Reviewers' comments:

Reviewer #1 (Remarks to the Author):

The major claims of the paper are that (1) a rugged landscape (with fluctuations in concentration amidst an increasing trend) is best traversed by bacteria relying on memory that has a time scale comparable to a cell's persistence, (2) a bacterial population that uses a short-term memory exhibits a greater drift velocity in a rugged landscape than one with no memory, and (3) a rugged landscape can segregate a bacterial population with relatively short memories yielding a non-Gaussian distribution.

The claims of the paper are consistent with a fundamental understanding of bacterial chemotaxis. For example, my interpretation of claim 2 is that fluctuations in concentration along a rugged landscape may lead to incorrect assessments of the overall gradient if bacteria rely only on instantaneous information. Thus, bacteria in a rugged landscape will change direction more frequently in response to fluctuations (i.e., errors in assessment) compared to a smooth landscape. This will lead to a decrease in the average run length which reduces the drift velocity of the population. Thus, it makes sense, if there are fluctuations in the signal, for bacteria to be able to integrate over some time period to get a more accurate read of the trend.

Memory has been a topic of great interest within the bacterial chemotaxis community over the past decade or so. What makes this work incredibly useful is ability to quantify the relative importance of memory for a given physical scenario. Because bacterial chemotaxis serves as a model system for understanding the stimulus-response mechanism in higher-order organisms, the outcomes reported in this paper will appeal to a broad audience. Furthermore, the outcomes will be of particular interest to researchers who want to design experiments to determine values for population scale parameters and need to know the scenarios under which the Keller-Segel model (which does not include memory) will be sufficient for analysis. The correction to the drift velocity provides a useful measure of the deviation from the Keller-Segel limit for these purposes.

The approach used by the authors illuminates the connection between the signal transduction mechanism internal to the bacteria and the external stimuli by using an agent-based model. Their multi-scale modeling approach provides useful relationships between parameters that describe single cell behavior and those that predict population-level behavior. The comparison between outcomes from the agent-based model, the Keller-Segel model and the analytical expressions are especially useful in understanding the limits of the classic Keller-Segel equation for bacterial chemotaxis. The consistency between the various approaches in predicting the limiting behavior associated with the Keller-Segel model provides confidence in the reported findings.

One conclusion of the authors is that memory is most important when the signal is noisy and the gradient is steep. However, I expect that under many natural scenarios gradients are shallow and not very noisy. It would be helpful for the authors to include examples of physical scenarios that correspond to a rugged landscape. Are rugged landscapes expected to occur frequently in nature where bacterial chemotaxis is exhibited? Perhaps, from stirring by flagellar rotation? Under turbulent flows? Within porous media?

Some other questions for the authors to consider: What is the basis for determining how rugged a landscape is? What are typical fluctuations that a cell may experience? Will these be the same for swimming in 1D vs 3D? In 1D a cell makes a 180° turn when it tumbles, while in 3D the turn angle of E. coli is near 70°. What are potential issues for generalizing the analysis to 3D-swimming and non-constant gradients?

Reviewer #2 (Remarks to the Author):

This manuscript considers the effect of memory on bacterial chemotactic navigation in noisy, shallow concentration gradients with spatial correlations. Memory is effectively achieved in bacterial chemotaxis by the process of adaptation. Simulating a 1D model, the authors compared the drift speeds of agents with memory to those of agents that respond only the local spatial gradient, a commonly used model that is recovered in various limits of the model with memory. One central result of the paper is that memory enhances the drift speed in noisy spatial gradients when the correlation length of the noise is comparable to or a bit shorter than the cell's run length. In the regime of a shallow gradient and noise that is small compared to the gradient steepness, the authors derived an approximate model of the contribution of memory to the drift speed, and it agrees well with simulations. The optimal memory time for each noise correlation length is derived, and the authors discuss how it arises from a balance between averaging out spatial noise and limiting the number of tumbles that occur during the averaging time. The results are interesting and new, but I have several main concerns.

Main:

The current understanding of why E. coli and other bacteria use some form of a run-and-reorient strategy to navigate is that they are too small to directly measure concentration gradients across their body due to diffusion. Instead, they measure spatial gradients from temporal changes in concentration during runs. This suggests that diffusion of most attractants should be fast enough to rapidly damp out concentration fluctuations in the environment that are smaller than one run length. But this is the regime where the authors' model predicts that memory produces the largest gain in drift speed. This makes me have doubts about the relevance of this mechanism for bacterial chemotaxis. Otherwise, the results are interesting and the work is rigorous. This mechanism could apply to other scenarios in which an agent navigates a noisy environment with gradient-like structure.

The authors derive an optimal memory time in the case that there is no noise in the gradient that comes out to be 1/4 of a run duration. I am struggling to reconcile this short optimal memory time with other work (using different models) that suggest longer memory time increases drift speed, in particular Dufour, Fu et al 2014 Plos Comp Bio. In Fig 2A of that paper, simulations of a model in 3D with some implementation of memory and a noiseless gradient had drift speed increasing with memory time even when the memory time was 10 times a typical run duration. I'm not sure whether this simulation was in the shallow gradient regime the authors study. The authors mention in the discussion that one possible reason for the difference from previous work is persistence of direction through a tumble, which is absent from the model in this manuscript and could make longer memory useful. But, to my knowledge, the theoretical predictions in Fig 2A of the same paper Dufour, Fu et al do not consider directional persistence through tumbles (their simulations do).

Minor:

Most bacteria sense the log of concentration and gradients of the log, for a large concentration range. But this doesn't substantially affect the authors' results.

There is a small error in the legend of Fig. 2b. The average drift speed of AB cells is averaged over the ensemble of simulated cells but not over gradient noise (as currently written) since there isn't any noise in this case.

At the bottom of page 8, first column, "studies" is repeated an extra time.

The last "and" of the Discussion became "ad".

Dear Reviewers,

We are pleased to resubmit the revised version of *"Beyond local gradient-sensing: memory effects in bacterial chemotactic navigation"* by Adam Gosztolai and Mauricio Barahona for publication in *Communications Physics*.

We thank the Reviewers for their insightful comments and the very careful reading of our manuscript.

In light of the Reviewer reports, and the journal requirements for a title without punctuation, we would like to make the slight change in the title to *"Cellular memory enhances bacterial chemotactic navigation in rugged environments"*.

We have responded below to the issues and suggestions raised and we have modified the text accordingly, as shown in the attached .diff file.

We believe that these modifications have further improved the clarity of the paper, and we hope that the manuscript will now be acceptable for publication.

Kind Regards,

Adam Gosztolai and Mauricio Barahona

**Reviewer 1**

The major claims of the paper are that (1) a rugged landscape (with fluctuations in concentration amidst an increasing trend) is best traversed by bacteria relying on memory that has a time scale comparable to a cell's persistence, (2) a bacterial population that uses a short-term memory exhibits a greater drift velocity in a rugged landscape than one with no memory, and (3) a rugged landscape can segregate a bacterial population with relatively short memories yielding a non-Gaussian distribution.

The claims of the paper are consistent with a fundamental understanding of bacterial chemotaxis. For example, my interpretation of claim 2 is that fluctuations in concentration along a rugged landscape may lead to incorrect assessments of the overall gradient if bacteria rely only on instantaneous information. Thus, bacteria in a rugged landscape will change direction more frequently in response to fluctuations (i.e., errors in assessment) compared to a smooth landscape. This will lead to a decrease in the average run length which reduces the drift velocity of the population. Thus, it makes sense, if there are fluctuations in the signal, for bacteria to be able to integrate over some time period to get a more accurate read of the trend.

Memory has been a topic of great interest within the bacterial chemotaxis community over the past decade or so. What makes this work incredibly useful is ability to quantify the relative importance of memory for a given physical scenario. Because bacterial chemotaxis serves as a model system for understanding the stimulus-response mechanism in higher-order organisms, the outcomes reported in this paper will appeal to a broad audience. Furthermore, the outcomes will be of particular interest to researchers who want to design experiments to determine values for population scale parameters and need to know the scenarios under which the Keller-Segel model (which does not include memory) will be sufficient for analysis. The correction to the drift velocity provides a useful measure of the deviation from the Keller-Segel limit for these purposes.

The approach used by the authors illuminates the connection between the signal transduction mechanism internal to the bacteria and the external stimuli by using an agent-based model. Their multi-scale modeling approach provides useful relationships between parameters that describe single cell behavior and those that predict population-level behavior. The comparison between outcomes from the agent-based model, the Keller-Segel model and the analyt-

ical expressions are especially useful in understanding the limits of the classic Keller–Segel equation for bacterial chemotaxis. The consistency between the various approaches in predicting the limiting behavior associated with the Keller–Segel model provides confidence in the reported findings.

We thank the Reviewer for their evaluation of our work, and for their insightful summary of the main points of our manuscript, upon which we have elaborated in the Introduction, Discussion and page 5.

One conclusion of the authors is that memory is most important when the signal is noisy and the gradient is steep. However, I expect that under many natural scenarios gradients are shallow and not very noisy. It would be helpful for the authors to include examples of physical scenarios that correspond to a rugged landscape. Are rugged landscapes expected to occur frequently in nature where bacterial chemotaxis is exhibited? Perhaps, from stirring by flagellar rotation? Under turbulent flows? Within porous media?

This remark is linked to the first Main Remark by Reviewer 2. We provide an answer to this point here as well as in the response to Reviewer 2 with additional details.

Small scale ruggedness in chemoattractant landscapes has different origins. As mentioned by both reviewers, ruggedness appears naturally as a result of random spatial inhomogeneities in the attractant landscape, e.g., when chemotaxis takes place in a porous or particulate medium. Several works have found that native microbial habitats are porous, or a mixture of dilute and degrading particulate matter exhibiting microscale gradient structures at scales smaller or comparable to the mean run-length of the bacteria that inhabit it. In these environments, bacteria experience the spreading pulse of chemoattractant as a ruggedness in dilute, shallow background gradients[1,2].

There is evidence for this effect in native *E. coli* environments: in soil, the length scales of the spatial inhomogeneities are estimated to vary between $\sim 2\,\mu$m–1 mm, according to Ref.[3]; in the digestive tract, the chemical environment is also thought to have fine-grained spatial structure[4]. The mean run length of *E. coli* is $\sim 20\mu$m.

[1] A. M. Hein et al. "Physical limits on bacterial navigation in dynamic environments". *J. R. Soc. Interface* 13.114 (2016), pp. 20150844–8.

[2] N. Blackburn, T. Fenchel, and J. Mitchell. "Microscale Nutrient Patches in Planktonic Habitats Shown by Chemotactic Bacteria". *Science* 282.5397 (1998), pp. 2254–2256.

[3] N. F. Ngom et al. "Extraction of three-dimensional soil pore space from microtomography images using a geometrical approach". *Geoderma* 163.1-2 (2011), pp. 127–134.

[4] F. J. H. Hol et al. "Bacterial predator–prey dynamics in microscale patchy landscapes". *Proc. R. Soc. B* 283.1824 (2016), pp. 20152154–9.

Such spatial inhomogeneities are also important for other organisms that display run and tumble behaviour. Some aquatic microbes also exhibit run-tumble motion albeit with higher run speeds of $\sim 100\,\mu$m/s and mean run lengths of 68–346 $\mu$m[5]. The habitats of such microbes have inhomogeneities arising from degrading particles and secretions of other organisms,[6] with commensurate length scales of $\sim$ 10–1000 $\mu$m. All these findings suggest that microscale structure can exist on scales smaller than a run length, which is where our model predicts enhanced drift velocity.

In addition, time-varying spatial ruggedness can also appear as a result of advection of the medium[5,7]. The stirring of a medium containing chemoattractants causes it to form a network of thin, elongated filaments[6]. Previous work[1] has shown that in aquatic environments the characteristic length scales of these turbulent conditions vary between $200 - 1000\,\mu$m, again commensurate with the run length of aquatic microbes (68–346 $\mu$m).

We have added the relevant references in the text and addressed this important point in the Discussion by describing the different scenarios when the natural landscape appears rugged for a chemotactic cell.

Some other questions for the authors to consider:

- What is the basis for determining how rugged a landscape is?

There are two aspects to ruggedness: amplitude and length scale. In our model, the importance of each of these two factors is measured relative to parameters of the cell as per our non-dimensionalisation.

The amplitude of the ruggedness *as perceived by the cell* is linearly proportional to $\beta\sigma_\eta$ (Eq. (S13) in SI), where $\beta$ is the signal gain and $\sigma_\eta$ is the amplitude of the landscape ruggedness. Our results also show that if the signal-to-noise ratio is large (Eq. (6)), the amplitude of the perceived ruggedness only enters as a factor of $(\beta\sigma_\eta)^2$ in the effect of correlations (Eq. (25)). Hence the effective size of the typical perceived fluctuations is scaled by the signal gain, which is a function of the number of receptors on the cell membrane, their interaction, and the biochemical cellular circuits that transduce chemotactic signals.

Regarding the length scale ($\mu$) of the landscape, one of our main conclusions is that the ruggedness is only important when $\mu \sim v_0\gamma$, where $v_0$ is the ballistic run speed and $\gamma$ is the memory. Therefore, the spatial fluctuations are only important when their length scale is comparable to the length of trajectory that is stored in memory over each run.

- What are typical fluctuations that a cell may experience?

[5] R. Watteaux, R. Stocker, and J. R. Taylor. "Sensitivity of the rate of nutrient uptake by chemotactic bacteria to physical and biological parameters in a turbulent environment". *J. Theor. Biol.* 387 (2015), pp. 120 –135.

[6] R. Stocker. "Marine microbes see a sea of gradients". *Science* 338.6107 (2012), pp. 628–633.

[7] J R Taylor and R Stocker. "Trade-offs of chemotactic foraging in turbulent water". *Science* 338.6107 (2012), pp. 675–679.

As discussed above, the size of the typical perceived fluctuations is $\beta\sigma_\eta$; hence it is modulated by the signal gain $\beta$, a function of biophysical and biochemical factors in the cell. Our work applies in the large signal-to-noise ratio regime (Eq. (16)). We thus consider the case when $\sigma_\eta/\alpha$ is small, i.e., when the amplitude of the spatial fluctuations ($\sigma_\eta$) is small relative to the gradient of chemoattractant ($\alpha$). For instance, in the agent-based simulations in Fig. 3, we have $\sigma_\eta/\alpha = 0.02$. Hence memory effects are relevant even when the fluctuations are small (measured relative to the gradient).

- `Will these be the same for swimming in 1D vs 3D? In 1D a cell makes a 180° turn when it tumbles, while in 3D the turn angle of E. coli is near 70°.`

The dimensionality of the motion and its effect on the directional persistence of swimming cells is an important consideration. In 1D, chemotactic cells turn 180° at each tumble; hence when traversing a static landscape, a cell experiences the 'anti-phase' signal over parts of successive 'up' and 'down' runs. This phenomenon cancels out part of the filtered signal and decreases the feedback from the environment. In 3D, cells turn by a smaller angle of 70°, thus allowing for higher positive feedback and a stronger nonlinear modulation of the tumbling rates, a phenomenon already present in the deterministic case with no ruggedness, as studied in Ref.[8]. It is expected that this increased directional persistence in 3D would allow the cells to integrate information for a longer time while maintaining the same heading, thus getting a better estimate of the underlying gradient. On the other hand, in dimensions above 1D, the rotation of the cell due to Brownian fluctuations needs to be taken into account. This means that despite the smaller turn angles the cell loses orientation in 3D on a timescale of a few seconds[9].

Our paper focuses on the role of filtering, and thus concentrates on the regime of small fluctuations where the nonlinear feedback is not important. Studying the nonlinear regime would be an important direction of future research and we have noted this in the Discussion.

[8] J. Long, S. W. Zucker, and T. Emonet. "Feedback between motion and sensation provides nonlinear boost in run-and-tumble navigation". *PLOS Comp. Biol.* 13.3 (2017), pp. 1–25.

[9] H. C. Berg and D. A. Brown. "Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking." *Nature* 239.5374 (1972), pp. 500–504.

- What are potential issues for generalizing the analysis to 3D-swimming and non-constant gradients?

Mathematically and computationally, the extension to 3D is relatively straightforward: the Keller-Segel drift velocity can be formulated in higher dimensions[10,11,12] and, likewise, the agent-based simulations can be extended to higher dimensions without much modification. In our paper, we have focused on the 1D case to facilitate the analytical insights that can be gained within this simplified setting.

More generally, however, the extension to 3D involves the consideration of additional phenomena. Firstly, as discussed above, 3D swimming leads to smaller turning angles yet with rotational fluctuations. Furthermore, these effects must be studied in conjunction with a model of the attractant landscape. For instance, patchy landscapes, which are frequently considered in the ecological theory[13], can give rise to complex navigation strategies[4] aimed at optimising the use of memory, swimming velocity and directional persistence. Studying these strategies in heterogeneous landscapes is an important point of future work.

Such increased behavioural complexity in 3D also raises challenges as to how to quantify chemotactic performance, which cannot be reduced just to the evaluation of the drift velocity but should involve other quantities (e.g., the amount of encountered chemoattractant) that capture the objective of the cell[14,11].

Regarding the extension to temporally changing gradients, we expect that our results will hold whenever the time scales of bacterial drift and environmental change are well separated, so that the environment appears static to the cell. This effect is accentuated by directional persistence, as the cell is less likely to sample the same location. In this scenario of slow chemoattractant fluctuations, we expect memory to be important for navigation. On the other hand, faster fluctuations would cause the environmental features to be less informative and memory-less, local gradient sensing becomes optimal[15].

We have emphasised and clarified all these points in the Discussion and we have pointed at these directions of future research.

[10] A. Celani and M. Vergassola. "Bacterial strategies for chemotaxis response". *Proc. Natl. Acad. Sci. USA* 107.4 (2010), pp. 1391–1396.

[11] R. Erban and H. Othmer. "From individual to collective behavior in bacterial chemotaxis". *SIAM J. Appl. Math* 65.2 (2004), pp. 361–391.

[12] Y. S. Dufour et al. "Limits of feedback control in bacterial chemotaxis". *PLOS Comp. Biol.* 10.6 (2014), pp. 1–11.

[13] W. K. Chang, D. VanInsberghe, and L. Kelly. "Towards a potential landscape framework of microbiome dynamics". *Preprint at https://biorxiv.org/content/early/2019/11/14/584201* (2019).

[14] A. Gosztolai, J. A. Carrillo, and M. Barahona. "Collective search with finite perception: transient dynamics and search efficiency". *Front. Phys.* 6 (2019), p. 153. ISSN: 2296-424X.

[15] M. Vergassola, E. Villermaux, and B. I. Shraiman. "'Infotaxis' as a strategy for searching without gradients". *Nature* 445 (2007), 406 EP –.

**Reviewer 2**

This manuscript considers the effect of memory on bacterial chemotactic navigation in noisy, shallow concentration gradients with spatial correlations. Memory is effectively achieved in bacterial chemotaxis by the process of adaptation. Simulating a 1D model, the authors compared the drift speeds of agents with memory to those of agents that respond only the local spatial gradient, a commonly used model that is recovered in various limits of the model with memory. One central result of the paper is that memory enhances the drift speed in noisy spatial gradients when the correlation length of the noise is comparable to or a bit shorter than the cell's run length. In the regime of a shallow gradient and noise that is small compared to the gradient steepness, the authors derived an approximate model of the contribution of memory to the drift speed, and it agrees well with simulations. The optimal memory time for each noise correlation length is derived, and the authors discuss how it arises from a balance between averaging out spatial noise and limiting the number of tumbles that occur during the averaging time. The results are interesting and new, but I have several main concerns.

Main:

The current understanding of why E. coli and other bacteria use some form of a run-and-reorient strategy to navigate is that they are too small to directly measure concentration gradients across their body due to diffusion. Instead, they measure spatial gradients from temporal changes in concentration during runs. This suggests that diffusion of most attractants should be fast enough to rapidly damp out concentration fluctuations in the environment that are smaller than one run length. But this is the regime where the authors' model predicts that memory produces the largest gain in drift speed. This makes me have doubts about the relevance of this mechanism for bacterial chemotaxis. Otherwise, the results are interesting and the work is rigorous.

We thank the Reviewer for this important remark, which overlaps with the first remark of Reviewer 1. We provide a response here, but please see also our response to Reviewer 1 for more details, and for the changes made in the manuscript to clarify this point.

As discussed in our response to Reviewer 1, there are several sources of spatial inhomogeneities (e.g., porous and particulate media as well as filamentous formations due to advection) that can have length scales commensurate with the bacterial run length. Please

see details above. In addition to such natural inhomogeneities, it is also possible to experimentally generate spatial ruggedness perceivable by the cell. Berg and Brown[16] reported enzymatically-generated gradients with an average correlation length of $\mu = 11.5\,\mu m$ (less than a run length), which Berg and Purcell[17] showed can be detected in $T = 0.0435$ s (also less than the typical tumbling time of $\sim 1\,s$ in *E. coli*). Hence such small-scale variations generated experimentally could be perceived by the cell.

We have added references and an extended discussion of these issues in the Discussion section of the paper.

`This mechanism could apply to other scenarios in which an agent navigates a noisy environment with gradient-like structure.`

We agree with this suggestion by the Reviewer. Memory-based search strategies are relevant in other biological contexts for higher animals. We have discussed some of these instances in previous work[14], including marine bacteria aggregating at food patches and rodent species that carry out search and exploration strategies. In addition, processes of visual search also rely on specific neurons in the early visual cortex with a response function commonly approximated by a bi-lobed function (c.f. Fig. 6 in Ref.[18]). From this perspective, saccadic eye movements could be thought of as navigation over the image with bi-lobed neurons using their memory to integrate the spatial correlations in the visual field during the visual search.

Mathematically, our work also has links to problems of optimisation over rugged landscapes, a topic of broad interest in many areas of physics, evolutionary biology, computer science and neural networks. Optimisation methods such as simulated annealing exploit local gradients to reach optima of the function combined with a random term to avoid getting trapped in local minima. Current algorithms have extended such methods by introducing geometric cost functions and optimal transport in the optimisation. Our work could therefore serve as an inspiration for the development of heuristic optimisation methods that use particles with memory to go beyond purely local gradients. However, future work would be needed to ascertain the significance of memory in the optimisation of functions in high dimensional spaces.

We have expanded the Discussion to mention these important future research avenues.

`The authors derive an optimal memory time in the case that there is no noise in the gradient that comes out to be 1/4 of a run duration. I am struggling to reconcile this short optimal memory time`

[16] D. A. Brown and H. C. Berg. "Temporal stimulation of chemotaxis in *Escherichia coli*". *Proc. Natl. Acad. Sci. USA* 71.4 (1974).

[17] H. C. Berg and E. M. Purcell. "Physics of chemoreception." *Biophys. J.* 20.2 (1977), pp. 193–219.

[18] E. H. Adelson and J. R. Bergen. "Spatiotemporal energy models for the perception of motion". *J. Opt. Soc. Am. A* 2.2 (1985), pp. 284–299.

with other work (using different models) that suggest longer memory time increases drift speed, in particular Dufour, Fu et al 2014 Plos Comp Bio. In Fig 2A of that paper, simulations of a model in 3D with some implementation of memory and a noiseless gradient had drift speed increasing with memory time even when the memory time was 10 times a typical run duration. I'm not sure whether this simulation was in the shallow gradient regime the authors study. The authors mention in the discussion that one possible reason for the difference from previous work is persistence of direction through a tumble, which is absent from the model in this manuscript and could make longer memory useful. But, to my knowledge, the theoretical predictions in Fig 2A of the same paper Dufour, Fu et al do not consider directional persistence through tumbles (their simulations do).

As the reviewer points out, there are some differences with the work in Dufour *et al* [13]. Specifically, the range of memory $\gamma$ in our paper is shorter; and our drift velocity $v_{\mathrm{KS}}$ (Eq. (6)) exhibits a maximum at a finite memory $\gamma^*$, rather than increasing monotonically with $\gamma$. These differences are due to the choice of kernel functions $K(t)$, as follows.

Our kernel $K(t)$ (Eq. 2) is the bi-lobed function (Fig. 1) corresponding to a three state linear filter (see SI) and is motivated by experimental measurements[19]. It is this bi-lobed function that originates the maximum of the drift velocity, due to the presence of both positive and negative contributions to the integral.

In contrast, Dufour et al. [13] describe the tumbling rate modulations based on the phosphorylation dynamics of the CheY protein through a model that takes account of the non-linear binding kinetics between receptor and chemoattractant molecules. The model is typically linearised assuming that the attractant concentration saturates the active receptor population and does not bind to the inactive receptor population. Under such conditions (see Eq. (2) in Ref. [13]), the tumbling rate is given by a linear function of an internal variable, $\lambda(t) = \lambda_0 - \beta F(t)$, where $F$ is governed by the simplified dynamical model

$$\frac{dF}{dt} = -\frac{1}{\tau}(F - F_0) + \omega S'(x),$$

where $\tau$ is the adaptation timescale (memory); $S'(x)$ is the gradient; and the constant $\omega = \pm\nu N$ is the product of the running velocity $\nu$ and the receptor gain $N$. It then follows that the tumbling response for the linearised Dufour model is given by

$$\lambda(t) = \lambda_0 - \int_{-\infty}^{t} K(t-u)S(x(u))du$$

[19] S. M. Block, J. E. Segall, and H. C. Berg. "Impulse responses in bacterial chemotaxis". *Cell* 31.1 (1982), pp. 215–226. ISSN: 0092-8674.

where the kernel function is a single exponential

$$K(t) = \frac{\beta}{\gamma} e^{-t/\gamma} \tag{1}$$

Since there is no negative contribution in the integral, the drift speed thus increases monotonically.

However, single exponential kernels do not possess the perfect adaptation property and, as shown in Ref. [10], are suboptimal for correlated stimuli. Hence bi-lobed kernels (such as the one we use in our work) are required to capture the experimental observations more precisely.

We have clarified the connection of our work to the model of Dufour *et al.* on page 2 in the manuscript and in a new section S1D in the SI.

Minor:

Most bacteria sense the log of concentration and gradients of the log, for a large concentration range. But this doesn't substantially affect the authors' results.

We thank the reviewer for this remark. Indeed, because of the concavity of the log function, the behaviour will be very similar. We have clarified this point in the Discussion.

- There is a small error in the legend of Fig. 2b. The average drift speed of AB cells is averaged over the ensemble of simulated cells but not over gradient noise (as currently written) since there isn't any noise in this case.
- At the bottom of page 8, first column, \studies" is repeated an extra time.
- The last \and" of the Discussion became \ad".

All corrected, with thanks.

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

I appreciate the thoughtful responses of the authors to my comments.

Reviewer #2 (Remarks to the Author):

The authors' additions to the Discussion have addressed my concerns about whether the small-scale chemical structures considered in this paper might occur in the real world.

I also appreciate the Discussion about how their 1D model differs from 3D models that also allow for directional persistence and rotational diffusion. However, I noticed that the authors made an error in their analysis of Dufour et al's model: that model is perfectly adapting and has a kernel that integrates to zero. To see this, when changes in concentration are absent, F always returns to F_0, bringing the tumble rate back to the same baseline. The exponential kernel that the authors identified in section D of the SI is one that operates on the time derivative of concentration, not on concentration (error in Eqn. S10). Integrating by parts, one can show that this is equivalent to a kernel acting on concentration that is proportional to a delta function minus a decaying exponential, which together integrate to zero. Therefore, imperfect adaptation cannot explain why drift increases with memory time in Dufour's model, while the authors' model exhibits an optimal memory time. I don't think it is necessary for the authors to work out the reason for this difference—clearly raising it in the Discussion, as they have done, is enough to me.

**Response to Reviewer comments** We addressed the remark of the Reviewer 2 by correcting the computation in Section S1D in Supplementary Note 1. It is now in full agreement with the advice of the Reviewer. Correspondingly, we changed the relevant sentence on page 2, Section 1 of the main text:

*"In Sect. S1D in Supplementary Note 1, we show that this model can be equivalently written as (1) with a single exponential kernel* $K(t) = (\beta/\gamma)e^{-t/\gamma}$*, which does not possess the perfect adaptation property."*

to the following sentence

*"In Sect. S1D in Supplementary Note 1, we show that this alternative model can be equivalently written in the form (1) but with a kernel* $K(t) = \beta(\delta(t) - (1/\gamma)e^{-t/\gamma})$ *with a singularity at* $t = 0$ *and a single decaying exponential, instead of the bi-lobed kernel (2)."*

No further technical changes to the manuscript were necessary.