POLITECNICO DI TORINO
Repository ISTITUZIONALE

Optimal metabolic strategies for microbial growth in stationary random environments

Original
Optimal metabolic strategies for microbial growth in stationary random environments / Muntoni, A.P., DE MARTINO, A.. - In: PHYSICAL BIOLOGY. - ISSN 1478-3967. - 20:3(2023), p. 036001. [10.1088/1478-3975/acc1bc]

Availability:
This version is available at: 11583/2976710 since: 2023-03-23T09:53:45Z

Publisher:
IOP

Published
DOI:10.1088/1478-3975/acc1bc

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

29 November 2026

IOP Publishing Phys. Biol. 20 (2023) 036001 https://doi.org/10.1088/1478-3975/acc1bc

Physical Biology

CrossMark

OPEN ACCESS

RECEIVED
21 October 2022

REVISED
28 January 2023

ACCEPTED FOR PUBLICATION
6 March 2023

PUBLISHED
21 March 2023

Original Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

Optimal metabolic strategies for microbial growth in stationary random environments

Anna Paola Muntoni[1,2] **and Andrea De Martino**[1,2,*]

[1] Politecnico di Torino, Turin, Italy
[2] Italian Institute for Genomic Medicine, Turin, Italy
[*] Author to whom any correspondence should be addressed.

E-mail: andrea.demartino@polito.it

Keywords: information theory, optimal growth strategies, metabolism, random media

Abstract

In order to grow in any given environment, bacteria need to collect information about the medium composition and implement suitable growth strategies by adjusting their regulatory and metabolic degrees of freedom. In the standard sense, optimal strategy selection is achieved when bacteria grow at the fastest rate possible in that medium. While this view of optimality is well suited for cells that have perfect knowledge about their surroundings (e.g. nutrient levels), things are more involved in uncertain or fluctuating conditions, especially when changes occur over timescales comparable to (or faster than) those required to organize a response. Information theory however provides recipes for how cells can choose the optimal growth strategy under uncertainty about the stress levels they will face. Here we analyse the theoretically optimal scenarios for a coarse-grained, experiment-inspired model of bacterial metabolism for growth in a medium described by the (static) probability density of a single variable (the 'stress level'). We show that heterogeneity in growth rates consistently emerges as the optimal response when the environment is sufficiently complex and/or when perfect adjustment of metabolic degrees of freedom is not possible (e.g. due to limited resources). In addition, outcomes close to those achievable with unlimited resources are often attained effectively with a modest amount of fine tuning. In other terms, heterogeneous population structures in complex media may be rather robust with respect to the resources available to probe the environment and adjust reaction rates.

1. Introduction

The standard theoretical view of bacterial growth posits that, in any growth medium, cells are capable of adjusting their metabolic degrees of freedom (i.e. the rates of metabolic reactions) within bounds dictated by thermodynamic and regulatory constraints (e.g. enzyme expression levels, reaction free energies, etc) so as to maximize their growth rate [1]. Besides evolutionary considerations, such a picture is supported by the fact that the expression levels of certain metabolic enzymes and of basic macromolecular machines like ribosomes actually appear to be tuned for growth-rate maximization in bacterial populations [2, 3]. On the other hand, the significant cell-to-cell variability in growth rates observed in experiments [4–8], together with the fact that constraints arising outside metabolism suffice to explain a large batch of empirical facts without assuming any growth-rate optimization [9, 10], appears to call for deeper insight into the notion of 'optimality' for bacterial growth.

Recent work has shown that the relationship between population growth and cell-to-cell variability is well described by a Maximum Entropy (MaxEnt) theory leading to a variable trade-off akin to the usual energy-entropy balance in statistical physics. More specifically, *E. coli* populations growing in carbon-limited media realize a close-to-optimal fitness-heterogeneity trade-off in rich media [11], while they seem to be less variable or slower-growing than optimal in poorer growth conditions [12]. Metabolic fluxes likewise appear to be better captured by accounting for such a trade-off than by a standard optimality assumption [13]. In each case, the balance between growth and variability is described by a finite (medium-dependent) 'temperature', where a

© 2023 The Author(s). Published by IOP Publishing Ltd

zero-temperature (resp. infinite-temperature) limit corresponds to maximal growth (resp. maximal variability). Save for a few general ideas derived from broad-brush models [14], what determines the 'temperature' (i.e. the fitness-heterogeneity balance) of actual microbial systems is still unclear.

High variability can naturally arise from unavoidable inter-cellular differences in gene expression levels or regulatory programs (e.g. cell cycle) [15]. In models of metabolism, this would lead, at the simplest level, to cell-dependent changes in the constraints under which growth is optimized. In this respect, cell-to-cell heterogeneity might be interpreted as 'optimality plus noise', and the 'temperature' described above would quantify, in essence, the noise strength. Importantly, though, there might be an inherent advantage in maintaining a diverse population, especially in environments that fluctuate (due e.g. to natural variability) or when cells have imperfect information about their growth medium (due e.g. to limits in precision caused by the high costs cells face to maintain and operate a sensing apparatus). These factors are not usually included in standard models of metabolic networks, which therefore cannot address the fitness benefits of heterogeneity.

The problem of growth maximization clearly becomes more subtle under uncertainty about environmental conditions, as the straightforward optimization that can be carried out in a perfectly known medium is no longer an option. Recipes for selecting the optimal growth strategy in uncertain environments are, however, provided by information theory. Theoretical work aimed at understanding how efficiently populations can harvest, process and exploit information about variable or unpredictable media has indeed shown that *bet-hedging* (i.e. maintaining a fraction of slow-growing cells even in rich media or sustaining a lower short-term growth to ensure faster long-term growth) can yield significant fitness gains in a wide variety of situations [16]. Biological implications of these results have been explored against several backdrops [17–19], albeit never specifically in the context of metabolism.

In this paper, inspired by the above studies as well as by [20, 21] (chapter 6), we look at a minimal, experiment-derived mathematical model to characterize optimal metabolic strategies for growth in uncertain environments, focusing for simplicity on static environments defined by the probability density of a single variable (the 'stress level'). Optimal strategies are parametrized by a 'temperature' that modulates the amount of information they encode about the growth medium or, loosely speaking, how precisely cells can match their phenotype to the external conditions in order to foster growth. We will show explicitly that, when the medium is sufficiently complex, optimal populations acquire a non-trivial phenotypic organization even when the metabolic strategy encodes the maximum possible amount of information about the external conditions. A broad spectrum of behaviours is uncovered upon varying the structure of the environment. Remarkably, however, the emerging scenarios are often robust to changes in the 'temperature'. This suggests that metabolic networks may yield outcomes close to globally optimal ones (at least in an information-theoretic sense) even when resources to probe the environment and adjust metabolic reactions are limited.

2. Results

2.1. Model of metabolism and growth

We consider a coarse-grained model of microbial growth metabolism in which each cell's metabolic strategy is described by just two quantities, namely the specific uptake (or inverse growth yield) q, quantifying the nutrient intake required to grow per unit of growth rate, and the biosynthetic expenditure ϵ, quantifying the proteome mass fraction to be devoted to metabolic enzymes per unit of growth rate. (In more detailed models of metabolic networks, the former quantity relates to the rate at which the limiting nutrient is imported, while the latter is proportional to a weighted sum of the absolute values of the fluxes through metabolic reactions [20, 22].) For *E. coli* growing in carbon-limited media it has been argued that, for given q and ϵ, the growth rate μ is well described by the formula [20]

$$\mu \simeq \frac{\phi}{w + sq + \epsilon}, \tag{1}$$

where $s \geqslant 0$ represents the level of nutritional stress to which the organism is subject, while $\phi > 0$ and $w > 0$ are constants representing respectively the fraction of proteome devoted to constitutively expressed proteins and the proteome share to be allocated to ribosome-affiliated proteins per unit of growth rate. (For glucose-limited *E. coli* growth, $\phi \simeq 0.48$ and $w \simeq 0.169$ h [9].) For our purposes, s can be assumed to be inversely proportional to the carbon level as argued in [22] (so $s \ll 1$ and $s \gg 1$ for carbon-rich and carbon-poor environments, respectively).

Let us assume that the stress level s in (1) is a homogeneous parameter whose value is controlled externally. If μ were to be maximized, the quantity $sq + \epsilon$ would have to be minimized. In carbon-limited *E. coli* growth, however, q and ϵ are subject to a trade-off such that high q implies low ϵ and vice versa [20]. Metabolic states with minimal biosynthetic expenditure are hence favoured in rich environments (small s), while states of minimal nutritional requirements prevail in poor media (large s). To make a concrete model, we draw inspiration from the fermentation/respiration duality that again characterizes *E. coli* growth in carbon-limited media and assume that the organism can regulate q and ϵ between two extreme

strategies, denoted by indices F and R, respectively, distinguished by the fact that $q_F > q_R$ (i.e. the specific nutrient requirements of F are higher than those of R) and $\epsilon_R > \epsilon_F$ (i.e. the specific expenditure for R is higher than for F). (Estimated values of the specific nutrient intake and the specific proteome mass fraction devoted to metabolic enzymes required for *E.coli* to grow on lactate under fermentation are $q_F \simeq 8\, g_{lac}/g_{DW}$ and $\epsilon_F \simeq 0.3$ h, respectively; the corresponding quantities under respiration are instead given by $q_R \simeq 5\, g_{lac}/g_{DW}$ and $\epsilon_R \simeq 0.6$ h [20]. In this work, we choose the representative values $q_F = 10$, $\epsilon_F = 0.1$, $q_R = 1$, $\epsilon_R = 1$, omitting the units as the specifics of the carbon source are immaterial for us. Nevertheless, with these choices the growth rate μ can be interpreted to be measured in 1/h.) We then describe the trade-off between q and ϵ by assuming that both depend on a single variable x whose value ranges between 0 and 1, such that

$$q(x) = q_R + (q_F - q_R)(1-x)^\nu, \tag{2}$$

$$\epsilon(x) = \epsilon_F + (\epsilon_R - \epsilon_F)x^\nu, \tag{3}$$

where $\nu > 1$ is a constant. The growth rate of the organism will then be given by

$$\mu(x,s) = \frac{\phi}{w + sq(x) + \epsilon(x)}. \tag{4}$$

By taking the derivative of $\mu(x,s)$ over x at fixed s, one finds that, for any given $q(x)$ and $\epsilon(x)$, μ is maximum when

$$s\frac{\partial q}{\partial x} + \frac{\partial \epsilon}{\partial x} = 0. \tag{5}$$

If $q(x)$ and $\epsilon(x)$ are given by (2) and (3), the maximum is achieved for $x = \widehat{x}(s)$, with

$$\widehat{x}(s) = \frac{s^n}{s^n + s_c^n}, \quad s_c = \frac{\epsilon_R - \epsilon_F}{q_F - q_R}, \quad n = \frac{1}{\nu - 1}. \tag{6}$$

This means that media with $s \ll s_c$ can be considered to be rich, while media with $s \gg s_c$ are effectively poor. (With our choices for the parameters, $s_c = 0.1$.) In turn, if μ is maximized, strategy F (i.e. $x = 0$) will be used in rich environments (where ϵ should be as small as possible) while strategy R (i.e. $x = 1$) will be used in poor ones (where q should be as small as possible), as shown in figure 1. As one modulates the stress level between these two extremes, growth is maximized by intermediate values of x (i.e. by strategies that use both F and R). Other choices generically lead to slower growth.

Notice that the $q - \epsilon$ trade-off gets stronger as ν approaches 1, when the corresponding optimal strategy is a step-like function. Conversely, it gets weaker and weaker as ν increases. For sakes of simplicity, in the following we shall always use the value $\nu = 3/2$, which qualitatively reproduces the trade-off reconstructed from empirical data [20]. A discussion of how results depend on ν (including the issue of why a specific value of ν may be evolutionarily preferred) is deferred to future work.

2.2. Optimizing growth in random environments: theoretical framework

Consider a microbe whose growth rate depends on x and s as in (4). For any fixed s, the cell can maximize μ by setting $x = \widehat{x}(s)$ (see (6)). Suppose, however, that s is a random variable with a prescribed probability density $p(s)$. In this case, cells face an uncertainty about the exact value of s they will encounter, although they have knowledge of the *ensemble* of environmental conditions in which they live (i.e. of $p(s)$). What is the optimal choice for x in this situation? The most sensible measure of performance is now arguably given by the mean growth rate

$$\langle \mu \rangle = \int ds p(s) \int dx p(x|s) \mu(x,s), \tag{7}$$

whose value depends on the conditional distribution $p(x|s)$ that describes the stochastic rule used by the cell to select x for any s. The question of optimality therefore concerns the optimal choice of $p(x|s)$. It turns out that this choice depends on the amount of information about s that is encoded in x. In particular, if

$$I = \int ds p(s) \int dx p(x|s) \log_2 \frac{p(x|s)}{p(x)}, \tag{8}$$

denotes the mutual information of x and s (in bits), and

$$p(x) = \int ds p(s) p(x|s), \tag{9}$$

then the optimal $p(x|s)$ is given by the solution of

$$\max_{p(x|s)} \langle \mu \rangle \quad \text{subject to} \quad I = \text{constant}, \tag{10}$$

i.e. (see appendix and [21])

$$p^\star(x|s) = \frac{p^\star(x)}{N(s,\beta)} e^{\beta \mu(x,s)}, \tag{11}$$

where

$$N(s,\beta) = \int dx p^\star(x) e^{\beta \mu(x,s)}, \tag{12}$$

while the 'inverse temperature' β is a Lagrange multiplier and $p^\star(x)$ denotes the probability density of x corresponding to the optimal choice. Equation (11) has a rather straightforward interpretation. When $\beta \to 0$ ('infinite temperature'), the choice of x becomes independent of s, which implies $I = 0$. As β increases, I increases (i.e. cellular responses encode more and more information about the environment) and $p^\star(x|s)$ tends to get more and more sharply peaked around $\widehat{x}(s)$. For $\beta \to \infty$ ('zero temperature'),

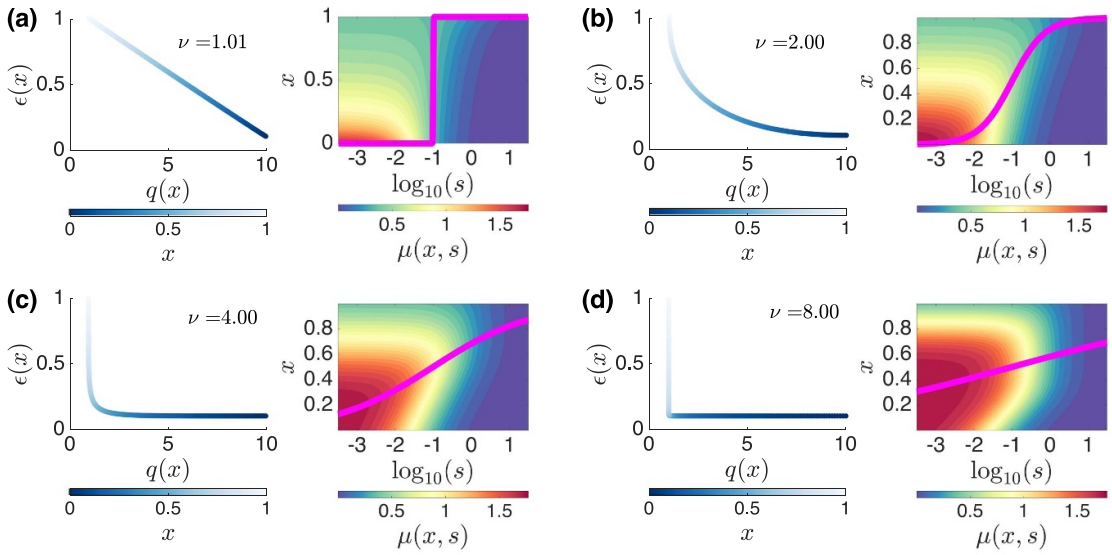

Figure 1. Representative scenarios for $q(x)$ vs $\epsilon(x)$ (equations (2) and (3), left plots in each panel) with the corresponding fitness landscape $\mu(x,s)$ (right plots), for $\nu = 1.01$ (panel (a)), 2 (b), 4 (c) and 8 (d). The values of x that maximize μ for each s, i.e. $\widehat{x}(s)$ (equation (6)), are represented by a magenta line. The constants characterizing the R and F strategies are set to $q_F = 10$, $q_R = 1$, $\epsilon_F = 0.1$ and $\epsilon_R = 1$.

in particular, one gets $p^\star(x|s) \simeq \delta[x - \widehat{x}(s)]$, so that cells respond to each instance of s by exactly choosing the value of x that maximizes μ. Such an 'infinite-precision' response requires I to be maximal. In other terms, the higher I, the higher $\langle\mu\rangle$. (Conversely, as β becomes more and more negative, $p^\star(x|s)$ tends to get more and more sharply peaked around the value of x that *minimizes* growth for any s. For sake of clarity, we shall however limit the following analysis to the case $\beta > 0$.) β is therefore a parameter by which one interpolates between the case of optimal response to any environmental cue and that in which the cell's metabolic strategy is completely insensitive to s. Note that, at optimality, the mean growth rate and mutual information given by

$$\langle\mu\rangle_\star = \int ds p(s) \int dx p^\star(x|s)\mu(x,s), \tag{13}$$

$$I^\star = \int ds p(s) \int dx p^\star(x|s) \log_2 \frac{p^\star(x|s)}{p^\star(x)}, \tag{14}$$

are both functions of β and are related by

$$I^\star = \frac{\beta}{\ln 2}\langle\mu\rangle_\star - \langle\log_2 N(s,\beta)\rangle_\star. \tag{15}$$

For different choices of $p(s)$ and β one can solve equations (9), (11) and (12) numerically by iteration from an initial guess. Starting (iteration $n = 0$) from uniform guesses for $p(x)$ and $p(x|s)$, we iterated equations (12), (9) and (11) up to numerical convergence, which we assumed to be achieved when

$$\max_x |p^{(n)}(x) - p^{(n-1)}(x)| < \sigma \quad \text{and} \tag{16}$$

$$\max_x |p^{(n)}(x|s) - p^{(n-1)}(x|s)| < \sigma, \tag{17}$$

where the index n denotes the iteration step while σ a numerical precision threshold. (A Matlab code implementing the above procedure is available from https://github.com/anna-pa-m/OptMetStrategy.) Solutions will clarify how optimal metabolic strategy and population structure change with β, i.e. with the amount of information about the environment encoded in x, in any given environmental condition (described by the chosen $p(s)$).

Before moving on, we note that the above setting suggests how an 'optimal' value of β may arise in this scenario. It is indeed reasonable to think that $I^\star$ is directly related to the quantity of cellular resources devoted to probing the environment and tuning metabolic reactions, and, in turn, that higher costs for sensing and metabolism may negatively affect the fitness if resources are limited. If one assumes for simplicity that such a cost reduces the growth performance by a constant amount c per bit of information encoded, the fitness faced by the organism can be written as

$$\mathcal{F} = \langle\mu\rangle_\star - cI^\star. \tag{18}$$

One easily sees that, contrary to $\langle\mu\rangle_\star$ and $I^\star$ (both of which increase steadily with β), $\mathcal{F}$ is maximum when $c = \partial\langle\mu\rangle_\star / \partial I^\star$, i.e. for

$$\beta = \beta^\star \equiv \frac{\ln 2}{c}. \tag{19}$$

In other terms, the costs associated with sensing and adjusting metabolism can lead to the existence of an optimal value of β, i.e. of an optimal level of trade-off between $\langle\mu\rangle$ and I, such that (expectedly) higher values of $\mathcal{F}$ are possible only if the cost of encoding information about the environment in metabolic

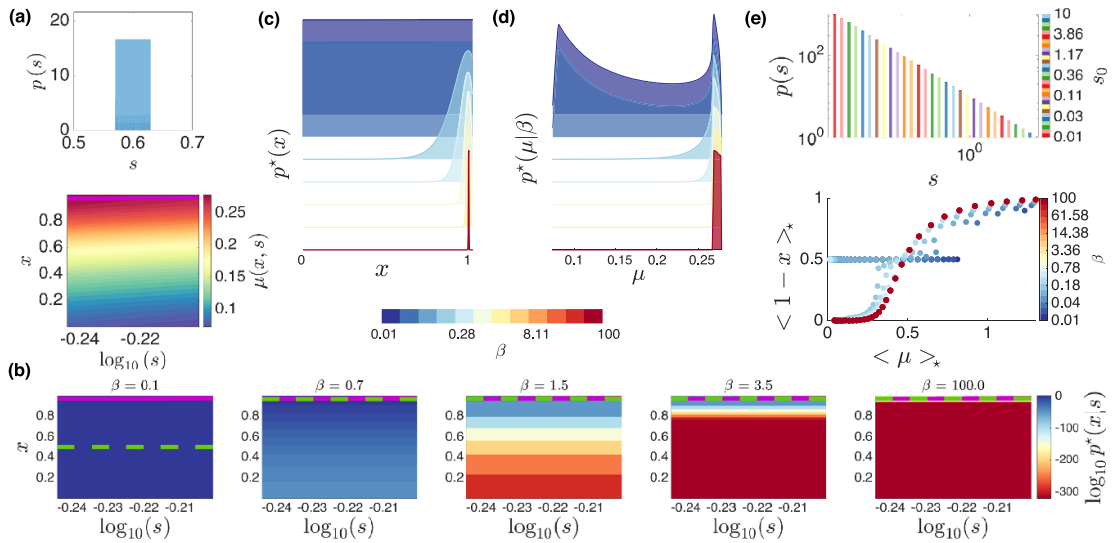

Figure 2. (a) Probability density of the stress level s in a controlled environment with $s_0 = 0.6 > s_c$ (top) and corresponding fitness landscape (bottom). Magenta line: growth-maximizing value of x for each s. (b) Optimal choice for $p^\star(x|s)$ for different values of β. Magenta line: growth-maximizing value of x for each s. Green line: mean value of x for the optimal growth strategy as a function of s. (c) Marginal probability density $p^\star(x)$ for different values of β. (d) Growth rate distribution at optimality, $p^\star(\mu|\beta)$, for different values of β. (e, top) Several stress distributions centered at different values of s_0. (e, bottom) Behavior of the quantity $\langle 1-x\rangle_\star$ as a function of the fitness for several values of β and s_0 (for fixed β, the fitness increases as s_0 decreases). A value of $\langle 1-x\rangle_\star \simeq 0$ indicates a strong preference for R (respiration); a value of $\langle 1-x\rangle_\star \simeq 1$ points instead towards massive use of F (fermentation).

strategies gets smaller. (A similar idea was discussed in a different context in [13].)

2.3. Tightly controlled stress levels

We start from the simple case where the stress level s is tightly regulated, as in a lab setting where the nutrient level is externally controlled with high precision. Specifically, we assume that $p(s)$ is uniform and centered at a value s_0, and that s can only vary by 5% with equal probability around s_0 (figure 2(a), top panel). The ensuing landscape of $\mu(x,s)$ is shown in figure 2(a), bottom panel. One sees that, with our choice of s_0 ($s_0 > s_c$, see (6)), growth is maximized for $x \simeq 1$ (R strategy) for all values of s (magenta line in figure 2(a)). Figure 2(b) reports the optimal strategy $p^\star(x|s)$ computed numerically for five different values of β. As expected, the distribution concentrates around the growth-maximizing value of x for large enough β, while it becomes more and more uniform over the $[0,1]$ interval (with the mean x, green dashed line, getting closer and closer to $1/2$) as β approaches zero. This gradual shift is also clearly visible in the behaviour of $p^\star(x)$, representing the distribution of metabolic strategies at optimality (figure 2(c)). Correspondingly, the growth-rate distribution $p^\star(\mu|\beta)$, which can be obtained numerically as

$$p^\star(\mu|\beta) = \int ds\, p(s) \int dx\, p^\star(x|s)\, \delta[\mu(x,s) - \mu], \tag{20}$$

is bimodal for small β but then tends to get more and more concentrated around the maximum achievable values of $\mu(x,s)$ as β increases (figure 2(d)). The residual variability seen at large β reflects the fact that the stress level can only be known up to a finite precision (see figure 2(a), top panel).

It is instructive to visualize together the results of a series of box-like environments $p(s)$ with different values of s_0, from very low to very high (see figure 2(e), top panel), a setup that mimics an ensemble of experiments performed at different, exogenously controlled nutrient levels. As soon as β is large enough, the mean response of the system (i.e. the mean value of x) changes from $x \simeq 0$ (F strategy) in lower-stress media (faster growth) to $x \simeq 1$ (R strategy) in higher-stress media (slower growth), see figure 2(e), bottom panel. This scenario recapitulates qualitatively the well-known respiration-fermentation crossover marked by the onset of acetate excretion that is observed in experimental *E. coli* populations [23]. Noticeably, however, it appears already for values β well below the reference scale given by $1/\mu_{\max}$ (with $\mu_{\max}$ the fastest growth rate achievable in this medium, approximately equal to 1/4 in this case). Only for values of β much smaller than this does the crossover not develop. In other terms, in these simple environments, the crossover appears to be a feature of the optimal stochastic response $p^\star(x|s)$ that is extremely robust over β.

2.4. Exponentially distributed stress levels

We now turn to the case in which the stress level is exponentially distributed. Specifically, we assume that $p(s) \propto \exp(-s/s_0)$ on the interval $[s_{\min}, s_{\max}]$

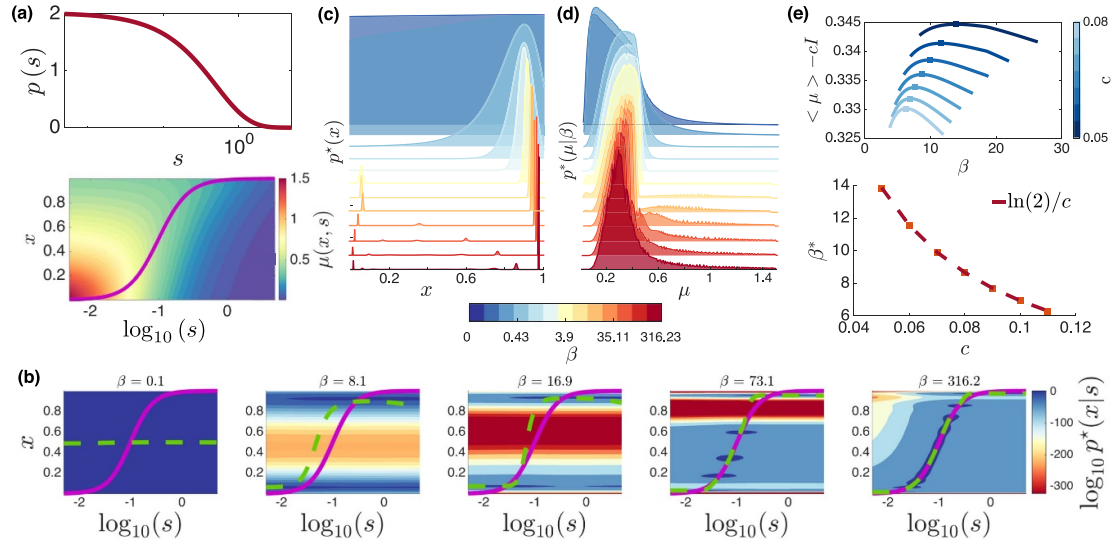

Figure 3. (a) Probability density of the stress level s in the 'exponential environment' (top) $p(s) \propto e^{-s/s_0}$ for $s_0 = 0.5$, $s \in [5 \cdot 10^{-3}, 5 \cdot 10^{0}]$ and corresponding fitness landscape (bottom). Magenta line: growth-maximizing value of x for each s. (b) Optimal choice for $p^\star(x|s)$ for different values of β. Magenta line: growth-maximizing value of x as a function of s. Green line: mean value of x for the optimal choice rule at different values of s. (c) Marginal probability density $p^\star(x)$ for different values of β. (d) Growth rate distribution at optimality, $p^\star(\mu|\beta)$, for different values of β. (e, top) Overall fitness $\mathcal{F}$ (equation (18)) as a function of β for several values of the information cost-per-bit c. (e, bottom) Optimal value of β ($\beta^\star$) versus c: numerical results (markers) vs theoretical prediction (equation (19), dashed line).

(figure 3(a), top panel). The corresponding fitness landscape $\mu(x,s)$ is shown in figure 3(a), bottom panel, along with the growth-maximizing curve $x = \widehat{x}(s)$ (magenta line). One sees that $\widehat{x}(s)$ now shifts continuously from 0 (F strategy) to 1 (R strategy) as the stress level increases. Figure 5(b) shows instead the optimal growth strategy $p^\star(x|s)$ for different values of β, with the green line representing the mean value of x at each s. As expected, for small enough β, $p^\star(x|s)$ is roughly uniform over $[0,1]$. Upon increasing β, however, the distribution acquires a structure that more and more closely concentrates around $\widehat{x}(s)$. The qualitatively correct crossover is however already observed at values of β of the order of $1/\mu_{\max}$. In turn, the distribution of x at optimality, $p^\star(x)$, acquires a very strong bimodal character as β increases (figure 3(c)). This clearly indicates that the coexistence of two subpopulations of fermenters ($x \simeq 0$) and respirators ($x \simeq 1$) provides the optimal metabolic response even in presence of limited resources to encode information about the environment into metabolic strategies, giving a quantitative representation to *bet-hedging* in the present model. Correspondingly, the growth-rate distribution (figure 3(d)) consistently displays large variability with a surprisingly weak dependence on β. Note that no bimodality is observed in $p^\star(\mu|\beta)$ despite the bimodality of $p^\star(x)$.

It is important to remark that large fluctuations persist even in the $\beta \to \infty$ limit, when the information about the stress levels encoded in the metabolic strategy is maximal. In other terms, the globally optimal growth strategy ($\beta \to \infty$) in an exponentially-distributed medium requires a large degree of heterogeneity in growth rates. To better visualize the robustness of these outcomes with respect to β, we display in figure 3(e) the re-scaled fitness $\mathcal{F}$ defined in (18) as a function of the 'inverse temperature'. As anticipated, $\mathcal{F}$ has a maximum at intermediate value of β, related to the cost-per-bit c of encoding information as shown in (19) (bottom panel): higher costs imply lower optimal values of β. Noticeably, though, with the realistic parameters we used, the maximum of the fitness function is extremely flat. This suggests that, at least in principle, metabolic networks could be capable of generating optimal growth profiles that are robust to the amount of information about the environment encoded in the growth strategy. In this respect, questions concerning the fine-tuning of the value of β might be less relevant for this problem than they likely are for the MaxEnt-related growth-heterogeneity trade-off discussed in [11–13]: relatively limited environmental cues could suffice to bias the distribution of metabolic strategies enough to generate a (stochastic) response close to the globally optimal one, i.e. that achievable by perfectly matching x to the stress levels.

2.5. Power-law distributed stress levels

A yet more complex scenario is that in which the distribution of stress levels has a power-law behaviour of the type $p(s) \propto 1/s$ on the interval $s \in [s_{\min}, s_{\max}]$ (figure 4(a), top panel, with the corresponding growth rate landscape $\mu(x,s)$ reported in the bottom panel). Results for this case are shown in figures 4(b)–(e). In brief, the optimal growth strategy $p^\star(x|s)$, which is uniform for sufficiently small β,

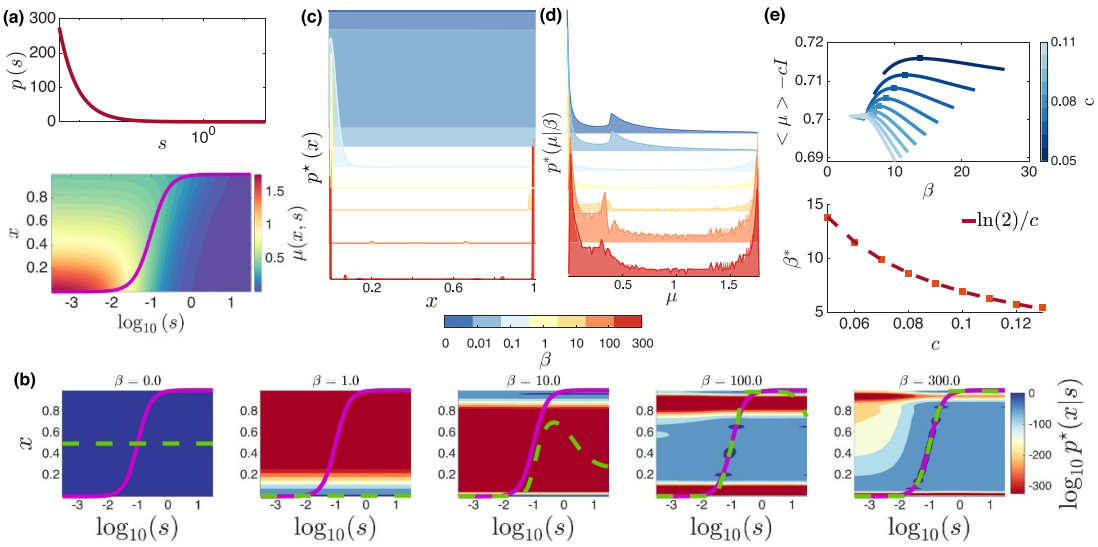

Figure 4. (a) Probability density of the stress level s in the 'power-law environment' (top) and corresponding fitness landscape (bottom). Here, $s_{\min} = 3 \cdot 10^{-4}$ and $s_{\max} = 3 \cdot 10^{1}$. Magenta line: growth-maximizing value of x for each s. (b) Optimal choice for $p^\star(x|s)$ for different values of β. Magenta line: growth-maximizing value of x as a function of s. Green line: mean value of x for the optimal choice rule at different values of s. (c) Marginal probability density $p^\star(x)$ for different values of β. (d) Growth rate distribution at optimality, $p^\star(\mu|\beta)$, for different values of β. (e, top) Overall metabolic objective function $\mathcal{F}$ (equation (18)) as a function of β for several value of the information cost-per-bit c. (e, bottom) Optimal value of β ($\beta^\star$) versus c: numerical results (markers) vs theoretical prediction (equation (19), dashed line).

matches the globally optimal one (i.e. $\widehat{x}(s)$, represented by the magenta line in figure 4(a)) more and more closely as β increases (figure 4(b)). The distribution of x at optimality (figure 4(c)) again underscores a non-trivial structuring of the population, which separates into well-defined groups of faster-growing fermenters ($x \simeq 0$) and slower-growing respirators ($x \simeq 1$). Remarkably, the separation already occurs effectively at values of β comparable with $1/\mu_{\max}$, with $\mu_{\max}$ the fastest growth rate achievable in the medium. At odds with the exponential case, however, the growth-rate distribution now displays bimodality robustly over β (see figure 4(d)). Finally, as in the exponential case, the value of β yielding the maximum fitness $\mathcal{F}$ decreases with the cost-per-bit c, and the maximum again appears to get flatter as c decreases (figure 4(e)), suggesting that limited resources may suffice to fine-tune β so as to achieve responses that are effectively optimal.

2.6. Two-state environments

Figure 5 displays results obtained for the case of a bimodal environment in which high-probability favourable conditions coexist with low-probability very adverse ones (figure 5(a)), as e.g. in experiments probing bacterial resistance to stress via antibiotic cycles [24]. Here, 95% of the distribution is concentrated around $s_0 = 0.05 < s_c$, whereas high stress levels take values around $s_1 = 0.8 > s_c$[3]. According to the growth-rate landscape (figure 5(a), bottom panel), while the low-cost strategy ($x = 0$) is favored in rich conditions, high-yield metabolism ($x = 1$) provides the optimal response under stress (magenta line in figure 5(a), bottom panel). Optimal adaptation to a strict bimodal environment therefore requires the ability to neatly separate the two phenotypes. Indeed, as β increases, the optimal strategy selection rule $p^\star(x|s)$ reproduces more and more closely the optimal response (figure 5(b)), while the population gradually acquires a structure in terms of x. As in previous cases, as soon as β is sufficiently large one observes the formation of two sub-populations (figure 5(c)): one formed by cells that use a mixed strategy with a stronger fermentative component (smaller x), the other formed by a small but robust group of strict respirators ($x \simeq 1$). This partitioning is reflected in turn in the distribution of growth rates (figure 5(d)), which displays a marked bimodal character even for large β, with lower growth rates associated with respirators. (It should however be noted that, in this environment, the background growth-rate landscape is effectively bimodal, as seen from the $\beta \to 0$ limit in figure 5(d).) These results paint a direct metabolic realization of the *bet-hedging* scenario, with respirators forming the group of 'persisters' [24].

Noticeably, the dependence on β in this environment appears to be stronger than in previous cases, suggesting that optimal adaptation to a two-state medium requires the encoding of a considerable amount of information about stress levels in the metabolic strategy. In view of this, it is reasonable to ask how the presence of a broad, low-probability background of stress levels along with the two 'peaks'

[3] The low-stress (high-stress) region is obtained by truncating in the interval $[0, 0.1]$ ($[0.7, 0.9]$) a normal density of mean 0.05 (0.8) and standard deviation 10^{-2} (10^{-2}). For $s \in (0.1, 0.7)$, $p(s) = 0$. In other terms, intermediate values of s are strictly prohibited.

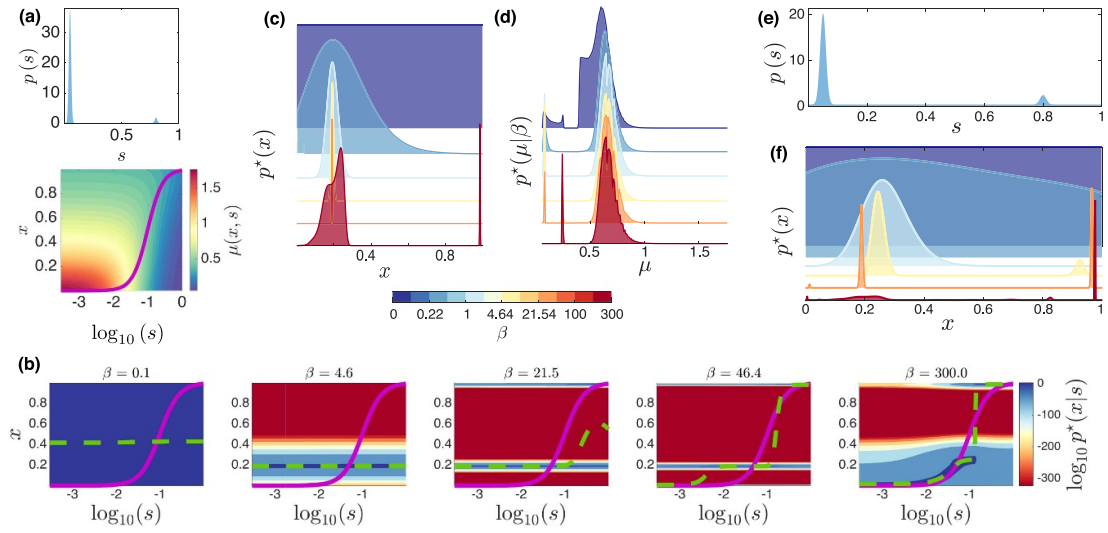

Figure 5. (a) Probability density of the stress level s in a two-state environment (top) and corresponding fitness landscape (bottom). Magenta line: growth-maximizing value of x for each s. (b) Optimal choice for $p^\star(x|s)$ for different values of β. Magenta line: growth-maximizing value of x for each s. Green line: mean value of x for the optimal choice rule at different as a function of s. (c) Marginal probability density $p^\star(x)$ for different values of β. (d) Growth rate distribution at optimality, $p^\star(\mu|\beta)$, for different values of β. (e) Alternative bimodal stress distribution with an extended background of stress levels with non-zero probability. (f) Corresponding growth rate distribution at optimality, $p^\star(\mu|\beta)$, for different values of β.

at low and high stress would change optimal adaptation. Remarkably, results shown in figure 5(e) suggest that this scenario mainly affects the faster-growing sub-population[4]. Indeed, as β grows, the optimal population tends to exploit intermediate values of x in a more and more refined way by depleting the sub-population at smaller values of x (figure 5(f), to be compared with figure 5(c)). In other terms, the presence of a robust sub-population of slower-growing respirators is crucial for optimal adaptation.

2.7. A closer look at optimal growth strategies

It is finally instructive to revisit the optimal phenotypic distributions found in the four types of environments discussed above in terms of the biologically relevant variables, namely the specific intake q and the specific proteome cost ϵ. The distribution of these quantities can be obtained numerically from the optimal growth strategy $p^\star(x|s)$ and the definitions (2) and (3) as

$$p^\star(q|\beta) = \int ds p(s) \int dx p^\star(x|s)\, \delta[q(x) - q], \quad (21)$$

$$p^\star(\epsilon|\beta) = \int ds p(s) \int dx p^\star(x|s)\, \delta[\epsilon(x) - \epsilon]. \quad (22)$$

Results are reported in figures 6(a)–(d). One clearly sees how, expectedly, optimal adaptation to a tightly controlled medium gradually concentrates the metabolic phenotypes around the growth-maximizing one as β increases (figure 6(a)). In all other (more complex) cases, however, the distributions of metabolic phenotypes effectively define clear sub-populations already for values of β comparable with $1/\mu_{\max}$. Based on the previous results, this process by itself can lead to an outcome close to the globally optimum one ($\beta \to \infty$) in an exponential environment (figure 6(b)) as well as in a power-law environment (figure 6(c)). Indeed in both cases phenotypic distributions do not change substantially as β is further increased. The bimodal environment again appears to be more sensitive to the value of β (i.e. to the amounts of resources available), as its corresponding phenotypic distribution undergoes considerable refinements as β gets larger and larger (figure 6(d)). In this respect, it is fair to say that bimodal media pose an especially complex adaptation problem for cells that strive to optimize growth.

3. Discussion

The standard theoretical framework to describe exponentially growing microbial populations mostly relies on the assumption that cells aim at maximizing their growth rate in any given medium. This idea should however be reconciled with the empirical fact that bacterial populations display remarkable cell-to-cell diversity even in terms of single-cell growth rates. Such a question is especially pressing in the context of metabolic networks, where optimality normally implies a maximal reduction of variability (e.g. as in flux-balance-analysis and related models). Here we attempt to bridge optimality and variability in metabolism through the fact that metabolic strategies need to encode information about the

[4] To obtain the probability density shown in figure 5(e) (top panel), we first removed the truncations used for figure 5(a), then added a uniform distribution in the interval [0, 1.0] multiplied by 0.8, and finally normalized the overall $p(s)$.

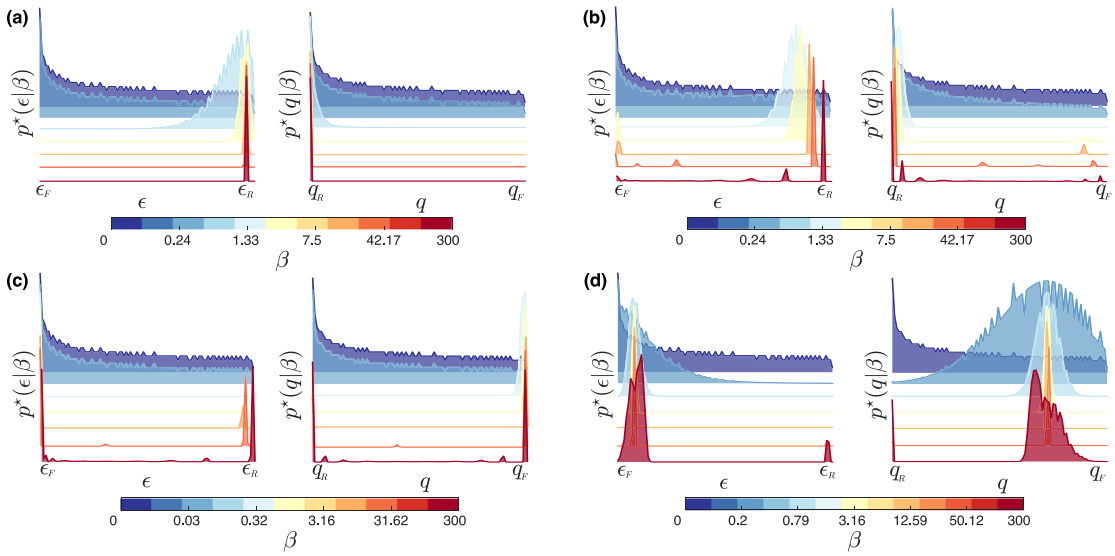

Figure 6. Numerical results for the probability densities of the specific proteome cost ϵ (left plot in each panel) and the specific uptake q (right plot in each panel) for different values of β. Panels (a)–(d) correspond respectively to the cases of a tightly controlled (a), exponential (b), power-law (c) or strict bimodal (d) stress distribution discussed in figures 2–5. The lower and upper bounds for ϵ and q are fixed to $\epsilon_F = 0.1$ and $q_R = 1$, and to $\epsilon_R = 1$ and $q_F = 10$, respectively.

growth medium. More precisely, optimal strategies for microbial growth in complex (random) media depend on the amount of information about the environment that can be encoded in the key metabolic variable(s). By framing this idea within a simple information-theory perspective, we have studied the optimal growth strategies for a coarse-grained but experiment-inspired model of growth metabolism subject to random stress levels. The key parameter $\beta \geqslant 0$ modulates, in essence, how precisely metabolic variables can be tuned in response to the stress levels or, equivalently for our purposes, the amount of cellular resources available to tune the metabolic response to the environmental conditions. By changing β, one passes from optimal strategies that are completely insensitive to the composition of the medium ($\beta = 0$) to strategies that are perfectly matched to the structure of the environment ($\beta \to \infty$). We have then analyzed the scenarios generated by these strategies in different types of environments. The general lesson is that, when cellular resources are limited, metabolic strategies for optimal growth require some degree of heterogeneity even in tightly controlled media. In addition, we find that, in realistic scenarios, outcomes close to those achievable with unlimited resources are often attained effectively with a modest amount of fine-tuning or, equivalently, with limited information about the environment. In other terms, optimal metabolic strategies for growth may be effectively robust with respect to the amounts of cellular resources available.

From a purely theoretical perspective, models addressing the adaptation of populations to unpredictable external conditions have been studied for several decades. In our view, despite lacking an explicit dynamics, our approach is conceptually related to that of [25], where heterogeneities in the phenotype distribution emerge as evolutionarily stable strategies to cope with fluctuating environments. Likewise, the 'exploration-exploitation' paradigm discussed in [17] leads to a similar outcome in terms of growth rate distributions, albeit in a highly stylized framework. A dynamical counterpart of the static scenario discussed here indeed shows that optimal adaptation can be defined even for dynamically varying stress levels (forthcoming). On the other hand, heterogeneity of growth rates (and indirectly of metabolic phenotypes) has been repeatedly observed in experiments probing actual cell populations subject to different stressors. Large variability in growth-rate distributions characterizes for instance microbial communities in fluctuating conditions [24, 26, 27], where heterogeneity induced e.g. by changes in environmental cues coexists with medium-independent phenotypic variability due to stochasticity in gene expression. By expanding the pool of genotypes present within a population, the latter plays a key role for its long-term adaptation [28]. Obtaining a direct quantitative picture of the the metabolic strategies of individual cells in a live population is instead experimentally challenging. Recent work integrating high-resolution techniques for local microenvironment sensing with statistical inference has however allowed to quantitatively assess the presence of different metabolic phenotypes in a population of cells that engineers its own environment [29]. Progress along these lines will hopefully shed more light on the distributions of biologically relevant variables in phenotypically heterogeneous populations.

The results we presented here suggest that, in any medium, fitness is maximized by an optimal distribution of phenotypes. To understand how close bacterial populations come to these optimal strategies, one needs to port our modeling framework to more realistic models of metabolic networks. It would be especially interesting, in our view, to compare the predictions based on optimal growth driven by information encoded about the environment with those obtained within the fitness-heterogeneity trade-off scenario of [11–13]. These approaches rely on the MaxEnt principle to characterize optimal phenotypic distributions in terms of a fitness-variability trade-off, and in turn provide a route to inferring how close real microbial populations are to optimality. So far, this issue has only been explored for populations living in tightly-controlled media [12]. Naturally, however, things may be environment-dependent. With proper data, and by integrating the current framework with the MaxEnt models introduced in [11], one could potentially establish how much of the observed variability is due to adaptation to complex environments versus stochastic bet-hedging.

A few other generalizations and open issues could be of further interest. First, we assumed the value of ν (equation (3)) to be given, representing a property of the metabolic and regulatory networks underlying the $q-\epsilon$ trade-off. Clearly, results depend on the choice of ν, albeit weakly if ν varies in a realistic range. Important and potentially insightful question however relate to (i) whether the value of ν can, in some sense, be optimized upon, and to (ii) how close empirical values of ν are to such optima. An information-theoretic framework similar to that employed here might be suited to tackle this issue. Second, in our formulation we have assumed that bacteria have knowledge of the ensemble of conditions in which they grow, i.e. of $p(s)$. This model can however be generalized to the case in which bacteria have to learn the ensemble over time. From an information theory perspective, a dynamical approach is likely to yield yet more insight into the idea of 'optimal bacterial growth'. Finally, it may be noted that our minimal model of growth completely disregards maintenance costs, namely the fact that cells can only grow when certain basal (and possibly growth-rate dependent) metabolic requirements are satisfied. While results presented here can be seen as the low-maintenance limit of this more general case, optimal strategies and growth rate distributions are bound to be affected by non-trivial maintenance costs, especially in heterogeneous media. A detailed study of these aspects would clarify the interplay between optimal growth, maintenance and stress levels in complex environments.

Data availability statement

The data that support the findings of this study are openly available at the following URL https://github.com/anna-pa-m/OptMetStrategy.

Appendix. Derivation of the optimal distribution

Our goal is to determine the conditional distribution $p(x|s)$ that maximizes the expectation value of the growth rate—computed with respect to the joint probability density of the degree of freedom x and of the stress s, i.e. $p(x,s)$– when the mutual information regarding the stress encoded in the conditional distribution is limited to a constant; that is, we want to solve

$$\max_{p(x|s)} \langle \mu \rangle \quad \text{subject to} \quad I = \text{constant}. \tag{A.1}$$

To this end, we introduce, within the formalism of Lagrange multipliers, the Lagrangian function

$$\begin{aligned}\mathcal{L}[p(x|s)] = &\int ds\, p(s) \int dx\, p(x|s)\, \mu(x,s) \\ &+ \gamma \left[\int ds\, p(s) \int dx\, p(x|s) \right. \\ &\left. \times [\log_2 p(x|s) - \log_2 p(x)] - I \right] \\ &+ \int ds \left[\lambda(s) \int dx\, p(x|s) - 1 \right],\end{aligned}$$

where γ is the Lagrange multiplier associated with the constraint on the mutual information between s and x, and $\lambda(s)$ ensures that for any value of s, the conditional distribution $p(x|s)$ is normalized to 1. Computing the functional derivative one gets

$$\begin{aligned}\frac{\delta \mathcal{L}}{\delta p(x|s)} = &\int ds' \int dx'\, \delta(x-x')\, \delta(s-s')\, p(s') \\ &\times \mu(x',s') + \gamma \left[\int ds' \int dx' p(s') \log_2 \right. \\ &\left. \times \frac{p(x'|s')}{p(x')} \delta(x-x')\, \delta(s-s') \right] \\ &+ \gamma \left[\int ds' \int dx' p(s')\, p(x'|s') \frac{1}{p(x'|s')} \right. \\ &\left. \times \delta(x-x')\, \delta(s-s') \right] \\ &- \gamma \left[\int ds' \int dx' p(s')\, p(x'|s') \frac{1}{p(x')} \frac{\delta p(x')}{\delta p(x|s)} \right] \\ &+ \int ds'\, dx'\, \lambda(s')\, \delta(s'-s)\, \delta(x-x').\end{aligned}$$

The remaining integrations involving Dirac δ-functions are easily performed, yielding

$$\frac{\delta \mathcal{L}}{\delta p(x|s)} = p(s)\,\mu(x,s) + \gamma\left[p(s)\log_2\frac{p(x|s)}{p(x)} + p(s) - I_0\right] + \lambda(s), \tag{A.2}$$

where

$$I_0 = \int ds' \int dx' p(s')\,p(x'|s')\,\frac{1}{p(x')}\,\frac{\delta p(x')}{\delta p(x|s)}. \tag{A.3}$$

Note however that $p(x) = \int ds\, p(x,s) = \int ds\, p(x|s)\,p(s)$. Therefore

$$\begin{aligned} I_0 &= \int ds' \int dx' p(s')\,p(x'|s')\,\frac{1}{p(x')}\int ds'' p(s'') \\ &\quad \times \delta(s''-s)\,\delta(x'-x) \\ &= p(s)\int dx'\frac{1}{p(x')}\int ds' p(s')\,p(x'|s')\,\delta(x-x') \\ &= p(s) \end{aligned}$$

Setting the derivative of the Lagrangian to zero leads to the condition

$$p(s)\left[\mu(x,s) + \gamma\log_2\frac{p(x|s)}{p(x)} + \gamma - \gamma\right] + \lambda(s) = 0, \tag{A.4}$$

from which one immediately gets

$$p^\star(x|s) \propto p^\star(x)\,e^{\beta\mu(x,s)+\beta\frac{\lambda(s)}{p(s)}}, \tag{A.5}$$

where here we have defined $\beta = -1/\gamma$. Upon imposing the normalization condition

$$\int dx\, p^\star(x|s) = 1 \qquad \forall s, \tag{A.6}$$

the value of $\lambda(s)$ can be straightforwardly determined to be

$$\lambda(s) = \beta^{-1}p(s)\log\left[\int dx\, p^\star(x)\,e^{\beta\mu(x,s)}\right]^{-1}. \tag{A.7}$$

Substituting this into (A.5) one arrives at equations (11) and (12), where $N(s,\beta) = \int dx\, p^\star(x)\,e^{\beta\mu(x,s)}$.

ORCID iD

Anna Paola Muntoni https://orcid.org/0000-0003-3937-3763

References

[1] Bruggeman F J, Planqué R, Molenaar D and Teusink B 2020 Searching for principles of microbial physiology *FEMS Microbiol. Rev.* **44** 821–44

[2] Jensen P R, Michelsen O and Westerhoff H V 1993 Control analysis of the dependence of escherichia coli physiology on the H(+)− ATPase *Proc. Natl Acad. Sci. USA* **90** 8068–72

[3] Bosdriesz E, Molenaar D, Teusink B and Bruggeman F J 2015 How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization *FEBS J.* **282** 2029–44

[4] Irwin P L, Nguyen L-H T, Paoli G C and Chen C-Y 2010 Evidence for a bimodal distribution of escherichia coli doubling times below a threshold initial cell concentration *BMC Microbiol.* **10** 1–4

[5] Kiviet D J, Nghe P, Walker N, Boulineau S, Sunderlikova V and Tans S J 2014 Stochasticity of metabolism and growth at the single-cell level *Nature* **514** 376–9

[6] Taheri-Araghi S, Bradde S, Sauls J, Hill N, Levin P, Paulsson J, Vergassola M and Jun S 2015 Cell-size control and homeostasis in bacteria *Curr. Biol.* **25** 385–91

[7] Kennard A S, Osella M J, Javer A, Grilli J, Nghe P, Tans S J, Cicuta P and Cosentino Lagomarsino M 2016 Individuality and universality in the growth-division laws of single *E. coli* cells *Phys. Rev.* E **93** 012408

[8] Jun S, Si F, Pugatch R and Scott M 2018 Fundamental principles in bacterial physiology: history, recent progress and the future with focus on cell size control: a review *Rep. Prog. Phys.* **81** 056601

[9] Scott M, Gunderson C W, Mateescu E M, Zhang Z and Hwa T 2010 Interdependence of cell growth and gene expression: origins and consequences *Science* **330** 1099–102

[10] Panlilio M, Grilli J, Tallarico G, Iuliani I, Sclavi B, Cicuta P and Cosentino Lagomarsino M 2021 Threshold accumulation of a constitutive protein explains *E. coli* cell-division behavior in nutrient upshifts *Proc. Natl Acad. Sci.* **118** e2016391118

[11] De Martino D, Capuani F and De Martino A 2016 Growth against entropy in bacterial metabolism: the phenotypic trade-off behind empirical growth rate distributions in *E. coli Phys. Biol.* **13** 036005

[12] Muntoni A P, Braunstein A, Pagnani A, De Martino D and De Martino A 2022 Relationship between fitness and heterogeneity in exponentially growing microbial populations *Biophys. J.* **121** 1919–30

[13] De Martino D, Andersson A , Bergmiller T, Guet C C and Tkačik G 2018 Statistical mechanics for metabolic networks during steady state growth *Nat. Commun.* **9** 2988–9

[14] De Martino D, Capuani F and De Martino A 2017 Quantifying the entropic cost of cellular growth control *Phys. Rev.* E **96** 010401

[15] Keren L, Van Dijk D , Weingarten-Gabbay S, Davidi D, Jona G, Weinberger A, Milo R and Segal E 2015 Noise in gene expression is coupled to growth rate *Genome Res.* **25** 1893–902

[16] Rivoire O and Leibler S 2011 The value of information for populations in varying environments *J. Stat. Phys.* **142** 1124–66

[17] De Martino A, Gueudre T and Miotto M 2019 Exploration-exploitation tradeoffs dictate the optimal distributions of phenotypes for populations subject to fitness fluctuations *Phys. Rev.* E **99** 012417

[18] Levien E, Min J, Kondev J and Amir A 2021 Non-genetic variability in microbial populations: survival strategy or nuisance? *Rep. Prog. Phys.* **84** 116601

[19] Dinis L, Unterberger J and Lacoste D 2022 Pareto-optimal trade-off for phenotypic switching of populations in a stochastic environment *J. Stat. Mech.* **2022** 053503

[20] Mori M, Marinari E and De Martino A 2019 A yield-cost tradeoff governs *Escherichia coli*'s decision between fermentation and respiration in carbon-limited growth *npj Sys. Biol. Appl.* **5** 1–9

[21] Bialek W 2012 *Biophysics: Searching for Principles* (Princeton, NJ: Princeton University Press)

[22] Mori M, Hwa T, Martin O C, De Martino A and Marinari E 2016 Constrained allocation flux balance analysis *PLoS Comp. Biol.* **12** e1004913

[23] Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson J R and Hwa T 2015 Overflow metabolism in *Escherichia coli* results from efficient proteome allocation *Nature* **528** 99–104

[24] Gefen O and Balaban N Q 2009 The importance of being persistent: heterogeneity of bacterial populations under antibiotic stress *FEMS Microbiol. Rev.* **33** 704–17
[25] Sasaki A and Ellner S 1995 The evolutionarily stable phenotype distribution in a random environment *Evolution* **49** 337–50
[26] Grimbergen A J, Siebring J, Solopova A and Kuipers O P 2015 Microbial bet-hedging: the power of being different *Curr. Opin. Microbiol.* **25** 67–72
[27] Morawska L P, Hernandez-Valdes J A and Kuipers O P 2022 Diversity of bet-hedging strategies in microbial communities—recent cases and insights *WIREs Mech. Dis.* **14** e1544
[28] Ackermann M 2015 A functional perspective on phenotypic heterogeneity in microorganisms *Nat. Rev. Microbiol.* **13** 497–508
[29] Onesto V *et al* 2023 Probing single-cell fermentation fluxes and exchange networks via pH-sensing hybrid nanofibers *ACS Nano* **17** 3313–23