The noisy edge of traveling waves
Supplementary Discussion

Oskar Hallatschek[1,*]

[1] *Biophysics and Evolutionary Dynamics Group*
Max Planck Institute for Dynamics and Self-Organization

(Dated: November 12, 2010)

APPENDIX A: STOCHASTIC DIFFERENTIAL EQUATION

In the main text, we introduced a model of branching random walks that are subject to a global constraint (see Eqs. (1,2,3)). Here, we show that up to higher order terms (h.o.t.), the dynamics of such a constrained branching random walk (CBRW) is described by

$$\begin{aligned} c_{t+\epsilon} - c_t &= \epsilon \left[(\mathcal{L} - 2u)\, c_t - \langle c_t \mid \left(\mathcal{L}^\dagger - 2u\right) u \rangle c_t \right] \\ &\quad + \sqrt{\epsilon} \left[\eta\sqrt{2c_t} - \langle u \mid \eta\sqrt{2c_t}\rangle c_t \right] \; . \end{aligned} \tag{A1}$$

As a corollary, the mean concentration field satisfies

$$\partial_t \bar{c}_t = (\mathcal{L} - 2u)\, \bar{c}_t - \overline{\langle c_t \mid (\mathcal{L}^\dagger - 2u)\, u \rangle c_t} \tag{A2}$$

in the limit $\epsilon \to 0$. Equation (A2) was crucial to our line of arguments in the main text (where it was cited as Eq. 4) as it made evident that the non-linear second term on the right hand side can be eliminated by a suitable choice of the selection function $u(x)$.

In order to prove Eq. (A1), we first combine the branching process Eq. (1) and the selection step Eq. (2) into a single equation of motion, and then eliminate eliminate λ, using the constraint Eq. (3).

Inserting Eq. (1) into Eq. (2) yields

$$c_{t+\epsilon} = (1 - \lambda) \left[c_t + \epsilon \mathcal{L} c_t + \sqrt{2\epsilon c_t}\eta \right] \tag{A3}$$

Now it is important to note that λ only has to be determined up to order $O(\epsilon)$,

$$\lambda = \sqrt{\epsilon}\lambda_s + \epsilon\lambda_d + \text{h.o.t} \; . \tag{A4}$$

The ansatz Eq. (A4) is chosen such that it leads to a Markov process for c_t in continuous time ($\epsilon \to 0$). It can be generally shown that such a process has deterministic $O(\epsilon)$ and stochastic $O(\epsilon^{1/2})$ contributions, and that higher order terms can be neglected in the limit $\epsilon \to 0$ [1]. The components λ_s and λ_d correspond to the stochastic and the deterministic part of the mortality rate, respectively. Inserting Eq. (A4) in Eq. (A3), one obtains

$$\begin{aligned} c_{t+\epsilon} - c_t &= \epsilon \left(\mathcal{L} c_t - \lambda_d c_t - \overline{\lambda_s \eta \sqrt{2c_t}} \right) \\ &\quad + \sqrt{\epsilon} \left(\eta\sqrt{2c_t} - \lambda_s c_t \right) + \text{h.o.t.} \end{aligned} \tag{A5}$$

Next, we eliminate λ using the constraint Eq. (3). To this end, we carry out the inner product of $u(x)$ and Eq. (A5), resulting in

$$\begin{aligned} 0 &= \epsilon \left[\langle u \mid \mathcal{L} c_t \rangle - \lambda_d - \overline{\lambda_s \langle u \mid \eta\sqrt{2c_t}\rangle} \right] \\ &\quad + \sqrt{\epsilon} \left[-\lambda_s + \langle u \mid \eta\sqrt{2c_t}\rangle \right] \end{aligned} \tag{A6}$$

By requiring the $O(\sqrt{\epsilon})$ and $O(\epsilon)$ terms in Eq. (A6) to vanish, one first infers both λ_s and λ_d,

$$\lambda_s = \langle u \mid \eta\sqrt{2c_t}\rangle \tag{A7}$$

$$\lambda_d = \langle u \mid \mathcal{L} c_t \rangle - \langle u \mid \overline{\lambda_s \eta \sqrt{2c_t}} \rangle \; . \tag{A8}$$

The average in the last term is evaluated as

$$\begin{aligned} \overline{\lambda_s \eta \sqrt{2c_t}} &= \overline{\langle u \mid \eta\sqrt{2c_t}\rangle \eta(x) \sqrt{2c_t}(x)} \\ &= 2 \int_{x'} u(x') \sqrt{c_t(x') c_t(x)} \; \overline{\eta(x)\eta(x')} \\ &= 2u(x) c_t(x) \; . \end{aligned} \tag{A9}$$

Note that underlying reason for obtaining a term linear in c_t for the average in Eq. (A9) is the fact that the noise is $\propto c^{1/2}$, which is ultimately a manifestation of the law of large numbers.

Inserting Eq. (A9) into Eq. (A8), we can summarize λ_d as

$$\lambda_d = \langle u \mid (\mathcal{L} - 2u)\, c_t \rangle \; . \tag{A10}$$

The stochastic dynamics Eq. (A1) of an entire time step now follows from inserting our expressions Eqs. (A7, A10) for λ_s and λ_d into Eq. (A5), resolving the average $\overline{\lambda_s \eta \sqrt{2c_t}}$ via Eq. (A9) and, finally, integration by parts:

$$\langle u \mid (\mathcal{L} - 2u)\, c_t \rangle = \langle c_t \mid \left(\mathcal{L}^\dagger - 2u\right) u \rangle \; . \tag{A11}$$

APPENDIX B: TIME DEPENDENT PROBLEMS

In the main text, we focused on branching random walks with time-independent generators $\mathcal{L}$. Here, we explicate how our formalism can be generalized to time dependent problems with a dynamic generator $\mathcal{L}(t)$. Interesting time-dependent problems include traveling waves in fluctuating environments or transients within fixed environments (e.g. How quickly does a population approach a fitness peak?).

*Electronic address: oskar.hallatschek@ds.mpg.de

For this generalization, we have to allow the selection function $u_t(x)$ as well as the "gauge" variable $x_0(t)$ to become time-dependent.

The fundamental time step of the time-dependent model may be summarized as follows: Suppose we have the concentration field $c_t(x)$ after selection at time t,

$$\langle u_t \mid c_t\rangle = 1 \; . \tag{B1}$$

To generate $c_{t+\epsilon}$, we first generate a branching random walk,

$$\tilde{c}_{t+\epsilon} - c_t = \epsilon \mathcal{L}_t c_t + \sqrt{2\epsilon c_t}\eta \; . \tag{B2}$$

The operator $\mathcal{L}_t$, shall contain the time-dependent gauge $x_0(t)$, which is adjusted later on to fix the mean population size,

$$\mathcal{L}_t = \mathcal{L}_t^0 - x_0(t) \; . \tag{B3}$$

The selection step now has to satisfy a slightly different constraint

$$\langle u_{t+\epsilon} \mid c_{t+\epsilon}\rangle = 1 \; . \tag{B4}$$

As in the main text, the constraint is enforced by a "fair" culling of a population fraction λ,

$$c_{t+\epsilon} = \tilde{c}_{t+\epsilon}(1-\lambda) \; . \tag{B5}$$

We next determine the stochastic PDE for c_t in a very similar way than in previous section for time-independent problems. The only new bits come from the fact that u_t is time-dependent. We require u_t to change deterministically,

$$u_{t+\epsilon} = u_t + \epsilon \partial_t u_t + o(\epsilon) \; , \tag{B6}$$

i.e. u_t has no stochastic component.

It can then be shown that the stochastic part and deterministic parts of λ are given by

$$\begin{aligned} \lambda_s &= \langle u_t \mid \eta\sqrt{2c_t}\rangle && \text{(B7)} \\ \lambda_d &= \langle c_t \mid \partial_t + \mathcal{L}_t^\dagger - 2u_t \mid u_t\rangle \; . && \text{(B8)} \end{aligned}$$

Notice that the stochastic λ_s is the same as in the time-independent case, however, the deterministic λ_d has acquired a time derivative acting on u_t.

The dynamics from c_t to $c_{t+\epsilon}$ may thus be summarized as

$$\begin{aligned} c_{t+\epsilon} - c_t &= \epsilon \left[\left(\mathcal{L}_t - 2u\right) c_t - \langle c_t \mid \left(\partial_t + \mathcal{L}_t^\dagger - 2u_t\right) u_t \rangle c_t \right] \\ &\quad + \sqrt{\epsilon}\left[\eta\sqrt{2c_t} - \langle u \mid \eta\sqrt{2c_t}\rangle c_t\right] \; . \end{aligned} \tag{B9}$$

The mean concentration field obeys

$$\partial_t \bar{c}_t = \left(\mathcal{L}_t - 2u\right)\bar{c}_t - \overline{\langle c_t \mid \left(\partial_t + \mathcal{L}_t^\dagger - 2u_t\right) u_t\rangle c_t} \tag{B10}$$

In order to obtain a solvable model, we apply (as in the main text) the trick to choose a selection function $u = u_*$ such that the non-linearity in the moment equation disappears,

$$\partial_t \bar{c}_t = \left(\mathcal{L}_t - 2u_*\right)\bar{c}_t \; . \tag{B11}$$

To this end, we have to choose

$$-\partial_t u_* = \left(\mathcal{L}_t^\dagger - 2u_*\right) u_* \; . \tag{B12}$$

Role of $x_0(t)$: Until now, the function $x_0(t)$ was entirely arbitrary. For each choice of $x_0(t)$, we obtain a different solvable model via Eqs. (B11, G4). These different solutions in general exhibit different expected population sizes. If we are interested in a solution with a constant mean population size $\overline{n}$, $x_0(t)$ has to be fixed such that

$$0 = \partial_t \overline{n} = \int_x \partial_t \bar{c}_t \; . \tag{B13}$$

Using Eq. (B11), this implies

$$\begin{aligned} 0 &= \int_x \left(\mathcal{L}_t - 2u_*\right)\bar{c}_t && \text{(B14)} \\ &= \int_x \left[\mathcal{L}_t^0 - x_0(t)\right]\bar{c}_t - 2 && \text{(B15)} \\ &= \left[\int_x \mathcal{L}_t^0 c_t\right] - x_0(t)\overline{n} - 2 \; . && \text{(B16)} \end{aligned}$$

In going from the first to the second line of the previous equation, we used the constraint $1 = \int_x u_* \bar{c}$, and inserted expression Eq. (D1) for $\mathcal{L}_t$.

If we insert, for instance, the operator $\mathcal{L}_{\text{evo}}$ from Eq. (7), we see that $x_0(t)$ is given by,

$$x_0(t) = -2\overline{n}^{-1} + \int_x x\bar{c}_t \; , \tag{B17}$$

which essentially represents the time dependent mean fitness for large population sizes.

APPENDIX C: INTERPRETATION OF THE TUNED MODEL

Among the three graphs depicted in Fig. 3 (main text), only the function $\bar{c}(x)$ has an obvious interpretation as the fitness wave profile. Using the results obtained in Ref. [2], one can also give an intuitive interpretation of the functions $u_*(x)$ and $g(x) \equiv u_*\bar{c}$. Both functions relate to the phenomenon of fixation. Imagine sampling an individual at position x and labeling it with an inheritable label (neutral mutation). As the dynamics proceeds, the abundance of this label will change due to number fluctuations and the fitness of its carriers. Eventually, this label will either go extinct, or become *fixed* in the population. The latter case occurs if the descendants of the labeled individual take over the population.

Fixation events are much more likely if the initially labeled individual belongs to the fitter part of the population. We thus expect the probability of fixation to be a steeply increasing function of x, similar to the function $u_*(x)$. Indeed, the interpretation of $u_*(x)$ is precisely that of a fixation probability of a particle at position x, which is derived below. It is interesting to note about the defining equation Eq. (5) of $u_*(x)$, that a very similar equation is well-known to describe the survival probability of an *unconstrained* random walk [3]. The only difference is the pre-factor of 2 instead of 1 in the non-linearity of Eq. (5).

The product $g(x) \equiv u_* \overline{c}$ also has an interesting interpretation in terms of a probability density. (Note that, as required for a probability density, the integral of $g(x)$ over x is equal to 1 by virtue of the constraint Eq. (3).) $g(x)$ represents the positional distribution function of the lucky particle whose descendants eventually will take over the population. Even though there must surely be a lucky one at any time, we cannot pinpoint to this individual, simply because the fixation event depends on random future events. Thus, the position of the "common ancestor of future generations" can only be described probabilistically, similar to the position of quantum mechanical particle. Also similar to quantum mechanics, the positional distribution function is the product of a bra and a ket. In fact the expression for $g(x)$,

$$g(x) = \frac{\overline{c}^2 e^{vx/D}}{\int_x \overline{c}^2 e^{vx/D}} \tag{C1}$$

has been hypothesized earlier in Ref. [2] based on a mean-field treatment of the fixation process inside traveling waves.

Reasoning along the lines of Ref. [2] allows us to justify the above interpretation of the function u_* as a fixation probability. Suppose, we attach inheritable neutral markers at time τ to a part of the population. Let us denote the concentration field corresponding to the labeled sub-population by $c_l(x,t)$, which has to be smaller than the total concentration field, $c(x,t)$, at any point. Then, the mean concentration field $\overline{c}_l(x,t)$ of labeled individuals will satisfy the same linear evolution equation Eq. (B11) as the total concentration field does, namely,

$$\partial_t \overline{c}_l(x,t) = (\mathcal{L}(t) - 2u_*)\,\overline{c}_l \; . \tag{C2}$$

Since the system is comprised of a finite number of particles, it is inevitable that the labeled sub-population eventually either goes extinct or takes over the whole population. Thus, with a certain fixation probability F, the labeled population will approach the total population $c_l \to c(x,t)$ on long times. On the other hand, $c_l \to 0$ with a probability $1-F$. This implies that the average of our labeled concentration field approaches

$$\overline{c}_l(x,t) \sim F\,\overline{c}(x,t) \qquad t \to \infty \; . \tag{C3}$$

The fixation probability F is a linear functional of the initial concentration field $c_l(\xi,\tau)$ of the labeled sub-population. In fact, we will show that F is given by

$$F = \int_\xi c_l(\xi,\tau) u_*(\xi,\tau) \tag{C4}$$

justifying our interpretation $u_*(\xi,\tau)$ as the fixation probability of a particle present at position ξ at time τ. As a sanity check, we notice that, due to the constraint Eq. (3), Eq. (C4) correctly predicts a fixation probability of $F=1$ when the whole population is labeled, $c_l = c$.

In order to proof Eq. (C4), we first express the solution to equation Eq. (C2) as

$$\overline{c}_l(x,t) = \int_\xi G(x,t;\xi,\tau) c_l(\xi,\tau) \tag{C5}$$

in terms of a Green's function $G(x,t;\xi,\tau)$, which is the solution of Eq. (C2) corresponding to a δ-function initial condition,

$$G(x,\tau;\xi,\tau) = \delta(x-\xi) \; . \tag{C6}$$

On long times, we may argue as above that G becomes proportional to $\overline{c}(x,t)$. However, viewed as a function of ξ and τ, one can also show that G must become proportional to $u_*(\xi,\tau)$. This follows from the fact that the Green's function not only satisfies the forward time equation Eq. (C2) but also a corresponding backward equation

$$-\partial_\tau G = \left(\mathcal{L}^\dagger(t) - 2u_*\right) G \; , \tag{C7}$$

where the operator on the right-hand-side acts on ξ. Incidentally, Eq. (C7) is the equation satisfied by $u_*(\xi,\tau)$, see Eq. (B14). Thus, for large time differences $t-\tau$, $G(x,t;\xi,\tau)$ must be proportional to both $u_*(x,t)$ and $\overline{c}(x,t)$.

$$G(x,t;\xi,\tau) \sim A u_*(\xi,\tau)\overline{c}(x,t) \qquad t\to\infty \; . \tag{C8}$$

Furthermore, the pre-factor A follows from the following computation

$$\begin{aligned}
\overline{c}(x,t) &= \int_\xi G(x,t;\xi,\tau) c(\xi,\tau) && \text{(C9)}\\
&= A\overline{c}(x,t) \int_\xi u_*(\xi,\tau) c(\xi,\tau) && \text{(C10)}\\
&= A\overline{c}(x,t) \; . && \text{(C11)}
\end{aligned}$$

Evidently, A has to be equal to 1. The first line in Eq. (C9) follows from the definition of the Green's function. In the second line, we inserted Eq. (C8) for G. In going from the second to the third line, we used the constraint Eq. (3).

Finally, our claim Eq. (C4) now follows from inserting Eq. (C8) with $A=1$ into Eq. (C5) and using Eq. (C3).

APPENDIX D: APPLICATION TO NOISY FISHER–KOLMOGOROV WAVES

In this section, we use our recipe of obtaining constraint branching random walk (CBRW) models with closed moment equations to study the important class of stochastic Fisher-Kolmogorov-Petrovsky-Piskunov (sFKPP) waves. To this end, we consider the Liouvillean

$$\mathcal{L}_{\mathrm{sFKPP}} = D\partial_x^2 + s\Theta(x) \; . \tag{D1}$$

The unit step function could equally be replaced by any other function, which saturates at the value 1 for large x. Such models control the number of particles that are in the tip of the wave by modulating the growth rates.

The moment equation of the CBRW model corresponding to $\mathcal{L}_{\mathrm{sFKPP}}$ is described by Eqs. (5, 9) and the constraint Eq. (3), which can be combined to give

$$0 = (D\partial_x^2 + s\Theta(x) + v\partial_x)\overline{c} - \frac{2\overline{c}^2 e^{vx/D}}{\int_x \overline{c}^2 e^{vx/D}} \; . \tag{D2}$$

As an aside, the relation between $\overline{c}$ and u_* requires that the denominator in the last term of Eq. (D2) is finite. This breaks down, when the speed parameter v is larger than the deterministic Fisher wave speed $v = 2\sqrt{Ds}$, which is the maximal achievable wave speed.

Let us introduce non-dimensional quantities

$$\begin{aligned} X &= x/\sqrt{D/s} && \text{(D3)}\\ V &\equiv v/\sqrt{Ds} && \text{(D4)}\\ \overline{C} &\equiv \overline{c}/\overline{N} \; . && \text{(D5)} \end{aligned}$$

In terms of these variables, the moment equation reads

$$0 = (\partial_X^2 + \Theta(X) + V\partial_X)\overline{C} - \frac{2\overline{C}^2 e^{VX}}{N_e \int_X \overline{C}^2 e^{VX}} \; , \tag{D6}$$

where we have introduced an "effective" population size

$$N_e \equiv \overline{N}\sqrt{Ds} \; . \tag{D7}$$

Notice that, for large N_e, the last non-linear term in the PDE in Eq. (D6) is relevant only for large $x \sim O(\ln N_e)$, where it becomes of order $O(1)$ due to the exponential. The net-effect of the nonlinearity may be mimicked by a cutoff in the growth rate at some position L, which we determine up to the order $O(\ln\ln N)$ in the following.

The position L of the effective cutoff can be determined by balancing the non-linearity in Eq. (D6) with the reaction term $\overline{C}\Theta(X)$ in Eq. (D6),

$$2\overline{C}e^{VL} \sim N_e \int_x \overline{C}^2 e^{VX} \; . \tag{D8}$$

To evaluate this condition, we need to refer to some known results on deterministic Fisher waves with a cutoff [4]: Their speed is, to leading order, given by

$$V \sim 2 - \frac{\pi^2}{L^2} \; . \tag{D9}$$

The wave profile of Fisher waves with cutoff satisfies for $0 \ll X \ll L$

$$\overline{C}e^{VX} \sim A\frac{L}{\pi}\sin \pi X/L \; , \tag{D10}$$

where A is an L (and N_e) independent constant. These results from the cutoff approach are summarized, for instance, in Ref. [4].

Inserting Eq. (D10) and Eq. (D9) into Eq. (D8) yields

$$e^L \sim N_e L^3 \tag{D11}$$

in the limit of large L. To evaluate the integral on the right hand side in Eq. (D8), we used that the contribution from $x > L$ of higher order. Solving Eq. (D11) for L, we find

$$L \sim \ln N_e + 3\ln\ln N_e \tag{D12}$$

up to the second leading order. The wave speed in the presence of a cutoff at position L now follows from Eq. (D9) as

$$v - v_F \sim \frac{\pi^2}{L^2} = \frac{\pi^2}{(\ln N)^2}\left(1 - \frac{3\ln\ln N}{\ln N}\right) \; . \tag{D13}$$

The mean-field theory augmented by a heuristic cutoff generates the first correction term. The second correction, however, requires a more subtle reasoning, which could so far only be based on heuristic assumptions[4]. For our CBRW model, the second leading order correction follows naturally from the $\ln\ln N$ correction to the position of the cutoff. This correction in L is, ultimately, the consequence of the non-local denominator of the cutoff term in Eq. (D6).

APPENDIX E: FLUCTUATIONS

As explained in the main text, most noisy traveling waves have fluctuating speeds and a constraint associated with population sizes. The fluctuations in the speed of the wave can be summarized by a wave diffusion constant D_{wave}.

Noisy waves with tuned non-linearities resemble a different statistical ensemble of the same problem: the wave speed is perfectly constant but the the mortality rate λ fluctuates to keep up with the constraint. As explained below, a diffusion constant D_λ can be associated with these fluctuations as well. In the limit of large population sizes, where both statistical ensembles seem to mirror the same universal dynamics, we expect both diffusion constants to be proportional to each other. We use this equivalence hypothesis to conjecture from an exact expression for D_λ the asymptotic diffusivity of Fisher waves, which is known, as well as the diffusivity of evolutionary waves, which represents a novel prediction.

1. Fluctuations in the mortality rate

In our model of constrained branching random walks (CBRWs), the mortality rate is on average constant,

$$\overline{\lambda} = \overline{\lambda_d} = \overline{\langle c_t | \left(\mathcal{L}^\dagger - 2u_*\right) u_* \rangle} \; . \tag{E1}$$

In fact, the deterministic component λ_d vanishes at any time. However, λ has a fluctuating component, λ_s, to which we can associate a diffusion constant by

$$D_\lambda \equiv \int_\tau \overline{\lambda(t)\lambda(t+\tau)}/2 = \overline{\lambda_s^2}/2 \tag{E2}$$

Inserting λ_s from Eq. (A7) and carrying out the average yields

$$D_\lambda = \langle u_* \mid u_* \overline{c} \rangle \, . \tag{E3}$$

If particles diffuse, u_* can be expressed in terms of $\overline{c}$ as in Eq. (9), so that we obtain

$$D_\lambda = \frac{\int_x \overline{c}^3 e^{+2vx/D}}{\left(\int_x \overline{c}^2 e^{+vx/D}\right)^2} \, . \tag{E4}$$

2. Wave diffusion

Fluctuations in λ of the tuned model have a different source than wave speed fluctuations of the constant population size model. Nevertheless, we conjecture that the amplitude of both fluctuations are proportional to each other

$$D_{\text{wave}} \sim 4D_\lambda (D/v)^2 \, , \tag{E5}$$

in the limit of large population sizes. Our argument for this conjecture is based on the observation that both types of fluctuations have very similar effects in the asymptotic limit. In the equation of motion, a change Δv in wave speed results in a term $+\Delta v \partial_x c$. This term approaches $2c\Delta v/v$ for large wave speeds, where the wave profile assumes

$$c \sim \varphi(x) e^{-vx/2D} \tag{E6}$$

with a slowly varying function $\varphi(x)$ in the bulk of the wave front. Thus, a change Δv acts like a change $\Delta\lambda \sim 2\Delta v/v$ in the mortality rate. This suggests that the asymptotic fluctuations in the speed, as measured by D_{wave}, are related to D_λ by Eq. (E5).

We would like to remark that our argument is based on the crucial assumption that the amplitude of the fluctuations in the mortality rate λ exhibits universal scaling for large population sizes notwithstanding the particular choice of $u(x)$. This allows us to choose the (unique) selection function $u_*(x)$, for which the mortality rate λ becomes uncorrelated in time. For any other choice of the selection function, λ exhibits temporal correlations (via its deterministic part λ_d) similar to the wave profile $c(x,t)$ and the wave speed in standard noisy traveling waves.

3. Fisher wave diffusivity

Next, we use our conjecture Eq. (E5) to determine the diffusion constant of noisy Fisher waves. To this end, we evaluate the integrals in Eq. (E4) using Eqs. (D9, D10) and Eq. (D12),

$$\begin{aligned}
\int_x \overline{c}^2 e^{+vx/D} &= \sqrt{\frac{D}{s}} N_e^2 \int_X \overline{C}^2 e^{+VX} && \text{(E7)}\\
&\sim N_e^2 \int_0^L dx \, (A\pi L \sin \pi x/L)^2 \\
&\sim N_e^2 L^3 \sim N_e^2 \ln^3 N_e \\
\int_x \overline{c}^3 e^{+2vx/D} &= \sqrt{\frac{D}{s}} N_e^3 \int_X \overline{C}^3 e^{+2VX} && \text{(E8)}\\
&\sim N_e^3 \int_0^L dX e^{VX/2} \, (A\pi L \sin \pi X/L)^3 \\
&\sim N_e^3 e^{VL/2} A^3 \int_0^\infty dX X^3 e^{-VX/2} \\
&\sim N_e^3 e^L \\
&\sim N_e^4 \ln^3 N_e \, .
\end{aligned}$$

Notice that the integral in Eq. (E7) has support throughout the bulk of the wave front, while the one in Eq. (E8) is peaked close to the cutoff. Replacing the integrals in Eq. (E4) by the above expressions and using conjecture Eq. (E5), we finally obtain

$$D_{\text{wave}} \sim \ln^{-3} N_e \tag{E9}$$

to leading order in $\ln N$. This prediction is in agreement with extensive simulations and a phenomenological theory reported in Ref. [4], which supports our conjecture Eq. (E5).

4. Diffusivity of evolutionary waves

We now proceed in the same way as above to conjecture the (yet unknown) diffusion constant of evolutionary waves.

Applying the cutoff approach to the equation of motion Eq. (10) along the lines of Ref. [5], one derives that the asymptotic behavior of the bulk density profile can be expressed in terms of the Airy function as

$$ce^{vx/2D} \sim v^{-3} e^{vL/2D} \text{Ai}(L + a_0 - x) \tag{E10}$$

where $\text{Ai}(x)$ is the Airy function with its first zero at $a_0 \approx -2.7781$, and

$$L \sim (v/2D)^2 - a_0 \tag{E11}$$

is the position of the cutoff (up to $O(2D/v)$). With the wave profile at hand, we can determine the asymptotics of the integrals appearing in Eq. (E4),

$$\begin{aligned}\int_x \overline{c}^2 e^{+vx/D} &= e^{vL/D} v^{-6} \int_0^L dx \mathrm{Ai}(L+a_0-x)^2 \\ &\sim e^{vL/D} v^{-6} \qquad \text{(E12)}\\ \int_x \overline{c}^3 e^{+2vx/D} &\sim v^{-9} e^{3vL/2D} \int_0^L dx e^{vX/2D} \mathrm{Ai}(L+a_0-x)^3 \\ &\sim e^{2vL/D} v^{-9} \int_0^\infty dx x^3 e^{-vx/2D} \\ &\sim e^{2vL/D} v^{-13} \; . \qquad \text{(E13)}\end{aligned}$$

Evaluating Eq. (E4) with the help of Eqs. (E12, E13) and using the conjecture Eq. (E5), we obtain

$$D_{\mathrm{wave}} \sim v^{-3} \; . \qquad \text{(E14)}$$

Note that this implies the scaling $D_{\mathrm{wave}} \sim \ln^{-1} N$ between diffusion constant and population size, since $v \sim \ln^{1/3} N$. The power law in Eq. (E14) is expected to hold only for large values of the control parameter $vD^{-2/3}$ introduced in the main text. For small speeds, on the other hand, we expect the diffusion constant to saturate, $D_{\mathrm{wave}} \to 1$.

APPENDIX F: DISCRETE FITNESS EFFECTS

For simplicity, our example of asexual adaptation was formulated assuming a continuous random walk of genotypes with diffusion constant D. For a given speed of adaptation v, this is only appropriate if

$$\frac{v\Delta x}{D} \ll 1 \; . \qquad \text{(F1)}$$

Intuitively, this condition comes about because of the exponential decay $\exp -vx/D$ of the frequency distribution in its tail. If Eq. (F1) is violated, the relative change in the frequency distribution from one fitness value to the next is of order 1, by which the continuous approximation becomes questionable. Assuming our asymptotic large N result $v \sim D^{2/3} \ln[ND^{1/3}]^{1/3}$, Eq. (F1) reads,

$$\Delta x D^{-1/3} \ln[ND^{1/3}]^{1/3} \ll 1 \; . \qquad \text{(F2)}$$

More concretely, we may assume that mutations have fitness effect $\Delta x = s$ and occur at a rate m. Then, the diffusion constant is given by $D \sim ms^2/2$. Inserting these expressions in Eq. (F2 yields

$$\left(\frac{s}{m}\right)^{1/3} \ln[N(ms^2)^{1/3}]^{1/3} \ll 1 \; . \qquad \text{(F3)}$$

If we are interested in the large N limit where the logarithm is larger then 1, we have thus to require $m \gg s$. At least for bacterial populations, this conditions seems to be violated.

We are thus looking for a description for the CBRW model using a discrete fitness lattice. To this end, we assume that jumps occur only between nearest neighbors; A is the jump rate $i \to i + \Delta x$, and B is the rate for the reverse jump. Furthermore, we suppose to have scaled time such that the effective diffusion constant $D = (A+B)\Delta x^2/2 = 1$. For this model, the u_* equation is obtained by using the discrete Laplacian

$$\partial_x^2 u_* \to A u_{*,i-1} - (A+B) u_{*,i} + B u_{*,i+1} \; , \qquad \text{(F4)}$$

where we introduced the notation $f_i \equiv f(i\Delta x)$. The relation between mean density and u_* changes to

$$\overline{c}_i = u_{*,i} \left(\frac{A}{B}\right)^i / \Gamma \; . \qquad \text{(F5)}$$

The proportionality constant follows from the constraint,

$$\Gamma = \sum_i u_{*,i}^2 \left(\frac{A}{B}\right)^i \; . \qquad \text{(F6)}$$

Solving the discrete model thus requires, in general, the solution of a nonlinear difference equation. Alternatively, one may look at the logarithm of c and solve a continuous equation, as it is done in the WKB method.

APPENDIX G: EVOLUTION OF LARGE POPULATIONS WITH CONSTANT FITNESS

In contrast to the adapting populations considered previously, we are now considering populations that have already “found” a (local) peak in the fitness landscape. We ask how one can describe the distribution growth rates in this quasi-steady state. In other words, we develop the basic frame work to calculate allele frequency distributions for large populations, in which the mean fitness is stationary.

For x =const., the selection function $u_*(x)$ satisfies

$$-\partial_t u_* = \left(\mathcal{L}_{\mathrm{evo}}^\dagger - x_0 - 2u_*\right) u_* = 0 \qquad \text{(G1)}$$

at stationarity. Since $u_*(x)$ has the interpretation of survival probabilities (Sec. C), it should be intuitively clear that $u_*(x)$ should uniformly go to 0 as the population size is send to infinity. Thus, the mean fitness x_0^∞ in an infinite population can be found by neglecting the nonlinearity,

$$\left(\mathcal{L}_{\mathrm{evo}}^\dagger - x_0^\infty\right) u_* = 0 \; . \qquad \text{(G2)}$$

We would like to remark that this equation has no solution for the model Eq. 7 of continually adapting populations (main text). In this case, the perturbation brought about by the fluctuations of a finite population was a singular perturbation. When the population is at a fitness peak, however, Eq. (G2) has a solution, and one may carry out a basic perturbation analysis to investigate the leading order effects of sampling noise. To this end, let us define

$$\mathcal{L}_\infty = \mathcal{L}_{\mathrm{evo}} - x_0^\infty \qquad \text{(G3)}$$

and study a population with slightly smaller mean fitness $x_0 = x_0^\infty - \delta$ ($\delta \ll 1$) than the mean fitness of an infinite one. In this case, we may expand

$$u_* = \delta \sum_n a_n f_n \tag{G4}$$

where f_n is the nth left eigenvector of the unperturbed $\mathcal{L}_\infty$

$$\mathcal{L}_\infty^\dagger f_n^L = \lambda_n f_n^L \ . \tag{G5}$$

Suppose, we have ordered the eigenvalues according to their magnitude. Then $\lambda_0 = 0$ by virtue of Eq. (G3). Performing the inner product of f_n^L and

$$0 = (\mathcal{L}_\infty^\dagger - \epsilon - u_*)u \tag{G6}$$

shows that to leading order in ϵ,

$$a_{n>0} = 0 \tag{G7}$$

$$a_0 = \frac{\epsilon}{2}\langle f_0 \mid f_0^2\rangle^{-1} \ . \tag{G8}$$

To leading order, we thus have

$$u_* = \frac{\epsilon f_0}{2\langle f_0 \mid f_0^2\rangle} \ . \tag{G9}$$

The mean population density $\overline{c} = b_0 f_0^R$ will accordingly be proportional to the right eigenvector f_0^R of $\mathcal{L}_\infty$ corresponding to eigenvalue 0. The pre-factor b_0 follows from the constraint, which reads

$$1 = \langle u \mid \overline{c}\rangle = \epsilon b_0 \frac{\langle f_0^L \mid f_0^R\rangle}{2\langle f_0^L \mid (f_0^L)^2\rangle} \ . \tag{G10}$$

Therefore,

$$\overline{c} = \frac{2\langle f_0^L \mid (f_0^L)^2\rangle}{\epsilon\langle f_0^L \mid f_0^R\rangle} f_0^R \tag{G11}$$

The mean total population size is given by

$$\overline{N} = \frac{2\langle 1 \mid f_0^R\rangle\langle f_0^L \mid (f_0^L)^2\rangle}{\epsilon\langle f_0^L \mid f_0^R\rangle} \ . \tag{G12}$$

[1] van Kampen NG (2001) *Stochastic processes in physics and chemistry* (Elsevier Science, Amsterdam).

[2] Hallatschek O, Nelson DR (2008) Gene surfing in expanding populations. *Theor. Pop. Biol.* 73:158–170.

[3] Derrida B, Simon D (2007) The survival probability of a branching random walk in presence of an absorbing wall. *Europhys. Lett.* 78:60006.

[4] Brunet E, Derrida B, Mueller AH, Munier S (2006) Phenomenological theory giving the full statistics of the position of fluctuating pulled fronts. *Phys. Rev. E* 73:056126.

[5] Cohen E, Kessler DA, Levine H (2005) Front propagation up a reaction rate gradient. *Phys. Rev. E* 72:066126.