The “friend and foe” of deterministic and stochastic cell-cell variations

Søren Vedel (svedel@nbi.dk)
Niels Bohr Institute, University of Copenhagen https://orcid.org/0000-0003-0405-4391
Andrej Košmrlj
Princeton University https://orcid.org/0000-0001-6137-9200
Harry Nunns
California Institute of Technology
Ala Trusina
University of Copenhagen https://orcid.org/0000-0003-1945-454X

Article

Keywords: PACS numbers, Asymmetric damage segregation (ADS), Population dynamics, Statistical physics, Biophysics

Posted Date: September 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-943216/v1

The "friend and foe" of deterministic and stochastic cell-cell variations

Søren Vedel,[1,2,*] Andrej Košmrlj,[3,4] Harry Nunns,[2,5] and Ala Trusina[2,†]

[1] *Niels Bohr International Academy, Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*
[2] *Center for Models of Life, Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*
[3] *Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA*
[4] *Princeton Institute for the Science and Technology of Materials, Princeton University, Princeton, NJ 08544, USA*
[5] *Division of Biology and Biological Engineering, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA*

(Dated: September 27, 2021)

By diversifying, cells in a clonal population can together overcome the limits of individuals. Diversity in single-cell growth rates allows the population to survive environmental stresses, such as antibiotics, and grow faster than undiversified population. These functional cell-cell variations can arise stochastically, from noise in biochemical reactions, or deterministically, by asymmetrically distributing damaged components. While each of the mechanism is well understood, the effect of the combined mechanisms is unclear. To evaluate the contribution of the deterministic component we mapped the growing population to the Ising model. Model results, confirmed by simulations and experimental data, show that cell-cell variations increase near-linearly with stress. As a consequence, we predict that the entropic gain — the gain in population doubling time compared to an "average" cell — is primarily stochastic at low stress but crosses over to deterministic at higher stresses. Furthermore, we find that while the deterministic component minimizes population damage, stochastic variations antagonize this effect. Together our results may help identifying stress-tolerant pathogenic cells and thus inspire novel antibiotic strategies.

PACS numbers: Asymmetric damage segregation (ADS) — Population dynamics — Statistical physics — Biophysics

INTRODUCTION

Despite identical genetic makeup, cells in a population are never the same. This phenotypic diversity is functional and allows clonal populations to outperform individual cells. As an example, just variations in *E. Coli* doubling times alone has been found to be sufficient for a population to grow faster than individual cells, a phenomenon referred to as entropic gain [1]. Variations in doubling times are also beneficial for dealing with sudden changes in the environment – stress – with slowly growing cells tending to be more stress-resilient [2] and by diversifying into fast-growing stress-sensitive and slow-growing stress-tolerant cells [3] populations can both grow fast in stress-free environment and survive sudden stresses (e.g. antibiotics). Other well-established functional outcomes of phenotypic diversity include improved multicellular migration [4] and functional roles in gene circuits [5].

In case of phenotypic diversity in *E. Coli* single-cell growth rates, the continuum of states is typically assumed to be generated by stochastic fluctuations in protein numbers [1, 6] or metabolism [7]. There is, however, also a deterministic component: Cells diversify through an ongoing process of asymmetrically redistributing damaged proteins at division known as Asymmetric Damage Segregation (ADS) (Fig. 1A) [4, 8, 9]. As cells with more damaged proteins grow slower [10–16], the cell-cell differences in damaged proteins result in differences in growth rates. This difference is amplified at every division along the lineages that consistently avoid or inherit the damage (illustrated by the green-shaded and top lineage in Fig. 1A, respectively). Between these extremes lie the lineages with all the other combinations of damage inheritance (Fig. 1B). Thus, ADS generates a structured population tree with correlated lineages (Fig. 1B) while increasing the population growth and decreasing the mean damage in the population [8, 11, 17].

Under physiological conditions cells experience both random and deterministic effects [1, 8, 11]. However, it remains unclear how the stochastic and deterministic components interact and together contribute to population fitness. Do the two sources of diversity interact positively by adding together or perhaps even amplifying the effects of each other? Or do they antagonize each other? In addition to their conceptual importance, answering these questions could help identifying stress-tolerant cells and thus inspire novel antibiotic strategies.

In pursuing these questions we combined experimental data on 39,221 cells, theoretical modeling and numerical simulations and arrived at three novel results. First, we find that despite noise dominating the experimental dis-

* Corresponding author: svedel@nbi.dk
† Corresponding author: trusina@nbi.dk

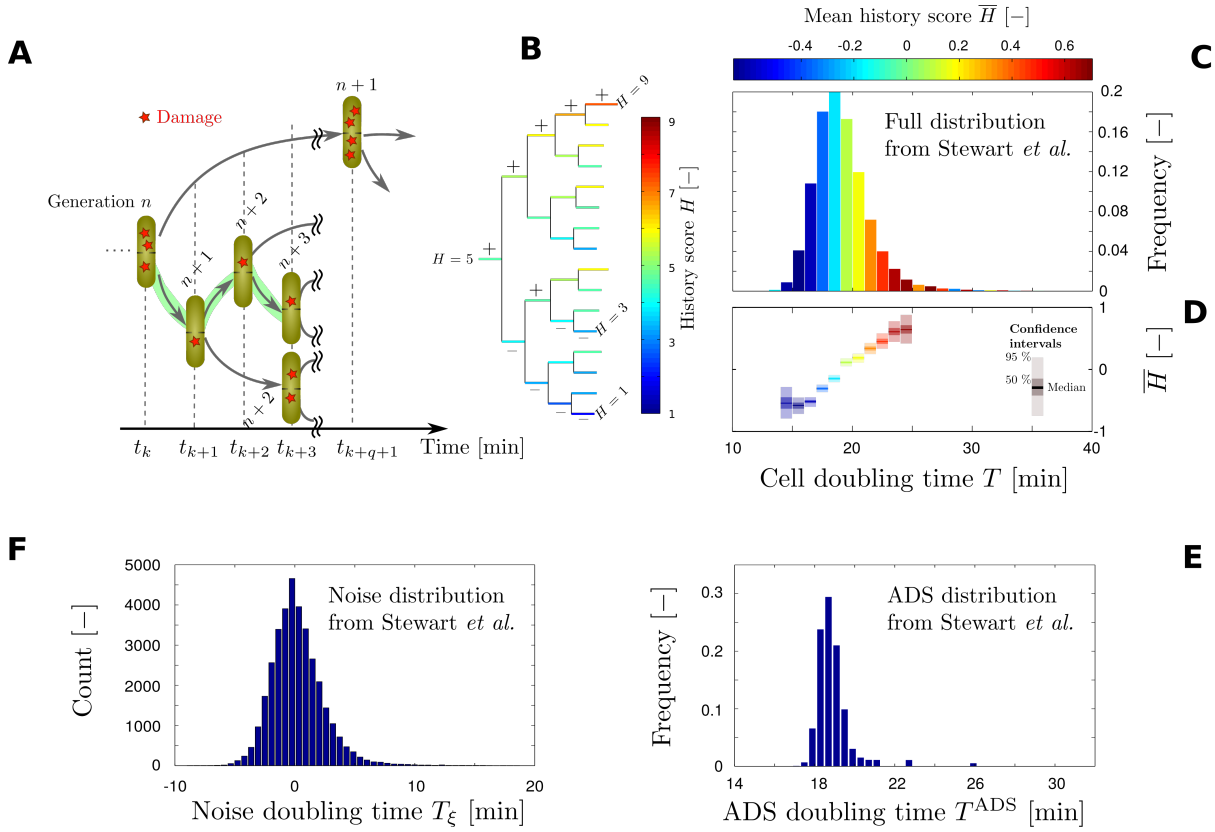

FIG. 1. **A**. Schematic illustration of damage redistribution in a population of dividing cells. The doubling time T of the cell is slowed down by the accumulated damage [10, 11] (Eq. (2)), and all cells acquire environmental damage (each red star in the schematic corresponds to λ amount of damage). At division, the damage is inherited asymmetrically resulting in difference in doubling time of the two sisters (Eq. (3)). Green shading marks the lineage where cells consequently evade damage. (Model parameters in the schematic: $a = 1$, $t_{\min} = 0$, $\mu\lambda = 1$). **B**. Damage-history scores H across a population tree generated by the ADS model (Eqs. (2) and (3)) with each cell (horizontal line) color-coded by its damage-history score. Low (blue) and high (red) scores correspond to low and high levels of damage. **C**. Distribution of doubling times from Stewart *et al.* [10]. Cells with similar doubling times are grouped together (bin size of 1 min). The color shows the average damage-history score in each bin $\overline{H}$. In line with ADS model, cells with higher damage-history score grow slower (see also SI Fig. S1). **D**. Bootstrap distributions of mean damage-history score illustrate that these results are statistically significant across bins with many cells (see also SI Fig. S2). **E**–**F**. The distribution of ADS doubling times (**E**) and noise (variations from average doubling time, **F**) extracted from the raw distribution in **C** via noise-filtering (Eq. (1)).

tribution of the cells' doubling times, the contributions from ADS and noise are statistically independent, so that the diversity from ADS creates a "scaffold" on top of which the diversity from random noise is added. Second, we show that the deterministic contribution from ADS can be described analytically by mapping the growing population of cells to the Ising model [18, 19]. Analytical results, confirmed by simulations and experimental data, suggest the nearly-linear relationship between population diversity and stress. Furthermore, we predict that the entropic gain crosses over from stochastic at low stress to deterministic at higher stresses. Third, we find that while stochastic and deterministic effects *add together* in reducing population doubling time, they *antagonize* each other in reducing population damage levels.

RESULTS

History of damage reveals the impact of ADS in data

The doubling times of *E. Coli* cells exhibit significant variation, and the shape of the distribution has been proposed to result from stochastic intracellular optimization of self-replication [6]. To quantify the contribution of deterministic ADS to shaping the distribution of doubling times, we first re-analysed experimental data by Stewart *et al.* [10] sampled over 35,049 individual cells (Fig. 1C).

If the contribution of deterministic ADS is significant, we would expect the cell's doubling time to correlate with its history of damage accumulation events [11]. To quantify the history of damage, we labeled each accumulation/evasion event by $+1/-1$(Fig. 1B) and defined the damage-history score H to be the sum of these labels

along each lineage (Fig. 1B; see Methods for details). Using pole-ages as a proxy of damage accumulation/evasion [11] – with older pole counting as +1, and newer pole as −1 – we assigned a damage-history score to each cell in the experimental dataset by Stewart *et al.*. Remarkably, this simple metric is sufficient to shed light on the underlying ADS-induced structure in the entire distribution of doubling times (Fig. 1C, D and Fig. S1). Confirming our hypothesis, the damage-history score, averaged over cells with similar doubling times, $\overline{H}$, indeed correlates with the cell's doubling time (Fig. 1D and Fig. S2). The trend in the average damage-history $\overline{H}$ furthermore reflects a typical pattern for individual damage-history scores, with the fraction of cells with positive H increasing with growing doubling times (Fig. S1B), and the opposite trend for cells with negative H.

History scores provide an illustrative heuristic for demonstrating the influence of ADS, but they do not account for the order of inheritance events. However, inheriting the old-pole sibling in the first and in last generations does not yield the same impact on the doubling time of the cell [11, 16]. Thus, to further characterize the nature of the deterministic and stochastic effects, we separated the stochastic contributions of noise from deterministic contribution of ADS using an extended version of our previously published method which accounts for the order of inheritance events [11] (Methods). In short, this method extends the history scores by grouping cells with the same lineage history ℓ (e.g. ℓ = ONNN for the lineage highlighted in green in Fig. 1A, while another group would have ℓ = NNON, where O and N correspond to the inherited old and new pole, respectively; these groups have different lineage histories but identical history scores). The median doubling time of each group is used as an estimate of the ADS contribution to the doubling time of cell i from group ℓ, $T^{\mathrm{ADS}}_{\ell,i}$ (Fig. 1E). The variation in doubling times within a group is then assumed to be caused by an *ADS-independent* stochastic noise $T^{\xi}_{\ell,i}$, such that the doubling time of cell i from group ℓ, $T_{\ell,i}$ can be represented as a sum of deterministic and stochastic components:

$$T_{\ell,i} = T^{\mathrm{ADS}}_{\ell,i} + T^{\xi}_{\ell,i}. \tag{1}$$

We found that the noise distribution (Fig. 1F) is statistically independent from the ADS distribution (contingency table; Fig. S4, Table S1 and Methods). In line with our previous results, we found that the distribution of T^{ADS} in dataset by Stewart et al.(Fig. 1E) is substantially more narrow than the full distribution (Fig. 1B), but – in accordance with ADS model [11] – remains right-skewed. We also confirmed that the damage-history scores go from low to high as the T^{ADS} increases (Fig. S3). Taken together, these results from full-population analyses indicate a significant impact of the damage segregation in shaping the population-distribution of doubling times by providing a deterministic "backbone" distribution which is widened further by random noise. This corroborates recent results by Proenca *et al.* showing that noise acts "on top of" ADS [8].

Analytical model for deterministic ADS population dynamics

To better understand the interaction of stochastic diversity and ADS we developed an analytical model for the deterministic model of ADS [11]. Most previous analytical treatments of growing populations focus on effects of stochastic dynamics either along lineages or the environment [20–25] and converge on entropy of lineages for increasing population growth, yet no comparable model exists for deterministic population processes such as ADS. In our model we sought to describe how single-cell ADS inheritance dynamics shape the growth of the population and variations in single-cell doubling time and damage.

We considered the simplest possible model for ADS which describes a clonal population growing in an environment with constant (sub-lethal) stress (Fig. 1A, [11]). In each generation n, each cell accumulates an amount of damage λ in response to the environmental stress, and in addition inherits $D_{\mathrm{inh}}(n)$ amount of damage from its mother that together add up to the total damage $D(n)$. The doubling time of the cell $T(n)$ is slowed down by the accumulated damage, resulting in the following single-cell model

$$\begin{aligned} D(n) &= \lambda + D_{\mathrm{inh}}(n), \\ T(n) &= t_{\min} + \mu D(n), \end{aligned} \tag{2}$$

where $t_{\min}$ (a constant) is a shortest-possible cell doubling time set by basal, non-ADS processes and μ is the proportionality constant between damage and doubling time. At division, the amount of inherited damage is determined by

$$D_{\mathrm{inh}}(n+1) = \frac{1 + s_n a}{2} D(n), \tag{3}$$

where $s_n = 1(-1)$ indicates whether a cell in generation n accumulates (evades) damage relative to its sibling (Fig. 2A, top panel). The amount of asymmetry is described by the parameter a ($0 \le a \le 1$).

Despite the simplicity of this single-cell model, deriving expressions for the emergent population dynamics is not straight forward because of the non-trivial coupling between cell-generations and time: Cells from different generations may be present at the same time, e.g. cells dividing at time t_{k+3} in Fig. 1A are in generations $n+2$ and $n+3$. We overcame this challenge by drawing on an unexpected connection between the non-equilibrium process of growing populations with ADS and equilibrium statistical physics of ferromagnetic particles (Ising model). In brief, the increment in population size $\mathrm{d}N_{\mathrm{cells}}(t) = N_{\mathrm{cells}}(t) - N_{\mathrm{cells}}(t - \mathrm{d}t)$ at time t is given by the number of lineages dividing exactly at time t. Thus, the the population growth rate can be obtained as the rate of growth of the total number of lineages $\Omega_L(t)$

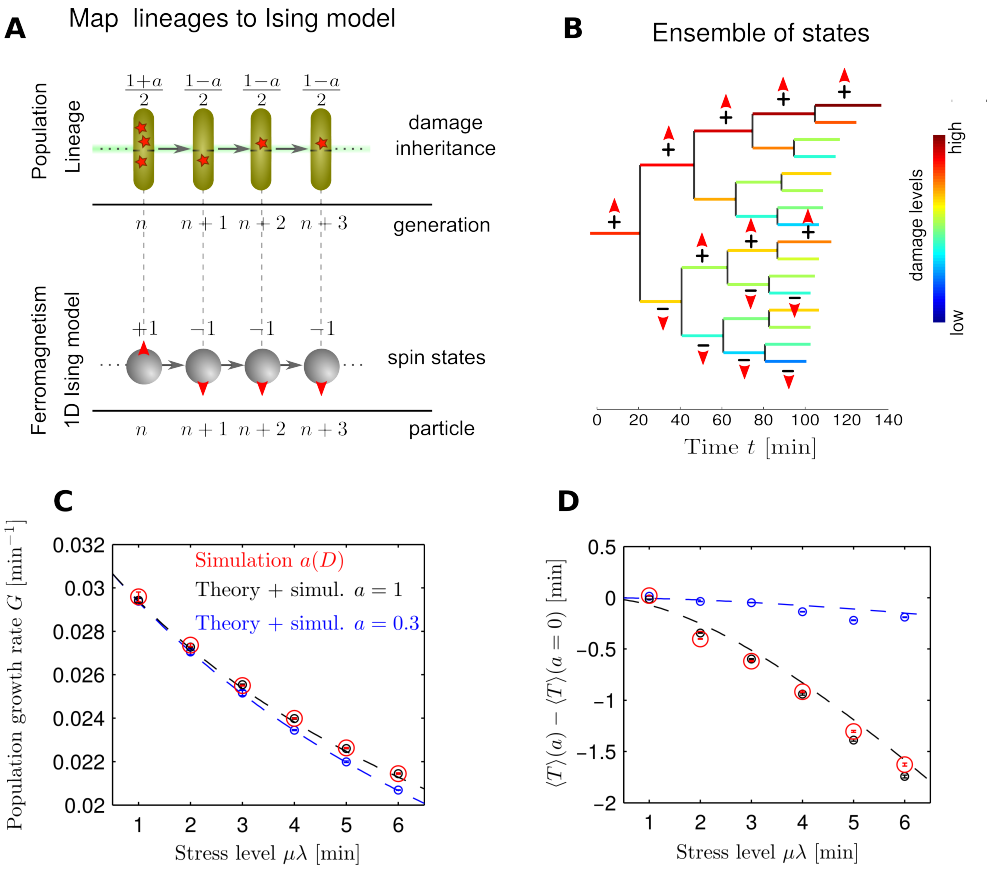

FIG. 2. Mapping of population with asymmetric damage segregation to statistical physics. **A**. Schematic illustration of the analogy between lineages in a growing population and 1D Ising model (illustrated lineage is the same highlighted in green in Fig. 1A). For a lineage, the history of damage inheritance corresponds to an unique configuration of spins on an Ising chain. The damage accumulation/evasion events shown above each cell (their fraction of inheritance of ancestor damage via Eq. (3)) correspond to the values of the spins (± 1, red arrows) in the analogous Ising chain. **B**. Each lineage in a population tree maps to a microstate in an Ising model. Select lineages are marked with the corresponding spin states as in panel A (note conceptual similarity to damage-history scores across population in Fig. 1B). **C**. Population growth rate as predicted by the theory (lines, Eq. (4)) agrees with numerical simulations (circles), as illustrated for $a = 0.3$ (blue) and $a = 1$ (black). Results for the reported damage-dependent asymmetry, $a(D)$ ([11], red) overlap with those for $a = 1$. Both $t_{\min} = 22$ min and $\mu\lambda$ are in relevant ranges for *E. Coli* under experimental conditions [11]. **D**. The population growth gain from ADS increases with stress. The growth gain is quantified by the difference of mean population doubling times $\langle T\rangle$ with and without ADS.

dividing at time t, which are the lineages where the total accumulated lineage time $T_{\rm tot}(n,\{s\}) = \sum_{i=1}^{n} T(i,\{s\})$ is equal to the elapsed time t, i.e. lineages satisfying $T_{\rm tot}(n,\{s\}) = \sum_{i=1}^{n} T(i,\{s\}) = t$. Mathematically, determining the number of dividing lineages $\Omega_L(t)$ is equivalent to the well-described problem of determining the number of states Ω with a given energy for a 1D system of ferromagnetic particles (Ising model) in equilibrium statistical physics (Fig. 2A), i.e. the number of possible ways the Hamiltonian $\mathcal{H}$ of the system evaluates to the energy level E, $\mathcal{H} = E$ (Fig. 2B and Table S2). In effect this maps each lineage to a particular microstate in the Ising model (Fig. 2B) and the total time for a lineage $\{s\}$ to reach n generations, $T_{\rm tot}(n,\{s\}) = \sum_{i=1}^{n} T_i$ plays the role of Hamiltonian $\mathcal{H}$. The full derivation is given in SI Sec. S1, including summary of all simplifying assumptions (Table S3) and assessment of their impact (Sec. S1.4.1).

To characterize the emergent population dynamics from deterministic ADS, we derived expressions for the population growth rate $G_{\rm ADS}$, the mean single-cell doubling time $\langle T\rangle_{\rm ADS}$ and the standard deviation of doubling times $\sigma_{T,\rm ADS}$ (Eq. 4, see SI Sec. S1 for derivation). It is important to note that we aimed to characterize the properties of the "active" part of the population. Thus only dividing cells are included when calculating population averages (see Sec. S4.5 for detailed argumentation). In other words, the properties of the growing cells, those that are between the division cycles, do not contribute to the averages.

$$
\begin{aligned}
G_{\mathrm{ADS}}(a,\lambda,\mu,t_{\min}) = \tilde{S}(a,\lambda,\mu,t_{\min}) &\approx \frac{8\lambda\mu + 4t_{\min} - 2\varphi(a,\lambda,\mu,t_{\min})}{a^2(9+2a^2)\lambda^2\mu^2},\\
\langle T\rangle_{\mathrm{ADS}}(a,\lambda,\mu,t_{\min}) &\approx t_{\min} + \mu\lambda\left[2 - \tilde{S}\frac{a^2\lambda\mu}{4}\left(a^2+9\right) + \tilde{S}^2a^4\lambda^2\mu^2\frac{27}{8}\right],\\
\sigma_{T,\mathrm{ADS}}(a,\lambda,\mu,t_{\min}) &\approx \frac{\lambda\mu a}{4}\sqrt{20 + a^2\left(4 + \tilde{S}\lambda\mu\left[-44 - 45\tilde{S}\lambda\mu + a^2\left\{2 + 9\tilde{S}\lambda\mu\right\}\right]\right)},\\
\varphi(a,\lambda,\mu,t_{\min}) &= \sqrt{2}\sqrt{8\lambda\mu t_{\min} + 2t_{\min}^2 - \lambda^2\mu^2\left[9a^2\ln 2 - 8 + a^4\ln 4\right]},
\end{aligned}
\tag{4}
$$

where $\tilde{S} = S/(k_{\mathrm{B}}t)$ is the normalized entropy of dividing lineages expressed in terms of the thermodynamic entropy S and the elapsed time t (SI Sec. S1.7) which is a measure of lineage diversity. Furthermore, to obtain a simpler expression for $\sigma_{T,\mathrm{ADS}}$ we assumed in the expression above $t_{\min} \gtrsim 15$ min. An approximate expression valid for all values of $t_{\min}$ is given in Eq. (S53). We furthermore use the short-hand notation $\tilde{S}$ instead of $\tilde{S}(a,\lambda,\mu,t_{\min})$ on the right hand-side.

Interestingly, all these population characteristics can be expressed in terms of the entropy of dividing lineages $\tilde{S}$ (population diversity), with the population growth rate G_{ADS} being mathematically identical to $\tilde{S}$ (SI Sec. S1.7). Thus, the entropic gain in the population growth rate by ADS alone (the increase in population growth rate driven by lineage entropy) $\Delta G_{\mathrm{ADS}} = G_{\mathrm{ADS}}(a) - G_{\mathrm{ADS}}(a=0)$ is equal to the increase in lineage entropy from ADS $\Delta G_{\mathrm{ADS}} = \tilde{S}(a) - \tilde{S}(a=0)$ (where $\tilde{S}(a=0)$ is given by SI Eq. (S58)). The reason is that growth of the population is determined by the number of dividing lineages (number of states Ω in the equivalent system), and so the rate of growth G is simply $\frac{1}{t}\ln\Omega(n,t)$ (n is the number of generations/particles), which is also the very definition of the normalized entropy $\tilde{S}$. This deep connection between population diversity and population growth, which is conceptually similar to the effects from stochastic variations presented by Wakamoto *et al.* [1, 22], establishes that population growth is driven by entropy of lineages.

We confirmed our theoretical predictions by comparing to an agent-based numerical simulation we introduced in Vedel *et al.* [11] (Fig. 2C, D and Fig. 3A). (See SI Sec. S3 for the details of the simulation algorithm and its numerical implementation). Simulations are in good numerical agreement with the analytical results, also for the reported case of damage-dependent asymmetry $a = a(D) \approx \frac{D^6}{D^6+\mathrm{constant}}$ [11].

Our theoretical results (Eq. (4), Fig. 2C, D) confirm the previous findings in Vedel *et al.* [11] that while the growth is reduced under higher stresses, asymmetric damage segregation ($a > 0$) ensures faster growth than symmetric ($a = 0$). Furthermore, the positive impact of ADS *grows* with stress $\lambda\mu$ (Fig. 2D), resulting in a dynamic, emergent population-benefit [11] endowing the population with an antifragile stress behavior [26]. This antifragility comes as a result of increased diversification under stress; cells can diversify more when they have more damage to redistribute. Here our analytical results predict that this diversity, measured as the width of the distribution of doubling times, increases linearly with stress $\sigma_{T,\mathrm{ADS}} \sim \lambda$ when $a = 1$ (Fig. 3A) and approximately linear for damage-dependent asymmetry $a(D)$, Fig. S5.

Experimental verification of ADS population dynamics

We verified this emergent behaviour by analyzing published data for 39,221 cells by Stewart *et al.* [10] and ourselves [11] for the stress-dependence of $\sigma_{T,\mathrm{ADS}}$ across three different stresses and three different strains . Using the ADS-filtered data, we fitted the deterministic ADS model to each experimental condition and obtained the experienced stress $\lambda\mu$ (Methods). Plotting the experimental values of $\sigma_{T,\mathrm{ADS}}$ against these experienced stresses, we found that the data across conditions was well-described by our theoretical predictions (Fig. 3B). The significantly lower levels of perceived stress in data by Stewart *et al.* [10], compared to our wild type strain [11], both grown at 37 °C, may reflect differences between growing cells in rich (Luria broth [10]) versus minimal (M63 [11]) media. In calculating the experimental values of $\sigma_{T,\mathrm{ADS}}$ we used all recorded cells in the growing population since this ensemble gives the same moments but higher cell count than using dividing cells at the last time point (see Methods, SI Sec. 4.5). This experimental validation of stress-induced deterministic diversification (Fig. 3B) together with the damage-structured population (Fig. 1C-E) strongly support the impact of the ADS on the population distribution and improved conditions for growth.

Synergistic contributions of stochastic and deterministic diversities to population growth

Having established and validated our population-level approach to ADS, we next focused on the central question of the interaction between the deterministic diversity and random noise. To investigate this we analyzed

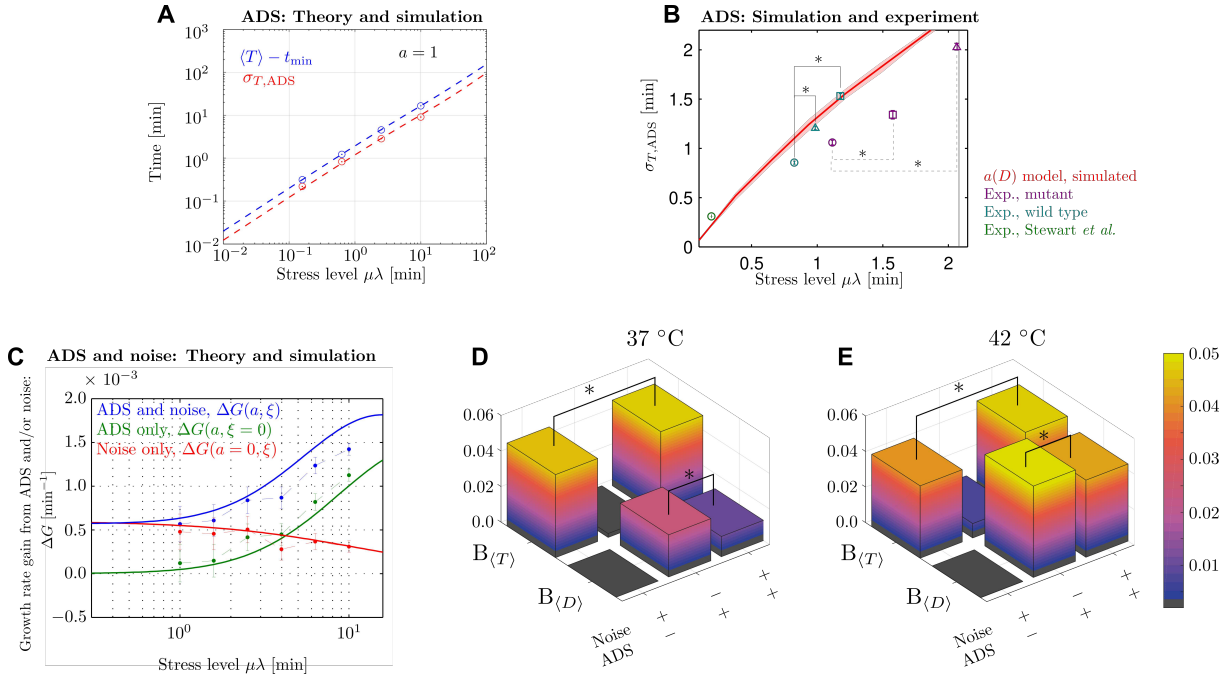

FIG. 3. The interplay of stochastic and deterministic sources of diversity under stress. **A**. Analytical predictions of linear increase of the effects of ADS with stress, as illustrated by $\langle T\rangle_{\mathrm{ADS}} - t_{\min}$ and $\sigma_{T,\mathrm{ADS}}$ (Eq. (4)). Dashed lines are analytical results, circles mark simulation results. Errorbars show sample error of the means from the simulations (Methods). The same trends hold for damage-dependent asymmetry, although exact linearity is not satisfied, see Fig. S5. **B** Comparison of model predictions for $\sigma_{T,\mathrm{ADS}}$ with experimental results across all stresses. Squares mark heat-stress, triangles antibiotic (kanamycin) stress and circle marks low UV stress (Stewart *et al.* [10]). Perceived experimental stress levels $\mu\lambda$ are inferred by fitting the model to the ADS-filtered data (Methods). Each mark represents the average over all cells with those particular experimental conditions, while the red line is the simulation results from 100 repeat simulations of $a(D)$-model (line indicates 50th percentile, with shaded region indicating 25th−75th percentile). **C** Synergistic contributions of noise and ADS to population growth rate. Points mark simulation results, while lines are the analytical results (Eq. (5)). The gain in population growth from noise alone,$\Delta G(a=0,\xi)$, is in green (Eq. (S65)); from ADS alone, $\Delta G(a=1,\xi=0)$, is in red (Eq. (S70)) and from combined noise and ADS, $\Delta G(a=1,\xi)$, is in blue (Eq. (5). Errorbars mark 95 % confidence interval on fitting exponential growth to the number of cells across 10 repeat simulations (SI Sec. S3.4.2.). **D**–**E** Benefits of doubling time, $B_{\langle T\rangle}$, and damage, $B_{\langle D\rangle}$, defined in Eq. (6), under various combinations of noise and ADS in a case of low (**D**) and medium (**E**) damage induced by heat stress [11]. While population benefits from both noise and ADS in reducing population doubling time (compare results for $B_{\langle T\rangle}$), noise antagonizes the beneficial effects of ADS in decreasing population damage (compare results for $B_{\langle D\rangle}$). For simulations with noise we show the mean across 10 realizations. Here the asymmetry is set to be damage-dependent, $a(D)$, however results are similar for $a = 1$. Asterisks indicate statistically significant differences with $p < 5.0\times10^{-8}$ (bootstrap, t-test).

the entropic gain for the population growth rate, i.e. the changes in population growth rate to modulations in lineage diversity $\tilde{S}$. In other words, we consider the increase in population growth rate $\Delta\tilde{G} = G(a,\xi) - G(a=0,\xi=0)$, where ξ denotes noise, and where lineage entropy is controlled via noise level ξ and asymmetry a. Using the Euler–Lotka model of population growth, we related the *population diversity* resulting from the combined existence of ADS and noise (with the probability density of doubling times $f_{\mathrm{full}}(T_{\mathrm{ADS}}, T_\xi)$) to *population growth* $G(a,\xi)$ [27–29]. Since the distributions of noise and ADS-induced doubling times are statistically independent (Table S1) we found a simplified expression for $G(a,\xi)$ by expanding each of these distributions in statistical cumulants and applying series inversion [30, 31] (SI Sec. S2.2). To account for the fact that noise amplitudes scale with stress (variance-to-mean ratio C_V of single-cell growth rates is conserved across stresses[32]), we furthermore scaled the noise amplitude relative to the mean doubling time $\langle T\rangle_{\mathrm{ADS}}$ (SI Sec. S2.1), resulting in the final expression

$$\begin{aligned}\Delta\tilde{G} &= G(a,\xi) - G(a=0,\xi=0)\\ &\approx \frac{\ln 2}{\langle T\rangle_{\mathrm{ADS}}} + \frac{1}{2}\left[\frac{\sigma^2_{T,\mathrm{ADS}}}{\langle T\rangle_{\mathrm{ADS}}} + C_V^2\langle T\rangle_{\mathrm{ADS}}\right]\left(\frac{\ln 2}{\langle T\rangle_{\mathrm{ADS}}}\right)^2\\ &\quad - \frac{\ln 2}{t_{\min}+2\mu\lambda},\end{aligned} \tag{5}$$

where $\langle T\rangle_{\mathrm{ADS}}, \sigma^2_{T,\mathrm{ADS}}$ are given by Eq. (4), $G(a=0,\xi=0) = \frac{\ln 2}{t_{\min}+2\mu\lambda}$ (SI Sec. S1.8) and $C_V = 0.23$ is inferred from the experimental data [11]. These analytical results (Eq. (5), Fig. 3C, blue line) capture the general trend of the simulated results (Fig. 3C, blue and magenta dots).

With all parameters constrained by the data, our analytical and simulation results show that the relative importance of the deterministic and stochastic processes in shaping the population growth changes with stress (green and red lines in Fig. 3C). At low stresses, the benefit from ADS (green) is small and noise (red) dominates, while ADS dominates for large stresses. The experimental estimates of the perceived stress $\mu\lambda$ (Fig. 3B) all fall below the cross-over point, $\mu\lambda_c \approx 4$, suggesting that noise dominates in the tested conditions. Furthermore, the estimates of the perceived stress allow us to read off the relative contributions of deterministic and stochastic components (Fig. 3C) for each of the experimental conditions. Thus for example, in mutant cells experiencing kanamycin stress with $\mu\lambda \sim 2$, ADS accounts for about 24 % of the gain in growth rate, $\Delta\tilde{S}$. Recently, Proenca *et al.* [8] estimated that in a stress-free environment deterministic sources contribute 22 % to the variance of the doubling times. Although the numbers are not directly comparable, our results agree on that noise dominates distribution of doubling times and as a consequence has a major contribution to the entropic gain in population growth rate. While the contribution from ADS is minor, it is significant (Fig. 3C-E) and may dominate under the experimental conditions where perceived stresses are higher.

Stochastic and deterministic effects antagonize each other in reducing population damage levels

We have previously shown that by concentrating damage in a few slowly growing cells, ADS decreases the average damage in a population [11]. We then asked how the measured levels of noise impact the ADS-induced reduction in population damage and mean doubling time.

To quantify the effects we considered how the population benefit from the diversity generated by either noise alone, ADS alone, or both ADS and noise combined. We use the case with $a = 0$ and "no noise" ($\xi = 0$) as a reference point and define benefits as the decrease in population damage and mean doubling times relative to their reference values:

$$B_{\langle T\rangle} = 1 - \frac{\langle T\rangle(a,\xi)}{\langle T\rangle(a=0,\xi=0)} \qquad (6)$$
$$B_{\langle D\rangle} = 1 - \frac{\langle D\rangle(a,\xi)}{\langle D\rangle(a=0,\xi=0)},$$

here $B_{\langle T\rangle}$ and $B_{\langle D\rangle}$ are the benefits from the diversity in the population-mean doubling time $\langle T\rangle$ and damage $\langle D\rangle$.

The results (Fig. 3D and E) for *E. Coli* doubling times mirrored our results for the entropic gain in population growth rate (Fig. 3C). The benefit in doubling time is dominated by noise $B_{\langle T\rangle} \sim 0.04$ and ADS provides a minor but statistically significant increase of ~ 0.01. The roles of ADS and noise are, however, distinctly different in reducing population damage. While ADS alone significantly reduces population damage, with $B_{\langle D\rangle} \sim 0.05$ at higher stress, noise antagonizes this effect decreasing the benefit by ~ 0.01 (Fig. 3D, E).

To understand these seemingly contradictory outcomes of noise, consider the dynamics along a single lineage. Noise interferes with the ordered chains of damage inheritance through the lineages so that it is no longer cells with the lowest amounts of damage which have shortest doubling times; even though the average noise contribution along each lineage is zero, the fastest-dividing cells — those which drive the population growth — will on average have more damage when noise is present. Consequently the average damage $\langle D\rangle$ is higher when noise is present, leading to lower benefits $B_{\langle D\rangle}$.

DISCUSSION

Single-cell studies on drug-tolerance suggest that there is a common theme across drugs and organisms: A small fraction of the population in a dormant state — cells growing slowly or completely arresting their growth — survive exposure to drugs [33, 34]. Several mechanisms have been proposed to explain the emergence of the tolerance: from diversity in lag times and growth rates to toxin-antitoxin regulatory networks [34, 35]. Similar studies in cancer cells are slowly emerging [33]. While these "executive" mechanisms are different, most are assumed to be a result of bet-hedging—initiated by stochastic fluctuations in protein numbers [3]. Alternate to random bet-hedging, the dormant cells (and thus drug-tolerance) may be generated by the deterministic processes such as asymmetric damage segregation [8, 11] or asymmetric partitioning of the multidrug efflux pump (AcrAB-TolC) towards the old pole [36].

Understanding the interplay between the stochastic and deterministic sources of diversity (of which tolerance-associated dormancy is a specific case) can influence the development of the novel antibiotic strategies. If the origin of the diversity is stochastic we can estimate how many but *not which* cells become tolerant, whereas if it is *deterministic*, we can, in principle, predict *both how many and which cells* are likely to become tolerant. Thus, for example, in the cases of asymmetric damage or AcrAB-TolC partitioning the pole-age, the size of protein aggregates may serve as a marker for dormancy and thus tolerance. In line with this, recent findings show that cells with aggregates are more likely to tolerate lethal stresses [37]. In this context, the uncovered correlation between damage-history score and doubling time suggest that slow growth in several consecutive generations may serve as another marker for old-pole and dormancy. Experimentally, this possible link between damage-history, dormancy and antibiotic tolerance can be assessed by fluorescent pulse-labeling of the cell wall [38].

Although simplistic at the levels of individual cells and lineages, the mathematical model describing the dynamic

interactions between the deterministic asymmetric damage segregation (ADS) and noise presented in this paper revealed intricate emergent population properties. Together with experimental validation across a large and diverse dataset (Fig. 3B), our modeling results suggest the following key results.

First, we analytically show that stress induces deterministic cell-cell variability through ADS which, in turn, results in deterministic changes to population growth (depends explicitly on the entropy of lineages $\tilde{S}$, Eq. (4)). Compared to most other studies, this variability is deterministic and correlated through generations, yet the effects on population growth are similar to those endowed by noise alone [1, 22].

Second, our analyses show that while operating on different time-scales, both the deterministic (ADS) and stochastic (noise) sources of diversity accelerate population growth (as compared to a non-diversified population) and their effects add up. It furthermore suggests that the "speed limit" on population growth proposed by Hashimoto *et al.* [1] can be circumvented if multiple sources of diversity are present.

Under varying stresses, the additive effect provides the population with the best of both contributions: under low stress the growth-benefit from diversity is dominated by the noise, while the benefits of ADS become more prevalent under higher stresses yet retaining the baseline growth benefit from noise. We find that in all the investigated experimental conditions the stress is not high enough for the deterministic component to dominate, however, even under the lowest perceived stress (low-level UV stress by Stewart *et al.*), the contribution of ADS is nonetheless significant (Fig 1C, D).

Third, while the deterministic and stochastic components both accelerate population growth, they have the opposing effects on population damage. Deterministic ADS reduces population damage because it effectively limits damage to a few slowly growing cells. Stochastic variations in cell's growth rates counteract this effect because they disrupt the coupling between cell's doubling times and damage levels (stochastically, cells with damage may grow fast and generate more damage-containing progenitors). This interplay between stochastic and deterministic variations suggest that to minimize both population doubling time and mean damage, there should be an optimal level of stochastic variations.

As our results are based on a generally applicable mathematical framework, we anticipate that these three key results will persist when other independent sources of variability (stochastic and deterministic) are at play (e.g. multivariate noise or other types of cellular damage, not limited to protein aggregates).

ACKNOWLEDGMENTS

We thank Eric J. Stewart for kindly sharing his data from [10].

METHODS

Experimental data

Previously published time-lapse experimental datasets of *E. Coli* with single-cell resolution containing single-cell growth rates and population pedigree were further analyzed for this paper [10, 11]. This data consisted of one dataset of more than 30,000 MG1655 cells from more than 90 different colonies [10], and 6 smaller datasets (each of 8-10 colonies, $\sim$ 2000 cells per dataset) of two different cell lines (MC4100ΔClpB with ClpB–YFP on a medium copy-number plasmid with lacI^{q} and expressed by added IPTG, and MC4100 with lacI^{q} (wild-type)) each exposed to both heat stress (37 °C to 42 °C) and antibiotic shock (0.0 μg/mL kanamycin to 0.5 μg/mL kanamycin, both at 37 °C) [11]. MG1655 cells were genetically modified to express yellow fluorescence and were excited every 2 or 4 minutes, while both MC4100 cell lines were imaged in brightfield channel every 3-4 min (and some cells also in fluorescent channel every 15 min).

Simulations

We extended our previously published agent-based simulation method for all results presented in this paper [11]. We present a more detailed description of the simulation method including convergence study and the use of low-amplitude noise to speed up convergence in SI Sec. S3. Briefly, the simulation method builds out the full population tree up to N_{gens} generations, and calculates the amount of accumulated damage and the doubling time for each cell by iteratively applying Eq. (2) (using either constant or damage-dependent asymmetry $a(D)$). From this we can obtain for each cell k in generation n the birth time $\tau_k^{\text{birth}}(n)$, the death time $\tau_k^{\text{death}}(n)$ and the damage $D_k(n)$. For simulations including random noise, this is added during the iterative application of Eq. (2) using $T(n) = t_{\min} + \mu D(n) + \xi_n$ with ξ_n being drawn at random from the experimental noise distribution (Fig. 1F) using MATLAB's built-in random number generator. We ran simulations to at least 22 generations unless explicitly stated otherwise, and ensured convergence for all results presented.

Statistical moments of population distributions — To infer population distribution moments ($\langle T\rangle_{\text{ADS}}$, $\sigma_{T,\text{ADS}}$, $\langle D\rangle$), the simulated population is interrogated at fixed times t_i, including only cells present at time t_i defined by $\tau_k^{\text{birth}}(n) \le t_i < \tau_k^{\text{death}}(n)$. Since the simulation starts from a single cell and therefore is strongly affected by low-count statistics to begin with, we waited for population moments to converge to stable values before sampling (details in SI Sec. S3). The instantaneous values of the moments at time t_i were obtained as the mean of all sampled values in the window $\tau_k^{\text{birth}}(n) \le t_i < \tau_k^{\text{death}}(n)$ and the standard error of the mean calculated using the same part of the simulated output (see details in SI Sec. S3).

Presented results in Figs. 2 and 3 were averaged over all sampling times after the equilibration period.

Benefits $B_{\langle T\rangle}$ and $B_{\langle D\rangle}$ — The benefits (Eq. (6)) are calculated using the estimates for the moments $\langle T\rangle$ and $\langle D\rangle$ described above. Statistical significance of increase in benefit (Fig. 3D, E) was calculated using two-sample t-tests.

Population growth rate and entropy of lineages — The population growth rate G was estimated by a linear fit to the logarithm of the simulated N_{cells} as a function of simulated time t assuming a first-order polynomial. As for the population moments the initial phase was ignored to allow for equilibration before fitting to obtain parameter values and their standard errors. For repeat simulations (Fig. 3C), average results were obtained by leveraging the Gaussian nature of inferred parameters from linear fits, see SI Sec. S3.4 for details.

Since the normalized entropy $\tilde{S} = \frac{S}{k_{\text{B}}t}$ is equal to the population growth rate for both ADS and noise (see SI Eq. (S34)), we obtain this from the simulations as described for the population growth rate, and the standard error for differences in $\tilde{S}$ in Fig. 3C were estimated using error propagation.

Theoretical modeling

Deterministic model for ADS — Full derivation of the ADS model is presented in SI Sec. S1 including derivation of the explicit yet surprising connection between the entropy of dividing lineages $\tilde{S}$ and population growth rate G_{ADS}.

Euler–Lotka model, cumulant expansion and noise model — Derivation of Eq. (5), detailed introduction and validity assessment of cumulant expansion for random noise and description of noise scaling across stress levels to satisfy single-cell constant growth rate variation to mean (as reported by [32]) are presented in SI Sec. S2. We used $C_V = 0.23$ obtained from the experimental data [11] in all results with the Euler–Lotka model involving noise.

Experimental data analysis

Further analysis details are presented in SI Sec. S4.

History scores — For each lineage, we analyzed the cell divisions i and assigned a value of $s_i = +1$ for each old-pole inheritance and $s_i = -1$ for each new-pole inheritance. The damage-history score H is the defined along each lineage as the sum of the s_is, $H = \sum_i s_i$.

Noise filtering: Splitting experimental single-cell doubling times into noise and ADS contributions — To estimate the effect of ADS we extended our previously published method [11] to take into account the entire sequence of inheritance events (i.e. entire lineage history ℓ) and thus only group the cells from repeat experiments that have exactly the same lineage history. Formally, each group forms a set of single-cell doubling times $L_\ell = \{T_{\ell,i} | i \text{ among sampled populations}\}$ for cells with a specified inheritance or lineage history sequence ℓ (e.g. $\ell =$ ONNN for a cell whose lineage history sequence starts with an old-pole inheritance in the first generation followed by new-pole inheritances in the last three inheritances; this lineage is highlighted in green in Figs. 1A and 2A). Notice that the lineage history ℓ and the experiment number i uniquely index the cell. Since the lineage history (inheritance history) ℓ for all cells in the group, L_ℓ, is the same, we estimate the ADS contribution to doubling time T_ℓ^{ADS} to be the median of L_ℓ; the variations about T_ℓ^{ADS} at the single-cell level are taken as noise. From this we obtain the contributions to the cell from experiment i from ADS ($T_{\ell,i}^{\text{ADS}}$) and noise ($T_{\ell,i}^{\xi}$),

$$\begin{aligned} T_{\ell,i}^{\text{ADS}} &= T_\ell^{\text{ADS}} = \text{median}\left(L_\ell\right), \\ T_{\ell,i}^{\xi} &= T_{\ell,i} - T_\ell^{\text{ADS}}. \end{aligned} \tag{7}$$

This results in the noise filtering indicated by Eq. (1). In the original version of this method published in [11] we grouped cells only based on the last two generations (i.e. the pole-age of the mother and the pole-age of the cell itself), whereas in the current version we are using the entire lineage inheritance history ℓ as basis for the grouping.

Inferring perceived stress in experiments and noise — Perceived stress $\mu\lambda$ of experiments were inferred by fitting the agent-based ADS simulation model to noise-filtered experimental doubling times (filtered using Eq. (1)). The model was fitted at all generations along all lineages using the noise-free simulations to obtain values for μ and λ. The residuals between the fit and actuals for each cell are the noise. At each point in the population tree, the median single-cell doubling time was used (Eq. (1)) as fitting target. For data from [10] we fixed t_{min} (cannot be inferred independently) and fitted μ and λ; for data from [11] we fixed μ for pairs of experiments with same stressor but different stress levels and allowed λ to take different values at each settings. More details given in SI Sec. S4.6.

Statistical independence of noise and ADS — We used a contingency table test to assess whether distributions of doubling times from ADS doubling and noise (Fig. 1E, F) are statistically independent. The null hypothesis is that the distributions are statistically independent and the alternative hypothesis that they are not. Briefly the test has been implemented as follows (see SI Sec. S4.4 for details): Following the filtering of Eq. (1), each of the two resulting distributions of doubling times for ADS and noise ($f_{\text{ADS}}(T^{\text{ADS}})$ and $f_\xi(T_\xi)$) are split into N_{ADS} and N_ξ bins, resulting in a $N_{\text{ADS}} \times N_\xi$ table of the number of noise counts (individual cells) $\mathcal{N}_{i,j}$ per ADS bin i and noise bin j. Assuming that noise is statistically independent of ADS we compared the expected number of cells per bin to the actually recorded number using a χ^2-statistic, and evaluated the P-value of the null hypothesis using the associated χ^2 distribution.

Ensembles for sampling: using all cells up to a time to represent a time snapshot — The benefits from ADS arise because cells with higher amounts of damage grow slower, thus resulting in an effective up-concentration of faster-growing cells with smaller amounts of damage at any instance in time. To properly measure that experimentally it is important that cells are sampled from a time snapshot when calculating population moments. However, we have found that the same population moments are found whether using the ensemble of all cells dividing up to a time t instead of the ensemble of cells present in a snaptshot taken at time t. We have used the larger ensemble to calculate $\sigma_{T,\mathrm{ADS}}$ for experimental data in Fig. 3B. More details on ensembles as well as a detailed analyses of their impact on population moments is provided in SI Sec. S4.5.

Statistical significance of increase in σ under different stresses — We used bootstrapping to assess statistical significance of increase in σ across different experimental conditions (Fig. 3B) since the underlying distributions are non-normal. Data was separately resampled with replacement under each of the two conditions to obtain estimates of means and calculate $\Delta\sigma = \sigma^{\text{high stress}} - \sigma^{\text{low stress}}$. Repeating this 10^3 times we obtained a normal distribution of differences which we compared to the normal distribution of differences of the associated null distributions from each original distribution (obtained by randomly shuffling cells across pole histories and calculating new ADS distributions using Eq. (1)) using a t-test. See SI Sec. S4.7 for more details.

[1] M. Hashimoto, T. Nozoe, H. Nakaoka, R. Okura, S. Akiyoshi, K. Kaneko, E. Kussell, and Y. Wakamoto, Proceedings of the National Academy of Sciences **113**, 3251 (2016).
[2] A. J. Lee, S. Wang, H. R. Meredith, B. Zhuang, Z. Dai, and L. You, Proceedings of the National Academy of Sciences **115**, 4069 (2018).
[3] O. Gefen and N. Q. Balaban, FEMS microbiology reviews **33**, 704 (2009).
[4] S. Vedel, S. Tay, D. M. Johnston, H. Bruus, and S. R. Quake, Proceedings of the National Academy of Sciences of the USA **110**, 129 (2013).
[5] A. Eldar and M. B. Elowitz, Nature **467**, 167 (2010).
[6] R. Pugatch, Proc. Nat. Acad. Sci. USA **112**, 2611 (2015).
[7] D. A. F. Fuentes, P. Manfredi, U. Jenal, and M. Zampieri, Nature Communications **12**, 3204 (2021).
[8] A. M. Proenca, C. U. Rang, C. Buetz, C. Shi, and L. Chao, Nature Communications **9**, 3722 (2018).
[9] J. Borgqvist, N. Welkenhuysen, and M. Cvijovic, Scientific Reports **10**, 1556 (2020).
[10] E. J. Stewart, R. Madden, G. Paul, and F. Taddei, PLoS Biology **3**, e45 (2005).
[11] S. Vedel, H. Nunns, A. Košmrlj, S. Semsey, and A. Trusina, Cell Systems **3**, 187 (2016).
[12] M. Coelho, A. Derelu, A. Haese, S. Kühn, L. Malinovska, M. E. DeSantis, J. Shorter, S. Alberti, T. Gross, and I. M. Tolić-Nørrelykke, Curr. Biol. **23**, 1844 (2013).
[13] N. Erjavec, M. Cvijovic, E. Klipp, and T. Nyström, Proc. Nat. Acad. Sci. USA **105**, 18764 (2008).
[14] L. Chao, PLoS Genetics **6**, e1001076 (2010).
[15] J. Winkler, A. Seybert, L. König, S. Pruggnaller, U. Haselmann, V. Sourjik, M. Weiss, A. S. Frangakis, A. Mogk, and B. Bukau, EMBO Journal **29**, 910 (2010).
[16] A. B. Lindner, R. Madden, A. Demarez, E. J. Stewart, and F. Taddei, Proc. Nat. Acad. Sci. USA **105**, 3076 (2008).
[17] M. Ackermann, L. Chao, C. T. Bergstrom, and M. Doebeli, Aging Cell **6**, 235 (2007).
[18] E. Ising, Z. Phys. **31**, 253 (1925).
[19] J. M. Yeomans, *Statistical mechanics of phase transitions* (Oxford, 1992).
[20] S. Leibler and E. Kussell, Proc. Nat. Acad. Sci. USA **107**, 13183 (2010).
[21] E. Kussell and S. Leibler, Science **309**, 2075 (2005).
[22] Y. Wakamoto, A. Y. Grosberg, and E. Kussell, Evolution **66**, 115 (2011).
[23] A. Amir, Phys. Rev. Lett. **112**, 208102 (2014).
[24] A. Marantan and A. Amir, Phys. Rev. E **94**, 012405 (2016).
[25] J. Lin and A. Amir, Phys. Rev. E **101**, 012401 (2020).
[26] N. N. Taleb, *Antifragile: Things that gain from disorder* (Random House, 2012).
[27] B. Charlesworth, Genetics **156**, 927 (2000).
[28] E. Levien, J. Kondev, and A. Amir, J. R. Soc. Interface **17**, 20190827 (2020).
[29] E. Levien, T. GrandPre, and A. Amir, Phys. Rev. Lett. **125**, 048102 (2020).
[30] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (U.S. Department of Commerce, 1972).
[31] S. Schindler, S. Tuljapurkar, J.-M. Gaillard, and T. Coulson, Theo. Popul. Biol. **82**, 244 (2012).
[32] A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, and N. Barkai, Nat. Genetics **38**, 636 (2006).
[33] R. H. Chisholm, T. Lorenzi, A. Lorz, A. K. Larsen, L. N. de Almeida, A. Escargueil, and J. Clairambault, Cancer research **75**, 930 (2015).
[34] A. Brauner, O. Fridman, O. Gefen, and N. Q. Balaban, Nature Reviews Microbiology **14**, 320 (2016).
[35] N. Q. Balaban, S. Helaine, K. Lewis, M. Ackermann, B. Aldridge, D. I. Andersson, M. P. Brynildsen, D. Bumann, A. Camilli, J. J. Collins, *et al.*, Nature Reviews Microbiology **17**, 441 (2019).
[36] T. Bergmiller, A. M. Andersson, K. Tomasek, E. Balleza, D. J. Kiviet, R. Hauschild, G. Tkačik, and C. C. Guet, Science **356**, 311 (2017).
[37] S. K. Govers, J. Mortier, A. Adam, and A. Aertsen, PLoS biology **16**, e2003853 (2018).
[38] T. S. Ursell, J. Nguyen, R. D. Monds, A. Colavin, G. Billings, N. Ouzounov, Z. Gitai, J. W. Shaevitz, and K. C. Huang, Proceedings of the National Academy of Sciences **111**, E1025 (2014).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SINatureCommsoriginalsubmission.pdf