

The effect of the fitness gradient on fixation probability

Corresponding Author: Professor Krishnendu Chatterjee

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This manuscript studies a model of evolution a mutant type in a population where the fitness of the mutant is location dependent. This kind of heterogeneous fitness is relevant as it has been observed in various real-world settings. There's been various biological studies of similar scenarios that have explored how the heterogeneity affects the rate of evolution.

The main result of the manuscript is rather unexpected. It shows that the (linear) heterogeneity is significantly beneficial, in terms of fixation probability, for the mutant up to a certain cutoff, beyond which it becomes significantly detrimental. This result might shed light on the causes of the aforementioned biological research results.

The manuscript provides both numerical evidence and rigorous analytical demonstration of the claimed behavior. The mathematical arguments are sound, clear, and well-written. They constitute a correct and complete proof of the claimed (and numerically demonstrated) behavior.

Minor comments:

- in section 4 of the supplementary material, it might be helpful to explain in more detail how the argument with the residual graph works. In particular, there seems to be a typo ($1+s/2$ instead of $s/2$) when describing the fitness in the residual graph.
- there is a broken reference near the end of section 4.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

In this paper, the authors study a one-dimensional Moran process where only the mutant feels a linear fitness gradient, while residents keep constant average-matching fitness. It shows three behaviors as the gradient grows with system size: near-neutral fixation for tiny gradients, a non-monotonic increase with a peak at intermediate gradients, and strong suppression for large gradients. Results also depend on where the mutant starts, with gradient-favoured sites driving the peak. Overall, the work clarifies when weak spatial variation can amplify fixation despite zero average advantage, challenging the usual "heterogeneity suppresses fixation" intuition.

The biological examples are apt but mapped to an idealised model, sometimes too tightly. A strictly 1D, isothermal line differs from 2D/3D experimental settings with diffusion fronts, so quantitative transfer is limited. The "mutant-only gradient" assumption is plausible yet narrow; many systems have both types responding to space. Also, end self-loops simplify boundaries unlike real fluxes/curvature. Net: the examples are illustrative, not literal—clearer limits would avoid "forcing the analogy."

I suggest a minor revision to address the following points:

- 1) Is the first inequality in Theorem 2 correct?
- 2) Consider giving each regime a precise asymptotic label at first mention (e.g., "negligible", "small", "large").

3) Add a sentence explaining why isothermality via end self-loops is mathematically convenient, and what changes if they are removed; this would pre-empt reader questions.

4) The Discussion briefly situates the work; adding a compact table or figure summarising how your results contrast with known "suppressor" findings under different heterogeneity models would sharpen the contribution.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

The authors have addressed all the feedback provided during the review process, enhancing the article.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

We thank the Editor for handling our manuscript and both Reviewers for their constructive comments. We have revised the manuscript accordingly. Below, we provide detailed responses to each review. Moreover, we append the manuscript with changes highlighted.

Reviewer #1 (Remarks to the Author):

This manuscript studies a model of evolution a mutant type in a population where the fitness of the mutant is location dependent. This kind of heterogeneous fitness is relevant as it has been observed in various real-world settings. There's been various biological studies of similar scenarios that have explored how the heterogeneity affects the rate of evolution.

The main result of the manuscript is rather unexpected. It shows that the (linear) heterogeneity is significantly beneficial, in terms of fixation probability, for the mutant up to a certain cutoff, beyond which it becomes significantly detrimental. This result might shed light on the causes of the aforementioned biological research results.

The manuscript provides both numerical evidence and rigorous analytical demonstration of the claimed behavior. The mathematical arguments are sound, clear, and well-written. They constitute a correct and complete proof of the claimed (and numerically demonstrated) behavior.

ANS: Thank you for the positive assessment of our work.

Minor comments:

- in section 4 of the supplementary material, it might be helpful to explain in more detail how the argument with the residual graph works. In particular, there seems to be a typo ($1+s/2$ instead of $s/2$) when describing the fitness in the residual graph.

ANS: We have fixed the typo. And we clarified the definition and the motivation behind the residual graph.

- there is a broken reference near the end of section 4.

ANS: We have fixed the broken reference.

Reviewer #2 (Remarks to the Author):

In this paper, the authors study a one-dimensional Moran process where only the mutant feels a linear fitness gradient, while residents keep constant average-matching fitness. It shows three behaviors as the gradient grows with system size: near-neutral fixation for tiny gradients, a non-monotonic increase with a peak at intermediate gradients, and strong suppression for large gradients. Results also depend on where the mutant starts, with gradient-favoured sites driving the peak. Overall, the work clarifies when weak spatial variation can amplify fixation despite zero average advantage, challenging the usual "heterogeneity suppresses fixation" intuition.

The biological examples are apt but mapped to an idealised model, sometimes too tightly. A strictly 1D, isothermal line differs from 2D/3D experimental settings with diffusion fronts, so quantitative transfer is limited. The "mutant-only gradient" assumption is plausible yet narrow; many systems have both types responding to space. Also, end self-loops simplify boundaries unlike real fluxes/curvature. Net: the examples are illustrative, not literal—clearer limits would avoid "forcing the analogy."

ANS: Thank you for a positive assessment of our work and for constructive comments. We have now added computational results for the 2D setting without self-loops to illustrate that our results are robust.

Change: We add section 6 in SI, together with Figure 1, which shows the fixation probability in a two-dimensional grid.

We mention in the main text, mainly on line 96-97; line 144-145; and line 252.

I suggest a minor revision to address the following points:

1) Is the first inequality in Theorem 2 correct?

ANS: The inequality is correct, but thanks for asking -- upon closer inspection, we found a minor error in the mathematical derivation (it occurred in the step where we bounded the geometric sum). We have now fixed the error (by bounding the corresponding terms in the subsequent step of the derivation), which even made the inequality slightly stronger.

Change: In SI, Section 4, page 5, we add computations for the geometric sum.
In the main text, the statement of Theorem 2 is tightened.

2) Consider giving each regime a precise asymptotic label at first mention (e.g., "negligible", "small", "large").

ANS: Great suggestion, we have done that in the abstract and clarified in the introduction.

Change: Abstract, Line 79-82.

3) Add a sentence explaining why isothermality via end self-loops is mathematically convenient, and what changes if they are removed; this would pre-empt reader questions.

ANS: The isothermality is convenient, because it simplifies the baseline regime $s=0$. When the graph is isothermal, the fixation probability in the baseline regime $s=0$ is equal to $1/N$, regardless of the starting position of the mutant. In the setting where self-loops are removed, the baseline fixation probability at $s=0$ becomes $2/(N+2)$ for mutants who start at either of the two endpoints of the line, and $1/(N+2)$ for mutants who start elsewhere. Thus, e.g. in Fig. 4(b) the curves would attain different values at $s=0$ (namely $2N/(N+2)$ for blue and purple, vs. $1/(N+2)$ for orange, green, and red). Our mathematical derivations suggest that the difference between the two settings (with or without self-loops) should decay exponentially quickly in the distance d from the boundary, so the overall trend should not be affected. This is confirmed by additional computational results for the 2D-grid, (see Section 6 in the Supplementary Information) where we did not use the self-loops. As suggested, we now also comment along those lines in the main text.

Change: We comment on loops on Line 118-121.

4) The Discussion briefly situates the work; adding a compact table or figure summarising how your results contrast with known "suppressor" findings under different heterogeneity models would sharpen the contribution.

ANS: Good idea. Briefly: Our results strongly contrast with the results of ref [32] that found that in well-mixed populations, mutant fitness heterogeneity monotonically reduces the mutant fixation probability. To a smaller extent, our results also contrast with refs [28,29] who found that in two-island models, mutant fitness heterogeneity monotonically increases the mutant fixation probability. None of those works identified the intriguing non-monotonic behavior that we find for the one-dimensional lattice.

Change: We have added the table as suggested (Table 1).