The Strehler-Mildvan correlation is nothing but a fitting artifact

Andrei E. Tarkhov[a,b,*], Leonid I. Menshikov[a,c], Peter O. Fedichev[a,d]

[a] *Gero LLC, Novokuznetskaya street 24/2, Moscow 119017, Russia,*
[b] *Skolkovo Institute of Science and Technology, Skolkovo Innovation Centre, Nobel street 3, Moscow 143026, Russia,*
[c] *Northern (Arctic) Federal University, 163002, Severnaya Dvina Emb. 17, Arkhangelsk, Russia,*
[d] *Moscow Institute of Physics and Technology, 141700, Institutskii per. 9, Dolgoprudny, Moscow Region, Russia*

Abstract

Gompertz empirical law of mortality is often used to parametrize survival fraction as a function of age with the help of mere two parameters: the Initial Mortality Rate (IMR) and the Gompertz exponent slope, inversely related to the Mortality Rate Doubling Time (MRDT). Furthermore, the two parameters are believed to be related through the Strehler-Mildvan (SM) correlation, which suggests that the more animals die during adolescence, the longer is the lifespan of the survived individuals. Even though there have been quite a few doubts expressed against the very existence of the correlation, it is still widely believed that the theory of ageing and mortality behind the SM correlation provides a mechanism-based explanation of Gompertz's law. In this Letter, we concentrate on uncertainties of identification of the Gompertz parameters by fitting from experimental survival records. We show, that whenever the number of animals in the experimental cohorts is insufficiently large, the use of least-squares fitting is rather a convenience for data visualization, than a robust procedure to amalgamate and smooth discrete data. We present an analytical explanation behind specific difficulties accompanying the fit once the average lifespan of the species exceeds MRDT. We demonstrate, that under such conditions a fit to the Gompertz law becomes unstable, and fails to produce a unique combination of the demographic parameters. In fact, one gets the whole degenerate manifold of the Gompertz parameters, which is nothing else but the line, corresponding to the proper approximate value of the average lifespan, and, at the same time, coincides with the SM correlation. We show that the employment of the age-independent Makeham mortality term does not resolve the degeneracy problem. Therefore, we have to conclude, that the average lifespan persists as the only stable feature, which can be reliably inferred from survival statistics in an ex-

*Corresponding author
Email address: at@gero.com (Andrei E. Tarkhov)

periment with a finite number of animals. The SM correlation, in this case, is nothing but an iso-average-lifespan degenerate manifold on the IMR versus α plane, is thus a fitting artifact, and its observation in a data set is an indication of insufficient statistical power, rather than a biological reality.

Keywords: aging parameters, Gompertz law, Gompertz-Makeham law, Strehler-Mildvan correlation, fitting parameters degeneracy

---

1. Introduction

Aging in most animals, including humans, leads to nearly an exponential increase of mortality $M(t) = M_0 \exp(\alpha t)$ with age t, the dependence commonly referred to as the Gompertz law [1]. It has been long assumed, based on both empirical evidence and theoretical arguments, that the Initial Mortality Rate, M_0 or IMR, and the Gompertz exponent slope, α, are universally related by the Strehler-Mildvan (SM) correlation

$$\ln M_0 - \ln K = -\alpha/B. \tag{1}$$

Here K and B are constants, introduced in [2], and α is inversely proportional to the Mortality Rate Doubling Time (MRDT). The SM correlation is not a trivial fact and is inconsistent with a recently observed scaling law, implying that the surviving fractions of nematodes worms under different environmental conditions or mutants with different lifespans can be cast into a universal function of a properly rescaled dimensionless age [3]. The observed temporal scaling cannot be consistent with the SM correlation since it implies the proportionality between IMR and α, rather than of lnIMR and α, as predicted by Strehler and Mildvan. In addition, the SM correlation does not hold well against a recently proposed model of aging, where the above-mentioned scaling law is a natural consequence of a genetic instability of regulatory gene networks in sufficiently long lived organisms [4]. We believe that the two observations together are sufficient to cast doubt on the biological origin of the SM correlation and investigate a possibility of its artifactual nature.

To make sense of the apparent contradiction, we re-analyzed the problem of the Gompertz parameters inference from experimental survival fraction curves (lifetables). To our surprise, we found out, that the commonly employed least squares procedure leads to an ill-defined non-linear optimization problem, which fails to yield a unique solution. In practical terms this means that one is always left with the family of parameters M_0 and α forming a degeneracy manifold, coinciding with the SM correlation with a few percent accuracy. The analytical treatment of the fitting difficulty provides a clear explanation of the apparent correlation: the exact analytical form of the SM relation is nothing else but a set of possible demographic parameters consistent with the experimentally derived average lifespan of the animals. This is why we have to conclude, that the SM correlation is nothing but a fitting artifact occurring whenever the number of

animals in experimental cohorts is not sufficiently large. The same conclusion probably holds for the biological explanations behind the SM correlation.

We further support our theoretical arguments by demonstrating the least squares procedure instability using survival plots, obtained in a recent series of experiments in C. elegans under the influence of lifespan modifying treatments [5]. We show, that the immediately apparent SM correlation in the experiment is a consequence of the fitting procedure uncertainty exacerbated by a small and finite number of animals in the cohorts. We show, that it is only possible to fit robustly the average lifespans of small cohorts, yet it is extremely difficult to yield reliable estimates of the Gompertz parameters unless the number of the animals in the experimental groups is large enough to produce a unique solution.

2. The Strehler-Mildvan correlation is a degenerate manifold

To investigate the influence of a factor, such as a therapy, or a mutation, on aging, one may want to estimate the effects of the experimental design conditions on aging model parameters, such as, for example, commonly used quantities M_0 and α of the Gompertz law. A natural way to achieve the goal is to analyze Kaplan-Meyer plots and fit an experimentally observed fraction of the animals surviving by the age t, $N(t)$, onto the model prediction. A typical behavior of the survival fraction is qualitatively depicted on Fig.1A. Usually, the average lifespan, $\bar{t}$, exceeds the mortality rate doubling time and the survival fraction is a very sharp, almost a step-wise function, dropping from 1 to 0 in a short interval of ages, $\triangle t/\bar{t} \ll 1$, around its mean value $\bar{t}$, so that $N(\bar{t}) = C \sim 0.5$. This means that $\bar{t} = t_{1/2}\left[1 + O\left(\triangle t/\bar{t}\right)\right]$. According to the Gompertz law, the mortality at a given age, t, is $M(t) = M_0 \exp(\alpha t)$, therefore the fraction of the animals alive by the same age is given by the expression $N_G\left(t|\alpha, M_0\right) = \exp\left[-M\left(e^{\alpha t} - 1\right)\right]$, where $M = M_0/\alpha \lesssim 1$.

To fit an experimentally observed dependence $N(t)$ to the Gompertz law, it is natural to minimize the approximation mean squared error. Normally, the surviving fraction is known at a series of values at discrete age-points. We will assume, that the observation points are sufficiently many in number and the number of the animals in the cohorts is large, so that the discrete set of the observations $N(t)$ is a smooth function of age. Under the circumstances, the objective function can be represented as the integral

$$J = \frac{1}{2}\int_0^\infty dt\left[N_G\left(t|\alpha, M\right) - N\left(t\right)\right]^2, \tag{2}$$

to be minimized with respect to the values of α and M_0. The full analysis of the optimization problem is presented in Appendix A. Let us summarize the results important for the discussion below. First of all, the Gompertz parameters are the solutions of the system of equations

$$\frac{\partial J}{\partial \alpha} = \frac{\partial J}{\partial M} = 0. \tag{3}$$

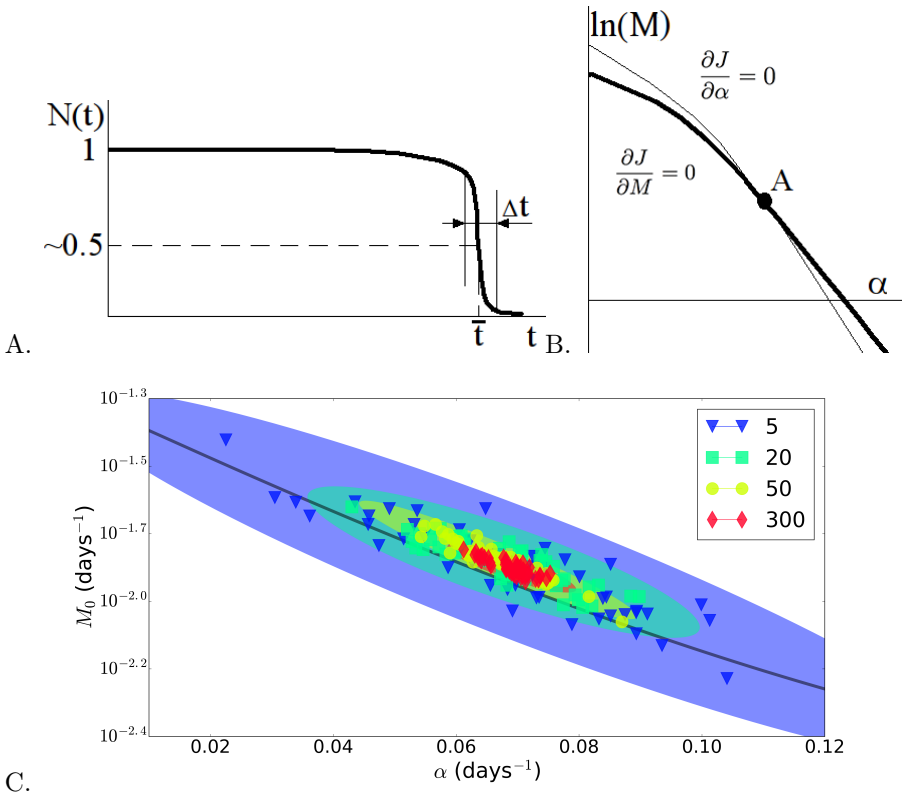

Figure 1: A. A schematic representation of a typical step-wise survival fraction curve; B. A graphic solution of Eqs. (3) in $\ln M$ and α plain. Here $\bar{t}$ is the average lifespan of the animals. The thick and the thin lines correspond to the vanishing derivatives of the objective function J with respect to the parameters M and α; C. The estimates of the Gompertz parameters M_0 versus α obtained by the least-squares fit (2) for the sub-populations of the wild type animals selected at random from the experimental data from [5] with the number of cohorts in sub-populations $N_{coh} = 5, 20, 50$ and 300. Each dot represents a fit using data from a single random sub-population, the colors mark N_{coh}. The black line represents an iso-average-lifespan curve for the wild-type average lifespan.

A graphical solution of the equations is represented in Fig.1B. The slopes of the curves representing the solutions of Eqs. (3) at the intersection point differ by a mere factor C_2/Λ^2, with $C_2 \sim 1$ being a numerical factor. For sufficiently long lived animals, $\Lambda \gtrsim 1$, and the curves get very close to each other near the intersection point, defining the solution. Since actual experimental data is intrinsically noisy, instead of a single and well-defined solution, one rather gets the whole degenerate set of possible of solutions, even if the biological conditions are exactly the same. Remarkably, each of the possible fits correspond to the correct value of the average lifespan, $t_{exp} = \int dt N(t)$. More specifically, every pair of IMR and α values, satisfying

$$t_G(\alpha, M) = t_{exp}, \tag{4}$$

can, in principle, be obtained from the analysis of different realizations of the same experiment. Since $t_{exp} \approx -\alpha^{-1} \ln M$ in the Gompertz model, the degeneracy manifold is approximately defined by the condition $\ln M_0 - \ln \alpha = -\alpha t_{exp}$, and therefore the condition (4) is practically indistinguishable from the SM correlation (1) whenever $\Lambda \gtrsim 1$. A more accurate approximation for the average lifespan $t_G(\alpha, M)$ from the Gompertz law and therefore the precise form of the SM correlation is derived in the appendix, see Eq. (B.1).

3. Survival statistics analysis example

To support our theoretical arguments and highlight the conclusions using realistic experimental data, we investigated the Kaplan-Meier plots from [5] experiment, an important example of a modern high-throughput screening for life-span modulating compounds. The dataset contains information on life histories in 1416 experiments of genetically homogeneous populations, actually the same strain of C. elegans, all obtained in the same laboratory. Each experimental cohort consists of roughly 10 worms, some experiments have been repeated to achieve higher significance.

In total, the dataset provides us with 768 control cohorts for the calculations. We employed these cohorts to estimate the uncertainty of the Gompertz fit parameters in the ideal case of a genetically homogeneous population under identical conditions in a particular laboratory. To show a purely artifactual nature of the SM correlation in this case, we imitated a typical aging population study with this data. A subset of N_{coh} cohorts (we used $N_{coh} = 5$, 20, 50 and 300) was selected at random from all the available aging cohorts, after that the survival plots of the chosen subset were averaged out to obtain a smooth characteristic survival curve, which was finally fitted to the Gompertz curve. This procedure was repeated a number of times to get a reliable estimate of the fitting parameters uncertainty for a chosen N_{coh}. Fig. 1C shows that the SM correlation exists even in this ideal experiment and it is linked to a subset size N_{coh}: the more cohorts one uses to calculate a smooth survival curve of a chosen sub-population, the less the degeneracy of the Gompertz parameters, hence we have to conclude that the nature of the SM correlation is purely

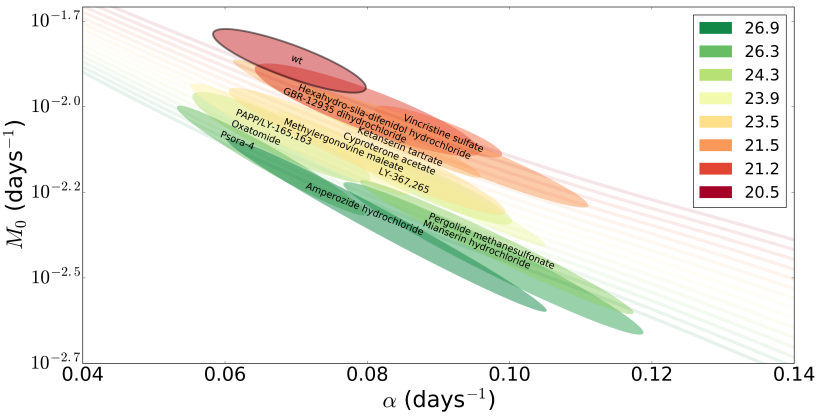

Figure 2: The IMR (M_0) versus the Gompertz exponent slope α estimates obtained by the least-squares procedure using lifetables from the C. elegans experiment [5]. The colored curves are the asymptotic expressions, the analytical solutions, corresponding to the iso-average-lifespan manifolds defined by Eq. B.2. Each colored ellipse (from green to red) corresponds to a single experiment, a cohort of worms under each of the several chosen treatments. The colors mark the average lifespans in days. The position and the size of the black-edged ellipse indicate the best Gompertz fit parameters and their uncertainty for the wild type and the treated groups (the estimates of the variance were obtained using models obtained from cohorts of smaller sizes with randomly selected animals).

artifactual since N_{coh} is the only varying parameter in this calculation. The SM correlation apparently arises due to the insufficient number of animals. The suppression of the fitting parameters fluctuations in the direction, orthogonal to the SM curves is the direct consequence of the degeneracy of the solutions of Eq. 3. This means, that, even though in theory, the optimization problem may have a unique solution, in practice, it is defined as an intersection point of the two curves with nearly coinciding slopes, which differ by a quadratically small in Λ term. Therefore, in a wide range of the fitting parameters there is efficiently only one equation, an iso-average-lifespan curve, for the determination of both parameters, hence there must exist some correlation between IMR and α, which is conventionally referred to as the SM correlation.

Of all the available treatments we have chosen medicines with the largest numbers of repeats and the smallest degeneracy of the Gompertz parameters, for which the number of replicas is adequate to resolve the effects of the treatments on lifespans. For each group of the experiments we averaged out survival curves for randomly chosen 2/3 of all the available replicas and attempted the Gompertz fit to obtain the parameters M_0 and α. We repeated the calculations by reshuffling the animals and estimated the parameters variance. The uncertainty of the results is represented by the sizes of the characteristic ellipses on M_0 versus α plane in Fig.2. To highlight the instability of the Gompertz fit parameters even for the absolutely homogeneous ageing cohorts, we plotted a black

ellipse corresponding to the wild-type worms after reshuffling in groups of 300 cohorts, but the degeneracy still leads to a relative error in the determination of the Gompertz parameters of nearly 25%. The iso-average-lifespan curves (the SM correlation curves) are plotted and colored according to the age as shown in the inset. For various animal groups, the average lifespan is, of course, slightly different, that is why the ellipses are to a small degree distributed along the direction orthogonal to the SM curves, whereas they are substantially stretched along the SM curves. This fact shows that small fluctuations of a survival curve for a homogeneous group of animal cohorts, due to a finite number of these cohorts, lead to the strong fluctuations of the fitting parameters along the SM curves, whereas fluctuations orthogonal to these SM curves are substantially suppressed.

4. Conclusions and Discussion

We conclude our analysis by the observation that a fit of experimental lifetables on an aging model, such as Gompertz law, may be very challenging and lead to flawed results. Our analytical treatment shows, that the solution of the fitting problem is not unique in practice because it is defined as an intersection point of the null curves of the derivatives of the fitting error with respect to the demographic parameters, having practically coinciding slopes. The SM correlation turns out to be a consequence of the fitting failure. In fact, the relation between the IMR and MRDT is nothing but the iso-average-lifespan curve. The exact form of the relation can be established analytically, though it is visually barely distinguishable from a straight line. Realistic noisy experimental data could be easily misinterpreted to produce such a correlation, which was, in fact, initially "observed" and and commonly referred to as the SM correlation [2] ever since. Therefore the immediate conclusion from our research is that the two-parametric Gompertz fit is an ill-defined procedure, and it should be used with care, especially if the number of animals in the groups is not very high.

Next to the solution of the minimization problem, at the intersection point A in Fig. 1B the derivatives of the objective function Eq. 2 with respect to the Gompertz variables $i = \alpha, M$ behave as $\alpha - \alpha_A \approx k_i \left(M - M_A\right)$, where k_i are the two numerical factors and are very close to each other. The problem is not limited to the specific choice of the aging model parametrization. One can, in principle, try to perform the same calculation using any other pair of quantities instead of M and α. For example, a possible choice could be the average lifespan $t_G \approx \ln\left(1/M\right)/\alpha$. In this case $\alpha - \alpha_A \approx K_i \left(t_G - t_G^A\right)$, where $K_i = k_i \left(dM/dt_G\right)_A$. The relative difference of the slopes stays the same under such a change of the variables, and remains to be quadratically small

$$\xi = \frac{K_1 - K_2}{K_1} = \frac{k_1 - k_2}{k_1} = O(\frac{1}{\Lambda^2}). \tag{5}$$

Exactly for this very same reason, the degeneracy of the solution cannot be lifted by a change of the fitting function, such as for example by the use of

mortality rate instead of survival fraction in the objective function. Indeed, as shown in Appendix A, the mortality rate function should be used with an extra care, since it can only be obtained by a numerical differentiation of experimental survival lifetables, which are by no means continuous and differentiable. The difficulty is even worse at advanced ages, when the number of the surviving animals becomes very small. We note, that the difficulties in the aging model parameters inference, highlighted in this Letter, are not limited only to the Gompertz law fit. The actual reason behind the large uncertainty of the results stems from the very abrupt change of the experimental surviving fraction near the average lifespan of the group.

A possibility of a superficial nature of the SM correlation has been already highlighted in by L. and N. Gavrilovs in [6, 7], where a more sophisticated aging model, the Gompertz-Makeham version of the mortality law, was invoked to produce a more stable fit. Later on the Makeham term was claimed to have a fundamental biological significance to describe some regimes of aging with a large age-independent component of mortality [8]. In Appendix C we prove, that the problem with the degeneracy of the solution does not go away if extra parameters, such as the age-independent mortality rate, are added into the procedure. Of course, the Gompertz-Makeham form of the mortality law may help to fit the experimental data in a different aging regime with a significant age-independent component in mortality rate, whenever it is required. However, the model fit with a limited number of animals in the cohort would give a degenerate set of solutions anyway.

In conclusion, let us use our analytical results to comment on the scale of the expected survival statistics analysis difficulties for different species. For experiments with C. elegans, for instance, we estimate $\Lambda \sim 2$, and therefore $\xi \sim 0.2$. It means that the Gompertz parameters α and M_0 can only be found using the regression solutions of Eq. 2 if a large, but still, a reasonable number of animals is included into the aging cohorts under the investigation. More specifically, what is important is the number of animals N_s in an age bin near the average age, where the steepest change in survival fraction occurs. To make a rough estimate of the required number of the animals, we assume the characteristic error of a mean experimental survival curve to be its standard deviation, which scales as $\sim O\left(\sqrt{N_s}\right)$. Since the null curves for the derivatives of the objective function differ in slopes by a factor $\sim O\left(1/\Lambda^2\right)$, we obtain the estimate $N_s = O\left(\Lambda^4\right)$. For simplicity, let us assume the proportionality coefficient to be of the order of one and is the same for all species. For C. elegans the size of the parameters spread of the wild-type cohorts in the Figure 2 suggests that ~ 300 animals in the middle-age bin may be sufficient to determine the Gompertz exponent with roughly 25% accuracy.

For human subjects, though, the requirements are much more challenging, since the Gompertz logarithm can be as large as $\Lambda \sim 8$, and therefore the minimum number of humans in the age bin from 50 to 70 years, where most of the subjects die, can be estimated to be as large as $N_c \sim 75000$. One can easily imagine, that the minimum required population of a country required for

a robust calculation is as large as a few million people. On top of that, realistic populations may be genetically inhomogeneous and live under very different environmental conditions along their life histories. These factors combined may prevent any solid analysis using any model as simple as the two-parameter Gompertz mortality law. We believe, that the outlined fitting difficulty on pair with limited statistical power due to the inclusion of countries with smaller populations into the analysis led to the spurious correlation. Indeed, the error in the parameter α determination of the order of 25% in the original paper [2], corresponds to $N_s \sim 100,000$ people and is in line with the expectation given by the spread of the countries populations.

After more than half a century after the publication of the original work [2], the idea of the SM correlation and biological picture behind it is deeply rooted in ageing studies (see e.g. [9, 10, 11, 12] as examples of recent works). Most notably, heterogeneity of ageing populations was recently investigated and invoked to explain the SM correlation [9, 13]. Remarkably, we show that the SM correlation can be obtained even in absolutely homogeneous populations of C. elegans worms, simply because the SM correlation in realistic experiments may be purely artifactual!

A number of theories, suggesting some physical or biological background behind the SM correlation, have been proposed since its discovery. In our work, we come to the conclusion that any biological interpretations should be used cautiously, especially those based on experimental data sets with an insufficient number of animals. Otherwise, one might misinterpret the SM correlation between the Gompertz parameters as the actual dependence of MRDT on IMR, whereas to obtain the real relation between these parameters one should rely only on sufficiently large data sets. In fact, the demographic parameters may be related indeed in a more advance aging model and therefore a more thorough analysis of this dependence is very much desired, but yet has to be provided.

Acknowledgements

We thank prof. Petrascheck for providing the raw survival data, the lifetables from the experiment [5], Natalia and Leonid Gavrilovs, Ivan Molodtsov and Boris Zhurov from Gero for helpful discussions, and help with visualization software. This work was supported by Gero LLC.

Appendix A. The SM correlation is a degenerate manifold

The solution for the regression problem in Eq. (2) can be found from the minimum conditions, $\partial J/\partial\alpha = 0$ and $\partial J/\partial M = 0$. Using the relations $\partial N_G/\partial\alpha = (t/\alpha)\partial N_G/\partial t$, and $\partial N_G/\partial M = (\alpha M)^{-1}\partial N_G/\partial t$, we obtain:

$$\frac{\partial J}{\partial \alpha} = \int_0^\infty dt \left[N_G(t) - N(t)\right] \frac{t}{\alpha} \frac{\partial N_G}{\partial t}, \tag{A.1}$$

and

$$\frac{\partial J}{\partial M} = \frac{1}{\alpha M} \int\limits_0^\infty dt \left[N_G(t) - N(t) \right] \frac{\partial N_G}{\partial t}. \tag{A.2}$$

If one fruitlessly tries to use mortality rate instead of survival fraction in order to eliminate the degeneracy, the resulting equations are:

$$\frac{\partial J}{\partial \alpha} = \int\limits_0^\infty dt \frac{\partial \ln (N_G/N)}{\partial t} \frac{t}{\alpha} \frac{\partial \ln N_G}{\partial t}, \tag{A.3}$$

$$\frac{\partial J}{\partial M} = \frac{1}{\alpha M} \int\limits_0^\infty dt \frac{\partial \ln (N_G/N)}{\partial t} \frac{\partial \ln N_G}{\partial t}. \tag{A.4}$$

Both $N(t)$ and $N_G(t)$ are step-like functions, $N(t) \approx \theta(\bar{t} - t)$ and $N_G(t) \approx \theta(t_G - t)$, hence there is no essential difference between the fit results of Eqs. A.3, A.4. As a result, for the arbitrary choice of the system variables, with a relative error $\sim \triangle t/\bar{t}$

$$t_G \approx \bar{t},$$

which is a form of correlation between the Gompertz fit parameters:

$$\alpha = \frac{1}{\bar{t}} \ln \left(\frac{\alpha}{M_0} \right).$$

The equation, of course, only holds if $\Lambda = \ln(\alpha/M_0) \gg 1$. It can be generalized approximately to any values of the parameter α/M_0 in a form

$$\alpha = \frac{1}{\bar{t}} \ln \left(1 + C_1 \frac{\alpha}{M_0} \right), \tag{A.5}$$

where $C_1 \sim 1$.

The difference between the slopes of the two solutions at an intersection point in Eq. 5 can be straightforwardly estimated as

$$\xi \sim \frac{1}{\Lambda^2}.$$

Therefore, the two solutions of the optimization problem in Eqs. A.1, A.2 are nearly equal with the identical quadratic precision

$$\frac{\partial J}{\partial \alpha} \propto \frac{\partial J}{\partial M}.$$

Appendix B. The Gompertz iso-average-lifespan curve

We have proven above, that the only parameter which is really robustly optimized by any gradient descent method is the average lifespan, which is for the Gompertz law is written accurately as

$$\bar{t}_G = \int_0^\infty dt N_G = \frac{\exp M}{\alpha} \Gamma\left(0, M\right), \tag{B.1}$$

where $\Gamma(0, M) = E_1(M)$, $M = M_0/\alpha$ is the upper incomplete gamma function (or a special function called the exponential integral) which has an expansion

$$E_1(M) = \Gamma(0, M) = -\gamma - \ln M - \sum_{k=1}^{\infty} \frac{(-M)^k}{k(k!)},$$

where $\gamma \approx 0.577$ is the Euler-Mascheroni constant.

The experimental average lifespan should be equal to the Gompertz average lifespan:

$$\bar{t} = \bar{t}_G = \frac{\exp M}{\alpha} \left(-\gamma - \ln M - \sum_{k=1}^{\infty} \frac{(-M)^k}{k(k!)} \right). \tag{B.2}$$

For $M \gg 1$ (large α) the solution is rather simple

$$\bar{t}_G \left(M \gg 1\right) = \frac{1}{\alpha} \left(\ln \frac{\alpha}{M_0} - \gamma \right).$$

Although the convergence of the series on the right side is known to be slow for arguments of larger modulus ($\alpha \to 0$), but we still can obtain a divergent series approximation

$$\bar{t}_G \left(M \ll 1\right) = \frac{1}{M_0} \left[1 - \frac{\alpha}{M_0} + \sum_{n=2}^{N-1} \left(-\frac{\alpha}{M_0} \right)^n n! \right] \tag{B.3}$$

To compare with the Strehler-Mildvan correlation, for $\ln M_0$ we now get the following asymptotic

$$\ln M_0 \left(\alpha \to 0\right) = \ln \left(1/\bar{t}\right) - \alpha \bar{t}, \tag{B.4}$$

$$\ln M_0 \left(\alpha \gg 0\right) = \ln \alpha - \alpha \bar{t} - \gamma. \tag{B.5}$$

To plot these correlation curves in Fig. 2 we simply join these asymptotic lines together with a spline line at the intersection point $\alpha_{is} = e^\gamma/\bar{t}$, that is why the curves look smooth as the exact solution, these curves look like straight lines at first sight). The actual origin of the Strehler-Mildvan correlation is a visual misinterpretation of an intricate degenerate manifold of the Gompertz fit parameters. Commonly used gradient descent methods are unable to find a unique optimal solution, because any pair of parameters M_0 and α on the iso-average-lifespan curve makes corresponding derivatives in a gradient descent method be zero with a quadratic precision.

Appendix C. The Gompertz-Makeham iso-average-lifespan curve

Gavrilov [6, 7] proposes the Strehler-Mildvan correlation to be a spurious correlation due to the neglect of an age-independent background component of mortality in the original work of Strehler and Mildvan (use of the Gompertz law instead of the Gompertz-Makeham law) [2]. Let us investigate the Gompertz-Makeham law properties in the same manner as above.

$$S_{GM}(t|\alpha, M_0, A) = \exp\left(-At + \frac{M_0}{\alpha}\left(1 - e^{\alpha t}\right)\right).$$

The exact average lifespan for the Gompertz-Makeham law is

$$\bar{t}_G = \int_0^\infty dt N_G = \frac{\exp\frac{M_0}{\alpha}}{\alpha}\left(\frac{M_0}{\alpha}\right)^{A/\alpha}\Gamma\left(-\frac{A}{\alpha}, \frac{M_0}{\alpha}\right),$$

which is indeed somewhat better then we had with the Gompertz law, because the Makeham parameter moves the centre of our expansion for the upper incomplete gamma function from the singular point near zero $\Gamma(0, \frac{M_0}{\alpha})$ to a better analytical point nearby non-zero point of $\Gamma\left(-\frac{A}{\alpha}, \frac{M_0}{\alpha}\right)$, therefore, one can expect the divergent series approximation as in Eq. B.3 to disappear, whereas we now have the three-parametric fitting procedure, for which the SM correlation will nevertheless exist. The degeneracy manifold in case of Gompertz's law arising from the vanishing slope between two optimization curves in case of the Gompertz-Makeham law emerges due to the shrinking slope between two optimization surfaces, therefore the Makeham parameter just determines the cross section of these surfaces, where the same degeneracy problem for the remaining two Gompertz parameters still persists. Even though the Makeham parameter A is well-defined as long as it is large enough, nevertheless for the common regime of human ageing it is not the case since a step-like behavior of survival curves is expected, hence the introduction of the Makeham parameter is of little consequence and the degeneracy is not eliminated.

References

[1] B. Gompertz, On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies, Philosophical transactions of the royal society of London 115 (1825) 513–583.

[2] B. L. Strehler, A. S. Mildvan, General theory of mortality and aging. a stochastic model relates observations on aging, physiologic decline, mortality, and radiation, Science See Saiensu 132 (1960) 14–21.

[3] N. Stroustrup, W. E. Anthony, Z. M. Nash, V. Gowda, A. Gomez, I. F. López-Moyado, J. Apfeld, W. Fontana, The temporal scaling of caenorhabditis elegans ageing, Nature (2016) 16550.

[4] V. Kogan, I. Molodtsov, L. I. Menshikov, R. J. S. Reis, P. Fedichev, Stability analysis of a model gene network links aging, stress resistance, and negligible senescence, Scientific reports 5 (2015) 13589.

[5] X. Ye, J. M. Linton, N. J. Schork, L. B. Buck, M. Petrascheck, A pharmacological network for lifespan extension in caenorhabditis elegans, Aging cell 13 (2) (2014) 206–215.

[6] L. A. Gavrilov, N. S. Gavrilova, The biology of life span: a quantitative approach., New York New York/Chur Switzerland Harwood Academic 1991., 1991.

[7] L. A. Gavrilov, N. S. Gavrilova, The reliability theory of aging and longevity, Journal of theoretical Biology 213 (4) (2001) 527–545.

[8] A. Golubev, Does makeham make sense?, Biogerontology 5 (3) (2004) 159–167.

[9] P. Wagner, et al., Vitality heterogeneity in the strehler-mildvan theory of mortality, Tech. rep., Max Planck Institute for Demographic Research, Rostock, Germany (2011).

[10] M. Finkelstein, Discussing the strehler-mildvan model of mortality, Demographic Research 26 (2012) 191–206.

[11] T. Li, J. J. Anderson, The strehler–mildvan correlation from the perspective of a two-process vitality model, Population studies 69 (1) (2015) 91–104.

[12] J. Shen, G. N. Landis, J. Tower, Multiple metazoan life-span interventions exhibit a sex-specific strehler–mildvan inverse relationship between initial mortality rate and age-dependent mortality rate acceleration, The Journals of Gerontology Series A: Biological Sciences and Medical Sciences (2016) glw005.

[13] K. Hawkes, K. R. Smith, J. K. Blevins, Human actuarial aging increases faster when background death rates are lower: a consequence of differential heterogeneity?, Evolution 66 (1) (2012) 103–114.