_{\zeta_2}\right) H - \delta(\zeta_1-\zeta_2)\frac{\gamma_f(n)}{n} H, \tag{121}$$

where we seek the smallest Λ or alternatively the largest eigenvalue of the operator on the right hand side.

We compute Λ perturbatively by treating $1/N$ as a small parameter. To the zeroth order, we can neglect the last term in Eq. (121) because it scales as $1/N$. Without the sink term, Eq. (119) admits a constant stationary solution ($H(t,\zeta) = \text{const}$), so the smallest decay rate is zero. Thus, the zeroth order solution of Eq. (121) reads

$$\begin{aligned} \Lambda &= 0, \\ R(\zeta_1, \zeta_2) &= 1, \end{aligned} \tag{122}$$

where we used the notation from quantum mechanics for the right eigenvector of $\mathcal{L}_{\zeta_1} + \mathcal{L}_{\zeta_2}$; see Ref. [88, 89]. Because $\mathcal{L}_\zeta$ contains terms linear in $\frac{\partial}{\partial\zeta}$, the operator in Eq. (121) is not Hermitian. Therefore, also need $\langle L|$, the left eigenvector of $\mathcal{L}_{\zeta_1} + \mathcal{L}_{\zeta_2}$, to compute the first order correction. It is not difficult to guess $\langle L|$ because it corresponds to the right eigenvector of the adjoint operator, and we already obtained the stationary distribution for $\partial S/\partial\tau = L_\zeta^+ S$ when we discussed the patterns of ancestry. Since $\mathcal{L}_{\zeta_1}$ and $\mathcal{L}_{\zeta_2}$ act on different variables, the sought-after eigen function is the product of the eigen functions of these two operators:

$$L(\zeta_1, \zeta_2) = n^2(\zeta_1) e^{\zeta_1 v/D} n^2(\zeta_2) e^{\zeta_2 v/D}. \tag{123}$$

The first order correction to Λ is given by the standard formula [74, 75]:

$$\Lambda = \frac{\int_{-\infty}^{+\infty} d\zeta_1 \int_{-\infty}^{+\infty} d\zeta_2 L(\zeta_1, \zeta_2) \delta(\zeta_1 - \zeta_2) \frac{\gamma_f(n)}{n} R(\zeta_1, \zeta_2)}{\int_{-\infty}^{+\infty} d\zeta_1 \int_{-\infty}^{+\infty} d\zeta_2 L(\zeta_1, \zeta_2) R(\zeta_1, \zeta_2)}. \tag{124}$$

We now use the expressions of $L(\zeta_1, \zeta_2)$ and $R(\zeta_1, \zeta_2)$ from Eqs. (122) and (123) and obtain the final result:

$$\Lambda = \frac{\int_{-\infty}^{+\infty} \gamma_f(n(\zeta)) n^3(\zeta) e^{\frac{2v\zeta}{D}} d\zeta}{\left(\int_{-\infty}^{+\infty} n^2(\zeta) e^{\frac{v\zeta}{D}} d\zeta\right)^2}. \tag{125}$$

which becomes identical to Eq. (4) in the main text upon substituting $n = N\rho$. This result was first obtained in Ref. [37].

For fully-pushed waves, all the integrals in Eq. (125) converge and one can obtain the dependence of Λ on model parameters by dimensional analysis. Specifically, each factor of n contributes a factor of N, and each $d\zeta$ contributes a width of the front (the integrands rapidly tend to zero in the bulk and the leading edge). In total, Λ is inversely proportional to the product of N and front width, i.e. to the number of individuals at the front. This result is quite intuitive because, in well-mixed populations, the rate of diversity loss scales as the total population size, and Λ^{-1} is often denoted as an effective population size [90]. Thus, the neutral evolution in a fully-pushed wave could be approximated by that in a well-mixed population consisting of all the organisms at the front. In

contrast, only the very tip of the front drives the evolutionary dynamics in semi-pushed and pulled waves.

Equation (125) also suggests that the deterministic approximation for $n(t,\zeta)$ that we made in Eq. (119) is asymptotically exact for fully-pushed waves. Indeed, the main contribution to the integrals in Eq. (125) comes for the interior regions of the front, where the fluctuations in n are small compared to the mean population density.

Finally, we note that one can avoid using the perturbation theory for non-Hermitian operators to derive Eq. (124). Specifically, one can recast $(\mathcal{L}_{\zeta_1}+\mathcal{L}_{\zeta_1})$ in a Hermitian form by finding a function $\beta(\zeta)$ such that $(\mathcal{L}_{\zeta_1}+\mathcal{L}_{\zeta_1})\beta(\zeta_1)\beta(\zeta_2)\Psi(t,\zeta_1,\zeta_2) = \beta(\zeta_1)\beta(\zeta_2)\mathcal{H}\Psi(t,\zeta_1,\zeta_2)$, where $\mathcal{H}$ is a Hermitian operator, which contains no terms linear in $\partial/\partial\zeta_1$ or $\partial/\partial\zeta_2$. Then, the substitution: $H(t,\zeta_1,\zeta_2) = \beta(\zeta_1)\beta(\zeta_2)\Psi(t,\zeta_1,\zeta_2)$ converts Eq. (121) into a Hermitian eigen-value problem.

The following equations summarize the main steps in this approach:

$$
\begin{aligned}
&\beta(\zeta) = n^{-1}(\zeta)e^{-\frac{\zeta v}{2D}},\\
&\mathcal{H} = D\frac{\partial^2}{\partial\zeta_1^2} - \left(\frac{D}{n(\zeta_1)}\frac{\partial^2 n(\zeta_1)}{\partial\zeta_1^2} + \frac{v}{n(\zeta_1)}\frac{\partial n(\zeta_1)}{\partial\zeta_1} + \frac{v^2}{4D}\right) + D\frac{\partial^2}{\partial\zeta_2^2} - \left(\frac{D}{n(\zeta_2)}\frac{\partial^2 n(\zeta_2)}{\partial\zeta_2^2} + \frac{v}{n(\zeta_2)}\frac{\partial n(\zeta_2)}{\partial\zeta_2} + \frac{v^2}{4D}\right),\\
&h_0(\zeta_1,\zeta_2) = \frac{1}{\beta(\zeta_1)}\frac{1}{\beta(\zeta_2)} = n(\zeta_1)e^{\frac{\zeta_1 v}{2D}}n(\zeta_2)e^{\frac{\zeta_2 v}{2D}},
\end{aligned}
\tag{126}
$$

where $h_0(\zeta)$ is the eigenvector corresponding to the zero eigenvalue. This eigenvector is easily obtained from the reverse transformation from H to Ψ and the fact that $H = \text{const}$ is the right eigenvector of the original operator, $(\mathcal{L}_{\zeta_1}+\mathcal{L}_{\zeta_1})$.

Since the eigenvalues of $\mathcal{H}$ coincide with the eigenvalues of $(\mathcal{L}_{\zeta_1}+\mathcal{L}_{\zeta_1})$, one can compute Λ by the standard formula:

$$
\Lambda = \frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h_0(\zeta_1,\zeta_2)\delta(\zeta_1-\zeta_2)\frac{\gamma_f(n(\zeta_1))}{n(\zeta_1)}h_0(\zeta_1,\zeta_2)d\zeta_1 d\zeta_2}{\int_{-\infty}^{+\infty} h_0^2(\zeta_1,\zeta_2)d\zeta_1 d\zeta_2}. \tag{127}
$$

It is now easy to see that Eqs. (126) and (127) lead to the same expression for Λ as in Eq. (125).

Backward-in-time analysis of lineage coalescence

To complement the forward-in-time analysis, we show how Λ can be computed by tracing ancestral lineages backward in time. One advantage of this approach is that it provides a more intuitive explanation of Eq. (125). The discussion of this approach closely follows Ref. [37].

We motivate the backward-in-time approach by considering how H can be estimated from its definition as the probability to sample two different genotypes. To determine whether the genotypes are different, we trace their ancestral lineages backward in time and observe that only two outcomes are possible: Either the lineages never interact with each other until they hit the initial conditions or

the lineages coalesce, i.e. converge on the same ancestor, at some point during the range expansion. In the former case, the probability to be different is determined by the initial heterozygosity of the population. In the latter case, the probability to be different is zero because we do not allow mutations. Thus, $H(t, \zeta_1^d, \zeta_2^d)$ is intimately related to the probability $S^{(2)}$ that two lineages sampled at time t at positions ζ_1^d and ζ_2^d have not coalesced up to time τ into the past and were present at ζ_1^a and ζ_2^a at time $t-\tau$; the superscripts distinguish between the positions of the descendants and the ancestors. To simplify the notation, we suppress descendent-related variables, drop the subscripts, and write this probability as $S^{(2)}(\tau, \zeta_1, \zeta_2)$. We keep the superscript to distinguish $S^{(2)}$ from S, which denotes the position of a single ancestral lineage.

The dynamical equation for $S^{(2)}$ can be derived either from the forward-in-time formulation for H or directly from the dynamics of ancestral lineages. The result reads

$$\frac{\partial S^{(2)}}{\partial \tau} = \left(\mathcal{L}^+_{\zeta_1} + \mathcal{L}^+_{\zeta_2}\right) S^{(2)} - \delta(\zeta_1 - \zeta_2)\frac{\gamma_f(n)}{n} S^{(2)}, \tag{128}$$

where the first term describes the motion of the two ancestral lineages, and the last term accounts for the lineage coalescence. As expected, the rate of coalescence events is inversely proportional to the local effective population size n/γ_f; see [57, 58, 90]. Because the linear operators on the right hand side of Eqs. (128) and (119) are adjoint to each other, their eigenvalues coincide. Therefore, the temporal decay of $S^{(2)}$ is exponential in τ with the decay rate equal to Λ.

The expression for Λ that we obtained previously (see Eq. (125)) is much easier to interpret in the backward-in-time formulation. To show this, let us rewrite Eq. (125) as

$$\Lambda = \int_{-\infty}^{+\infty} \frac{\gamma_f(n(\zeta))}{n(\zeta)} S^2(\zeta) d\zeta, \tag{129}$$

where we used Eq. (38) to express Λ in terms of $S(\zeta)$, the stationary distribution of the location of a single ancestral lineage. We can now see that the effective coalescence rate, Λ, is given by the sum of the local coalescence rates, γ_f/n, weighted by the probability that two lineages are present at the same location, S^2. Thus, the first order perturbation theory is equivalent to assuming that the positions of the two ancestral lineages are uncorrelated with each other and distributed according to their stationary distribution $S(\zeta)$.

The last results also clarifies the difference between pulled, semi-pushed, and fully-pushed waves. For pulled waves, $S(\zeta)$ is peaked at the leading edge and, since the coalescent rate peaks at the same location, the neutral evolution is driven by the very tip of the front. In semi-pushed waves, $S(\zeta)$ is peaked in the interior of the front, but the $1/n$ increase in the coalescence rates at the front is sufficiently strong to keep all coalescent events at the front edge. Finally, in fully-pushed waves, the decay of $S^2(\zeta)$ at the front is stronger than the increase in the coalescence rates, and most coalescence events occur in the interior of the front. Thus, the focus of diversity is located in the interior of the front in fully-pushed waves, but at the front edge in pulled and semi-pushed waves.

Explicit results for Λ in exactly solvable models and connection

In the regime of fully-pushed waves, we can evaluate Λ explicitly for the exactly solvable models

introduced in section I. Specifically, we find that

$$\Lambda = \frac{\gamma_f^0}{4\pi N}\sqrt{\frac{g_0}{2D}}\frac{1-4\rho^*}{\rho^*}\tan 2\pi\rho^* \tag{130}$$

for the model specified by Eq. (24) with $\gamma_f(n)$ that does not depend on n and is equal to γ_f^0. Note that the choice of $\gamma_f(n)$ is not specified by the deterministic model of population growth and needs to be determined from the microscopic dynamics, from phenomenological considerations, or empirically. For most commonly used models, $\gamma_f(n)$ is a constant. In our simulations, this constant is $1/(a\tau)$, where τ is the generation time and a is the spatial scale over which genetic drift is correlated. However, different γ_f are possible. For example, $\gamma_f = \gamma_f^0(1-n/N)$ could be appropriate for models that allow no births or deaths once the population has reached the carrying capacity. In such models $\gamma_f(N) = 0$, and genetic drift operates only at the front.

For completeness, we also provide the results for other models and different choices of $\gamma_f(n)$:

$$\Lambda = \frac{\gamma_f^0}{6\pi N}\sqrt{\frac{g_0}{2D}}\frac{(1-2\rho^*)(1-4\rho^*)}{\rho^*}\tan 2\pi\rho^* \tag{131}$$

for the model specified by Eq. (24) with $\gamma_f = \gamma_f^0(1-n/N)$,

$$\Lambda = \frac{\gamma_f^0}{4\pi N}\sqrt{\frac{r_0 B}{2D}}(B+4)\tan\frac{2\pi}{B} \tag{132}$$

for the model of cooperative growth defined in the main text with $\gamma_f = \gamma_f^0$;

$$\Lambda = \frac{\gamma_f^0}{6\pi N}\sqrt{\frac{r_0 B}{2D}}\frac{(B+2)(B+4)}{B}\tan\frac{2\pi}{B} \tag{133}$$

for the model of cooperative growth defined in the main text with $\gamma_f = \gamma_f^0(1-n/N)$.

IX. Cutoffs for deterministic and fluctuating fronts

The integrands that appear in the expressions for Δv, $D_{\rm f}$, and Λ diverge near the front edge in pulled and semi-pushed waves. Since there are no organisms sufficiently far ahead of the front, these divergences are technical artifacts that do not represent the actual dynamics of the traveling wave. For example, in our calculation of Λ, the divergence appears because we approximate the wave front by the stationary, deterministic profile, $n(\zeta)$, from Eq. (13). In this section, we show how to remove these divergences by applying a cutoff at large ζ. The value of the cutoff, ζ_c, scales as $\ln(N)/q$ for fluctuating fronts, but as $\ln(N)/k$ for deterministic fronts with $\gamma_n = 0$; k and q are given in Eq. (19). The derivation of $\zeta_c \sim \ln(N)/q$ is the main new result in this section.

Cutoff for deterministic fronts
A cutoff for the growth rate was first introduced in the context of pulled waves [1]. The primary motivation for the cutoff was to compute the corrections to the wave velocity and to resolve the velocity-selection problem, i.e. to explain why the simulations of discrete entities never exhibit waves with velocities greater than v_{F} even though such solutions are possible in the continuum limit.

The naive argument for a cutoff is that there should be no growth in areas where the average number of individuals falls below one per site in lattice models or one per typical dispersal distance in models with continuous space. We denote the relevant spatial scale, i.e. the distance between lattice sites or the dispersal distance, by a, so the cutoff density is $1/a$. Since, at such low densities, the front shape is well approximated by the asymptotic solution $n \sim Ne^{-k\zeta}$, the value of the cutoff is given by

$$\zeta_c = \frac{1}{k}\ln(Na). \tag{134}$$

While this cutoff regularizes all the integrals and captures the gross effects of the stochastic dynamics, it is not quantitatively accurate for fluctuating fronts. Previous studies showed significant differences between the predictions of Eq. (134) and simulations and argued that the factor multiplying $\ln(N)$ in Eq. (134) should be different from $1/k$ [33]. The main goal of this section is to derive the correct cutoff for fluctuating fronts.

Before proceeding with fluctuating fronts, however, it is important to point out that Eq. (134) prescribes the correct cutoff for deterministic models with discrete entities [91]. In such models, the main effect of discreteness is simply the absence of growth for $n < 1/a$, and, therefore, Eq. (134) does apply. Our simulations show clear differences in the scaling of Δv and Λ with N for deterministic and fluctuating fronts (Fig. 13). Moreover, these differences are explained entirely by the different cutoffs that one needs to apply for fluctuating and deterministic fronts.

Cutoff for fluctuating fronts
Analysis of fluctuating fronts is a challenging problem that is typically addressed by matching the nonlinear quasi-deterministic dynamics at the bulk of the front and the linear, but stochastic dynamics at the front edge [24]. Recently, a more rigorous approach has been developed in Refs. [56, 81], which relies on modifying the reaction-diffusion model to ensure that the hierarchy of moment equations closes exactly. The details of this approach are sufficiently technical and tangential to the main issues discussed in this paper, so we do not discuss them here. Instead, we refer the readers to Ref. [56] for a self-contained presentation of the new method.

For our purpose, the most useful result from Ref. [56] is that the deterministic equation for the steady-state density profile needs to be modified as

$$Dn'' + vn' + r(n)n - \frac{\gamma_n(n)n^2 e^{\frac{v\zeta}{D}}}{\int_{-\infty}^{+\infty} n^2 e^{\frac{v\zeta}{D}} d\zeta} = 0; \tag{135}$$

see Eq. (10) in Ref. [56].

The only difference between Eq. (135) and Eq. (13) is an additional term, which, as we show below, effectively imposes a cutoff on the growth rate. To quantify the magnitude of the new term, it is convenient to define a ratio between the terms due to front fluctuations and population growth:

$$E(\zeta) = \frac{\frac{\gamma_n(n) n^2 e^{\frac{v\zeta}{D}}}{\int_{-\infty}^{+\infty} n^2 e^{\frac{v\zeta}{D}} d\zeta}}{r(n) n} = \frac{\gamma_n(n) n e^{\frac{v\zeta}{D}}}{r(n) \int_{-\infty}^{+\infty} n^2 e^{\frac{v\zeta}{D}} d\zeta}. \tag{136}$$

Note that the first three terms in Eq. (135) are of the same order at the front, so any one of them could be used to define E.

Since we are only interested in the behavior of E near the front edge, E can be further simplified as

$$E(\zeta) \approx \frac{\gamma_n(0) n e^{\frac{v\zeta}{D}}}{r(0) \int_{-\infty}^{+\infty} n^2 e^{\frac{v\zeta}{D}} d\zeta} = \frac{\gamma_n(0) v \rho e^{\frac{v\zeta}{D}}}{r(0) D N I}, \tag{137}$$

where, in the last equality, we made the dependence on all dimensional quantities explicit by using $\rho = n/N$ and introducing a non-dimensional integral

$$I = \int_{-\infty}^{+\infty} \rho^2(\zeta) e^{\frac{v\zeta}{D}} d\left(\frac{v\zeta}{D}\right). \tag{138}$$

To obtain the scaling behavior of E for large ζ, we approximate ρ as $e^{-k\zeta}$ and replace v/D by $k + q$ (see Eq. (19))

$$E \sim e^{-\zeta(k - v/D)} \sim e^{q\zeta}. \tag{139}$$

Note that q is defined in Eq. (18) as the decay rate for the solution of Eq. (17) that is inconsistent with the boundary conditions. Therefore, q corresponds to the unphysical part of the solution for ρ and does not directly enter the asymptotic behavior of the population density. In the following, q is often used instead of v to make the formulas more compact.

We now determine the cutoff ζ_c for all subtypes of traveling waves. The main idea is to check whether the solution of Eq. (13) is consistent with the addition of a term due to front fluctuations in Eq. (135). If the solution is consistent, then no cutoff is necessary. If the solution is not consistent, then it can be valid only up to some critical ζ, which acts as an effective cutoff.

Cutoff for pushed waves expanding into a metastable state
When the invaded state is metastable, the low-density growth rate is negative, and, therefore, $q < 0$;

see Eq. (19). In consequence, $E \to 0$ as $\zeta \to +\infty$, and front fluctuations have a negligible effect on wave dynamics. Thus, no cutoff is necessary, i.e. $\zeta_c = +\infty$.

Cutoff for pushed waves expanding into an unstable state
When the invaded state is unstable, q is positive, and E diverges as $\zeta \to +\infty$. This contradicts Eq. (135), where all terms need to cancel out. To satisfy the equation, ρ must decay faster than $e^{-k\zeta}$ at the front beyond some critical ζ_c, so that E never becomes much greater than one. The value of ζ_c is then determined by the solution of $E(\zeta_c) = 1$ with the deterministic approximation for ρ. Hence, we substitute $\rho \sim e^{-k\zeta}$ in Eq. (137) and find that

$$\zeta_c = \frac{1}{q} \ln \left(N \frac{Dr(0)I}{v\gamma_n(0)} \right) = \frac{1}{q} \ln \left(\frac{N}{k+q} \frac{r(0)I}{\gamma_n(0)} \right) \sim \frac{1}{q} \ln N. \tag{140}$$

This is the most important result of this section because it determines the novel scaling behavior of Δv, D_{f}, and Λ in semi-pushed waves. To the best of our knowledge, Eq. 140 is a new finding.

Since $q < k$, fluctuating fronts have a larger ζ_c than deterministic fronts. In consequence, the normalized cutoff density $\rho_c = e^{-k\zeta_c}$ scales as $N^{-k/q}$ for fluctuating fronts rather than as N^{-1}. The applicability of the continuum theory for $\rho < 1/N$ seems counter-intuitive, not only because the expected number of organisms at a site falls below one, but also because fluctuations should appreciably modify the profile density at least when $\rho < 1/\sqrt{N}$, i.e. well before the naive $1/N$ cutoff. The key problem with this argument is that it assumes a continual front and neglects the possibility that a sufficiently large group of organisms can occasionally expand well ahead of the deterministic front [21, 92]. Such front excursions prevent a sharp cutoff in the population density below $1/N$. More importantly, they significantly amplify both genetic drift and front wandering. In the continuum theory, this increase in fluctuations is captured by greater ζ_c, which increases both Λ and D_{f}.

Cutoff for pulled waves
For pulled waves, $q > 0$, and we obtain the value of the cutoff by solving $E(\zeta_c) = 1$ just as for semi-pushed waves:

$$\gamma_n(0)\rho(\zeta_c)e^{\frac{v\zeta_c}{D}} = r(0)N \int_{-\infty}^{\zeta_c} \rho^2(\zeta)e^{\frac{v\zeta}{D}} d\zeta. \tag{141}$$

Note, however, that there are two important differences in the calculation for pulled compared to semi-pushed waves. First, the integral I diverges and needs to be cut off at ζ_c. Second, the front shape also acquires corrections due to the cutoff and needs to be determined self-consistently. This sensitivity of the front shape originates from the degeneracy $k = q$ that occurs in pulled waves. In the continuum limit, this degeneracy modifies the scaling of ρ from $e^{-k_{\mathrm{F}}\zeta}$ to $k_{\mathrm{F}}\zeta e^{-k_{\mathrm{F}}\zeta}$. For a fluctuating front, however, the wave velocity deviates slightly from v_{F}, and the correction to the front shape is different.

The shape of the front can be obtained by setting the growth rate to zero for $\zeta > \zeta_c$ and solving the resulting equation for $\rho(\zeta)$; see Refs. [1, 33, 35]. The result reads

$$\rho(\zeta) \sim \frac{k_{\rm F}\zeta_c}{\pi} \sin\left(\pi \frac{\zeta_c - \zeta}{\zeta_c}\right) e^{-k_{\rm F}\zeta}. \tag{142}$$

Upon substituting this result in Eq. (141), we find the following condition on ζ_c:

$$e^{k_{\rm F}\zeta_c} \sim \frac{1}{2\pi^2} \frac{r(0)N}{\gamma_n(0)k_{\rm F}} (k_{\rm F}\zeta_c)^3, \tag{143}$$

where we shifted ζ_c in the argument of the sine by $1/k_{\rm F}$, which does not change the asymptotic scaling of the exponential term, but avoids setting the left hand side to zero. To solve Eq. (143), we treat $(k_{\rm F}\zeta_c)^3$ as a small perturbation compared to $\frac{r(0)N}{\gamma_n(0)k_{\rm F}}$ and obtain that, up to additive numerical factors, the leading behavior of ζ_c is given by

$$\zeta_c = \frac{1}{k_{\rm F}} \ln\left(\frac{r(0)N}{\gamma_n(0)k_{\rm F}}\right) + \frac{3}{k_{\rm F}} \ln\left[\ln\left(\frac{r(0)N}{\gamma_n(0)k_{\rm F}}\right)\right] \sim \frac{1}{k_{\rm F}} \ln N + \frac{3}{k_{\rm F}} \ln\ln N. \tag{144}$$

Note that the second order term cannot be neglected in the calculation of Δv, $D_{\rm f}$, and Λ because these quantities have terms that scale both linearly and exponentially with ζ_c. Equation (144) was first motivated phenomenologically in Ref. [35] and then derived more rigorously in Ref. [56].

X. Scaling of Δv, $D_{\rm f}$, and Λ in pulled, semi-pushed, and fully-pushed waves

In this section, we synthesize the results of the perturbation theory for Δv, $D_{\rm f}$, and Λ and supplement them with an appropriate cutoff ζ_c when needed. We show that pushed waves consist of two distinct classes. In fully-pushed waves, the fluctuations scale as $1/N$ consistent with the central limit theorem, but, in semi-pushed waves, non-trivial power scaling occurs. The exponent of this power law depends only on $v/v_{\rm F}$ and is the same for Δv, $D_{\rm f}$, and Λ. For completeness, also provide the corresponding results for pulled waves and models with deterministic fronts. While most results of the perturbation theory are not new, the synthesis of these results, the application of an appropriate cutoff, and the discovery of semi-pushed waves are novel contributions of this paper.

The main results of the perturbation theory are given by Eq. (117) for the correction to wave velocity, by Eq. (87) for the effective diffusion constant of the front, and by Eq. (125) for the rate of diversity loss. All of these equations, have a similar form and contain a ratio of two integrals. Both integrals converge for $\zeta \to -\infty$, but they could diverge for $\zeta \to +\infty$. At the front, the integrands in the numerator scale as $e^{-(3k-2v/D)\zeta}$, and the integrands in the denominators scale as $e^{-(2k-v/D)\zeta}$; therefore, the integrals in the denominators always converge when the integrals in the numerators converge. Since the ratio of k to v/D depends on the degree to which the growth is cooperative, the integrals could change their behavior as cooperativity is varied. A change from convergence to divergence in either of the integrals corresponds to a transitions between different classes of waves. Below we consider each class separately.

$1/N$ scaling in fully-pushed waves
The class of fully-pushed waves is defined by the requirement that all integrals converge. In this

case, the perturbation theory is well-posed without a cutoff and provides not only the scaling, but also the exact values of Δv, D_{f}, and Λ. Straightforward dimensional analysis shows that all three quantities scale as $1/N$, i.e. the central limit theorem holds.

The convergence of integrals requires that kD/v is greater than $2/3$. For waves expanding into an metastable state, $kD/v > 1$ (see Eq. (19)), so these waves are always fully-pushed. For expansions into an unstable state, it is convenient to express the convergence condition only as a function of v using Eq. (19):

$$v \geq \frac{3}{2\sqrt{2}} v_{\mathrm{F}}. \tag{145}$$

where $v_{\mathrm{F}} = 2\sqrt{Dr(0)}$ is the linear spreading velocity. We emphasize that v_{F} serves only as convenient notation for $2\sqrt{Dr(0)}$; in particular, the wave is not pulled, and the wave velocity is greater than v_{F}.

Equation (145) immediately implies that not every pushed wave is fully-pushed. Indeed, only $v > v_{\mathrm{F}}$ is required for a wave to be pushed, which is a weaker condition than Eq. (145). Because v/v_{F} increases with cooperativity, fully-pushed waves occur once cooperativity exceeds a certain threshold.

We can also express the condition that $kD/v > 2/3$ in terms of k and q using Eq. (19). Because $v/D = k + q$, this convergence condition is equivalent to $k > 2q$. For all pushed waves, $k > q$, but a stronger inequality is required for fully-pushed waves. Note that $q < 0$ for waves propagating into a metastable state, so $k > q$ is satisfied.

Finally, we discuss the effects of a cutoff derived in the previous section. For expansions into a metastable state, $\zeta_c = +\infty$, i.e. no cutoff is necessary. For expansions into an unstable state, the theory suggest a finite cutoff: $\zeta_c \sim \ln(N)/q$. Note, however, that the application of this cutoff in the formulas for Δv, D_{f}, and Λ only produces subleading corrections to the $1/N$ scaling because convergent integrals are insensitive to small changes in their upper limit of integration.

N^α scaling in semi-pushed waves

We now proceed to the second class of pushed waves, for which the integrals in the numerators diverge, i.e. $v < \sqrt{9/8} v_{\mathrm{F}}$. We term the waves in this class semi-pushed because the fluctuations at the front make a significant contribution to their dynamics. Note that the integrals in the denominators converge for all pushed waves because $kD/v > 1/2$; see Eq. (19).

To estimate the scaling of Δv, D_{f}, and Λ, we cut off the integrals in the numerators at $\zeta_c \sim \ln(N)/q$ and find that three quantities scale as N^α with α given by

$$\alpha = -2 \frac{\sqrt{1 - v_{\mathrm{F}}^2/v^2}}{1 - \sqrt{1 - v_{\mathrm{F}}^2/v^2}}. \tag{146}$$

The details of this calculation for Λ are summarized below

$$\Lambda = \frac{1}{N}\frac{\int_{-\infty}^{\zeta_c}\gamma_f(\rho)\rho^3(\zeta)e^{\frac{2v\zeta}{D}}d\zeta}{\left(\int_{-\infty}^{+\infty}\rho^2(\zeta)e^{\frac{v\zeta}{D}}d\zeta\right)^2} \sim \frac{1}{N}\frac{\gamma_f(0)\int_{-\infty}^{+\zeta_c}e^{-(k-2q)\zeta}d\zeta}{\left(I\frac{D}{v}\right)^2} \sim \frac{(k+q)^2\gamma_f(0)}{I^2N}\frac{e^{-\frac{(k-2q)}{q}\ln\left(\frac{N}{k+q}\frac{r(0)I}{\gamma_n(0)}\right)}}{k-2q}$$
$$\sim N^{-\frac{k-q}{q}}I^{-\frac{k}{q}}\gamma_f(0)[\gamma_n(0)]^{\frac{k-2q}{q}}[r(0)]^{-\frac{k-2q}{q}}(k+q)^{\frac{k}{q}}(k-2q)^{-1} \sim N^{-\frac{k-q}{q}} \sim N^{-2\frac{\sqrt{v^2-v_{\mathrm{F}}^2}}{v-\sqrt{v^2-v_{\mathrm{F}}^2}}}. \tag{147}$$

where we used Eqs. (19), (125), (138), and (140). The calculations for D_{f} and Δv are essentially the same.

Logarithmic scaling in pulled waves
The remaining possibility is that the integrals diverge in both numerators and denominators. This is the case for pulled waves because $v_{\mathrm{F}} = 2k_{\mathrm{F}}D$. To compute the asymptotic scaling of D_{f} and Λ, we use (87) and (125) together with the cutoff from Eq. (144) and the profile shape from Eq. (142). The results read

$$\begin{aligned} D_{\mathrm{f}} &\sim \frac{\gamma_n(0)}{k_{\mathrm{F}}^2}\ln^{-3}\left(\frac{N}{k_{\mathrm{F}}}\right), \\ \Lambda &\sim \gamma_f(0)\ln^{-3}\left(\frac{N}{k_{\mathrm{F}}}\right). \end{aligned} \tag{148}$$

This completes our discussion of different scaling regimes in fluctuating fronts.

Scaling with N in deterministic fronts
Some of our results for fluctuating fronts depend on the specific choice of the cutoff $\zeta_c = \ln(N)/q$. This cutoff is different from the naive expectation that $\zeta_c = \ln(N)/k$ because occasional fluctuations establish a small population far ahead of the deterministic front. To understand the effect of such fluctuations, we now examine the properties of deterministic fronts, where $\gamma_n = 0$, but there is a cutoff on the growth rate below $\rho_c \sim 1/N$. Since deterministic fronts do not fluctuate, their diffusion constant is zero. Genetic drift, however, occurs even without any fluctuations in the total population size, so the rate of diversity loss is well-defined. Therefore, we focus on the scaling of Λ with N in this subsection.

Our analysis of fully-pushed waves remains unchanged because all the integrals converge, and a cutoff is not required. Thus, $\Lambda \sim N^{-1}$ for fully-pushed waves with or without demographic fluctuations. Moreover, the transition point between fully-pushed and semi-pushed waves remains the same because it depends on the behavior of the integrands in Eq. (125) rather than on the value of the cutoff.

For semi-pushed waves, ζ_c does enter the calculation and changes the value of α. For deterministic fronts, we find that

$$\alpha_{\text{deterministic}} = -2\frac{k-q}{k} = -\frac{4\sqrt{1-v_{\text{F}}^2/v^2}}{1+\sqrt{1-v_{\text{F}}^2/v^2}}. \tag{149}$$

Similarly to our results for the fluctuating fronts, $\alpha_{\text{deterministic}}$ approaches 0 and -1 near the transitions to pulled and fully-pushed waves. Within the class of semi-pushed waves, however, $\alpha_{\text{deterministic}}$ is less than α for fluctuating fronts ($|\alpha_{\text{deterministic}}| > |\alpha_{\text{fluctuating}}|$), that is genetic drift is amplified by front fluctuations.

For pulled waves, we find that

$$\Lambda \sim \gamma_f(0) \ln^{-6}(N/k_{\text{F}}), \tag{150}$$

which further supports the fact that genetic drift is weaker without front fluctuations.

The $\ln^{-6} N$ scaling was previously suggested for the diffusion constant of pulled waves based on the incorrect application of the naive cutoff [36]. Moreover, simulations that limited the extent of demographic fluctuations indeed observed that $D_{\text{f}} \sim \ln^{-6} N$ [91].

Comparison of Δv due to a cutoff in growth and due to demographic fluctuations
We close this section by comparing the results for Δv from the calculation based on the growth rate cutoff (Eq. (74)) and the calculation based on treating demographic noise as a perturbation (Eq. (117)). The former calculation always requires a cutoff even for fully-pushed waves and, therefore, does not signal the transition between different wave classes. In contrast, the latter calculation shows that $\Delta v \sim 1/N$ for fully-pushed waves and that a cutoff is necessary for semi-pushed waves. Both equations predict the same scaling of Δv with N in semi-pushed waves provided the correct cutoff $\zeta_c = \ln(N)/q$ is used for fluctuating fronts. For fully-pushed waves, however, only Eq. (117) agrees with the $1/N$ scaling observed in simulations.

For deterministic fronts, only Eq. (74) is valid, and it predicts that

$$\Delta v_{\text{deterministic}} \sim N^{-(k-q)/k} \sim N^{-2\frac{\sqrt{1-v_{\text{F}}^2/v^2}}{1+\sqrt{1-v_{\text{F}}^2/v^2}}}. \tag{151}$$

Note that the exponent is different from the one in Eq. (149). Thus, in deterministic fronts, Δv and Λ depend on N differently. In particular, $\Delta v \sim 1/\sqrt{N}$ at the transition between semi-pushed and fully-pushed waves, and the $1/N$ scaling occurs only when $r(0) = 0$, i.e. at the transition between propagation into unstable and metastable states. For waves propagating into a metastable state, Δv decays faster than $1/N$; such scaling is not possible for fluctuating fronts.

XI. Precise definitions of the foci of growth, ancestry, and diversity

In this section, we consolidate the results obtained above on the spatial distribution of growth, ancestry, diversity processes within the wave front. We also provide the precise definitions of the

foci of growth, ancestry, and diversity.

The spatial distribution of the per capita growth rate is given by

$$\text{growth distribution} \sim r(n(\zeta)). \tag{152}$$

The mode of this distribution is the focus of growth. For monotonically decrseasing $r(n)$, the focus of growth is at the very edge of the front, i.e. at $\zeta = \zeta_c$, but it is in the interior of the front otherwise.

The spatial distribution of the most recent common ancestor of the entire population at the front is given by Eq. (38):

$$\text{ancestry distribution} \sim n^2(\zeta)e^{v\zeta/D}. \tag{153}$$

The mode of this distribution is the focus of ancestry, which is the most likely location of the most recent common ancestor. The focus of ancestry is located at $\zeta = \zeta_c$ for pulled waves and in the interior of the front for pushed waves.

To characterise the contribution of the different regions of the front to genetic diversity, we consider the spatial distribution of the locations where two ancestral lineages coalesce. That is we consider the spatial location of the most recent common ancestor of two randomly sampled individuals. From Eq. (129), it follows that this distribution is given by

$$\text{diversity distribution} \sim \frac{\gamma_f(n(\zeta))}{n(\zeta)} S^2(\zeta) \sim \gamma_f(n(\zeta))n^3(\zeta)e^{2v\zeta/D}. \tag{154}$$

The mode of this distribution is the focus of ancestry. The focus of diversity is located at $\zeta = \zeta_c$ for pulled and semi-pushed waves and in the interior of the front for fully-pushed waves.

XII. Computer simulations

In this section, we explain the details of our computer simulations and the subsequent data analysis.

Interpretation of the simulations as the Wright-Fisher model with vacancies
Deterministic migration between patches followed by the Wright-Fisher sampling provides one of the most efficient ways to simulate population dynamics. In its standard formulation, the Wright-Fisher model cannot simulate population growth because it assumes that the population size is fixed at the carrying capacity. To overcome this difficulty, we generalized the Wright-Fisher model by considering the number of vacancies, i.e. the difference between the carrying capacity N and the total population density n, as the abundance of an additional species. With this modification, the total abundance of the two genotypes can increase at the expense of the number of vacancies.

Following Ref. [37], the growth of the population was modeled by introducing a fitness difference

between the vacancies and the actual species. Specifically, the fitness of the two genotypes was set to $w_i = 1$ and the fitness of the vacancies was set to $w_\mathrm{v} = 1 - r(n)/(1 - n/N)$. The probability to sample genotype i was then proportional to the ratio of w_i to the mean fitness of the population $\bar{w} = n/N + w_\mathrm{v}(N - n)/N = 1 - r(n)$, which explains why we used $1/(1 - r(n)\tau)$ instead of $1 + r(n)\tau$ in Eq. (10).

Simulations of deterministic fronts
We also simulated range expansions without demographic fluctuations ($\gamma_n = 0$), but with genetic drift. In these simulations, the total population density was updated deterministically:

$$n(t + \tau, x) = \lfloor (p_1 + p_2)N \rfloor \,, \tag{155}$$

where p_i are the same as in Eq. (10), and $\lfloor y \rfloor$ denotes the floor function, which is equal to the greast integer less than y. The abundances of the two neutral genotypes were then determined by Binomial sampling with $n(t + \tau, x)$ trials and $p_i/(p_1 + p_2)$ probability of choosing genotype i.

For all simulations $m = 0.25$ and $r_0 = g_0 = 0.01$ were used, unless noted otherwise.

Boundary and initial conditions
The most direct approach to simulating a range expansion is to use a stationary habitat, in which the range expansion proceeds from one end to the other. This approach is however expensive because the computational times grows quadratically with the duration of the simulations. Instead, we took advantage of the fact that all population dynamics are localized to the vicinity of the expansion front and simulated only a region of 300 patches comoving with the expansion. Specifically, every simulation time step, we shifted the front backward if the total population inside the simulation array n_array exceeded $150N$, i.e. half of the maximally possible population size. The magnitude of the shift was equal to $\lfloor (n_\mathrm{array} - 150N)/N \rfloor + 1$. The population density in the patches that were added ahead of the front was set to zero, and the number of the individuals moved outside the box was storted, so that we could compute the total number of individuals n_tot in the entire population including both inside and outside of the simulation array. Our choice of 300 patches in the simulation array was sufficient to enusure that at least one patch remained always unoccupied ahead of the expansion front and that the patches shifted outside the array were always at the carrying capacity.

We initialized all simulations by leaving the right half of the array unoccupied and filling the left half to the carrying capacity. In each occupied patch, we determined the relative abundance of the two neutral genotypes by sampling the binomial distribution with N trials and equal probabilities of choosing each of the genotypes.

Duration of simulations and data collection
To ensure that we can access the exponential decay of the average heterozygosity, simulations were carried out for $2\sqrt{N}$ generations for pulled waves and for N generations for pushed waves. Although, for pulled waves, the expected timescale of heterozygosity decay is $\ln^3 N$, we chose a longer duration of simualtions to account for possible deviations from this asymptotic scaling. In all simulations, the minimal simulation time was set to 10^4 time steps.

For each simulation, we saved the total population size n_{tot} and the population heterozygosity h at 1000 time points evenly distributed across the simulation time. These were used to compute $\mathrm{Var}X_{\mathrm{f}}$ and H by averaging over 1000 independent simulation runs.

Computing front velocity
The velocity of the front was measured by fitting n_{tot}/N to $vt + \mathrm{const}$. For this fit, we discarded the first 10% of the total simulation time (1000 generations for the shortest runs) to account for the transient dynamics. The length of the transient is the largest for pulled waves and is specified by the following result from Ref. [1]:

$$v(t) = v_{\mathrm{F}} \left[1 - \frac{3}{4r_0 t} + \mathcal{O}\left(t^{-3/2}\right) \right]. \tag{156}$$

Thus, discarding time points prior to $t \sim 1/r_0$ was sufficient to eliminate the transient dynamics in all of our simulations.

Computing the diffusion constant of the front
To measure D_{f}, we discarded early time points as described above and then fitted $\mathrm{Var}\{n_{\mathrm{tot}}/N\}$ to $2D_{\mathrm{f}}t + \mathrm{const}$.

Computing heterozygosity and the rate of its decay
For each time point, the heterozygosity h was computed as follows

$$h(t) = \frac{1}{300} \sum_x \frac{2n_1(t,x)n_2(t,x)}{[n_1(t,x) + n_2(t,x)]^2}, \tag{157}$$

where the sum is over x within the simulation array. The average heterozygosity H was then obtained by averaging over independent simulation runs. To compute Λ we fitted $\ln H$ to $-\Lambda t + \mathrm{const}$.

The transient, non-exponential, decay of H lasted much longer compared to the transient dynamics of v and $\mathrm{Var}X_{\mathrm{f}}$; in addition, our estimates of H had large uncertainty for large t because only a few simulation runs had non-zero heterozygosity at the final time point. To avoid these sources of error, we restricted the analysis $H(t)$ to $t \in (t_{\mathrm{i}}, t_{\mathrm{f}})$. The value of t_{f} was chosen such that at least 50 simulations had non-zero heterozygosity at $t = t_{\mathrm{f}}$. The value of t_{i} was chosen to maximise the goodness of fit (R^2) between the fit to $H \sim e^{-\Lambda t}$ and the data subject to the constraint that $t_{\mathrm{f}} - t_{\mathrm{i}} > 1000$. The latter constraint ensured that we had a sufficient number of uncorrelated data points to carry out the fitting procedure.

Computing the scaling exponents for D_{f}, Λ, and $v - v_{\mathrm{d}}$
To quantify the dependence of D_{f}, Λ, and $v - v_{\mathrm{d}}$ on N, we fitted a power-law dependence using linear regression on log-log scale. Because the power-law behavior is only asymptotic and did not match the results for low N, the exponents were calculated using the data only for $N > 10^4$.

For the velocity corrections, we also needed to determine the value of v_{d}. This was done by

maximizing the goodness of fit (R^2) between the simulation results and theoretical predictions.

XIII. Supplemental results and figures

In this section, we present additional simulation data that further supports and clarifies the conclusions made in the main text. Of particular interest is the comparison between deterministic and fluctuating fronts and the results for an alternative model of an Allee effect that can describe propagation into a metastable state (strong Allee effect).

Figure 8 shows the data that we used to conclude that fluctuations in semi-pushed waves exhibit different scaling behavior compared to pulled and fully-pushed waves. Figure 9 demonstrates that the perturbation theory accurately predicts not only the scaling with N, but also the exact values of D_f and Λ for fully-pushed waves.

The scaling properties of fully-pushed waves that propagate into a metastable state are shown in Fig. 10. This figure also illustrates the transition from pulled to semi-pushed and then to fully-pushed waves in an alternative model of an Allee effect.

Figure 11 shows that the transition between different wave classes can also be detected from the small corrections to the wave velocity due to demographic fluctuations.

Genetic drift in deterministic fronts is examined in Fig. 12, and Fig. 13 compares the scaling behavior of Λ with N in deterministic vs. fluctuating fronts.

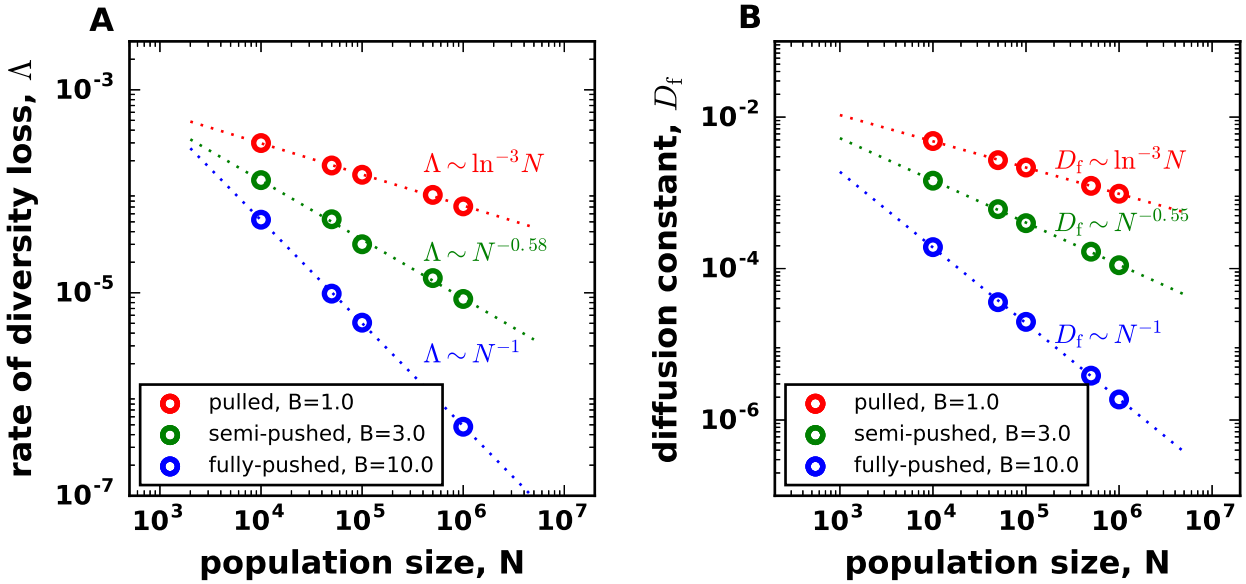

Figure 8. Dependence of D_f and Λ on N for all three classes of waves. Circles show the results from the simulations, and dashed lines show the fits of the expected asymptotic scaling: power law for pushed waves and $\ln^{-3} N$ for pulled waves. The simulation results are for the growth rate specified by Eq. (3).

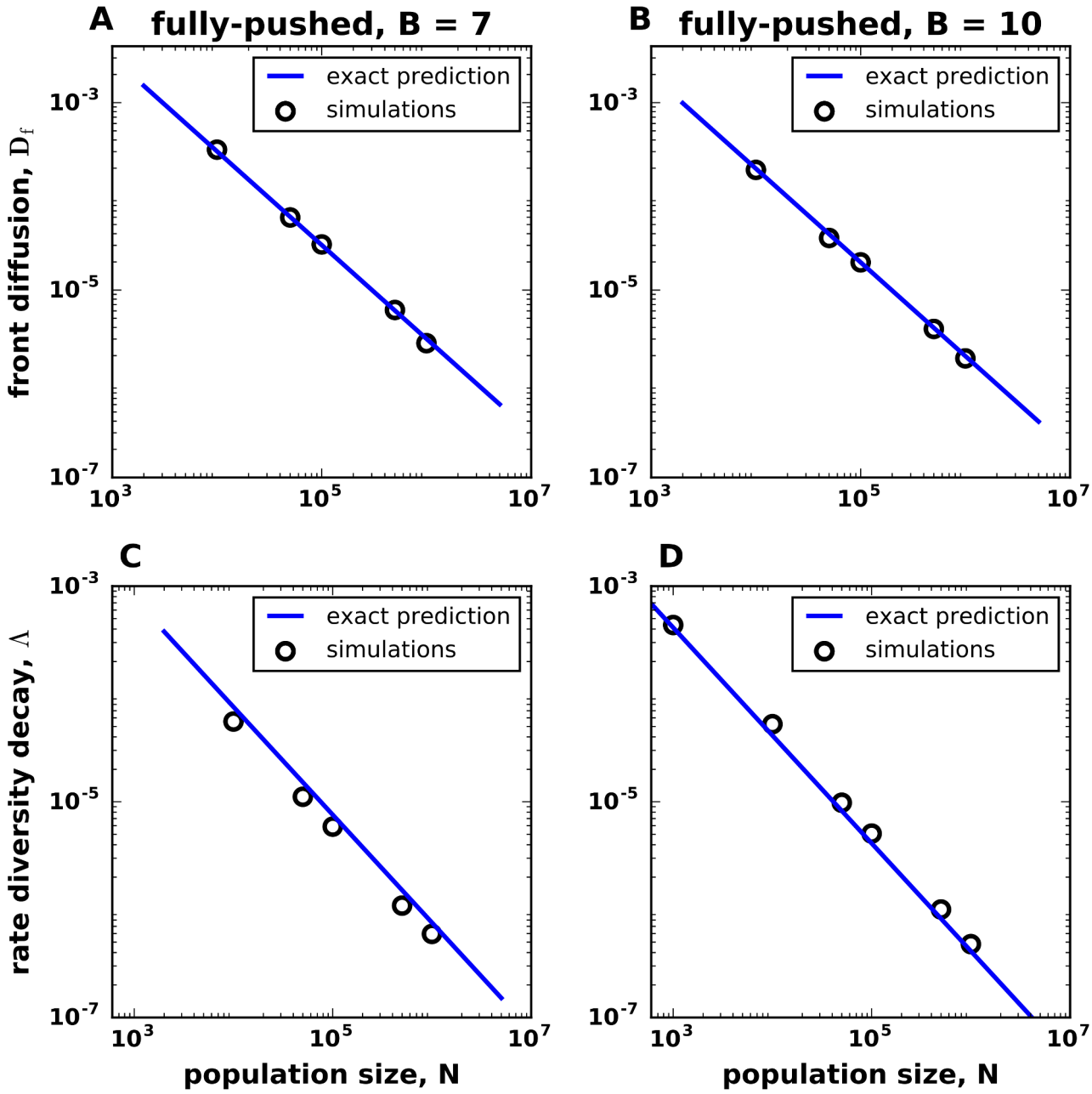

Figure 9. For fully-pushed waves, theoretical predictions agree with the simulation results without any fitting parameters. The diffusion constant of the front D_f and the rate of diversity loss Λ are shown with circles for two values of cooperativity B. Both values of B are greater than the minimal cooperativity required for fully-pushed waves. The solid lines are the predictions of the perturbation theory: Eq. (96) for D_f and Eq. (132) for Λ. The simulation results are for the growth rate specified by Eq. (3).

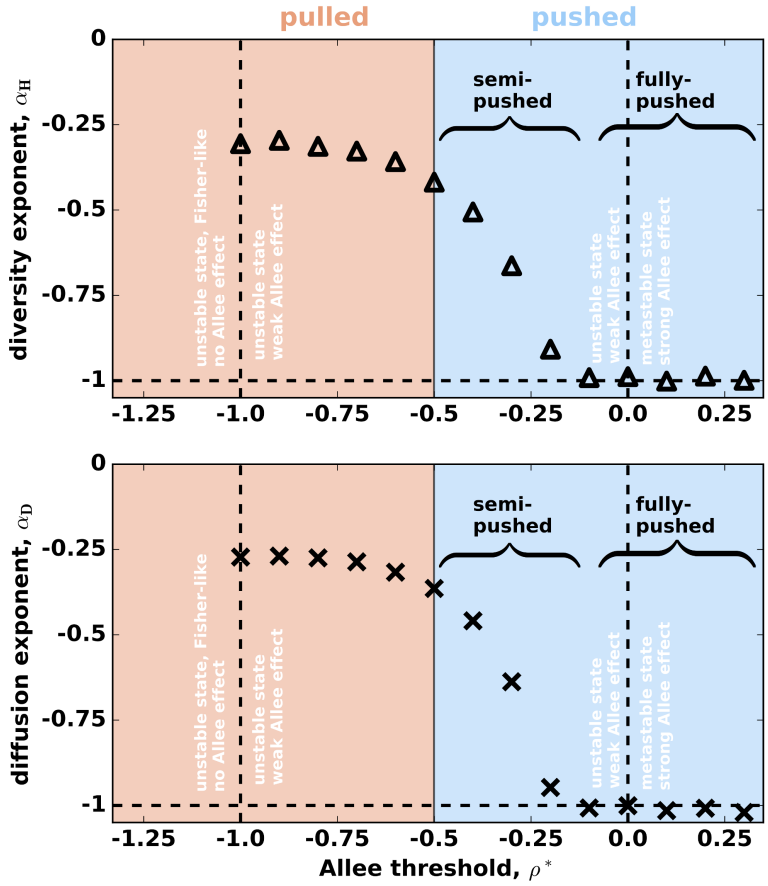

Figure 10. Waves propagating into a metastable state are fully-pushed. This figures shows the scaling exponents α_D and α_H for an alternative model of an Allee effect specified by Eq. (24). A strong Allee effect is possible in this model, so the waves can propagate both into an unstable and metastable states depending on whether the Allee threshold is negative or positive. Note that, both the diffusion constant of the front and the rate of diversity loss scale as $1/N$ for fully-pushed waves irrespective of the stability of the invaded state. The transition between pulled and semi-pushed waves occurs at $\rho^* = -0.5$ and between semi-pushed and fully-pushed waves at $\rho^* = -0.25$.

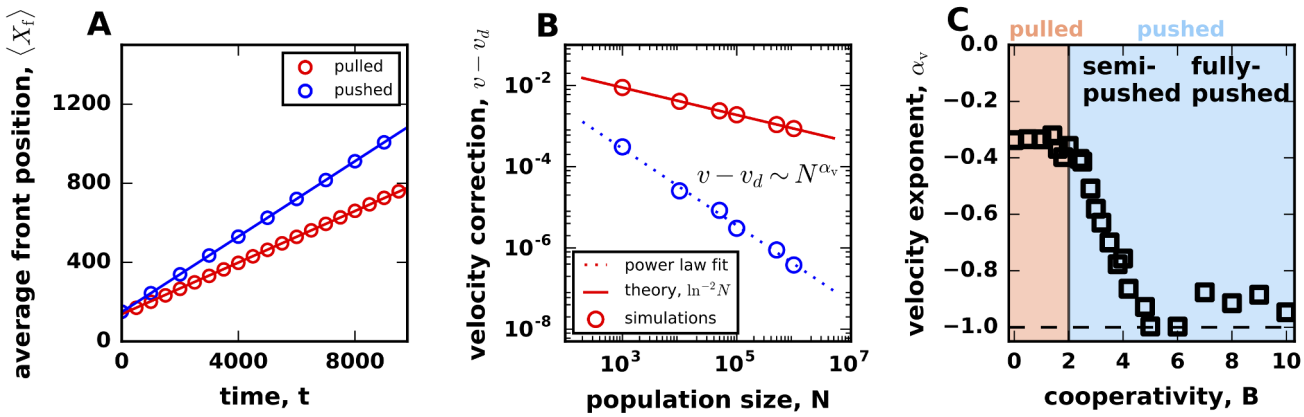

Figure 11. Correction to wave velocity also shows a transition between semi-pushed and fully-pushed waves. (A) The average position of the front increases linearly with time for both pulled and pushed expansions. **(B)** The velocity of the front is reduced by demographic fluctuations below its deterministic value v_d. For pulled waves, $v - v_d \sim \ln^{-2} N$, while, for pushed waves, $v - v_d$ decreases as a power law $N^{-\alpha_v}$. **(C)** The dependence of α_v on cooperativity is the same as for α_D and α_H. In particular, $v(N)$ clearly shows that the class of pushed waves consists of two subclasses: fully-pushed waves with $v - v_d \sim 1/N$ and semi-pushed waves with $\alpha_D \in (-1, 0)$. Similar to other figures, the limited range of N in simulations makes logarithmic scaling for pulled waves also consistent with a power-law with a small negative exponent. The simulation results are for the growth rate specified by Eq. (3).

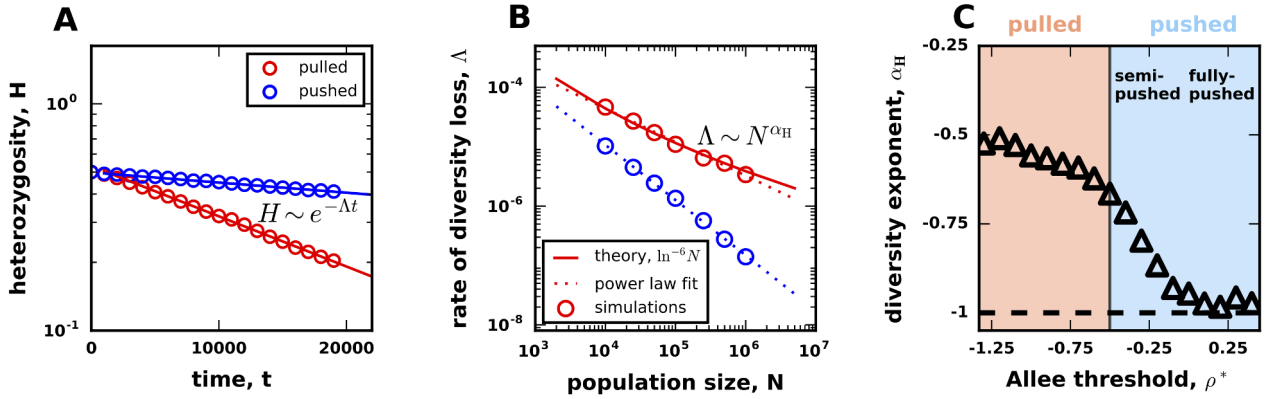

Figure 12. The rate of diversity decay for deterministic fronts. (A) The average heterozygosity, H, is a measure of diversity equal to the probability to sample two distinct genotypes in the population and is well-defined for both fluctuating and deterministic fronts. The decay of genetic diversity is exponential in time: $H \sim e^{-\Lambda t}$ for both pulled and pushed waves. **(B)** For deterministic pulled waves, $\Lambda \sim \ln^{-6} N$ from Eq. (150), while, for fully-pushed waves, $\Lambda \sim N^{-1}$ from Eq. (4). To quantify the dependence of Λ on N, we fit $\Lambda \sim N^{\alpha_H}$. The dashed red line shows that even though α_H should equal 0 for pulled waves, the limited range of N results in a different value of $\alpha_H \approx -0.51$. **(C)** The dependence of the scaling exponent on cooperativity identifies the same three classes of waves as in Fig. 5C for fluctuating fronts. The transitions between the wave classes occur at the same values of B, but the values of the exponents for semi-pushed waves are slightly different.

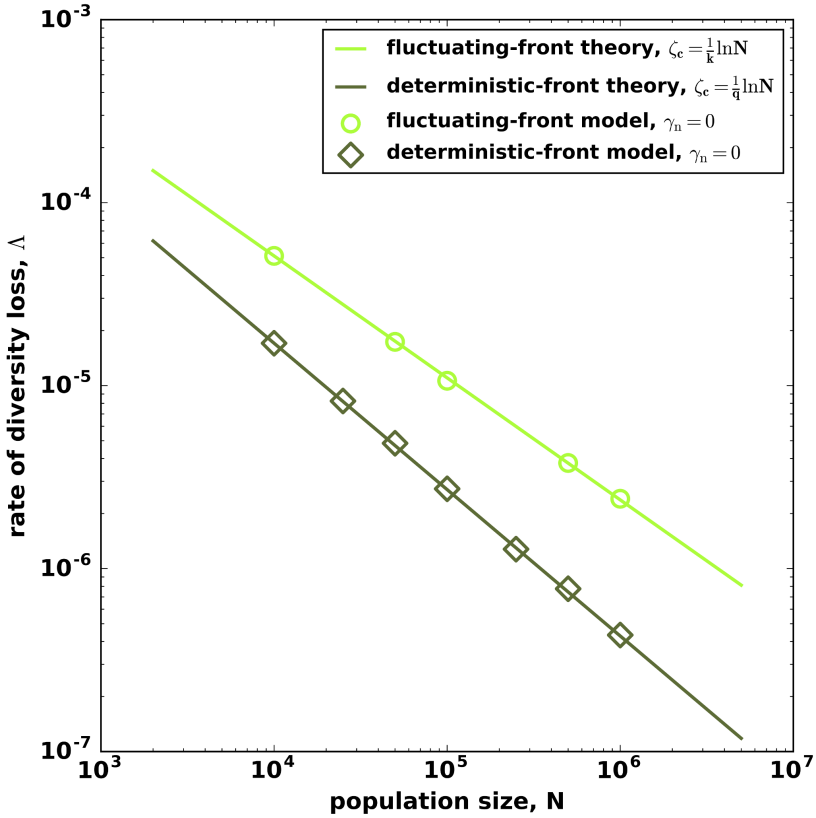

Figure 13. Comparison between deterministic and fluctuating fronts. The circles show the dependence of Λ on N for both stochastic and deterministic simulations of population fronts, and the lines show the corresponding theoretical predictions from Eqs. (146) and (149). It is clear that the theory correctly captures the contribution of the front fluctuations to the rate of genetic drift. For this figure, we chose the value of cooperativity well within the class of semi-pushed waves to avoid the contribution of the crossover behavior near the transition to pulled and fully-pushed waves. The simulation results are for the growth rate specified by Eq. (24) with $\rho^* = -0.3$.

References

[1] Van Saarloos, W. Front propagation into unstable states. *Physics Reports* **386**, 29–222 (2003).

[2] Cross, M. C. & Hohenberg, P. C. Pattern formation outside of equilibrium. *Reviews of modern physics* **65**, 851 (1993).

[3] Aranson, I. S. & Kramer, L. The world of the complex Ginzburg-Landau equation. *Reviews of Modern Physics* **74**, 99 (2002).

[4] Sachdev, P. L. *Nonlinear diffusive waves* (Cambridge University Press, 1987).

[5] Barenblatt, G. I. *Scaling, self-similarity, and intermediate asymptotics: dimensional analysis and intermediate asymptotics*, vol. 14 (Cambridge University Press, 1996).

[6] Douglas, J. F., Efimenko, K., Fischer, D. A., Phelan, F. R. & Genzer, J. Propagating waves of self-assembly in organosilane monolayers. *Proceedings of the National Academy of Sciences* **104**, 10324–10329 (2007).

[7] Fox, J. M. & Whitesides, G. M. Warning signals for eruptive events in spreading fires. *Proceedings of the National Academy of Sciences* **112**, 2378–2383 (2015).

[8] Ramaswamy, S., Toner, J. & Prost, J. Nonequilibrium fluctuations, traveling waves, and instabilities in active membranes. *Physical review letters* **84**, 3494 (2000).

[9] Murray, J. D. *Mathematical Biology* (Springer, 2003).

[10] Korolev, K. S. Evolution arrests invasions of cooperative populations. *Physical Review Letters* **115**, 208104 (2015).

[11] Nelson, P. *Biological physics* (WH Freeman New York, 2004).

[12] Takamatsu, T. & Wier, W. Calcium waves in mammalian heart: quantification of origin, magnitude, waveform, and velocity. *The FASEB Journal* **4**, 1519–1525 (1990).

[13] Weiner, O. D., Marganski, W. A., Wu, L. F., Altschuler, S. J. & Kirschner, M. W. An actin-based wave generator organizes cell motility. *PLoS Biol* **5**, e221 (2007).

[14] Chang, J. B. & Ferrell Jr, J. E. Mitotic trigger waves and the spatial coordination of the xenopus cell cycle. *Nature* **500**, 603–607 (2013).

[15] Ishihara, K., Korolev, K. S. & Mitchison, T. J. Physical basis of large microtubule aster growth. *eLife* **5**, e19145 (2016).

[16] Pálsson, E. *et al.* Selection for spiral waves in the social amoebae dictyostelium. *Proceedings of the National Academy of Sciences* **94**, 13719–13723 (1997).

[17] Hastings, A. *et al.* The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters* **8**, 91–101 (2005).

[18] Fagan, W. F., Lewis, M. A., Neubert, M. G. & Van Den Driessche, P. Invasion theory and biological control. *Ecology Letters* **5**, 148–157 (2002).

[19] Bocedi, G., Peer, G., Heikkinen, R. K., Matsinos, Y. & Travis, J. M. Projecting species range expansion dynamics: sources of systematic biases when scaling up patterns and processes. *Methods in Ecology and Evolution* **3**, 1008–1018 (2012).

[20] Martín, P. V., Bonachela, J. A., Levin, S. A. & Muñoz, M. A. Eluding catastrophic shifts. *Proceedings of the National Academy of Sciences* **112**, E1828–E1836 (2015).

[21] Hallatschek, O. & Korolev, K. Fisher waves in the strong noise limit. *Physical Review Letters* **103**, 108103 (2009).

[22] Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–1342 (2013).

[23] Tsimring, L. S., Levine, H. & Kessler, D. A. RNA virus evolution via a fitness-space model. *Physical review letters* **76**, 4440 (1996).

[24] Rouzine, I. M., Wakeley, J. & Coffin, J. M. The solitary wave of asexual evolution. *Proceedings of the National Academy of Sciences* **100**, 587–592 (2003).

[25] Good, B. H., Rouzine, I. M., Balick, D. J., Hallatschek, O. & Desai, M. M. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proceedings of the National Academy of Sciences* **109**, 4950–4955 (2012).

[26] Neher, R. A. & Hallatschek, O. Genealogies of rapidly adapting populations. *Proceedings of the National Academy of Sciences* **110**, 437–442 (2013).

[27] Brunet, É. & Derrida, B. How genealogies are affected by the speed of evolution. *Philosophical Magazine* **92**, 255–271 (2012).

[28] Lewis, M. A. Finding the sweet spot for invasion theory. *Proceedings of the National Academy of Sciences* **113**, 6819–6820 (2016).

[29] Tanaka, H., Stone, H. A. & Nelson, D. R. Spatial gene drives and pushed genetic waves. *bioRxiv* 126722 (2017).

[30] Stokes, A. On two types of moving front in quasilinear diffusion. *Mathematical Biosciences* **31**, 307–315 (1976).

[31] Paquette, G., Chen, L.-Y., Goldenfeld, N. & Oono, Y. Structural stability and renormalization group for propagating fronts. *Physical review letters* **72**, 76 (1994).

[32] Kessler, D. A. & Levine, H. Velocity selection in dendritic growth. *Physical Review B* **33**, 7867 (1986).

[33] Kessler, D. A., Ner, Z. & Sander, L. M. Front propagation: precursors, cutoffs, and structural stability. *Physical Review E* **58**, 107 (1998).

[34] Gandhi, S. R., Yurtsev, E. A., Korolev, K. S. & Gore, J. Range expansions transition from pulled to pushed waves as growth becomes more cooperative in an experimental microbial population. *Proceedings of the National Academy of Sciences* **113**, 69226927 (2016).

[35] Brunet, E., Derrida, B., Mueller, A. & Munier, S. Phenomenological theory giving the full statistics of the position of fluctuating pulled fronts. *Physical Review E* **73**, 056126 (2006).

[36] Panja, D. Effects of fluctuations on propagating fronts. *Physics Reports* **393**, 87–174 (2004).

[37] Hallatschek, O. & Nelson, D. R. Gene surfing in expanding populations. *Theoretical Population Biology* **73**, 158–170 (2008).

[38] Roques, L., Garnier, J., Hamel, F. & Klein, E. K. Allee effect promotes diversity in traveling waves of colonization. *Proceedings of the National Academy of Sciences* **109**, 8828–8833 (2012).

[39] Fisher, R. A. The wave of advance of advantageous genes. *Annals of Eugenics* **7**, 355–369 (1937).

[40] Kolmogorov, A. N., Petrovsky, N. & Piscounov, N. S. A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem. *Moscow University Bulletin of Mathematics* **1**, 1 (1937).

[41] Skellam, J. G. Random dispersal in theoretical populations. *Biometrika* 196–218 (1951).

[42] Lewis, M. & Kareiva, P. Allee dynamics and the spread of invading organisms. *Theoretical Population Biology* **43**, 141–158 (1993).

[43] Veit, R. R. & Lewis, M. A. Dispersal, population growth, and the allee effect: dynamics of the house finch invasion of eastern north america. *The American Naturalist* **148**, 255–274 (1996).

[44] Courchamp, F., Clutton-Brock, T. & Grenfell, B. Inverse density dependence and the Allee effect. *Trends in Ecology & Evolution* **14**, 405–410 (1999).

[45] Kramer, A. M., Dennis, B., Liebhold, A. M. & Drake, J. M. The evidence for allee effects. *Population Ecology* **51**, 341 (2009).

[46] Dai, L., Vorselen, D., Korolev, K. S. & Gore, J. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* **336**, 1175–1177 (2012).

[47] Tobin, P. C. *et al.* The role of allee effects in gypsy moth, lymantria dispar (l.), invasions. *Population Ecology* **51**, 373–384 (2009).

[48] Aronson, D. G. & Weinberger, H. G. *Nonlinear diffusion in population genetics, combustion*

and nerve propagation Lectures Notes Math, vol. 446 (Springer, New York, 1975).

[49] Fife, P. C. & McLeod, J. B. The approach of solutions of nonlinear diffusion equations to travelling front solutions. *Archive for Rational Mechanics and Analysis* **65**, 335–361 (1977).

[50] Vlad, M. O., Cavalli-Sforza, L. L. & Ross, J. Enhanced (hydrodynamic) transport induced by population growth in reaction–diffusion systems with application to population genetics. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10249–10253 (2004).

[51] Marculis, N. G., Lui, R. & Lewis, M. A. Neutral genetic patterns for expanding populations with nonoverlapping generations. *Bulletin of Mathematical Biology* 1–25 (2017).

[52] Hallatschek, O., Hersen, P., Ramanathan, S. & Nelson, D. R. Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences* **104**, 19926–19930 (2007).

[53] Korolev, K. S., Xavier, J. B., Nelson, D. R. & Foster, K. R. A quantitative test of population genetics using spatiogenetic patterns in bacterial colonies. *The American Naturalist* **178**, 538–552 (2011).

[54] Meerson, B., Sasorov, P. V. & Kaplan, Y. Velocity fluctuations of population fronts propagating into metastable states. *Physical Review E* **84**, 011147 (2011).

[55] Nullmeier, J. & Hallatschek, O. The coalescent in boundary-limited range expansions. *Evolution* **67**, 1307–1320 (2013).

[56] Hallatschek, O. The noisy edge of traveling waves. *Proceedings of the National Academy of Sciences* **108**, 1783–1787 (2011).

[57] Kingman, J. F. C. The coalescent. *Stochastic processes and their applications* **13**, 235–248 (1982).

[58] Wakeley, J. Coalescent theory. *Roberts & Company* (2009).

[59] Gillespie, J. H. *Population genetics: a concise guide* (JHU Press, 2010).

[60] Neher, R. A., Kessinger, T. A. & Shraiman, B. I. Coalescence and genetic diversity in sexual populations under selection. *Proceedings of the National Academy of Sciences* **110**, 15836–15841 (2013).

[61] Desai, M. M., Walczak, A. M. & Fisher, D. S. Genetic diversity and the structure of genealogies in rapidly adapting populations. *Genetics* **193**, 565–585 (2013).

[62] Pateman, R. M., Hill, J. K., Roy, D. B., Fox, R. & Thomas, C. D. Temperature-dependent alterations in host use drive rapid range expansion in a butterfly. *Science* **336**, 1028–1030 (2012).

[63] Dornelas, M. *et al.* Assemblage time series reveal biodiversity change but not systematic loss. *Science* **344**, 296–299 (2014).

[64] Roman, J. & Darling, J. A. Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution* **22**, 454–464 (2007).

[65] Lee, C. E. Evolutionary genetics of invasive species. *Trends in ecology & evolution* **17**, 386–391 (2002).

[66] Sexton, J. P., McKay, J. K. & Sala, A. Plasticity and genetic diversity may allow saltcedar to invade cold climates in north america. *Ecological Applications* **12**, 1652–1660 (2002).

[67] Dlugosch, K. & Parker, I. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular ecology* **17**, 431–449 (2008).

[68] Shine, R., Brown, G. P. & Phillips, B. L. An evolutionary process that assembles phenotypes through space rather than through time. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 5708–5711 (2011).

[69] Travis, J. M. & Dytham, C. Dispersal evolution during invasions. *Evolutionary Ecology Research* **4**, 1119–1129 (2002).

[70] Phillips, B. L., Brown, G. P., Greenlees, M., Webb, J. K. & Shine, R. Rapid expansion of the cane toad (bufo marinus) invasion front in tropical australia. *Austral Ecology* **32**, 169–176 (2007).
[71] Gray, M. E., Sappington, T. W., Miller, N. J., Moeser, J. & Bohn, M. O. Adaptation and invasiveness of western corn rootworm: intensifying research on a worsening pest. *Annual review of entomology* **54**, 303–321 (2009).
[72] Korolev, K. S. The fate of cooperation during range expansions. *PLoS computational biology* **9**, e1002994 (2013).
[73] Petrovskii, S. V. & Li, B.-L. *Exactly solvable models of biological invasion* (CRC Press, 2005).
[74] Risken, H. *The Fokker-Planck equation: Methods of Solution and Applications* (Springer, Berlin and Heidelberg, 1989).
[75] Gardiner, C. *Handbook of Stochastic Methods* (Springer, New York, 1985).
[76] Gradshteyn, I. & Ryzhik, I. *Table of Integrals, Series, and Products* (Academic Press, New York, 1980).
[77] Constable, G. W. & McKane, A. J. Models of genetic drift as limiting forms of the lotka-volterra competition model. *Physical review letters* **114**, 038101 (2015).
[78] Chotibut, T. & Nelson, D. R. Evolutionary dynamics with fluctuating population sizes and strong mutualism. *Physical Review E* **92**, 022718 (2015).
[79] Blythe, R. A. & McKane, A. J. Stochastic models of evolution in genetics, ecology and linguistics. *Journal of Statistical Mechanics: Theory and Experiment* **2007**, P07018 (2007).
[80] Korolev, K. S., Avlund, M., Hallatschek, O. & Nelson, D. R. Genetic demixing and evolution in linear stepping stone models. *Reviews of modern physics* **82**, 1691 (2010).
[81] Hallatschek, O. & Geyrhofer, L. Collective fluctuations in the dynamics of adaptation and other traveling waves. *Genetics* **202**, 1201–1227 (2016).
[82] Brunet, E. & Derrida, B. Shift in the velocity of a front due to a cutoff. *Physical Review E* **56**, 2597 (1997).
[83] Mikhailov, A. S., Schimansky-Geier, L. & Ebeling, W. Stochastic motion of the propagating front in bistable media. *Physics Letters A* **96**, 453–456 (1983).
[84] Armero, J., Casademunt, J., Ramírez-Piscina, L. & Sancho, J. M. Ballistic and diffusive corrections to front propagation in the presence of multiplicative noise. *Physical Review E* **58**, 5494 (1998).
[85] Rocco, A., Casademunt, J., Ebert, U. & van Saarloos, W. Diffusion coefficient of propagating fronts with multiplicative noise. *Physical Review E* **65**, 012102 (2001).
[86] Balasuriya, S. & Gottwald, G. A. Wavespeed in reaction–diffusion systems, with applications to chemotaxis and population pressure. *Journal of mathematical biology* **61**, 377–399 (2010).
[87] Oksendal, B. *Stochastic differential equations: an introduction with applications* (Springer Science & Business Media, 2013).
[88] Merzbacher, E. *Quantum Mechanics* (Wiley, 1997).
[89] Landau, L. D. & Lifshits, E. M. *Quantum Mechanics Non-relativistic Theory: Transl. from the Russian by JB Sykes and JS bell. 2d Ed.* (Pergamon Press, 1965).
[90] Gillespie, J. *Population genetics: a concise guide* (Johns Hopkins University Press, 2004).
[91] Moro, E. Numerical schemes for continuum models of reaction-diffusion systems subject to internal noise. *Physical Review E* **70**, 1–4 (2004).
[92] Brunet, É., Derrida, B., Mueller, A. H. & Munier, S. Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization. *Physical Review E* **76**, 041104 (2007).