

## *E. coli* chemotaxis is information-limited

H.H. Mattingly<sup>1,2,†</sup>, K. Kamino<sup>1,2,†</sup>, B.B. Machta<sup>3,4,\*</sup>, T. Emonet<sup>1,2,3,\*</sup>

### Affiliations:

<sup>1</sup> Department of Molecular, Cellular, and Developmental Biology, Yale University.

<sup>2</sup> Quantitative Biology Institute, Yale University.

<sup>3</sup> Department of Physics, Yale University.

<sup>4</sup> Systems Biology Institute, West Campus, Yale University.

† Equal contribution.

\* Correspondence to: [thierry.emonet@yale.edu](mailto:thierry.emonet@yale.edu) and [benjamin.machta@yale.edu](mailto:benjamin.machta@yale.edu)

### Abstract

Organisms must acquire and use environmental information to guide their behaviors. However, it is unclear whether and how information quantitatively limits behavioral performance. Here, we relate information to behavioral performance in *Escherichia coli* chemotaxis. First, we derive a theoretical limit for the maximum achievable gradient-climbing speed given a cell's information acquisition rate. Next, we measure cells' gradient-climbing speeds and the rate of information acquisition by the chemotaxis pathway. We find that *E. coli* make behavioral decisions with much less than the 1 bit required to determine whether they are swimming up-gradient. However, they use this information efficiently, performing near the theoretical limit. Thus, information can limit organisms' performance, and sensory-motor pathways may have evolved to efficiently use information from the environment.

## Main text

Organisms' survival depends on their ability to perform behavioral tasks. These tasks can be challenging because they require that the organism measure signals in its environment and respond appropriately, in real time. Information theory is a natural language for quantifying the fidelity of measurements and responses, but it is unclear how an abstract quantity like information might limit an organism's performance at real-world tasks. Past studies have used information theory to understand the maximum amount of information biological systems can acquire and transmit about environmental signals<sup>1-6</sup> and have shown that they can approach certain theoretical limits<sup>6-9</sup>. But high information transfer is not sufficient for high performance because not all of the information contained in the signal is behaviorally relevant, and not all of it is appropriately acted on<sup>10</sup>. What limits does information place on performance, and how efficiently do organisms use the information they acquire relative to these limits?

We address these questions using one of the best-understood behaviors in biology: bacterial chemotaxis. The bacterium *Escherichia coli* alternates between "runs," which propel the cell forward, and "tumbles," which randomly reorient its swimming direction<sup>11</sup> (Fig. 1). Along its path, *E. coli* continuously senses the concentration  $c(t)$  of chemoattractant it encounters using transmembrane receptors. Relative changes in concentration<sup>12</sup>  $s(t) = \frac{d}{dt} \log(c)$  along the cell's swimming trajectory, which we define to be the "signal", induce changes in activity  $a(t)$  of receptor-associated CheA kinases. CheA activity goes on to modulate transitions in the cell's motor behavior  $m(t)$  between run and tumble states via a signal transduction pathway<sup>13,14</sup>. If the relative concentration of attractant is increasing ( $s(t) > 0$ ), the cell tends to run for longer on average, thereby biasing its random motion up the gradient<sup>11</sup>. However, noise in sensing and signal transduction corrupt the signal<sup>15-18</sup>. Since the goal of chemotaxis is to climb chemical gradients, performance can be quantified by the cell's drift velocity  $v_d$  up a static gradient. *E. coli* chemotaxis has been studied extensively, but the amount of information a cell acquires about chemical signals and its relationship to chemotactic performance are unclear.

Quantifying information transfer during behavioral tasks presents new challenges. First, bacteria continuously make decisions about whether to tumble. Thus, unlike most other studies in biological systems<sup>1-3,5,6</sup>, we cannot use the one-shot or instantaneous mutual information<sup>19</sup> between signal  $s(t)$  and behavior  $m(t)$  as a measure of information transfer. Instead, we need an information rate<sup>20</sup>. Second, the signals are generated by the cell's own motion in the gradient<sup>21</sup>. Thus, a natural extension, the mutual information rate  $\dot{M}I$  between signal  $s(t)$  and behavior  $m(t)$ , is unsuitable because signal-motor correlations make it non-zero even for non-responsive cells (Fig. 1AB). We address these challenges by isolating the information that flows from signal to behavior. The mutual information rate can be decomposed into the sum of two directed information terms<sup>22</sup> (SI), the transfer entropy<sup>23</sup> rates from behavior to signal  $\dot{I}_{m \rightarrow s}$  and from signal to behavior  $\dot{I}_{s \rightarrow m}$ , or  $\dot{M}I = \dot{I}_{m \rightarrow s} + \dot{I}_{s \rightarrow m}$ . The first term  $\dot{I}_{m \rightarrow s}$  quantifies the feedback of behavior onto signal. The second term,  $\dot{I}_{s \rightarrow m}$ , which we refer to here as the information rate, in bits/s, measures the directed influence of the signal on behavior and must be nonzero for the cell to climb the gradient.

We first used rate-distortion theory<sup>6,24,25</sup> to derive the theoretical bound on performance imposed by information transfer. To do this, we constructed a model of run-and-tumble navigation that is abstracted from *E. coli*'s molecular implementation (Fig. 1A; SI). During runs, cells swim with an effective, constant speed  $v_0$  (Fig. 1A) and lose direction with rotational diffusion coefficient  $D_r$ . During tumbles, they reorient randomly with directional persistence  $\alpha$ . In the absence of a gradient, cells switch from run to tumble and vice versa with constant rates  $\lambda_{R0}$  and  $\lambda_T$ . In a gradient, tumbles are instead initiated at a rate  $\lambda_R(\{s\})$ , which depends on the past history of signals  $\{s\}$ . Throughout, we consider navigation in shallow gradients, where information is most likely to be a limitation on chemotactic performance.

Using this model, we derived how performance  $v_d$  and information transfer  $\dot{I}_{s \rightarrow m}$  depend on the tumble rate response  $\lambda_R(\{s\})$ . While any response to signals implies information transfer, it does not necessarily imply high drift speed. We derived the maximum drift speed  $v_d$  possible given an information rate  $\dot{I}_{s \rightarrow m}$  by optimizing over responses  $\lambda_R(\{s\})$ . In the SI, we show that in this model the optimal response only depends on the current rate of change of log-concentration,  $s(t)$ . Because bacteria necessarily make comparisons of concentrations over a finite time to infer  $s(t)$  from stochastic arrivals of ligand molecules at the cell surface<sup>15</sup>, they can't implement the optimal behavioral response. Nevertheless, the performance achieved by any response is bounded by (Fig. 1C):

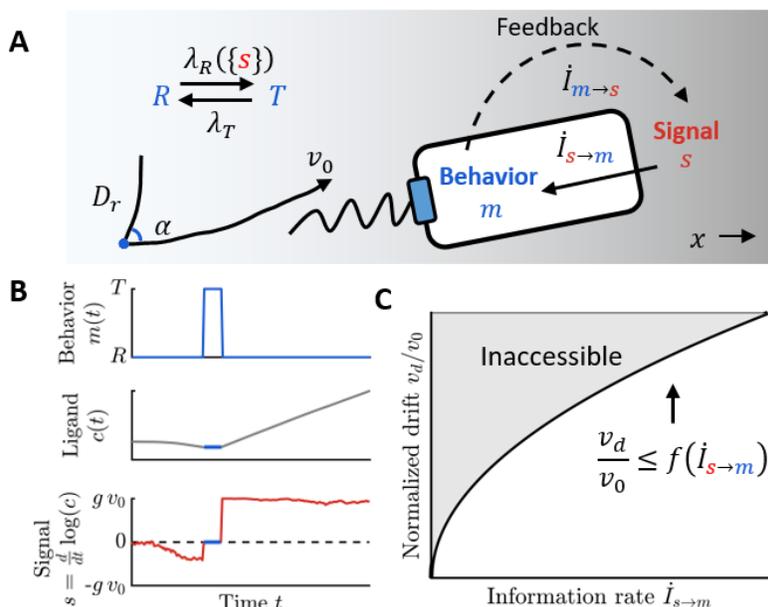
$$v_d/v_0 \leq F(\boldsymbol{\theta}) \left( \frac{\log(2) \dot{I}_{s \rightarrow m}}{12 D_r} \right)^{\frac{1}{2}}, \text{ where } 0 \leq F(\boldsymbol{\theta}) \leq 1, \boldsymbol{\theta} = \{\lambda_{R0}, \lambda_T, \alpha\}. \quad (1)$$

$F(\boldsymbol{\theta}) = \frac{(1-\alpha) \lambda_{R0}}{(1-\alpha) \lambda_{R0} + 2 D_r} \left( 8 \frac{D_r}{\lambda_{R0}} P_{run} \right)^{\frac{1}{2}}$  is a dimensionless function of the behavioral parameters (see Fig. 1C for a definition). This expression makes rigorous the intuition that information transfer sets a limit on how fast a cell can climb a gradient<sup>15,26,27</sup>.

To quantify the bound above, we measured the rotational diffusion coefficient  $D_r$  and behavioral parameters  $\boldsymbol{\theta}$  from trajectories of swimming *E. coli* cells (Figs. S1, S2). Individual cells in a clonal population exhibit nongenetic differences in behavioral parameters  $\boldsymbol{\theta}$ <sup>28-30</sup>, which in *E. coli* are highly correlated with  $P_{run}$ , the fraction of time the cell is running  $P_{run}$ <sup>28,30</sup> (Fig. S1,S2). From this data, we find  $F(\boldsymbol{\theta}) = 0.531 \pm 0.005$  for the median phenotype ( $\pm$  one standard error; parameters are in Table S1). Perhaps surprisingly, the bound predicts that run-and-tumble navigation is theoretically possible with very small information rates: a hundredth of a bit per second is sufficient to climb gradients at  $\sim 6\%$  of the run speed. This is far less than the 1 bit per run ( $\sim 1$  bit/s) required to distinguish whether concentration is currently increasing or decreasing before every tumble decision<sup>20</sup>.

Our central questions are: how much sensory information do *E. coli* acquire, and how efficiently do they use that information? Eqn. 1 was derived by considering a model of *E. coli* in which signaling details have been coarse-grained away. But information about signals is acquired upstream in the *E. coli* chemotaxis pathway and must be communicated to the motors. Since this signal transduction process can only lose information<sup>31</sup>, the information available at any

intermediate step within the signaling pathway also places a limit on the cell's performance, as in Eqn. 1. Intuitively, in an information-efficient cell, most sensory information acquired upstream would be preserved at the motors and contribute to gradient-climbing. If so, the cell should climb gradients at speeds approaching the theoretical bound (Fig 1C). Alternatively, cells could acquire abundant information about the signal but either lose much of it during transfer to behavior or fail to act on it appropriately. For these cells, performance would be further away from the bound. To quantify where *E. coli* lie on this spectrum, we define information efficiency,  $\eta$ , as the ratio of the cell's chemotactic performance  $v_d$  to the maximum performance possible with the amount of sensory information it acquires, dictated by Eqn 1.



**Figure 1. Information flow from signal to swimming behavior sets a limit on chemotaxis performance.**

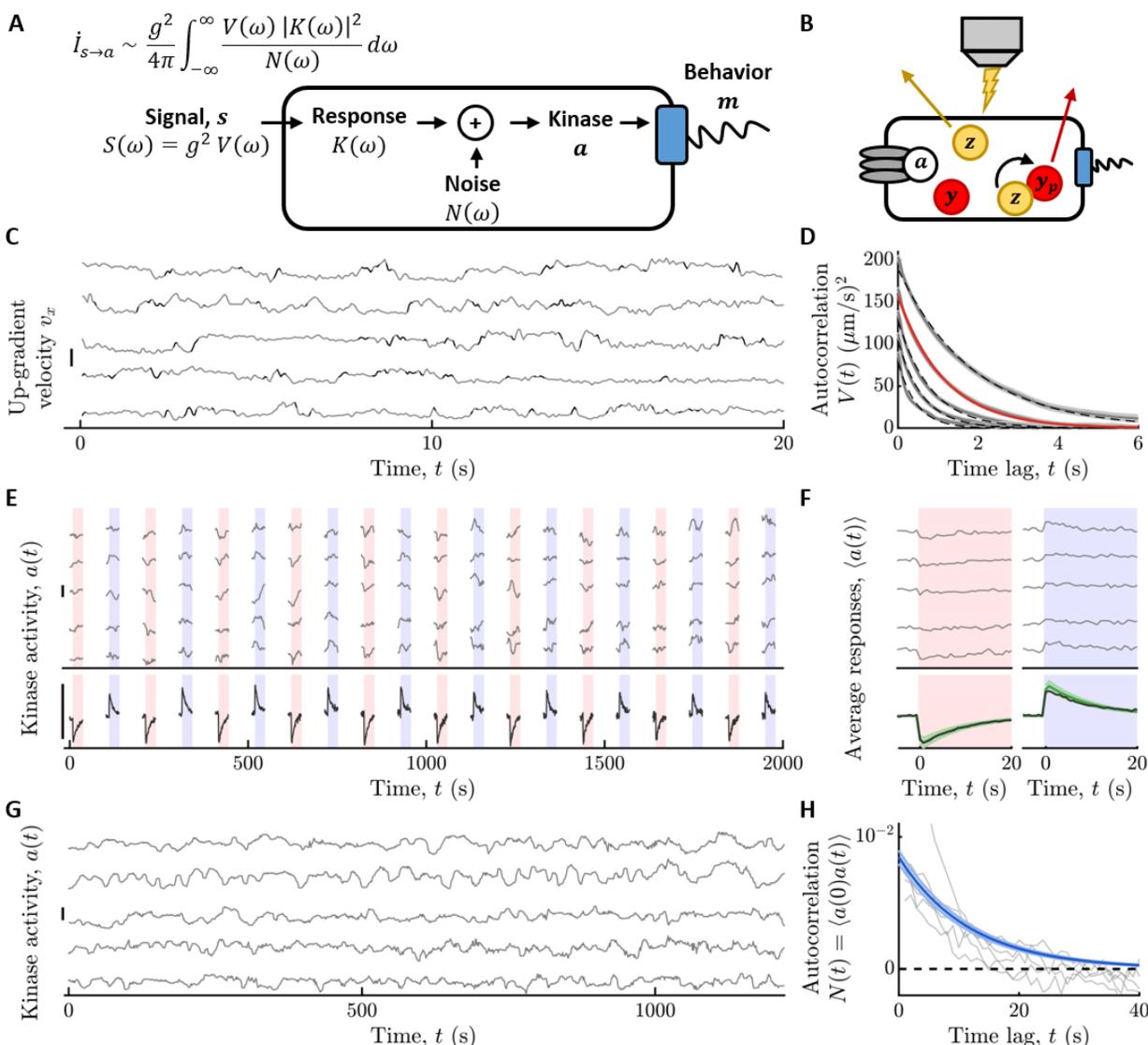
**A)** A cell navigates chemical gradients by sensing relative changes  $s(t) = \frac{d \log(c)}{dt}$  in attractant concentration  $c$  over time. The cell stochastically transitions between behavioral states  $m$ : a run state  $R$ , in which the cell swims with constant speed  $v_0$  and rotational diffusion  $D_r$ ; and a tumble state  $T$ , in which it reorients randomly with directional persistence  $\alpha$  (see Fig. S1). The cell responds to past signals  $\{s\}$  by changing its transition rate  $\lambda_R(\{s\})$  from run to tumble. The average tumble rate  $\lambda_{R0}$  and run rate  $\lambda_T$  define the fraction of time the cell spends in the run state,  $P_{run} = \frac{\lambda_T}{\lambda_{R0} + \lambda_T}$ . The cell's motion creates the signal that it experiences (dashed arrow). The transfer entropy rate from behavior to signal  $\dot{I}_{m \rightarrow s}$  quantifies this information flow, which is non-zero even for a non-chemotactic cell. Signal affects the cell's behavior through  $\lambda_R(\{s\})$  (solid arrow), and the transfer entropy rate from signal to behavior  $\dot{I}_{s \rightarrow m}$  quantifies this information flow needed to perform chemotaxis. **B)** Time courses of behavior  $m(t)$  (top), concentration  $c(t)$  (middle), and signal  $s(t)$  (bottom) for the trajectory shown in (A). In static gradients, signals arise entirely from the cell's motion,  $s(t) = g v_x(t)$ , where  $g = d \log(c) / dx$  is the steepness of the log-concentration gradient and  $v_x$  is the cell's up-gradient velocity. **C)** The information sent from signal to behavior  $\dot{I}_{s \rightarrow m}$  sets an upper limit on chemotaxis performance, defined as the up-gradient drift speed  $v_d$  relative to  $v_0$ . With information rate  $\dot{I}_{s \rightarrow m}$ , the cell's drift speed is bounded by Eqn. 1.

Quantifying *E. coli*'s information efficiency requires measuring the rate at which cells acquire information during chemotaxis. The activity  $a(t)$  of receptor-associated CheA kinases is among the first steps of *E. coli*'s signaling pathway that computes concentration changes<sup>14</sup>. Therefore, we consider the rate of information transfer from signals to kinase activity  $\dot{I}_{s \rightarrow a} \geq \dot{I}_{s \rightarrow m}$  to be the cell's information acquisition rate. To quantify  $\dot{I}_{s \rightarrow a}$  we measured the signal statistics cells experience during navigation, as well as the response and noise properties of the kinases in immobilized cells<sup>20,32</sup> (Fig. 2). This approach eliminates the feedback from kinase activity onto the signals cells experience, allowing us to combine these measurements to estimate the information acquisition rate  $\dot{I}_{s \rightarrow a}$  during navigation (Fig. 2A; SI).

The input signal statistics are characterized by their power spectrum  $S(\omega)$ . It is often difficult to know the natural signal statistics an organism experiences<sup>7,33</sup>. But during bacterial chemotaxis in static gradients, the signal is generated from the cell's own motion in the gradient. Thus, the signal power spectrum is  $S(\omega) = g^2 V(\omega)$ , where  $V(\omega)$  is the power spectrum of the cell's velocity along the gradient direction and  $g = d \log(c) / dx$  is the steepness of the log-concentration gradient. Furthermore, in shallow gradients, the statistics of the cell's motion are nearly identical to those in the absence of a gradient (SI). To quantify the  $x$ -velocity autocorrelation function  $V(t)$ , we tracked individual swimming trajectories in a constant 100  $\mu\text{M}$  background of the attractant  $\alpha$ -methyl-aspartate (MeAsp) (Fig. 2CD; Fig. S2; Methods; SI;  $\sim 10^4$  s of total trajectory time in the bin of median  $P_{run}$ , 7 s average trajectory duration). For different values of  $P_{run}$ , we fit each of the measured  $V(t)$  with decaying exponentials,  $V(t) = a_v e^{-\lambda_{tot} |t|}$ . For the median phenotype, the best-fit parameters were  $a_v = 157.1 \pm 0.5$  ( $\mu\text{m/s})^2$  and  $\lambda_{tot} = 0.862 \pm 0.005$   $\text{s}^{-1}$ .  $V(\omega)$  was then computed by taking the Fourier transform of  $V(t)$  (Fig. S7).

Having quantified the input statistics, we turned to measuring the response and noise properties of CheA kinase activity (SI) using Förster resonance energy transfer (FRET) between the kinase's substrate CheY and the phosphatase CheZ inside single cells<sup>17,18,34</sup> (Fig. 2B). In shallow gradients, cells experience weak signals, and therefore the average kinase response depends linearly on recent signals. In this regime, the response is fully characterized by how it amplifies different frequencies in the signal,  $K(\omega)$ , or equivalently, the response to an impulse (delta function) of signal,  $K(t)$ . To measure the impulse response  $K(t)$  to MeAsp, we used a microfluidic device that allows us to rapidly switch (in  $\sim 100$  ms) the concentration of attractant delivered to hundreds of immobilized cells<sup>34</sup> (Figs. S3). To ensure cells were in the log-sensing regime<sup>12,35</sup>, we first adapted them to a background of 100  $\mu\text{M}$  MeAsp. We then delivered 10 positive and 10 negative 10% step changes of MeAsp concentration (corresponding to delta functions of signal  $s$ ) (Fig. 2E; Methods), small enough that the cells' responses were in the linear regime<sup>36,37</sup> (Fig. S5). Individual cell responses exhibited a stereotypical shape (Fig. 2F) that was well-described by a phenomenological model  $K(t) = G e^{-t/\tau_2} (1 - e^{-t/\tau_1}) H(t)$ , where  $G$  is the gain,  $\tau_1$  is the rise time,  $\tau_2$  is the adaptation time, and  $H(t)$  is the Heaviside step function. We fit this model to each cell's average responses to the positive and negative stimuli simultaneously and then determined the population median values of the parameters ( $\pm$  one standard error;  $n = 442$  cells) (Table S1; Fig. S5):  $G = 1.73 \pm 0.03$ ,  $\tau_1 = 0.22 \pm 0.01$  s, and  $\tau_2 = 9.9 \pm 0.3$  s. The value of  $\tau_1$  that we inferred includes the

CheA kinase response time, the stimulus concentration switching time, and the kinetics of CheY/CheZ binding, making it longer than the actual kinase response time. The kinase response time alone has been measured before to be  $\tau_1 \sim 0.05$  s<sup>38</sup>, but is not directly accessible using this FRET system. After verifying that our results are not sensitive to the value of  $\tau_1$  (SI; Fig. S8), we used the literature value in our estimate of the information rate.



**Figure 2. Measuring the rate of information transfer from signal to intracellular kinase.** **A)** Information rate  $I_{s \rightarrow a}$  from signal  $s$  to kinase activity  $a$  depends on the signal power spectrum  $S(\omega)$ , the kinase frequency response  $K(\omega)$ , and the kinase noise power spectrum  $N(\omega)$ . The signal is  $s(t) = g v_x(t)$ , where  $g$  is the gradient steepness and  $v_x$  is the cell's up-gradient velocity.  $S(\omega) = g^2 V(\omega)$ , where  $V(\omega)$  is the power spectrum of  $v_x$ . **B)** Kinase activity  $a$  was quantified from the FRET between the substrate of the kinase, CheY-mRFP, and its phosphatase, CheZ-mYFP (Methods; SI; Figs. S3-S6). **C)** (Gray) Individual cells'  $v_x(t)$  in a uniform concentration of 100  $\mu$ M MeAsp. (Black: tumbles; scale bar 40  $\mu$ m/s.) **D)** (Gray) Average autocorrelation of  $v_x$ ,  $V(t)$ , for  $P_{run} = 0.93, 0.89, 0.84, 0.79, 0.74$  (top to bottom; throughout, shading is  $\pm$  one standard error; black dashed lines are exponential fits to  $V(t) =$

$a_v \exp(-\lambda_{tot} |t|)$ ). (Red) Best fit to the population median bin,  $P_{run} \sim 0.89$  (Fig. S2; Table S1; SI). From this we computed  $V(\omega)$ . **E**) Immobilized cells were delivered 10  $\mu\text{M}$  steps, up (red shading) and down (blue shading) (background 100  $\mu\text{M}$  MeAsp). (Top, gray) Kinase activity  $a(t)$  for five cells (here and in (G)), smoothed with 10<sup>th</sup> order median filter, and scale bars represent  $\Delta a = 0.3$ ; Methods; SI; Fig. S5). (Bottom, black) Population average  $a(t)$  ( $n = 442$  cells). **F**) Single-cell average (top, gray) and population-average (bottom, black) response functions  $K(t)$  to positive and negative stimuli. (Green)  $K(t) = G \exp(-t/\tau_2)(1 - \exp(-t/\tau_1)) H(t)$ , where  $H(t)$  is the Heaviside step function and  $G$ ,  $\tau_1$ , and  $\tau_2$  are the population median parameters extracted from fits to single-cell responses. From this we computed  $K(\omega)$ . **G,H**) (Gray) Kinase activity (G) and corresponding autocorrelations (H) in single cells adapted to a constant background of 100  $\mu\text{M}$  MeAsp (Methods; SI; Fig. S6). (Blue, H)  $N(t) = \langle a(t)a(0) \rangle = \sigma_n^2 \exp(-|t|/\tau_n)$ , where  $\sigma_n$  and  $\tau_n$  are the population median parameters extracted from fits to single-cell traces ( $n = 262$  cells). From this we computed  $N(\omega)$ .  $V(\omega)$ ,  $K(\omega)$ , and  $N(\omega)$  shown in (Fig. S7).

We quantified the statistics of noise in kinase activity by measuring FRET in single cells in a constant background of 100  $\mu\text{M}$  MeAsp (Fig. 2G; Fig. S6). These fluctuations were well-approximated by an Ornstein-Uhlenbeck process, consistent with previous measurements<sup>16,18</sup>. Using Bayesian filtering (SI), we inferred the single-cell parameters of the noise model directly from the time series. These parameters determined the noise autocorrelation function  $N(t) = \sigma_n^2 e^{-|t|/\tau_n}$  (Fig. 2H; SI), from which we computed the power spectrum  $N(\omega)$  (Fig. S7). The median parameter values in the population ( $\pm$  one standard error;  $n = 262$  cells) were  $\sigma_n = 0.092 \pm 0.002$  AU, the long-time standard deviation of the noise, and  $\tau_n = 11.75 \pm 0.04$  s, the correlation time (Table S1). Note that these measurements include the effects of all noise sources upstream of the kinases, including Poisson single-molecule arrivals at the cells' transmembrane receptors<sup>15</sup>.

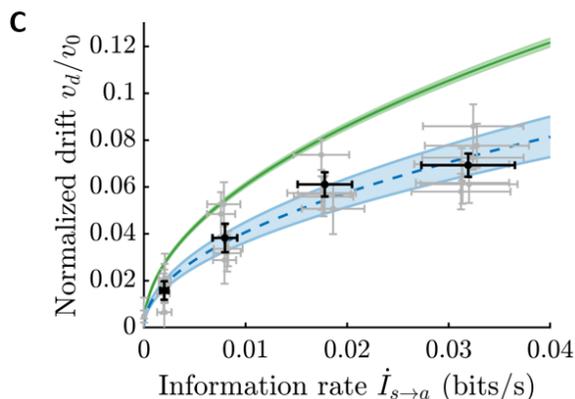
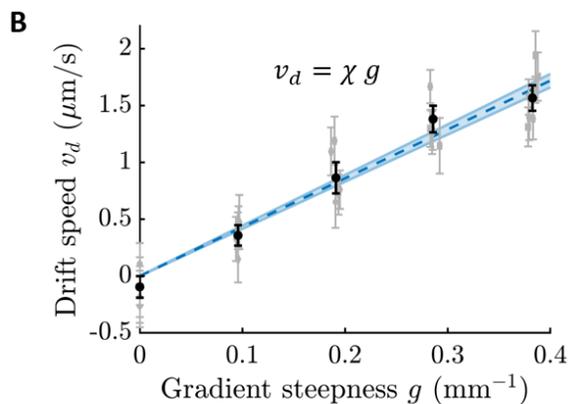
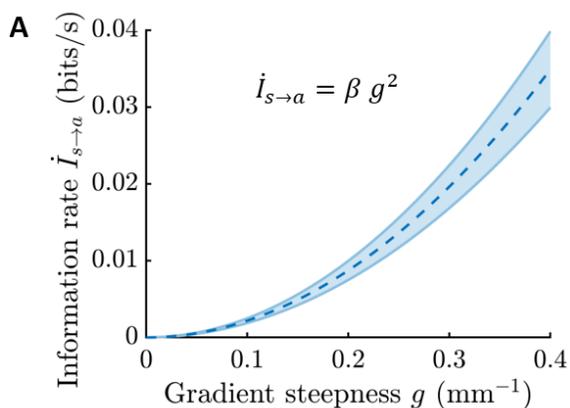
With the input statistics, response function, and noise, we then computed the information rate from the signal to kinase activity  $\dot{I}_{s \rightarrow a}$  (Fig. 3A). Since the signal power is proportional to  $g^2$ , information rate is, as well:  $\dot{I}_{s \rightarrow a} = \beta g^2$ . Using our measurements above, we estimate that the *E. coli* chemotaxis system transfers information to the kinases at a rate  $\beta = 0.22 \pm 0.03$  bits/s per  $\text{mm}^{-2}$  of squared gradient steepness (SI). Thus, in shallow gradients, where concentration varies on millimeter to centimeter length scales, cells only get on the order of  $10^{-2}$  bits/s. The bound in Eqn. 1 predicts that this should be sufficient for a run-and-tumble navigator to climb gradients at a few percent of its swimming speed. However, it is unclear how much of this information is communicated to the motors and used to navigate.

To determine how efficiently *E. coli* use this information to navigate, we measured their drift speeds by tracking individual cells' motion in gradients of varying steepness. Static, linear MeAsp gradients were constructed (Methods) in a background concentration of 100  $\mu\text{M}$ , with length scales ranging from 10 mm ( $g = 0.1 \text{ mm}^{-1}$ ) to 2.5 mm ( $g = 0.4 \text{ mm}^{-1}$ ). From  $>10^5$  seconds of trajectories in each gradient condition, we estimated the average drift speed  $v_d$  as the time-averaged up-gradient velocity over all cells in each experiment (SI). As expected from theory in shallow gradients, the drift speeds increased linearly with gradient steepness  $v_d = \chi g$ , with a

proportionality constant of  $\chi \sim 4.30 \pm 0.15 \mu\text{m/s}$  per  $\text{mm}^{-1}$  of gradient steepness (Fig. 3B; Fig. S9), consistent with previous measurements<sup>39</sup>.

With measurements of both the information acquisition rate and the performance, we are now in a position to quantify *E. coli*'s information efficiency,  $\eta$ . For each gradient  $g$ , we plotted the drift speed  $v_d(g)$  against the information rate  $\dot{I}_{s \rightarrow a}(g)$  (blue curve in Fig. 3C). On the same plot, we show the maximum drift speeds, given by the bound in Eqn. 1 (green curve in Fig. 3C). The ratio of these two curves is the information efficiency,  $\eta = v_d(g) / \left[ F(\theta) \left( \frac{\log(2) \dot{I}_{s \rightarrow a}(g)}{12 D_r} \right)^{\frac{1}{2}} \right]$ . We find that

*E. coli* achieve an efficiency of  $\eta = 0.66 \pm 0.05$ —that is, they climb gradients at  $\sim 66\%$  of the maximum possible speed given the rate at which their kinases acquire information about the



gradients, despite having to infer concentration changes over a finite time. The observation that *E. coli*'s efficiency is order 1 implies that much of the information about the signal contained in kinase activity is preserved in behavior and used to navigate. Thus, *E. coli* cells efficiently use information from environmental signals to perform chemotaxis, indicating that information is likely difficult for their receptor-associated kinases to acquire, limiting their performance.

**Figure 3. *E. coli* use information efficiently to navigate.**

**A)** The rate of information transfer from signal to kinase activity depends on gradient steepness  $g$ :  $\dot{I}_{s \rightarrow a} = \beta g^2$  ( $\beta = 0.22 \pm 0.03$  bits/s /  $\text{mm}^{-2}$ ; Fig. 2; SI). Throughout, shading and error bars indicate  $\pm$  one standard error. **B)** Chemotactic performance as a function of gradient steepness  $g$ , in a background of  $100 \mu\text{M}$  MeAsp. Gray dots: average drift speeds in individual experiments. Black dots: averages over experiments. Error bars on  $g$  are smaller than the markers. Population average drift speed increases linearly with gradient steepness,  $v_d = \chi g$  ( $\chi = 4300 \pm 150 \mu\text{m}^2/\text{s}$ ; blue dashed line and shading; SI). **C)** From measurements of *E. coli* cells' information rates and chemotactic drift speeds, we compared their performance to the theoretical bound (Eqn. 1). Green: predicted maximum performance given information acquisition rate  $\dot{I}_{s \rightarrow a}$  (Eqn. 1). Blue: measured performance  $v_d/v_0$  ( $v_0 = 22.61 \pm 0.07 \mu\text{m/s}$ ) versus information rate  $\dot{I}_{s \rightarrow a}$ , obtained by eliminating  $g$  from the fits of  $v_d(g) = \chi g$  and  $\dot{I}_{s \rightarrow a}(g) = \beta g^2$  to plot

$v_d/v_0 = \chi/v_0 (\dot{I}_{s \rightarrow a}/\beta)^{1/2}$ . Black and gray dots are data points from (B). Taking the ratio of the blue and green curves, we find that *E. coli* achieve drift speeds within  $66 \pm 5\%$  of the theoretical limit.

Many studies of information theory in biology have focused on the maximum amount of information signaling pathways can <sup>1-6,40-42</sup>. Here, we instead ask how much information an organism needs, and how much it uses, while carrying out a task, by studying information transfer in its functional context. Achieving high efficiency requires the signaling cascade to acquire, transmit, and act on information that is relevant to their behavioral task <sup>10</sup>, but in many cases, which bits are relevant is not clear *a priori*. By using rate-distortion theory <sup>6,24,25</sup>, we find that the bits that are relevant to bacterial chemotaxis are those that indicate how fast the attractant concentration is currently changing (see SI). This is demonstrated by our finding that maximum efficiency is achieved when the behavioral response  $\lambda_R(\{s\})$  is instantaneous. Furthermore, previous works measured one-shot information transfer, in bits, by biochemical networks <sup>1-6,40-42</sup>. Here, we measured the rate at which *E. coli* transfer information, in bits/s, between time-varying inputs and outputs, with natural input statistics driven by the cells' own motion. Combining this with measurements of *E. coli*'s chemotactic performance and comparing to the theoretical limit (Eqn. 1), we found that most of the information about concentration changes that flows through CheA is both relevant to and used for navigation.

The information-performance bound in Eqn. 1 depends on the cell's swimming behavior and the physical properties of the environment in which it swims. We quantified this bound for a typical wild type cell swimming in liquid, but individual cells have different behavioral parameters  $\theta$  <sup>16,28,30,29</sup>, and the distribution of swimming phenotypes in the population depends on growth conditions. This raises the question of what is the maximum performance that can be achieved with a given information rate by any phenotype. Maximizing  $F(\theta)$  in Eqn. 1 with respect to  $\theta$  (SI), we find that the optimal agent changes direction by tumbling at the same rate as rotational diffusion <sup>21,43,44</sup> (Fig. S10). The median phenotype in our conditions tumbles more frequently than this. However, in other environments, such as semisolid agar, in which reorientations are required to escape traps <sup>45</sup>, more frequent tumbling may be optimal.

*E. coli*'s proximity to the theoretical limit raises the question of how information is transmitted so reliably from continuous kinase activity into discrete transitions at the flagellar motors. Additionally, while stochastic arrival of ligand molecules at the cell's surface <sup>15,46</sup> reduce how efficiently *E. coli* can use information by forcing them to integrate and respond to signals over a finite time, they also place an upper limit on the amount of relevant information a cell can possibly acquire. Future work will clarify what this fundamental bound on sensing is and how close *E. coli* are to it.

Information transfer is not necessarily the end-goal of biological tasks, but it is needed to perform many of them. Our results suggest organisms may be under selective pressure to efficiently use information from environmental cues to perform tasks necessary for their survival.

## Methods

### Strains and plasmids

The strain used for the FRET experiments is a derivative of *E. coli* K-12 strain RP437 (HCB33), a gift of T. Shimizu, and described in detail elsewhere<sup>18,34</sup>. In brief, the FRET acceptor-donor pair (CheY-mRFP and CheZ-mYFP) is expressed in tandem from plasmid pSJAB106<sup>18</sup> under an isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG)-inducible promoter. The glass-adhesive mutant of FliC (FliC\*) was expressed from a sodium salicylate (NaSal)-inducible pZR1 plasmid<sup>18</sup>. The plasmids are transformed in VS115, a *cheY cheZ fliC* mutant of RP437<sup>18</sup> (gift of V. Sourjik). The crosstalk coefficient for spectral bleedthrough was measured using a strain expressing CheZ-YFP from a plasmid, and that for cross-excitation was measured using a strain expressing CheY-mRFP from a plasmid, both of which are gifts from T. Shimizu. RP437, the direct parent of the FRET strain and also a gift from T. Shimizu, was used to measure behavioral parameters and chemotactic drift speeds. A mutant that can't tumble due to an in-frame deletion of the *cheY* gene, VS100 (gift of V. Sourjik), was used to measure the rotational diffusion coefficient  $D_r$ . All strains are available from the authors upon request.

### Cell preparation

Single-cell FRET microscopy and cell culture was carried out essentially as described previously<sup>18,34</sup>. In brief, cells were picked from a frozen stock at -80°C and inoculated in 2 mL of Tryptone Broth (TB; 1% bacto tryptone, 0.5 % NaCl) and grown overnight to saturation at 30°C and shaken at 250 RPM. Cells from a saturated overnight culture were diluted 100X in 10 mL TB and grown to OD600 0.45-0.47 in the presence of 100  $\mu$ g/ml ampicillin, 34  $\mu$ g/ml chloramphenicol, 50  $\mu$ M IPTG and 3  $\mu$ M NaSal, at 33.5°C and 250 RPM shaking. Cells were collected by centrifugation (5 min at 5000 rpm, or 4080 RCF) and washed twice with motility buffer (10 mM KPO<sub>4</sub>, 0.1 mM EDTA, 1  $\mu$ M methionine, 10 mM lactic acid, pH 7), and then were resuspended in 2 mL motility buffer. Cells were left at 22°C for 90 minutes before loading into the microfluidic device. All experiments, FRET and swimming, were performed at 22-23°C.

For swimming and chemotaxis experiments, cells were prepared identically. Saturated overnight cultures were diluted 100X in 5 mL of TB. After growing to OD600 0.45-0.47, 1 mL of cell suspension was washed twice in motility buffer with 0.05% w/v of polyvinylpyrrolidone (MW 40 kDa) (PVP-40) added. Washes were done by centrifuging the suspension in an Eppendorf tube at 1700 RCF (4000 RPM in this centrifuge) for 3 minutes. After the last wash, cells were resuspended with varying concentrations of MeAsp (see below).

### Microfluidic device fabrication and loading for FRET measurements

Microfluidic devices for the FRET experiments<sup>34</sup> were constructed from polydimethylsiloxane (PDMS) on 24 x 60 mm cover glasses (#1.5) following standard soft lithography protocols<sup>47</sup>. Briefly, the master molds for the device were created with a negative SU-8 photoresist on 100-mm silicon wafers. Approximately 16- $\mu$ m-high master molds were created. To fabricate the device, the master molds were coated with a 5-mm-thick layer of degassed 10:1 PDMS:curing agent mixture (Sylgard 184, Dow Corning). The PDMS layer was cured at 80 °C for 1 hour, and then cut and

separated from the wafer, and holes were punched for the inlets and outlet. The punched PDMS layer was further cured at 80 °C for > 2 hours. Then, the PDMS was cleaned with transparent adhesive tape (Magic Tape; Scotch) followed by rinsing with (in order) isopropanol, methanol, and Millipore-filtered water. The glass was rinsed with (in order) acetone, isopropanol, methanol, and Millipore-filtered water. The PDMS device was tape-cleaned an additional time before the surfaces of the device and coverslip were treated in a plasma bonding oven (Harrick Plasma). After 1 min of exposure to plasma under vacuum, the device was laminated to the coverslip and then baked at 80°C hotplate for > 30 min to establish a covalent bond.

Sample preparation in the microfluidic device was conducted as follows. Five inlets of the device (Fig. S3) were connected to reservoirs (Liquid chromatography columns, C3669; Sigma Aldrich) filled with motility buffer containing various concentrations of  $\alpha$ -methyl-aspartate (MeAsp) through polyethylene tubing (Polythene Tubing, 0.58 mm id, 0.96 mm od; BD Intermedic). The tubing was connected to the PMDS device through stainless steel pins that were directly plugged into the inlets or outlet of the device (New England Tubing). Cells washed and suspended in motility buffer were loaded into the device from the outlet and allowed to attached to the cover glass surface via their sticky flagella by reducing the flow speed inside the chamber. The pressure applied to the inlet solution reservoirs was controlled by computer-controlled solenoid valves (MH1; Festo), which rapidly switched between atmospheric pressure and higher pressure (1.0 kPa) using a source of pressurized air. Only one experiment was conducted per device.

#### Single-cell FRET imaging system

FRET imaging in the microfluidic device was performed using an inverted microscope (Eclipse Ti-E; Nikon) equipped with an oil-immersion objective lens (CFI Apo TIRF 60X Oil; Nikon). YFP was illuminated by an LED illumination system (SOLA SE, Lumencor) through an excitation bandpass filter (FF01-500/24-25; Semrock) and a dichroic mirror (F01-542/27-25F; Semrock). The fluorescence emission was led into an emission image splitter (OptoSplit II; Cairn) and further split into donor and acceptor channels by a second dichroic mirror (FF580-FDi01-25x36; Semrock). The emission was then collected through emission bandpass filters (FF520-Di02-25x36 and FF593-Di03-25x36; Semrock) by a sCMOS camera (ORCA-Flash4.0 V2; Hamamatsu). RFP was illuminated in the same way as YFP except that an excitation bandpass filter (FF01-575/05-25; Semrock) and a dichroic mirror (FF593-Di03-25x36; Semrock) were used. An additional excitation filter (59026x; Chroma) was used in front of the excitation filters. To synchronize image acquisition and the delivery of stimulus solutions, a custom-made MATLAB program controlled both the imaging system (through the API provided by Micro-Manager<sup>48</sup>) and the states of the solenoid valves.

#### Procedure for measuring the linear response functions

All experiments were performed in a background MeAsp concentration of  $c_0 = 100 \mu\text{M}$ . Measurements were made in single cells. First, the FRET level at minimum kinase activity was measured by delivering a saturating stimulus (1 mM MeAsp plus 100  $\mu\text{M}$  serine<sup>49</sup>) for 10 seconds. Immediately afterwards, the FRET level at maximum kinase activity was measured by delivering motility buffer with no attractant (0  $\mu\text{M}$  MeAsp, 0  $\mu\text{M}$  serine) for 5 seconds. When cells are

adapted to 100  $\mu\text{M}$  MeAsp, removing all attractant is sufficient to elicit a maximal response<sup>18,36</sup>. Donor excitation interval (i.e., measurements of  $I_{DD}$  and  $I_{DA}$ ; see SI) was 0.5 seconds and acceptor excitations (i.e., measurements of  $I_{AA}$ ; See SI) were done before and after the set of donor excitations. After this, the concentration of MeAsp was returned to the background  $c_0$ , and no serine was delivered to the cells for the rest of the experiment. Imaging was then stopped and cells were allowed to adapt to the background for 120 seconds.

After this, a series of stimuli were delivered to the cells in the microfluidic device (see Figure 2E for stimulus protocol). Importantly, the cells were only illuminated and imaged for part of the experiment in order to limit photobleaching. First, cells were imaged for 7.5 seconds in the background concentration  $c_0$ . Then, the concentration of MeAsp was shifted up to  $c_+ = 110 \mu\text{M}$  for 30 seconds and imaging continued. Donor excitation interval was 0.75 seconds and acceptor excitations were done before and after the set of donor excitations. After this time, imaging was stopped and the MeAsp concentration returned to  $c_0$  for >60 seconds to allow cells to adapt. Then, the same process was repeated, but this time shifting MeAsp concentration down to  $c_- = 90 \mu\text{M}$ . Alternating up and down stimuli were repeated 10 times each.

FRET levels at minimum and maximum kinase activity were measured again at the end of the experiment. The whole imaging protocol lasted <2200 seconds. In total, cells spent <60 minutes in the device, from loading to the end of imaging. Analyses of these data are described in the SI.

#### Procedure for measuring the noise statistics

Spontaneous fluctuations in kinase activity were also measured in a background MeAsp concentration of  $c_0 = 100 \mu\text{M}$ . Measurements were made in single cells. FRET levels at minimum and maximum kinase activity were measured at the beginning and the end of each experiment, as described above. As above, after these measurements, imaging was then stopped and cells were allowed to adapt to the background for 120 seconds. After this, cells were imaged for about 1200 seconds. Throughout, donor excitations (i.e., measurements of  $I_{DD}$  and  $I_{DA}$ ; see SI) were done every 1.0 second, except when it was interrupted by acceptor excitations (i.e., measurements of  $I_{AA}$ ; see SI), which were conducted every 100 donor excitations. The whole imaging protocol lasted <1400 seconds. In total, cells spent about < 60 minutes in the device, from loading to the end of imaging. Analyses of these data are described in the SI.

#### Procedure to measure swimming and behavioral parameters

After the second wash, cells were centrifuged again and resuspended in motility buffer containing 100  $\mu\text{M}$  MeAsp. Then, the cell suspension was diluted to an OD600 of 0.00025. The cell suspension was then loaded into  $\mu$ -Slide Chemotaxis devices (ibidi; Martinsried, Germany), the same type of device used to create static gradients, described below. However, instead of tracking cells in the gradient region, we tracked their swimming in one of the large reservoirs, which are roughly 750  $\mu\text{m}$  deep. 1000-s movies of swimming cells were recorded on a Nikon Ti-E Inverted Microscope using a CFI Plan Fluor 4X objective (NA 0.13). This objective's depth of field is about  $\pm 18 \mu\text{m}$ , much shorter than the depth of the chamber. Adjusting the focal plane to the middle of the chamber made cells that were swimming near the ceiling or floor of the device, which could

experience hydrodynamic interactions that affect their behavior<sup>50</sup>, not visible in the movie. At the same time, this lower magnification objective allowed us to collect relatively longer swimming trajectories. Movies were captured around 30 minutes after loading cells into the chamber to mimic the gradient experiments below. Images here and below were captured using a sCMOS camera (ORCA-Flash4.0 V2; Hamamatsu). Analyses of these data are described in the SI. Five biological replicates were done for behavioral parameter measurements, and four biological replicates were done for measuring  $D_r$ .

### Procedure to measure chemotactic drift speeds

Chemotaxis experiments were performed in  $\mu$ -Slide Chemotaxis devices (ibidi; Martinsried, Germany). These devices generate a linear gradient between two concentration reservoirs that is stable for a long time. After the second wash, the cell suspension was split into two Eppendorf tubes, 0.5 mL each. After one more centrifugation, one tube of cells was resuspended in 1 mL of motility buffer with 100  $\mu$ M of MeAsp, to be injected into the “low-concentration reservoir”, and the other was resuspended in 1 mL of motility buffer with 2  $\mu$ M of fluorescein and varying concentrations of attractant, to be injected into the “high-concentration reservoir”. Cells in both tubes were diluted to OD 0.001 for each experiment. Loading cells in both reservoirs ensured that the concentration of cells throughout the experimental device was approximately uniform. This limited the effects of potential biases that could arise from observing a finite field of view.

Using a background concentration of at least 100  $\mu$ M MeAsp ensured that the cells were in the log-sensing regime<sup>12</sup>. The “high” concentrations of MeAsp used were 110.5  $\mu$ M, 122.1  $\mu$ M, 135.0  $\mu$ M, and 149.2  $\mu$ M. With 1 mm separating the two reservoirs, these concentrations produced linear gradients that approximated shallow exponential gradients with steepness of roughly:  $g = \{0.1, 0.2, 0.3, 0.4\} \text{ mm}^{-1}$ .  $g$  was calculated from  $g = \frac{\Delta \log c}{\Delta x}$ , where  $\Delta \log c$  is the difference in log concentrations between the two reservoirs, and  $\Delta x$  is the distance between them. This is exactly the average steepness of log-concentration across with the width of the channel. To see this, the steady state concentration profile is linear,  $c(x) = \frac{\Delta c}{\Delta x}(x - x_0) + c_0$ , where  $\Delta c$  is the difference in concentration between the two reservoirs,  $x_0$  is the midpoint between them, and  $c_0$  is the concentration at  $x = x_0$ . From this, the gradient of log concentration depends on position  $x$  and can be computed from  $g(x) = \frac{d \log(c(x))}{dx} = \frac{1}{\frac{1}{c_0 \Delta x} x + 1} \frac{1}{c_0} \frac{\Delta c}{\Delta x}$ , where we have defined a reference frame where  $x_0 = 0$ . Averaging over the channel by integrating over  $x$  from  $-\Delta x/2$  to  $\Delta x/2$  and dividing by  $\Delta x$  gives,  $\langle g \rangle = \frac{\log(1 + \frac{1}{2} \frac{\Delta c}{c_0}) - \log(1 - \frac{1}{2} \frac{\Delta c}{c_0})}{\Delta x} = \frac{\log(c_0 + \frac{1}{2} \Delta c) - \log(c_0 - \frac{1}{2} \Delta c)}{\Delta x} = \frac{\Delta \log c}{\Delta x} = g$ . Close to the low-concentration reservoir,  $g(x)$  is larger than  $g$ , and vice versa near the high-concentration reservoir, but these errors are small and approximately cancel each other out when we average drift speeds of cells across the channel.

To load the device, first the reservoirs were sealed with the manufacturer’s tabs. Cell suspension with 100  $\mu$ M MeAsp was injected into the channel where the gradient would form. Excess liquid in the inlets was removed. Then one tab from each reservoir was removed, and the gradient channel

was sealed with tabs. The left reservoir was then fully unsealed, and the right reservoir was sealed with tabs. 60-65  $\mu\text{L}$  of cell suspension with 100  $\mu\text{M}$  MeAsp was injected into the left reservoir, and then both inlets of that reservoir were sealed with tape or tabs. Care was taken to make sure there were no bubbles in reservoir at the inlets. Then, the right reservoir was unsealed, and 60-65  $\mu\text{L}$  of cell suspension with the higher concentration of MeAsp was injected. A timer was then immediately started. The right reservoir was then sealed.

Cells were imaged by phase contrast with a CFI Plan Fluor 10X objective (NA 0.30). The depth of the gradient region of the device is 70  $\mu\text{m}$ , and the depth of field of the objective is about  $\pm 4$   $\mu\text{m}$ . Focusing on the middle of the chamber with this objective filtered out cells that could be interacting with the ceiling or floor surfaces. Images of fluorescein were taken every 5 minutes using a CFI Plan Fluor 4X objective (NA 0.13) through a YFP filter cube (Chroma 49003), illuminated by a LED (SOLA SE, Lumencor) with an exposure time of 100 ms. Since the diffusivity of fluorescein is similar to (slightly lower than) that of MeAsp (MW of fluorescein is 376 kDa; MW of MeAsp is 147 kDa; Sigma Aldrich), we used fluorescein as an indication of when the attractant gradient was stable and linear in the observation region between the two reservoirs, as has been done before<sup>51,52</sup>. Once the fluorescein profile was stable for several time points (typically around 50-60 minutes after loading), a 1000-second phase contrast movie was recorded at 20 FPS using the 10X phase contrast objective. Before the recording, the transmitted light illumination was adjusted to minimize the number of saturated pixels. After the recording, an additional image of the fluorescein profile was recorded, and the cells were observed to check that they were still swimming normally. Analyses of these data are described below. At least five biological replicates were performed for each gradient steepness.

**Acknowledgements:** We thank Jeremy Moore and Xiaowei Zhang for help setting up the experimental assays. We also thank Katja Taute and Marianne Grognot for helpful discussions about the gradient experiments. We thank Pieter Rein ten Wolde for providing detailed feedback on an earlier version of this manuscript, as well as Artur Wachtel, Isabella Graf, and Damon Clark for providing comments. We acknowledge Tom Shimizu and Victor Sourjik for bacteria strains and Rafael Gomez-Sjoberg, Microfluidics Lab, for providing information and software to control the solenoid valves in the microfluidic setup.

**Funding:** HM, KK, and TE were funded by NIH R01s GM106189 and GM138533. HM was funded by NIH F32 GM131583. TE and BM were funded by a Yale PEB Seed Grant. BM was funded by Simons Investigator Award 624156 and NIH R35 GM138341.

**Contributions:** HM and KK contributed equally to this work. HM, KK, BM and TE designed the research. HM and BM derived the theoretical bound with inputs from TE and KK. HM performed the experiments tracking bacteria. KK performed the single cell FRET experiments. HM, KK, and TE validated the data. HM, KK, BM, and TE discussed the data analysis. HM and KK performed the data analysis. HH, KK, BM and TE wrote the initial draft and all revisions.

**Competing interests:** Authors declare no competing interests.

## References

1. Cheong, R., Rhee, A., Wang, C. J., Nemenman, I. & Levchenko, A. Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science* **334**, 354–358 (2011).
2. Uda, S. *et al.* Robustness and Compensation of Information Transmission of Signaling Pathways. *Science* **341**, 558–561 (2013).
3. Selimkhanov, J. *et al.* Accurate information transmission through dynamic biochemical signaling networks. *Science* **346**, 1370–1373 (2014).
4. Granados, A. A. *et al.* Distributed and dynamic intracellular organization of extracellular information. *PNAS* **115**, 6088–6093 (2018).
5. Bowsher, C. G. & Swain, P. S. Environmental sensing, information transfer, and cellular decision-making. *Current Opinion in Biotechnology* **28**, 149–155 (2014).
6. Tkačik, G. & Bialek, W. Information Processing in Living Systems. *Annual Review of Condensed Matter Physics* **7**, 89–117 (2016).
7. Laughlin, S. A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch., C, Biosci.* **36**, 910–912 (1981).
8. Rieke, F., Warland, D. & Bialek, W. Coding Efficiency and Information Rates in Sensory Neurons. *EPL* **22**, 151–156 (1993).
9. Petkova, M. D., Tkačik, G., Bialek, W., Wieschaus, E. F. & Gregor, T. Optimal Decoding of Cellular Identities in a Genetic Network. *Cell* **176**, 844–855.e15 (2019).
10. Bialek, W., De Ruyter Van Steveninck, R. R. & Tishby, N. Efficient representation as a design principle for neural coding and computation. in *2006 IEEE International Symposium on Information Theory* 659–663 (2006). doi:10.1109/ISIT.2006.261867.
11. Berg, H. C. *E. coli in motion*. (Springer, 2004).
12. Lazova, M. D., Ahmed, T., Bellomo, D., Stocker, R. & Shimizu, T. S. Response rescaling in bacterial chemotaxis. *PNAS* **108**, 13870–13875 (2011).
13. Tu, Y. Quantitative Modeling of Bacterial Chemotaxis: Signal Amplification and Accurate Adaptation. *Annual Review of Biophysics* **42**, 337–359 (2013).
14. Parkinson, J. S., Hazelbauer, G. L. & Falke, J. J. Signaling and sensory adaptation in Escherichia coli chemoreceptors: 2015 update. *Trends in Microbiology* **23**, 257–266 (2015).
15. Berg, H. C. & Purcell, E. M. Physics of chemoreception. *Biophysical Journal* **20**, 193–219 (1977).
16. Korobkova, E., Emonet, T., Vilar, J. M. G., Shimizu, T. S. & Cluzel, P. From molecular noise to behavioural variability in a single bacterium. *Nature* **428**, 574–578 (2004).
17. Colin, R., Rosazza, C., Vaknin, A. & Sourjik, V. Multiple sources of slow activity fluctuations in a bacterial chemosensory network. *eLife* **6**, e26796 (2017).

18. Keegstra, J. M. *et al.* Phenotypic diversity and temporal variability in a bacterial signaling network revealed by single-cell FRET. *eLife* **6**, e27455 (2017).
19. Shannon, C. E. A Mathematical Theory of Communication. *Bell System Technical Journal* **27**, 379–423 (1948).
20. Tostevin, F. & ten Wolde, P. R. Mutual Information between Input and Output Trajectories of Biochemical Networks. *Phys. Rev. Lett.* **102**, 218101 (2009).
21. Long, J., Zucker, S. W. & Emonet, T. Feedback between motion and sensation provides nonlinear boost in run-and-tumble navigation. *PLOS Computational Biology* **13**, e1005429 (2017).
22. Hartich, D., Barato, A. C. & Seifert, U. Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a Maxwell’s demon interpretation. *J. Stat. Mech.* **2014**, P02016 (2014).
23. Schreiber, T. Measuring Information Transfer. *Phys. Rev. Lett.* **85**, 461–464 (2000).
24. Shannon, C. E. Coding Theorems for a Discrete Source With a Fidelity Criterion. *IRE Nat. Conv. Rec.* 142–163 (1959).
25. Andrews, B. W. & Iglesias, P. A. An Information-Theoretic Characterization of the Optimal Gradient Sensing Response of Cells. *PLOS Computational Biology* **3**, e153 (2007).
26. Becker, N. B., Mugler, A. & ten Wolde, P. R. Optimal Prediction by Cellular Signaling Networks. *Phys. Rev. Lett.* **115**, 258103 (2015).
27. Clausznitzer, D., Micali, G., Neumann, S., Sourjik, V. & Endres, R. G. Predicting Chemical Environments of Bacteria from Receptor Signaling. *PLOS Computational Biology* **10**, e1003870 (2014).
28. Spudich, J. L. & Koshland, D. E. Non-genetic individuality: chance in the single cell. *Nature* **262**, 467–471 (1976).
29. Masson, J.-B., Voisinne, G., Wong-Ng, J., Celani, A. & Vergassola, M. Noninvasive inference of the molecular chemotactic response using bacterial trajectories. *PNAS* **109**, 1802–1807 (2012).
30. Waite, A. J., Frankel, N. W. & Emonet, T. Behavioral Variability and Phenotypic Diversity in Bacterial Chemotaxis. *Annual Review of Biophysics* **47**, 595–616 (2018).
31. Sigtermans, D. Towards a Framework for Observational Causality from Time Series: When Shannon Meets Turing. *Entropy* **22**, 426 (2020).
32. Shannon, C. E. Communication in the Presence of Noise. *Proceedings of the IRE* **37**, 10–21 (1949).
33. Rieke, F., Bodnar, D. A. & Bialek, W. Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **262**, 259–265 (1995).
34. Kamino, K., Keegstra, J. M., Long, J., Emonet, T. & Shimizu, T. S. Adaptive tuning of cell sensory diversity without changes in gene expression. *Science Advances* (2020).
35. Kalinin, Y. V., Jiang, L., Tu, Y. & Wu, M. Logarithmic Sensing in Escherichia coli Bacterial Chemotaxis. *Biophysical Journal* **96**, 2439–2448 (2009).
36. Sourjik, V. & Berg, H. C. Receptor sensitivity in bacterial chemotaxis. *Proceedings of the National Academy of Sciences* **99**, 123–127 (2002).

37. Shimizu, T. S., Tu, Y. & Berg, H. C. A modular gradient-sensing network for chemotaxis in *Escherichia coli* revealed by responses to time-varying stimuli. *Molecular Systems Biology* **6**, 382 (2010).
38. Francis, N. R. *et al.* Subunit Organization in a Soluble Complex of Tar, CheW, and CheA by Electron Microscopy. *J. Biol. Chem.* **277**, 36755–36759 (2002).
39. Colin, R., Zhang, R. & Wilson, L. G. Fast, high-throughput measurement of collective behaviour in a bacterial population. *Journal of The Royal Society Interface* **11**, 20140486 (2014).
40. Potter, G. D., Byrd, T. A., Mugler, A. & Sun, B. Dynamic Sampling and Information Encoding in Biochemical Networks. *Biophysical Journal* **112**, 795–804 (2017).
41. Anders, A., Ghosh, B., Glatter, T. & Sourjik, V. Design of a MAPK signalling cascade balances energetic cost versus accuracy of information transmission. *Nature Communications* **11**, 3494 (2020).
42. Razo-Mejia, M. *et al.* First-principles prediction of the information processing capacity of a simple genetic circuit. *Phys. Rev. E* **102**, 022404 (2020).
43. Berg, H. C. & Brown, D. A. Chemotaxis in *Escherichia coli* analysed by Three-dimensional Tracking. *Nature* **239**, 500–504 (1972).
44. Celani, A. & Vergassola, M. Bacterial strategies for chemotaxis response. *PNAS* **107**, 1391–1396 (2010).
45. Wolfe, A. J. & Berg, H. C. Migration of bacteria in semisolid agar. *PNAS* **86**, 6973–6977 (1989).
46. Mora, T. & Wingreen, N. S. Limits of Sensing Temporal Concentration Changes by Single Cells. *Phys. Rev. Lett.* **104**, 248101 (2010).
47. Qin, D., Xia, Y. & Whitesides, G. M. Soft lithography for micro- and nanoscale patterning. *Nature Protocols* **5**, 491–502 (2010).
48. Edelstein, A., Amodaj, N., Hoover, K., Vale, R. & Stuurman, N. Computer Control of Microscopes Using  $\mu$ Manager. *Current Protocols in Molecular Biology* **92**, 14.20.1-14.20.17 (2010).
49. Sourjik, V. & Berg, H. C. Functional interactions between receptors in bacterial chemotaxis. *Nature* **428**, 437–441 (2004).
50. Frymier, P. D., Ford, R. M., Berg, H. C. & Cummings, P. T. Three-dimensional tracking of motile bacteria near a solid planar surface. *PNAS* **92**, 6195–6199 (1995).
51. Colin, R., Zhang, R. & Wilson, L. G. Fast, high-throughput measurement of collective behaviour in a bacterial population. *Journal of The Royal Society Interface* **11**, 20140486 (2014).
52. Waite, A. J. *et al.* Non-genetic diversity modulates population performance. *Molecular Systems Biology* **12**, 895 (2016).