

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

L.H.J.Krah*¹ and R. Hermsen¹

¹Dept. of Biology, Theoretical Biology and Bioinformatics, University of Utrecht

*corresponding author: l.h.j.krah@uu.nl

Submitted: August 28th, 2020

Abstract

In bacterial cells, protein expression is a highly stochastic process. At the same time, physiological variables such as the cellular growth rate also fluctuate significantly. A common intuition is that, due to their relatively high noise amplitudes, proteins with a low mean expression level are the most important causes of these fluctuations on a larger, physiological scale. Noise in highly expressed proteins, whose stochastic fluctuations are relatively small, is often ignored. In this work, we challenge this intuition by developing a theory that predicts the contribution of a protein's expression noise to the noise in the instantaneous, cellular growth rate. Using mathematical analysis, we decomposed the contribution of each protein into two factors: the noise amplitude of the protein, and the sensitivity of the growth rate to fluctuations in that protein's concentration. Next, we incorporated evolution, which has shaped the mean abundances of growth-related proteins to optimise the growth rate, causing protein abundances, but also cellular sensitivities to be non-random. We show that in cells that grow optimally fast, the growth rate is most sensitive to fluctuations in highly abundant proteins. This causes such proteins to overall contribute strongly to the noise in the growth-rate, despite their low noise levels. The results are confirmed in a stochastic toy model of cellular growth.

Keywords— Stochasticity, Gene Expression, Noise Propagation, Cellular Growth

Stochasticity is inherent to the nature of gene expression (39; 10; 30). Stochastic variation in the copy numbers of proteins is observed even under constant external conditions, and between individual cells in a population of isogenic bacteria. The amplitude of this noise in protein expression can be quantified in terms of the coefficient of variation (CV), defined as the standard deviation divided by the mean. Multiple studies, both theoretical and experimental, consider the scaling of the coefficient of variation squared (CV^2) with mean expression levels ($\mathbb{E}[X]$) (38; 1). For proteins with a low mean expression, noise is dominated by the intrinsic stochasticity (29; 35) of the chemical reactions involved and CV^2 scales as $1/\mathbb{E}[X]$ (33; 12; 5; 38; 10; 43). For higher mean expression, noise levels decrease and eventually reach a plateau, where fluctuations in gene expression are dominated

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

by extrinsic noise, such as division noise or environmental noise. Because of their larger noise levels, lowly expressed proteins are commonly assumed to be particularly important drivers of fluctuations observed in variables at the cellular level, such as the growth rate. At the same time the effects of the relatively small fluctuations of highly abundant proteins have largely been neglected.

Indeed, noisy gene expression is commonly accepted as the dominant mechanism behind the strong phenotypic divergence that has been observed in populations of genetically identical cells (31). In an exponentially growing population of cells, the growth rate of individual cells is distributed surprisingly broadly (37). Since cellular growth rate is often considered as an important proxy for bacterial fitness, the growth rate –and its fluctuations– have received much attention (16; 18; 15; 32).

In a recent study, correlations between a single protein’s expression level and the cellular growth rate have been explicitly measured for living *E. coli* cells. This showed that fluctuations in the protein’s concentration propagated, via their effect on metabolism, to the growth rate (19). In earlier work, our group proposed a mathematical framework to describe these experimental results. In it, noise in the expression of all proteins in the cell together caused cellular fluxes to fluctuate and, therewith, also propagated to the growth rate (20). We introduced so-called Growth Control Coefficients (GCCs) that quantify, to first order, the sensitivity of the cellular growth rate to changes in the expression of a particular protein species - how much “control” this protein species exerts over the growth rate. Using this theory and the underlying GCCs, we were able to reproduce the characteristics of the observed noise propagation (20).

Another line of research has focused on the remarkable ability of bacteria to tune their protein levels in order to grow, under many external conditions, at a near-optimal rate (3; 42; 7; 23; 34; 13). Optimal gene expression and growth has also been an important and fruitful assumption in countless modelling studies and techniques concerning deterministic growth, including Flux Balance Analysis (27; 6; 14). However, so far this observed optimality has never been taken into account when discussing noise propagation from protein expression to growth rate in single cells.

In this work, we therefore consider systems for which evolution has shaped protein expression to be optimal for fast growth. We then obtain predictions for the contribution of each protein to the noise in the growth rate as a function of its mean expression only. The main result, directly opposing common intuitions, is that proteins with a high mean expression are most important for the noise at the cellular level. The argument is, in short, that a protein’s contribution to noise in the growth rate does not only depend on the protein’s noise level, but also on the sensitivity of the growth rate to that protein’s fluctuations – its GCC. We show that when protein expression levels are optimised for growth, abundant proteins bear more growth control than lowly expressed ones. Consequently, noise in highly expressed protein species affects the growth rate strongly, causing them to contribute strongly to noise in the growth rate despite their low noise levels. Lastly, in a simple stochastic toy model of gene expression and cellular growth, we show these results to be robust for noise amplitudes that resemble those found inside living cells.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

Results

To grow, bacteria need to express a certain set of (metabolic) proteins. Together, these proteins create a metabolic flux used to build cellular components and new proteins. Since the cell's growth rate is limited by this metabolic flux, noise in the expression levels of the proteins involved propagates through the metabolic network to affect the growth rate (19). For any fixed external environment, we therefore assume the existence of an unknown function $\mu(\mathbf{X})$ that describes the instantaneous rate of cellular growth as a function of the copy numbers of all proteins, \mathbf{X} . The growth-rate is thus a deterministic function of stochastic variables.

To quantify how fluctuations in the expression of protein i affect growth rate μ , we use Growth Control Coefficients (20):

$$C_i^\mu \equiv \left(\frac{X_i}{\mu} \frac{\partial \mu}{\partial X_i} \right) \Big|_{\mathbb{E}[\mathbf{X}]} \quad (1)$$

As we will show below, these GCCs offer a way to decompose and analyse the noise in the growth rate in terms of contributions by each of the noisy components, the proteins.

To arrive at a comprehensive and useful noise decomposition, we adhere to two simplifying assumptions. First, noise levels are assumed to be small, so that all protein abundances are close to their means. The growth rate can then be approximated as a linear function of the protein levels. Secondly, fluctuations in all protein species are assumed to be independent. In bacterial cells, this is certainly not the case. However in the case of correlated fluctuations, noise contributions can not be uniquely defined (4; 24). Indeed, when two proteins correlate, and their joint fluctuations affect growth, the attribution of the noise contribution to either protein is arbitrary. Therefore, we here present the simplified case where all protein abundances are uncorrelated so that noise contributions can be uniquely defined and understood intuitively.

Under these assumptions noise in the growth rate, CV_μ^2 , can be approximated as a sum of contributions from all protein species:

$$CV_\mu^2 \approx \sum_i CV_i^2 (C_i^\mu)^2, \quad (2)$$

where CV_i is the coefficient of variation of the copy number of protein species i (see Appendix for the derivation).

In this equation, each protein's contribution consists of two factors. The first factor is no surprise: the proteins' coefficient of variation quantifies the fluctuations in that particular protein. The second factor is the protein's GCC, that quantifies how strongly these fluctuations actually affect the cellular growth rate.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

Distribution of Growth Control Coefficients

To further quantify which proteins are important for noise in the growth rate, we need to gain more insight into how growth control is distributed among proteins. This distribution is not arbitrary due to three properties of the GCCs, which are discussed below.

Sum Rule

Firstly, the sum of the GCCs equals zero (20):

$$\sum_i C_i^\mu = 0. \quad (3)$$

This sum rule originates from the so-called *intensity* of the growth rate: if all protein copy numbers inside a cell are increased by the same factor, the cellular flux increases, but the mass increases as well, such that the cell's growth rate (mass increase per mass) stays the same. That the growth rate is to a good approximation an intensive variable has been shown in multiple experiments (19; 37) and is a common modelling assumption (6). Moreover, it is analogous to the assumption that the metabolic flux and cellular mass are *extensive* variables, which is used in Metabolic Control Analysis to derive a similar sum rules for fluxes (17).

H-proteins

Secondly, there is a set of proteins, here called *H*-proteins, that are crucial for the cell's survival, but do not contribute to metabolism or cellular growth. This set *H* includes 'house-keeping proteins' participating in, *e.g.*, stress-response, immunity, and DNA damage repair; in bio-engineering, *H* may also contain engineered pathways. In wild type *E. coli*, the *H*-sector comprises an estimated 25 – 40% of the total protein mass (28).

Even when *H*-sector proteins are not toxic or otherwise harmful to the cell, their control on the growth rate will still be negative. Indeed, while they do not contribute to the metabolic flux, their synthesis does take up resources that otherwise could go to growth-related enzymes. The GCC of such *H*-sector proteins has earlier been calculated (20) to be equal to their mass fraction: $C_{i \in H}^\mu = -\phi_i$. Here we write $\phi_i := \mathbb{E}[X_i] / \sum_j \mathbb{E}[X_j]$ for the proteome mass fractions ϕ_i of each protein species *i* and ignore, for simplicity of notation, that different proteins have different masses. The total *H* sector is denoted as ϕ_H .

Together with the sum rule, the presence of the *H*-sector has important consequences for the distribution of GCCs: because some proteins have negative GCCs, others must have a positive GCC.

Optimal Growth

Lastly, natural selection tends to favour cells that on average grow faster, which drives the (mean) expression levels of all proteins to be (near)-optimal for growth (42; 7). We here show that this also affects the GCC of the optimised proteins.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

To do so, evolution is treated mathematically as a constrained optimisation problem, where the mean growth rate, $\mathbb{E}[\mu]$, is optimised under two constraints. First, the cell's density is kept constant. Second, only a fixed fraction of the proteome ($1 - \phi_H$) can be allocated towards proteins related to growth. To maximise the growth rate, only the protein abundances inside this fraction can be tuned by evolution while the total size must stay the same.

Using the method of Lagrange multipliers on a linearisation of μ , such an evolutionary process can indeed be shown to affect the distribution of growth control between metabolic proteins (for a formal derivation see Appendix). Intuitively, one can reason that in the optimal state all partial derivatives of the growth rate must be equal: if the growth rate would increase more upon increasing $\mathbb{E}[X_i]$ than upon increasing $\mathbb{E}[X_j]$, increasing the expression of i at the expense of j would increase the growth rate, and hence the growth rate would not be optimal. In terms of the GCCs, this observation translates to:

$$C_{i \notin H}^{\mu*} = \left(\frac{\phi_H}{1 - \phi_H} \right) \phi_i^*, \quad C_{i \in H}^{\mu} = -\phi_i. \quad (4)$$

Here, the asterisks indicate that the equality is only valid under optimality.

For cells optimised for growth, Eq. 4 reveals two important properties. First, growth control is shared between all metabolic proteins: there is no single growth-limiting protein. Secondly, and most importantly, enzymes with a higher mean expression level have a proportionally larger control on the growth rate.

Combining all factors

The use of these insights regarding the distribution of the GCCs in cells growing optimally, combined with the experimentally observed scaling of coefficients of variation of protein levels, results in a prediction for the contribution of each protein to the noise in the growth rate. Moreover, we can analyse how this contribution scales with $\mathbb{E}[X]$, the mean expression level of that particular protein. We write κ_i^P for the predicted noise contribution of protein species i , which is defined as the protein's relative contribution in Eq. 2:

$$\kappa_i^P := \frac{(C_i^\mu)^2 CV_i^2}{CV_\mu^2} \approx \frac{(C_i^\mu)^2 CV_i^2}{\sum_j (C_j^\mu)^2 CV_j^2} \quad (5)$$

Inspired by experimental data (38; 43), the intrinsic noise component is assumed to be inversely proportional to mean abundance:

$$CV_i^2 = F / \mathbb{E}[X_i], \quad (6)$$

with a fixed Fano Factor F . Ignoring the noise plateau caused by extrinsic noise sources, Eq. 6 then sets the noise levels of all protein species. Note that for the highly expressed proteins, noise levels are thus deliberately underestimated, resulting in a conservative estimate for their contribution to

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

noise in the growth rate. To now analyse noise propagation in optimally growing cells, Eqs. 4 and 6 are inserted in Eq. 5. This results in:

$$\kappa_i^{P^*} = \frac{\phi_H}{1 - \phi_H} \phi_i^* \quad (7)$$

This equation is the pivotal finding of this study. It predicts that, in cells whose expression levels are optimised for growth, proteins with a high mean expression contribute most strongly to fluctuations in the growth rate.

Stochastic Toy Model

We set out to test the above predictions in a toy model. To do so, we defined a highly simplified model of a growing cell with stochastic protein expression levels. To mimic the effects of evolution, we then employed an optimisation algorithm to search for the mean protein expression levels that optimise the mean growth rate of such cells. Next, we characterised the noise propagation in such optimised cells to verify the predictions of Eq. 4 and 7.

The model cell consists of a metabolic pathway of five reactions in a linear pathway that imports an external metabolite (m_1) and converts it to biomass (Fig. 1A). Each reaction is catalysed by a single enzyme species and inhibited by its own product. Additionally, a sixth protein species is expressed that is not metabolically active, representing the H -sector. Given the abundances \mathbf{X} of all proteins, the instantaneous growth rate μ is defined as the steady state flux through the pathway divided by the total number of expressed proteins, including the H -sector. Note that the growth rate depends non-linearly on all protein abundances. (For more details, see Appendix.)

The abundances \mathbf{X} themselves are stochastic: each X_i is distributed according to a Gamma distribution (12) characterised by mean $\mathbb{E}[X_i]$ and a Fano factor F . The Fano factor is chosen the same for all proteins, consistent with Eq. 6, and sets the overall noise amplitude in the cell.

It is useful to distinguish three levels of description of the model cells: their kinotype, genotype, and phenotype. We introduce the ‘kinotype’ as the set of reaction parameters that fully characterise the enzymes in a cell’s metabolic network: kinetic rates, Michaelis–Menten constants and inhibition parameters of the five reactions. We define a cell’s ‘genotype’ as the mean abundances of all protein species. Lastly, a cell’s ‘phenotype’ is given by the vector of the current protein abundances and the co-occurrent growth rate. The phenotype is therefore a stochastic, multidimensional variable whose probability distribution depends on the genotype.

The optimisation algorithm used to mimic evolution searches for the optimal genotype for a given and fixed kinotype. The optimal genotype is the one that generates phenotypes with, on average, the highest growth rate. During an evolutionary trajectory, the genotype is repeatedly subjected to mutations that are subsequently either rejected or accepted. Mutations increase or decrease the mean abundance of one particular protein species, after which the mean expression of all other metabolic proteins is adjusted such that the total cell density remains fixed ($\sum_i \mathbb{E}[X_i] = \Omega = 10^4$ in all simulations, for more details see Appendix). Mutations are accepted only if they increase the

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

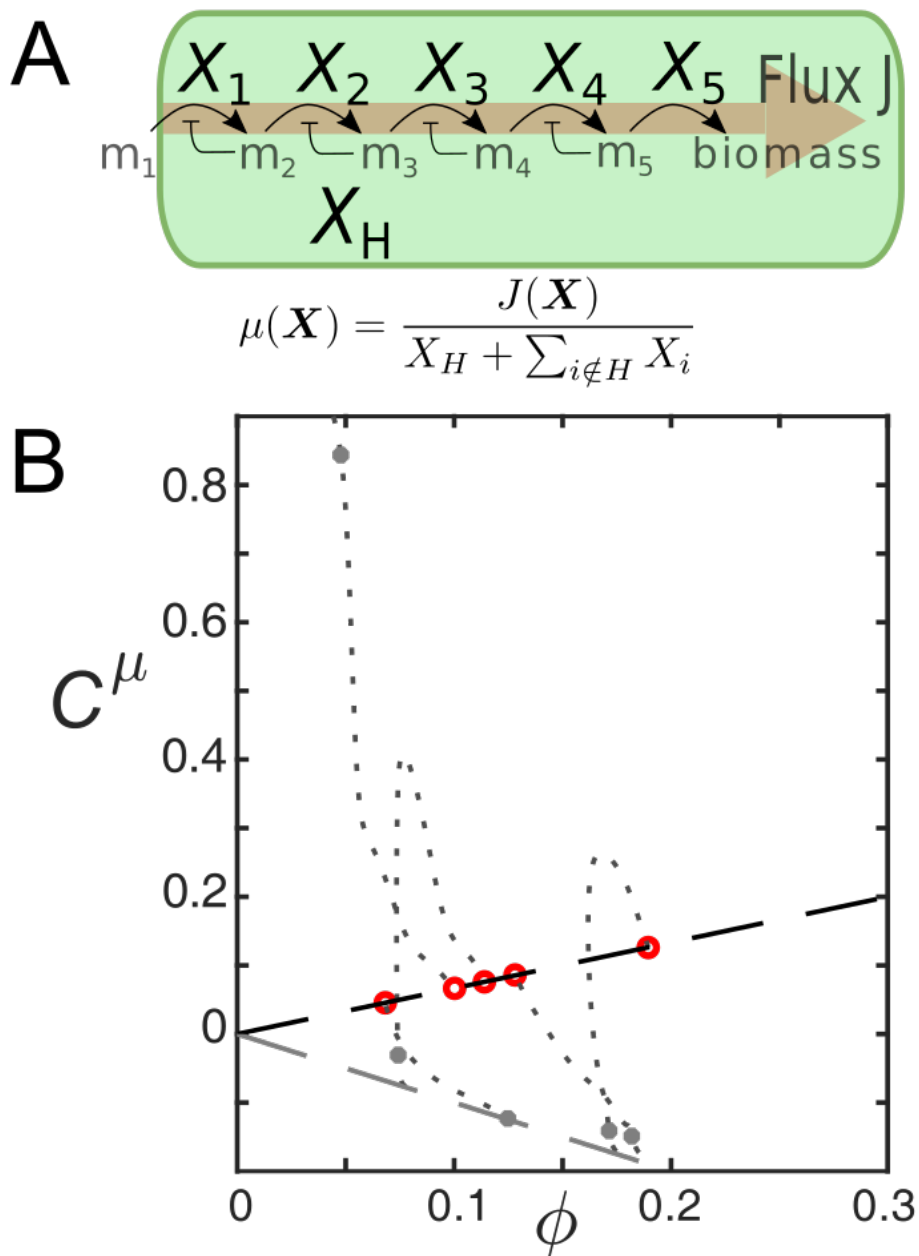


Figure 1: (A) Representation of the stochastic toy model. A cell expresses five metabolic protein species, each catalysing a single reaction in a linear reaction chain that imports and converts a fixed external metabolite, m_1 , into biomass. The growth rate is defined as the steady state flux through the network, divided by the total number of expressed proteins, including the H -sector protein. (B) Example trajectory during the optimisation of a single kinotype. Grey dotted lines represent the change of the GCCs during the optimisation process. Red point are the values of the GCCs in the optimal genotype, matching the predicted scaling (dashed line). See Tab. 1 for parameters.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

mean growth rate. The mean growth rate of a genotype is determined by sampling many phenotypes generated by that genotype.

After the optimal genotype has been found, the noise contribution of each protein species can be measured by sampling and analysing many phenotypes. Using a method adopted from (4) to measure these noise contributions, they can then be compared with our prediction, κ^P (Eq. 7).

Lastly, the whole process above was repeated for many kinotypes (randomly sampled; see Appendix), resulting in different optimal genotypes.

Low-Noise Regime

First we study the system in the low-noise regime (using $F = 1$, $\Omega = 10^4$), where our results (Eq. 2 and 7) are expected to hold exactly. When noise levels are this low, the mean growth rate $\mathbb{E}[\mu]$ is well approximated by the growth rate in the vector of mean abundances, $\mu(\mathbb{E}[\mathbf{X}])$. Instead of using an undirected, slow stochastic optimisation algorithm, the growth rate was therefore optimised with a deterministic gradient-based hill climb algorithm (for details see Appendix).

During each step of the optimisation process, we measured the GCCs of the metabolic proteins to observe how they adjust during optimisation. A representative example of such a trajectory is shown in Fig. 1B (dotted grey lines). Early in the optimisation process, when genotypes are still far from optimal, often one particular protein species is strongly limiting growth ($C^\mu \approx 1$, indicating that the growth rate could be improved by increasing this protein's expression level). In contrast, the expression of other metabolic proteins is too high; those proteins have a negative GCC, indicating that most of their expression is a burden to the cell. Eventually, as fitter genotypes are found, growth control becomes shared among all proteins (Fig. 1B, dotted grey lines). When eventually the optimal genotype is found, the predicted positive scaling between a protein's GCC and its mean abundance is obtained (Fig. 1B, red points, and Eq. 4).

Repeating the same process for multiple kinotypes ($n = 10$) confirms the generality of the positive scaling between C^μ and ϕ after optimisation (Fig. 2A, red points). Again, note that in an early stage of the optimisation process the distribution of the GCCs is markedly different (Fig. 2A, grey dots).

Next, we measured for each kinotype the noise contributions of each protein in the optimal genotype, and, for comparison, in a non-optimal genotype. For all kinotypes, the noise contributions in the optimal genotypes neatly follow our prediction (Fig. 2B, red point). In contrast, for non-optimal genotypes, noise contributions are dominated by only a few lowly expressed proteins (Fig. 2B, grey dots). The mean abundance of these proteins is too low, causing both their GCC and their CV to be large, resulting in a large noise contribution.

This analysis clearly highlights the fundamental role of evolution in shaping noise propagation properties: only in evolved cells that grow at an (almost) optimal rate a positive scaling exist between κ and ϕ .

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

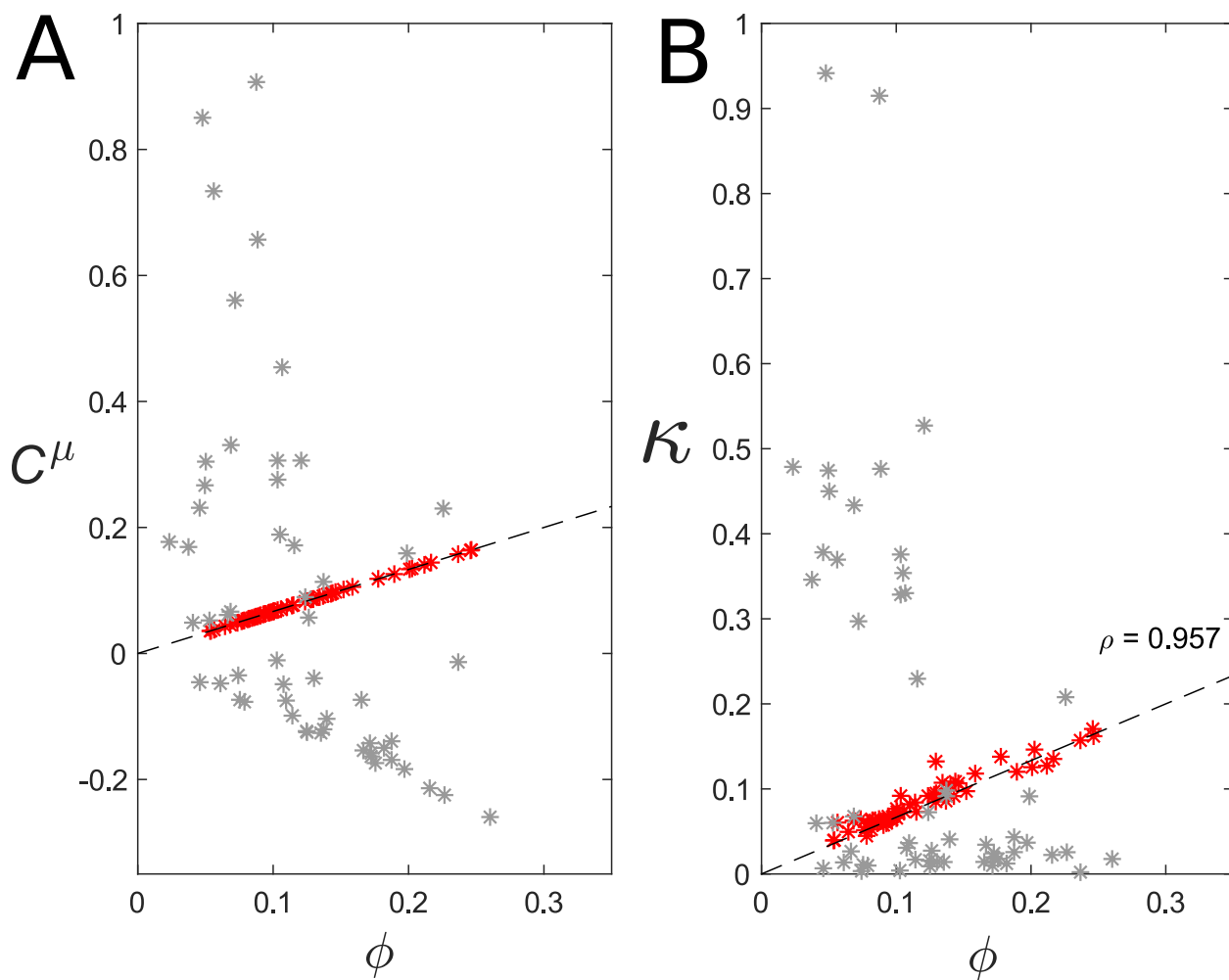


Figure 2: Perfect prediction in the case of linear noise for 10 different kinotypes. (A) Grey dots are GCCs in genotypes after 150 optimisation steps, which are not yet optimal. Red points are the GCCs in the optimal genotypes. Dashed lines are predictions for the values of GCCs for metabolic (positive) or H-proteins (negative) (B) Measured noise contributions compared to predicted noise contributions (dashed line). Red points are measured in the optimal genotype, grey points in the non-optimal genotypes.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

High-Noise Regime

Next we repeat the above analysis for larger noise levels that match those observed in living bacteria ($F = 10$, $\Omega = 10^4$, resulting in CVs up to 0.2, Fig. S4C). In this regime, the non-linearity of the growth rate becomes important and might affect which genotype is optimal. Below, we therefore distinguish two genotypes for each kinotype: the ‘Low-Noise’ (LN) genotype, which is optimal in the low-noise regime, and the ‘High-Noise’ (HN) genotype’, which evolved in the high-noise regime. To efficiently find the HN genotype of each kinotype, we used the LN genotypes as the starting conditions of the evolutionary algorithm.

For some kinotypes, the resulting HN genotypes indeed differed significantly from the LN ones. This can be understood as follows. When noise levels are low, control over the growth rate is shared between the proteins, and no single protein species is rate-limiting. When noise levels increase while the genotype is kept the same, large fluctuations and non-linear effects can increase the probability that in some phenotypes a single protein species becomes the sole rate-limiting step. Indeed, some genotypes that were optimal in the low-noise regime generated many slow-growing phenotypes when noise levels were high (Fig. S1B). The low growth rates were often caused by a single protein species whose phenotypic expression was too low (Fig. S1A) and therefore became a bottle-neck.

From an allocation point of view, increasing the mean expression of lowly abundant, rate-limiting proteins is cheap: few additional resources are needed to cause a relatively large effect. Indeed, the mean expression of potential bottlenecks increased during evolution, but mainly for lowly abundant protein species (Fig. S1B and Fig. S2).

Unsurprisingly, the distribution of the GCCs is also different in the evolved HN genotype (Fig. 3A, red points). GCCs are by definition linear measures and as soon as the non-linearity of the growth rate become relevant, Eq. 4 is not expected to hold exactly anymore. Interestingly, however, the positive scaling between ϕ and C^μ remains, and becomes, if anything, even steeper. Again, this makes sense: when noise levels increase, lowly expressed proteins are, due to their larger CV, more likely to fluctuate down to levels that strongly limit growth (Fig. S2A). Increasing the mean expression of these proteins will reduce their GCC (Fig. S2B-C), while increasing the GCC of the other proteins via the sum rule (Eq. 3). The net result is an increase in the slope in Figure 3A.

As growth control and noise propagation are strongly intertwined, much of the same reasoning can be applied to the noise contributions. The positive scaling between κ and a protein’s mean expression remained in the high-noise regime (Fig. 3B, red points). At the start of the evolutionary trajectory, when the genotype is still the LN genotype, some lowly abundant protein species are dominant noise contributors (Fig 3B, grey points top-left corner). During evolution those proteins obtained a slightly higher expression, reducing both their GCC and their noise contribution κ (Fig. 3B, red points and Fig. S2B-C), and at the same time increasing the contributions of the other proteins species.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

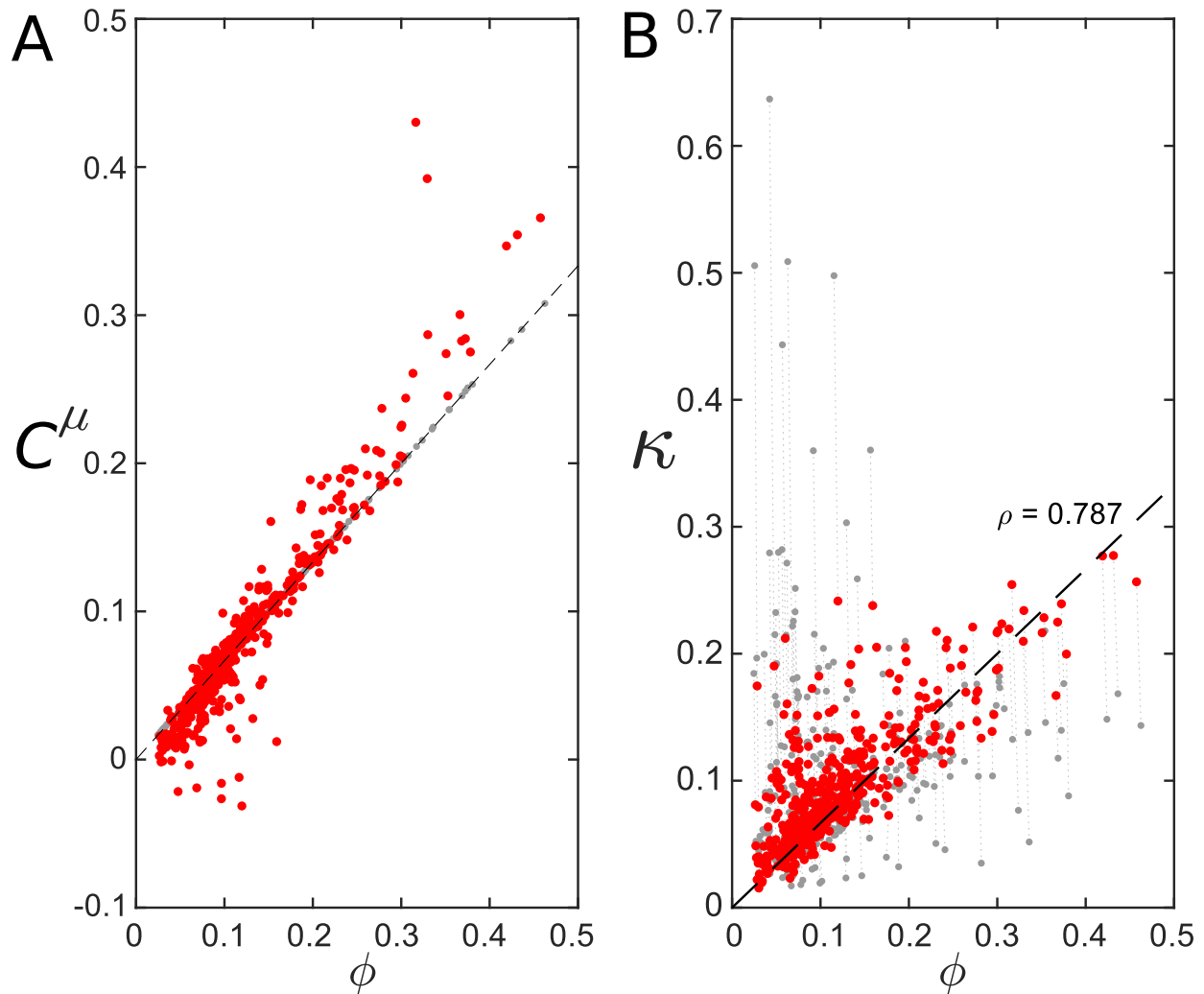


Figure 3: For 100 kinotypes, Low-noise genotypes (grey dots) compared with high-noise genotypes (red dots). Dashed lines are theoretical predictions for metabolic proteins. (A) Growth Control Coefficients. (B) Measured noise contributions.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

Discussion

In this article we argued that highly expressed proteins play an important role in system-level noise properties. In summary, our argument is that although abundant proteins have low noise levels, their GCCs become large when the cell is optimised for growth. The product of a protein's CV and GCC -an indicator for the protein's noise contribution- is predicted to be proportional to the protein's mean abundance. In a simplistic toy model of cellular growth, we show this proportionality to persist even when noise levels increase and the growth rate is no longer linear.

Although the toy model assumed particular enzyme dynamics, all mathematical predictions were derived without any assumptions concerning the underlying biochemistry. This implies that the strong contribution of highly expressed proteins to the noise in the growth rate is a general property of evolved biological systems.

Furthermore, while this paper only treated a metabolic and an H -sector, our analysis can be extended to a cell with additional proteomic sectors, such as the catabolic and ribosomal sectors(21), assuming that their allocation is regulated (26; 44) and they are independently optimised for their specific contribution towards growth (42). On an even smaller metabolic scale, the derived positive scaling between κ and ϕ will also be present in a single metabolic pathway as long as its protein abundances are optimised, while its total protein abundance is constrained.

The presence of cellular constraints is crucial for our results. In this study, a tight constraint was imposed by assuming that the H -sector is completely static. However, a small relaxation of this constraint -*e.g.*, assuming that the allocation towards the H -sector has to be within a certain range- should yield similar results. Also other allocation constraints, *e.g.* a maximum density of membrane proteins (36), will result in a similar positive scaling between κ and ϕ .

We point out that, besides the growth rate, other cellular traits, such as stress response or antibiotic resistance, are important for bacterial fitness as well. Interestingly, noise in these traits can be analysed in the same way as noise in the cellular growth rate. Therefore, the argument suggests that noise in any intensive trait that has been optimised during the bacteria's evolutionary history should be dominated by highly expressed proteins.

In this paper we focused on ranking the noise contribution of different protein species and ignored other noise sources in the cell. However, other noise sources, *e.g.* external fluctuations or division and partition noise, also contribute significantly to variance in system-level variables (25). Therefore, the total variance of the growth rate is underestimated in our simulations, preventing any quantitative predictions for the noise contributions.

Moreover, additional noise sources could cause protein abundances to become correlated (11; 25). Such correlations will remain a problem when discussing noise propagation from the protein-level to cellular growth, because as soon as protein abundances are correlated, noise contributions cannot be uniquely defined. One way to circumvent this problem is to use a fine-grained model description, *e.g.* at the level of individual chemical reactions, in which the noise sources are inherently uncorrelated (22; 41), or to adopt a meta-modelling approach (24). That said, the method

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

presented here, by directly relating noise contributions to natural abundances, allows for a more intuitive interpretation of noise contributions.

The results discussed above add to the realisation that global cellular constraints have intricate consequences for the overall physiology of evolved cells, from noisy gene expression (20), to metabolism (36; 14) and growth (44). It highlights the holistic nature of noise propagation via the sum rule for the GCCs (Eq. 3). The sum rule specifically could have important consequences for biotechnology: tinkering with a specific part of the cell affects noise propagation properties of the entire system. For example, most synthetic proteins or pathways do not contribute to growth, but instead create by-products. Such pathways will thus be part of the H-sector, and increase the GCC -and therewith the propagation- of all metabolic proteins.

The stochastic toy model moreover revealed a trade-off between efficient resource allocation and robust metabolism (8). Genotypes that are optimal in the low-noise regime allocate resources efficiently, but lack metabolic robustness at higher noise levels. Cells with different levels of expression noise therefore required different genotypes to grow, on average, the fastest (Fig. 3). Similar observations were also made in a recent experiment in yeast (32). Together, these observations can have consequences for FBA-like techniques, where optimal growth states have so far been calculated mostly deterministically, *i.e.*, optimising the growth rate in the mean expression levels.

We conclude that noise in gene expression -and its propagation towards the growth rate- needs to be considered when discussing optimal growth, but also vice versa: optimal enzyme expression affects noise propagation properties in such a way that abundant protein species become most relevant for noise on a system level.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

References

- [1] A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O’Shea, Y. Pilpel, and N. Barkai. Noise in protein expression scales with natural protein abundance. *Nature Genetics*, 38(6):636–643, June 2006.
- [2] J. Berkhout, E. Bosdriesz, E. Nikerel, D. Molenaar, D. de Ridder, B. Teusink, and F. J. Bruggeman. How biochemical constraints of cellular growth shape evolutionary adaptations in metabolism. *Genetics*, 194(2):505–512, June 2013.
- [3] E. Bosdriesz, D. Molenaar, B. Teusink, and F. J. Bruggeman. How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *The FEBS Journal*, 282(10):2029–2044, 2015. _eprint: <https://febs.onlinelibrary.wiley.com/doi/pdf/10.1111/febs.13258>.
- [4] C. G. Bowsher and P. S. Swain. Identifying sources of variation and the flow of information in biochemical networks. *Proceedings of the National Academy of Sciences*, 109(20):E1320–E1328, May 2012.
- [5] L. Cai, N. Friedman, and X. S. Xie. Stochastic protein expression in individual cells at the single molecule level. *Nature*, 440(7082):358–362, Mar. 2006.
- [6] H. de Jong, S. Casagrande, N. Giordano, E. Cinquemani, D. Ropers, J. Geiselman, and J.-L. Gouzé. Mathematical modelling of microbes: metabolism, gene expression and growth. *Journal of The Royal Society Interface*, 14(136):20170502, Nov. 2017.
- [7] E. Dekel and U. Alon. Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436(7050):588–592, July 2005.
- [8] S. Donati, T. Sander, and H. Link. Crosstalk between transcription and metabolism: how much enzyme is enough for a cell? *WIREs Systems Biology and Medicine*, 10(1):e1396, 2018. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/wsbm.1396>.
- [9] H. Dourado and M. J. Lercher. An analytical theory of balanced cellular growth. *Nature Communications*, 11(1):1–14, Mar. 2020. Number: 1 Publisher: Nature Publishing Group.
- [10] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain. Stochastic Gene Expression in a Single Cell. *Science*, 297(5584):1183–1186, Aug. 2002.
- [11] J. Fernandez-de Cossio-Diaz, R. Mulet, and A. Vazquez. Cell population heterogeneity driven by stochastic partition and growth optimality. *Scientific Reports*, 9(1):1–7, June 2019. Number: 1 Publisher: Nature Publishing Group.
- [12] N. Friedman, L. Cai, and X. S. Xie. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Physical Review Letters*, 97(16):168302, Oct. 2006.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

- [13] A. Goelzer, V. Fromion, and G. Scorletti. Cell design in bacteria as a convex optimization problem. *Automatica*, 47(6):1210–1218, June 2011.
- [14] D. H. d. Groot, C. v. Boxtel, R. Planqué, F. J. Bruggeman, and B. Teusink. The number of active metabolic pathways is bounded by the number of cellular constraints at maximal metabolic rates. *PLOS Computational Biology*, 15(3):e1006858, Mar. 2019. Publisher: Public Library of Science.
- [15] M. Hashimoto, T. Nozoe, H. Nakaoka, R. Okura, S. Akiyoshi, K. Kaneko, E. Kussell, and Y. Wakamoto. Noise-driven growth rate gain in clonal cellular populations. *Proceedings of the National Academy of Sciences*, 113(12):3251–3256, Mar. 2016.
- [16] S. Iyer-Biswas, C. S. Wright, J. T. Henry, K. Lo, S. Burov, Y. Lin, G. E. Crooks, S. Crosson, A. R. Dinner, and N. F. Scherer. Scaling laws governing stochastic growth and division of single bacterial cells. *Proceedings of the National Academy of Sciences*, 111(45):15912–15917, Nov. 2014.
- [17] H. Kacser, J. A. Burns, H. Kacser, and D. A. Fell. The control of flux. *Biochemical Society Transactions*, 23(2):341–366, May 1995.
- [18] A. S. Kennard. Individuality and universality in the growth-division laws of single *E. coli* cells. *Physical Review E*, 93(1), 2016.
- [19] D. J. Kiviet, P. Nghe, N. Walker, S. Boulineau, V. Sunderlikova, and S. J. Tans. Stochasticity of metabolism and growth at the single-cell level. *Nature*, 514(7522):376–379, Oct. 2014.
- [20] I. T. Kleijn, L. H. J. Krah, and R. Hermsen. Noise propagation in an integrated model of bacterial gene expression and growth. *PLOS Computational Biology*, 14(10):e1006386, Oct. 2018. Publisher: Public Library of Science.
- [21] S. Klumpp, Z. Zhang, and T. Hwa. Growth-rate dependent global effects on gene expression in bacteria. *Cell*, 139(7):1366, Dec. 2009.
- [22] M. Komorowski, J. Miękisz, and M. H. Stumpf. Decomposing Noise in Biochemical Signaling Systems Highlights the Role of Protein Degradation. *Biophysical Journal*, 104(8):1783–1793, Apr. 2013.
- [23] N. E. Lewis, K. K. Hixson, T. M. Conrad, J. A. Lerman, P. Charusanti, A. D. Polpitiya, J. N. Adkins, G. Schramm, S. O. Purvine, D. Lopez-Ferrer, K. K. Weitz, R. Eils, R. König, R. D. Smith, and B. Palsson. Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Molecular Systems Biology*, 6:390, July 2010.
- [24] G. Li, H. Rabitz, P. E. Yelvington, O. O. Oluwole, F. Bacon, C. E. Kolb, and J. Schoendorf. Global Sensitivity Analysis for Systems with Independent and/or Correlated Inputs. *The Journal of Physical Chemistry A*, 114(19):6022–6032, May 2010. Publisher: American Chemical Society.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

- [25] S. Modi, C. A. Vargas-Garcia, K. R. Ghusinga, and A. Singh. Analysis of Noise Mechanisms in Cell-Size Control. *Biophysical Journal*, 112(11):2408–2418, June 2017.
- [26] D. Molenaar, R. v. Berlo, D. d. Ridder, and B. Teusink. Shifts in growth strategies reflect tradeoffs in cellular economics. *Molecular Systems Biology*, 5(1):323, Jan. 2009.
- [27] M. Mori, T. Hwa, O. C. Martin, A. De Martino, and E. Marinari. Constrained Allocation Flux Balance Analysis. *PLoS computational biology*, 12(6):e1004913, 2016.
- [28] E. J. O’Brien, J. Utrilla, and B. O. Palsson. Quantification and Classification of E. coli Proteome Utilization and Unused Protein Costs across Environments. *PLOS Computational Biology*, 12(6):e1004998, June 2016.
- [29] J. Paulsson. Summing up the noise in gene networks. *Nature*, 427(6973):415–418, Jan. 2004.
- [30] A. Raj and A. van Oudenaarden. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell*, 135(2):216–226, Oct. 2008.
- [31] J. M. Raser and E. K. O’Shea. Noise in Gene Expression: Origins, Consequences, and Control. *Science*, 309(5743):2010–2013, 2005. Publisher: American Association for the Advancement of Science.
- [32] J. M. Schmiedel, L. B. Carey, and B. Lehner. Empirical mean-noise fitness landscapes reveal the fitness impact of gene expression noise. *Nature Communications*, 10(1):1–12, July 2019. Number: 1 Publisher: Nature Publishing Group.
- [33] V. Shahrezaei and S. Marguerat. Connecting growth with gene expression: of noise and numbers. *Current Opinion in Microbiology*, 25:127–135, June 2015.
- [34] C. Solem, D. Petranovic, B. Koebmann, I. Mijakovic, and P. R. Jensen. Phosphoglycerate Mutase Is a Highly Efficient Enzyme without Flux Control in *Lactococcus lactis*. *Journal of Molecular Microbiology and Biotechnology*, 18(3):174–180, 2010. Publisher: Karger Publishers.
- [35] P. S. Swain, M. B. Elowitz, and E. D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800, Oct. 2002. Publisher: National Academy of Sciences Section: Biological Sciences.
- [36] M. Szenk, K. A. Dill, and A. M. R. de Graff. Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Systems*, 5(2):95–104, Aug. 2017.
- [37] S. Taheri-Araghi, S. Bradde, J. T. Sauls, N. S. Hill, P. A. Levin, J. Paulsson, M. Vergassola, and S. Jun. Cell-size control and homeostasis in bacteria. *Current biology: CB*, 25(3):385–391, Feb. 2015.
- [38] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, 329(5991):533–538, July 2010.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

- [39] M. Thattai and A. v. Oudenaarden. Stochastic Gene Expression in Fluctuating Environments. *Genetics*, 167(1):523–530, Jan. 2004.
- [40] P. Thomas, H. Matuschek, and R. Grima. Intrinsic Noise Analyzer: A Software Package for the Exploration of Stochastic Biochemical Kinetics Using the System Size Expansion. *PLOS ONE*, 7(6):e38518, June 2012.
- [41] P. Thomas, G. Terradot, V. Danos, and A. Y. Weiße. Sources, propagation and consequences of stochasticity in cellular growth. *Nature Communications*, 9(1):1–11, Oct. 2018.
- [42] B. D. Towbin, Y. Korem, A. Bren, S. Doron, R. Sorek, and U. Alon. Optimality and sub-optimality in a bacterial growth law. *Nature Communications*, 8(1):1–8, Jan. 2017.
- [43] L. Wolf, O. K. Silander, and E. van Nimwegen. Expression noise facilitates the evolution of gene regulation. *eLife*, 4:e05856, June 2015. Publisher: eLife Sciences Publications, Ltd.
- [44] C. You, H. Okano, S. Hui, Z. Zhang, M. Kim, C. W. Gunderson, Y.-P. Wang, P. Lenz, D. Yan, and T. Hwa. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306, Aug. 2013.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

Appendix and Supplementary Figures

Derivation of Eq. 7

Here we derive equations 2 and (therewith) 7.

Linearisation of μ

Let the growth rate μ be an *intensive* function of the protein copy numbers \mathbf{X} . Assuming that protein copy numbers are close to some linearisation point $\tilde{\mathbf{X}}$, we can write to first order:

$$\mu(\mathbf{X}) \approx \mu(\tilde{\mathbf{X}}) \left(1 + \sum_i \left(\frac{X_i}{\mu} \frac{\partial \mu}{\partial X_i} \right) \Big|_{\tilde{\mathbf{X}}} \left(\frac{X_i - \tilde{X}_i}{\tilde{X}_i} \right) \right) \quad (8)$$

Decomposition of CV_μ

When the noise in protein copy numbers is uncorrelated, it is straightforward to calculate CV_μ . First, using the definition of the Growth Control Coefficients of Eq. 1 we can simplify equation 8 by setting $\tilde{\mathbf{X}}$ equal to the mean expression levels, $\mathbb{E}[\mathbf{X}]$:

$$\mu(\mathbf{X}) \approx \mu(\mathbb{E}[\mathbf{X}]) \left(1 + \sum_i C_i^\mu \left(\frac{X_i - \mathbb{E}[X_i]}{\mathbb{E}[X_i]} \right) \right). \quad (9)$$

Then, we use the basic properties of the variance, *i.e.*, for any two uncorrelated stochastic variables Y_1 and Y_2 , and scalars a, b , $\text{Var}\{aY_1 + b + Y_2\} = a^2 \text{Var}\{Y_1\} + \text{Var}\{Y_2\}$, to write:

$$\text{Var}\{\mu\} \approx \mu(\mathbb{E}[\mathbf{X}])^2 \left(\sum_i (C_i^\mu)^2 \frac{\text{Var}\{X_i\}}{\mathbb{E}[X_i]^2} \right). \quad (10)$$

In the regime where Eq. 8 holds, we can also read from Eq. 9 that $\mathbb{E}[\mu] \approx \mu(\mathbb{E}[\mathbf{X}])$. This allows us to further simplify Eq. 10:

$$\frac{\text{Var}\{\mu\}}{\mathbb{E}[\mu]^2} =: CV_\mu^2 \approx \sum_i (C_i^\mu)^2 CV_i^2, \quad (11)$$

which is Eq. 2 (note that it's also a special case of (40)).

Optimisation

Next, we optimise the *mean* expression levels of the metabolic proteins to achieve the maximal *mean* growth rate. We assume that a fixed number of proteins has to be allocated for other things than

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

growth (the H -sector proteins), and constrain the average total cell mass $\sum_i \mathbb{E}[X_i]$ to a fixed value Ω . The optimisation can then be written as maximising $\mathbb{E}[\mu]$ over $\mathbb{E}[\mathbf{X}_{\neq H}]$, with the constraint $\sum_i \mathbb{E}[X_i] = \Omega$:

$$\text{Max}_{\mathbb{E}[\mathbf{X}_{\neq H}]} \left[\mathbb{E}[\mu] \mid \sum_{i \notin H} \mathbb{E}[X_i] = \Omega - \mathbb{E}[X_H] \right].$$

To find the optimal genotype, we use the Lagrange Multiplier method, which in this case takes the form

$$\mathcal{L} = \mathbb{E}[\mu] - \lambda \left(\sum_{i \notin H} \mathbb{E}[X_i] - (\Omega - \mathbb{E}[X_H]) \right),$$

where we need to solve $\nabla \mathcal{L} = 0$. This gives:

$$\left. \frac{\partial \mathbb{E}[\mu]}{\partial \mathbb{E}[X_i]} \right|_{\mathbb{E}[\mathbf{X}]^*} = \lambda, \text{ for all } i \notin H, \quad (12)$$

$$\sum_{i \notin H} \mathbb{E}[X_i]^* = \Omega - \mathbb{E}[X_H]. \quad (13)$$

Here the asterisks indicate the optimal value of that variable. Using Eq. 8 we can calculate $\mathbb{E}[\mu]$ explicitly:

$$\mathbb{E}[\mu] \approx \mu(\tilde{\mathbf{X}}) \left(1 + \sum_i \left(\frac{X_i}{\mu} \frac{\partial \mu}{\partial X_i} \right) \bigg|_{\tilde{\mathbf{X}}} \frac{\mathbb{E}[X_i] - \tilde{X}_i}{\tilde{X}_i} \right),$$

which allows us to calculate the partial derivatives of Eq. 12:

$$\frac{\partial \mathbb{E}[\mu]}{\partial \mathbb{E}[X_i]} = \mu(\tilde{\mathbf{X}}) \left(\frac{X_i}{\mu} \frac{\partial \mu}{\partial X_i} \right) \bigg|_{\tilde{\mathbf{X}}} / \tilde{X}_i, \text{ for all } i \notin H.$$

Next, we make a particular choice for the point of linearisation $\tilde{\mathbf{X}}$ by setting it equal to optimal genotype, $\mathbb{E}[\mathbf{X}]^*$. In this vector, the partial derivatives equal λ (Eq. 12):

$$\lambda = \left. \frac{\partial \mathbb{E}[\mu]}{\partial \mathbb{E}[X_i]} \right|_{\mathbb{E}[\mathbf{X}]^*} = \frac{\mu^* C_i^{\mu^*}}{\mathbb{E}[X_i]^*} \quad (14)$$

Here, the asterisk denotes that the equation is only valid at the vector with optimal mean expression levels. (Similar arguments were exploited in (2) and (9), although without the inclusion of the H -sector and stochastic gene expression.) Lastly, we can use Eq. 13, together with the sum rule for

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

the GCCs (Eq. 3), to calculate λ explicitly:

$$\begin{aligned} 0 &= \sum_i C_i^{\mu^*} = \sum_{i \notin H} \frac{\lambda}{\mu^*} \mathbb{E}[X_i]^* + \sum_{i \in H} C_i^{\mu^*} \\ &= \frac{\lambda}{\mu^*} (\Omega - \mathbb{E}[X_H]) - \frac{\mathbb{E}[X_H]}{\Omega}, \end{aligned}$$

and thus

$$\lambda = \frac{\mu^* \mathbb{E}[X_H]}{\Omega(\Omega - \mathbb{E}[X_H])} = \frac{\phi_H}{\Omega(1 - \phi_H)}, \quad (15)$$

using to the definition of $\phi_i := \mathbb{E}[X_i] / \Omega$. Finally, by combining Eq. 14 and 15 we find the Growth Control Coefficient for all proteins that do not belong to the H -sector in the optimal growth state:

$$C_i^{\mu^*} = \frac{\phi_H}{(1 - \phi_H)} \phi_i^* \quad (16)$$

Stochastic Simulation

In our toy model, we simulated a linear chain of 5 proteins, where the flux through the first 4 steps is given by:

$$v_i = \frac{k_{\text{cat},i} X_i m_i}{k_m + m_i + m_{i+1}/k_{i+1}} \quad (17)$$

Here, k_{cat} is the reaction rate, k_m is the Michaelis-Menten constant, and k is the inhibition constant. To define a kinotype, we sample k_{cat} and k_m uniformly from the interval $[0.1, 6.1]$ and k from $[1, 7]$. The external metabolite, m_1 is set to 10 and kept constant, simulating a fixed environment. The fifth protein in the reaction chain creates biomass and is not inhibited. The initial genotype is constructed by sampling 5 uniform numbers, $\mathbf{x} \in [0, 1]^5$ and setting $\mathbb{E}[\mathbf{X}]_{\text{initial}} = \mathbf{x} \phi_H / \sum_j x_j \Omega$, with $\phi_H = 0.4$ (the proteomic fraction allocated to the H -sector) and $\Omega = 10^4$, the (mean) total cell size.

Genotypes were mutated in two different ways, depending on the noise amplitude chosen in the simulation. In the low noise regime, the next copy numbers (of all but the H -protein) were determined with a gradient-based hill climbing algorithm that uses the current GCCs:

$$\mathbf{X}_{t+1} = \mathbf{X}_t + \epsilon \delta \mathbf{X}_t / |\delta \mathbf{X}_t|^2, \quad \delta \mathbf{X}_t = \frac{\Omega}{5} \left(\mathbf{C}^{\mu_t} / \phi_t - \sum_{j \notin H} C_{j,t}^{\mu} / \phi_{j,t} \right),$$

where ϵ is small (0.0002). This algorithm changes the genotype in the direction of the steepest growth-rate increase. In the high noise regime, changes in the genotype are stochastic, where one of the metabolic proteins is chosen at random and its mean abundance is changed according to a normal distributed percentage (mean zero, variance 5%), after which the entire genotype is renormalised to enforce a fixed mean cell size of Ω . A mutant genotype is only accepted if it yields a higher mean

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

growth rate. The evolutionary process is terminated when 100 mutants have been declined and repeated 15 times. From these 15 evolutionary trajectories, the genotype with the highest mean growth rate was chosen to be the HN genotype.

A phenotype is then sampled according to six independent Gamma distributions, with means defined by the genotype ($\mathbb{E}[\mathbf{X}]$ for the metabolic proteins and $\phi_H\Omega$ for the H -sector) and variance $F\mathbb{E}[X]$. Note that although the mean expression levels are always equal to Ω , the total number of proteins expressed in a particular phenotype is not.

To next calculate the growth rate for a specific phenotype, we then integrate the system of ODEs for all metabolites (for $i \in \{2, 3, 4, 5\}$, $\dot{m}_i = v_{i-1} - v_i = 0$) until steady state (10^6 time steps in Matlab 2019b). Note that the kinetics chosen (Eq. 17) enforce that for all possible phenotypes a steady state exists. In the steady state, where all fluxes are equal, we define $v_i = v_{i-1} = J$, and:

$$\mu := J / \left(\sum_i X_i \right)$$

After sampling $2 \cdot 10^4$ phenotypes, we measured noise contributions as follows. First, we divided the distribution for each protein in 100 bins, and calculated the mean growth rate in each bin (sampling extra if less than 100 phenotypes fell in a particular bin). Afterwards, we calculated the weighted variance between these growth rates. This is an approximation of a conceptual decomposition method from Swain *et al.* (4):

$$\kappa_i = \frac{\text{Var} \{ \mathbb{E} [\mu | X_i] \}}{\text{Var} \{ \mu \}}. \quad (18)$$

This method is a first order approximation of a full Global Sensitivity Analysis (24). The approximation is only valid if the sum of all contributions is close to unity. This is indeed the case (see Fig. S4A). Moreover, sampling errors in κ are small (Fig. S4B). All Matlab codes are available upon request.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

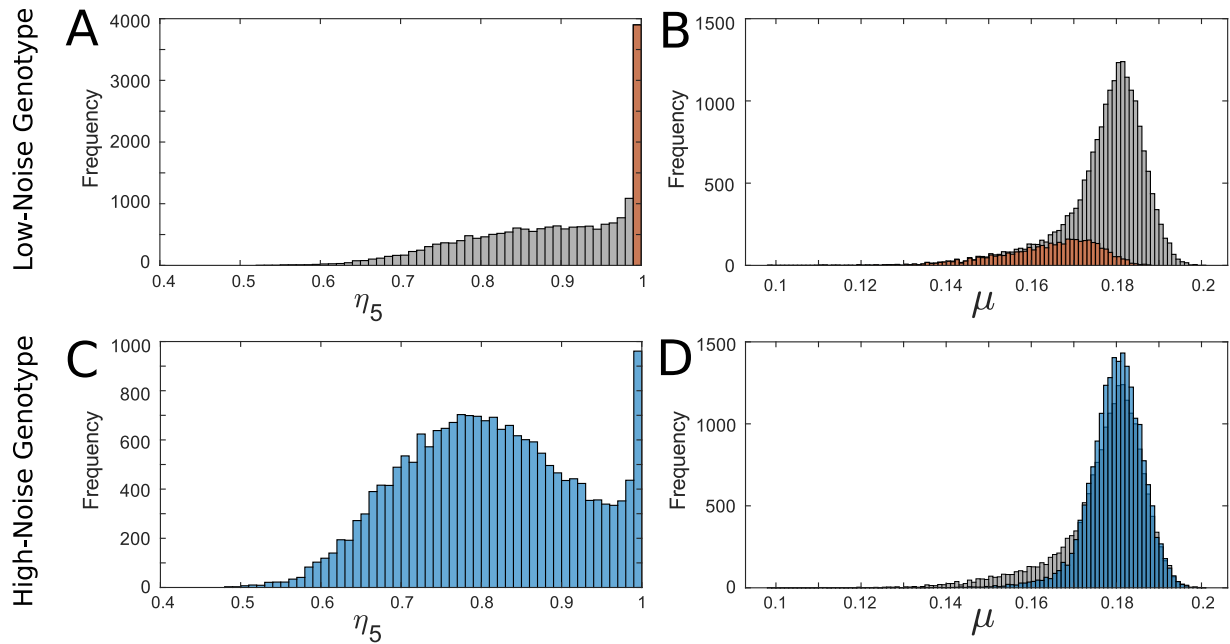


Figure S1: Example of highly skewed distributions in an optimised kinotype. (We picked this kinotype because it most clearly showed the effect of evolution on the distributions.) **(A)** Distribution of the efficiency of the fifth protein ($\eta_i := \frac{k_{\text{cat},i} X_i}{\mu \sum_j X_j}$). **(B)** Distribution of growth rates for the optimal Low-Noise genotype in the high noise regime. Red areas in (A) and (B) correspond to the same phenotypes. **(C)** The same as (A), but for the evolved High-Noise genotype. **(D)** the same as (B), but for the evolved, High-Noise genotype. Grey distribution is the distribution in the LN-genotype for comparison.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

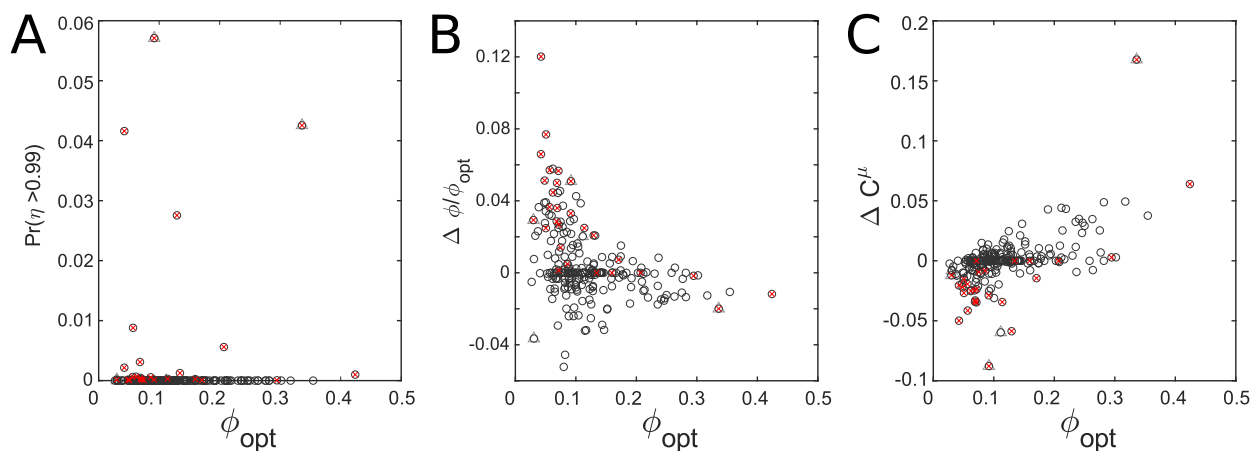


Figure S2: **(A)** Probability a protein's efficiency is very close to unity, calculated over $2 \cdot 10^4$ phenotypes. Although a higher efficiency on first glance seems good, an efficiency close to unity indicates that this protein might be limiting growth. All proteins for with $p(\eta > 0.99) > 0$ are marked with a red cross. Two outliers (triangle) protein species from the same kinotype. **(B)** Changes in genotype when noise levels increase, as a function of the mean ϕ in the optimal genotype. Red crosses are those proteins that had a non-zero probability of getting an efficiency close to one in the optimal genotype with high noise. Note that the highly expressed proteins with $p(\eta > 0.99) > 0$ are not increased, probably because this required the allocation too much additional resource. **(C)** Changes in mean genotype coincide with a change in C^μ .

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

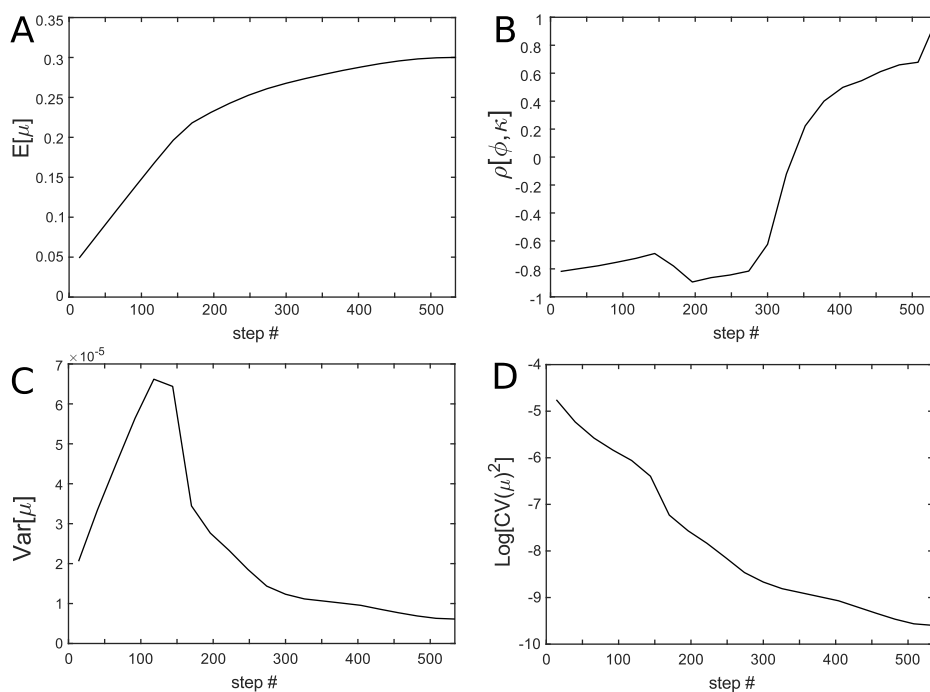


Figure S3: Example of the optimisation algorithm (gradient-based hill climb algorithm). **(A)** Growth rate increases each step. **(B)** Correlation coefficient increases and switches sign during optimisations. **(C)** Variance decreases during most parts of the optimisation process. **(D)** CV_μ^2 decreases. Parameters of this example kinotype are equal to figure 1B.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

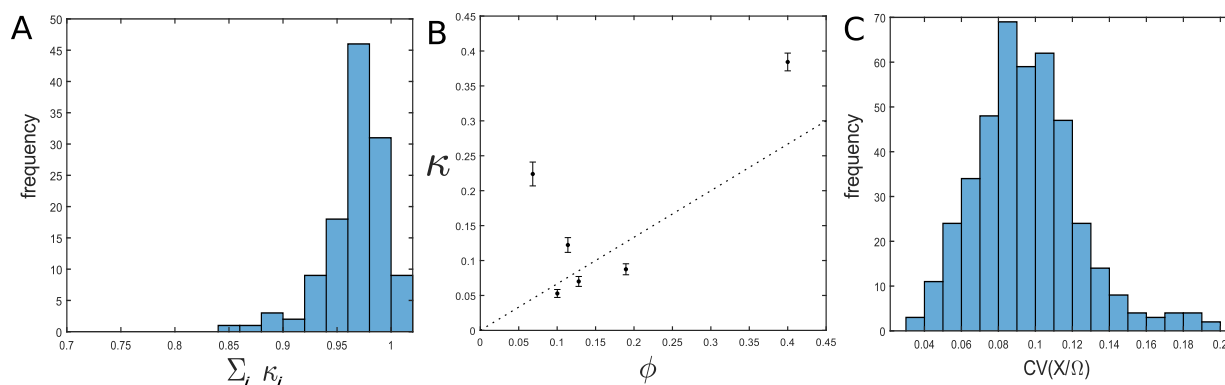


Figure S4: Examination of the noise decomposition method and noise levels. **(A)** The sum of all first order noise contributions is close to unity, indicating that a first order Global Sensitivity Analysis captures the variance contributions well. **(B)** Noise contributions for the example kinotype in the LN genotype in the high noise regime. Error bars indicate 2 sd over 125 repeated sampling of $2 \cdot 10^4$ phenotypes. Sampling errors in κ are within reasonable bounds. **(C)** Distribution of CVs as encountered in HN genotypes.

Abundant proteins contribute strongly to noise in the growth rate of optimally growing bacteria

Table 1: Kinetic parameters, k_{cat} and inhibition constant (k_{inhi}) for the five metabolic protein species in the kinotype used to create figure Fig. 1 and S3. Additionally, the initially sampled (relative) protein abundances are given.

Protein	k_{cat}	k_{inhi}	ϕ_{initial}
1	3.4048	0.2153	0.19177
2	4.3489	4.3831	0.0080544
3	1.8454	3.6435	0.18137
4	3.1650	6.5335	0.084103
5	5.4577	1.8425	0.1347