Demographic Consequences of Damage Dynamics in Single-Cell Ageing

Murat Tuğrul* and Ulrich K Steiner
Institute of Biology, Freie Universität Berlin, Königin-Luise-Str. 1-3, 14195 Berlin, Germany
(Dated: May 23, 2023)

Ageing is driven by the accumulation of diverse types of damage that leads to a decline in function over time. In unicellular organisms, in addition to this damage accumulation within individuals, asymmetric partitioning of damage at cell division might also play a crucial role in shaping demographic ageing patterns. Despite empirical single-cell studies providing quantitative data at the molecular and demographic level, a comprehensive understanding of how cellular damage production and partition propagate and influence demographic patterns is still lacking. To address this gap, we present a generic and flexible damage model using a stochastic differential equation approach which incorporates stochastic damage accumulation and asymmetric damage partitioning at cell divisions. We provide an analytical approximation linking cellular and damage parameters to demographic ageing patterns. Interestingly, we observe that the lifespan of the cells follows an inverse-gaussian distribution whose statistical properties can be expressed with cellular and damage parameters, as well as easily inferred from empirical single-cell data. Furthermore, we demonstrate how stochasticity (noise) in damage production and asymmetry in damage partitioning contribute to shaping lifespans. Applying the model to empirical *E.coli* data reveals non-exponential scaling in mortality rates, which cannot be captured by classical Gompertz-Makeham models. Additionally, we highlight the essential role of stochastic division times in shaping lifespans. Our findings provide a deeper understanding of how fundamental processes contribute to cellular damage dynamics and generate demographic patterns. The generic nature and flexibility of our damage model offer a valuable framework for investigating ageing in diverse biological systems.

INTRODUCTION

Ageing disrupts cellular integrity and leads to functional decline in cells and individuals and affects fitness components like reproduction and survival [López-Otín *et al.* 2013]. Although the consequences of ageing are ranging from evolution to medicine, a thorough quantitative and mechanistic understanding is lacking that links molecular to demographic levels. Molecular roots of ageing are considered due to damage accumulation at cells disrupting cellular processes where damage types and sources can be diverse including DNA oxidation, protein aggregation, mitochondrial dysfunction, mutations, etc. [Gladyshev 2016, Schumacher *et al.* 2021]. Cellular damage drives the ageing of organisms and consequently shapes demographic patterns, which are surprisingly diverse across species and populations [Jones *et al.* 2014]. Molecular studies of ageing mostly focused on genetics and environments and provided important knowledge [López-Otín *et al.* 2013] but there appears to be also a neglected stochastic component in ageing [Steinsaltz *et al.* 2020]. Individual differences (heterogeneity) in ageing is observed in genetically identical twin and inbred animal studies [Finch and Kirkwood 2000]. At cellular level, inherently stochastic processes such as gene expression [Elowitz *et al.* 2002, Raj and Oudenaarden 2008, Patange *et al.* 2018] and asymmetric partition of cellular content at cell divisions [Huh and Paulsson 2011a,b, Shi *et al.* 2020] might produce stochasticity in cellular damage accumulations which in turn result in individual differences. Deciphering the quantitative relationship between stochastic damage dynamics and individual heterogeneity in ageing requires high-throughput single-cell studies and mathematical models.

In their seminal work, Stewart *et al.* [2005] showed that *E.coli*, a morphologically symmetric dividing microbe, is also susceptible to reproductive ageing. This illustrates that ageing goes beyond asymmetric and sexually reproducing organisms and provides us with a model organism to study individual heterogeneity in ageing with available molecular, genetic and imaging techniques. Bacterial ageing has been extensively studied after Stewart's seminal work (for review, see [Steiner 2021]). Experimental studies mostly elaborated on cellular growth and division in ageing by studying external stress conditions [Wang *et al.* 2010, Rang *et al.* 2011, 2012, Łapińska *et al.* 2019] and molecular factors such as chaperones associated with protein aggregation [Winkler *et al.* 2010, Proenca *et al.* 2019]. Limited theoretical models of bacterial ageing focused on various aspects such as the optimality of asymmetric damage partitioning for maximising fitness [Watve *et al.* 2006, Evans and Steinsaltz 2007]; the existence of distinct reproductive growths [Chao 2010, Blitvić and Fernandez 2020], and damage repair [Clegg *et al.* 2014]. It was only recently that longer experiments were also conducted to study mortality (cell death) events of bacterial mother cells which inherit the older cell-wall during divison [Jouvet *et al.* 2018, Steiner *et al.* 2019]. These studies also confirmed chronological ageing in bacteria by showing that mortality (i.e. lifespan, survival) profiles deviate from random mortality process. However, it

* murat.tugrul@fu-berlin.de

is still not clear how stochastic damage dynamics leads to individual differences in ageing and produces such demographic patterns.

Existing mathematical models for bacterial ageing mostly rested within an equilibrium assumption and did not exploit underlying dynamical aspects influencing ageing and demography patterns. In demography, lifespan-related data, e.g. how mortality chances are changing with age, have been traditionally interpreted with phenomenological models such as the Gompertz-Makeham model. This does not only fail to capture any non-exponential mortality rates but usually does not help understand damage mechanisms as the connection between molecular and demographic parameters cannot be known. Anderson [2000] and Weitz and Fraser [2001] independently developed stochastic models of vitality using a stochastic differential equation (SDE) approach which contributed to our understanding of ageing patterns such as mortality plateaus in various organisms. To our knowledge, a similar SDE modelling framework has not been adopted for single-cell or bacterial ageing, or remained minimal in applications [Yang *et al.* 2023]. Such direct extensions especially require dealing with asymmetric damage partition, i.e. sudden jumps in damage, which might play a crucial role in single-cell ageing. Here we contribute to developing such a general framework by adapting jump-diffusion type SDEs to single-cell and bacterial ageing. We construct a simple illustrative SDE to model damage dynamics along mother bacterial cells and derive its lifespan-related demographic characteristics. We then explore how demographic outcomes are influenced by initial conditions and different parameters of cellular damage: production rate and noise, and partition asymmetry. Lastly, we use a dataset for *E.coli* with mortality events [Steiner *et al.* 2019] to demonstrate how such a model can help us interpret bacterial ageing.

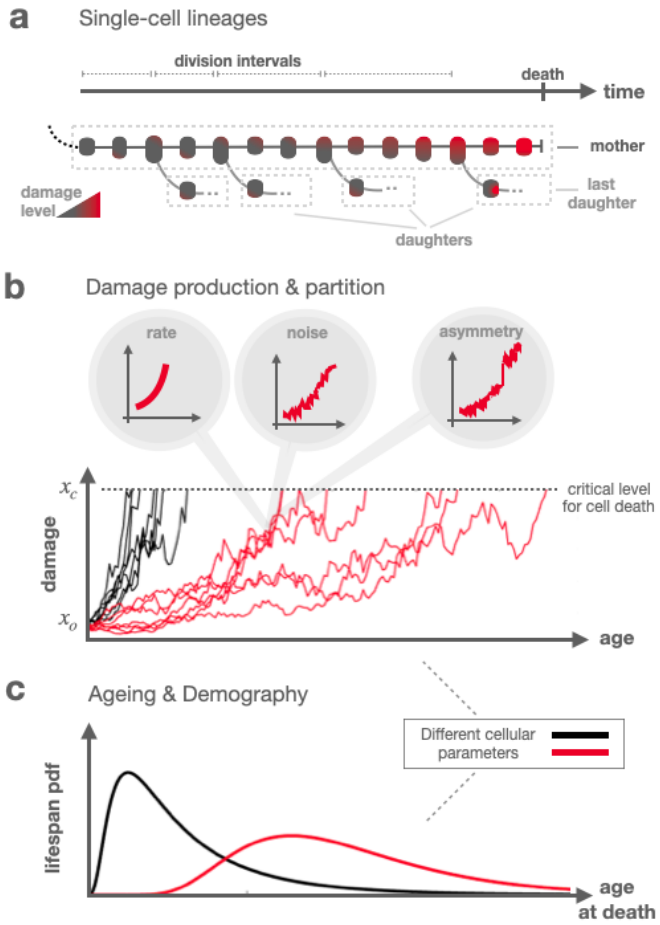

FIG. 1. **a)** We model the damage and ageing dynamics of single-cell mother (old pole) lineages, that typically receives more damage in cell divisions. **b)** We consider a stochastic differential equation framework to model damage dynamics taking account of damage production and partition mechanisms. Each path in the figure exemplifies a damage trajectory realisation in a mother lineage where black and red colourings correspond to different cellular parameter choices. Trajectories start from the same initial damage and terminate at different times at crossing a critical damage level which defines the time of death, i.e. the lifespan. **c)** The parameters settings define the damage dynamics that in turn dictate the ageing dynamics and shape demographic patterns like the lifespan distribution.

MODEL

We build a modelling framework based on the assumption that single-cells deteriorate over time due to the accumulation and transmission of *damage* entities which upon reaching a critical value cause cell death (mortality). We consider a damage density measure, i.e. number of damage bodies per cell volume, including the cell growth, dilution and repair processes. In this manuscript, we focus on modelling the cellular damage dynamics along mother cell lineages which receive the older cell wall at cell divisions and typically receive more damage than the daughter cells (**Fig. 1-a**). Our damage dynamics model considers two important processes: damage production and damage partition (**Fig. 1-b**). The damage production encapsulates a deterministic net rate and a stochastic noise for random fluctuations. The damage partition describes the sudden jumps due to asymmetric cell content (cytoplasm) partition during a division event whose timings are stochastic and drawn from a fixed distribution. Once the parameters of damage dynamics are set, individual stochastic life trajectories are determined by imposing initial and final (absorbing) damage levels (**Fig. 1-b**). In this manuscript, we focus on a special case where all the parameters are constants. In other words, damage will influence the lifespan but will not affect cellular growth and division mechanisms in our model. This choice is mostly for the sake of mathematical derivations and serves as a standard stochastic model to link cellular

TABLE I. Key Notations

Notation	Description
x_t	stochastic variable for damage density at time t
x_o	initial damage
x_c	critical damage
ε	damage production rate
σ	noise strength in damage production
γ	asymmetry level (0 for symmetry)
r	mean division rate
σ_r	standard deviation of division rate
W_t	Wiener (Brownian) process
J_t	Jump process
s	survival probability
f	lifespan probability density
h	mortality rate
VMR	Variance-to-Mean Ratio (Fano factor)

damage and demography. Future works may relax this assumption and explore the model extensively.

In mathematical terminology, we consider a jump-diffusion dynamics [Merton 1976, Kou 2002, Kou and Wang 2003] which, in its most general form, is expressed by the following stochastic differential equation (SDE)

$$dx_t = \varepsilon(x_t,t)\,x_t\,dt \;+\; \sigma(x_t,t)\,x_t\,dW_t \;+\; \gamma(x_t,t)\,x_t\,dJ_t. \tag{1}$$

Here, the damage density x along a lineage realisation at time t is a stochastic variable and denoted as x_t (**Fig. 1-b**). $\varepsilon\,(x_t,t)$ is the deterministic damage production rate. $\sigma\,(x_t,t)$ is the strength of noise which is modelled with a Wiener (Brownian) process where W_t is a gaussian random variable with mean and variation as 0 and t, respectively. The timing of division events and direction/magnitude of damage jumps are modelled with the random process J_t and the parameter $\gamma(x_t,t)$, respectively. The $\gamma(x_t,t)$ is the sudden damage jump (asymmetry) parameter at cell division where $\gamma=0$ is for symmetric division and $1>\gamma>0$ is for a mother cell receiving more damage than daughter cell. During a cell division, damage density x_t will jump to a new damage $x_t\,(1+\gamma)$ along the mother cell lineage, whereas a new daughter cell will be produced with damage density $x_t\,(1-\gamma)$. Note that $-1<\gamma<0$ corresponds to the mother's rejuvenation, i.e. receiving less damage (our modelling is also valid for this regime but the results will not be explored here). Cell damage level starts from an initial level x_0 at time $t=0$ and the cell mortality is defined by the first crossing of a critical (absorbing) level x_c. Note also that time and age are identical for this manuscript therefore t refers to both of them.

Under constant parameter choice, the SDE is reduced to a simpler form:

$$dx_t = \varepsilon\,x_t\,dt \;+\; \sigma\,x_t\,dW_t \;+\; \gamma\,x_t\,dJ_t, \tag{2}$$

where an explicit solution is known as

$$x_t = x_o\,e^{(\varepsilon-\frac{1}{2}\sigma^2)\,t+\sigma W_t}\;\prod_{j=1}^{n_t}(1+\gamma). \tag{3}$$

with n_t being a random variable for the total number of division (jump) events until time t in a realisation. The following iterative equation can be used for simulations of this jump-diffusion dynamics.

$$x_{t+\Delta t} = x_t\,e^{(\varepsilon-\frac{1}{2}\sigma^2)\,\Delta t+\sigma\sqrt{\Delta t}\,Z_t}\,(1+\gamma)^{dn_t} \tag{4}$$

where Z_t is a random number from a unit normal distribution and the jump process takes $dn_t=1$ or 0 at each discrete time step depending on the jumping process J_t. One can simulate many individual paths and obtain a statistical description but we further seek an analytical expression and rewrite the last term of the exact solution in Eq. 3 as $\exp\big(\log(\prod_{j=1}^{n_t}(1+\gamma))\big) = \exp\big(\sum_{j=1}^{n_t}\log(1+\gamma)\big) = \exp\big(n_t\log(1+\gamma)\big)$. Eq. 3 is therefore transformed into

$$x_t = x_o\,e^{(\varepsilon-\frac{1}{2}\sigma^2)\,t+\sigma W_t}\,e^{n_t\log(1+\gamma)} \tag{5}$$

If we assume the number of divisions per unit time (i.e. division rate) follows a gaussian distribution $\mathcal{N}(r,\sigma_r^2)$, then n_t follows also a gaussian distribution $\mathcal{N}(rt,\sigma_r^2 t)$. This assumption allows us to rewrite the above expression in an approximate compact solution form

$$x_t = x_o\,e^{\widetilde{\varepsilon}\,t+\widetilde{\sigma}W_t} \tag{6}$$

where the modified drift and diffusion terms are

$$\widetilde{\varepsilon} = \varepsilon - \frac{1}{2}\sigma^2 + r\log(1+\gamma) \tag{7}$$

$$\widetilde{\sigma} = \sqrt{\sigma^2+\sigma_r^2\,(\log(1+\gamma))^2} \tag{8}$$

Therefore, the original SDE in Eq. 2 is approximated to the following geometric Brownian motion

$$dx_t = (\widetilde{\varepsilon}+\frac{1}{2}\widetilde{\sigma}^2)\,x_t\,dt \;+\; \widetilde{\sigma}\,x_t\,dW_t \tag{9}$$

We can also express it in arithmetic Brownian form as $d(\ln x_t/x_o) = \widetilde{\varepsilon}\;dt + \widetilde{\sigma}\,dW_t$, from which we obtain the probability density function for damage level $p(x,t)$ by applying the Fokker-Planck equation with initial and boundary conditions of $p(x,t=0)=\delta(x-x_o)$ and $p(x=x_c,t)=0$ as

$$p(x,t|x_o,x_c,\widetilde{\varepsilon},\widetilde{\sigma}) = \frac{1}{\sqrt{2\pi\widetilde{\sigma}^2 t}}\Big[\exp\Big(-\frac{\big(\ln(x/x_o)-\widetilde{\varepsilon}\,t\big)^2}{2\widetilde{\sigma}^2\,t}\Big) - \exp\Big(\frac{2\ln(x_c/x_o)\,\widetilde{\varepsilon}}{\widetilde{\sigma}^2} - \frac{\big(\ln(x/x_o)-2\ln(x_c/x_o)-\widetilde{\varepsilon}\,t\big)^2}{2\widetilde{\sigma}^2\,t}\Big)\Big]. \tag{10}$$

Expressing the probability density function for lifespans $f(t)$ is equivalent to determining the first passage time to the absorbing level x_c. Interestingly it follows a distribution of an inverse gaussian (IG) as

$$f(t|x_o,x_c,\widetilde{\varepsilon},\widetilde{\sigma}) = \frac{\ln(x_c/x_o)}{\sqrt{2\pi\,\widetilde{\sigma}^2\,t^3}}\exp\Big(\frac{-\big(\widetilde{\varepsilon}\,t-\ln(x_c/x_o)\big)^2}{2\,\widetilde{\sigma}^2\,t}\Big) \tag{11}$$

Following Weitz and Fraser [2001], we make a connection from *molecular* (cellular) to *macroscopic* (demographic) parameters by rewriting this 4 parameter expression in terms of 2 parameters as

$$f(t|\mu,\tau) = \sqrt{\frac{\tau}{2\pi t^3}}\exp\Big(\frac{-\tau\big(t-\mu\big)^2}{2\mu^2 t}\Big) \tag{12}$$

where μ and τ are respectively the mean and shape parameters and can be expressed in terms of cellular parameters as

$$\mu = \ln(x_c/x_o)/\widetilde{\varepsilon} \tag{13}$$

$$\tau = \Big(\frac{\ln(x_c/x_o)}{\widetilde{\sigma}}\Big)^2 \tag{14}$$

The variance of IG distribution is μ^3/τ which is in terms of cellular parameters

$$Var = \ln(x_c/x_o)\,\widetilde{\sigma}^2/\widetilde{\varepsilon}^3 \tag{15}$$

We observe that $\ln(x_c/x_o)$ plays a role of time scaling for the mean and variance of the lifespan distribution. The variance-to-mean ratio (VMR, also named the index of dispersion or Fano factor) of lifespans therefore becomes a damage-free measure, i.e.

$$\mathrm{VMR} = \big(\widetilde{\sigma}/\widetilde{\varepsilon}\big)^2 \tag{16}$$

It can be helpful to interpret empirical data and molecular mechanisms as initial and critical damage levels can be difficult to measure. It quantifies the dispersion level of the lifespans and compares it to an underlying Poisson-like mechanism where VMR $= 1$ (i.e. as if all death events are due to an external mortality). VMR < 1 and VMR > 1 indicate under- and over-dispersed distributions, respectively.

The IG nature of lifetimes was first discovered in hazard analysis by Chhikara and Folks [1977] but the importance in demography was highlighted by Anderson [2000] and Weitz and Fraser [2001], to our knowledge. Thanks to the IG form, we can express the survival probability density easily as

$$s(t|\mu,\tau) = 1 - \frac{1}{2}\Big(\mathrm{erfc}\,\big(-\sqrt{\frac{\tau}{2t}}\,(t/\mu-1)\big) + \exp(2\tau/\mu)\,\mathrm{erfc}\,\big(\sqrt{\frac{\tau}{2t}}\,(t/\mu+1)\big)\Big) \tag{17}$$

where erfc is the complementary error function. The mortality rate function, i.e. $h(t) = \frac{f(t)}{s(t)} = \frac{-d\log s(t)}{dt}$ can also be expressed as

$$h(t|\mu,\tau) = \frac{2\sqrt{\frac{\tau}{2\pi t^3}}\exp\big(\frac{-\tau(t-\mu)^2}{2\mu^2 t}\big)}{2 - \Big(\mathrm{erfc}\,\big(-\sqrt{\frac{\tau}{2t}}\,(t/\mu-1)\big) + \exp(2\tau/\mu)\,\mathrm{erfc}\,\big(\sqrt{\frac{\tau}{2t}}\,(t/\mu+1)\big)\Big)} \tag{18}$$

Lastly, we also integrate a damage-independent mortality term into our model by adding a constant killing rate h_o to the mortality rate function. The modified functions for mortality rate, survival and lifespan distribution respectively become

$$\widetilde{h}(t|\mu,\tau) = h(t|\mu,\tau) + h_0. \tag{19}$$

$$\widetilde{s}(t|\mu,\tau) = s(t|\mu,\tau)\,e^{-h_0\,t} \tag{20}$$

$$\widetilde{f}(t|\mu,\tau) = f(t|\mu,\tau)e^{-h_o\,t} + s(t|\mu,\tau)\,h_0\,e^{-h_0\,t} \tag{21}$$

Note that the obtained mortality rate function is different than the classical phenomenological Gompertz-Makeham law of mortality where it has a simple exponential form

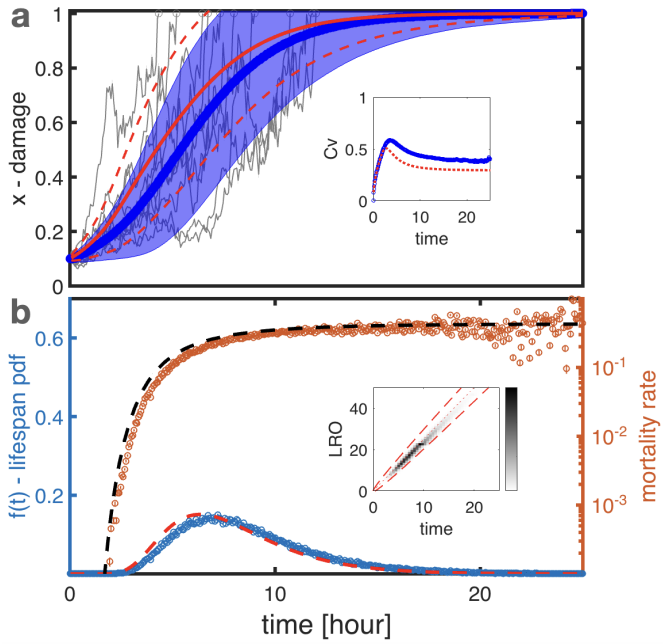

FIG. 2. The dynamics of damage and ageing according to our model is demonstrated with parameters $x_o = 0.1$, $x_c = 1.0$, $\varepsilon = 0.1$, $\sigma^2/\varepsilon = 1.0$, $\gamma = 0.1$, $r = 2.5$, $\sigma_r = 0.5$ by using both the simulations of the exact solution ($N = 10^5$, $\Delta t = 1/15$) and the analytic expression of the approximation method. **a)** The grey damage trajectories are randomly selected realisations in the simulation. The blue circles and shading correspond to the mean and one standard deviation of 10^5 realisations (the death cells are fixed to x_c for this calculation). The red solid and dashed curve stand for the mean and one standard deviation obtained from the analytic expression (Eq 10) by numerical integration. The inset plots the coefficient of variation of the damage levels for the living cells (i.e. the death cells are excluded) where the blue circles and the dashed curves correspond the the simulation and analytical results. **b)** Left y-axis: the lifespan probability density function is shown with the light blue circles (simulations) and the red dashed curve (analytic expression, Eq. 11). Right y-axis: the same data is represented with the hazard function (mortality rate) measure with the brown circles (simulations) and the black dashed curve (analytic expression, Eq. 18). The inset shows the lifetime reproductive output, LRO, (i.e. the total number of offspring of an individual until death) of a mother lineage that ends (mother cell dies) at time t. As parameterised, LRO, it is distributed along the linear curve rt (red dotted line) with a gaussian dispersion of a variation $\sigma_r^2 t$ (red dashed line: ± 3 sd).

with three parameters (α, β, λ):

$$h_{GM}(t) = \alpha e^{\beta t} + \lambda. \tag{22}$$

RESULTS

Biologically relevant parameter ranges

We obtain our results considering biologically relevant parameter ranges oriented on bacterial species. Division rate r in bacteria varies depending on conditions where *E.coli* in rich media have $r \sim 3.0$ divisions per hour whereas under natural conditions it is much slower, including frequent growth arrest. The damage partition asymmetry is controlled by parameter γ assuring that mothers take over more damage. We consider the full range of asymmetry levels, i.e. $\gamma = 0 - 1$ where 0 refers to full symmetry and 1 full asymmetry at cell divisions. $\gamma \sim 1$ can be a very rare situation requiring an active energy-dependent process, as has been suggested for *S. cerevisae* [Kaeberlein 2010]. The damage production rate ε and fluctuation σ are hard to estimate and probably are integrated into complex gene expression dynamics and gene regulatory networks taking into account of transcription, translation, decay and dilution mechanisms that are, in addition, influenced by the environment. We start with the condition that σ is related to ε with a constant noise factor. Here we consider a relative noise measure similar to the variance-to-mean (also called Fano factor or noise-to-signal) as $F_f = \sigma^2/\varepsilon$, which takes values from no noise ($F_f = 0$) to high noise ($F_f = 2$). The net damage rate may range from $\varepsilon = 0$, i.e. when unwanted damage entities are efficiently eliminated or diluted, to an extremely high value where the damage level quickly escalates resulting in sudden death before the cell divides again.In order to set a numeric value for the critical damage production rate, we can refer to Eq. 13 where the case of $\varepsilon \ll r\log(1+\gamma)$ states that the dominant damage drift factor is the asymmetric division process. Considering a typical value of division rate $r = 2 - 3$ and maximal asymmetry $\gamma = 1$ we argue that $\varepsilon \sim 1$ reflects high damage production. We consider $\varepsilon \gg 1$ for extreme damage rate, referring to a sudden death case which is trivial and not interesting for ageing studies, but might be of relevance for evaluating the efficacy of anti-infective substances that do not allow single cells to escape treatment. We consider $\varepsilon \sim 0.1$ for an example of a low damage production rate. The logarithmic ratio between the critical and initial damage level $\ln(x_c/x_o)$ scales time (see Eq. 13 and Eq. 15 for the mean and variance of lifespan distribution), so whenever suitable we present our results with this time scaling unit.

Analytical expressions vs numerical simulations for the dynamics of damage and ageing

In **Fig 2** we show an example of the combined dynamics of damage and ageing for particular parameter choices. The results are obtained both from our analytical approximation approach and from numerical simulations obtained using the discrete-time version of the explicit solution for the SDE model (Eq. 4). **Fig 2-a** illustrates how cellular damage level, and its substantial among lineage variance in damage levels, changes over time from a defined initial value until reaching a critical value defining the time of death in our model. The demographic statistics are shown in **Fig 2-b** with a lifespan distribution that follows an inverse-gaussian dis-

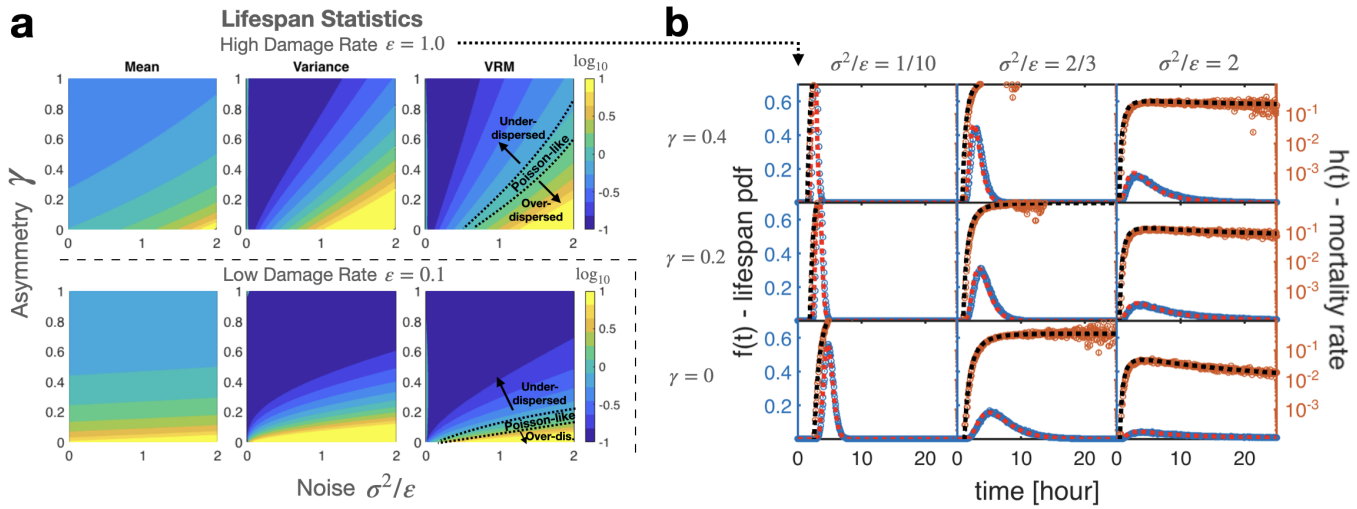

FIG. 3. **a)** Our analytic expression maps molecular damage parameters to demographic statistics. The figure here shows how relative noise in damage production (measured with σ^2/ε) and damage partition asymmetry level γ determine the lifespan statistics of mean, variance and variance-to-mean ratio (VMR). Settings for this figure are $r = 2.5$ and $\sigma_r = 0.5$, and we plot the results (contours) with $\ln x_c/x_o$ scaling in $\log_{10}$ scale. The top and bottom panel rows are for high and low damage rates, respectively ($\varepsilon = 1.0$ and $= 0.1$). In the VMR plots we show the $VMR \sim 1$ region corresponding to Poisson-like dispersion as well as under- and over-dispersion regions ($VMR < 1$ and > 1, respectively). **b)** For high damage rate ($\varepsilon = 1.0$, , top panels of a) and selected σ^2/ε and γ values, the lifespan distributions and the mortality rate curves are shown as in Fig. 2 ($x_o = 0.01$, $x_c = 1.0$, $N = 10^5$, $\Delta t = 1/15$).

tribution and a mortality rate measure that exemplifies a rapid early-age increase followed by a constant probability of death at older ages. Additionally, **Fig 3-b** highlights the demographic fates for different parameter ranges. The small difference between numerical simulations and analytic approximation is partially due to time-discretisation as seen in the error convergence investigation in SI (Fig. 6). Overall, our approximate mathematical solution is in close agreement with the numerical simulations obtained with the more explicit solution and proves to be insightful to understand how molecular parameters influence demographic patterns.

Lifespan statistics for different asymmetry and noise levels

Our analytical derivation provides an exact mapping from cellular damage parameters to demographic statistics where lifespan distribution follows an inverse gaussian function whose mean and shape parameter are expressed in terms of cellular parameters (i.e. Eqs. 13, 14). Using this mapping, we can ask how asymmetry in damage partition and noise in damage production influence the lifespan distribution. In **Fig 3-a** we show how different lifespan statistics (mean, variation, variance-to-mean) change with the noise and asymmetry parameters in low and high damage rate cases. For the case of high damage production, both noise and asymmetry influence the lifespan distribution significantly. Higher asymmetry facilitates quicker damage deposition along mother lineages and therefore shortens both the mean and variation of lifespan. Similarly increase in damage production noise will result in a drastic shift and spread of lifetimes. For the case of low damage production, the noise factor affects less, and a quick damage accumulation and early mortality is only possible if there is significant partition asymmetry, otherwise, cell lineages are likely observed as very long living, even as immortal if they start from initially low damage.

As a damage-free measure, the variance-to-mean ratio (VMR) of the lifespan distribution provides additional insights into the underlying mechanisms. The dispersion level of a standard Poisson process is a reference point as VMR = 1. We can draw a direct link from molecular relative noise level ($F_f = \sigma^2/\varepsilon$) to observed lifespan dispersion by adapting the Eq. 16 as

$$\mathrm{VMR} = \frac{(F_f + \sigma_r^2/\varepsilon \log(1+\gamma))^2}{\left(1 - \frac{1}{2}F_f + r/\varepsilon \log(1+\gamma)\right)^2} \qquad (23)$$

In case of complete damage partition symmetry (i.e. $\gamma = 0$), the above expression is reduced to $\mathrm{VMR} = \left(2F_f/(2-F_f)\right)^2$. Interestingly this sets a threshold damage noise level as $F_f = \sigma^2/\varepsilon = 2/3$ which leads to a lifespan distribution as if a Poisson-like mortality mechanism holds. **Fig 3-a** indicates which cellular parameter regions exhibit a Poisson-like behaviour as if lifespans are generated by a constant and external killing mechanism (see **Fig 3-b** for flat mortality rates corresponding this VMR ~ 1 region). VMR > 1 and VMR < 1 parameter regions correspond to over- and under-dispersed lifespan distributions with respect to a Poisson process.

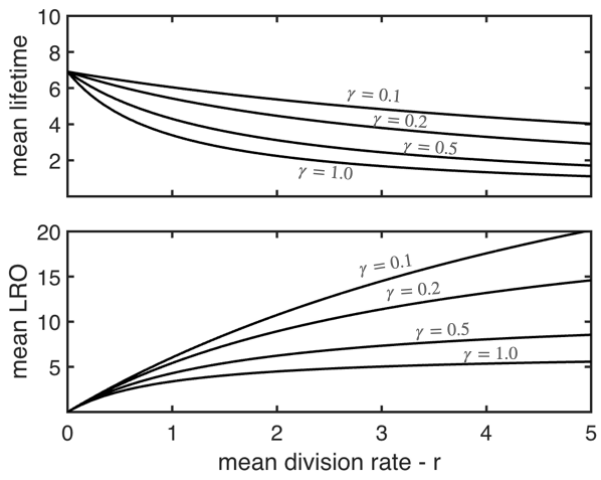

FIG. 4. A demonstration of how expected lifetime and lifetime reproductive output change with mean division rate r. Different asymmetry parameters $\gamma = 0.1,\ 0.2, 0.5, 1.0$ are used for each corresponding curves. Other fixed parameters are $x_o = 0.01$, $x_c = 1.0$, $\varepsilon = 1.0$, $\sigma^2/\varepsilon = 2/3$.

Fig 3-b shows the actual lifespan distribution and mortality rate curves for selected parameters in high damage rate in our model and exemplifies the capability of producing different mortality rate scalings. They demonstrate the cases of a rapidly increasing scaling, an increasing scaling with a plateau and a first increasing then decreasing scaling where the last is not a possible pattern in classical exponential models such as in the Gompertz-Makeham model but observed in nature [Jones *et al.* 2014, Steiner *et al.* 2019, Weitz and Fraser 2001].

Mean division rate vs mean lifespan

Furthermore, our model provides also a direct relationship between mean division rate and mean lifespan (Eq. 13). Keeping all other parameters fixed and having an asymmetric partition case ($\gamma > 0$), a higher division rate will result in higher damage accumulation in mother lineages causing earlier death events, i.e. shorter lifespans, as demonstrated in **Fig 4-a**. This observation is in agreement with the classical pattern observed between the reproduction and longevity [Kirkwood and Rose 1991]. However, reducing the mean division rate is likely not a permanent escape from mortality in an evolutionary strategy as it trades off with the fitness components such as the lifetime reproductive output, i.e. the number of offspring in the entire lifespan (**Fig 4-b**).

Application to empirical data:

Using our model we revisit previously published single-cell *E.coli* ageing data where the authors [Steiner *et al.* 2019] reported mortality and division events of mother and last daughter cell lineages (a last daughter is the last offspring of a mother cell lineage, see **Fig. 1-a**). In **Fig. 5-a**, we plot the demographic statistics (i.e. the lifetime, survivorship and mortality rate) of this dataset. The mother and the last daughter lineages exhibit clearly distinct ageing features as originally emphasised by the authors. The lifespans of the mother lineages show a spread distribution (average lifetime: $\sim$ 13 hours) and the mortality rate curve exhibits a late-age plateau after an early-age exponential increase. In contrast, the lifetime distribution of the last daughter lineages exhibits a sharp decreasing exponential decline (i.e. majority of the last daughters die in the first hours, average lifetime: $\sim$ 7 hours). Dissimilar initial damage between these two lineage cohorts was considered the main factor behind the intriguing difference, however, mechanistic damage models are needed for better elucidation. Therefore, we fit both our damage model (using the modified version including a constant external killing rate, i.e. Eq. 21), as well as the Gompartz-Makeham model (Eq. 22) for a comparison, to the lifespan datasets using a maximum likelihood method (see the caption of **Fig. 5** for the details of fitting). Our damage model catches the mother lineages's non-exponential mortality curve that cannot be captured by a Gompartz-Makeham model. It also better encapsulates the excess of last daughter deaths at early ages.

In our model, the logarithmic ratio of critical and initial damage levels $\ln(x_c/x_o)$ is directly proportional to the average lifetimes. In a simple scenario where all other parameters are kept constant, an expected increase in the initial damage of the daughter lineages in comparison to initial damage of the mother lineages can decrease the average lifespan significantly. But does our model support the initial damage argument as the only fact underlying the difference in the observed lifespan distributions? The variance-to-mean (VMR) measures the dispersion of lifespans and according to our model is a damage-free measure (Eq.16). If the observed differences depend only on the differences in the initial damage, we expect to see a similar dispersion level. However, the calculated dispersion values are highly distinct, i.e. VMR $= 3.6$ and 8.2 respectively for mother and last daughter lineages (based on Variation/Mean in the lifespan dataset directly). These values represent slightly over-dispersed and highly over-dispersed distributions in comparison to a standard Poisson distribution. Note that if our model's fitted distribution is used in VMR calculations, the results even indicate a bigger difference in dispersion characteristics (mother: VMR $= 2.1\,(1.3 - 3.7)$ and last daughter: VMR $= 29.5\,(13.2 - 63.0)$). Therefore, other factors are expected to contribute to the differences in lifespan statistics in light of our model.

Upon close inspection of our derived expression for VMR in our damage model, i.e. Eq. 16, we see that the mean and variation of division rates influence the dispersion. Therefore we checked the reported average division rates along each lineage in this dataset as shown in **Fig. 5-b** for both mother and last daughter lineages. It is striking that average division rates are very sim-

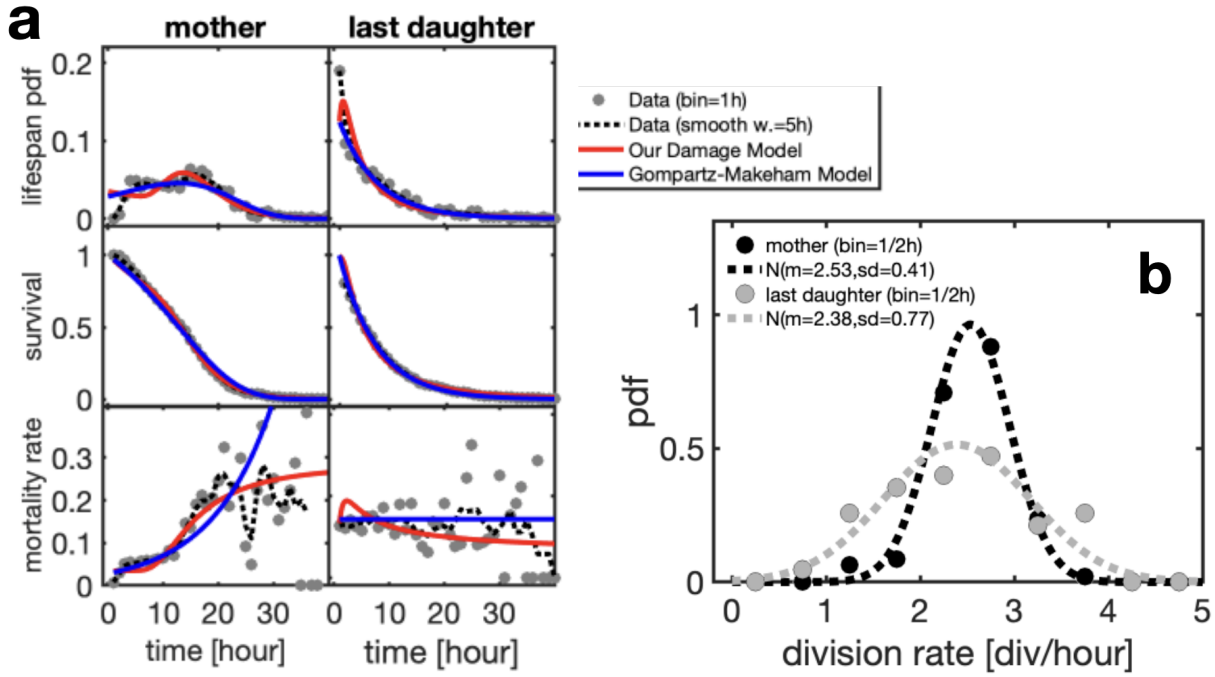

FIG. 5. Demographic data for *E.coli* mother and last daughter lineages [Steiner *et al.* 2019]. **a)** We replot the lifespan data as lifetime pdf, survival probability and mortality rate for both the mother ($N = 516$) and last daughter ($N = 516$) lineages. Mean $\pm$ standard deviation of lifetimes: 13.4 ± 6.9 for the mothers and 7.0 ± 7.6 for the last daughters. Note that we added 1.4 hour to the reported lifetimes to compensate for pre-recorded time durations (see SI of Steiner et al 2019 et al.). The grey circles and the dashed curves are for the binned data (1 hour) and smoothed data (5 hours). The red and blue curves are for the fitted models using maximum likelihood estimation (with 95% CI), respectively, for our damage model (mother: $\hat{\mu} = 18.3\,(17.2{-}194)$, $\hat{\tau} = 159.1\,(101.7{-}216.5)$, $\hat{h}_o = 0.035\,(0.027{-}0.042)$; last daughter: $\hat{\mu} = 12.8\,(9.7{-}15.9)$, $\hat{\tau} = 5.5\,(4.0{-}7.1)$, $\hat{h}_o = 0.042\,(0.032 - 0.052)$) and the Gompartz-Makeham model (mother: $\hat{\alpha} = 0.03(0.02 - 0.04)$, $\hat{\beta} = 0.09(0.07 - 0.10)$, $\hat{\lambda} = 0.0(-0.01 - 0.01)$; last daughter: $\hat{\alpha} = 0.14(0.09 - 0.19)$, $\hat{\beta} = 0.0(-0.01 - 0.01)$, $\hat{\lambda} = 0.0(-0.04 - 0.04)$). Note that we confined the constant killing rate parameter (h_o and λ) estimation of the last daughters into the obtained 95%CI estimatates of the mothers. **b)** The division rate histograms ($bin = 1/2$ hour) and the fitted gaussian pdfs are shown for the mother (black, $N = 93$) and last daughter (gray, $N = 85$) lineages. The mean value for both distributions is very similar: $2.53\,(2.45 - 2.62)$ and $2.38\,(2.22 - 2.55)$ div/hour but the variations differ substantially: sd$= 0.41\,(0.36 - 0.48)$ vs $0.77\,(0.67 - 0.91)$ div/h (the parentheses are 95%CI in estimations).

ilar ($\sim$ 2.4 – 2.5 div/hour), probably due to general cellular restrictions as mentioned above, but the variation in division rates are very distinct with a ratio of $\sigma_r^{2(l)}/\sigma_r^{2(m)} = 3.5$. This ratio is in agreement with the observed VMR increase in the last daughter lineages according to our Eq. 16. Although average division rates are similar, more widespread division timings might have contributed to dispersed damage accumulation patterns via asymmetric damage partition in different stochastic realisations of the last daughters, therefore, providing an over-dispersed lifetime distribution with a substantial number of early deaths as interestingly observed for the last daughter lineages. Although this is not a rigorous quantitative justification due to unknown damage parameters (i.e. ε and σ) in this dataset, our model is still instructive to guide the interpretation of the empirical data.

DISCUSSION

Single-cell ageing studies, which successfully produced remarkable quantitative data in recent years, need also complementary mathematical models to draw the causative link between cellular damage and ageing at demographic levels. However, we still need a better understanding of how damage characteristics affect the patterns of ageing and demography. This especially requires a bridging from the stochasticity at the cellular level to individual heterogeneity at a larger scale, even when genetic and environmental factors are fixed. To fill such a gap, we introduced a generic damage model using stochastic differential equations (SDEs) that is flexible to hold and explore further various damage facts or assumptions. Our damage model takes the parameters for damage production rate and noise strength as well as damage partition asymmetry and outputs stochastic damage trajectory along a mother cell lineage. For this

manuscript, we focused on a constant parameter case of the SDE (Eq. 2) and successfully obtained interpretable analytical solutions for important measures like the probability density of damage (Eq. 10) and the probability density of lifespans (Eq. 12) as a function of age and in terms of molecular damage parameters. Our analytical expressions successfully match numerical simulations (as exemplified in **Fig 2** and **Fig 3-b**) and prove to be therefore instructive in an understanding of how *molecular* damage dynamics scale up to *macroscopic* demographic patterns. Remarkably, our modelling with jump-diffusion SDE results in the lifespan distribution having an inverse-gaussian (IG) form (Eq.12) where its statistics can be expressed explicitly with cellular damage parameters as well as inferred from empirical lifetime data using standard methods like maximum likelihood. This allows one to connect molecular damage parameters to demographic measures in a direct way, providing more mechanistic estimates in cellular processes. IG lifespan distribution and its potential application in demographic data were already discussed by Anderson [2000] and Weitz and Fraser [2001] with their arithmetic Brownian SDE modelling of vitalities for classical demographic data. Our modelling successfully extended their approach to bacterial or single-cell ageing by incorporating damage partition asymmetry. As highlighted in **Fig 3**, we showed how asymmetry and relative noise affect the lifespan distribution where particular arrangements for these parameters can yield in different lifespan characteristics. This not only determines the early and late life mortality expectations but also shapes the distribution with respect to a standard Poisson-like process (as similar, or over- and under-dispersed lifespans). Such diverse lifespan behaviours likely alter the population and evolutionary dynamics as well. Apart from the mathematical contributions we developed, our modelling can be insightful for quantitative understanding in bacterial ageing as such in immortality conditions of bacterial cells [Wang *et al.* 2010, Rang *et al.* 2011, 2012, Łapińska *et al.* 2019]. We quantitatively exemplified in **Fig 3** that if the bacterial cells start from damage-free initial levels and are kept under low damage production conditions, one can observe very large lifespans of a mother lineage unless a very extreme asymmetry exists. Typical demographic models such as the Gompartz-Makeham model are largely phenomenological not providing mechanistic interpretations and also are limited to exponential rates for mortality missing for example a mortality plateau pattern as observed in the bacterial lifespan data [Steiner *et al.* 2019] and captured by our model as revealed in **Fig 5-a**. In this empirical data application, we also revealed in **Fig 5-b** how stochasticity (different variation levels) in division times can be affected in mother and last daughter cell lineages with different lifespan characteristics. This might indicate a different transmission of heterostasis to offspring born to older individuals, probably triggering the noise in division decisions in cells.

Our approach directly models the stochastic dynamics of cellular damage and is not bound to an equilibrium assumption as largely held in previous bacterial ageing modellings [Watve *et al.* 2006, Evans and Steinsaltz 2007, Chao 2010, Blitvić and Fernandez 2020], therefore, it is more promising to explore or address various phenomena responsible for individual (cell-to-cell) heterogeneity observed in ageing. Recent studies with single-cells show rich intrinsic cell-to-cell variation in gene expression [Elowitz *et al.* 2002, Foreman and Wollman 2020, Urchueguía *et al.* 2021]. One possible application is to relate damage to gene regulatory context and make a connection from cell-to-cell variation in gene expression. The relation between gene expression and ageing has been investigated mostly based on transcriptome analyses on humans and other model organisms. The importance of down-regulation of genes encoding mitochondrial proteins; downregulation of the protein synthesis machinery; dysregulation of immune system genes; reduced growth factor signalling; constitutive responses to stress and DNA damage; dysregulation of gene expression and mRNA processing have been already underlined [Frenk and Houseley 2018]. However, to our knowledge, no connection between stochastic dynamics of gene expression and individual heterogeneity in ageing has been investigated. Our modelling can be also insightful for designing experiments to understand the context of cellular damage better. Simultaneous measurements of both cell fates and damage-related gene expression signals are promising [Sampaio *et al.* 2022] where we can utilise the stochastic features of gene regulation to understand the mechanisms and feed our modelling choices.

The analyses in this manuscript were limited to a constant parameter case in the generic stochastic differential equation setting. Future applications can relax this assumption and explore, at least with numerical simulations, biologically more plausible scenarios. Time or damage-dependent parameters can be integrated into our model as damage production rate, noise and asymmetry levels, as well as division rate, might alter with cellular damage levels. As we also see in our application to the empirical data (**Fig 5**), division rate characteristics are likely influenced by damage. Such extension will not only make our modelling framework more comparable to existing models by others [Chao 2010, Blitvić and Fernandez 2020] but also will provide elaborative results on evolutionary trade-offs between reproduction and longevity as extensively discussed in evolution of ageing [Kirkwood and Rose 1991]. Our assumption on a critical damage level defining the time of cell death can also be altered in future applications, for example, by using a killing rate depending on the damage level such as in the form of a linear function [Evans and Steinsaltz 2007], or of a Hill function as typically observed in cellular environments due to the thermodynamics of finite particles. A more

direct extension of our modelling and analyses would be to project damage dynamics into population dynamics. This will require following the damage dynamics of an entire population descending from a single *mother* cell where each newborn daughter is itself a mother lineage starting from a different initial damage. Albeit challenging in rigorous calculations, approximations may bring us valuable grounds for answering population-related and evolutionary questions such as damage distribution in population or the optimal damage characteristics for a higher fitness [Watve *et al.* 2006, Evans and Steinsaltz 2007]. This theoretical achievement will complement the recent attempts to collect data simultaneously at single-cell tracking and batch population [Bakshi *et al.* 2021] and will advance our evolutionary understanding.

ACKNOWLEDGEMENTS

We thank Mathias Franz, Audrey M. Proenca and Shripad Tulkapurkar for comments on a previous draft of the manuscript. US was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 430170797 as a Heisenberg Fellow. MT was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 430174701 and by the European Union under the Marie Skłodowska-Curie Actions' European Postdoctoral Fellowship grant agreement – 101069035.

REFERENCES

C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, and G. Kroemer, Cell **153**, 1194 (2013), publisher: Elsevier.
V. N. Gladyshev, Aging Cell **15**, 594 (2016).
B. Schumacher, J. Pothof, J. Vijg, and J. H. J. Hoeijmakers, Nature **592**, 695 (2021), number: 7856 Publisher: Nature Publishing Group.
O. R. Jones, A. Scheuerlein, R. Salguero-Gómez, C. G. Camarda, R. Schaible, B. B. Casper, J. P. Dahlgren, J. Ehrlén, M. B. García, E. S. Menges, P. F. Quintana-Ascencio, H. Caswell, A. Baudisch, and J. W. Vaupel, Nature **505**, 169 (2014).
D. Steinsaltz, M. D. Christodoulou, A. A. Cohen, and U. K. Steiner, in *Encyclopedia of Biomedical Gerontology*, edited by S. I. S. Rattan (Academic Press, Oxford, 2020) pp. 386–394.
C. E. Finch and T. B. L. Kirkwood, *Chance, Development, and Aging* (Oxford University Press, 2000).
M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Science **297**, 1183 (2002), publisher: American Association for the Advancement of Science.
A. Raj and A. v. Oudenaarden, Cell **135**, 216 (2008), publisher: Elsevier.
O. Patange, C. Schwall, M. Jones, C. Villava, D. A. Griffith, A. Phillips, and J. C. W. Locke, Nature Communications **9**, 5333 (2018).
D. Huh and J. Paulsson, Proceedings of the National Academy of Sciences **108**, 15004 (2011a), publisher: Proceedings of the National Academy of Sciences.
D. Huh and J. Paulsson, Nature Genetics **43**, 95 (2011b), number: 2 Publisher: Nature Publishing Group.
C. Shi, L. Chao, A. M. Proenca, A. Qiu, J. Chao, and C. U. Rang, Proceedings of the Royal Society B: Biological Sciences **287**, 20200569 (2020), publisher: Royal Society.
E. J. Stewart, R. Madden, G. Paul, and F. Taddei, PLOS Biology **3**, e45 (2005), publisher: Public Library of Science.
U. K. Steiner, Frontiers in Cell and Developmental Biology **9** (2021), 10.3389/fcell.2021.668915, publisher: Frontiers.
P. Wang, L. Robert, J. Pelletier, W. L. Dang, F. Taddei, A. Wright, and S. Jun, Current Biology **20**, 1099 (2010), publisher: Elsevier.
C. U. Rang, A. Y. Peng, and L. Chao, Current Biology **21**, 1813 (2011), publisher: Elsevier.
C. U. Rang, A. Y. Peng, A. F. Poon, and L. Chao, Microbiology **158**, 1553 (2012), publisher: Microbiology Society.
U. Łapińska, G. Glover, P. Capilla-Lasheras, A. J. Young, and S. Pagliara, Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **374**, 20180442 (2019).
J. Winkler, A. Seybert, L. König, S. Pruggnaller, U. Haselmann, V. Sourjik, M. Weiss, A. S. Frangakis, A. Mogk, and B. Bukau, The EMBO Journal **29**, 910 (2010).
A. M. Proenca, C. U. Rang, A. Qiu, C. Shi, and L. Chao, PLOS Biology **17**, e3000266 (2019), publisher: Public Library of Science.
M. Watve, S. Parab, P. Jogdand, and S. Keni, Proceedings of the National Academy of Sciences **103**, 14831 (2006), publisher: National Academy of Sciences Section: Biological Sciences.
S. N. Evans and D. Steinsaltz, Theoretical Population Biology **71**, 473 (2007).
L. Chao, PLOS Genetics **6**, e1001076 (2010), publisher: Public Library of Science.
N. Blitvić and V. I. Fernandez, Journal of Theoretical Biology **502**, 110331 (2020).
R. J. Clegg, R. J. Dyson, and J.-U. Kreft, BMC Biology **12**, 52 (2014).
L. Jouvet, A. Rodríguez-Rojas, and U. K. Steiner, Oikos **127**, 728 (2018).
U. K. Steiner, A. Lenart, M. Ni, P. Chen, X. Song, F. Taddei, J. W. Vaupel, and A. B. Lindner, Evolution **73**, 847 (2019).
J. J. Anderson, Ecological Monographs **70**, 445 (2000), publisher: Ecological Society of America.
J. S. Weitz and H. B. Fraser, Proceedings of the National Academy of Sciences **98**, 15383 (2001), publisher: National Academy of Sciences Section: Biological Sciences.
Y. Yang, O. Karin, A. Mayo, A. Bar, X. Song, P. Chen, A. L. Santos, A. B. Lindner, and U. Alon, "Damage dynamics in single E. coli and the role of chance in the timing of cell death," (2023).
R. C. Merton, Journal of Financial Economics **3**, 125 (1976).
S. G. Kou, Management Science **48**, 1086 (2002), publisher: INFORMS.
S. G. Kou and H. Wang, Advances in Applied Probability **35**, 504 (2003), publisher: Cambridge University Press.
R. S. Chhikara and J. L. Folks, Technometrics **19**, 461 (1977), publisher: Taylor & Francis.
M. Kaeberlein, Nature **464**, 513 (2010), number: 7288 Publisher: Nature Publishing Group.

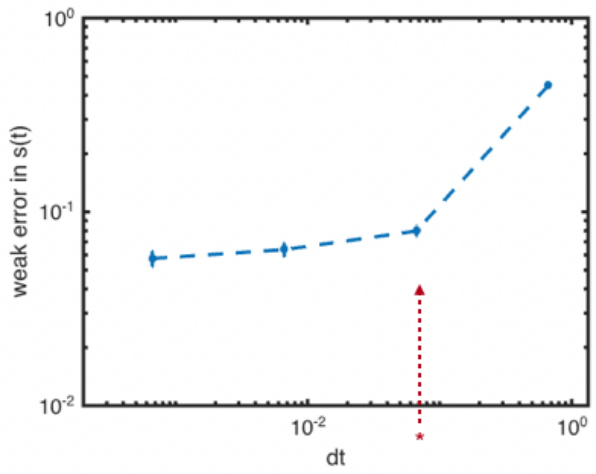

FIG. 6. Error convergence is presented here. The simulations with $N = 10^4$ realisations are conducted with the same parameter selection as in Fig. 2 for different Δt values (the asterix shows the value ($\Delta t = 1/15$) used in the main text). The weak error is calculated by obtaining $abs(Max_t(s_{anal,t} - s_{simul,t}))$. The error bars show the mean and standard obtained by repeating the simulations 10 times.

T. B. L. Kirkwood and M. R. Rose, Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **332**, 15 (1991), publisher: Royal Society.
R. Foreman and R. Wollman, Molecular Systems Biology **16**, e9146 (2020), publisher: John Wiley & Sons, Ltd.
A. Urchueguía, L. Galbusera, D. Chauvin, G. Bellement, T. Julou, and E. v. Nimwegen, PLOS Biology **19**, e3001491 (2021), publisher: Public Library of Science.
S. Frenk and J. Houseley, Biogerontology **19**, 547 (2018).
N. M. V. Sampaio, C. M. Blassick, V. Andreani, J.-B. Lugagne, and M. J. Dunlop, Proceedings of the National Academy of Sciences **119**, e2115032119 (2022), publisher: Proceedings of the National Academy of Sciences.
S. Bakshi, E. Leoncini, C. Baker, S. J. Cañas-Duarte, B. Okumus, and J. Paulsson, Nature Microbiology **6**, 783 (2021), number: 6 Publisher: Nature Publishing Group.

I- SUPPORTING INFORMATION