Run and tumble navigation can be very fast

J. Long$^{1,2}$, S. W. Zucker$^{3,4}$, and T. Emonet$^{*2,1}$

$^1$Department of Physics, Yale University, New Haven, CT 06520, USA.
$^2$Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520, USA.
$^3$Department of Computer Science, Yale University, New Haven, CT 06520, USA.
$^4$Department of Biomedical Engineering, Yale University, New Haven, CT 06520, USA.

May 11, 2016

Abstract

The ability to navigate environmental gradients is often critical for survival. When gradients are shallow or noisy, most organisms move by alternating straight motion (runs) with random reorientations (tumbles). Navigation is achieved by transiently reducing the probability to tumble when attractant signal increases. One drawback of this strategy is that occasional runs in the wrong direction reduce progress up the gradient. Here we discovered a positive feedback regime inherent in this strategy that strongly mitigates this problem. In an attractant field, motion up the gradient reduces tumble probability, which further boosts drift up the gradient. This positive feedback can drive large fluctuations in the internal state of the organism away from its mean, resulting in long runs in favorable directions but short ones otherwise. In this new regime the organism achieves a "ratchet-like" gradient climbing behavior unexpected from mean field theory, and drift speeds much faster than previously believed possible.

*Email: thierry.emonet@yale.edu.
Introduction

Navigating gradients requires detecting the direction in which signals of interest increase. When the size of the navigator is small compared to the length scale of the gradient, local directional information becomes unreliable. In this case, cells\textsuperscript{1–4}, worms\textsuperscript{5,6}, larvae\textsuperscript{7}, and robots\textsuperscript{8}, adopt run-and-tumble strategies similar in many ways to the well characterized chemotactic behavior of \textit{Escherichia coli}. During runs the organism moves approximately straight at a constant speed, which enables it to collect information about the direction of the gradient. During tumbles the speed is nil while the organism reorients randomly to be able to explore a new direction (Fig. 1a). Rapid transduction of the signal to the motility apparatus is accomplished via an internal state variable, $F$, which in turn controls the probability to run $r(F)$. Following an increase in attractant the probability to run increases transiently, then adapts back due to a negative feedback (Fig. 1b). Using this approach an organism can perform a biased “random walk” by lengthening runs when it moves in the desired direction, and progress is made by averaging motion over many runs and tumbles. The efficiency of this process depends on the structure of the environment as perceived by the organism, the memory length of the organism, and the physical constraints on the movement of the organism.

We encapsulate these factors into three timescales to reveal a diversity of dynamical regimes (Fig. 1). The first time scale of interest is that of the change in perceived signal. For simplicity we will consider static environments, in which change in perceived signal is solely due to the motion of the organism (Fig. 1a). We define the timescale of the perceived environment, $t_E$, as the perceived gradient length scale divided by the run speed. We will show that $t_E$ varies inversely with the strength of the positive feedback. The second important timescale is the duration of the memory of the organism, $t_M$, that corresponds to the relaxation time of the adaptation mechanism (Fig. 1b). We will show that $t_M$ varies inversely with the strength of the negative feedback mediated by adaptation. The interplay between these two feedbacks — positive and negative — will be instrumental in determining the dynamics of gradient tracking. Finally, the direction decorrelation time, $t_D$, quantifies how long the organism can move before losing its original direction due to tumbling or other physical constraints such as rotational Brownian motion (Fig. 1a). In effect, $t_D$ limits the usable duration of runs and memory. From these three timescales one can define the ratios $\tau_E = t_E/t_M$ and $\tau_D = t_D/t_M$. The first one quantifies the relative strength of the negative and positive feedback. The second one measures how well the memory length $t_M$ matches the direction decorrelation time $t_D$.

Classical studies by Keller & Segel and others\textsuperscript{9–19} considered what happens when adaptation is fast ($\tau_E \gg 1$). In this case the negative feedback dominates the dynamics. The internal state of the organism $F$ exhibits small fluctuations around its mean $\langle F \rangle$, which in turn barely deviates from $F_0$, the adapted value of $F$ in the absence of gradient. When $\tau_E \gtrsim 1$, the negative feedback is not as strong but still dominates over the positive feedback, keeping the fluctuating value of $F$ close to its mean. However, $\langle F \rangle$ now deviates from $F_0$ because adaptation is not fast enough to fully compensate the average rate of increase of the signal introduced by the average drift up the gradient\textsuperscript{20,21}. Related phenomena were beautifully demonstrated by stimulating cells with exponential ramps\textsuperscript{22,23} and sinusoidal signals\textsuperscript{24}. When the negative feedback dominates, motion resembles a biased random walk, with a moderate drift speed resulting from averaging over many runs of similar length, several of which are oriented in the wrong direction.

In this paper we show that there is another regime of the run-and-tumble gradient ascent strategy wherein much higher drift velocities can be achieved. This regime takes place when the positive feedback dominates over the negative feedback, $\tau_E \ll 1$. A significant difference from the classical studies emerges because the distribution of the internal state $F$ develops a long tail, which precludes the use of mean field approaches. The gradient-climbing strategy becomes highly effective so that, in a sense, it resembles a ratchet: the organism cruises along favorable directions and quickly turns away from unfavorable ones. Performance is
**Figure 1 | Different dynamical regimes of bacterial chemotaxis.**  
(a) Run and tumble trajectory $X(t)$ of a cell climbing up a gradient of attractant $\nabla C$. $\theta$ is the angle between the direction of motion and that of the gradient. During runs (tumbles), speed is $v_0$ (zero) and rotational diffusion is $D_R$ ($D_T$), resulting in the direction decorrelation time $t_{D}$. 
(b) Input, internal, and behavioral variables: a sudden step in the external concentration $\Delta C$ is amplified by the receptors into a change $\sim N\Delta C/C$ in their free energy $F$ above the adapted level $F_0$ before the negative feedback mechanism brings it back at rate $t^{-1}_M$. The motors further amplify the signal by a factor $H$ to cause a change $\sim r(1-r)NH\Delta C/C$ in the probability to run. 
(c) Drift velocity $V_D$ of simulated *E. coli* cells swimming in static exponential gradients as a function of $\tau_E = t_E/t_M$ and $\tau_{D0} = t_D(t_0)/t_M$. Black dashed line: horizontal: $\tau_{D0} = 1$ fixed while $\tau_E = 0.1$ (green dot), $\tau_E = 1$ (blue dot), and $\tau_E = 3$ (red dot); vertical: $\tau_E = 0.1$ fixed while $\tau_{D0} = 1$ (green dot) and $\tau_{D0} = 0.1$ (orange dot). White dots: sampling of realistic parameters for 16,000 individual *E. coli* cells from a wild type isogenic population near the bottom of the gradient. Black dots: the same cells near the top of the gradient. Black/white contour lines: same as the black/white dots but for $D_R$ and $D_T$ 8 times smaller, i.e. assuming cells are twice as long (see Methods). 
(d) Main result: classical (red) vs. rapid climbing (green) trajectories. 5 sample trajectories (thin solid curves) and the mean over $10^4$ sample trajectories (thick lines) of position $x = X/(v_0 t_M)$ as a function of time $\tau = t/t_M$ for cells in a positive-$(\tau_E = 0.1)$ and negative-dominated $(\tau_E = 3)$ regime. Colors correspond to the green and red dots in a. See Methods for simulation details.
highest when the duration of the memory matches the direction decorrelation time \( \tau_D \simeq 1 \). (We consider cases corresponding to the colored dots in Fig. 1c.) Because of phenotypic heterogeneity\(^{25–27} \), cell speed, cell length, and signal transduction gain will likely vary from cell-to-cell even in an isogenic population. Using \textit{Escherichia coli} chemotaxis as an example, we predict that within a wild type population there will be individual cells that operate in the positive feedback regime. Hence, some cells will race up the gradient much faster than the rest of the population.

### Results

#### Exploration of the dynamical regimes of the run-and-tumble gradient ascent. For concreteness we consider the well characterized bacterial chemotaxis system in \textit{Escherichia coli}\(^{1,28–31} \) (see Supplementary Note 1). \textit{E. coli} swims by rotating flagella\(^{32–34} \). Switching between runs and tumbles follows Poisson statistics with transition rates \( \lambda_R(F) \) and \( \lambda_T(F) \),\(^{1,35–38} \) causing the state of the cell to rapidly fluctuate around the probability to run \( r(F) \)

\[
r(F) = \frac{\lambda_T(F)}{\lambda_R(F) + \lambda_T(F)} \simeq \frac{1}{1 + \exp(-HF)},
\]

with characteristic time scale \( t_S = 1/(\lambda_R + \lambda_T) \) (Supplementary Note 2). Here \( F \) represents the free energy of the receptor clusters (normalized by \( k_B T \)) and \( H \) is the input-output gain of the flagellar motors\(^{39} \). During runs the speed of the cell is constant \( \|\dot{X}\| = v_0 \) and the direction of motion is subject to rotational Brownian motion with diffusion coefficient \( D_R^{-1} \). During tumbles the speed is nil and reorientation is modeled with rotational diffusion \( D_T > D_R \)\(^{40} \). Taken together these two processes cause the cell to lose its original direction at the rate \( t_D^{-1} = (n-1)(rD_R + (1-rD_T)) \) where \( n = 2, 3 \) for two- and three-dimensional motion respectively.

Like many other sensory systems, \textit{E. coli} responds to relative changes along its trajectory \( \dot{X}(t) \) in the external concentration \( C(X,t) \) of an attractant\(^{41–43} \), while simultaneously adapting back to the steady state \( F_0 \) over the time scale \( t_M \)\(^{22,23,44–48} \). These features of the organism’s internal dynamics are conveniently encapsulated in the basic equation\(^{23,31,37,49} \) (Supplementary Note 3)

\[
\dot{F} = -\frac{F - F_0}{t_M} + N \left( \frac{\partial}{\partial t} + \dot{X} \cdot \nabla \right) \ln \frac{1 + C/K_i}{1 + C/K_a},
\]

where \( N \) is the gain of the receptor systems\(^{50,51} \), and \( K_i \ll K_a \) are the dissociation constants of the receptors limiting the cell’s range of sensitivity\(^{23,52–55} \).

The first term on the right hand side of equation (2) is the negative feedback. Because the probability to run \( r \) increases monotonically with \( F \), it too relaxes to its adapted state \( r_0 = r(F_0) \) with the time scale \( t_M \). The second term on the right of equation (2) shows that \( F \), and therefore \( r \), rises whenever \( C \) increases, either locally or due to the cell moving up a gradient. In a static gradient, this term is proportional to the cell velocity along the gradient; but is zero when the cell tumbles. Thus, the resulting increase in \( F \) (and \( r \)) is proportional to the probability to run \( r \), creating a positive feedback.

In this paper we will examine the dynamics of \( r \) taking place over multiple runs and tumble events, i.e. we will focus on time scales longer than \( t_S \). Taking the limit \( t_S \ll t_E, t_M, t_D \), the positive and negative feedback can be encapsulated in the following equation derived from equations (1) and (2) (see Methods and
Supplementary Note 4)

\[
\dot{r} \simeq r (1 - r) \left( \frac{-f(r) - f_0}{t_M} + \frac{rs}{L/(NHv_0)} \right).
\]  

(3)

For simplicity we start with a static exponential gradient with length scale \( L = 1/||\nabla \ln C|| \) and concentration within the sensitivity range of the cell \( K_i \ll C \ll K_a \). Here \( s = \cos \theta \) denotes the direction of motion relative to that of the gradient, and \( f = HF = -\ln (1/r - 1) \) denotes the rescaled free energy of the receptors. Two timescales naturally appear, \( t_M \) and \( t_E = L/(NHv_0) \), which encapsulate the magnitude of the negative feedback (\( \propto 1/t_M \)) and positive feedback (\( \propto 1/t_E \)). From these definitions, we define the following non-dimensional timescales:

\[
\tau_E = \frac{t_E}{t_M} = \frac{L}{t_M NHv_0},
\]

\[
\tau_D(r) = \frac{t_D}{t_M} = \frac{1/t_M}{(n - 1)(rD_R + (1 - r)D_T)}.
\]  

(4)

Here \( \tau_E \) quantifies the ratio between the negative and positive feedbacks. \( \tau_D \) is the ratio between the direction decorrelation time, which depends on the weighted average of the two rotational Brownian processes, and the length of the memory. One would expect different dynamical regimes depending on whether \( \tau_E \) and \( \tau_D \) are larger or smaller than one.

We first explored these differences by performing numerical simulations of the full stochastic model of bacterial chemotaxis for cells in an exponential gradient of MeAsp. We used a previously described stochastic simulator that reproduces well experimental measurements of the behavior of \( E. coli \) cells (see Methods). Calculating the steady state drift velocity from \( 10^4 \) sample trajectories for a range of \( \tau_E \) and \( \tau_D \) values reveals that cells climb the gradient much faster when the positive feedback \( \tau_E < 1 \) rather than the negative feedback \( \tau_E > 1 \) dominates (Fig. 1c). The drift velocity also depends on \( \tau_D \) and peaks when the direction decorrelation time is on the same order as the memory duration \( \tau_D \simeq 1 \).

These differences are reflected in the structure of the cell trajectories. When the negative feedback dominates cell trajectories exhibit runs in both the positive and negative directions with only a slight bias in their durations, resulting in a slow average progress up the gradient (Fig. 1d red). In contrast when the positive feedback dominates runs are long in the right direction and short otherwise, resulting in a much faster average drift up the gradient (Fig. 1d green). In this case matching the memory of the cell with the direction decorrelation time is important. If the cell loses direction too quickly \( \tau_D \) it cannot take full advantage of the elongated runs in the correct direction (Supplementary Figure 1).

**The positive feedback drives large fluctuations in the internal state of the organism.** To better understand the origin of the fast drift velocity when the positive feedback dominates, we examined the relationship between drift velocity \( V_D \) and the statistics of the internal state of the cell \( f \) while it is climbing the gradient at steady state. Non-dimensionalizing time and space, we define \( P(x, f, s, \tau) \), the probability that at time \( \tau = t/t_M \) the cell is at position \( x = X/(v_0 t_M) \) along the gradient with internal state \( f = HF \) and moving in the direction \( s = \cos \theta \). Starting from the master equation describing the run and tumble dynamics of an individual cell, we derive the following Fokker-Planck equation, valid for time scales longer than an individual run or tumble, by taking the limit \( t_S \ll t_M, t_D, t_E \) (see Methods):
The positive feedback drives large fluctuations in the internal state of the organism. (a) Marginal probability distribution of the internal variable at steady state $p(f)$; solid line: numerical solution of the Fokker-Planck equation (5); dashed line: sampled distribution from agent-based simulation. The colors indicate different parameter cases as in Fig. 1c: $\tau_{D0} = 1$, $\tau_E = 0.1$ (green), $\tau_E = 1$, $\tau_{D0} = 1$ (blue), $\tau_E = 3$, $\tau_{D0} = 1$ (red), and $\tau_E = 0.1$, $\tau_{D0} = 0.1$ (orange). Inset: zoomed view of $p(f)$ for blue and red; black curves are second order analytical approximations (see Methods).

(b) Drift velocity $V_D$ as a function of $\tau_E$ while keeping $\tau_{D0} = 1$ fixed; solid line: numerical integration of equations (5) and (6); dashed line: agent-based simulations; dash-dot line: mean field theory; dotted line: Keller-Segel. See Methods for details.

Here

$$\partial_x P = -\partial_f \left(\left(- (f - f_0) + \frac{r_s}{\tau_E}\right) P\right) + \frac{\hat{L}_s P}{(n-1) \tau_D} - \partial_x (r_s P).$$

(5)

Here $\hat{L}_s = (1 - s^2)^{\frac{3-n}{2}} \partial_s \left((1 - s^2)^{\frac{n-1}{2}} \partial_s\right)$ is the rotational diffusion operator on the $(n-1)$-sphere, and $r$ and $\tau_D$ are functions of $f$. Because the cells are log-sensing and we are considering a static exponential gradient, they experience a constant external “field”. Thus, at steady state the solution becomes independent of $x$; integrating over $s$ and $f$ we obtain the constant drift velocity (Supplementary Note 5)

$$V_D = \langle rs \rangle = \tau_E \langle f - f_0 \rangle.$$  

(6)

Here $\langle \cdot \rangle$ represents the average over $s$ and $f$ and the bar indicates steady state. Equation (6) shows that the drift velocity depends linearly on the first moment of the internal state $f$, i.e. on the shape of the marginal steady state distribution $\overline{p}(f) = \int P ds$. To calculate $\overline{p}(f)$ we solve (5) at steady state by expanding the solution in orthonormal eigenfunctions of the operator $\hat{L}_s$ with the first two coefficients being the marginal distribution $\overline{p}(f)$ and the first order moment $\overline{p}_1(f)/\sqrt{n} = \int P s ds$ (see Methods and Supplementary Note 6).

When the negative feedback dominates ($\tau_E \ll 1$) the distribution of $f$ becomes sharply peaked and nearly Gaussian, with variance $\sigma^2 = \tau_{D0}^2 r_0^2 / n \tau_E^2$ around the first moment; it barely deviates from the adapted state $f_0$ (Fig. 2a red, see Methods and Supplementary Note 7). Substituting into equation (6) yields known mean field theory results: $V_D \approx r_0 \tau_{D0} / (n \tau_E (1 + \tau_{D0}))$ where $\tau_{D0} = d \tau_D / df$ at $f_0$ (Fig. 2b). When $\tau_E \approx 1$ the negative feedback is not as strong and the distribution of $f$ widens proportionally to the strength of the positive feedback $\sigma \propto \tau_E^{-1}$ (Fig. 2a blue).
These approximations no longer hold when the positive feedback dominates \( (\tau_E \ll 1) \). \( p(f) \) now exhibits large deviations (Fig. 2a green) bounded below and above by \( f_U \) and \( f_L \) respectively. These can be derived from the Fokker-Planck equation (5) (see Methods)

\[
\begin{align*}
    f_L - f_0 &= -\frac{r(f_L)}{\tau_E} \\
    f_U - f_0 &= \frac{r(f_U)}{\tau_E}.
\end{align*}
\]

(7)

As \( \tau_E \) becomes small the lower bound decreases as \( f_L \to \ln \tau_E \), whereas the upper bound increases as \( f_U \to 1/\tau_E \) (see Methods), causing the distribution to become very asymmetric. Thus, mean field theory becomes inadequate.

In this regime, keeping the direction of motion long enough \( (\tau_D \gtrsim 1) \) in the positive direction allows the internal variable to reach the upper limit \( f \to f_U \), accumulating significant distribution near \( f_U \) (Fig. 2a green). A short direction decorrelation time, by comparison, shifts the peak of the distribution back to \( f_0 \) (Fig. 2a orange), so that drift velocity becomes smaller (Supplementary Figure 1).

We verified that \( \bar{p}(f) \) obtained from equation (5) in the limit \( t_S \ll t_E, t_M, t_D \) is a good approximation that captures well the relevant dynamics as long as \( \tau_D \) is not too small by plotting it against the distribution of \( f \) obtained from the numerical simulations with non-zero \( t_S \) described earlier (Fig. 2a). Integrating \( p(f) \) according to equation (6) predicts well the drift velocity over the entire range of feedback strengths, including in the region where the positive feedback dominates \( (\tau_E < 1) \). Note: in this regime Keller-Segel and mean field approximations (see Methods) fail (Fig. 2b).

Non-modal dynamics generate long runs uphill but short ones otherwise. In the fast gradient climbing regime \( (\tau_E \ll 1) \) runs are long in the correct direction but short otherwise (Fig. 1b). To gain mechanistic insight into this striking difference we examined the Langevin equations equivalent to the Fokker-Planck equation (5). Defining \( v = rs \) as the instantaneous speed along the gradient, we change variables from \((f, s, x)\) to \((r, v, x)\) and obtain (Supplementary Note 8):

\[
\begin{align*}
    \frac{dr}{d\tau} &= r(1-r) \left( -(f(r) - f_0) + \frac{v}{\tau_E} \right) \\
    \frac{dv}{d\tau} &= v(1-r) \left( -(f(r) - f_0) + \frac{v}{\tau_E} \right) - \frac{v}{\tau_D(r)} + \sqrt{\frac{r^2 - v^2}{\tau_D(r)}} \eta(\tau) \\
    \frac{dx}{d\tau} &= v,
\end{align*}
\]

(8)

where \( \eta(\tau) \) denotes Gaussian white noise with \( \langle \eta(\tau_1)\eta(\tau_2) \rangle = \delta(\tau_1 - \tau_2) \).

The nullclines of the deterministic component, \( v = \tau_E(f - f_0) \) and \( v = 0 \) (dashed lines in Fig. 3a,d), intersect only once at the fixed point \((r, v) = (r_0, 0)\), which is stable because the eigenvalues of the corresponding relaxation matrix

\[
\begin{bmatrix}
    -1 & (1 - r_0) r_0 / \tau_E \\
    0 & -1 / \tau_D_0
\end{bmatrix}
\]

(9)

are both negative (see Methods). However, fluctuations introduced by rotational diffusion and tumbling prevent the system from settling at the fixed point (heat maps in Fig. 3). Instead, the system constantly
deviates from or relaxes towards it; net drift along the gradient is achieved by spending more time on average in the region of positive velocity $v > 0$.

The stochastic excursions in the $(r; v)$-plane away from the fixed point exhibit distinctive trajectories depending on the value of $\tau_E$. When the positive feedback dominates ($\tau_E \ll 1$; Fig. 3a,b,c) the eigenvectors of the relaxation matrix, $(1,0)^T$ and $(\frac{(1-r_0)^{\tau_0}}{\tau_E \tau_D(1)}, \frac{\tau_0}{\tau_D(1)})^T$, are highly non-orthogonal. As a result, linear deviations from the fixed point can grow transiently\textsuperscript{57} to feed the non-linear positive feedback ($v^2$ term second line in equation (8)) leading to large deviations. Importantly, this only happens for runs that start in the correct direction. If instead the cell starts running in the wrong direction the linear deviation does not grow (Fig. 3c). This asymmetry arises because the $v^2$ term is always positive. Similar selective amplification properties are observed in neuronal networks, where non-normal dynamics enables the network to respond to certain signals while ignoring others (including noise)\textsuperscript{58,59}. Thus, a cell running in the correct direction is aided by the positive feedback, which pushes its internal dynamics towards the upper right corner of the phase plane where the probability to run $r = 1$ and the speed $v = 1$. If, instead, a cell starts running in the wrong direction ($v < 0$), its internal dynamical state is pushed back into the high noise region near the fixed point where it will rapidly pick a new direction (Fig. 3b).
In contrast, when the negative feedback dominates ($\tau_E \gtrsim 1$; Fig. 3d,e,f), the eigenvectors of the relaxation matrix become nearly orthogonal. Linear deviations from the fixed point cannot grow and simply relax to the fixed point regardless of the initial direction of the run. This key difference between the positive and negative feedback regimes are reflected in the flow field (white lines in Fig. 3a,d). Thus, in the negative feedback regime, stochastic excursions from the fixed point are mostly dominated by the decorrelation time $\tau_D$ with only slight bias in duration between runs up and down the gradient.

Receptor saturation, varying gradient length scales, and trade-offs. So far our analysis assumed a constant exponential gradient at concentrations within the sensitivity range of a log-sensing organism. However, for *E. coli* equation (2) indicates that the sensitivity to methyl-aspartate is logarithmic only for concentrations between $K_i \ll C \ll K_a$. When $C < K_i$ the sensitivity becomes linear, whereas when $C > K_a$ the receptors saturate. Thus, as a cell approaches a high concentration source its sensitivity decreases, which effectively increases the value of $\tau_E$. Simulations in an exponential gradient show that this effect results in an eventual slow-down as the cell approaches the source (Fig. 4a-c).

Realistic gradients are typically limited in spatial extent and often are not exponential, in which case $L$ and therefore $\tau_E$ is different in different regions. $L$ is long near the source in a linear gradient, for example, and decreases linearly with distance from the source. Simulations show that the cell initially climbs the gradient fast but later slows down as the gradient length scale $L$ increases and $\tau_E$ increases (Fig. 4d-f). On the contrary, for a static localized source in three dimensions, $L$ is short near the source but increases linearly with distance from it (see Methods). Thus, $\tau_E$ decreases and the cell accelerates as it approaches the source (Fig. 4g-i). Comparing cells in various dynamical regimes (different values of $\tau_E$) across these different gradients suggests that a lower value of $\tau_E$ results in a faster gradient ascent.

When entering a food gradient, it is natural to try to climb as fast as possible. However this strategy could create a problem: the longer runs implied by the positive feedback mechanism could propel the accelerating *E. coli* beyond the nutrient source. This is the case in Fig. 4e, where cells with the lowest $\tau_E$ (green) reach the source first but overshoot slightly; they settle, on average, at a further distance than those with intermediate $\tau_E$ (blue). In general, natural environments will be complex, with a variety of different sources and gradients, implying different parameter domains will be optimal for *E. coli* at different times. Such phenotypic diversity may well confer an advantage.

Fast chemotactic cells in wild type isogenic population of *E. coli*. Motivated by our findings we asked whether in a wild type population of *E. coli* some cells are likely to be in the positive-feedback-dominated regime where $\tau_E = L / (NHv_0t_M) \ll 1$. This would happen when the gradient is steep ($L$ short), the gain of the receptor cluster $N$ or of the motors $H$ is large, the velocity of the cell $v_0$ is large, or the cell memory $t_M$ is long. These parameters are likely to vary from cell-to-cell due to non-genetic phenotypic heterogeneity, known to be large in *E. coli*. Using a recent model of phenotypic diversity in *E. coli* chemotaxis that reproduces available experimental data on the laboratory strain RP437 climbing a linear gradient of methyl-aspartate, we ran simulations of 16,000 wild type-like cells in that experiment. For each cell we calculated $t_E$, $t_M$, and $t_D$ based on the parameter input and their locations in the known gradient to show their scatter. Initially, the cells are near the bottom of the gradient and many cells are in the positive feedback dominated region (white dots in Fig. 1a). As the cells climb the gradient $\tau_E$ becomes larger and the regime shifts to higher $\tau_E$ (black dots in Fig. 1a).
Figure 4 | Environmental context, length scales, and receptor saturation. (a-c) Exponential gradient. (a) Schematic of a gradient of methyl-aspartate $C = C_0 \exp \left( -R/L \right)$ with length scale $L = 1000 \mu m$ and source concentration $C_0 = 10 \text{ mM}$. Contour lines show logarithmically spaced concentration levels in units of $mM$. Contours spacing illustrates $L$ constant. (b) The mean trajectory over $10^4$ E. coli cells of the position $X$ (in real units $\mu m$) as a function of time $t$ (in s) when receptor saturation is taken into account. Initial values of $\tau_E$ are 0.1 (green), 1 (blue) or 3 (red). The shadings indicate standard deviations. The labels on the right axis show the concentration in $mM$ at each position. The effect of receptor saturation becomes visible around $X = 4000 \mu m$ (dashed lines) where the concentration of methyl-aspartate is $5.5 \text{ mM}$. (c) Corresponding time trajectories of the values of $\tau_E$ at mean positions. (d-f) Linear gradient. Similar to a-c but for $C = C_1 - aR$ where the source concentration is $C_1 = 1 \text{ mM}$ and decreases linearly at rate $a = 0.0001 \text{ mM}/\mu m$ with distance $R$ from the source. Contours spacing decreases with distance from the source (at the top), illustrating decreasing $L$. Diagonally filled region indicates decreasing $\tau_E$ over time. (g-i) Localized source. Similar to a-c but for $C = C_2R_0/R$, $R > R_0$ where the source concentration ($C_2 = 1 \text{ mM}$) is constant within a ball of radius $R_0 = 100 \mu m$ and decreases with radial distance as $1/R$ away from the source. Contours spacing increases away from the source (at the origin), illustrating increasing $L$. Diagonally filled region indicates decreasing $\tau_E$ over time. See Methods for details.
In these simulations we assumed that all the cells had the same length. However, as cells grow and divide, their length will change on average by a factor 2. Because the rotational diffusion coefficients $D_R$ and $D_T$ depend non-linearly on cell length $^{38,68}$ (see Methods) the cells are likely to move within the parameter space of Fig. 1a during their lifetime. Repeating the simulation with cells twice as long (black contours in Fig. 1a) reveals that as cells mature they should have a higher chance to be in the faster parameter regime.

It is interesting that not all wild cells are in the fast $\tau_E < 1$ regime. A likely reason is that RP437 is a laboratory strain that was selected by swimming in agar plates. In agar, cells can get trapped in the agar mesh, requiring a tumble to escape the trap. As a result, runs that are too long become detrimental $^{69}$. Thus cells likely face a trade-off wherein long memory, $t_M$, is better in liquid but could become detrimental in agar. Again, in alternating environments phenotypic heterogeneity might be advantageous $^{26,27,56,66,67}$.

**Discussion**

It is common to make simplifying assumptions to facilitate analysis, but we do not believe that these limit our findings. Although we assumed log-sensing receptors and a constant adaptation time for most of our derivations (equation (2)), we showed with simulations that our results hold when the limited range of sensitivity of the receptors is taken into account. Likewise, we verified that the results remain valid over a wide range of the adapted probability to run $r_0$, as well as when we include the asymmetric rates of adaptation exhibited by *E. coli* in response to step up and down in attractant $^{70}$ (see Supplementary Figure 2 for details).

In this work we considered what happens when the environmental gradients are static (or slowly changing), and revealed the capability for very rapid response. This could be particularly important in rapidly changing gradients such as those generated by turbulent flows and lysing organisms in the sea $^{3,64}$. Our findings imply that organisms using a run-and-tumble strategy to climb gradients of interest could do so rapidly and efficiently when their internal state undergoes large asymmetric fluctuations away from the mean. Such fluctuations arise from suppression (facilitation) of tumbles following motions up (down) the gradient, which corresponds to benefits (costs) associated with the gradient ascent. The large asymmetry in the distribution of internal states boosts the drift velocity significantly by dramatically increasing the benefit-to-cost ratio. We showed how this is driven by the positive feedback inherent to the run-and-tumble strategy, and how it contrasts sharply with the negative feedback-dominated regime previously described.

Although we used *E. coli* bacterial chemotaxis as an example, we expect our findings to be generally applicable to other biased “random walk” strategies exhibited by organisms when local directional information is unreliable. In essence, any stochastic navigation strategy requires a memory, $t_M$, to make temporal comparisons, a reorientation mechanism, $t_D$, to sample new directions, and external information, $t_E$, relayed to decision-making circuitry through motion and signal amplification. Our theoretical contribution showed the (surprisingly) diverse behavioral repertoire that is possible by having these work in concert. In retrospect, perhaps this should not be surprising given the diverse environments in which running-and-tumbling organisms can thrive.

**Methods**

*Derivation of equation (3)-(5) in the fast switching limit.* Define $P_R$ and $P_T$ the probability distributions at time $t$ to be running or tumbling at position $X$ along the gradient in direction $s = \cos \theta$ relative to gradient direction with internal variable
\[ f = H F. \] They satisfy a two-state master equation model (Supplementary Note 4)

\[
\partial_t P_R = -\partial_f \left( \left( -\frac{f - f_0}{t_M} + \frac{NHv_0}{L} s \right) P_R \right) + D_R \tilde{L}_s P_R - \frac{1 - r}{t_S} P_R + \frac{r}{t_S} P_T - s \partial_X (v_0 P_R)
\]

\[
\partial_t P_T = -\partial_f \left( -\frac{f - f_0}{t_M} P_T \right) + D_T \tilde{L}_s P_T + \frac{1 - r}{t_S} P_R - \frac{r}{t_S} P_T,
\]

(10)

where \( r = \lambda_T / (\lambda_R + \lambda_T) \), \( t_S = 1 / (\lambda_R + \lambda_T) \), and \( \tilde{L}_s \) is the polar angle part of the rotational diffusion operator on the \((n-1)\)-sphere.

If we assume the switching terms with \( t_S^{-1} \) in equation (10) dominate, the probabilities to be running and tumbling equilibrate on a much faster timescale than the other ones. Therefore we can let \( P = P_R + P_T \) and can approximate the actual probability to run as its expected value \( P_R / P \approx r \). Adding the two equations above yields the Fokker-Planck equation:

\[
\partial_t P \approx -\partial_f \left( \left( -\frac{f - f_0}{t_M} + \frac{rs}{L/NHv_0} \right) P \right) + (rD_R + (1 - r) D_T) \tilde{L}_s P - rs \partial_X (v_0 P).
\]

(11)

This is equivalent to a system of Langevin equations. Considering \( \partial r / \partial f = r (1 - r) \) the internal variable dynamics (the first term) gives equation (3) which defines \( t_E \). The angular dynamics (the second term) defines \( t_D \).

Using the time scale definitions in (4) and non-dimensionalizing time \( \tau = t / t_M \) and position \( x = X \newcommand{v}{(v_0 t_M)} \), we obtain the Fokker-Planck equation (5).

**Derivation of the analytical solution to Fokker-Planck equation (5) by angular moment expansion.** In 3D the eigenfunctions of the axially symmetric diffusion operator \( \tilde{L}_s \) are the Legendre polynomials, so we can expand the distribution in a Fourier-Legendre series (see Supplementary Note 6). The first two orders are the marginal distribution \( p(f) = \int_{-1}^{1} P(f, s) ds \) and the first order moment \( p_1 / \sqrt{n} = \int_{-1}^{1} s P(f, s) ds \). In general, each order satisfies

\[
\partial_{\tau} p_k = -\partial_f \left( -\frac{f - f_0}{\tau_E} p_k + \frac{r(f)}{\tau_E} \hat{s}_{kl} p_l \right) + \frac{\lambda_k}{(n - 1) \tau_D} (f) p_k - \partial_x (\hat{s}_{kl} p_l),
\]

(12)

where \( \hat{s}_{kl} = \sqrt{\frac{k(k+n-3)}{(2k+n-4)(2k+n-2)}} \delta_{k-1,l} + \sqrt{\frac{(k+1)(k+n-2)}{(2k+n-2)(2k+n)}} \delta_{k+1,l} \) (summation over \( l \) implied) is an operator relating neighboring orders. It comes from the positive feedback term. When \( n = 3 \) it is \( \hat{s}_{kl} = \frac{k}{\sqrt{4k^2 - 1}} \delta_{k-1,l} + \frac{k+1}{\sqrt{4(k+1)^2 - 1}} \delta_{k+1,l} \).

When \( \tau_D \ll 1 \) we neglect the 2nd and higher orders to close the infinite series of moment equations. At steady state the approximation gives the analytical solution

\[
\bar{p}(f) = \frac{1}{Z} \exp \left( -\frac{1}{n} \frac{f - f_0}{\tau_E} \hat{s} \partial_{\tau} (f_1) \right),
\]

(13)

where \( Z \) is a normalization constant.

**Derivation of the distribution \( \bar{p}(f) \) when the negative feedback dominates.** Integrating equation (13) and assuming \( |f - f_0| \) small, we can write the solution as (Supplementary Note 7)

\[
\bar{p}(f) = C e^{-\frac{(f-f_0)^2}{2\sigma^2}} (1 + G),
\]

(14)

where \( \sigma^2 = \frac{r_0^2 \tau_D \delta \sqrt{v}}{n \tau_E} \), \( C \) is a normalization factor and \( G \) is a polynomial in \( (f - f_0) \), giving a small correction to the Gaussian distribution. Explicitly integrating the above distribution we recover previous mean field theory results (see details in Supplementary Note 7)

\[
V_D = \frac{r_0 \sqrt{\tau_D}}{n \tau_E (1 + \tau_D)},
\]

(15)
where \( \tau' D_0 = d\tau D / df \) is evaluated at \( f_0 \).

**Bounds of the distribution** \( p(f) \). The first term in equation (5) says the flux in \( f \)-space is non-negative provided \(- (f - f_0) + rs / \tau_E > 0 \), or, noting \( s \leq 1 \)
\[
f \leq f_0 + r(f)s/\tau_E \leq f_0 + r(f)/\tau_E.
\]
Thus the upper bound \( f_U \) of the distribution \( p(f) \) is achieved at equality. Similarly, the lower bound \( f_L \) is achieved when we take equal signs of
\[
f \geq f_0 + r(f)s/\tau_E \geq f_0 - r(f)/\tau_E,
\]
noting \( s \geq 1 \). This gives equation (7).

When \( \tau_E \) becomes small we note \( f_U, L \) deviates far away from \( f_0 \) as \( 1/\tau_E \to \infty \). Using the definition \( r = 1/(1 + \exp(-f)) \), from equation (7) we write
\[
f_{L,U} = \mp \frac{1}{\tau_E (1 + \exp(-f_{L,U}))}.
\]
The plus sign gives \( \exp(-f_U) \ll 1 \) and \( f_U \approx 1/\tau_E \). The minus sign gives \( \exp(-f_L) \gg 1 \) and \( f_L = -\exp(f_L)/\tau_E \). Taking logarithm, the latter gives \( f_L = \ln(|f_L|/\tau_E) \approx \ln \tau_E \).

**Linear response near the fixed point of the Langevin system** (8). Consider the region in \( r-v \) plane near the fixed point
\[
r = r_0 + \delta r \\
v = 0 + \delta v.
\]
Substitute this into the deterministic component of the Langevin equation (8) to obtain
\[
\frac{d}{d\tau} \begin{bmatrix} \delta r \\ \delta v \end{bmatrix} = \begin{bmatrix} -1 & (1 - r_0) r_0 / \tau_E \\ 0 & -1/\tau_D0 \end{bmatrix} \begin{bmatrix} \delta r \\ \delta v \end{bmatrix}.
\]
The eigenvalues and corresponding eigenvectors of the matrix describing linear response are:
\[
\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{for eigenvalue} \quad -1;
\]
\[
\begin{bmatrix} (1 - r_0) r_0 / \tau_E \\ \tau_D0 / \tau_E \end{bmatrix} \quad \text{for eigenvalue} \quad -\frac{1}{\tau_D0}.
\]
When \( \tau_E \) is large, \( (1 - r_0) r_0 / \tau_E \ll 1 \) and the eigenvectors are almost orthogonal. When \( \tau_E \) is small, \( (1 - r_0) r_0 / \tau_E \ll 1 \) and the eigenvectors are not orthogonal.

**Receptor saturation.** The second term in equation (2) is, using \( L = 1/||\nabla \ln C|| \),
\[
\frac{N_{t0}}{L} \left( \frac{1}{K_i/C + 1} - \frac{1}{K_a/C + 1} \right). \quad (16)
\]
The factor in the parenthesis is always less than 1, although it is close to 1 when \( K_i \ll C \ll K_a \). As \( C \) increases this factor contributes to a smaller receptor gain and a corresponding larger \( \tau_E \) (Fig. 4b).

**Varying gradient length scales.** For a linear gradient \( C = C_1 - aX \), we see \( L = 1/||\nabla \ln C|| = C/a = C_1/a - X \), i.e. the gradient length scale decreases linearly with distance away from the source.

For a localized source we assume that the concentration profile satisfies a standard diffusion equation \( \partial_t C = \nabla^2 C \) without decay. The spherically symmetric solution in 3D is \( C \propto 1/R \) where \( R \) is the radial distance. Thus \( L = 1/||\nabla \ln C|| = R \), i.e. the gradient length scale increases linearly with distance away from the source.
Rotational diffusion coefficient. An *E. coli* can be modeled as an ellipsoid with major axis parallel to the direction of motion. During runs and tumbles the change in running direction is associated with rotation about its minor axis. With semi-major axis $a$ and semi-minor axis $b$, the rotational diffusion coefficient is

$$D_R = \frac{k_B T}{8 \pi \eta a^3 / 3},$$

where $k_B$ is Boltzmann’s constant, $T$ is the temperature, and $\eta$ is the viscosity of the fluid. For a long cell $\ln \frac{2a}{b}$ does not change significantly compared to the cubic dependence on $a^3$ in the denominator, so the diffusion coefficient is about 8 times smaller when cells are twice as long. We assume a proportionate diffusion coefficient during a tumble and let $D_T$ also scale 8 times smaller.

Simulation methods. The agent-based simulations were performed using Euler’s method to integrate a standard model of *E. coli* chemotaxis\cite{20,21,23,31,37} (see Supplementary Methods). Parameter values are in Supplementary Table 1.

In Fig. 1c heat map, $10^4$ sample trajectories that were each 650 seconds long — sufficient for the cells to reach steady state — were generated for each parameter condition. The concentrations were assumed to be in the log-sensing range throughout the domain, and in the numerical integration $\ln C$ was used instead of $\ln \left(1 + \frac{C}{K_i}\right) / \left(1 + \frac{C}{K_a}\right)$. The drift velocity $V_D$ was calculated by fitting the linear part of the mean trajectory. In Fig. 1d the first 50 s were removed to avoid the start up transient.

In Fig. 1c the black/white dots are generated using experimentally relevant simulation parameters previously described\cite{56}. The time scale values of each cell depend on the cell parameters and the gradient length scale $L$ as in equation (4). Near the bottom of the gradient $L = 1500 \mu m$ and near the top of the gradient $L = 4800 \mu m$.

In Fig. 4 $\ln \left(1 + \frac{C}{K_i}\right) / \left(1 + \frac{C}{K_a}\right)$ was used. In Fig. 4b,d,f the $\tau_E$ calculation considered receptor saturation according to (16) as well as the varying gradient length scales, with $C$ and $L$ evaluated at mean positions. Note this is not the average $\tau_E$ over the population.

In Fig. 3 the Langevin trajectories were generated using Euler’s method to integrate equation (8).

Numerical methods and data analysis. Figure 2 used the simulations in Fig. 1c. The steady state $\bar{p}(f)$ from agent-based simulation was calculated from the histogram of all the internal values of the $10^4$ individual cells between $\tau = 10$ and $\tau = 20$, sampled at regular steps of $\tau = 0.01$. $V_D$ was calculated as in Fig. 2a.

Numerical solutions of the Fokker-Planck equation (5) were obtained by expanding the distribution in angles, as in equation (12), and keeping the first 10 orders. The steady state $\bar{p}(f)$ was found by solving an initial value problem using the NDSolve function in Mathematica, with $10^4$ spatial points and integration time up to $\tau = 10$. Further orders, finer grid, and longer integration times were checked to ensure solution accuracy. $V_D$ was calculated from equation (6).

In Fig. 2a inset, the black curves show the approximate Gaussian distribution in equation (14). The explicit expression is in Supplementary Note 7.

In Fig. 2b Keller-Segel refers to equation (15), similarly expressed in previous research\cite{10,11,15,18}, where negative feedback dominates ($\tau_E \gg 1$). MFT refers to cases where negative feedback is not as strong ($\tau_E \gtrsim 1$) so that the mean internal variable deviates noticeably from $f_0$. Thus, $\tau_D$, and $\tau_D'$ in equation (15) were evaluated at the steady state mean $\langle f \rangle$ instead of $f_0$. This was combined with equation (6) to find $V_D$\cite{20,21}.

References

1. Berg, H. C. & Brown, D. A. Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature* 239, 500-504 (1972).
2. Blackburn, N., Fenchel, T. & Mitchell, J. Microscale nutrient patches in planktonic habitats shown by chmotactic bacteria. *Science* **282**, 2254-2256 (1998).
3. Stocker, R. Marine microbes see a sea of gradients. *Science* **338**, 628-633 (2012).
4. Agliari, E. *et al.* Cancer-driven dynamics of immune cells in a microfluidic environment. *Sci. Rep.* **4**, 6639 (2014).
5. Pierce-Shimomura, J. T., Morse, T. M. & Lockery, S. R. The fundamental role of pirouettes in *Caenorhabditis elegans* chemotaxis. *J. Neurosci.* **19**, 9557-9569 (1999).
6. Albrecht, D. R. & Bargmann, C. I. High-content behavioral analysis of *Caenorhabditis elegans* in precise spatiotemporal chemical environments. *Nat. Methods* **8**, 599-606 (2011).
7. Gomez-Marin, A., Stephens, G. J. & Louis, M. Active sampling and decision making in *Drosophila* chemotaxis. *Nat. Commun.* **2**, 441 (2011).
8. Taylor-King, J. P., Franz, B., Yates, C. A. & Erban, R. Mathematical modelling of turning delays in swarming robots. *IMA J. Appl. Math.* **80**, 1454-1474 (2015).
9. Keller, E. F. & Segel, L. A. Traveling bands of chemotactic bacteria: a theoretical analysis. *J. Theor. Biol.* **30**, 235-248 (1971).
10. Lovely, P. S. & Dahlquist, F. W. Statistical measures of bacterial motility and chemotaxis. *J. Theor. Biol.* **50**, 477-496 (1975).
11. Schnitzer, M. J Theory of continuum random walks and application to chemotaxis. *Phys. Rev. E* **48**, 2553-2568 (1993).
12. Spiro, P. A., Parkinson, J. S. & Othmer, H. G. A model of excitation and adaptation in bacterial chemotaxis. *Proc. Natl. Acad. Sci. USA* **94**, 7263-7268 (1997).
13. Strong, S. P., Freedman, B., Bialek, W. & Koberle, R. Adaptation and optimal chemotactic strategy for *E. coli*. *Phys. Rev. E* **57**, 4604-4617 (1998).
14. De Gennes, P. G. Chemotaxis: the role of internal delays. *Eur. Biophys. J.* **33**, 691-693 (2004).
15. Erban, R. & Othmer, H. G. From individual to collective behavior in bacterial chemotaxis *SIAM J. Appl. Math.* **65**, 361-391 (2004).
16. Clark, D. A. & Grant, L. C. The bacterial chemotactic response reflects a compromise between transient and steady-state behavior. *Proc. Natl. Acad. Sci. USA* **102**, 9150-9155 (2005).
17. Tailleur, J. & Cates, M. E. Statistical mechanics of interacting run-and-tumble bacteria. *Phys. Rev. Lett.* **100**, 218103 (2008).
18. Celani, A. & Vergassola, M. Bacterial strategies for chemotaxis response. *Proc. Natl. Acad. Sci. USA* **107**, 1391-1396 (2010).
19. Saragosti, J., Calvez, V., Bournaveas, N., Perthame, B., Buguin, A. & Silberzan, P. Directional persistence of chemotactic bacteria in a traveling concentration wave. *Proc. Natl. Acad. Sci. USA* **108**, 16235-16240 (2011).
20. Si, G., Wu, T., Ouyang, Q. & Tu, Y. Pathway-based mean-field model for *Escherichia coli* chemotaxis. *Phys. Rev. Lett.* **109**, 048101 (2012).
21. Dufour, Y., Fu, X., Hernandez-Nunez, L. & Emonet, T. Limits of feedback control in bacterial chemotaxis. *PLoS Comput. Biol.* **10**, e1003694 (2014).
22. Block, S. M., Segall, J. E. & Berg, H. C. Adaptation kinetics in bacterial chemotaxis. *J. Bacteriol.* **154**, 312-323 (1983).
23. Shimizu, T. S., Tu, Y. & Berg, H. W. A modular gradient-sensing network for chemotaxis in *Escherichia coli* revealed by responses to time-varying stimuli. *Mol. Syst. Biol.* **6**, 382-395 (2010).
24. Zhu, X. *et al.* Frequency-dependent *Escherichia coli* chemotactic behavior. *Phys. Rev. Lett.* **108**, 128101 (2012).
25. Spudich, J. L. & Koshland, D. E., Jr. Non-genetic individuality: chance in the single cell. *Nature* **262**, 467-471 (1976).
26. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183-1186 (2002).
27. Ackermann, M. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol.* **13**, 497-508 (2015).
28. Berg, H. C. Motile behavior of bacteria. *Phys. Today* **53**, 24-29 (2000).
29. Wadhams, G. H. & Armitage J. P. Making sense of it all: bacterial chemotaxis. *Nat. Rev. Mol. Cell Biol.* **5**, 1024-1037 (2004).
30. Sourjik, V. & Wingreen, N. S. Responding to chemical gradients: bacterial chemotaxis. *Curr. Opin. Cell Biol.* **24**, 262-268 (2012).
31. Tu, Y. Quantitative modeling of bacterial chemotaxis signal amplification and accurate adaptation. Annu. Rev. Biophys. 42, 337-359 (2013).
32. Berg, H. C. & Anderson, R. A. Bacteria swim by rotating their flagellar filaments Nature 245, 380-382 (1973).
33. Silverman, M. & Simon, M. Flagellar rotation and the mechanism of bacterial motility. Nature 249, 73-74 (1974).  
34. Larsen, S. H., Reader, R. W., Kort, E. N., Tso, W. & Adler, J. Change in direction of flagellar rotation is the basis of the chemotactic response in Escherichia coli. Nature 249, 74-77 (1974).
35. Segall, J. E., Manson, M. D. & Berg, H. C. Signal processing times in bacterial chemotaxis. Nature 296, 855-857 (1982).
36. Tu, Y. & Grinstein, G. How white noise generates power-law switching in bacterial flagellar motors. Phys. Rev. Lett. 94, 208101 (2005).
37. Sneddon, M. W., Pontius, W. & and Emonet, T. Stochastic coordination of multiple actuators reduce latency and improves chemotactic response in bacteria. Proc. Natl. Acad. Sci. USA 109, 805-810 (2012).
38. Taute, K. M., Gude, S., Tans, S. J. & Shimizu, T. S. High-throughput 3D tracking of bacteria on a standard phase contrast microscope. Nat. Commun. 6, 8776 (2015).
39. Cluzel, P., Surette., M & Leibler, S. An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. Science 287, 1652-1655 (2000).
40. Saragosti, J., Silberzan, P. & Buguin A. Modeling E. coli tumbles by rotational diffusion. Implications for chemotaxis. PLoS One 7, e35412 (2012).
41. Adler, J. Chemotaxis in bacteria. Science 153, 708-716 (1966).
42. Dahlquist, F. W., Lovely, P. & Koshland, D. E. Quantitative analysis of bacterial migration in chemotaxis. Nat. New Biol. 236, 120-123 (1972).
43. Kalinin, Y. V., Jiang, L., Tu, Y. & Wu, M. Logarithmic sensing in Escherichia coli bacterial chemotaxis. Biophys. J. 96, 2439-2448 (2009).
44. Springer, M. S., Goy, M. F. & Adler, J. Protein methylation in behavioral control mechanisms and in signal transduction. Nature 280, 279-284 (1979).
45. Segall, J. E., Block, S. M. & Berg, H. C. Temporal comparisons in bacterial chemotaxis. Proc. Natl. Acad. Sci. USA 83, 8987-8991 (1986).
46. Barkai, N. & Leibler, S. Robustness in simple biochemical networks. Nature 387, 913-917 (1997).
47. Alon, U., Surette, M. G., Barkai, N. & Leibler, S. Adaptation in bacterial chemotaxis. Nature 397, 168-171 (1999).
48. Lazova, M. D., Ahmed, T., Bellomo, D., Stocker, R. & Shimizu, T. S. Response rescaling in bacterial chemotaxis. Proc. Natl. Acad. Sci. USA 108, 13870-13875 (2011).
49. Celani, A., Shimizu, T. S. & Vergassola M. Molecular and functional aspects of bacteria chemotaxis. J. Stat. Phys. 144, 219-240 (2011).
50. Bray, D., Levin, M. D. & Morton-Firth C. J. Receptor clustering as a cellular mechanism to control sensitivity. Nature 393, 85-88 (1998).
51. Sourjik, V. & Berg, H. C. Receptor sensitivity in bacterial chemotaxis. Proc. Natl. Acad. Sci. USA 99, 123-127 (2002).
52. Adler, J. Chemoreceptors in bacteria. Science 166, 1588-1597 (1969).
53. Adler, J. Chemotaxis toward amino acids in Escherichia coli. J. Bacteriol. 112, 315-326 (1972).
54. Endres, R. G. & Wingreen, N. S. Precise adaptation in bacterial chemotaxis through “assistance neighborhoods”. Proc. Natl. Acad. Sci. USA 103, 13040-13044 (2006).
55. Hansen, C. H., Endres, R. G. & Wingreen, N. S. Chemotaxis in Escherichia coli: a molecular model for robust precise adaptation. PLoS Comput. Biol. 4, e1 (2008).
56. Frankel, N. W., Pontius, W., Dufour, Y. S., Long, J., Hernandez-Nunez, L. & Emonet, T. Adaptability of non-genetic diversity in bacterial chemotaxis. eLife 10.7554/eLife.03526 (2014).
57. Schmid, P. J. Nonmodal Stability Theory. Annu. Rev. Fluid Mech. 39, 129-162 (2007).
58. Murphy, B. K. & Miller, K. D. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. Neuron 61, 635-648 (2009).
59. Hennequin, G., Vogels, T. P. & Gerstner, W. Non-normal amplification in random balanced neural networks. Phys. Rev. E 86, 011909 (2012).
60. Colin, R., Zhang, R. & Wilson, L. G. Fast, high-throughput measurement of collective behavior in a bacterial population. \textit{J. R. Soc. Interface} \textbf{11}, 20140486 (2014).

61. Mello, B. A. & Tu, Y. Quantitative modeling of sensitivity in bacterial chemotaxis the role of coupling among different chemoreceptor species. \textit{Proc. Natl. Acad. Sci. USA} \textbf{100}, 8223-8228 (2003).

62. Sourjik, V. & Berg, H. C. Functional interactions between receptors in bacterial chemotaxis. \textit{Nature} \textbf{428}, 437-441 (2004).

63. Keymer, J. E., Endres, R. G., Skoge, M., Meir, Y. & Wingreen, N. S. Chemosensing in \textit{Escherichia coli}: two regimes of two-state receptors. \textit{Proc. Natl. Acad. Sci. USA} \textbf{103}, 1786-1791 (2006).

64. Hein, A. M., Brumley, D. R., Carrara, F., Stocker, R. & Levin, S. A. Physical limits on bacterial navigation in dynamic environments. \textit{J. R. Soc. Interface} \textbf{13}, 20150844 (2016).

65. Clausznitzer, D., Micali, G., Neumann, S., Sourjik, V. & Endres, R. G. Predicting chemical environments of bacteria from receptor signaling. \textit{PLoS Comput. Biol.} \textbf{10}, e1003870 (2014).

66. Kussell, E. & Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. \textit{Science} \textbf{309}, 2075-2078 (2005).

67. Acar, M., Mettetal, J. T. & van Oudenaarden, A. Stochastic switching as a survival strategy in fluctuating environments. \textit{Nat. Genet.} \textbf{40}, 471-475 (2008).

68. Berg, H. C. \textit{Random walks in biology} (Princeton Univ. Press, Princeton, 1983).

69. Wolfe, A. J. & Berg, H. C. Migration of bacteria in semisolid agar. \textit{Proc. Natl. Acad. Sci. USA} \textbf{86}, 6973-6977 (1989).

70. Clausznitzer, Diana, Oleksiuk, O., Løvdok, L., Sourjik, V. & Endres, R. G. Chemotactic response and adaptation dynamics in \textit{Escherichia coli}. \textit{PLoS Comput. Biol.} \textbf{6}, e1000784 (2010).

\textbf{Acknowledgements}

The authors would like to thank D. Clark, X. Fu, A. Waite, S. Kato, N. Frankel, D. C. Vural, and Y. Dufour for valuable discussions. Many simulations and data analysis routines were run on the clusters maintained by the Yale Center for High Performance Computing. This study was supported by the Paul G. Allen Family Foundation grant 11562 and by the National Institutes of Health grant 1R01GM106189.

\textbf{Author contributions}

T.E. and S.Z. designed the research. J.L. and T.E. derived the model. J.L., T.E. and S.Z. wrote the paper.

\textbf{Additional information}

\textbf{Competing financial interests:} The authors declare no competing financial interests.
Supplementary Information for
“Run and tumble navigation can be very fast”

J. Long$^{1,2}$, S. W. Zucker$^{3,4}$, and T. Emonet$^{*2,1}$

$^1$Department of Physics, Yale University, New Haven, CT 06520, USA.
$^2$Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520, USA.
$^3$Department of Computer Science, Yale University, New Haven, CT 06520, USA.
$^4$Department of Biomedical Engineering, Yale University, New Haven, CT 06520, USA.

May 11, 2016

*Email: thierry.emonet@yale.edu.
Supplementary Figure 1 | Effect of changing $\tau_D$. 5 sample trajectories (thin solid curves) and the mean over $10^4$ sample trajectories (thick lines) of non-dimensionalized position $x = X/(v_0 t_M)$ as a function of time $\tau = t/t_M$ for cells with matching (green $\tau_D = 1$) and non-matching (orange $\tau_D = 0.1$) reorientation time corresponding to the green and orange circles in Fig. ??c in the main text (same methods).
Supplementary Figure 2 | Robustness of results. (a) The same as Fig. ??c except with \( r_0 = 0.5 \). (b) The same as Fig. ??c except without assuming receptors in log-sensing range, i.e. \( \ln \frac{1+C/K_i}{1+C/K_a} \) is used rather than \( \ln C \). (c) The same as b, but additionally implements adaptation asymmetry, where the adaptation rate \( t_m^{-1} \) in equation (S4) depends on \( m^3 \). Here the adaptation is 3 times faster when \( m > m(C) \) than when \( m < m(C) \), and \( t_M \) is defined as the time scale when \( m < m(C) \).
### Fixed Parameters

| Name  | Definition         | Value    | Reference                  |
|-------|--------------------|----------|----------------------------|
| \( \epsilon_0 \) | Equation (S2)     | 6        | Shimizu et al. 2010\(^2\)  |
| \( \epsilon_1 \) | Equation (S2)     | -1       | Shimizu et al. 2010\(^2\)  |
| \( N \)   | Equation (S3)     | 6        | Shimizu et al. 2010\(^2\)  |
| \( K_i \)  | Equation (S3)     | 0.0182 \(mM\) | Shimizu et al. 2010\(^2\) |
| \( K_a \)  | Equation (S3)     | 3 \(mM\) | Shimizu et al. 2010\(^2\)  |
| \( \alpha \) | \( Y = \alpha a \) | 6        | Sneddon et al. 2012\(^3\) |
| \( K \)    | Equation (S6)     | 3.06 \(mM\) | Sneddon et al. 2012\(^4\)  |
| \( \epsilon_2 \) | Equation (S6)     | 40       | Sneddon et al. 2012\(^3\)  |
| \( \epsilon_3 \) | Equation (S6)     | 40       | Sneddon et al. 2012\(^3\)  |
| \( \omega \) | Equation (S6)     | 1.3 \(s^{-1}\) | Sneddon et al. 2012\(^3\) |
| \( v_0 \) | Running speed     | 20 \(\mu m/s\) | Sneddon et al. 2012\(^3\) |
| \( \delta \) | \( r = \frac{1}{1+\exp(-H(F-\delta))} \) | -0.04    | Equation (S8)               |
| \( H \)    | \( r = \frac{1}{1+\exp(-H(F-\delta))} \) | 4.9      | Equation (S8)               |
| \( r_0 \)  | Adapted probability to run | 0.8 | This study                  |

### Parameters in Fig. ??

| Name  | Definition                   | Value     | Reference                  |
|-------|------------------------------|-----------|----------------------------|
| \( t_M \) | Memory                      | 10 \(s\)  | Dufour et al. 2014\(^4\)   |
| \( D_R \)  | Rotational diffusion (run) | 0.00061 – 0.19 \(s^{-1}\) | Let \( \tau_D \) vary from 10\(^{-1.5}\) to 10\(^1\) |
| \( D_T \)  | Rotational diffusion (tumble)| 0.023 – 7.1 \(s^{-1}\) | Keep \( D_T/D_R \approx 37\(^5,6\) |
| \( L \)    | Gradient length scale       | 186 – 58800 \(\mu m\) | Let \( \tau_E \) vary from 10\(^{-1.5}\) to 10\(^1\) |

### Parameters in Fig. ??

| Name  | Definition                  | Value     | Reference                  |
|-------|------------------------------|-----------|----------------------------|
| \( L \)    | Initial gradient length scale | 1000 \(\mu m\) | This study                  |
| \( C_0 \)  | Exponential gradient source | 10 \(mM\) | This study                  |
| \( C_1 \)  | Linear gradient source      | 1 \(mM\) | This study                  |
| \( C_2 \)  | Localized source            | 1 \(mM\) | This study                  |
| \( a \)    | Slope of the linear gradient| 0.0001 \(\mu M/\mu m\) | This study                  |
| \( R_0 \)  | Spherical food source radius | 100 \(\mu m\) | This study                  |
| \( t_M \)  | Memory                      | 0.54 – 17 \(s\) | Let initial \( \tau_E \) vary from 3 to 0.1 |
| \( D_R \)  | Rotational diffusion (run)  | 0.0036 – 0.11 \(s^{-1}\) | Keep \( \tau_D = 1 \) fixed |
| \( D_T \)  | Rotational diffusion (tumble)| 0.13 – 4.2 \(s^{-1}\) | Keep \( D_T/D_R \approx 37\(^5,6\) |
Supplementary Notes

Supplementary Note 1: Model of bacterial chemotaxis. We use a standard model of *E. coli* chemotaxis in which the cell relays information from the external environment to the flagellar motor through a signalling cascade triggered by ligand-binding receptors. The receptors are described by a two-state model where the activity $a$ is determined by the free energy difference $F$ (in units of $k_B T$) between the active and inactive states

$$ a = \frac{1}{1 + e^F}. \quad (S1) $$

Two terms, $F = F_m(m) + F_C(C)$, contribute to the free energy. The former depends on the methylation level $m$ of the receptor

$$ F_m(m) = \epsilon_0 + \epsilon_1 m, \quad (S2) $$

with constants $\epsilon_0$ and $\epsilon_1$, while the latter depends on the ligand concentration $C$

$$ F_C(C) = N \ln \frac{1 + C/K_i}{1 + C/K_a}. \quad (S3) $$

where $K_a$ and $K_i$ are the ligand dissociation constants of the active and inactive states and $N$ is the gain of the receptor complex.

As the external environment signals changes in $F_C(C)$, the receptor adapts via methylation and demethylation to control $F_m(m) = F - F_C(C)$, trying to maintain an activity level $a_0$ independent of the environment. The methylation level follows

$$ \frac{dm}{dt} = -\frac{m - m_0(C)}{t_M}, \quad (S4) $$

where $t_M$ is the adaptation time and $m_0(C)$ is the methylation level that maintains $a_0$. Using equations (S1)-(S3), we see

$$ m_0(C) = \frac{1}{\epsilon_1} \left( F_0 - N \ln \frac{1 + C/K_i}{1 + C/K_a} - \epsilon_0 \right), \quad \text{with} \quad F_0 = -\ln \left( \frac{1}{a_0} - 1 \right). \quad (S5) $$

Assuming the signal transduction is fast, the receptor activity determines the concentration of a key protein CheY-P, $Y(a) = \alpha a$ with $\alpha$ constant. CheY-P then binds to the flagellar motor complex, modulating its switching between rotating clockwise (CW) or counterclockwise (CCW), which are Poisson processes with rates from CCW to CW as $\lambda_{CCW}(Y)$ and from CW to CCW as $\lambda_{CW}(Y)$

$$ \lambda_{CCW,CW}(Y) = \omega \exp \left( \mp G(Y) \right), \quad \text{with} \quad G(Y) = \frac{\epsilon_2}{4} - \frac{\epsilon_3}{2} \frac{1}{1 + K/Y}, \quad (S6) $$

where $\omega$, $\epsilon_2$, $\epsilon_3$, and $K$ are constants.

The motor state of each flagellum determines its conformation to be in one of normal, curly-1, and semi-coiled. The flagellar conformations in turn determine the cell’s motility state (run or tumble). For a single-flagellum cell, the motor states CCW and CW corresponds to the motility states run and tumble, and we can write $\lambda_R = \lambda_{CCW}$ and $\lambda_T = \lambda_{CW}$ as the switching rates from run to tumble and from tumble to run, respectively.
Supplementary Note 2: Derivation of equation (1) in the main text. Given constant Poisson switching rates $\lambda_R$ and $\lambda_T$, the probability to be running is determined by $r = \frac{\lambda_T}{\lambda_R + \lambda_T}$ and the characteristic time scale is $t_S = \frac{1}{\lambda_R + \lambda_T}$\textsuperscript{9}. When the rates change over time we can say the probability to run tends to follow this value. From equation (S6) and the definition $Y = \alpha \epsilon = \alpha / (1 + e^F)$ we can write

$$r(F) = \frac{1}{1 + \frac{2R}{\lambda_T}} = \frac{1}{1 + \exp \left( - \frac{\epsilon_2}{2} + \frac{\epsilon_3}{1 + \alpha (1 + e^F)} \right)}.$$  \hspace{2cm} (S7)

Thus the probability to run $r$ is a monotonically increasing function of the free energy $F$. Its shape is almost identical to the standard sigmoidal function $1/(1 + \exp (-H (F - \delta)))$ with a scaling $H$ and shift $\delta$. We match them by linearly expanding near $\delta$

$$r(F) \approx \frac{1}{1 + \Delta} + \frac{\Delta \epsilon_3 K e^{-\delta}}{(1 + \Delta)^2 (1 + \frac{K}{\alpha} (1 + e^{-\delta}))^2} (F - \delta) + O((F - \delta)^2),$$

where

$$\Delta = \exp \left( - \frac{\epsilon_2}{2} + \frac{\epsilon_3}{1 + \frac{K}{\alpha} (1 + e^{-\delta})} \right),$$

and

$$\frac{1}{1 + \exp (-H (F - \delta))} \approx \frac{1}{2} + \frac{1}{4} H (F - \delta) + O((F - \delta)^2).$$

To match the zeroth order we must have $\Delta = 1$. This yields $\delta$ to match the first order to get $H$

$$\delta = - \ln \left( \frac{(2\epsilon_3/\epsilon_2 - 1) \alpha}{K - 1} \right)$$

$$H = \epsilon_3 e^{-\delta} \left( \frac{\epsilon_2}{2 \epsilon_3} \right)^2 \frac{K}{\alpha}.$$ \hspace{2cm} (S8)

For the parameters chosen we have $\delta \approx 0$ (see Supplementary Table). So we can write equation (??) in the main text.

Supplementary Note 3: Derivation of equation (2) in the main text. We substitute the definitions (S2), (S3), and (S5) into the adaptation dynamics equation (S4) and get

$$\frac{dF}{dt} - N \frac{d}{dt} \ln \left( 1 + \frac{C}{K_i} \right) = - \frac{F - F_0}{t_M},$$

where $d/dt$ is taken along the path of the cell, so $d/dt = \partial / \partial t + \hat{X} \cdot \nabla$. Substitute this into the above we obtain equation (??) in the main text.

Supplementary Note 4: Derivation of the master equation model. We define $P_R(X, \hat{u}, F, t)$ and $P_T(X, \hat{u}, F, t)$ as the probability distributions at time $t$ to be running or tumbling at position $X$ in direction $\hat{u}$ with internal variable $F$. A two-state stochastic model\textsuperscript{9} is given by

$$\partial_t P_R = - \partial_F \left( F |_{\text{run}} P_R \right) + \nabla_\hat{u}^2 (D_R P_R) - \lambda_R P_R + \lambda_T P_T - \nabla \cdot (v_0 \hat{u} P_R)$$

$$\partial_t P_T = - \partial_F \left( F |_{\text{tumble}} P_T \right) + \nabla_\hat{u}^2 (D_T P_T) + \lambda_R P_R - \lambda_T P_T,$$ \hspace{2cm} (S9)
where $\dot{F}|_{\text{run,tumble}}$ are defined by equation (29) in the main text with constant speed $v_0$ in a run and zero speed in a tumble. Let $s = \hat{u} \cdot \hat{X} = \cos \theta$. In a static exponential gradient with length scale $L = 1/||\nabla \ln C||$, the internal dynamics satisfies

$$
\dot{F}|_{\text{run}} = -\frac{F - F_0}{t_M} + \frac{Nv_0}{L} s
$$

$$
\dot{F}|_{\text{tumble}} = -\frac{F - F_0}{t_M}.
$$

(S10)

On the other hand, the angular dynamics undergoes rotational diffusion with operator $\nabla_u^2$. In $n$-dimensional space we iteratively write down the Laplace-Beltrami operator\(^10\) as

$$
\nabla_{S^{n-1}}^2 = (\sin \theta)^{2-n} \partial_\theta ((\sin \theta)^{n-2} \partial_\theta) + (\sin \theta)^{-2} \nabla_{S^{n-2}}^2.
$$

Since $0 < \theta < \pi$, with change of variable $s = \cos \theta$ we can write $\sin \theta = \sqrt{1 - s^2}$ and $\partial_\theta = -\sqrt{1 - s^2} \partial_s$. Then the polar angle part as

$$
\hat{L}_s = (1 - s^2)^{\frac{3-n}{2}} \partial_s \left( (1 - s^2)^{-\frac{n+1}{2}} \partial_s \right).
$$

(S11)

Since the gradient varies in one direction only we focus on motion in the gradient direction and integrate the probability over all other directions. Thus $\nabla \cdot \hat{u} = s \partial_x$ and $\nabla_u^2 = \hat{L}_s$. Using $f = H F$, $r = \lambda_T / (\lambda_R + \lambda_T)$ and $t = 1/ (\lambda_R + \lambda_T)$, we get

$$
\partial_t P_R = -\partial_f \left( \left( -\frac{f - f_0}{t_M} + \frac{Nv_0}{L} s \right) P(R) \right) + D_R \hat{L}_s P_R - \frac{1-r}{t_s} P_R + \frac{r}{t_s} P_T - s \partial_x (v_0 P_R)
$$

$$
\partial_t P_T = -\partial_f \left( -\frac{f - f_0}{t_M} P_T \right) + D_T \hat{L}_s P_T + \frac{1-r}{t_s} P_R - \frac{r}{t_s} P_T.
$$

(S12)

**Supplementary Note 5: Derivation of $V_D$, equation (6) in the main text.** At steady state we eliminate $\partial_t$ and $\partial_s$ and integrate over $s$, the Fokker-Planck equation (??) in the main text becomes

$$
0 = -\partial_f \left( -(f - f_0) \int \bar{P} w(s) ds + \frac{r(f)}{\tau_E} \int s \bar{P} w(s) ds \right),
$$

where the bar indicates steady state. By the boundary conditions that $P \to 0$ at $\pm \infty$, we must have

$$
r(f) \int s \bar{P} w(s) ds = \tau_E (f - f_0) \int \bar{P} w(s) ds.
$$

(S13)

From the third term on the right of the Fokker-Planck equation (??) in the main text, the spatial flux is $r(f)s$ and the drift velocity is its average over the distribution. Thus we get the drift velocity as equation (??) in the main text

$$
V_D = \langle rs \rangle = \int \int r(f) s \bar{P} w(s) ds df = \tau_E \int \int (f - f_0) \bar{P} w(s) ds df = \tau_E \langle f - f_0 \rangle.
$$

(S14)
Supplementary Note 6: Derivation of the analytical solution to Fokker-Planck equation (??) by angular moment expansion. The eigenvalue problem of the angular operator $\hat{L}_s$, defined in equation (S11), is

$$(1 - s^2)y'' - (n - 1) sy' = \lambda y.$$  

We identify this as the Gegenbauer differential equation\textsuperscript{11}, with eigenfunctions the Gegenbauer polynomials $C_k^{(n/2-1)}(s)$ and the corresponding eigenvalues $\lambda_k = -k(k+n-2)$. When $n = 3$ they are Legendre polynomials with eigenvalues $-k(k+1)$. The first few are

$$
\begin{align*}
C_0^{(n/2-1)}(s) &= 1 \\
C_1^{(n/2-1)}(s) &= (n-2)s \\
C_2^{(n/2-1)}(s) &= \frac{n-2}{2}(ns^2 - 1).
\end{align*}
$$

The Gegenbauer polynomials are orthogonal in the sense that

$$
\int_{-1}^{1} C_k^{(n/2-1)}(s) C_l^{(n/2-1)}(s) \left(1 - s^2\right)^{\frac{n-3}{2}} ds = N_k^{(n/2-1)},
$$

where the normalization constants are $N_k^{(n/2-1)} = \frac{\pi^{2-n}(k+n-3)!}{k!(2k+n-2)(\Gamma(n/2-1))^2}$. When $n = 3$ they are $\frac{2}{2k+1}$, those of Legendre polynomials.

The weight in the integration above is consistent with the geometry on an $(n-1)$-sphere $S^{n-1}$, whose the volume element are iteratively defined\textsuperscript{10} as

$$
d_{S^{n-1}}\omega = (\sin \theta)^{n-2} \, d\theta d_{S^{n-2}}\omega.
$$

After a change of variable $s = \cos \theta$ and integrating over all remaining dimensions, we see that any integration of $s$ should carry a weight

$$
w(s) ds = \left(1 - s^2\right)^{\frac{n-3}{2}} ds.
$$

From the above, we can write any function of $s$, in particular the probability distribution $P$, as a series of Gegenbauer polynomials. When $n = 3$ this is the Fourier-Legendre Series.

$$
P(x, f, s, \tau) = \sum_{k=0}^{\infty} p_k(x, f, \tau) \frac{C_k^{(n/2-1)}(s)}{\sqrt{N_0^{(n/2-1)} N_k^{(n/2-1)}}} \\
= \frac{1}{N_0^{(n/2-1)}} \left( p_0 + p_1 \sqrt{ns} + p_2 \sqrt{\frac{n+2}{n-1} \frac{ns^2 - 1}{2}} + \cdots \right),
$$

where

$$
p_k(x, f, \tau) = \int_{-1}^{1} \frac{N_0^{(n/2-1)}}{N_k^{(n/2-1)}} C_k^{(n/2-1)}(s) P(x, f, s, \tau) \left(1 - s^2\right)^{\frac{n-3}{2}} ds.
$$
The extra factor $\sqrt{N_0^{(n/2-1)}}$ ensures $p_0 = \int_{-1}^{1} P(1 - s^2)^{n-3} ds$ is the same as the marginal distribution. From now on we denote the marginal distribution $p(f) = p_0(f)$. Also, from this definition $p_1 = \sqrt{n} \int_{-1}^{1} sP(1 - s^2)^{n-3} ds$.

Substitute the expansion (S17) into the Fokker-Planck equation (??) in the main text and use the orthogonality equation (S15), we obtain

$$\partial_{x} p_k = - \partial_{f} \left( -(f - f_0) p_k + \frac{r(f) \hat{s}_{kl} p_l}{\tau_{E}} \right) + \frac{\lambda_{k}}{(n - 1) \tau_{D}(f)} p_k - \partial_{x} (\hat{s}_{kl} p_l) , \quad (S18)$$

where $\hat{s}_{kl} = \sqrt{\frac{k+1}{2(2k+n-4)(2k+n-2)}} \delta_{k-1,l} + \sqrt{\frac{(k+1)(k+n-2)}{(2k+n-2)(2k+n)}} \delta_{k+1,l}$ (summation over $l$ implied) is an operator relating neighboring orders. It comes from the positive feedback term. When $n = 3$ it is $\hat{s}_{kl} = \frac{k}{\sqrt{4(k+1)^2-1}} \delta_{k-1,l} + \frac{k+1}{\sqrt{4(k+1)^2-1}} \delta_{k+1,l}$. The first few equations are

$$\begin{align*}
\partial_{x} p &= - \partial_{f} \left( -(f - f_0) p + \frac{r(f)}{\tau_{E}} \right) - r(f) \partial_{x} \left( \frac{p}{\sqrt{n}} \right) \\
\partial_{x} p_1 &= - \partial_{f} \left( -(f - f_0) p_1 + \frac{r(f)}{\tau_{E}} \left( \frac{1}{\sqrt{n}} p + \sqrt{\frac{2(n-1)}{n(n+2)} p_2} \right) \right) \\
&\quad - \frac{1}{\tau_{D}(f)} p_1 - r(f) \partial_{x} \left( \frac{1}{\sqrt{n}} p + \sqrt{\frac{2(n-1)}{n(n+2)} p_2} \right) \\
\partial_{x} p_2 &= - \partial_{f} \left( -(f - f_0) p_2 + \frac{r(f)}{\tau_{E}} \left( \sqrt{\frac{2(n-1)}{n(n+2)} p_1} + \sqrt{\frac{3n}{(n+2)(n+4)} p_3} \right) \right) \\
&\quad - \frac{2n}{(n-1) \tau_{D}(f)} p_2 - r(f) \partial_{x} \left( \sqrt{\frac{2(n-1)}{n(n+2)} p_1} + \sqrt{\frac{3n}{(n+2)(n+4)} p_3} \right) .
\end{align*} \quad (S19)$$

In the definition of $\hat{s}_{kl}$, when $k \gg 1$ the non-zero entries approach a constant $1/2$. This means for large $k$ the coefficients $p_k$ in equation (S18) evolve similarly except that higher orders decay with faster rates $k k_{n-2} / (n-1) \tau_{D}$. Therefore when $\tau_{D0}$ is small we can neglect the 2nd and higher orders, which closes the infinite series of moment equations and leaves two equations concerning the zeroth and first marginal moments in $s$, $p(x,f,\tau)$ and $p_1(x,f,\tau)$ respectively. The analytical solution at steady state can be found as

$$\bar{p}(f) = \frac{1}{Z} \frac{r(f)}{\tau_{E}} \exp \left( - \int_{f}^{f_0} \frac{f - f_0}{\tau_{D}(f)} df \right) , \quad (S20)$$

where $Z$ is a normalization constant.

**Supplementary Note 7: Derivation of the distribution $\bar{p}(f)$ when the negative feedback dominates.**

We expand the steady state solution in orders of $\frac{1}{\tau_{E}} \ll 1$ and $\tau_{D0} \ll 1$. We first write the steady state distribution (S20) as
\[ p(f) = \frac{1}{Z} \frac{r(f)}{\tau_E} \left( \frac{1}{n} \left( \frac{r(f)}{\tau_E} \right)^2 - (f - f_0)^2 \right) \exp \left( - \int f f_1 - f_0 \frac{r(f_1)}{\tau_E} \left( \frac{1}{n} \left( \frac{r(f_1)}{\tau_E} \right)^2 - (f_1 - f_0)^2 \right) \right) \]  
(S21)

\[ = B(f) \exp \left( - \int f A(f_1) df_1 \right). \]

We write the Taylor expansion of the integrand in the exponent

\[ A(f) = \frac{f - f_0}{\tau_D(f)} \left( \frac{1}{n} \left( \frac{r(f)}{\tau_E} \right)^2 - (f - f_0)^2 \right) - \frac{n \tau_E^2}{r_0^2 \tau_D} (f - f_0) + \left( 1 + O\left( \frac{1}{\tau_E} \right) \right) \sum_{m=1}^{\infty} \frac{n^{m+1} \tau_E^{2m+2}}{r_0^{2m+2} \tau_D} (f - f_0)^{2m+1} \]

\[ - \left( 1 + O\left( \frac{1}{\tau_E} \right) \right) \sum_{m=1}^{\infty} \frac{n^{m-2m} \tau_E^{2m} r_0}{r_0^2 \tau_D} \left( 2m + \frac{r_0 \tau_D^2}{r_0^2 \tau_D} \right) (f - f_0)^{2m} \]

where \( \tau' = d/df \). We see that if we define

\[ \sigma^2 = \frac{n \tau_E^2 \tau_D}{r_0^2}, \]  
(S22)

the first term in \( A(f) \) will give \( - \int A(f) df = - \frac{(f - f_0)^2}{2 \sigma^2} + \ldots \). If we can show that the rest of the terms are small when \( \frac{1}{\tau_E} < 1 \) and \( \tau_D < 1 \), we can write \( p(f) \) as a small deviation from a Gaussian.

Indeed, if we consider the integration range \( |f - f_0| \sim \sigma \sim O(\sqrt{\tau_D}/\tau_E) \), we can write

\[ - \int_{f_0}^{f} A(f_1) df_1 = - \frac{(f - f_0)^2}{2 \sigma^2} - \sum_{m=1}^{\infty} \frac{\tau_D^m}{2m+2} \frac{(f - f_0)^{2m+2}}{\sigma^{2m+2}} \]

\[ + \sum_{m=1}^{\infty} \frac{r_0 \tau_D^{m-1}}{r_0^2 \tau_D} \left( 2m + \frac{r_0 \tau_D^m}{r_0^2 \tau_D} \right) \frac{(f - f_0)^{2m+1}}{\sigma^{2m}} + O\left( \frac{1}{\tau_E^2} \right) \]

\[ = - \frac{(f - f_0)^2}{2 \sigma^2} + \frac{r_0^'}{r_0} \left( 2 \frac{r_0 \tau_D^3}{r_0^2 \tau_D} \right) \frac{(f - f_0)^3}{\sigma^2} - \frac{\tau_D}{4} \frac{(f - f_0)^4}{\sigma^4} \]

\[ + \frac{r_0^' \tau_D}{r_0} \frac{(f - f_0)^5}{\sigma^4} + O\left( \frac{1}{\tau_E^2} \right) + O(\tau_D^2), \]

and similarly

\[ B(f) = \frac{n \tau_E}{r_0} \left( 1 - \frac{r_0^'}{r_0} (f - f_0) + \tau_D \frac{(f - f_0)^2}{\sigma^2} - 3 \tau_D \frac{r_0^' (f - f_0)^3}{\sigma^2} + O\left( \frac{1}{\tau_E^2} \right) + O(\tau_D^2) \right) \]

Substitute the above back into equation (S21), taking care of the orders of all cross terms, we obtain
\[ p(f) = C e^{-\frac{(f-f_0)^2}{2\sigma^2}} \left( 1 - \frac{r'_0}{r_0} (f-f_0) + \tau_{D0} \frac{(f-f_0)^2}{\sigma^2} \right) \]
\[ + \frac{r'_0}{3r_0} \left( 2 - 9\tau_{D0} + \frac{r_0 \tau'_{D0}}{r'_0 \tau_{D0}} (f-f_0)^3 \right) \frac{1}{\sigma^2} - \tau_{D0} \frac{(f-f_0)^4}{4 \sigma^4} \]
\[ + \frac{r'_0}{60r_0} \left( 103 + 32 \frac{r_0 \tau'_{D0}}{r'_0 \tau_{D0}} \right) \tau_{D0} \frac{(f-f_0)^5}{\sigma^4} \]
\[ - \frac{r'_0}{12r_0} \left( 2 + \frac{r_0 \tau'_{D0}}{r'_0 \tau_{D0}} \right) \tau_{D0} \frac{(f-f_0)^7}{\sigma^6} + O\left(\frac{1}{\tau_E^2}\right) + O\left(\tau_{D0}^2\right). \]  

(S23)

We notice from equation (S21) that the range of distribution is bounded by \( f_L \) and \( f_U \), defined by

\[ f_L - f_0 = -\frac{1}{\sqrt{n}} \frac{r(f_L)}{\tau_E}, \]
\[ f_U - f_0 = \frac{1}{\sqrt{n}} \frac{r(f_U)}{\tau_E}. \]

Since \( \sigma = \frac{r_0 \tau_{D0}}{\sqrt{n} \tau_E} \ll \frac{r_0}{\sqrt{n} \tau_E} \), we see that the integration range is much larger than the standard deviation of the Gaussian factor, and thus can be considered from \(-\infty\) to \(\infty\). Therefore we get the normalization constant

\[ C = \frac{1}{1 + \frac{\tau_{D0}}{4} + O\left(\frac{1}{\tau_E}\right) + O\left(\tau_{D0}^2\right)}. \]  

(S24)

Substitute equations (S23) and (S24) into equation (S14) and carry out the integrals

\[ V_D = \frac{r_0 \tau_{D0}}{n \tau_E} \left( 1 - \frac{3}{4} \tau_{D0} \right) \frac{r'_0 + r_0 \tau'_{D0}}{1 + \frac{\tau_{D0}}{4} + O\left(\frac{1}{\tau_E}\right) + O\left(\tau_{D0}^2\right)} + O\left(\tau_{D0}^2\right). \]  

(S25)

Finally, noticing that by the definition of \( \tau_D \) in equation (???) in the main text

\[ \tau_D(f) = \tau_{D0} \frac{r_0 D_R + (1-r_0) D_T}{r(f) D_R + (1-r(f)) D_T}, \]

we can get

\[ \tau'_{D0} = \tau_{D0} \frac{D_T - D_R}{r_0 D_R + (1-r_0) D_T} r'_0. \]

Therefore

\[ r'_0 + r_0 \frac{\tau'_{D0}}{\tau_{D0}} = \tau_{D0} \frac{D_T}{D_T - D_R} \left( \frac{r_0 D_R + (1-r_0) D_T}{D_T - D_R} + r_0 \right) \]
\[ = \frac{\tau'_{D0}}{\tau_{D0}} \frac{D_T}{D_T - D_R}. \]
Taking $D_T \gg D_R$, we put this back into equation (S25) and get

$$V_D = \frac{r_0 \tau_D'}{n\tau_E} \left[ 1 - \frac{3}{4} \frac{\tau_D}{\tau_D'} + O\left(\frac{1}{\tau_E}\right) + O\left(\frac{\tau_D}{\tau_D'}\right) \right] \quad (S26)$$

When converted back to real units ($t$ instead of $\tau = t/t_M$), the highest-order term is identical, except for notations, to equation (3) in Dufour et al.\textsuperscript{4} obtained from a different approach. It can also be reduced to equation (12) in Si et al.\textsuperscript{7} by assuming a high running probability $r \approx 1$ and a long memory ($t_M \gg t_{0m}$). It agrees with equation (6.24) in Erban & Othmer\textsuperscript{12} and equation (16) in Franz et al.\textsuperscript{13} with appropriate inclusion of rotational diffusion.

**Supplementary Note 8: Derivation of Langevin equations (8).** The geometric weight factor $w(s)$ in equation (S16) for angular integration complicates the issue to get the Langevin system from the Fokker-Planck equation. To derive $s$-dynamics, we start with the angular part of the Fokker-Planck equation (??) in the main text

$$\partial \tau P = \frac{\hat{L}_s P}{(n-1)\tau_D(f)} + \ldots$$

Multiplying an arbitrary function $A(s)$ and integrating over all dimensions, we obtain

$$\int dx \int df \int_{-1}^{1} w(s)ds A(s) \partial \tau P = \int dx \int df \int_{-1}^{1} w(s)ds A(s) \frac{(1 - s^2)^{3-n}}{(n-1)\tau_D(f)} \partial s \left( (1 - s^2)^{\frac{n-1}{2}} \partial s P \right)$$

$$= - \int dx \int df \int_{-1}^{1} w(s)ds \frac{sP}{\tau_D(f)} \partial s A(s)$$

$$+ \int dx \int df \int_{-1}^{1} w(s)ds \frac{(1 - s^2)P}{(n-1)\tau_D(f)} \partial^2 s A(s).$$

To apply the standard result of equivalence between Fokker-Planck equations and Langevin equations, we need to change the measure in $s$-space to unity. This prompts the definition $Q(s,t) = w(s) \int P(y, f, s, t)dx df$ so that the above becomes

$$\int_{-1}^{1} ds A(s) \partial s Q = - \int_{-1}^{1} ds \frac{sQ}{\tau_D(f)} \partial s A(s) + \int_{-1}^{1} ds \frac{(1 - s^2)Q}{(n-1)\tau_D(f)} \partial^2 s A(s)$$

$$= \int_{-1}^{1} ds A(s) \partial s \frac{sQ}{\tau_D(f)} + \int_{-1}^{1} ds A(s) \partial^2 s \frac{(1 - s^2)Q}{(n-1)\tau_D(f)},$$

where we integrated by parts and discarded boundary terms. Since $A(s)$ is an arbitrary function, we can write down the Fokker-Planck equation

$$\partial_s Q = \frac{sQ}{\tau_D(f)} + \frac{(1 - s^2)Q}{(n-1)\tau_D(f)},$$

12
which is equivalent\textsuperscript{9} to the Langevin equation

\[
\frac{ds}{d\tau} = -\frac{s}{\tau_D(r)} + \sqrt{\frac{2(1 - s^2)}{(n - 1) \tau_D(r)}} \eta(\tau). \tag{S27}
\]

where \(\eta(\tau)\) denotes the Gaussian white noise with \(\langle \eta(\tau_1)\eta(\tau_2) \rangle = \delta(\tau_1 - \tau_2)\).

The other two variables follow standard results\textsuperscript{9} from the Fokker-Planck equation \(\text{(??)}\) in the main text

\[
\begin{align*}
\frac{df}{d\tau} &= -(f - f_0) + \frac{r(f)s}{\tau_E}, \\
\frac{dy}{d\tau} &= r(f)s.
\end{align*} \tag{S28}
\]

Now we change variables according to the definitions \(r(f) = 1/(1 + \exp(-f))\) and \(v = rs\), and derive from the above dynamics in equation (S27) and (S28) to get the Langevin equation \(\text{(??)}\) in the main text

\[
\begin{align*}
\frac{dr}{d\tau} &= r(1 - r) \left( -(f(r) - f_0) + \frac{v}{\tau_E} \right) \\
\frac{dv}{d\tau} &= v(1 - r) \left( -(f(r) - f_0) + \frac{v}{\tau_E} \right) - \frac{v}{\tau_D(r)} + \sqrt{\frac{2(v^2 - v^2)}{(n - 1) \tau_D(r)}} \eta(\tau) \\
\frac{dy}{d\tau} &= v,
\end{align*} \tag{S29}
\]

where \(f(r) = -\log(1/r - 1)\).
Supplementary Methods

Agent-based simulation. The simulation is performed using Euler method. At each time, each cell moves forward or stays according to its motility state (run or tumble), which also determines whether its direction change with rotational diffusion coefficients $D_R$ or $D_T$. At the new position the ligand concentration determines the adapted methylation level in equation (S5). Then the new methylation is updated as equation (S4), which determines the receptor activity $a(t)$ according to equations (S1)-(S3). This in turn determines the switching rates according to equation (S6) and $Y(t) = \alpha a(t)$, and the motor state is changed if a random number drawn exceeds the probability to switch over the time step. After that the motor state determines the flagellum state and subsequently the cell motility state with rules and parameters as in 3, completing one time step.
Supplementary References
1. Clausznitzer, Diana, Oleksiuk, O., Løvdok, L., Sourjik, V. & Endres, R. G. Chemotactic response and adaptation dynamics in *Escherichia coli*. *PLoS Comput. Biol.* 6, e1000784 (2010).
2. Shimizu, T. S., Tu, Y. & Berg, H. W. A modular gradient-sensing network for chemotaxis in *Escherichia coli* revealed by responses to time-varying stimuli. *Mol. Syst. Biol.* 6, 382-395 (2010).
3. Sneddon, M. W., Pontius, W. & Emonet, T. Stochastic coordination of multiple actuators reduce latency and improves chemotactic response in bacteria. *Proc. Natl. Acad. Sci. USA* 109, 805-810 (2012).
4. Dufour, Y., Fu, X., Hernandez-Nunez, L. & Emonet, T. Limits of feedback control in bacterial chemotaxis. *PLoS Comput. Biol.* 10, e1003694 (2014).
5. Berg, H. C. *Random walks in biology* (Princeton Univ. Press, Princeton, 1983).
6. Saragosti, J., Silberzan, P. & Buguin A. Modeling *E. coli* tumbles by rotational diffusion. Implications for chemotaxis. *PLoS One* 7, e35412 (2012).
7. Si, G., Wu, T., Ouyang, Q. & Tu, Y. Pathway-based mean-field model for *Escherichia coli* chemotaxis. *Phys. Rev. Lett.* 109, 048101 (2012).
8. Tu, Y. Quantitative modeling of bacterial chemotaxis signal amplification and accurate adaptation. *Annu. Rev. Biophys.* 42, 337-359 (2013).
9. Gardiner, C. W. *Handbook of stochastic methods for physics, chemistry and the natural sciences* (Springer Berlin Heidelberg, Berlin, 2004).
10. Jost, J. *Riemannian geometry and geometric analysis* (Springer Berlin Heidelberg, Berlin, 2011).
11. Abramowitz, M. & Stegun, I. *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (Dover Publications, Mineola, 1964).
12. Erban, R. & Othmer, H. G. From individual to collective behavior in bacterial chemotaxis *SIAM J. Appl. Math.* 65, 361-391 (2004).
13. Franz, B., Xue, C., Painter, K. J. & Erban, R. Travelling waves in hybrid chemotaxis models. *Bull. Math. Biol.* 76, 377-400 (2014).