# ESTIMATING MUTATION RATES UNDER HETEROGENEOUS STRESS RESPONSES

A PREPRINT

**Lucy Lansch-Justen**[1*], **Meriem El Karoui**[2,3], **and Helen K. Alexander**[1,3†]
[1]Institute of Ecology and Evolution, School of Biological Sciences, University of Edinburgh, UK
[2]Institute of Cell Biology, School of Biological Sciences, University of Edinburgh, UK
[3] Centre for Engineering Biology, University of Edinburgh, UK

September 5, 2023

**ABSTRACT**

Exposure to environmental stressors, including certain antibiotics, induces stress responses in bacteria. Some of these responses increase mutagenesis and, this way, potentially accelerate resistance evolution. Many studies report increased mutation rates under stress, often using the standard lab approach of fluctuation assays to estimate the mutation rates. However, many stress responses are heterogeneously expressed in bacterial populations, which existing estimation methods have not yet addressed. We develop a population dynamic model that considers heterogeneous stress responses (subpopulations of cells with the response *off* or *on*) and derive the mutant count distribution arising in fluctuation assays under this model. We then implement a computational method to estimate the mutation-rate increase specifically associated with the induction of the stress response. Using simulated mutant count data, we show that our inference method allows for accurate and precise estimation of the mutation-rate increase, provided induction of the response also reduces the division rate, as is the case for the SOS response. Interestingly, we find that existing methods assuming a homogeneous stress response can also accurately infer the increase in population mean mutation rate. Without additional experiments, it is not generally possible to distinguish between a model with heterogeneously expressed stress responses and a model in which mutants have a differential fitness compared to non-mutants.

***Keywords*** stress-induced mutagenesis · fluctuation assay · phenotypic heterogeneity · mathematical model · antibiotics · inference method

## 1 Introduction

Bacteria are commonly exposed to adverse conditions, such as starvation, sub-optimal temperatures or toxins, including antibiotics. To cope with these conditions, bacteria have evolved a range of stress responses that enhance viability under stress. Some of these response pathways also increase mutagenic mechanisms by, for example, increasing the expression of error-prone polymerases or down-regulating error-correcting enzymes [11, 9]. It has been proposed that this so-called 'stress-induced mutagenesis' in bacterial cells could accelerate the evolution of populations that are poorly adapted to their environment [5, 33]. Consequently, inhibiting bacterial stress responses has been suggested to prevent antibiotic resistance evolution [7, 32, 40].

Several studies report increased mutation rates in bacterial populations exposed to sub-lethal antibiotic concentrations [39, 14, 18, 2, 24, 16, 31]. The typical approach to measuring microbial mutation rates and how they increase under stress is to use a fluctuation assay [8, 25]: The standard experiment involves growing parallel cultures under permissive conditions while mutations occur randomly and subsequently plating the cultures onto selective

*lucy.lanju@googlemail.com
†helen.alexander@ed.ac.uk

plates. The mutation rate is estimated from the distribution of the number of mutants on the plates. To quantify the mutation-rate increase associated with stress, the standard experiment is repeated, but this time, the cultures are exposed to stress during the growth phase. Then, the *stress-induced* mutation rate is estimated and compared with the mutation rate under permissive conditions. However, stress impacts the growth of bacterial cells in several ways, which are not accounted for in commonly applied estimation methods, potentially leading to biased estimates of the mutation rate. For instance, increased cell death leads to an overestimation of the mutation rate [10]. Another effect that has not yet been addressed is cell-to-cell heterogeneity in stress responses.

In recent years, single-cell experiments have revealed extensive heterogeneity in the expression of stress responses [29, 30, 22, 35, 34, 38, 37, 19, 6]. Heterogeneity can arise for various reasons, including stochastic gene expression of stress response proteins, especially when present in small numbers [30, 22, 35], phenotypic variability in the stability of key regulators [21], or micro-environmental variation in cell-to-cell interactions [6]. In general, positive and negative feedback loops are common features of stress response regulatory networks, which generate characteristic dynamical behaviours, including cell-to-cell variation [27]. In some cases, a subpopulation of cells showing elevated stress responses has been directly associated with a higher rate of DNA mismatches or higher mutant frequency [30, 22, 35, 38, 31, 37]. In addition to mutagenic mechanisms, stress responses can alter cell division and death rates. For example, the widely studied SOS response, which leads to the transcriptional induction of approximately 40 genes after exposure to DNA damage, involves inhibition of cell division and induction of error-prone polymerases that could increase mutation rate [12, 3]. Single-cell studies using fluorescent reporters for the SOS response have revealed that its expression is highly heterogeneous in many bacterial species [29, 22, 19]. It has been shown that this heterogeneity can be approximated as a bimodal response [19] with some bacteria expressing low levels of SOS under exposure to DNA damage whilst others show a very high level of expression. Overall, heterogeneously expressed stress responses impact both bacterial population dynamics and mutational input during the growth phase of a fluctuation assay, and it is unclear whether estimation methods that neglect heterogeneity in stress responses produce reliable results.

In this study, we present a population dynamics model that considers cell-to-cell heterogeneity in stress responses. Motivated by the SOS response, we describe two discrete subpopulations of cells, where induction of the stress response is associated with both a higher mutation rate and a lower division rate. We derive the resulting mutant count distribution in the total population and implement a computational method to estimate the mutation-rate increase associated with the induction of the stress response. We test the performance of our method using stochastic simulations of fluctuation assays under permissive and stressful conditions, including robustness to biologically realistic model deviations such as mutant fitness costs and cell death. We also apply formal model comparison to assess whether cell-to-cell heterogeneity could be detected from fluctuation assays alone.

## 2 Model & Methods

### 2.1 Population dynamics with heterogeneous stress responses

We model the population dynamics of bacterial cultures during the growth phase of a fluctuation assay, (i) under permissive conditions and (ii) under exposure to some form of stress that induces a mutagenic stress response in the cells. Our model considers cell-to-cell heterogeneity in the expression of the stress response. In particular, for simplicity, we assume that under stress, the population can be divided into two subpopulations: one with a low expression level of the stress response (here referred to response switched *off* even if strictly speaking the response is not fully off but very low) and the other with a high expression level (here referred to response switched *on*). This *heterogeneous-response model* differs from models used in currently available mutation-rate estimation methods, which allow incorporation of stress responses only in a homogeneous manner, i.e. all cells responding to the stress in the same way (*homogeneous-response model*).

In the following, we describe the full heterogeneous-response model we use in the simulations to test our inference algorithm and the approximations we make to derive the mutant count distributions used in the inference.

#### 2.1.1 Heterogeneous-response model

We consider a bacterial population previously growing under permissive conditions and now exposed to stress. Exposure to stress can change the division rate and cause cell death, which we assume to be absent without stress. Moreover, cells induce a stress response with a constant switching rate $\alpha$, leading to the emergence of a response-*on* subpopulation. Inducing the stress response alters the mutation rate of the cells but potentially also their division and/or death rates. We assume that, as long as the stress persists, cells do not switch the response *off* again.

The above-described population dynamics result in the following set of coupled differential equations for the population sizes of the non-mutant response-*off* and response-*on* subpopulations, $n_{off}$ and $n_{on}$, respectively

$$\dot{n}_{off} = (\gamma^s_{off} - \delta_{off} - \alpha)n_{off}, \tag{1}$$

$$\dot{n}_{on} = \alpha n_{off} + (\gamma_{on} - \delta_{on})n_{on}, \tag{2}$$

Here, $\gamma^s_{off}$ is the division rate of the response-*off* subpopulation under stress (which can be different than under permissive conditions, $\gamma^p_{off}$), $\delta_{off}$ the death rate of the response-*off* subpopulation, and $\gamma_{on}$ and $\delta_{on}$ the division and death rate of the response-*on* subpopulation. The solution to these equations is given by

$$n_{off}(t) = n_{off}(0)\mathrm{e}^{(\gamma^s_{off} - \delta_{off} - \alpha)t} \tag{3}$$

$$n_{on}(t) = \frac{\alpha n_{off}(0)}{\gamma^s_{off} - \delta_{off} - \alpha - (\gamma_{on} - \delta_{on})}\left(\mathrm{e}^{(\gamma^s_{off} - \delta_{off} - \alpha)t} - \mathrm{e}^{(\gamma_{on} - \delta_{on})t}\right) + n_{on}(0)\mathrm{e}^{(\gamma_{on} - \delta_{on})t}. \tag{4}$$

with $n_{off}(0)$ being the initial population size of the response-*off* subpopulation and $n_{on}(0)$ the number of response-*on* cells before the exposure to the stress.

This approach assumes that the response-*on* subpopulation can be treated deterministically. We test the validity of this assumption using stochastic simulations; we simulate switching *on* of the response as a time-inhomogeneous Poisson process and the growth dynamics of the response-*on* subpopulation as a stochastic birth-death branching process. Then, we compare the resulting population size with equation 4. We find that this assumption is valid for a wide range of switching rates and division rates of response-*on* cells and that it holds for zero and small initial sizes of the response-*on* subpopulation (Appendix Figure 6).

In the simulations to test our inference algorithm, we use equations 3 and 4 for the dynamics of the non-mutant population, setting $n_{on}(0) = \frac{\alpha}{\gamma^s_{off} - \delta_{off}}$. For the inference itself, however, we approximate equation 4 as

$$\hat{n}_{on}(t) \approx \frac{\alpha n_{off}(0)}{\gamma^s_{off} - \delta_{off} - \alpha - (\gamma_{on} - \delta_{on})}\mathrm{e}^{(\gamma^s_{off} - \delta_{off} - \alpha)t}. \tag{5}$$

because this allows us to derive an analytical expression for the mutant count distribution later. This approximation is valid when the initial population size of the response-*on* subpopulation is comparably small, $n_{on}(0) \ll n_{off}(0)$, and its growth is slower than the growth of the response-*off* subpopulation, $\gamma_{on} - \delta_{on} \ll \gamma^s_{off} - \delta_{off} - \alpha$. As a consequence of this approximation, the total population grows exponentially with a *population growth rate* of

$$\lambda^s = \gamma^s_{off} - \delta_{off} - \alpha, \tag{6}$$

and the response-*on* subpopulation makes up a constant fraction of

$$f_{on}(t) = \frac{n_{on}(t)}{n_{on}(t) + n_{off}(t)} = \frac{\alpha}{\gamma^s_{off} - \delta_{off} - (\gamma_{on} - \delta_{on})} =: f^*_{on} \tag{7}$$

For the inference, we assume that the fraction of the response-*on* subpopulation at the end of the growth phase, $f_{on}$, is a good approximation of $f^*_{on}$.

Defining the relative switching rate as $\tilde{\alpha} := \frac{\alpha}{\gamma^s_{off} - \delta_{off}}$ and the differential fitness of response-*on* compared to response-*off* cells as $r_{on} := \frac{\gamma_{on} - \delta_{on}}{\gamma^s_{off} - \delta_{off}}$, we obtain

$$f_{on} = \frac{\tilde{\alpha}}{1 - r_{on}} \quad \Rightarrow \quad \tilde{\alpha} = f_{on}(1 - r_{on}). \tag{8}$$

In the simulations and the inference, we consider mutations in the response-*off* and the response-*on* subpopulation to occur according to time-inhomogeneous Poisson processes and treat the dynamics of the resulting mutants stochastically. Mutations arise in each subpopulation at rates $\nu^s_{off} n_{off}(t)$ and $\nu_{on} n_{on}(t)$ (simulation), or $\nu_{on}\hat{n}_{on}(t)$ (inference), respectively. Note that $\nu$ describes the mutation rate to the phenotype of interest selected on the plates (mutations per cell per unit time). In the simulations, we set the duration $t_f$ of the growth phase such that the expected number of *mutations* (not *mutants*) is constant across the considered parameter ranges (Appendix equations 1, 2).

In a fluctuation assay, the selection on the plates is chosen to be unrelated to the stress and the stress response. Therefore, mutants are expected to respond to the stress similarly to non-mutants and can be divided into response-*off* and response-*on* subpopulations. Nonetheless, mutations might affect the fitness during the growth phase,

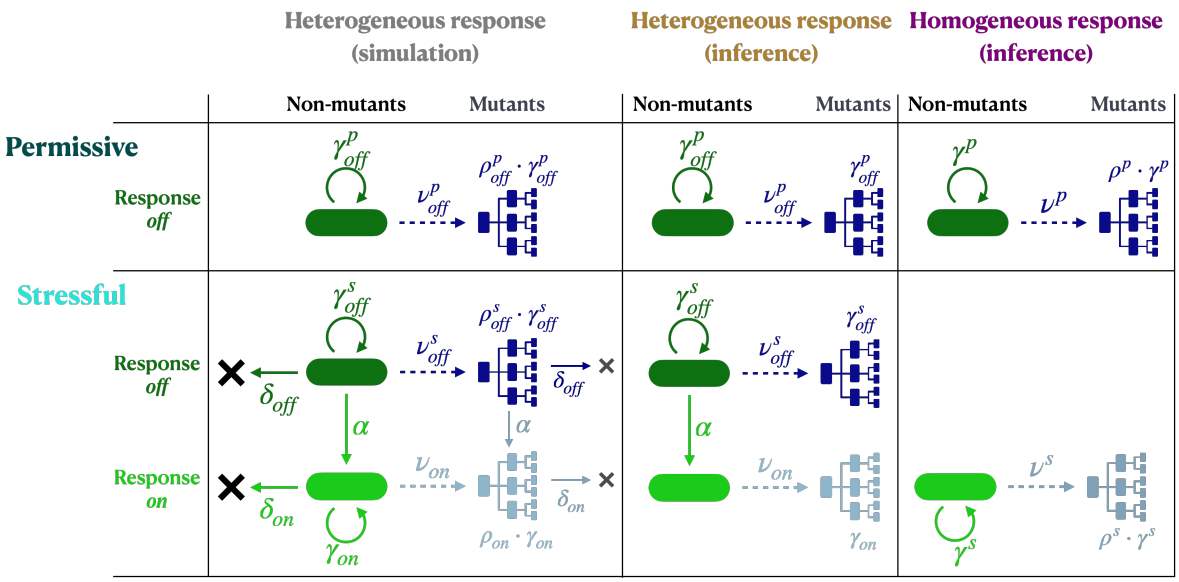


Figure 1: Schematic illustrating the heterogeneous-/homogeneous-response models used in simulations and the inference.

which we implement by an altered division rate of mutants. Together, this results in a two-type branching process with birth rates $\rho^s_{off}\gamma^s_{off}$ and $\rho_{on}\gamma_{on}$, death rates $\delta_{off}$ and $\delta_{on}$, and switching from response-*off* to response-*on* at rate $\alpha$.

When testing our inference algorithm, we simulate the mutant dynamics using the above-described two-type branching process. For the inference, however, we neglect the induction of the stress response in the mutants. For mathematical convenience, we assume that mutations have no fitness effect and consider switching as a reduction in the division rate of response-*off* mutants by $\alpha$ instead (birth rate equal to $\gamma^s_{off} - \alpha$). This approximation allows us to derive an analytical expression for the mutant count distribution because the dynamics of response-*off* and response-*on* mutants are now independent birth-death branching processes and the total mutant count distribution is given by the sum of the mutant contribution of the response-*off* and response-*on* subpopulation as two independent random variables.

Figure 1 summarises the full heterogeneous-response model used in the simulation, the version of the heterogeneous-response model approximated for the inference and the homogeneous-response model used in current inference methods as a comparison.

The complete documentation of all population dynamic functions can be found in the file called **population_dynamics.jl** at `https://github.com/LucyL-J/Quantifying-SIM`.

### 2.1.2 Mutant count distributions

We use the results from [23], who derive the mutant count distribution for the general case of an exponentially growing population with initial population size $N_i$, population growth rate $\lambda$, mutation rate $\nu$, mutants with birth rate $b$ and death rate $d$ and for 100% plating efficiency. They calculate the probability-generating function $G(z)$ of the mutant count distribution when the population reaches a final population size $N_f$. For $N_f \gg N_i$ (neglecting initial population size effects), and defining the per-generation mutation rate as $\mu := \frac{\nu}{\lambda}$, the differential fitness of mutants as $\rho := \frac{b-d}{\lambda}$ and the extinction probability of mutants as $\epsilon := \frac{d}{b}$, G(z) is given by

$$G(z) = \exp\left[-\mu N_f\,(1-\epsilon)\,F\left(\begin{matrix}1,\ \frac{1}{\rho}\\ 1+\frac{1}{\rho}\end{matrix};\ \frac{z-\epsilon}{z-1}\right)\right] \tag{9}$$

with $F$ being the hypergeometric function. Moreover, the authors [23] give a recursive expression for the probabilities $p_m$ to observe $m$ mutants, which allows for the numerically relatively inexpensive calculation of the likelihood of observing empirical mutant count data.

Using equation 9, we can substitute the corresponding parameters to obtain the mutant count distributions

under permissive/stressful conditions for the heterogeneous-response model. Importantly, this assumes that all mutants, including response-*on* cells, can form a colony on the selective plates.

As done in currently available methods, we assume that under permissive conditions, there is no response-*on* subpopulation nor cell death ($\epsilon = 0$). In addition, we consider the simplifying case that mutants have the same division rate as non-mutants ($\rho = 1$). This results in the following mutant count distribution under permissive conditions:

$$G^p_{\text{het}}(z) = \exp\left[\mu_{off} N^p_f \frac{(1-z)\log(1-z)}{z}\right], \tag{10}$$

where $\mu_{off} := \frac{\nu^p_{off}}{\gamma^p_{off}}$ describes the *per-division* mutation rate, which equals the *per-generation* rate as, under permissive conditions, the population growth is solely determined by cell division.

Under stressful conditions, on the other hand, the population growth rate depends on the switching rate $\alpha > 0$ via

$$\lambda^s = \gamma^s_{off} - \delta_{off} - \alpha = \gamma^s_{off}(1-\epsilon_{off})(1 - f_{on}(1-r_{on})) \tag{11}$$

which we have re-written as a function of the division rate $\gamma^s_{off}$ and the extinction probability $\epsilon_{off} := \frac{\delta_{off}}{\gamma^s_{off}}$ of response-*off* cells, and the fraction $f_{on}$ and differential fitness $r_{on}$ of the response-*on* subpopulation. Importantly, we assume that the per-division mutation rate of response-*off* cells is the same under stressful as under permissive conditions, $\frac{\nu^s_{off}}{\gamma^s_{off}} = \frac{\nu^p_{off}}{\gamma^p_{off}}$. Then, the per-generation mutation rates of response-*off* and response-*on* cells are given by

$$\frac{\nu^s_{off}}{\lambda^s} = \frac{\nu^s_{off}}{\gamma^s_{off}(1-\epsilon_{off})(1-f_{on}(1-r_{on}))} = \frac{\mu_{off}}{(1-\epsilon_{off})(1-f_{on}(1-r_{on}))} \tag{12}$$

$$\frac{\nu_{on}}{\lambda^s} = \frac{\nu_{on}}{\gamma^s_{off}(1-\epsilon_{off})(1-f_{on}(1-r_{on}))} = \frac{\mu_{on}}{(1-\epsilon_{off})(1-f_{on}(1-r_{on}))} \tag{13}$$

with $\mu_{on} := \frac{\nu_{on}}{\gamma^s_{off}}$. For response-*on* cells, we have that the differential fitness of mutants $\rho_{on}$ (compared to the population growth rate) can be expressed via

$$\rho_{on} = \frac{\gamma_{on} - \delta_{on}}{\lambda^s} = \frac{r_{on}}{1 - f_{on}(1-r_{on})}. \tag{14}$$

The total mutant count is the sum of the independent contributions from the response-*off* and response-*on* subpopulations. Therefore, the probability generating function of the distribution of the total mutant count under stressful conditions, $G^s_{\text{het}}(z)$, is given by the product of $G^s_{off}(z) \cdot G^s_{on}(z)$, where

$$G^s_{off}(z) = \exp\left[\frac{\mu_{off}}{1-f_{on}(1-r_{on})}(1-f_{on})N^s_f \frac{(1-z)\log\left(\frac{1-z}{1-\epsilon_{off}}\right)}{z-\epsilon_{off}}\right] \tag{15}$$

$$G^s_{on}(z) = \exp\left[-\frac{\mu_{on}}{1-f_{on}(1-r_{on})}\frac{1-\epsilon_{on}}{1-\epsilon_{off}} f_{on} N^s_f F\left(\begin{matrix}1, \frac{1-f_{on}(1-r_{on})}{r_{on}} \\ 1 + \frac{1-f_{on}(1-r_{on})}{r_{on}}\end{matrix}; \frac{z-\epsilon_{on}}{z-1}\right)\right]. \tag{16}$$

In the case of $\gamma_{on} = 0$, the contribution to the mutant count from the response-*on* subpopulation follows a Poisson distribution and equation 16 reduces to

$$G^s_{on}(z) = \exp\left[\frac{\mu_{on}}{1-f_{on}(1-r_{on}) - r_{on}} \frac{1}{1-\epsilon_{off}} f_{on} N^s_f (z-1)\right]. \tag{17}$$

Note, that in this case $r_{on} = -\frac{\delta_{on}}{\gamma^s_{off} - \delta_{off}} \leq 0$.

The homogeneous-response model assumes that all cells respond to the stress similarly. The stress response results in a change in the per-generation mutation rate, $\mu^p \to \mu^s$, and potentially the differential fitness of mutants compared to non-mutants, $\rho^p \to \rho^s$. Therefore, the probability-generating functions under permissive/stressful conditions are given by

$$G^p_{\text{hom}} = \exp\left[-\mu^p N^p_f F\left(\begin{matrix}1, \frac{1}{\rho^p} \\ 1 + \frac{1}{\rho^p}\end{matrix}; \frac{z}{z-1}\right)\right] \tag{18}$$

$$G^s_{\text{hom}} = \exp\left[-\mu^s N^s_f (1-\epsilon^s) F\left(\begin{matrix}1, \frac{1}{\rho^s} \\ 1 + \frac{1}{\rho^s}\end{matrix}; \frac{z-\epsilon^s}{z-1}\right)\right] \tag{19}$$

assuming again that there is no cell death under permissive conditions.

### 2.2 Estimation of the mutation-rate increase under stress

We use joint maximum likelihood estimation to find the set of model parameters for which the empirically observed mutant count data are most likely. We consider the permissive and the stressful condition concurrently and jointly infer the respective parameters. The joint inference crucially relies on our assumption that, even though the division rate might change under stress, the *per-division* mutation rate of response-*off* cells is constant under permissive/stressful conditions. Then, we define a log-likelihood function

$$\mathcal{L}(x^p, x^s|y) \tag{20}$$

that is a measure for the probability of observing the mutant counts $x^p$ and $x^s$ under permissive and stressful conditions, respectively, for a given model and model parameters $y$. It is calculated using the probabilities of the above-described mutant count distributions. Then, we use an optimisation algorithm implemented in the Julia [4] package Optim.jl (`https://julianlsolvers.github.io/Optim.jl/stable/`), to find the parameters that maximise this log-likelihood function. The initial parameters for the optimisation are calculated using the generating function method described in [15].

#### 2.2.1 Inferred and measured parameters

The mutant count distributions under permissive/stressful conditions derived for the heterogeneous-response model, equations 10, 15 and 16, depend on the following parameters:

- $\mu_{off}$: The per-division mutation rate of response-*off* cells is inferred.
- $\mu_{on}$: The per-division mutation rate of response-*on* cells is also inferred. Note that this rate is given in 'mutations in response-*on* cells per divisions in response-*off* cells' because the division rate of response-*on* can also be zero; this also makes the mutation rates $\mu_{off}$ and $\mu_{on}$ directly comparable.
- $N_f^{p,s}$: The final population sizes under permissive and stressful conditions are taken as known quantities in our inference and would normally be measured separately, typically through plating on non-selective media and colony counting.
- $r_{on}$: The differential fitness of response-*on* compared to response-*off* cells is an optional inference parameter. By default, it is set to $r_{on} = 0$, but it can also be inferred (Results section 3.3) or set to another value if an estimate of $r_{on}$ is available (Appendix Figure 9).
- $f_{on}$: The fraction of the response-*on* subpopulation at the end of the growth phase is either taken as a known quantity, which would have to be measured in a separate experiment (e.g. microscopy or flow cytometry), or inferred (Results section 3.4).
- $\epsilon_{off}$, $\epsilon_{on}$: The extinction probabilities of response-*off* cells under stress and of response-*on* cells are set to zero in the inference, corresponding to no cell death. We test the robustness of this assumption (Results section 3.2).

Ultimately, we are interested in quantifying the relative increase in mutation rate associated with induction of the stress response, that is, $\frac{\mu_{on}}{\mu_{off}}$. Alternatively, the increase in population mean mutation rate $\Delta\bar{\mu} = \frac{(1-f_{on})\mu_{off}+f_{on}\mu_{on}}{\mu_{off}}$ can be calculated.

The mutant count distributions under permissive and stressful conditions derived for the homogeneous-response model, equations 18 and 19, depend on the following parameters:

- $\mu^{p,s}$: The per-generation mutation rates under permissive/stressful conditions are inferred.
- $N_f^{p,s}$: The final population sizes are, as for the heterogeneous-response model, taken as known quantities.
- $\rho^{p,s}$: The differential fitness of mutants compared to non-mutants is an optional inference parameter, implemented, for example, in [28]. By default, it is set to $\rho^{p,s} = 1$. Alternatively, it can be considered a joint inference parameter, $\rho^p = \rho^s$, or two separate inference parameters, $\rho^p$ and $\rho^s$.
- $\epsilon$: The extinction probability is set to zero in the inference because we neglect cell death, in common with most existing approaches, but see [10].

Ultimately, the quantity of interest is the mutation-rate increase associated with the stress, that is, $\frac{\mu^s}{\mu^p}$. Note, however, that both $\mu^p$ and $\mu^s$ are population-wide rates and $\frac{\mu^s}{\mu^p}$ is equivalent to the increase in population mean mutation rate under stress.

The complete documentation of all inference algorithms can be found in the file called **inference.jl** at `https://github.com/LucyL-J/Quantifying-SIM`.

### 2.3 Testing the inference algorithms using stochastic simulations

To test the inference algorithms, we perform stochastic simulations of pairs of fluctuation assays, one under permissive and the other under stressful conditions. Then, we infer the respective parameters under different models and compare them to the true underlying parameters; see Figure 1 for schematics of the models used in simulation and inference.

#### 2.3.1 Accuracy and precision

First, we evaluate the performance of the estimation of the mutation-rate increase associated with the stress response. For this purpose, we simulate fluctuation assays under the heterogeneous-response model without cell death or differential mutant fitness and for (a) $\gamma_{on} = 0$, with varying $\{\mu_{on},\, \alpha\}$ (Results section 3.1), and (b) varying $\gamma_{on} > 0$ (Results section 3.3). Then, we infer the underlying model parameters under the same model as simulated (with the only difference that switching of mutants and initial population size effects are neglected in the inference, as described above). In case (a), we set the relative fitness of response-*on* cells, $r_{on}$, to zero, and in case (b), we consider both setting $r_{on} = 0$ and inferring $r_{on}$. In either case, we assume that separate estimates of $N_f^{p,s}$ and $f_{on}$ are available, and we take these as fixed values.

We simulate $R = 100$ pairs of fluctuation assays, with $c = 50$ parallel cultures per fluctuation assay for each parameter set. We evaluate the accuracy of the estimation by calculating the relative error of the median across the $R$ replicates,

$$\mathrm{RE} = \frac{\mathrm{median}\left(\left\{\frac{\mu_{on}}{\mu_{off}}\right\}_{\mathrm{est}}\right) - \left(\frac{\mu_{on}}{\mu_{off}}\right)_{\mathrm{true}}}{\left(\frac{\mu_{on}}{\mu_{off}}\right)_{\mathrm{true}}}. \tag{21}$$

Here, a positive or negative relative error implies over- or under-estimation, respectively. We also evaluate the precision of the estimation by calculating the coefficient of variation across the $R$ replicates,

$$\mathrm{CV} = \frac{\mathrm{std}\left(\left\{\frac{\mu_{on}}{\mu_{off}}\right\}_{\mathrm{est}}\right)}{\mathrm{mean}\left(\left\{\frac{\mu_{on}}{\mu_{off}}\right\}_{\mathrm{est}}\right)}. \tag{22}$$

#### 2.3.2 Robustness to model deviations

Next, we evaluate the robustness of the estimation to model deviations. For this purpose, we simulate fluctuation assays under a more complex version of the heterogeneous-response model than is used in the inference. We simulate with (a) cell death in three different forms: $\{\delta_{off} = \delta_{on} > 0\}$, $\{\delta_{off} > 0,\, \delta_{on} = 0\}$, or $\{\delta_{off} = 0,\, \delta_{on} > 0\}$; and (b) mutants having a different fitness than non-mutants. In the inference, however, we neglect these effects and evaluate whether this introduces any biases to the estimates (Results section 3.2). Again, we assume that separate estimates of $N_f^{p,s}$ and $f_{on}$ are available.

#### 2.3.3 Model selection: heterogeneous versus homogeneous response

Finally, we determine whether it is possible to identify the heterogeneity of stress responses from mutant count data alone. In this case, we suppose we do not have separate experimental data showing heterogeneity and, therefore, do not have an estimate of $f_{on}$. For this purpose, we simulate fluctuation assays under the heterogeneous-response model for varying $\gamma_{on}$. Then, we infer the respective model parameters assuming (a) the heterogeneous-response model with $r_{on} = 0$ and $f_{on}$ as an inference parameter; and (b) the homogeneous-response model with either $\rho^{p,s} = 1$ as a fixed value, $\rho^p = \rho^s$ as a joint inference parameter, or $\rho^p$ and $\rho^s$ as two separate inference parameters (Results section 3.4, Appendix section A.6). To decide which model fits the observed mutant count data best, we use the Akaike information criterion (AIC), defined as

$$\mathrm{AIC} = 2(k - \mathcal{L}) \tag{23}$$

where $k$ is the number of inferred parameters of the model. We select the model with the lowest AIC. However, if the difference in AIC between the heterogeneous- and the best-fitting homogeneous-response model is within $\pm 2$, we say that neither of the models can be selected because their AICs are comparable.

We also evaluate the models' performances in estimating the relative increase in population mean mutation

| Symbol | Condition | Definition | Numerical value |
|---|---|---|---|
| $\gamma^p_{off}$ | Permissive | Division rate response-*off* cells | $1\,h^{-1}$ |
| $\rho^p_{off}$ | Permissive | Differential mutant fitness response-*off* cells | 1 |
| $\nu^p_{off}$ | Permissive | Mutations per division response-*off* cells | $10^{-8}\,h^{-1}$ |
| $\gamma^s_{off}$ | Stressful | Division rate response-*off* cells | $1\,h^{-1}$ |
| $\rho^s_{off}$ | Stressful | Differential mutant fitness response-*off* cells | 1 |
| $\nu^s_{off}$ | Stressful | Mutations per division response-*off* cells | $10^{-8}\,h^{-1}$ |
| $\delta_{off}$ | Stressful | Death rate response-*off* cells | $0\,h^{-1}$ |
| $\gamma_{on}$ | Stressful | Division rate response-*on* cells | $0\,h^{-1}$ |
| $\rho_{on}$ | Stressful | Differential mutant fitness response-*on* cells | 1 |
| $\nu_{on}$ | Stressful | Mutations per division response-*on* cells | $10^{-6}\,h^{-1}$ |
| $\delta_{on}$ | Stressful | Death rate response-*on* cells | $0\,h^{-1}$ |
| $\alpha$ | Stressful | Switching rate | $0.05\,h^{-1}$ |

Table 1: Parameters used in the simulations unless explicitly specified otherwise. The switching rate for the SOS response in *E. coli* has been estimated using single-cell imaging [19]. The mutation rate is based on rifampicin resistance, a selective marker commonly used in fluctuation assays. In *E. coli*, the number of point mutations conferring rifampicin resistance has been estimated to be 79 [13] and the per-nucleotide mutation rate between $0.2 \cdot 10^{-10}$ and $5 \cdot 10^{-10}$ [20].

rate. In case (a), the population mean mutation rate increase is calculated as $\Delta\bar{\mu} = \frac{(1-f_{on})\mu_{off}+f_{on}\mu_{on}}{\mu_{off}}$ and in (b) as $\Delta\bar{\mu} = \frac{\mu^s}{\mu^p}$.

The complete documentation of the computational analysis can be found in the **README** at `https://github.com/LucyL-J/Quantifying-SIM`.

## 3 Results

We aim to estimate the increase in mutation rate associated with the induction of the stress response when this response is heterogeneously expressed across the bacterial population. In particular, we consider cases in which the population can be divided into two discrete subpopulations; one with a low expression level (response *off*) and the other with a high expression level (response *on*). The key principle of our method is to jointly infer the mutation rates from mutant count data obtained from a pair of fluctuation assays under permissive and stressful conditions. For the latter, we need to disentangle the contributions from the response-*off* and response-*on* subpopulations. The success of this method relies on the changing shape of the mutant count distribution under stress, which occurs if there is a highly mutating but slowly dividing response-*on* subpopulation.

To evaluate the performance of our method we use simulated mutant count data so that we can compare the estimated parameters with the true underlying ones. First, we explore how the accuracy and precision of our method depend on the model parameters by simulating and inferring under the same model. Then, we test the robustness of our method to model deviations by simulating under a more complex model than used in the inference. Finally, we determine under what conditions the heterogeneous-response model can be distinguished from the homogeneous-response model, which is assumed in currently available methods, by inferring under both models and comparing how well they fit simulated data.

In all simulations, we set the initial population size to $10^4$, the initial fraction of the response-*on* subpopulation to $\frac{\alpha}{\gamma^s_{off}-\delta_{off}}$ which is a lower bound for the equilibrium fraction of the response-*on* subpopulation, and the number of parallel cultures to $50$. Moreover, we consider the duration of the growth phase such that the expected number of mutations equals one. This way, the resulting number of resistant mutant colonies on each selective plate is usually within an experimentally countable range of zero to a couple hundred (Supplementary Material equations 27, 28). Other parameters used in the simulations are summarised in Table 1.

### 3.1 Estimating the mutation-rate increase

First, we evaluate our method's performance in the best-case scenario; that is, we simulate and infer under the same model, beginning with the simplest model of heterogeneous stress responses without cell death, with mutant fitness equal to non-mutant fitness, and with zero division rate of response-*on* cells. We simulate $R = 100$ pairs of fluctuation assays under permissive/stressful conditions using a parameter range of $\nu_{on} \in [10^{-5},\ 10^{-8}]$ and $\alpha \in [0.001,\ 0.1]$ in $h^{-1}$.

For the inference, we consider the same model as used in the simulations with the only exception that, in the inference, we neglect switching *on* the stress response in mutants and initial population size effects, see Model & Methods. Moreover, we assume that the final population sizes $N_f^{p,s}$ and the fraction of the response-*on* subpopulation $f_{on}$ are known from separate experimental measurements.

For each of the $R = 100$ sets of mutant count data, we infer the mutation rates $\mu_{off}$ and $\mu_{on}$ and calculate the mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$. To quantify the accuracy of our method, we calculate the relative error of the median of our estimated mutation-rate increases (equation 21). Additionally, we use the coefficient of variation (equation 22) across the estimates to quantify our method's precision.

Comparing the estimated with the true mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$, we find that the accuracy and precision improve with increasing mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$ and relative switching rate $\tilde{\alpha}$ (Figure 2B, 2C). For example, for a relative switching rate of $\tilde{\alpha} = 0.01$, as $\frac{\mu_{on}}{\mu_{off}} \geq 100$, $75\%$ of estimates lie within $2-$fold of the true mutation-rate increase and the estimation is unbiased (Figure 2A). For smaller $\frac{\mu_{on}}{\mu_{off}}$, on the other hand, the variation in the estimates becomes large. Nonetheless, if the mutation-rate increase is estimated to be $> 100$, the true increase is very likely to be $> 25$.

The composite parameter $H := \frac{\mu_{on}}{\mu_{off}} \frac{f_{on}}{1-f_{on}}$, which is a measure for the mutation-rate heterogeneity in the population, determines our method's performance. This can be seen from the mutant count distribution derived in Model & Methods. For the simplest model considered here, the probability-generating function of the mutant count distribution under stress, obtained from equations 15 and 17 is given by

$$G^s_{\text{het}}(z) = \exp\left[\overbrace{\mu_{off} N_f^s \frac{(1-z)\log(1-z)}{z}}^{\text{Response-}off\text{ contribution}} + \overbrace{\mu_{on} \frac{f_{on}}{1-f_{on}} N_f^s (z-1)}^{\text{Response-}on\text{ contribution}}\right]. \tag{24}$$

The mutation-rate heterogeneity determines the shape of the distribution as for $H \gg 1$, the contribution of the response-*on* subpopulation is dominating. In contrast, for $H \ll 1$, the response-*on* subpopulation contributes very little to the total mutant count.
Overall, in the best-case scenario and for the parameter regime considered here, a mutation-rate heterogeneity of $H \sim \mathcal{O}(1)$ is sufficient for an accurate and precise estimate of the mutation-rate increase.

### 3.2 The impact of cell death

Our model takes heterogeneity in the stress response into account but neglects other potential consequences of the stress, for example, cell death. Therefore, we test whether the occurrence of cell death, if neglected in the inference model, causes a biased estimate as shown previously [10]. For this purpose, we simulate fluctuation assays under an extended model of heterogeneous stress responses with cell death. We consider the cases that either (a) only response-*off* cells are affected by cell death, (b) only response-*on* cells are affected, or (c) all cells are affected equally, using parameter ranges of $\delta_{off}, \delta_{on} \in [0.0,\ 0.5]$ in $h^{-1}$.

As before, we repeat the simulations $R = 100$ times and assume that the final population sizes and the fraction of the response-*on* subpopulation are known. From the resulting mutant count data, we infer the mutation rates $\mu_{off}$ and $\mu_{on}$ from which, in turn, we calculate the estimated mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$. We compare the estimated value with the true value to determine any biases caused by neglecting cell death in the inference.

Interestingly, we find that any biases depend on which subpopulation is affected by cell death. If only response-*off* cells are affected by cell death, the mutation-rate increase is overestimated for large death rates (Figure 3A). On the other hand, if only response-*on* cells are affected, the mutation-rate increase is underestimated for large death rates (Figure 3B). If the total population is affected by cell death, the estimation of the mutation-rate increase remains unbiased; however, the variation in the estimates increases for large death rates (Figure 3C). From the contribution of the response-*on*

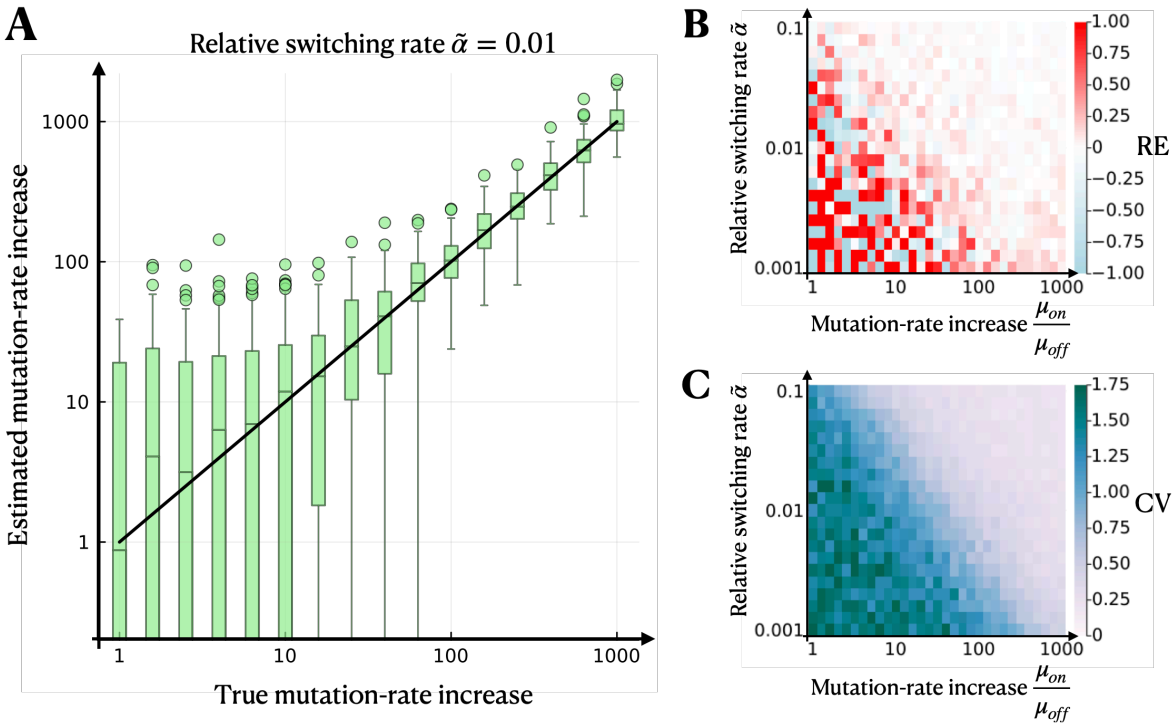


Figure 2: Quantification of the accuracy and precision of our inference method using simulated mutant count data. We simulate $R = 100$ pairs of fluctuation assays using the simplest model of heterogeneous stress responses (without cell death or differential mutant fitness, and with zero division rate of response-*on* cells) and infer the mutation rates $\mu_{on}$ and $\mu_{off}$ assuming the same model in the inference. **A** Estimated compared to true mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$ for a range of values of $\frac{\mu_{on}}{\mu_{off}}$ and a relative switching rate of $\tilde{\alpha} = 0.01$. **B** Relative error of estimated compared to true mutation-rate increase for a range of $\frac{\mu_{on}}{\mu_{off}}$ and $\tilde{\alpha}$ with over-/underestimation indicated in red/blue and **C** coefficient of variation across estimates. The parameter ranges used in the simulations are $\nu_{on} \in [10^{-5},\ 10^{-8}]$ and $\alpha \in [0.001,\ 0.1]$ in $h^{-1}$.

subpopulation to the mutant count given in equation 16, it can be seen that the effects of cell death in response-*off* and response-*on* cells partly cancel each other out. This result also holds for other parameter values (Appendix Figure 7).

We test another biologically realistic model deviation: the fitness of mutants differs from non-mutants during the growth phase (Appendix Figure 8). We find that neglecting this effect in the inference leads to a slight underestimation of the mutation-rate increase when mutants have a fitness cost. On the other hand, if mutations bring a fitness advantage (for example, due to cross-resistance), the mutation-rate increase is slightly overestimated.

### 3.3 Estimation when response-*on* cells have a non-zero division rate

So far, we considered response-*on* cells not to divide at all, motivated by the SOS response. However, the division rate of response-*on* cells might be non-zero, especially if cells are exposed to a very low level of DNA damage (in the case of SOS) or for other stress responses. By default, our method sets the relative fitness of response-*on* cells to zero, $r_{on} = 0$, but it also allows us to estimate it as an inference parameter. In the following, we evaluate the performance of our method in inferring $r_{on} > 0$ and how this impacts the estimated mutation-rate increase. We simulate under the heterogeneous-response model with a non-zero division rate of response-*on* cells, considering a parameter range of $\gamma_{on} \in [0,\ 1]$ in $h^{-1}$.

Again, we repeat the simulations $R = 100$ times and assume that the final population sizes and the fraction of the response-*on* subpopulation are known. In the inference, we consider two different cases. In one case, we infer both the mutation rates $\mu_{off}$ and $\mu_{on}$, and the relative fitness of response-*on* cells $r_{on}$. Alternatively, we infer only $\mu_{off}$ and $\mu_{on}$ while setting $r_{on} = 0$. We calculate the estimated mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$ in both cases and compare it to the

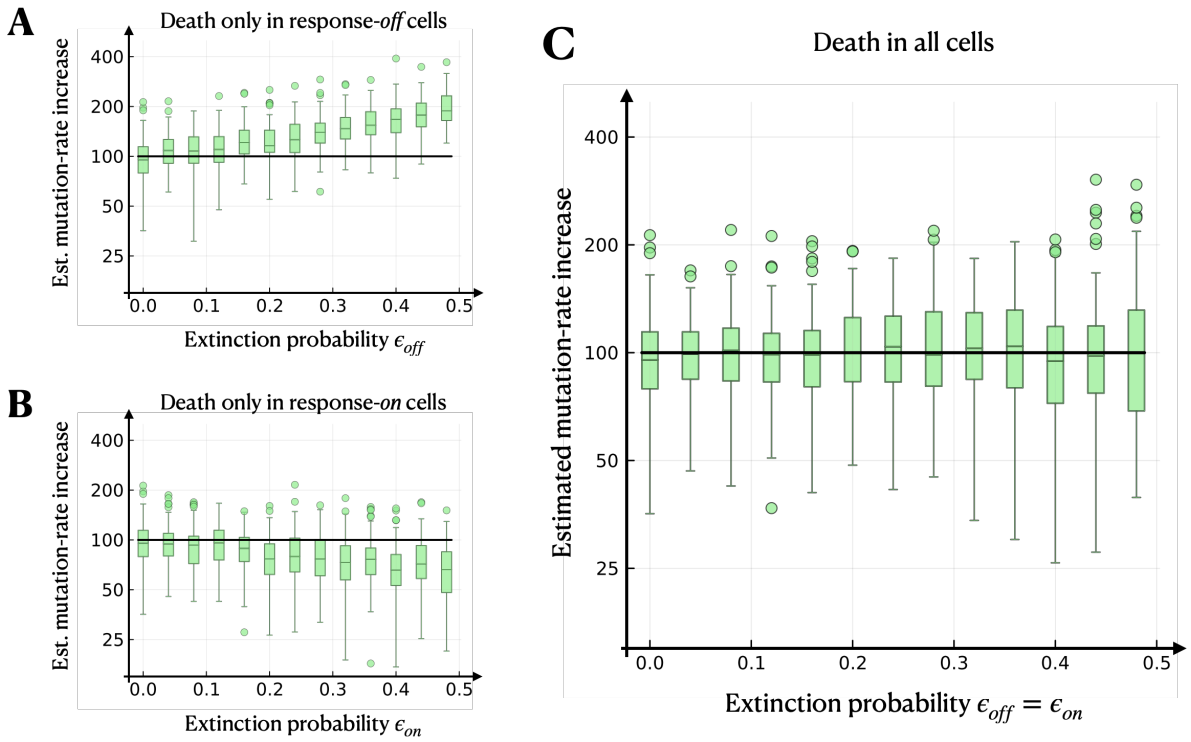


Figure 3: Impact of cell death on the estimation of the mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$. We simulate $R = 100$ pairs of fluctuation assays using the heterogeneous-response model (without differential mutant fitness and with zero division rate of response-*on* cells), but with cell death. However, we neglect cell death to infer $\mu_{off}$ and $\mu_{on}$. The black solid lines indicate the true mutation-rate increase used in the simulations. **A** Estimated mutation-rate increase when only response-*off* cells are affected by cell death, **B** when only response-*on* cells are affected by cell death, and **C** when all cells are affected by cell death equally. The parameter ranges used in the simulations are $\delta_{off}, \delta_{on} \in [0.0,\ 0.5]$ in $h^{-1}$.

true value. In the first case, we also compare the estimated with the true relative fitness of response-*on* cells.

We find that the estimation of the mutation-rate increase remains accurate for small to moderate relative fitness $r_{on}$. For large $r_{on} \to 1$, on the other hand, the mutation-rate increase is underestimated, yet more accurate and precise if $r_{on}$ is also inferred (Figure 4A). However, the estimate of $r_{on}$ itself is not very precise, and is underestimated for large $r_{on}$ (Figure 4B).
We also evaluate the performance when $r_{on}$ is set to the true value in the inference. Interestingly, this increases the accuracy and precision of the estimate of $\frac{\mu_{on}}{\mu_{off}}$ only slightly compared to when $r_{on}$ is inferred (Appendix Figure 9). The reason for this lies in the approximation (equation 5) made to derive the mutant count distribution, which is no longer valid for large $r_{on}$.

### 3.4 Model selection between heterogeneous and homogeneous response

In many cases, it may not be known *a priori* whether the stress response is heterogeneously expressed across the population or whether, in contrast, all cells respond similarly. We want to determine whether distinguishing these two models is possible using mutant count data from fluctuation assays. To do so, we simulate fluctuation assays under the heterogeneous-response model for a range of relative fitness of response-*on* cells $r_{on}$ (but without cell death or differential mutant fitness). In the inference, we consider both the heterogeneous-response model (setting $r_{on} = 0$) and the homogeneous-response model without or with differential mutant fitness, and compare how well they fit the data. We use the same simulation data as in the previous section ($R = 100$ simulations, parameter range $\gamma_{on} \in [0,\ 1]$ in $h^{-1}$).

Now, we assume that the final population sizes $N_f^{p,s}$ are known but not the fraction of the response-*on* subpopulation $f_{on}$. Therefore, when using the heterogeneous-response model in the inference, we infer $f_{on}$ in addition to the mutation rates of response-*off* and response-*on* cells, $\mu_{off}$ and $\mu_{on}$. From this, we calculate the increase in population

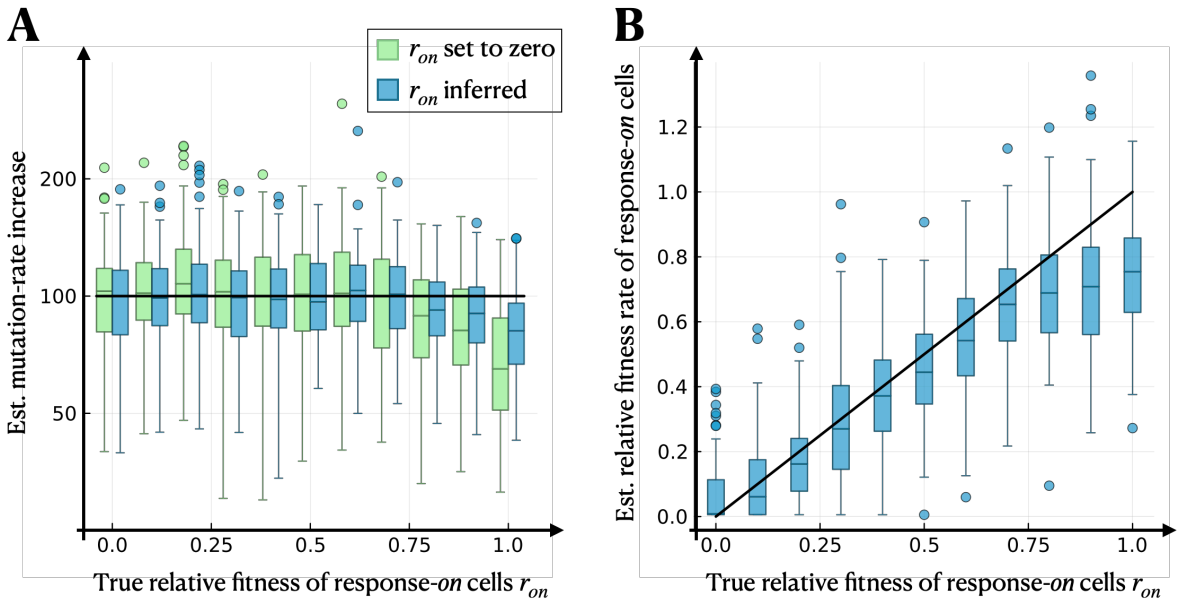


Figure 4: Estimation when response-*on* cells have a non-zero division rate. We simulate $R = 100$ pairs of fluctuation assays using the heterogeneous-response model (without cell death or differential mutant fitness) with $r_{on} \geq 0$ being the relative fitness of response-*on* cells compared to response-*off* cells. We consider two cases for the inference: (i) setting $r_{on}$ to zero and only inferring $\mu_{off}$ and $\mu_{on}$, and (ii) inferring $r_{on}$ in addition to the two mutation rates. **A** Estimated mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$ for both cases and a range of relative fitness of response-*on* cells. The solid black line indicates the true value of $\frac{\mu_{on}}{\mu_{off}}$. **B** Estimated compared to true relative fitness of response-*on* cells in inference case (ii). The parameter range used in the simulations is $\gamma_{on} \in [0,\ 1]$ in $h^{-1}$.

mean mutation rate $\Delta\bar{\mu} = \frac{(1-f_{on})\mu_{off}+f_{on}\mu_{on}}{\mu_{off}}$. Having $f_{on}$ as an inference parameter prevents us from having $r_{on}$ as yet another inference parameter because this renders the likelihood surface too complex for numerically efficient estimation. When using the homogeneous-response model, on the other hand, we infer the population-wide mutation rates under permissive and stressful conditions ($\mu^p$ and $\mu^s$, respectively), and, optionally, the differential mutant fitnesses ($\rho^p$ and $\rho^s$), and calculate the resulting increase in population-wide mutation rate $\frac{\mu^s}{\mu^p}$.

We compare the estimated population mean mutation rate from each inference model with the true value (Figure 5A). Moreover, we perform model selection using the AIC, which considers how well the model reproduces the data and penalises the number of model parameters. We compare (i) the heterogeneous-response model (inference parameters: $\mu_{off}$, $\mu_{on}$ and $f_{on}$) with (ii) the homogeneous-response model without differential mutant fitness (inference parameters: $\mu^p$ and $\mu^s$) and (iii) the homogeneous-response model with differential mutant fitness (inference parameters: $\mu^p$, $\mu^s$, $\rho^p$ and $\rho^s$).

We find that using the heterogeneous-response model, the estimation of the increase in population mean mutation rate is unbiased over the range of relative fitness of response-*on* cells, $r_{on}$ (Figure 5A). However, the estimates show greater variation for larger values of $r_{on}$. In contrast, inference using the homogeneous-response model without differential mutant fitness estimates a slightly smaller value for the increase in population mean mutation rate. This estimate improves when allowing for a differential mutant fitness, being as accurate as the estimation using the heterogeneous-response model. In this case, the differential mutant fitness under the permissive condition is estimated to be around $\rho^p \approx 1$, the true value (Figure 5B). For the stressful condition, on the other hand, a mutant cost is inferred, which is increasingly severe for smaller relative fitness of response-*on* cells (Figure 5B).

Comparing model fit, we find that the heterogeneous response model is selected only in a minority of cases (Figure 5C). The best case is when the relative fitness of response-*on* cells is small, where the heterogeneous-response model is selected for slightly less than 50% of the simulations. The homogeneous-response model with differential mutant fitness is selected in ≈10% of simulations, while no model is clearly preferred for the remaining ≈40%. As $r_{on}$ increases, the heterogeneous-response model gives a poorer fit to the mutant count data because we assume $r_{on} = 0$

in the inference. For intermediate values of $r_{on}$, the homogeneous-response model with differential mutant fitness is selected for the majority of the simulations. Eventually, for large $r_{on}$, the homogeneous-response model without differential mutant fitness is selected in most cases, as it considers two fewer inference parameters than the model with differential mutant fitness.

We also consider the homogeneous-response model with differential mutant fitness values constrained to be equal under both stressful and permissive conditions (Appendix Figure 10). The estimate of $\Delta\bar{\mu}$ is less accurate using this model. Moreover, for small $r_{on}$, the heterogeneous-response model is preferred for most simulations.

## 4 Discussion

Since its introduction 80 years ago, the standard model behind the fluctuation assay has been extended numerous times to overcome limitations and make it more biologically realistic. Extensions particularly relevant for quantifying stress-induced mutagenesis include taking into account cell death [10, 36] and differential mutant fitness [42]. The latter can become important when resistance allowing growth on the selective plates also confers an advantage to the stressor, for example, due to cross-resistance; or when mutants carry a fitness cost.

In this study, we addressed a so-far overlooked limitation of fluctuation analysis: heterogeneity in the expression of stress responses, which has recently been demonstrated by single-cell studies. Our population dynamic model considers response-*off* and response-*on* subpopulations, which allows us to estimate the relative increase in mutation rate associated with the induction of the stress response, $\frac{\mu_{on}}{\mu_{off}}$.

We tested our estimation method with simulated mutant count data, which confirmed accurate and precise estimation of $\frac{\mu_{on}}{\mu_{off}}$ for sufficiently large mutation-rate heterogeneity $H = \frac{f_{on}}{1-f_{on}} \frac{\mu_{on}}{\mu_{off}}$ (Results section 3.1). $H$ depends on the mutation-rate increase itself and the fraction of the response-*on* subpopulation $f_{on}$. While $\frac{\mu_{on}}{\mu_{off}}$ is inherent to the stress response, $f_{on}$ could potentially be increased through experimental design. For example, increasing the antibiotic concentration has been shown to increase the rate of switching on the stress responses and, with it, the fraction of the response-*on* subpopulation [19]. Our results suggest that mutation rate estimates would therefore be more accurate at higher antibiotic doses. Increasing the level of stress could, however, also increase cell death. We neglect cell death in our inference, but we showed that our method is robust to this model deviation up to moderate death rates when cell death affects response-*off* and response-*on* subpopulations equally (Results section 3.2).

Our approach to quantifying stress-induced mutagenesis assumes that the expression of the stress response is bimodal and can reasonably be modelled as either switched *off* or *on*. This expression pattern has been observed for the SOS response [19] but is somewhat of a simplification. For some other stress responses, it might be more appropriate to model the expression as a continuous distribution. We expect, however, that this increase in model complexity would make parameter inference more challenging. We also neglect response-*on* cells switching the response *off* so long as the stress remains present during the comparably short growth phase of a fluctuation assay. This assumption implies that the model cannot capture pulse-like mutagenesis as observed under DNA alkylation damage [34] and oxidative stress [26].

To be able to derive an analytical expression for the mutant count distribution, we make a series of approximations, the most important one being that cells with stress response switched *on* have a net growth rate ($\gamma_{on} - \delta_{on}$) much lower than that of *off* cells ($\gamma^s_{off} - \delta_{off} - \alpha$). For the SOS response, this approximation is valid, as induction of the response inhibits cell division. However, it might be violated for other stress responses, particularly if they protect cells from death, resulting in $\delta_{on} < \delta_{off}$. In this case, our approximation would no longer be valid and therefore, parameter estimation using our method would be less accurate.

Provided response-*on* cells have low relative fitness compared to response-*off* cells ($r_{on} = \frac{\gamma_{on}-\delta_{on}}{\gamma^s_{off}-\delta_{on}} \ll 1$), we found that it is sometimes possible to detect a signal of cell heterogeneity. That is, the heterogeneous-response model is selected by AIC over standard (homogeneous-response) models when fitted to simulated data but is only clearly preferred in about half of the experiments (Results section 3.4). However, as $r_{on}$ increases, the above-described approximation breaks down and the (approximate) heterogeneous model provides a poorer fit, even though data was simulated under the (exact) heterogeneous model. Already at moderate values of $r_{on}$, a homogeneous-response model with differential mutant fitness under permissive and stressful conditions ($\rho^p$ and $\rho^s$, respectively) fits the data comparably well or better.

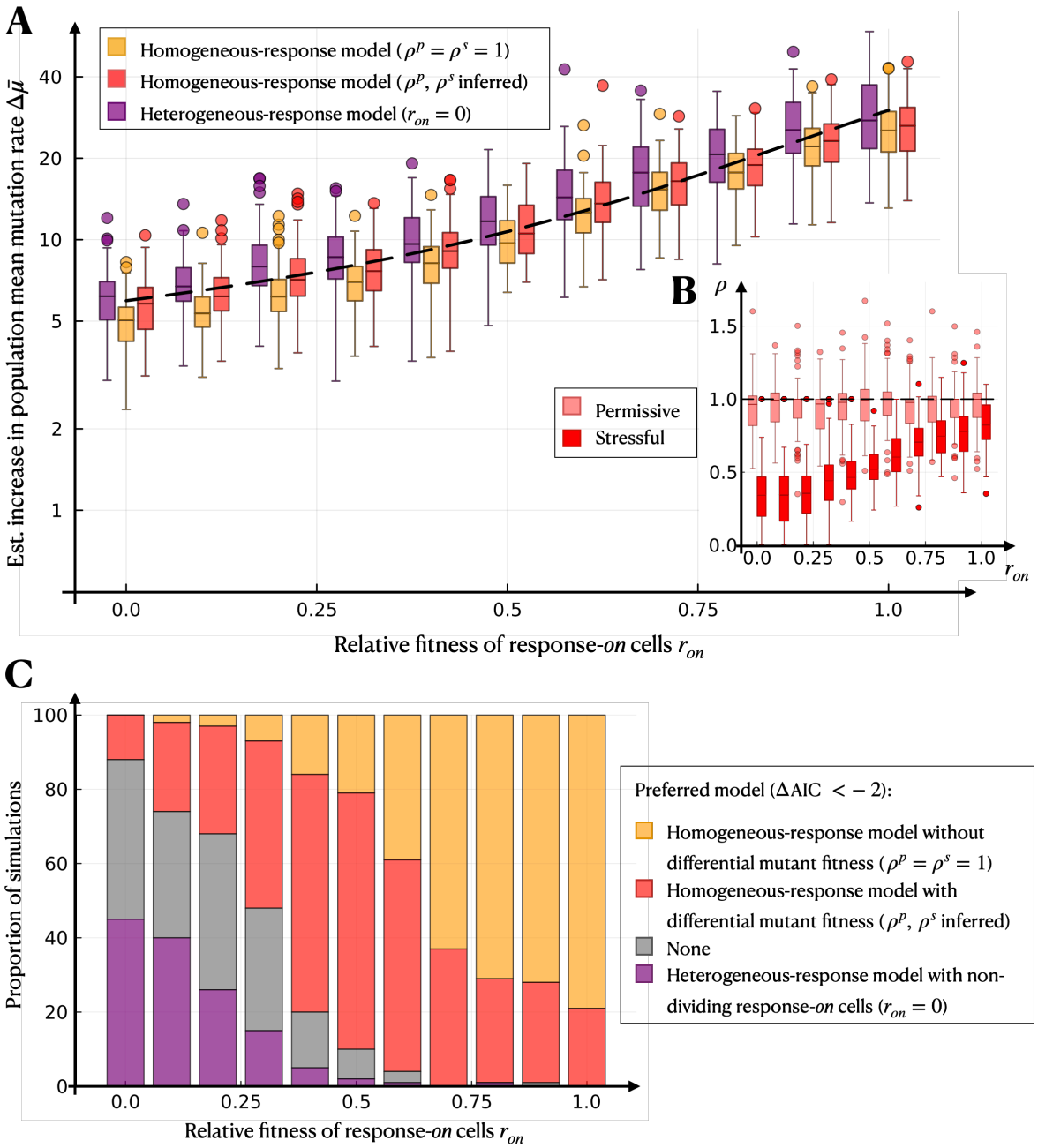


Figure 5: Model selection between the heterogeneous-response and the homogeneous-response model. We simulate $R = 100$ pairs of fluctuation assays under the heterogeneous-response model (without cell death or differential mutant fitness) for a range of relative fitness of response-*on* cells, $r_{on}$. In the inference, we use (i) the heterogeneous-response model (shown in purple) to infer the mutation rates $\mu_{off}$ and $\mu_{on}$ plus the fraction of the response-*on* subpopulation $f_{on}$ (but set $r_{on} = 0$), (ii) the homogeneous-response model without differential mutant fitness ($\rho^p = \rho^s = 1$) (shown in orange) and (iii) the homogeneous-response model with differential mutant fitness (shown in red) to infer the population-wide mutation rates under permissive and stressful conditions, $\mu^p$ and $\mu^s$, and, for (iii), additionally $\rho^p$ and $\rho^s$. **A** Estimated increase in population mean mutation rate (given by $\frac{(1-f_{on})\mu_{off}+f_{on}\mu_{on}}{\mu_{off}}$ in the heterogeneous model and $\frac{\mu^s}{\mu^p}$ in the homogeneous model). **B** Estimated differential mutant fitness under permissive and stressful conditions in inference case (iii). The black dashed lines indicate the true values. **C** Model selection using the AIC. The parameter range used in the simulations is $\gamma_{on} \in [0,\ 1]$ in $h^{-1}$.

Taken together, these results indicate that it will likely be difficult to infer heterogeneity in the stress response from fluctuation assay data alone; rather, alternative biological mechanisms (specifically, a mutant fitness cost) can explain the data equally well. Therefore, these data should be complemented with additional experiments to support or rule out alternative models. For example, mutants arising in the fluctuation assay can be isolated to measure their fitness and check whether fitness values estimated from the homogeneous model fit ($\rho^p$ and $\rho^s$) are reasonable, or to further constrain the model fitting. If there is reason to suspect heterogeneity in the stress response, experimenters can test this directly by engineering fluorescent reporters into the bacterial strain of interest and measuring the response on a single-cell level, e.g. by flow cytometry or microscopy. These experiments would provide an independent estimate of the fraction of response-*on* subpopulation to further constrain the heterogeneous response model. In reality, multiple factors causing deviation from the standard fluctuation assay model (e.g. heterogeneous responses, differential mutant fitness, and cell death) are likely to be operating simultaneously. Since it is not feasible to reliably estimate a large number of parameters from fluctuation assay data alone, separate experiments become important to decide which deviation(s) are most relevant to incorporate into the fluctuation analysis or to constrain the number of fitted parameters.

Interestingly, both the heterogeneous-response model and homogeneous-response model with differential mutant fitness perform equally well in estimating the increase in population mean mutation rate (Results section 3.4). This implies that previous studies of stress-induced mutagenesis using fluctuation assays likely estimated the increase in population mean mutation rate accurately. However, they may not capture the true magnitude of the increase in mutagenesis if only a subpopulation induces the stress response. With our new method, it is possible to estimate the mutation-rate increase explicitly associated with the induction of the stress response. Estimating not only the increase in population mean but also variation in mutation rate could be important for parameterising evolutionary models, for instance, predictions of antibiotic resistance evolution. Theoretical modelling suggests that single-locus adaptation can be accurately captured by the population mean mutation rate, but cell-to-cell variation can speed up multi-locus adaptation [1].

In summary, we have presented and validated a new method for inferring stress-induced increases in mutation rate from fluctuation assays. In future work, we aim to incorporate this method into user-friendly tools for application to experimental data, similar to existing R packages [28, 41] and web tools [17, 15, 25] for fluctuation analysis. Future work could also adapt our method to other heterogeneously-expressed stress responses that do not meet the assumptions we modelled on the SOS response.

## Author contributions

LLJ, MEK and HKA conceptualized the study. LLJ derived the mathematical results, performed the simulations and drafted the paper under the advice and supervision of MEK and HKA. All authors critically revised the manuscript and gave approval of the final version for submission.

## Acknowledgements

The authors are grateful for helpful feedback on the mathematical results received from Tibor Antal and for the discussion and inspiration provided by the Alexander and El Karoui labs. This work was supported by the UKRI Biotechnology and Biological Sciences Research Council (BBSRC) grant number BB/T00875X/1 and a University of Edinburgh Principal's Career Development PhD Scholarship to LLJ, a Wellcome Trust Investigator Award 205008/Z/16/Z to MEK, and a Royal Society University Research Fellowship URF/R1/191269 to HKA.

## Conflict of interest

The authors declare no conflict of interest.

## Data Accessibility

The complete annotated documentation of the computational analyses of this study will be archived upon publication of the paper and is currently deposited at `https://github.com/LucyL-J/Quantifying-SIM`.

# Appendix A

## A.1 Response-*on* non-mutants can be treated deterministically

In Model & Methods section 2.1, we derive the population sizes of the response-*off* and response-*on* non-mutants (equations 3 and 4) assuming that the response-*on* subpopulation can be treated deterministically. However, for small initial population sizes $n_{on}(0)$, this assumption might not hold. Therefore, we test its validity using stochastic simulations; we simulate switching *on* of the response as a time-inhomogeneous Poisson process, the growth dynamics of the response-*on* subpopulation as a stochastic birth-death branching process until it reaches a size of $N = 10^4$ and according to equation 4 afterwards. Then, we compare the resulting population size with equation 4 at an early/late time point: $t_1$ the expected time of the first mutation (with an exponential waiting time) in the response-*off* subpopulation given by

$$t_1 = \frac{\Gamma\left(0, \frac{\mu_{off} n_{off}(0)}{\gamma^s_{off} - \delta_{off} - \alpha}\right)}{\gamma^s_{off} - \delta_{off} - \alpha}, \tag{25}$$

and the time when the response-*off* subpopulation reaches a size of $N = 10^9$ given by

$$t_N = \frac{\log\left(\frac{10^9}{n_{off}(0)}\right)}{\gamma^s_{off} - \delta_{off} - \alpha}. \tag{26}$$

For two initial population sizes of the response-*on* subpopulation, (i) $n_{on}(0) = 0 \cdot n_{off}(0)$ and (ii) $n_{on}(0) = \frac{\alpha}{\gamma^s_{off} - \delta_{off}} \cdot n_{off}(0)$ which is a lower bound for the equilibrium fraction of the response-*on* subpopulation, we calculate the error relative to the median of $R = 100$ simulations and the coefficient of variation across the simulations (Figure 6). We find that both relative error and coefficient of variation are larger at the earlier than the later time point, and that they increase with decreasing relative switching rate $\tilde{\alpha}$ and increasing relative fitness of response-*on* cells $r_{on}$. In general, equation 4 tends to slightly overestimate the size of the response-*on* subpopulation (indicated in red), but the error is sufficiently small to justify our approximation.

## A.2 Calculation of the duration of the growth phase

In all simulations to test our inference method, we set the duration $t_f$ of the growth phase such that the expected number of *mutations*, $m$ (not *mutants*) equals one.

$$\text{Permissive conditions: } \mathbb{E}[m] = \int_0^{t_f} \nu_{off} n_{off}(t)\, \mathrm{d}t \stackrel{!}{=} 1, \quad (\text{with } \alpha = \delta_{off} = 0) \tag{27}$$

$$\text{Stressful conditions: } \mathbb{E}[m] = \int_0^{t_f} \nu_{off} n_{off}(t) + \nu_{on} n_{on}(t)\, \mathrm{d}t \stackrel{!}{=} 1 \tag{28}$$

This way, the resulting number of resistant mutant colonies on each selective plate is comparable across the considered parameter ranges and usually within an experimentally countable range of zero to a couple hundred. For example, the duration of simulated fluctuation assays under permissive conditions is set to $t_f \approx 9.21\, h$.

## A.3 The impact of cell death

In Results section 3.2, we show that cell death impacts the estimation of the mutation-rate increase depending on which subpopulation is affected by death. However, when all cells are affected equally, the variation in the estimates increases but it remains largely unbiased. This result also holds for different values of $\tilde{\alpha}$ (Figure 7). A possible explanation for this result is that death in response-*off* cells causes an overestimation of the per-division mutation rate $\mu_{on}$ because, as pointed out in [10], in the presence of cell death, the number of cell divisions to reach the final population sizes is underestimated. At the same time, death in response-*on* cells can lead to the extinction of response-*on* mutant lineages, causing an underestimation of $\mu_{on}$. These two effects cancel out, leading to an unbiased estimate overall. The estimation of $\mu_{off}$ is not noticeably impacted by cell death due to the joint inference under stressful and permissive conditions, as for the latter, there is no cell death.

## A.4 The impact of differential mutant fitness

In addition to cell death, we test the robustness of our method to the presence of a differential mutant fitness which is neglected in the inference. For this purpose, we simulate fluctuation assays under an extended model of heterogeneous

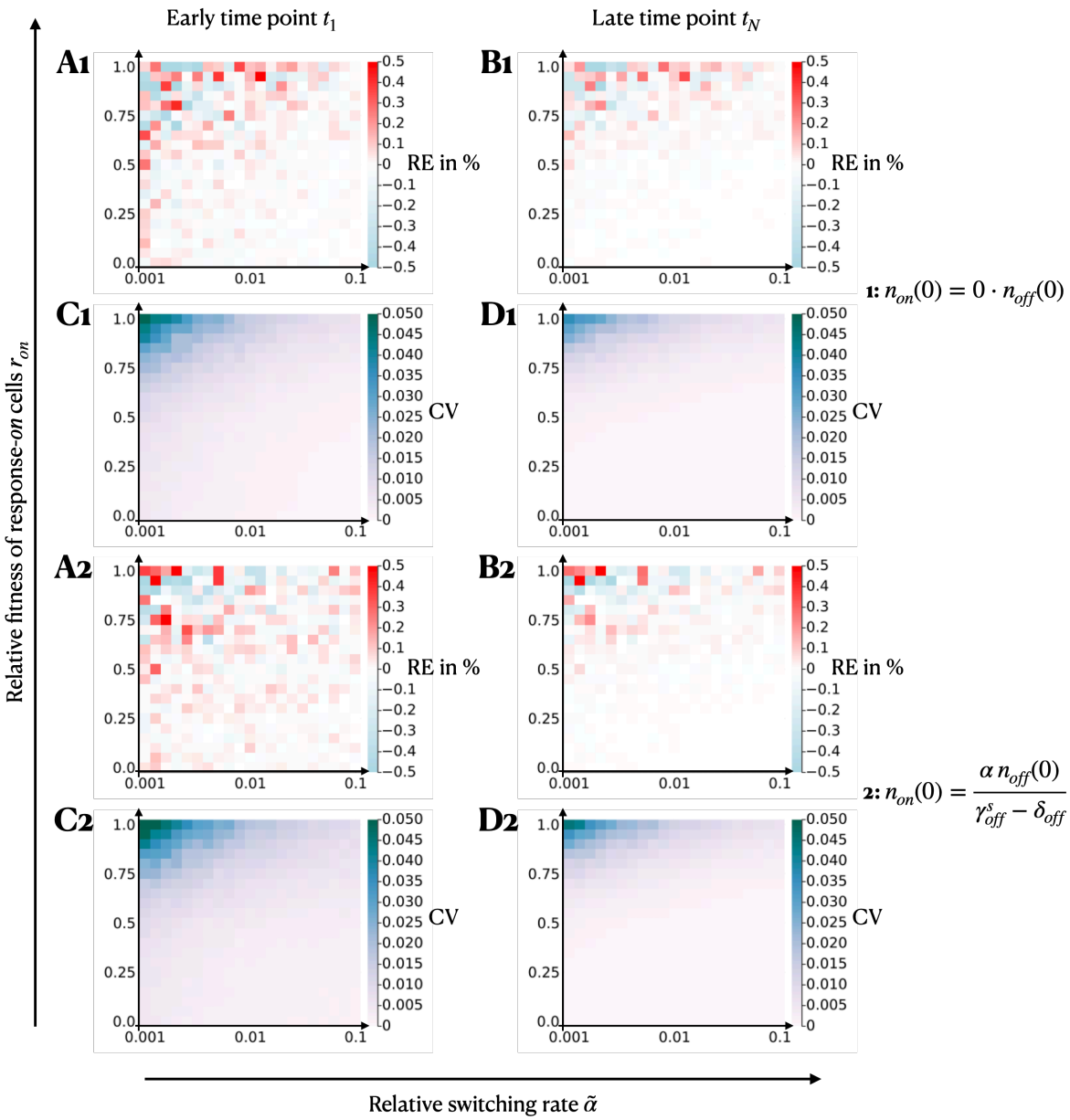


Figure 6: Deterministic compared to stochastic dynamics of response-*on* non-mutants for initial population sizes of the response-*on* subpopulation of (**1**) $n_{on}(0) = 0 \cdot n_{off}(0)$ and (**2**) $n_{on}(0) = \frac{\alpha}{\gamma^s_{off} - \delta_{off}} \cdot n_{off}(0)$. We simulate switching *on* of the response as a time-inhomogeneous Poisson process and the dynamics of response-*on* non-mutants as stochastic birth-death processes ($R = 100$ simulation runs). Then, we compare the resulting population sizes with equation 4 at an early/late time point. **A** Relative error of the median in % at $t_1$, **B** relative error of the median in % at $t_N$, **C** coefficient of variation across the simulations at $t_1$ and **D** coefficient of variation across the simulations at $t_N$. The parameters used in the simulations are $N^s_0 = 10^4$, $\gamma^s_{off} = 1$, $\delta_{off} = \delta_{on} = 0$, $\alpha \in [0.001,\ 0.1]$, $\gamma_{on} \in [0.0,\ 1.0]$.

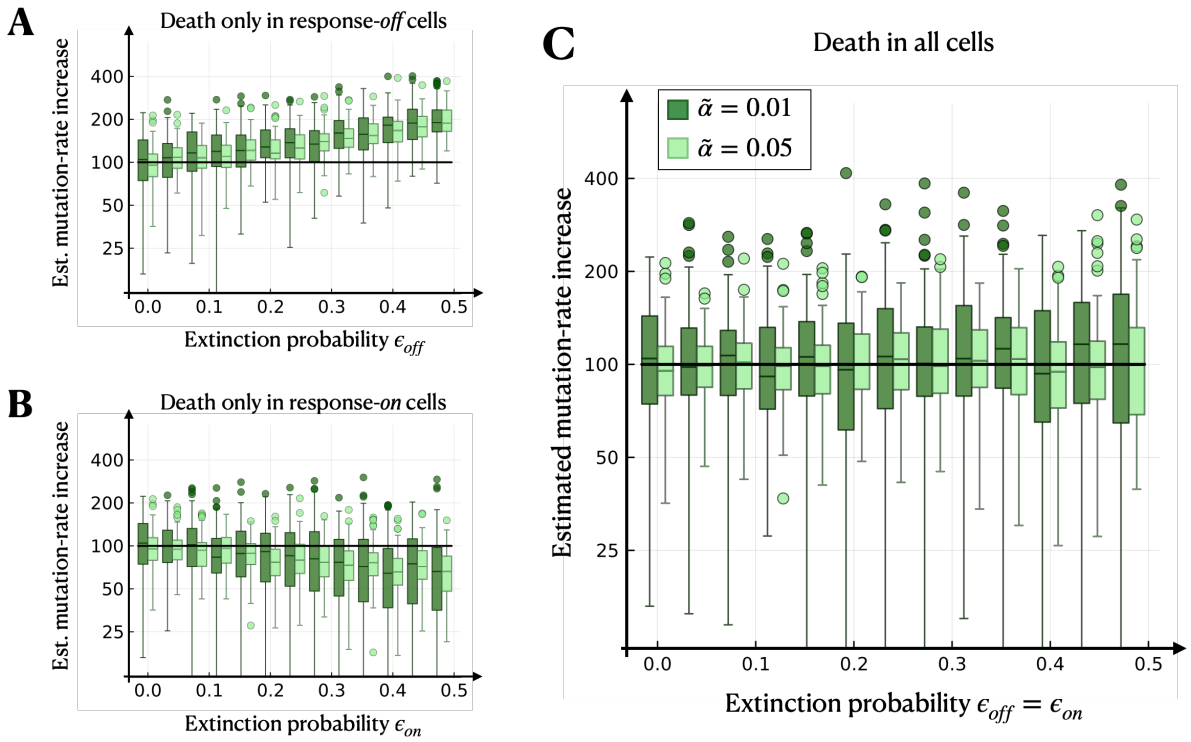


Figure 7: Impact of cell death on the estimation of the mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$. We simulate $R = 100$ pairs of fluctuation assays using the heterogeneous-response model (without differential mutant fitness and with zero division rate of response-*on* cells) extended by cell death, but neglect cell death in the model used to infer $\mu_{off}$ and $\mu_{on}$. The black solid lines indicate the true mutation-rate increase used in the simulations. **A** Estimated mutation-rate increase when only response-*off* cells are affected by cell death, **B** when only response-*on* cells are affected by cell death and **C** when all cells are affected by cell death equally. The parameter range used in the simulations is $\alpha \in \{0.01, 0.05\}$ in $h^{-1}$, $\delta_{off} \in [0.0,\ 0.5]$, $\delta_{on} \in [0.0,\ 0.5]$ in $h^{-1}$.

stress responses with differential mutant fitness using a parameter range of $\rho^p_{off} = \rho^s_{off} \in [0.0,\ 1.5]$. Note that we consider a differential mutant fitness only in response-*off* cells as the response-*on* cells have a zero division rate in this set of simulations and are, therefore, not affected by a differential mutant fitness.
We repeat the simulations $R = 100$ times and assume that the final population sizes and the fraction of the response-*on* subpopulation are known. From the resulting mutant count data, we infer the mutation rates $\mu_{off}$ and $\mu_{on}$ from which, in turn, we calculate the estimated mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$. We compare the estimated value with the true value to determine any biases caused by neglecting the differential mutant fitness in the inference (Figure 8). We find that neglecting this effect in the inference leads to a slight overestimation of the mutation-rate increase when mutants have a fitness cost ($\rho_{off} < 1$). On the other hand, if mutations bring a fitness advantage ($\rho_{off} > 1$), the mutation-rate increase is slightly underestimated. The latter could be the result of, for example, cross-resistance or because the same antibiotic is used as a stressor and on the selective plates [42]. However, the bias remains small across the tested parameter range.

### A.5 Estimation when response-*on* cells have a known non-zero division rate

In Results section 3.3, we evaluate the performance of our method in estimating the mutation-rate increase when response-*on* cells have a non-zero division rate for the cases that (i) the non-zero division rate is neglected in the inference (setting $r_{on} = 0$) and (ii) $r_{on} > 0$ is inferred additionally. Here, we consider a third case: $r_{on}$ is set to the true value, which could, for example, be measured in microfluidics experiments using time-lapse microscopy and image analysis to estimate the division rate of cells identified as response-*on* . Interestingly, we find that setting $r_{on}$ to the true value hardly improves the estimation of the mutation-rate increase (Figure 9). The reason for this lies in the approximation of the size of the response-*on* subpopulation (equation 5) which assumes that $\gamma_{on} - \delta_{on} \ll \gamma^s_{off} - \delta_{off} - \alpha$ and is no longer valid for large $r_{on} = \frac{\gamma_{on} - \delta_{on}}{\gamma^s_{off} - \delta_{off}} \to 1$. Therefore, since we derive the mutant count distribution based on

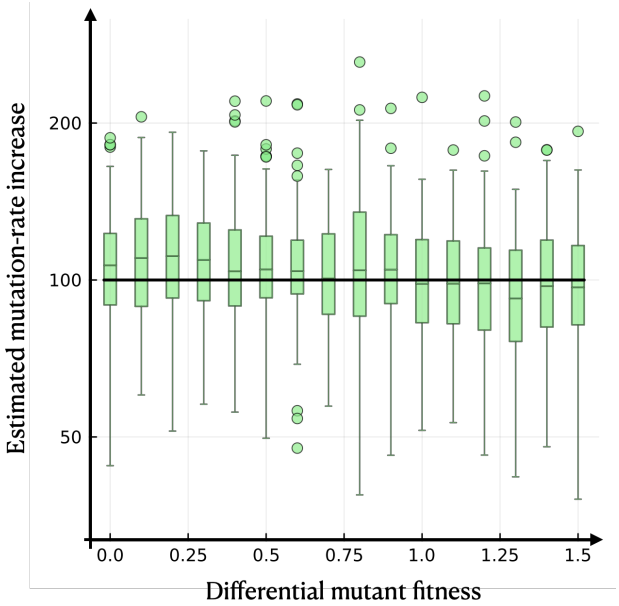


Figure 8: Impact of a differential mutant fitness on the estimation of the mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$. We simulate $R = 100$ pairs of fluctuation assays using the heterogeneous-response model (without cell death and with zero division rate of response-*on* cells) with mutants having a different division rate than non-mutants, but neglect this effect in the model used to infer $\mu_{off}$ and $\mu_{on}$. The black solid line indicates the true mutation-rate increase. The parameter range used in the simulations is $\rho^p_{off} = \rho^s_{off} \in [0.0,\ 1.5]$.

equation 5, setting $r_{on}$ to the true value still results in a biased estimate. This suggests that, given the available inference method, it is not worthwhile to obtain a precise estimate of $r_{on}$, besides validating that it is small.

### A.6 Model selection between heterogeneous and homogeneous response with constrained mutant fitness

In Results section 3.4, we simulate under the heterogeneous-response model for a range of relative fitness of response-*on* cells $r_{on}$ (but without cell death or differential mutant fitness) and perform model selection between (i) the heterogeneous-response model (setting $r_{on} = 0$ and inferring the fraction of response-*on* cells, $f_{on}$, in addition to the mutation rates) and (ii) the homogeneous-response model without differential mutant fitness (setting $\rho^p = \rho^s = 1$) or (iii) the homogeneous-response model with differential mutant fitness ($\rho^p$ and $\rho^s$ inferred in addition to the mutation rates). Moreover, we evaluate the models' performance in estimating the increase in population mean mutation rate. Here, we repeat this analysis, but for model (iii), we constrain the differential mutant fitness to be equal under stressful as under permissive conditions, $\rho^p = \rho^s$.

We find that inferring a constrained differential mutant fitness does not improve the estimation of the increase in population mean mutation rate $\Delta\bar{\mu}$ compared to setting $\rho^p = \rho^s = 1$ (Figure 5A). The differential mutant fitness itself ($\rho^p = \rho^s$) is estimated to be increasingly severe when the true relative fitness of response-*on* cells used in the simulations is smaller (Figure 10B).

Comparing model fit, we find that the heterogeneous-response model is selected in most cases when the relative fitness of response-*on* cells $r_{on}$ is small (Figure 5C). For large values of $r_{on}$, on the other hand, the homogeneous-response model without differential mutant fitness is selected in most cases. The homogeneous-response model with constrained differential mutant fitness is selected in only a few of the simulations, with the highest percentage of $\sim 25\%$ for intermediate values of $r_{on}$.

### A.7 Estimation when the fraction of the response-*on* subpopulation is unknown

In Results section 3.4, we assume that the fraction of response-*on* cells is unknown in the inference using the heterogeneous-response model. In this case, we set the relative fitness of response-*on* cells to $r_{on} = 0$ because otherwise, the likelihood surface would be too complex for numerically efficient inference. As a result, the pgf of the mutant count distribution under stress used in the inference reduces to a simpler form that depends only on the composite parameter $\mu_{on}\frac{f_{on}}{1-f_{on}}$ (equation 24). This implies that the heterogeneous-response model is non-identifiable because the

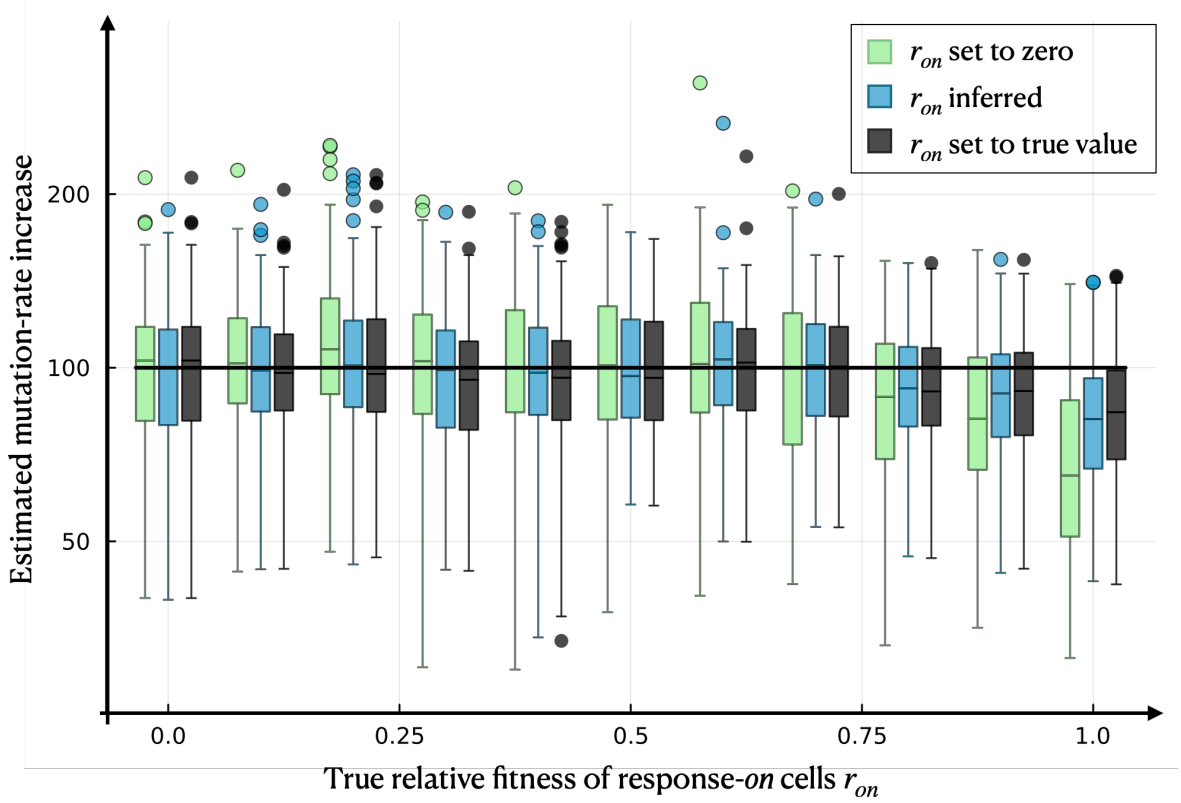


Figure 9: Estimated mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$ when response-*on* cells have a non-zero division rate. We simulate $R = 100$ pairs of fluctuation assays using the heterogeneous-response model (without cell death or differential mutant fitness) with $r_{on} \geq 0$ being the relative fitness of response-*on* cells compared to response-*off* cells. We consider three cases for the inference: (i) setting $r_{on}$ to zero and only inferring $\mu_{off}$ and $\mu_{on}$, (ii) inferring $r_{on}$ in addition to the two mutation rates, and (iii) inferring $\mu_{off}$ and $\mu_{on}$ while setting $r_{on}$ to the true value. The solid black line indicates the true value of $\frac{\mu_{on}}{\mu_{off}}$. The parameter range used in the simulations is $\gamma_{on} \in [0.0,\ 1.0]$ in $h^{-1}$.

parameters $\mu_{on}$ and $f_{on}$ cannot be inferred individually. We find that estimating these parameters separately is heavily biased (Figure 11). Nonetheless, unbiased estimation of the increase in population mean mutation rate $\frac{(1-f_{on})\mu_{off}+f_{on}\mu_{on}}{\mu_{off}}$ is possible (Figure 10A).

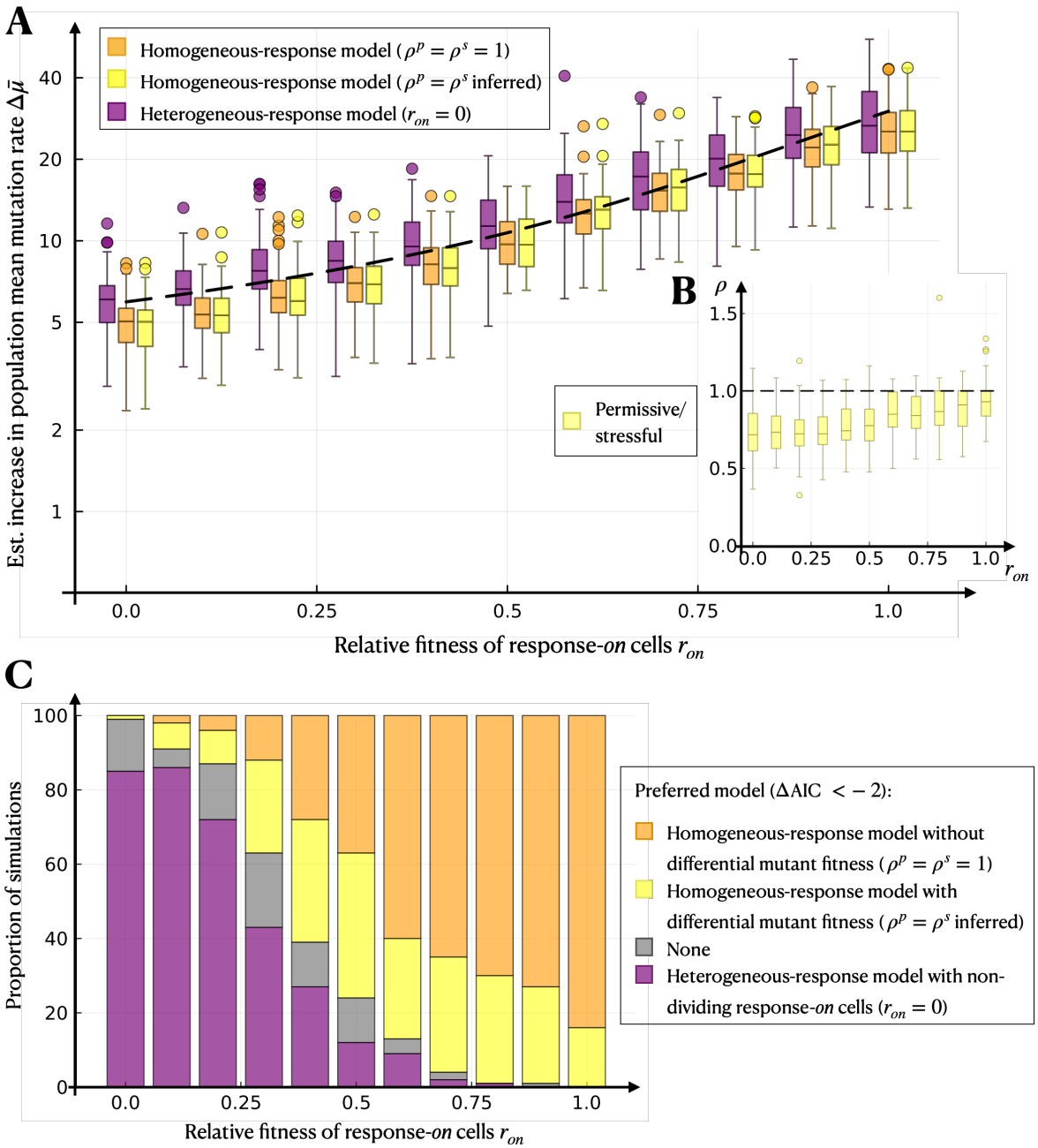


Figure 10: Model selection between the heterogeneous-response and the homogeneous-response model. We simulate $R = 100$ pairs of fluctuation assays under the heterogeneous-response model (without cell death or differential mutant fitness) for a range of relative fitness of response-*on* cells, $r_{on}$. For the inference, we use three models: (i) the heterogeneous-response model to infer the mutation rates $\mu_{off}$ and $\mu_{on}$ plus the fraction of the response-*on* subpopulation $f_{on}$ (but set $r_{on} = 0$), (ii) the homogeneous-response model without differential mutant fitness ($\rho^p = \rho^s = 1$) and (iii) the homogeneous-response model with (joint) differential mutant fitness ($\rho^p = \rho^s$) to infer the population-wide mutation rates under permissive/stressful conditions, $\mu^p$ and $\mu^s$, and additionally $\rho^p = \rho^s$ for model (iii). **A** Estimated increase in population mean mutation rate (given by $\frac{(1-f_{on})\mu_{off}+f_{on}\mu_{on}}{\mu_{off}}$ in the heterogeneous model and $\frac{\mu^s}{\mu^p}$ in the homogeneous models), where the dashed black line indicates the true value. **B** Estimated differential mutant fitness under permissive and stressful conditions for model (iii), where the dashed black line indicates the true value. **C** Model selection using the AIC. The parameter range used in the simulations is $\gamma_{on} \in [0.0,\ 1.0]$ in $h^{-1}$.

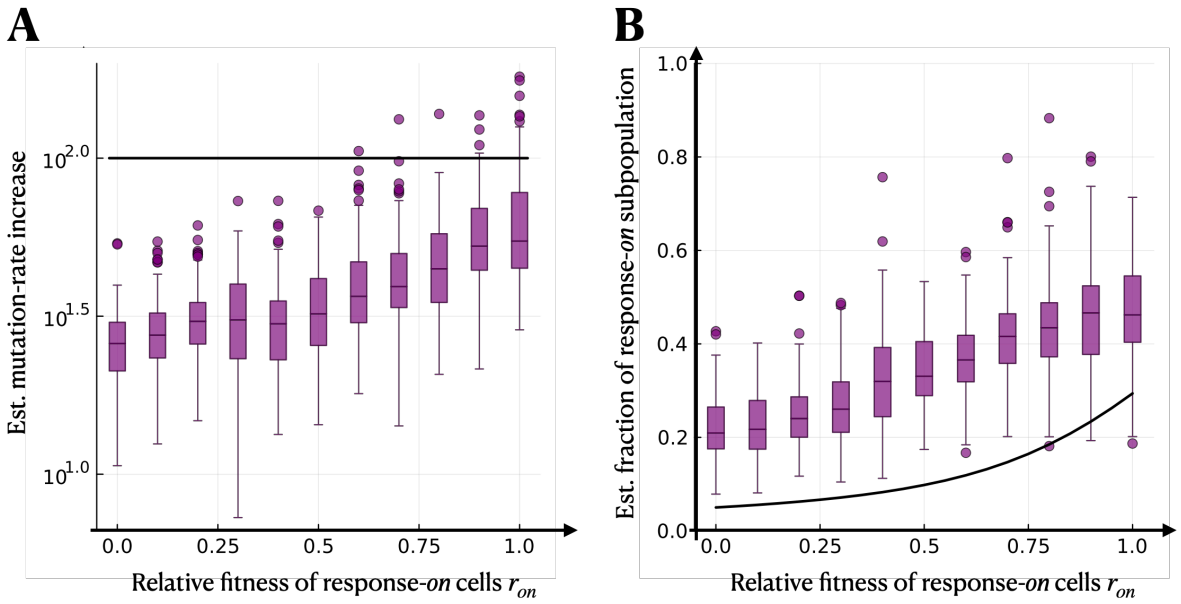


Figure 11: Estimation when the fraction of the response-*on* subpopulation is unknown. We simulate $R = 100$ pairs of fluctuation assays under the heterogeneous-response model (without cell death or differential mutant fitness) for a range of relative fitness of response-*on* cells, $r_{on}$. In the inference, we use the heterogeneous-response model to infer the mutation rates $\mu_{off}$ and $\mu_{on}$ plus the fraction of the response-*on* subpopulation $f_{on}$ (but set $r_{on} = 0$), from which we calculate the mutation-rate increase $\frac{\mu_{on}}{\mu_{off}}$. **A** Estimated mutation-rate increase and **B** estimated fraction of response-*on* subpopulation, $f_{on}$. The true values are indicated in solid black lines. The parameter range used in the simulations is $\gamma_{on} \in [0.0,\ 1.0]$ in $h^{-1}$.

## References


[1] ALEXANDER, H. K., MAYER, S. I., AND BONHOEFFER, S. Population Heterogeneity in Mutation Rate Increases the Frequency of Higher-Order Mutants and Reduces Long-Term Mutational Load. *Molecular biology and evolution 34*, 2 (2017), 419–436.

[2] BAHAROGLU, Z., AND MAZEL, D. Vibrio cholerae Triggers SOS and Mutagenesis in Response to a Wide Range of Antibiotics: a Route towards Multiresistance. *Antimicrobial Agents and Chemotherapy 55*, 5 (may 2011), 2438–2441.

[3] BAHAROGLU, Z., AND MAZEL, D. SOS, the formidable strategy of bacteria against aggressions. *FEMS Microbiology Reviews 38*, 6 (nov 2014), 1126–1145.

[4] BEZANSON, J., EDELMAN, A., KARPINSKI, S., AND SHAH, V. B. Julia: A Fresh Approach to Numerical Computing. *SIAM Review 59*, 1 (jan 2017), 65–98.

[5] BJEDOV, I., TENAILLON, O., GÉRARD, B., SOUZA, V., DENAMUR, E., RADMAN, M., TADDEI, F., AND MATIC, I. Stress-Induced Mutagenesis in Bacteria. *Science 300*, 5624 (may 2003), 1404–1409.

[6] CHOUDHARY, D., LAGAGE, V., FOSTER, K. R., AND UPHOFF, S. Phenotypic heterogeneity in the bacterial oxidative stress response is driven by cell-cell interactions. *Cell Reports 42*, 3 (mar 2023), 112168.

[7] CIRZ, R. T., AND ROMESBERG, F. E. Controlling mutation: Intervening in evolution as a therapeutic strategy. *Critical Reviews in Biochemistry and Molecular Biology 42*, 5 (2007), 341–354.

[8] FOSTER, P. L. Methods for determining spontaneous mutation rates. *Methods in enzymology 409* (2006), 195–213.

[9] FOSTER, P. L. Stress-Induced Mutagenesis in Bacteria. *Critical Reviews in Biochemistry and Molecular Biology 42*, 5 (jan 2007), 373–397.

[10] FRENOY, A., AND BONHOEFFER, S. Death and population dynamics affect mutation rate estimates and evolvability under stress in bacteria. *PLOS Biology 16*, 5 (may 2018), 1–23.

[11] FRIEDBERG, E. C., WALKER, G. C., SIEDE, W., WOOD, R. D., SCHULTZ, R. A., AND ELLENBERGER, T. Mutagenesis and Translesion Synthesis in Prokaryotes. In *DNA Repair and Mutagenesis*. ASM Press, 2006, pp. 509–568.

[12] FRIEDBERG, E. C., WALKER, G. C., SIEDE, W., WOOD, R. D., SCHULTZ, R. A., AND ELLENBERGER, T. The SOS Responses of Prokaryotes to DNA Damage. In *DNA Repair and Mutagenesis*. ASM Press, 2006, pp. 463–508.

[13] GARIBYAN, L., HUANG, T., KIM, M., WOLFF, E., NGUYEN, A., NGUYEN, T., DIEP, A., HU, K., IVERSON, A., YANG, H., AND MILLER, J. H. Use of the rpoB gene to determine the specificity of base substitution mutations on the Escherichia coli chromosome. *DNA Repair 2*, 5 (2003), 593–608.

[14] GILLESPIE, S. H., BASU, S., DICKENS, A. L., O'SULLIVAN, D. M., AND MCHUGH, T. D. Effect of subinhibitory concentrations of ciprofloxacin on Mycobacterium fortuitum mutation rates. *Journal of Antimicrobial Chemotherapy 56*, 2 (2005), 344–348.

[15] GILLET-MARKOWSKA, A., LOUVEL, G., AND FISCHER, G. bz-rates: A web tool to estimate mutation rates from fluctuation analysis. *G3: Genes, Genomes, Genetics 5*, 11 (2015), 2323–2327. `https://www.lcqb.upmc.fr/bzrates`.

[16] GUTIERREZ, A., LAURETI, L., CRUSSARD, S., ABIDA, H., RODRÍGUEZ-ROJAS, A., BLÁZQUEZ, J., BAHAROGLU, Z., MAZEL, D., DARFEUILLE, F., VOGEL, J., AND MATIC, I. $\beta$-lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity. *Nature Communications 4* (2013).

[17] HALL, B. M., MA, C. X., LIANG, P., AND SINGH, K. K. Fluctuation anaLysis calculator: A web tool for the determination of mutation rate using Luria-Delbück fluctuation analysis. *Bioinformatics 25*, 12 (2009), 1564–1565. `https://lianglab.brocku.ca/FALCOR/`.

[18] HENDERSON-BEGG, S. K., LIVERMORE, D. M., AND HALL, L. M. Effect of subinhibitory concentrations of antibiotics on mutation frequency in Streptococcus pneumoniae. *Journal of Antimicrobial Chemotherapy 57*, 5 (2006), 849–854.

[19] JARAMILLO-RIVERI, S., BROUGHTON, J., MCVEY, A., PILIZOTA, T., SCOTT, M., AND EL KAROUI, M. Growth-dependent heterogeneity in the DNA damage response in Escherichia coli. *Molecular Systems Biology 18*, 5 (2022), 1–14.

[20] JEE, J., RASOULY, A., SHAMOVSKY, I., AKIVIS, Y., R. STEINMAN, S., MISHRA, B., AND NUDLER, E. Rates and mechanisms of bacterial mutagenesis from maximum-depth sequencing. *Nature 534*, 7609 (jun 2016), 693–696.

[21] JONES, E. C., AND UPHOFF, S. Single-molecule imaging of LexA degradation in Escherichia coli elucidates regulatory mechanisms and heterogeneity of the SOS response. *Nature Microbiology* (jun 2021).

[22] KAMENŠEK, S., PODLESEK, Z., GILLOR, O., AND ŽGUR-BERTOK, D. Genes regulated by the Escherichia coli SOS repressor LexA exhibit heterogenous expression. *BMC Microbiology 10*, i (2010).

[23] KELLER, P., AND ANTAL, T. Mutant number distribution in an exponentially growing population. *Journal of Statistical Mechanics: Theory and Experiment 2015*, 1 (jan 2015), P01011.

[24] KOHANSKI, M. A., DEPRISTO, M. A., AND COLLINS, J. J. Sublethal Antibiotic Treatment Leads to Multidrug Resistance via Radical-Induced Mutagenesis. *Molecular Cell 37*, 3 (feb 2010), 311–320.

[25] KRAŠOVEC, R., RICHARDS, H., GOMEZ, G., GIFFORD, D. R., MAZOYER, A., AND KNIGHT, C. G. Measuring microbial mutation rates with the fluctuation assay. *Journal of Visualized Experiments 2019*, 153 (2019), 1–9. `https://adrimaz.shinyapps.io/shinyflan/`.

[26] LAGAGE, V., CHEN, V., AND UPHOFF, S. Adaptation delay causes a burst of mutations in bacteria responding to oxidative stress. *EMBO reports 24*, 1 (2023).

[27] LAGAGE, V., AND UPHOFF, S. Pulses and delays, anticipation and memory: seeing bacterial stress responses from a single-cell perspective. *FEMS microbiology reviews 44*, 5 (2020), 565–571.

[28] MAZOYER, A., DROUILHET, R., DESPRÉAUX, S., AND YCART, B. Flan: An R package for inference on mutation models. *R Journal 9*, 1 (2017), 334–351.

[29] MCCOOL, J. D., LONG, E., PETROSINO, J. F., SANDLER, H. A., ROSENBERG, S. M., AND SANDLER, S. J. Measurement of SOS expression in individual Escherichia coli K-12 cells using fluorescence microscopy. *Molecular Microbiology 53*, 5 (jul 2004), 1343–1357.

[30] MRAK, P., PODLESEK, Z., VAN PUTTEN, J. P., AND ŽGUR-BERTOK, D. Heterogeneity in expression of the Escherichia coli colicin K activity gene cka is controlled by the SOS system and stochastic factors. *Molecular Genetics and Genomics 277*, 4 (2007), 391–401.

[31] PRIBIS, J. P., GARCÍA-VILLADA, L., ZHAI, Y., LEWIN-EPSTEIN, O., WANG, A. Z., LIU, J., XIA, J., MEI, Q., FITZGERALD, D. M., BOS, J., AUSTIN, R. H., HERMAN, C., BATES, D., HADANY, L., HASTINGS, P., AND ROSENBERG, S. M. Gamblers: An Antibiotic-Induced Evolvable Cell Subpopulation Differentiated by Reactive-Oxygen-Induced General Stress Response. *Molecular Cell 74*, 4 (may 2019), 785–800.e7.

[32] PRIBIS, J. P., ZHAI, Y., HASTINGS, P. J., AND ROSENBERG, S. M. Stress-Induced Mutagenesis, Gambler Cells, and Stealth Targeting Antibiotic-Induced Evolution. *mBio 13*, 3 (2022).

[33] TENAILLON, O., DENAMUR, E., AND MATIC, I. Evolutionary significance of stress-induced mutagenesis in bacteria. *Trends in Microbiology 12*, 6 (2004), 264–270.

[34] UPHOFF, S. Real-time dynamics of mutagenesis reveal the chronology of DNA repair and damage tolerance responses in single cells. *Proceedings of the National Academy of Sciences 115*, 28 (jul 2018), E6516–E6525.

[35] UPHOFF, S., LORD, N. D., OKUMUS, B., POTVIN-TROTTIER, L., SHERRATT, D. J., AND PAULSSON, J. Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation. *Science 351*, 6277 (mar 2016), 1094–1097.

[36] VASSE, M., BONHOEFFER, S., AND FRENOY, A. Ecological effects of stress drive bacterial evolvability under sub-inhibitory antibiotic treatments. *ISME Communications 2*, 1 (2022).

[37] VINCENT, M. S., AND UPHOFF, S. Cellular heterogeneity in DNA alkylation repair increases population genetic plasticity. *Nucleic Acids Research 49*, 21 (dec 2021), 12320–12331.

[38] WOO, A. C., FAURE, L., DAPA, T., AND MATIC, I. Heterogeneity of spontaneous DNA replication errors in single isogenic Escherichia coli cells. *Science Advances 4*, 6 (2018), 2–10.

[39] WU, Y. L., SCOTT, E. M., LI WAN PO, A., AND TARIQ, V. N. Development of resistance and cross-resistance in Pseudomonas aeruginosa exposed to subinhibitory antibiotic concentrations. *Apmis 107*, 6 (1999), 585–592.

[40] ZHAI, Y., PRIBIS, J. P., DOOLING, S. W., GARCIA-VILLADA, L., MINNICK, P. J., XIA, J., LIU, J., MEI, Q., FITZGERALD, D. M., HERMAN, C., HASTINGS, P. J., COSTA-MATTIOLI, M., AND ROSENBERG, S. M. Drugging evolution of antibiotic resistance at a regulatory network hub. *Science advances 9*, 25 (2023), eadg0188.

[41] ZHENG, Q. rSalvador: An R package for the fluctuation experiment. *G3: Genes, Genomes, Genetics 7*, 12 (2017), 3849–3856.

[42] ZHENG, Q. Estimation of Rates of Non-neutral Mutations When Bacteria are Exposed to Subinhibitory Levels of Antibiotics. *Bulletin of Mathematical Biology 84*, 11 (2022), 1–18.