Maximum entropy approach to reliability

Yi-Mu Du,[1] Yu-Han Ma,[2] Yuan-Fa Wei,[3] Xuefei Guan,[1,*] and C. P. Sun[1,2,†]
[1]*Graduate School of China Academy of Engineering Physics, Beijing 100193, China*
[2]*Beijing Computational Science Research Center, Beijing 100193, China*
[3]*Institute of Systems Engineering, China Academy of Engineering Physics, Mianyang 621999, China*

(Received 28 September 2019; published 3 January 2020)

The aging process is a common phenomenon in engineering, biological, and physical systems. The hazard rate function, which characterizes the aging process, is a fundamental quantity in the disciplines of reliability, failure, and risk analysis. However, it is difficult to determine the entire hazard function accurately with limited observation data when the degradation mechanism is not fully understood. Inspired by the seminal work pioneered by Jaynes [Phys. Rev. **106**, 620 (1956)], this study develops an approach based on the principle of maximum entropy. In particular, the time-dependent hazard rate function can be established using limited observation data in a rational manner. It is shown that the developed approach is capable of constructing and interpreting many typical hazard rate curves observed in practice, such as the bathtub curve, the upside down bathtub, and so on. The developed approach is applied to model a classical single function system and a numerical example is used to demonstrate the method. In addition its extension to a more general multifunction system is presented. Depending on the interaction between different functions of the system, two cases, namely reducible and irreducible, are discussed in detail. A multifunction electrical system is used for demonstration.

DOI: 10.1103/PhysRevE.101.012106

I. INTRODUCTION

The aging process is a common phenomenon in many engineering, biological, and physical systems. Modeling the aging process is of critical importance for the evaluation of reliability, availability, and safety of a system, and thus is widely concerning. However, it is difficult to describe the aging process microscopically due to its inherent complex and stochastic nature. Consequently the lifetime of a component or system, as the final outcome of the aging process, is also random. The hazard rate function characterizing the time-dependent nature of the aging process plays an important role in reliability engineering [1]. Knowing the hazard rate function allows for prediction of the lifetime distribution and therefore the failure probability can be evaluated. Early and systematical attempts in construction of the hazard function can be traced back to the reliability study of military radar systems [2]. The statistical analysis of component life testing data has since become an engineering practice to estimate the hazard rate function. The bathtub shape of the hazard rate function is then widely observed and the mechanism of such a shape is empirically explained.

To describe the above mentioned concept in detail, consider a single function system with an uncertain lifetime T as a random variable. The survival function, $F(t)$, is defined as the probability that the lifetime T is greater than t, i.e., $F(t) = \Pr(T \geqslant t)$, where $t \in (0, +\infty)$. The corresponding probability density function (PDF) for lifetime $T = t$ is obtained as $p(t) \equiv -dF(t)/dt$ with a normalizing constant $\int_0^\infty p(t)dt = F(0) - F(\infty) = 1$. The hazard rate function $x(t)$ is then defined as

$$x(t) \equiv -\frac{1}{F(t)}\frac{dF(t)}{dt} = \frac{p(t)}{F(t)}. \tag{1}$$

The hazard rate is the failure rate of the system at time t conditioning on the system at t is still functioning. Empirical evidences show that for a number of systems, the hazard rate function exhibits the so-called bathtub shape or U shape (the solid curve shown in Fig. 2). The earliest bathtub shaped hazard rate function appears in an actuarial life-table analysis in 1693 [3,4]. The bathtub shape of the hazard rate function implies that the failure mechanism of the system may be divided into three phases: the infant mortality, random failures, and wear-out failures.

Although the bathtub shape of hazard rate functions is widely observed in many realistic components and systems, the lack of underlying physical theory supporting the interpretation leads to some criticisms [4]. Moreover, the hazard rate functions that are not in bathtub shape are also observed in some biological systems [5] and electronic systems [6]. Beside the areas of engineering science and biology, the aging processes (especially in materials) are also widely concerned in the area of material physics [7–14], including, but not be limited to, the rupture of the fibrous materials [9,12], the cracking of heterogeneous material [10,13], and physical aging of colloidal glass [14].

The perspectives on the aging problems may be different to engineers and physicists. The former usually focuses on the degradation of the functioning systems and lifetime distributions. The latter concerns the physical and statistical laws in aging processes. The physical aging usually means the dynamical relaxation of the systems [15–19]. The connections

*xfguan@gscaep.ac.cn
†cpsun@csrc.ac.cn

between the aging processes in different areas are also signified in, for example, Ref. [20] where a direct link between the biological aging and the physical aging is established.

The hazard rate function is usually difficult to obtain due to the lack of precise understanding of the physical mechanism and the limitation of information access. Statistical analysis is one of the most common approaches to estimate the hazard rate curves using lifetime testing data. A sufficient amount of lifetime data are required to ensure a reliable estimation of the hazard rate curve. A few methods have been reported to fit the hazard rate curves based on Weibull distributions, Lindley distributions, exponential distributions, and its variants [21–23]. The modeling of the commonly seen bathtub shaped failure rate curve is reviewed in Ref. [24] and some recent development regarding this topic is discussed in Refs. [25,26]. The performance of these methods relies on the sufficiency of the experimental data and the choice of distributions. Another approach to construct the hazard rate function is based on modeling the physical mechanism of the aging process [7–9]. This approach can provide a clear physical picture of the aging process and a close agreement with experimental data is observed for relatively simple systems. For systems with complex aging mechanisms this approach is difficult to apply. Recently, entropy-based methods are also reported to tackle the problems in the field of reliability [27–36]. In Refs. [27–29], the authors considered remaining lifetime and corresponding time-dependent residual entropy, and investigated the lifetime distributions to maximize the residual entropy in several conditions.

Although many studies have been reported on estimation of the hazard rate curve, an approach built upon fundamental principles of physics to construct the hazard rate function is rarely seen. In addition, to the best of the authors' knowledge, there is no analytical modeling work reported on multifunction systems consisting of several individual components. In view of this, the goal of this study is to develop an approach to reliability problems based on the fundamental principle of maximum entropy (MaxEnt), allowing for the construction of the hazard rate function in a rational manner given available data. This study extends the MaxEnt method to the time-dependent cases. The principle of MaxEnt and the least-action principle is compared to formulate the equation of motion for the most probable hazard rate, i.e., the Euler-Lagrange equation. Another focus of this study is to reveal the underlying connection between different shapes of the hazard rate and the information processing. Both the model-dependent and the data-dependent hazard rate shapes are investigated in this study. In addition, the modeling of multifunction systems consisting of multiple individual components is made using the developed approach. A linear assumption is incorporated to cope with the complexity introduced by multiple functions in the system. Using the linear assumption, the hazard rate of the system can be represented as a multidimensional matrix. Depending on the interaction between each individual function, two cases, namely the reducible one and the irreducible one, are derived using the proposed method. It demonstrates a viable means for the bottom-up modeling of hazard rate functions of complex systems.

This paper is organized as follows. In Sec. II, the principle of MaxEnt is briefly reviewed. The variational technique is applied to MaxEnt to formulate the dynamical equation of the hazard rate. The equation of motion is solved to obtain the most probable hazard rate function. In Sec. III, the modeling of the hazard rate function for single function systems is presented. Using a double-moment constraint, the modeling problem is recast to a MaxEnt inference problem. It is shown that the resulting model is capable of producing three different shapes of the hazard rate curves: the bathtub shape, the upside down bathtub shape, and the monotonically increasing shape. The data dependence of the shapes is demonstrated by using different model parameters. Furthermore, using a quadruple-moment constraint allows the resulting model to produce other shapes of the hazard rate function. Numerical experiments of a single-function electrical system is used to validate the effectiveness of the method. In Sec. IV, the modeling of multifunction systems where the correlation is among the components is presented. A two-lamp circuit is used to demonstrate the modeling of multifunction systems. Finally, conclusions are drawn in Sec. V.

II. MAXENT METHOD

A. Principle of maximum entropy (MaxEnt)

The concept of entropy was originally introduced by Clausius to identify the reversible and irreversible processes in thermodynamics. According to the view of Clausius, the second law of thermodynamics states that the entropy of a thermal-isolation system never decreases. Boltzmann, then, associated the entropy to statistical quantities, laying out the foundation of the modern statistical physics. Shannon [37] developed the information theory which is based on a quantity called Shannon entropy, $S = -\sum_i p_i \log p_i$, as a measure of information loss or the ignorance of an observer. Jaynes [38,39] combined the insights of both statistical physics and information theory, and established the fundamental logic of probabilistic inference. In his point of view, statistical physics is only to do probabilistic inferences from limited information based on the principle of MaxEnt. Jaynes [40] later compared the MaxEnt method with other methods of inference, and pointed out that it is applicable to inference problems with a well-defined hypothesis space and noiseless incomplete data.

The principle of MaxEnt states that the most probable distribution function is the one that maximizes the entropy given testable information. A PDF $p(t)$ and its hazard rate function $x(t)$ relate to each other through

$$p(t) = x(t) \exp\left[-\int_0^t x(t')dt'\right]. \tag{2}$$

Note that Eq. (2) is just the solution to Eq. (1). Using Eq. (2), the entropy S for a given PDF $p(t)$ can be expressed as

$$\begin{aligned} S &= -\int_0^\infty p(t) \ln p(t) dt \\ &= -\int_0^\infty \dot{X} \exp(-X) \ln[\dot{X} \exp(-X)] dt, \end{aligned} \tag{3}$$

where $X = \int_0^t x(t')dt'$ and $\dot{X} = x(t)$. The entropy in Eq. (3) is called the differential entropy. A constant term $\ln(dt)$ has been neglected, and thus the term $\ln p(t)$ has an abnormal dimension. In Eq. (3) the measure in integral is set to dt. In

general, the measure should be properly chosen according to the prior knowledge.

Experimental data can be encoded into pieces of testable information using statistical moments of any observable time-dependent quantities $g_i(t)$, $i = 1, 2, \ldots$, i.e., $\bar{g}_i = \int_0^\infty p(t) g_i(t) dt$, $i = 1, 2, \ldots$. To maximize the entropy based on the testable information, the variational term of the entropy can be used,

$$\begin{aligned}\delta\bar{S} &= \delta\bigg(-\int_0^\infty p(t)\ln p(t)dt - \alpha\int_0^\infty p(t)dt \\ &\quad - \sum_i \beta_i \int_0^\infty g_i(t)p(t)dt\bigg) \\ &= -\delta\bigg(\int_0^\infty \dot{X}\,\exp(-X)\bigg\{\ln[\dot{X}\,\exp(-X)] \\ &\quad + \alpha + \sum_i \beta_i g_i(t)\bigg\}dt\bigg) = 0, \end{aligned} \tag{4}$$

where α and β_i, $i = 1, 2, \ldots$ are Lagrange multipliers.

B. Entropy as the action of hazard rate dynamics

The entropy in the above equation is the function of the time dependent variable $X = \int_0^t x(t')dt'$. To determine the most probable hazard rate one may adopt the variational principle. In this sense the concept of MaxEnt is similar to the least-action principle.

One can see that the quantity $\tilde{S} = -\bar{S}$ is similar to the action which governs the hazard rate dynamics. The "velocity" $\dot{X}$ which minimizes the action is the most probable hazard rate. We rewrite Eq. (4) in Lagrangian form explicitly to have

$$\mathcal{L}[X, \dot{X}, t] = \dot{X}\,\exp(-X)\bigg\{\ln[\dot{X}\,\exp(-X)] + \alpha + \sum_i \beta_i g_i(t)\bigg\}. \tag{5}$$

To maximize the action, X satisfies the Euler-Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{X}}\right) - \frac{\partial\mathcal{L}}{\partial X} = 0, \tag{6}$$

which is expressed explicitly as

$$\ddot{X} - \dot{X}^2 + \dot{X}\sum_i \beta_i \dot{g}_i(t) = 0. \tag{7}$$

The Lagrange multiplier α is related to the normalization and does not appear in the above equation because the function $\int \dot{X}\,\exp(-X)dt$ is nothing but a constant. This equation of motion governs the most probable hazard rate varying with time. It should be noted that the constraints used for method development are based on the lifetime statistics. The constraints can be loosely seen as pieces of information encoding some features over the entire lifespan. Therefore, the constraints given above are not varying with time; therefore, the resulting Lagrange multipliers are not time-varying quantities. However when the constraints are time varying, the resulting Lagrange multipliers will be time dependent. For example, the following constraint is time dependent and can result in a time-dependent Lagrange multiplier,

$$\int_{t_i}^{t_{i+1}} p(t)dt = c_i,$$

where $i = 1, 2, \ldots, t_i < t_{i+1}$ for any i, and c_i can be evaluated from the empirical probability. It is worth mentioning that the method itself can deal with both the time-dependent and time-independent constraints, as the underlying principle of the MaxEnt only provides a mechanism to process the given constraints as long as they are testable information.

Equation (5) should not be overlooked. Although on the surface it appears as an algebraic manipulation, the significance of the established dynamics should be emphasized here. The hazard rate is originally a quantity in the statistical sense, and the dynamical equations are usually governed by the action in physical theories. Equation (5) relates the hazard rate function to the dynamics of time-dependent aging processes. In this sense, the hazard rate becomes a bridge that links the statistical and the dynamical aspects of aging processes. In other words, the entropy here is not only the measure of uncertainty but also becomes the action that governs the hazard rate dynamics. This remark is established based solely on one fundamental physical principle, MaxEnt. Consequently, the MaxEnt method provides a natural and rational way to treat the problem of reliability more physically.

C. Most probable hazard rate and distribution

Take the moments of lifetime as the constraints, i.e., $g_i = t^i$, and use the initial condition of $g_i(0) = 0$ $(i = 1, 2, \ldots)$. Solve Eq. (7) with certain initial conditions to obtain

$$\dot{X} = \frac{x_0 \exp\big(-\sum_i \beta_i g_i\big)}{1 - x_0\int_0^t \exp\big(-\sum_i \beta_i g_i\big)dt'}, \tag{8}$$

where $\dot{X}(0) = x_0$ is the initial hazard rate. The function in Eq. (8) has a singularity point denoted by $t_{\max}$ with the following condition:

$$x_0\int_0^{t_{\max}} \exp\left(-\sum_i \beta_i g_i\right)dt' = 1. \tag{9}$$

Note that the definition of hazard rates Eq. (1) shows that the hazard rates are always non-negative. The term $\dot{X}$ in Eq. (8) is negative when $t > t_{\max}$. It implies that only $t \leqslant t_{\max}$ is allowable. This can be directly proved by verifying the normalization condition. Substituting Eq. (8) into Eq. (2), for $t \leqslant t_{\max}$, the PDF $p(t)$ becomes

$$p(t) = x_0 \exp\left(-\sum_i \beta_i g_i\right). \tag{10}$$

Equation (9) is nothing but a normalization condition, and the singularity $t_{\max}$ is the maximum lifetime. The detailed derivations and explanations of the maximum lifetime are presented in Appendix A.

The above discussion shows that the physically allowable hazard rate is

$$x \equiv \dot{X} = \begin{cases} \frac{x_0 \exp(-\sum_i \beta_i g_i)}{1 - x_0\int_0^t \exp(-\sum_i \beta_i g_i)dt'}, & 0 \leqslant t < t_{\max}, \\ 0, & t \geqslant t_{\max}, \end{cases} \tag{11}$$

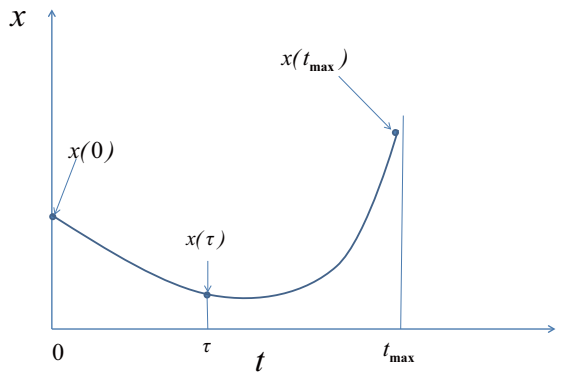

FIG. 1. Illustration of alternative additional constraints. $x(0), x(t_{\max})$ are the alternative additional constraints at the boundary. One can, in general, use the hazard rate at an arbitrary time $x(\tau)$ to obtain the solution.

where $\dot{X}(0) = x_0$ is the initial hazard rate and $t_{\max}$ is the maximum lifetime. It follows from Eq. (2) that

$$p(t) = \begin{cases} x_0 \exp\left[-\sum_i \beta_i g_i(t)\right], & 0 \leqslant t < t_{\max}, \\ 0, & t > t_{\max}. \end{cases} \tag{12}$$

The normalization of $p(t)$ indicates that the maximum lifetime $t_{\max}$ is related to the initial hazard rate x_0 through Eq. (9) and vice versa.

The hazard rate increases with time when t approaches $t_{\max}$ given that $t_{\max}$ is finite. This feature plays an important role in the bathtub shaped hazard rate function in the following two aspects. First consider the initial hazard rate x_0 and the moment $\bar{g}_i$ are observed information, and the maximum lifetime $t_{\max}$ and other features of the distribution are quantities to be inferred. The following example illustrates this setting: the maximum lifetime of humans is the variable of interest to be inferred based on the information of infant mortality and the average lifetime. Conversely given that the maximum lifetime $t_{\max}$ and the moment $\bar{g}_i$ are observed information, and the initial hazard rate $x(0)$ and other features of the distribution are quantities to be inferred, a different inference problem is formulated. For example, given the maximum lifetime $t_{\max}$ or the average lifetime and its limit, an estimation on the initial hazard rate can be made. In general these local information, such as initial hazard rate x_0 and the maximum lifetime $t_{\max}$, serves as the boundary condition of hazard rate dynamics, and provides the necessary constraint in solving the hazard rate function using MaxEnt. One can, in general, use the hazard rate at an arbitrary time to do the inference. However, the results of the inference usually depend on the choice of such additional constraints, and are discussed in the next section.

The different choices of additional constraints are illustrated in Fig. 1. It is worth mentioning that the most probable distribution of the lifetime can be obtained from the hazard rate $\dot{X}$ by utilizing Eq. (2) or the equivalent Eq. (12). It can also be determined directly using MaxEnt with necessary constraints such as the maximum lifetime or an initial hazard rate. It can be realized from Fig. 1 that the resulting distribution functions using the two considerations are equivalent if the observation error is omitted. The proof is presented in Appendix A. This property is useful for the self-consistency test of the method, which is discussed later.

III. SINGLE FUNCTION SYSTEMS

Following Eq. (1), the survival function $F(t)$ of a single function system decreases with time as

$$\frac{dF}{dt} = -x(t)F(t), \tag{13}$$

where $F(t) = \int_0^t p(t')dt'$ is a monotonically decreasing function and $x(t) \geqslant 0$ is the hazard rate function.

A. Double-moment constraints

Recall that $g_i(t), i = 1, 2, \ldots$ denotes the random variables which can be observed from experiments. Usually only a small number of samples are available in practical problems. In this case only the first order and second order moments are reliable. Denote the first order and second order moments as $g_1(t) = t$ and $g_2(t) = t^2$, respectively. Omitting the term t in $x(t)$ and representing $x(0)$ as x_0 only for simplicity, the equation of motion of the hazard rate becomes

$$\dot{x} - x^2 + x(\beta_1 + 2\beta_2 t) = 0, \tag{14}$$

where β_1 and β_2 are the Lagrange multipliers for the first and second order moments, respectively. Given the initial condition $x(0) = x_0 > 0$, the solution to Eq. (14) is

$$x = \begin{cases} \dfrac{x_0 \exp(-\beta_1 t - \beta_2 t^2)}{1 - x_0 \int_0^t \exp(-\beta_1 t' - \beta_2 t'^2)dt'}, & 0 \leqslant t < t_{\max}, \\ 0, & t \geqslant t_{\max}. \end{cases} \tag{15}$$

The normalization requires

$$x_0 \int_0^{t_{\max}} \exp(-\beta_1 t - \beta_2 t^2)dt = 1. \tag{16}$$

Because $t_{\max} \in [0, \infty)$ and $\exp(-\beta_1 t - \beta_2 t^2) > 0$, one has

$$x_0 \int_0^{\infty} \exp(-\beta_1 t - \beta_2 t^2)dt \geqslant 1. \tag{17}$$

This is the necessary and sufficient condition to guarantee that the solution is physically viable. If the above condition is violated, the resulting PDF cannot be normalized and thus is improper.

Conditions violating the inequality in Eq. (17) only occur when $\beta_2 > 0$. The inequality always holds because the integration $\int_0^{\infty} \exp(-\beta_1 t - \beta_2 t^2)dt$ is divergent when $\beta_2 < 0$. In particular, for $\beta_2 > 0$, the equality implies the resulting PDF is a truncated Gaussian distribution.

Consider the triplet of (x_0, β_1, β_2) which satisfies Eq. (17); the most probable hazard rate function $x(t)$ can yield three different types of shapes. Table I presents the three possible shapes and corresponding domains of parameters. The detailed derivations are presented in Appendix B. In the case of $\beta_2 < 0, \beta_1 > x_0$, the hazard rate functions are in bathtub shapes. In the case of $\beta_2 > 0, x_0 \int_0^{\infty} \exp(-\beta_1 t - \beta_2 t^2)dt = 1$, the hazard rate functions are in upside down bathtub shapes. The hazard rate functions are associated with truncated Gaussian distributions. In other physically viable cases, i.e.,

TABLE I. Shapes of hazard-rate curves and the corresponding domains of parameters.

	$x_0 \geqslant \beta_1$	$x_0 < \beta_1$
$\beta_2 < 0$	monotonically increasing	bathtub
$\beta_2 > 0$	$x_0 \int_0^\infty \exp(-\beta_1 t - \beta_2 t^2)dt = 1$ upside down bathtub	$x_0 \int_0^\infty \exp(-\beta_1 t - \beta_2 t^2)dt > 1$ monotonically increasing

$\beta_2 > 0, x_0 \int_0^\infty \exp(-\beta_1 t - \beta_2 t^2)dt > 1$, the hazard rate functions are monotonically increasing curves.

Figure 2 presents results of Eq. (14) with an initial condition $x_0 = 1$ for illustration. For example, the hazard rate function associated with parameters $\beta_1 = -0.1$ and $\beta_2 = 1.5989$ is a monotonically increasing curve, satisfying the condition of $\beta_2 > 0$ and $x_0 \int_0^\infty \exp(-\beta_1 t - \beta_2 t^2)dt > 1$. The parameters of the other two curves (the bathtub shape and the upside down bathtub shape) also satisfy the conditions yielding the two shapes. It should be noted that the quantities which can directly be obtained from experimental data are not the parameters of β_1 and β_2 but the moments of $\bar{t}$ and $\bar{t^2}$.

Note that the number of the possible types of shapes depends on the constraints. The double-moment (the first and second order moments) constraint can yield hazard rate functions in bathtub shapes, upside down bathtub shapes, and the monotonically increasing curves. Figure 3 shows how the shape of the curves varies with the first order and second order moments. The numerical results show that if the point $(\bar{t^2}, \bar{t})$ is located in the black region in Fig. 3, the resulting hazard rate function function exhibits a bathtub shape; otherwise, the hazard rate function is a monotonically increasing curve (in the gray region). In this case the maximum lifetime is 1, i.e., a finite quantity, and there is no hazard rate function in upside down bathtub shapes.

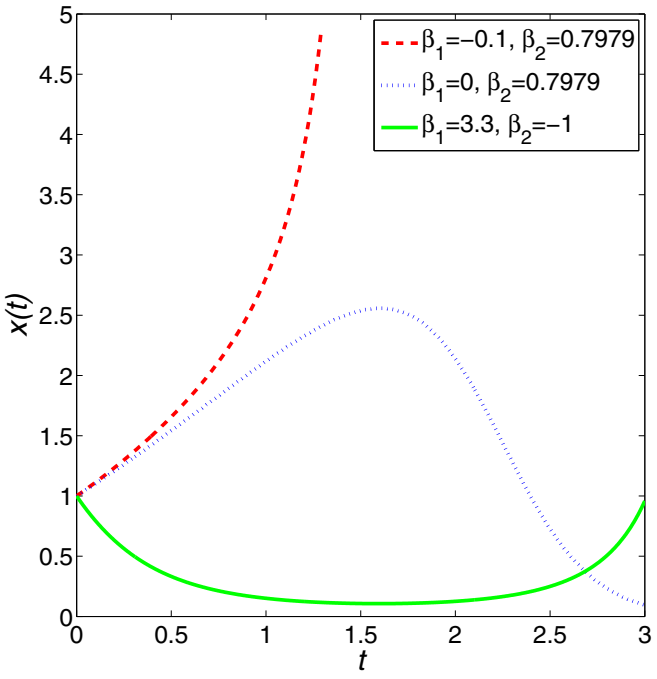

FIG. 2. Hazard rate function $x(t)$ using a double-moment constraint with different (β_1, β_2) parameters, showing three different shapes. The initial hazard rate x_0 is set to 1. The parameters associated with the three shapes are only for illustration purposes.

B. Quadruple-moment constraints

The analysis of the hazard rate function with a quadruple-moment constraint is similar to the previous one with a double-moment constraint. Equation (7) implies that for a quadruple-moment constraint the equation of motion of the hazard rate function is written as

$$\dot{x} - x^2 + x(\beta_1 + 2\beta_2 t + 3\beta_3 t^2 + 4\beta_4 t^3) = 0, \qquad (18)$$

where β_i, $i = 1, \ldots, 4$ are parameters that need to be solved given the moment constraints. The resulting hazard rate function $x(t)$ of the above equation has seven types of shapes at most. The proof is presented in Appendix B. Besides the three types of shapes shown in the double-moment case, the quadruple-moment constraint is capable of generating W- and upside down W-, N- and upside down N-shaped curves. Figure 4 shows the additional possible four types of shapes of the hazard rate function allowed by Eq. (18). The parameters used in Fig. 4 are only for illustration purposes.

C. Illustration of MaxEnt: A circuit model

To demonstrate the performance of the MaxEnt method, the aging problem of an electrical system is presented in this section. As shown in Fig. 5 the electrical system contains the

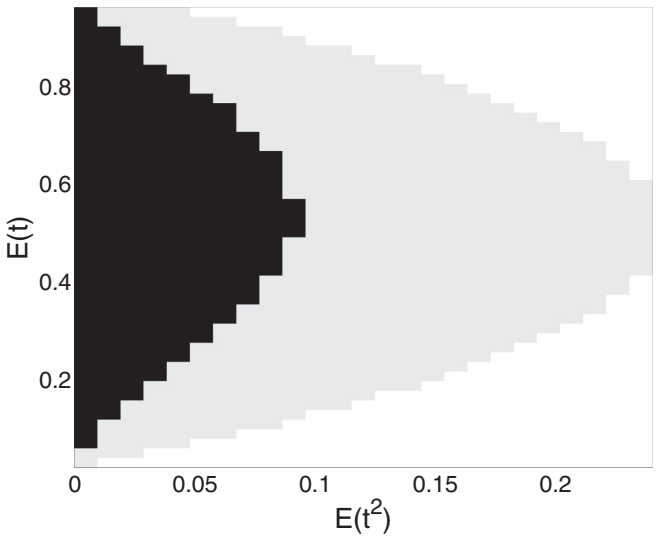

FIG. 3. Two regions of the hazard rate functions with different first order moments $E(t) = \bar{t}$ and second order moments $E(t^2) = \bar{t^2}$. The maximum lifetime $t_{\max}$ is set to 1. The black region denotes the domain generating monotonically increasing curves. The gray region denotes the domain generate bathtub shapes.

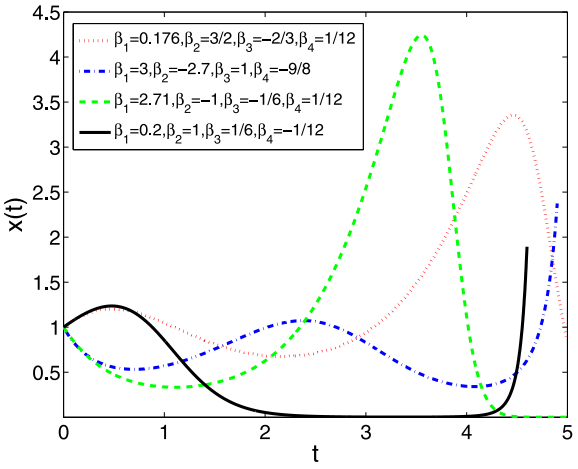

FIG. 4. Four additional shapes of the hazard rate function $x(t)$ with a quadruple-moment constraint. The initial hazard rate x_0 is set to 1. The parameters associated with the four shapes are only for illustration purposes. These parameters satisfy the condition of $x_0 \int_0^\infty \exp(-\beta_1 t - \beta_2 t^2 - \beta_3 t^3 - \beta_4 t^4) dt \geqslant 1$.

battery source, the lamp and the conducting wire. The term E denotes the electromotive force of the source. The internal resistances of the lamp and the conducting wire are denoted by R and r, respectively. The resistance of the conducting wire increases due to the heat generation. For demonstration purposes, the regime where the temperature T is much higher than the Debye temperature is considered, and the resistance $r \propto T$ [41]. The aging process due to the heat generation of the wire can be modeled with the thermodynamical consideration by

$$C\frac{dT}{dt} = \lambda r\left(\frac{E}{R+r}\right)^2, \tag{19}$$

where C is the heat capacity of the material of the conducting wire. The right hand side of the above equation denotes the remaining heat generated by the wire, and λ is a factor which depends on the heat conduction between the wire and the environment. With the high temperature approximation $r \propto T$, the above equation becomes

$$\frac{dr}{dt} = \tilde{\lambda} r\left(\frac{E}{R+r}\right)^2, \tag{20}$$

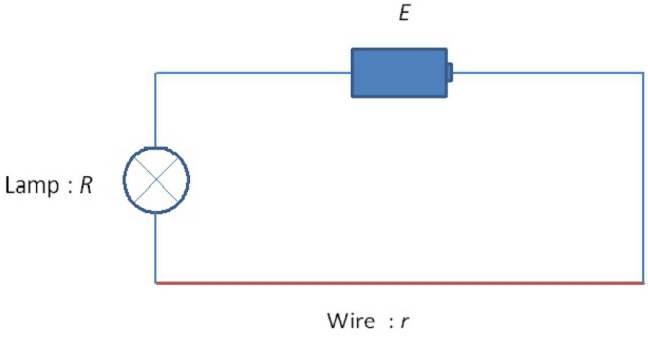

FIG. 5. Schematic of an electrical system consisting of the source, the lamp, and the conducting wire.

where $\tilde{\lambda}$ is the effective ratio parameter. This equation describes the behavior of the time-dependent resistance $r(t)$. One can see that the efficiency $\eta = R/(r+R)$ of the circuit decreases in the aging process. The system failure criterion is defined as $\eta \leqslant 50\%$. The aging process is stochastic when $r(0)$ and $\tilde{\lambda}$ are considered random variables. Numerical experiments are made to generate data representing the actual observed quantities. The initial resistance $r(0)$ follows a PDF of $p[r(0)] \propto \exp[-\Delta r(0)]$ in the interval [0.033,1]. The ratio term $\tilde{\lambda}$ is considered as a uniform distribution in $[1, \Lambda)$. The first order and second order moments are used as constraints. To investigate the performance of the method under different distributions of $r(0)$ and $\tilde{\lambda}$, different combinations of (Λ, Δ) are used. The Kolmogorov distance [42] between the estimated distribution p_e and the simulated distributions p_s is used as a measure to evaluate the performance of MaxEnt. The distance is defined as

$$\epsilon = \int dt \frac{1}{2} |p_e(t) - p_s(t)|. \tag{21}$$

Results of the performance under different $r(0)$ and $\tilde{\lambda}$ distributions are shown in Fig. 6(a). The hazard rate function results associated with two arbitrarily chosen combinations of (Λ, Δ) are presented in Fig. 6(b).

Results in Fig. 6 also show that there are regions where the performance in terms of Kolmogorov distances is not available. In these regions the double-moment constraints are conflicted with the initial value of the hazard rate. In addition, the estimated hazard rate function is compared with the analytical one used to generate the simulated observation data; however, in practice the actual (analytical) hazard rate function is rarely known *a priori*. These two issues are due to the fact that subjective choices of the evidences lead to subjective results of the inference, and this fact is also related to the correctness and consistency of the inference.

D. Self-consistency test of the MaxEnt method

To avoid the conflict between moments and the initial value of the hazard rate, the hazard rate value at a proper time $x(\tau), \tau > 0$, instead of $x(0)$, can be used. To illustrate the influence of the initial value to the estimation result, consider the region labeled as C in Fig. 6(a), which corresponds to parameters $\Delta = 0.3$ and $\Lambda = 2.4$. The hazard rate functions obtained using four arbitrarily chosen times $x(\tau), \tau > 0$ as constraints are presented in Fig. 7. It shows that different local information can lead to slightly different results.

The second issue relates to the correctness and consistency of the inference. For any nontrivial hazard rate estimation problems the true hazard rate is unknown; otherwise, it will not be a problem at the first place. Therefore, measures are needed to ensure the results obtained using the developed method are reliable. One measure for the evaluation is to compare results obtained using different choices of the additional constraints. This measure can be made using Kolmogorov distance between two sets of results. The comparison can loosely be seen as a self-consistency test. The average distance $\bar{d}$ between a series of resulting PDFs with additional constraints

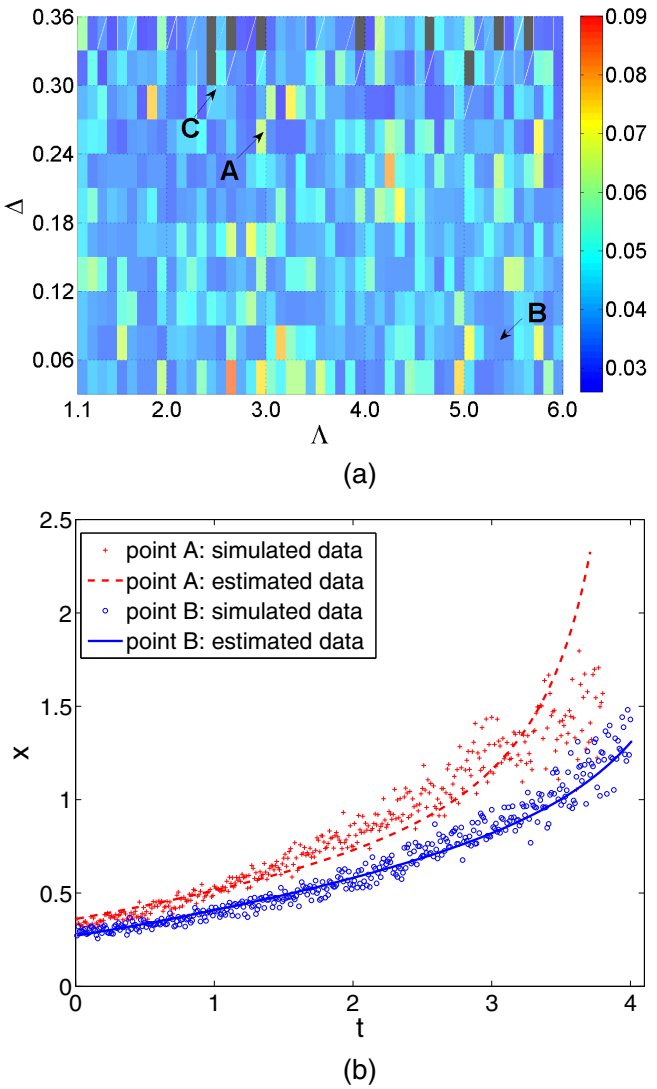

FIG. 6. (a) Results of Kolmogorov distance ϵ between the estimated distribution and the simulated distributions with different values of Δ and Λ and (b) results of the hazard rate function $x(t)$ associated with point "A" ($\Delta = 0.27, \Lambda = 3$) and point "B" ($\Delta = 0.09, \Lambda = 5.4$) in (a) for illustration.

of $x(\tau)$ is defined as

$$\bar{d} = \frac{1}{2 \times n!} \sum_{i,j=1}^{n} \int dt |p_{\mathrm{e}}^{(i)} - p_{\mathrm{e}}^{(j)}|. \tag{22}$$

Similarly the average distances $\bar{\epsilon}$ between the resulting PDFs and the analytical PDFs is

$$\bar{\epsilon} \equiv \frac{1}{n} \sum_{i} \epsilon_i = \frac{1}{2 \times n} \sum_{i=1}^{n} \int dt \left| p_{\mathrm{e}}^{(i)} - p_{\mathrm{s}} \right|. \tag{23}$$

Four different proper additional constraints for $x(\tau)$, $\tau = (0.1, 0.79, 1.02, 1.97)$, are arbitrarily chosen to obtain $x(t)$ for three points (A, B, and C) in Fig. 6(a). For each of the points, four resulting PDFs $p_{\mathrm{e}}^{(i)}, i = 1, 2, 3, 4$ are obtained. Results of the four resulting PDFs are used to evaluate $\bar{d}$ and $\bar{\epsilon}$ using Eq. (22) and Eq. (23), respectively. Results are

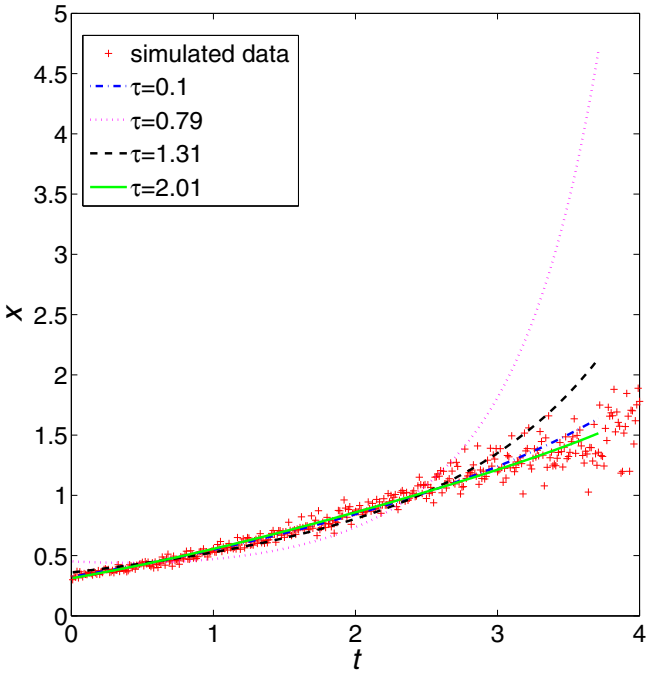

FIG. 7. Inference results of the hazard rate functions obtained using different $x(\tau), \tau > 0$ as constraints. The same moment constraints are used.

presented in Table II. These numerical results indicate that $\bar{\epsilon}$ is positively correlated with $\bar{d}$, and thus the self-consistency test is effective. In this section, the MaxEnt-based method for modeling the hazard rate functions of single-function systems is presented. It also serves as a basis for modeling the hazard rate functions of multifunction systems which are discussed next.

IV. MULTIFUNCTION SYSTEMS

A multifunction system is defined as a system providing multiple functions, and each of the functions has its own failure mechanism. The failure mechanisms of the functions can be isolated or interacted. Consequently, the lifetimes associated with these functions can be independent or correlated. Consider a system having n functions. The survival probability of the $i(i = 1, 2, \ldots, n)$th function is denoted by $F_i(t)$. Given that the hazard rate functions of the system are independent of the survival probabilities, the dynamic equations of the system reduce to a set of linear differential equations. Denote the hazard rate functions of the system consisted of n functions as a $n \times n$ matrix χ. Extending the definition of the hazard rate of the single function system in

TABLE II. Results of $\bar{d}$ and $\bar{\epsilon}$ for points A, B, and C in Fig. 6 using four constraints of $x(\tau), \tau = 0.1, 0.79, 1.02, 1.97$.

Point	$\bar{d}$	$\bar{\epsilon}$
A	0.0150	0.0533
B	0.0408	0.1089
C	0.0164	0.0715

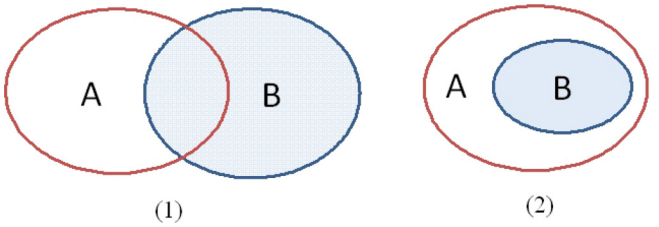

FIG. 8. Schematic of the reducible and irreducible cases for a double-function system. The ellipse with the red borderline denotes the event that the function a works, and the blue ellipse denotes the event that the function b works. (1) The reducible case where the event that the function a works and the event that the function b works are not strongly correlated. In this case, the intersection between the sets A and B does not affect the reducibility of the matrix χ. (2) The irreducible case where the function b works only if the function a also works, i.e., $B \subseteq A$.

Eq. (1) to a multidimensional case, the dynamic equations of the system can be written using a matrix form as

$$\frac{dF}{dt} = -\chi F, \tag{24}$$

where $F = (F_1, F_2, \ldots, F_n)^{\mathrm{T}}$ and χ is a $n \times n$ hazard rate matrix.

Note that the above equation is obtained based on the condition that the hazard rate x or χ only depends on time t and is independent of F. For single function systems, the $x(t)$ and $F(t)$ are independent variables, because there is no physical interaction or statistical correlation in sampling during the aging processes. For multifunction systems, it should be regarded as an approximation for the hazard rate. A sufficient condition for the linear assumption to hold is presented in Appendix D.

It is worth mentioning that the difference between a single-variable hazard rate x and a multidimensional hazard rate matrix χ is not only the dimension. The term x is only related to the probability distribution (if one knows the PDF, then x can directly be calculated and vice versa); however, the term χ also encodes the information of the interaction among different functions in the aging process. Depending on the interaction, two cases are discussed below.

A. Reducibility of hazard-rate matrices

The nonincreasing nature of the survival probability $F_i(t)$ implies that

$$\frac{dF_i}{dt} \leqslant 0, \tag{25}$$

and dF_i/dt and F_i approach zero simultaneously. These two basic properties will result in a constraint on the hazard rate matrix χ.

To illustrate this consider the following simplest case: a double-function system shown in Fig. 8. To avoid confusion, the two functions are labeled by a and b. The ellipses A and B represent the domains that functions a and b work, respectively.

These are two typical cases of the double-function system. The first one shown in Fig. 8(1) denotes the case that a and b are not fully correlated. In this case the system has four states listed in Table III, labeled by 1,2,3,4. Similarly, the second case presented in Fig. 8(2) denotes that if the function b works, then function a must work, i.e., $B \subseteq A$. In this case the system has three states listed in Table IV, labeled by 1,2,3. A typical example for this case is a series-parallel electric circuit presented in Fig. 10, where the two functions of the system are denoted by the two lamps. One can see that if lamp 2 works lamp 1 must work and thus this system contains only three states. It can be shown that for the first case the hazard matrix χ must be diagonal or said reducible. For the second case χ is an upper triangular matrix or said irreducible.

TABLE III. Four states of a double-function system in a reducible case.

States	Function a	Function b
1	Work	Work
2	Work	Break down
3	Break down	Work
4	Break down	Break down

Consider the first case; without loss of generality the initial state of the system is assumed to be 2 in Table III. This initial state corresponds to the following condition:

$$\begin{pmatrix} F_a(0) \\ F_b(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{26}$$

Because $F_b(t)$ is nonincreasing, $F_b(t) = 0$. Since dF_b/dt and F_b approach zero simultaneously, one has

$$\frac{dF_b}{dt} = \chi_{ba} F_a(t) = 0$$

and $\chi_{ba}(t) = 0$. Similarly, assuming the initial state of the system is state 3 in Table III, the initial condition is

$$\begin{pmatrix} F_a(0) \\ F_b(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{27}$$

Because $F_a(t)$ is nonincreasing, $F_a(t) = 0$. As dF_a/dt and F_a approach zero simultaneously, one gets

$$\frac{dF_a}{dt} = \chi_{ab} F_b(t) = 0$$

and $\chi_{ab}(t) = 0$. Therefore, in this case χ is diagonal.

The second case shown in Fig. 8(2) indicates $F_a(t) \geqslant F_b(t)$. Assuming the system is initially at state 2 in Table III, the initial condition is

$$\begin{pmatrix} F_a(0) \\ F_b(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{28}$$

TABLE IV. Three states of a double-function system in an irreducible case.

States	Function a	Function b
1	Work	Work
2	Work	Break down
3	Break down	Break down

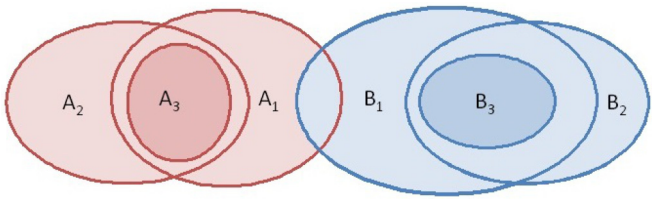

FIG. 9. Schematic of reducible cases for an n-function system. The ellipses A_i and B_i $(i = 1, 2, 3)$ represent the sets that functions a_i and b_i work, respectively. Here $A_3 \subseteq A_1 \cap A_2$ and $B_3 \subseteq B_1 \cap B_2$. The hazard rate matrix is block diagonal. The intersection between the sets A_i and $B_j (i, j = 1, 2, 3)$ does not affect the reducibility of the matrix χ.

Because $F_b(t)$ is nonincreasing, $F_b(t) = 0$ and $\chi_{ba}(t) = 0$. However, χ_{ab} may not necessarily be zero, because the condition, that dF_a/dt and F_a approach zero simultaneously, always holds when $F_a(t) \geqslant F_b(t)$. Therefore, in this case χ is an upper triangular matrix.

For the reducible case of an n-function system, we label the functions by several classes $a, b, c, d, \ldots$. In addition, for each of the classes, e.g., the class a, there are individual functions labeled by $a_1, a_2, \ldots$. The state of a function in a only depends on that particular function or other functions in the same class. This relationship is illustrated in Fig. 9. In this way the classes in the system are independent, resulting in a block diagonal matrix, i.e.,

$$\chi = \begin{pmatrix} \chi^{(a)} & 0 & \cdots \\ 0 & \chi^{(b)} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \tag{29}$$

where $\chi^{(a)}, \chi^{(b)}, \ldots$ denotes the hazard rate matrices of the function classes $a, b, \ldots$, respectively. $\chi^{(a)}, \chi^{(b)}, \ldots$ are irreducible matrices. For the irreducible case of an n-function system, there is at least one function whose set belongs to that of all other functions.

Next the estimation of hazard rate matrices χ using the MaxEnt method for reducible and irreducible cases is presented.

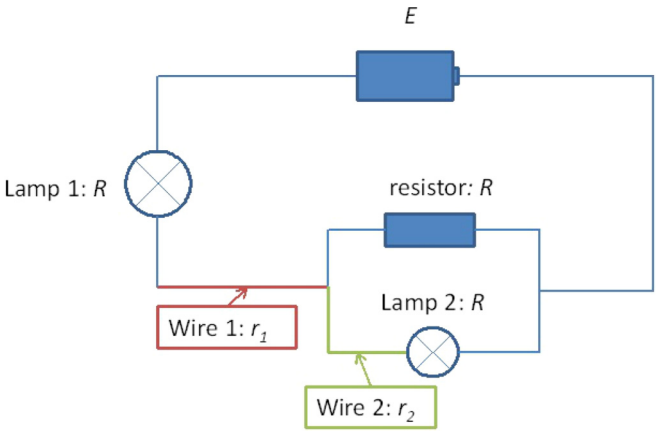

FIG. 10. Schematic of the two-lamp circuit problem. The two functions are labeled by the two lamps. The series-parallel circuit indicates that the hazard rate matrix for this system is irreducible.

B. MaxEnt for the reducible case

Using the above double-function system, the lifetimes of a and b are denoted as t_a and t_b, respectively. For the reducible case, the hazard rate matrix is diagonal:

$$\chi = \begin{pmatrix} \chi_{aa} & 0 \\ 0 & \chi_{bb} \end{pmatrix}. \tag{30}$$

The two elements can directly be calculated from the marginal distributions of t_1 and t_2 as

$$\begin{aligned} \dot{\chi}_{aa} - \chi_{aa}^2 - \chi_{aa} \frac{d \ln p_a(t)}{dt} = 0, \\ \dot{\chi}_{bb} - \chi_{bb}^2 - \chi_{bb} \frac{d \ln p_b(t)}{dt} = 0, \end{aligned} \tag{31}$$

where p_a and p_b are the marginal distributions of t_1 and t_2. The related joint distribution of $p(t_a, t_b)$ is also derived and shown in Appendix C. In this way, the double-function system can be decomposed into two uncorrelated single-function systems.

C. MaxEnt for the irreducible case

The irreducible cases correspond to systems that cannot be decomposed into isolated functions. The lifetimes of individual functions are correlated with each other due to the logical relationship among these functions. Mathematically the irreducible cases yield hazard rate matrices with off-diagonal terms, e.g., for a double-function system

$$\chi = \begin{pmatrix} \chi_{aa} & \chi_{ab} \\ 0 & \chi_{bb} \end{pmatrix}. \tag{32}$$

The detailed derivations of the irreducible case are presented in Appendix D. To demonstrate the main idea, one may consider a system for which the following equation holds:

$$\frac{d}{dt}\begin{pmatrix} F_a \\ F_b \end{pmatrix} = -\begin{pmatrix} \chi_{aa} & \chi_{ab} \\ 0 & \chi_{bb} \end{pmatrix}\begin{pmatrix} F_a \\ F_b \end{pmatrix}. \tag{33}$$

A typical example is a Markovian aging process. Note that there are three states as shown in Table IV. The probability for the system in the jth state at time t is denoted by $q(j, t)$. The Markovian property leads to

$$\frac{dq(j,t)}{dt} = \sum_i \kappa_{ji} p(i,t), \tag{34}$$

where κ is the transfer matrix and κ_{ji} is the probability of transition from state i to state j. Appendix D shows that Eq. (33) holds when $\kappa_{21} = 0$. In this case,

$$\begin{aligned} \chi_{aa} &= \kappa_{23}, \\ \chi_{ab} &= \kappa_{13} - \kappa_{23}, \\ \chi_{bb} &= \kappa_{13} + \kappa_{12}. \end{aligned} \tag{35}$$

It should be noted that the elements of κ except κ_{21} in the above equation are unknown quantities and are to be estimated by MaxEnt.

Recall that the lifetime of function a cannot be less than that of function b. This constraint can be incorporated into the joint distribution using the step function $\theta(t_a - t_b)$. The second constraint is the physical condition that the probability for the two functions breaking down simultaneously is finite

when $\kappa_{13} \neq 0$. The condition of $\kappa_{13} \neq 0$ is the probability of transition from the normal state to the failure state. This constraint can be incorporated into the joint distribution using the delta function $\delta(t_a - t_b)$. The detailed proof is given in Appendix D. Based on the above constraints, the joint distribution is written

$$\begin{aligned} p(t_a, t_b) = &(1-\mu)p_{\mathrm{N}}(t_a, t_b)\theta(t_a - t_b) \\ &+ \mu p_{\mathrm{A}}(t_a)\delta(t_a - t_b), \end{aligned} \tag{36}$$

where $\theta(\cdot)$ denotes the step function and $\delta(\cdot)$ denotes the delta function. In the above equation p_{N} and

$$p_{\mathrm{A}} = \frac{1}{\mu}\kappa_{13} \exp\left(\int_0^t \kappa_{22} dt\right) \tag{37}$$

are both normalized PDFs; $0 < \mu < 1$ is the total probability that the two functions break down simultaneously. p_{N} and p_{A} can be determined by MaxEnt as

$$\begin{aligned} \delta\bar{S} = &-\iint_{\mathcal{D}} p_{\mathrm{N}}(t_a, t_b)[\ln p_{\mathrm{N}}(t_a, t_b) + \alpha_{\mathrm{N}}]dt_a dt_b \\ &- \iint_{\mathcal{D}} p_{\mathrm{N}}(t_a, t_b)\left[\sum_i \beta_{\mathrm{N},i} h_{\mathrm{N},i}(t_a, t_b)\right] dt_a dt_b \\ &- \int p_{\mathrm{A}}(t)\left[\ln p_{\mathrm{A}}(t) + \alpha_{\mathrm{A}} + \sum_{i'} \beta_{\mathrm{A},i'} g_{\mathrm{A},i'}(t)\right] dt \\ = &\,0, \end{aligned} \tag{38}$$

where $\mathcal{D}$ denotes that the domain of $t_a > t_b$, α_{N}, $\beta_{\mathrm{N},i}$, α_{A}, $\beta_{\mathrm{A},i'}$ are the Lagrangian multipliers and p_{N} and p_{A} are independent functions.

The connection between the hazard rate matrix and the marginal distributions can then be constructed as the following equations:

$$\begin{aligned} &\frac{d\chi_{aa}}{dt} - \chi_{aa}^2 \frac{p_{\mathrm{N},a} - p_{\mathrm{N},b}}{p_{\mathrm{N},a}} - \chi_{aa}\frac{d\,\ln p_{\mathrm{N},a}}{dt} = 0, \\ &\frac{d\chi_{bb}}{dt} - \chi_{bb}^2 - \chi_{bb}\frac{d\,\ln p_b}{dt} = 0, \\ &\chi_{ab} = \mu\chi_{bb}\frac{p_{\mathrm{A}}}{p_b} - \chi_{aa}, \end{aligned} \tag{39}$$

where $p_{\mathrm{N},a}$ and $p_{\mathrm{N},b}$ are the marginal distributions related to the PDF of $p_{\mathrm{N}}(t_a, t_b)$ and $p_b = (1-\mu)p_{\mathrm{N},b} + \mu p_{\mathrm{A}}$ is the marginal distribution related to the joint PDF of $p(t_a, t_b)$. The solution to the above equation is the resulting hazard rate matrix of the system.

D. Two-lamp circuit model

A two-lamp circuit shown in Fig. 10 is used to represent a double-function system for demonstration. The two functions of the system are denoted by the two lamps. One can see that if lamp 2 works lamp 1 must work. This dependence indicates the system is irreducible. The degradation driving factor in the problem is the heat generation of the two wires.

In the model the physical degradation is also related to the temperature of the wires. The heat conductance between the wires is also considered in this two-lamp circuit model. The equation of heat conductance is assumed as

$$\begin{aligned} \frac{dT_1}{dt} &= -\tilde{\Upsilon}T_1 + \tilde{\Upsilon}T_2, \\ \frac{dT_2}{dt} &= -\tilde{\Upsilon}T_2 + \tilde{\Upsilon}T_1, \end{aligned} \tag{40}$$

where $T_{1(2)}$ denote the temperature of wire 1(2) and $\tilde{\Upsilon}$ is assumed as a material dependent constant. For demonstration purposes, the regime where the temperatures are much higher than the Debye temperature is considered. The resistance is proportional to the temperature [41], i.e., $r_1, r_2 \propto T$, and the resistances follow the equation

$$\begin{aligned} \frac{dr_1}{dt} &= -\Upsilon r_1 + \Upsilon r_2 + \Omega_1, \\ \frac{dr_2}{dt} &= -\Upsilon r_2 + \Upsilon r_1 + \Omega_2, \end{aligned} \tag{41}$$

where the first two terms in the right hand side represent the heat conduction between the two wires. The term Υ is assumed as a material dependent constant, and the term $\Omega_{1(2)}$ represents the heat generation of the two wires. The heat generation of wire 1(2) is proportional to its power $I_{1(2)}^2 \times r_{1(2)}$, where $I_{1(2)}$ denote the electric current in wire 1(2). Using Kirchhoff's law, the electric current in the circuit model can be directly calculated, and is presented in Appendix E. The heat generation variables read

$$\begin{aligned} \Omega_1 &= \tilde{\Lambda}_1 r_1 \left(\frac{(2R + r_2)}{3R^2 + 2(r_1 + r_2)R + r_1 r_2}\right)^2, \\ \Omega_2 &= \tilde{\Lambda}_2 r_2 \left(\frac{R}{3R^2 + 2(r_1 + r_2)R + r_1 r_2}\right)^2, \end{aligned} \tag{42}$$

where R is the overall resistance of the lamps and the resistor. $\tilde{\Lambda}_{1(2)}$ is the effective ratio parameter. For illustration purposes, let $2\tilde{\Lambda}_1 = \tilde{\Lambda}_2 \equiv \tilde{\Lambda}$.

The terms $p_{\mathrm{N},1}$, $p_{\mathrm{N},2}$, and p_{A} can be estimated separately. Note that in this case Eq. (36) indicates that $p_{\mathrm{N},1}(0) = 0$. This prior knowledge leads to the breaking down of the problem. To avoid that a nonlinear measure of the integral of t, i.e., $dt \to t\,dt$, is added. This modification of the measure is equivalent to the maximization of the cross entropy. The proof is as follows:

$$\int (-p\,\ln p)\nu(t)dt \equiv -\int p\nu(t)\ln\frac{p\nu(t)}{\nu(t)}dt, \tag{43}$$

where $\nu(t)$ is the proper measure. Recall the constraint of $\int p\nu\,dt = 1$. The variational term of the entropy is

$$\begin{aligned} &\delta\left[\int (-p\,\ln p)\nu\,dt - \alpha\int p\nu\,dt\right] \\ &= \delta\left[-\int p\nu\,\ln\frac{p\nu}{\nu}dt - \alpha\int p\nu\,dt\right] \\ &= \delta\left[-\int p\nu\,\ln\frac{p\nu}{\exp(-\alpha_\nu)\nu}dt - \alpha_p\int p\nu\,dt\right] \\ &= \delta\left[-\int \tilde{p}\,\ln\frac{\tilde{p}}{p_0}dt - \alpha_p\int \tilde{p}\,dt\right], \end{aligned} \tag{44}$$

where $\alpha = \alpha_p + \alpha_\nu$ and α_p are Lagrange multipliers, $p_0 \equiv \exp(-\alpha_\nu)\nu(t)$ is a prior distribution which satisfies $\exp(-\alpha_\nu)\int \nu(t)dt = 1$, and $\tilde{p} \equiv p\nu$ is the variational PDF. Note that this modification requires prior knowledge about hazard rates of the two functions and its time derivations near $t = 0$.

To demonstrate the effectiveness of the method, numerical experiments are made to simulate actual experimental data. Parameters of $R = 1$, $\Upsilon = 0.01$, and ($\Delta = 0.09$, $\Lambda = 2$) are used to generate simulated data. In this setting the numerical calculation yields $\mu \approx 0.1266$. In estimation $\chi(\tau = 0.32)$ is used as the additional constraint. The estimated results are shown in Fig. 11, which are close to the simulated data. These results indicate that the method combining the linear assumption with the double-moment constraint is effective to analyze a multifunction system.

V. CONCLUSIONS AND DISCUSSIONS

Inspired by statistical mechanics in physics, as a theory of statistical inference, a MaxEnt approach to reliability is developed in this study, allowing for constructing the hazard rate function in a rational manner. The hazard rate function is a fundamental quantity in the disciplines of reliability and risk analysis, characterizing the aging process of a system. In particular, the time-dependent hazard rate can fully describe the dynamics of the aging system. The basic idea of the developed method is to recast an estimation problem to a probabilistic inference problem using the principle of maximum entropy. The most probable hazard rate function is the one that maximizes the information entropy. Information such as observed data in terms of statistical moments are used as constraints to obtain the most probable hazard rate function.

It is shown that different shapes of hazard rate functions, such as the widely observed bathtub shape in engineering, upside down bathtub shape in biological system, and W/N shapes can all be interpreted as the most probable hazard rate under certain constraints. In addition to the single-function system, the multifunction system consisting of multiple individual isolated and/or correlated functions is investigated by extending the hazard rate function to a multidimensional hazard rate matrix. For a system with isolated functions, it can be reduced to a set of independent hazard rate functions and yield a block diagonal hazard rate matrix. For a system with correlated functions, the interaction terms yield off-diagonal terms in the hazard rate matrix and the system is the so-called irreducible system. The overall method is demonstrated using numerical examples, and the effectiveness of the method is verified for both single- and multifunction systems.

The application of the proposed methods in general involves the following steps. (1) Process the observed lifetime data as testable information, e.g., calculate the statistical moments; (2) choose several time points $\tau_1, \tau_2, \ldots$ and evaluate the corresponding hazard rate $x(\tau_1), x(\tau_2), \ldots$ from the lifetime data; (3) use the hazard rate at one chosen time point and obtain the parameters, i.e., the Lagrange multipliers, using the testable information as constraints; (4) perform consistency check following step (3) using other time points.

To understand the "physical" meaning of parameters, the information beyond lifetime is necessary. In this paper, the approach is independent of specific systems. This is based on the assumption that no other information is accessed except to the information of lifetime. In realistic applications, the system may contain other information, e.g., lifetime data on components, structures, and their correlations and so on. The information can result in different hazard rate functions when it is used as constraints. Moreover, the linear approximation is applied in the multifunction cases. The linear approximation may not be valid and the justification must be carefully made.

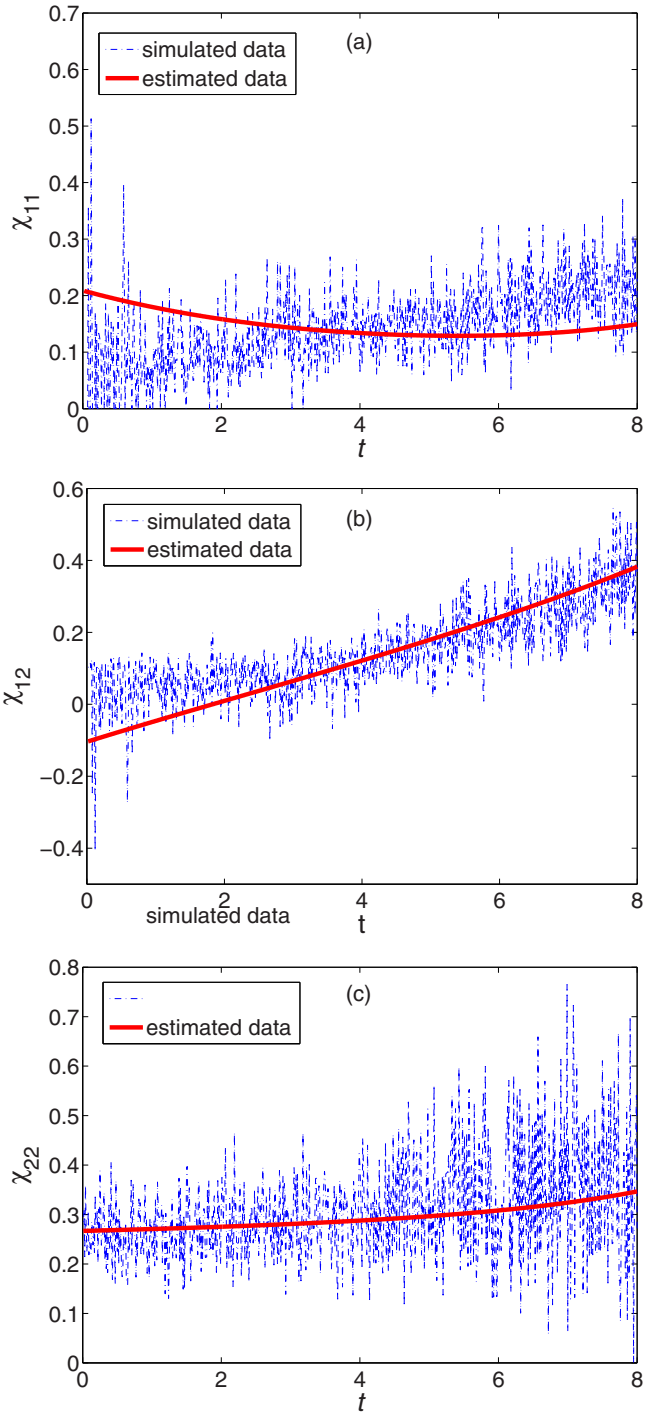

FIG. 11. Comparisons of the estimated components of the hazard rate matrix with the simulated data.

The features of structures and correlations are also related to nonlinearity of multicomponent systems. How to fuse the hierarchical information using the proposed method should to be further investigated.

In addition, Eq. (7) generates the PDFs belonging to the exponential family. For specific cases, to generate PDFs which do not belong to the exponential family, different equations of motion should be considered. There are at least two ways to achieve this. One is to adopt a nonlinear measure in the integrals, which has been discussed in the previous section. Another is to construct a distribution $F(t) = \exp(-X)\phi(X)$ where $\phi(X)$ is a suitable function satisfying $dF/dt \geqslant 0, \phi(0) = 1$, such as $F(t) = \exp(-X)(1 + X)$.

The significance of the developed method lies in the fact that (1) it provides a rational instrument to construct the hazard rate function consistently given any available information from experimental data, (2) it provides a statistical mechanics-based approach to interpreting the generation of different shapes of the hazard rate curves observed in the field of reliability over the past few decades, and (3) it provides a theoretical bridge linking the reliability engineering to one of the most fundamental principles, MaxEnt, in physics. This study lays out a possible pathway to the enlightened goal in Ref. [36], which is "the reliability theory is a new science."

ACKNOWLEDGMENTS

The work in this study was supported by National Basic Program of China (Grant No. 2016YFA0301201), NSFC (Grants No. 11534002 and No. 51975546), NSAF (Grants No. U1730449, No. U1530401, No. U1930402, and No. U1930403), and Science Challenge Project (Grant No. TZ2018007). The support is greatly acknowledged. The authors would like to thank the anonymous reviewers for their constructive comments.

APPENDIX A: MOST PROBABILITY DISTRIBUTION AND THE EQUATION OF MOTION OF THE HAZARD RATE FUNCTION

In general, the entropy is given as

$$S = -\int_{t_{\min}}^{t_{\max}} p(t)\ln p(t)dt, \tag{A1}$$

where $t_{\min}$ and $t_{\max}$ are the minimum and the maximum lifetime, respectively. We maximize the entropy with moment constraints using variations

$$\delta\bar{S} = \delta\left\{-\int_{t_{\min}}^{t_{\max}} dt' p(t')\ln p(t') - \beta_i\int_{t_{\min}}^{t_{\max}} dt' g_i(t')p(t') - \alpha\int_0^{t_{\max}} dt'(t')p(t')\right\} = 0, \tag{A2}$$

where

$$\int_{t_{\min}}^{t_{\max}} dt' g_i(t')p(t'), \tag{A3}$$

with $i = 1, 2, \ldots$, are the moments from observation.

The most probable distribution is

$$p(t) = \begin{cases} \frac{1}{Z}\exp\left[-\sum_i \beta_i g_i(t)\right], & t_{\min} \leqslant t \leqslant t_{\max}, \\ 0, & \text{else}, \end{cases} \tag{A4}$$

where

$$Z = \int_{t_{\min}}^{t_{\max}} \exp\left[-\sum_i \beta_i g_i(t)\right]dt \tag{A5}$$

is the partition function or so-called normalizing constant in probability.

The motion equation in Eq. (7) can be also derived from the most probable distribution. We use

$$p(t) = x\,\exp(-X), \tag{A6}$$

where

$$X(t) = \int_{t_{\min}}^{t} x(t')dt', \tag{A7}$$

to obtain

$$x(t)\exp[-X(t)] = \frac{1}{Z}\exp\left[-\sum_i \beta_i g_i(t)\right]. \tag{A8}$$

Its derivation with respect to time t reads

$$\dot{x}\,\exp(-X) - x^2\exp(-X) = \frac{1}{Z}\exp\left[-\sum_i \beta_i g_i(t)\right]\left[-\sum_i \beta_i \dot{g}_i(t)\right]. \tag{A9}$$

Namely,

$$\dot{x} - x^2 + \sum_i \beta_i \dot{g}_i(t)x = 0 \tag{A10}$$

is obtained. By assuming $t_{\min} = 0$ and $g_i(t)$ satisfy $g_i(0) = 0$, the initial condition is given by $x(0) = 1/Z$. Finally, the solution to Eq. (7) is

$$x(t) = \frac{x_0\exp\left(-\sum_i \beta_i g_i\right)}{1 - x_0\int_0^t \exp\left(-\sum_i \beta_i g_i\right)dt'}. \tag{A11}$$

We define a new function $\tilde{x} = x\,\exp[\sum_i \beta_i g_i(t)]$ and substitute it into Eq. (7) to have

$$\dot{\tilde{x}} = \exp\left[-\sum_i \beta_i g_i(t)\right]\tilde{x}^2. \tag{A12}$$

Solving Eq. (A12) we obtain

$$\frac{1}{\tilde{x}_0} - \frac{1}{\tilde{x}_t} = \int_0^t \exp\left[-\sum_i \beta_i g_i(t')\right]dt'. \tag{A13}$$

The initial condition is $\tilde{x}(0) = \exp[-\sum_i \beta_i g_i(0)]x_0 = x_0$. From the above equation and $x = \tilde{x}\,\exp[-\sum_i \beta_i g_i(t)]$, one can verify Eq. (A11).

APPENDIX B: NUMBER OF INFLECTION POINTS

The most probable hazard rate function constructed using a double-moment constraint can have at most one point of inflection. The proof is given below.

Rewrite Eq. (14) as

$$\dot{x} = x(x - \beta_1 - 2\beta_2 t), \tag{B1}$$

where $\dot{()}$ denotes $d()/dt$. For $x > 0$ the points of inflection must be located at $x = \beta_1 + 2\beta_2 t$. For convenience we define a function $y(t) = \beta_1 + 2\beta_2 t$. Consider the function $x(t) - y(t)$, where $x(t)$ is the hazard rate function which satisfies Eq. (14). The time derivative of $x(t) - y(t)$ is

$$\dot{x} - \dot{y} = x(x - y) - 2\beta_2, \tag{B2}$$

where the hazard rate function satisfies $x(t) > 0$.

For $\beta_2 < 0$, if $x(t_0) - y(t_0) \geqslant 0$ with $t_0 > 0$ then the above equation implies that $x(t) - y(t) > 0$ with arbitrary $t > t_0$, i.e., the region $x(t) \geqslant y(t)$ is the absorption domain for Eq. (14) with $\beta_2 < 0$. Therefore, there is at most one point of inflection for $\beta_2 < 0$. More precisely, for $\beta_2 < 0, x_0 > 0 > \beta_1$ and for $\beta_2 < 0, x_0 \geqslant \beta_1 > 0$, there is no point of inflection. For $\beta_2 < 0, 0 < x_0 < \beta_1$ there is one point of inflection and the function is in bathtub shapes.

Similarly, for $\beta_2 > 0$, if $x(t_0) - y(t_0) \leqslant 0$ with $t_0 > 0$ then the above equation implies that $x(t) - y(t) < 0$ with arbitrary $t > t_0$, i.e., the region $x(t) \leqslant y(t)$ is the absorption domain for Eq. (14) with $\beta_2 > 0$. Therefore, there is also at most one point of inflection for $\beta_2 > 0$. More precisely, for $\beta_2 > 0, 0 < x_0 \leqslant \beta_1$ and for $\beta_2 > 0, x_0 \int_0^\infty \exp(-\beta_1 t - \beta_2 t^2)dt > 1$, there are no points of inflection; for $\beta_2 > 0$, $x_0 \int_0^\infty \exp(-\beta_1 t - \beta_2 t^2)dt = 1$ there is one point of inflection and the function is in upside down bathtub shapes.

The hazard rate for $\beta_2 > 0, x_0 \leqslant \beta_1$ is monotonically decreasing; however, this condition violates Eq. (17). From the above discussion, the shapes of $x(t)$ can be divided into three types, i.e., the monotonically increasing, the bathtub shape, and the upside down bathtub shape.

It can be shown, as follows, that the hazard rate function constructed using an n-moment constraint contains at most $n - 1$ inflection points. For convenience, let $y(t) = \sum_{n=1} n\beta_n t^{n-1}$ and the time derivative of $x(t) - y(t)$ reads

$$\dot{x} - \dot{y} = x(x - y) - \sum_{n=2} n(n-1)\beta_n t^{n-2}. \tag{B3}$$

The location of the absorption domain depends on the sign of $\dot{y} = \sum_{n=2} n(n-1)\beta_n t^{n-2}$. The function $y(t)$ here has at most $n - 2$ inflection points, i.e., the function $\dot{y}$ has at most $n - 2$ zero points.

We denote $t_1, t_2, \ldots, t_{n-2}$ as the zero points and $0 < t_1 < t_2 < \cdots < t_{n-2}$. The sign of $\dot{y}$ in the time intervals $(t_i, t_{i+1}), i = 1, 2, \ldots, n-3$ and intervals $[0, 1), (t_{n-2}, \infty)$ remains unchanged. In any one of these time intervals, $x(t)$ has at most one inflection point, and therefore $x(t)$ has a total of $n - 1$ inflection points at most. In addition, the n-moment constraint is capable of generating $2n - 1$ types of shapes. The monotonically decreasing curve has been excluded because it violates the normalization condition.

APPENDIX C: JOINT AND THE MARGINAL DISTRIBUTIONS FOR REDUCIBLE AND IRREDUCIBLE CASES IN A DOUBLE-FUNCTION SYSTEM

In the reducible cases, the marginal distributions related to the joint PDF of $p(t_1, t_2)$ can be obtained by MaxEnt,

$$\begin{aligned}\bar{S} = &- \int_{t_{a,\min}}^{t_{a,\max}} \int_{t_{b,\min}}^{t_{b,\max}} p(t_1, t_2)[\ln p(t_1, t_2) - \alpha]dt_1 dt_2 \\ &- \sum_i \beta_i \int_{t_{a,\min}}^{t_{a,\max}} \int_{t_{b,\min}}^{t_{b,\max}} h_i(t_1, t_2)p(t_1, t_2)dt_1 dt_2,\end{aligned} \tag{C1}$$

where α and $\beta_i, i = 1, 2, \ldots$ are the Lagrangian multipliers and $h_i(t_1, t_2)$ are the correlation functions. We maximize $\bar{S}$ to obtain the joint PDF for $t_i \in [t_{i,\min}, t_{i,\max}], i = a, b$ as

$$p(t_a, t_b) = \frac{1}{Z} \exp\left[-\sum_i \beta_i h_i(t_a, t_b)\right], \tag{C2}$$

where $Z = \int_{t_{a,\min}}^{t_{a,\max}} \int_{t_{b,\min}}^{t_{b,\max}} p(t_a, t_b)dt_a dt_b$ is the partition function. The marginal distributions are given as

$$\begin{aligned}p_a(t) &= \int_{t_{b,\min}}^{t_{b,\max}} p(t, t_b)dt_b, \\ p_b(t) &= \int_{t_{a,\min}}^{t_{a,\max}} p(t_a, t)dt_a.\end{aligned} \tag{C3}$$

In the irreducible cases, we maximize the entropy in Eq. (38) and obtain

$$\begin{aligned}p_{\mathrm{N}}(t_a, t_b) &= \frac{1}{Z_{\mathrm{N}}} \exp\left[-\sum_i \beta_{\mathrm{N},i} h_{\mathrm{N},i}(t_a, t_b)\right], \\ p_{\mathrm{A}}(t) &= \frac{1}{Z_{\mathrm{A}}} \exp\left[-\sum_i \beta_{\mathrm{A},i} g_{\mathrm{N},i}(t)\right].\end{aligned} \tag{C4}$$

The marginal distributions $p_i(t) = p_{\mathrm{N},i}(t) + p_{\mathrm{A},i}(t), i = a, b$ are given as

$$\begin{aligned}p_{\mathrm{N},a}(t) &= (1 - \mu) \int_{t_{b,\min}}^{t_{b,\max}} dt_b p_{\mathrm{N}}(t, t_b), \\ p_{\mathrm{N},b}(t) &= (1 - \mu) \int_{t_{a,\min}}^{t_{a,\max}} dt_a p_{\mathrm{N}}(t_a, t), \\ p_{\mathrm{A},a}(t) &= p_{\mathrm{A},b}(t) = \mu p_{\mathrm{A}}(t),\end{aligned} \tag{C5}$$

where μ is defined as before.

APPENDIX D: MARKOVIAN AGING PROCESS

We consider a Markovin process, and denote the probability of the trajectory $(i_1, t_1; i_2, t_2; \ldots)$ as $q(i_1, t_1; i_2, t_2; \ldots)$. The Markovian approximation is

$$\begin{aligned}&q(j_{m+1}, t_{m+1}|j_m, t_m; j_{m-1}, t_{m-1}; \ldots; j_0, t_0) \\ &\quad = q(j_{m+1}, t_{m+1}|j_m, t_m),\end{aligned} \tag{D1}$$

where the conditional probability $q(A|B)$ denotes the probability of event A conditional on event B. The diffusion equation

is

$$\frac{dq(j,t)}{dt} = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_i [q(j,t|i,t-\Delta t)-\delta_{i,j}]q(i,t-\Delta t) = \sum_i \kappa_{ji}p(i,t). \tag{D2}$$

The joint PDF $p(t_a,t_b)$ is related to the trajectory probability. In the irreducible cases, one has

$$p(t_a,t_b)\Delta t^2 = q(3,t_a;2,t_a-\Delta t;2,t_a-2\Delta t;\ldots;2,t_b;1,t_b-\Delta t;1,t_b-2\Delta t;\ldots;1,0), \tag{D3}$$

where $q(3,t_a;2,t_a-\Delta t;\ldots)$ is the probability of the trajectory. It is shown that

$$\begin{aligned} p(t,t)\Delta t^2 &= q(3,t_a;1,t_a-\Delta t;1,t_a-2\Delta t;\ldots;1,t_b;1,t_b-\Delta t;1,t_b-2\Delta t;\ldots;1,0)\\ &= q(3,t_a|1,t_a-\Delta t)q(1,t_a-\Delta t;1,t_a-2\Delta t;\ldots;1,t_b;1,t_b-\Delta t;1,t_b-2\Delta t;\ldots;1,0)\\ &= q(3,t_a|1,t_a-\Delta t)q(1,t_a-\Delta t|1,t_a-2\Delta t)\ldots q(1,\Delta t|1,0)\\ &= \kappa_{13}(1+\kappa_{22}\Delta t)^{t/\Delta t}\Delta t. \end{aligned} \tag{D4}$$

Furthermore,

$$\begin{aligned} p(t,t-\Delta t) &= q(3,t_a;2,t_a-\Delta t;1,t_a-2\Delta t;\ldots;1,t_b;1,t_b-\Delta t;1,t_b-2\Delta t;\ldots;1,0)\frac{1}{\Delta t^2}\\ &= q(3,t_a|2,t_a-\Delta t)q(2,t_a-\Delta t;1,t_a-2\Delta t;\ldots;1,t_b;1,t_b-\Delta t;1,t_b-2\Delta t;\ldots;1,0)\frac{1}{\Delta t^2}\\ &= q(3,t_a|2,t_a-\Delta t)q(2,t_a-\Delta t|1,t_a-2\Delta t)\ldots q(1,\Delta t|1,0)\frac{1}{\Delta t^2}\\ &= \kappa_{23}\kappa_{12}(1+\kappa_{22}\Delta t)^{t/\Delta t-1}\Delta t^2\frac{1}{\Delta t^2}\\ &= \kappa_{23}\kappa_{12}(1+\kappa_{22}\Delta t)^{t/\Delta t-1}. \end{aligned} \tag{D5}$$

The above two equations imply that

$$p(t,t) = \lim_{\Delta t\to 0}\frac{\kappa_{13}}{\Delta t}\exp\left(\int_0^t \kappa_{22}dt'\right) = \lim_{t'\to t}\delta(t-t')\kappa_{13}\exp\left(\int_0^t \kappa_{22}dt''\right). \tag{D6}$$

Comparing the above equation to Eq. (36), one has

$$\mu p_{\rm A} = \kappa_{13}\exp\left(\int_0^t \kappa_{22}dt'\right). \tag{D7}$$

We rewrite the diffusion equation as

$$\begin{aligned} \frac{d}{dt}q_1 &= I^{(2\to 1)} - I^{(1\to 3)} + I^{(3\to 1)} - I^{(1\to 2)},\\ \frac{d}{dt}q_2 &= -I^{(2\to 1)} + I^{(1\to 2)} + I^{(3\to 2)} - I^{(2\to 3)},\\ \frac{d}{dt}q_3 &= I^{(2\to 3)} - I^{(3\to 2)} - I^{(3\to 1)} + I^{(1\to 3)}, \end{aligned} \tag{D8}$$

where $I^{(i\to j)}(t) = \kappa_{ji}p(i,t)$ is the current of probability from i to j at time t.

The survival functions of a and b are denoted as F_a and F_b, respectively. We denote the probability that the system is in state i, $i=1,2$ without previously being in state 3 as G_i. Such a condition leads to $F_a = G_1 + G_2$, and G_1, G_2, F_b satisfy the following equations:

$$\begin{aligned} \frac{d}{dt}G_1 &= -\frac{G_1}{q_1}[I^{(1\to 2)} + I^{(1\to 3)}] + \frac{G_2}{q_2}I^{(2\to 1)},\\ \frac{d}{dt}G_2 &= -\frac{G_2}{q_2}[I^{(2\to 1)} + I^{(2\to 3)}] + \frac{G_1}{q_1}I^{(1\to 2)},\\ \frac{d}{dt}F_b &= -\frac{F_b}{q_2}[I^{(1\to 2)} + I^{(1\to 3)}]. \end{aligned} \tag{D9}$$

We rewrite the above equation as

$$\frac{d}{dt}\begin{pmatrix} G_1\\ G_2\\ F_b\end{pmatrix} = \begin{pmatrix} -\kappa_{12}-\kappa_{13} & \kappa_{21} & 0\\ \kappa_{12} & -\kappa_{21}-\kappa_{23} & 0\\ 0 & 0 & -\kappa_{12}-\kappa_{13}\end{pmatrix}\begin{pmatrix} G_1\\ G_2\\ F_b\end{pmatrix}. \tag{D10}$$

If $\kappa_{21}=0$, the above equation is reduced to

$$\frac{d}{dt}\begin{pmatrix} F_a\\ F_b\end{pmatrix} = \begin{pmatrix} -\kappa_{23} & \kappa_{23}-\kappa_{13}\\ 0 & -\kappa_{13}-\kappa_{12}\end{pmatrix}\begin{pmatrix} F_a\\ F_b\end{pmatrix}. \tag{D11}$$

The hazard rate matrix is

$$\chi = -\begin{pmatrix} -\kappa_{23} & \kappa_{23}-\kappa_{13}\\ 0 & -\kappa_{13}-\kappa_{12}\end{pmatrix}, \tag{D12}$$

where the elements of the matrix are to be determined by MaxEnt.

The result of Eq. (39) is obtained as follows. Considering the initial condition $F_a(0)=1$, $F_b(0)=1$, we rewrite Eq. (D11) in integral form as

$$\begin{aligned} F_a &= \exp(-\chi_{aa})\left[1-\int_0^t \chi_{ab}\exp(\chi_{aa}-\chi_{bb})dt'\right],\\ F_b &= \exp(-\chi_{bb}). \end{aligned} \tag{D13}$$

With $F_i = \int_0^t p_i(t')dt'$, $i = a, b$, we take the time derivative of the above equation to have

$$p_a = \chi_{aa}\exp(-\chi_{aa})\left[1 - \int_0^t \chi_{ab}\exp(\chi_{aa} - \chi_{bb})dt'\right] + \chi_{ab}\exp(-\chi_{bb}),$$
$$p_b = \chi_{bb}\exp(-\chi_{bb}), \tag{D14}$$

where $p_a = p_{\mathrm{N},a}(t) + p_{\mathrm{A},a}(t)$ and $p_b = p_{\mathrm{N},b}(t) + p_{\mathrm{A},b}(t)$ are the marginal distributions shown in Eq. (C5). Combining the above equations with Eq. (D7) and taking the time derivative again, one can verify Eq. (39). Given $t_{a,\min} = t_{b,\min} = 0$ and $t_{a,\max} = t_{b,\max} = t_{\max}$, the solution to Eq. (39) is finally obtained as

$$\chi_{aa} = -\frac{p_{\mathrm{N},a}}{\int_0^t dt'(p_{\mathrm{N},a} - p_{\mathrm{N},b})},$$
$$\chi_{bb} = \frac{p_b}{1 - \int_0^t dt' p_b},$$
$$\chi_{ab} = \frac{\mu p_{\mathrm{A}}}{1 - \int_0^t dt' p_b} + \frac{p_{\mathrm{N},a}}{\int_0^t dt'(p_{\mathrm{N},a} - p_{\mathrm{N},b})}. \tag{D15}$$

APPENDIX E: POWER OF THE WIRES IN THE TWO-LAMP CIRCUIT MODEL

Let $P_{1(2)}$ denote the power of wire 1(2). The term E denotes the electromotive force of the source. The currents in wires 1 and 2 are

$$I_1 = E/\left[R + r_1 + \frac{R(R + r_2)}{2R + r_2}\right] = \frac{E(2R + r_2)}{3R^2 + 2(r_1 + r_2)R + r_1 r_2},$$
$$I_2 = I_1\frac{R}{2R + r_2} = \frac{ER}{3R^2 + 2(r_1 + r_2)R + r_1 r_2}. \tag{E1}$$

Consequently,

$$P_1 = I_1^2 r_1 \propto \left(\frac{2R + r_2}{3R^2 + 2(r_1 + r_2)R + r_1 r_2}\right)^2 r_1,$$
$$P_2 = I_2^2 r_2 \propto \left(\frac{R}{3R^2 + 2(r_1 + r_2)R + r_1 r_2}\right)^2 r_2. \tag{E2}$$

We introducing an effective coefficient term to the above equation to obtain Eq. (42).

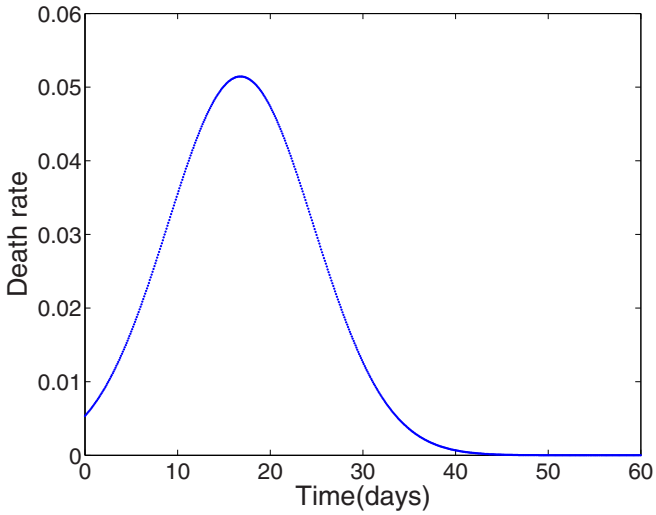

FIG. 12. Estimated death rate curve for fruit fly.

APPENDIX F: BIOLOGICAL EXAMPLE: FRUIT FLY

Using the data in Table III in Ref. [43], the first and second moments of the lifetime of the fruit fly are obtained as $E(t) = 17.12$, $E(t^2) = 349.73$, and $t_{\max} = 100$. Using the double-moment model, the estimated death rate of the fruit fly is presented in Fig. 12. The death rate curve is in upside down bathtub shape, which coincides qualitatively with the result in Ref. [43].

APPENDIX G: DETAILS IN NUMERICAL CALCULATION

To obtain the solution to the equation of motion numerically, a fourth order Runge-Kutta method with a time step of 0.01 is used with the chosen parameters shown in the main body of the text. To sample the lifetime data in simulation of the single- and two-lamp model, the plain Monte Carlo method is used (100 000 samples for each case). To estimate the Lagrange multipliers, the standard Newton-Raphson method is applied with a convergence criterion such that all relative errors are less than 1×10^{-4}.

[1] M. Rausand and A. Høyland, *System Reliability Theory: Models, Statistical Methods, and Applications* (John Wiley & Sons, New York, 2004).
[2] D. M. Boodman, J. Oper. Res. Soc. Am. **1**, 39 (1953).
[3] E. Halley, Philos. Trans. R. Soc. London **17**, 596 (1693).
[4] G.-A. Klutke, P. C. Kiessler, and M. A. Wortman, IEEE T. Reliab. **52**, 125 (2003).
[5] O. R. Jones *et al.*, Nature (London) **505**, 169 (2014).
[6] K. L. Wong, Qual. Reliab. Eng. Int. **5**, 29 (1989).
[7] B. D. Coleman, J. Appl. Phys. **29**, 968 (1958).
[8] W. I. Newman and S. L. Phoenix, Phys. Rev. E **63**, 021507 (2001).
[9] S. Santucci, L. Vanel, and S. Ciliberto, Phys. Rev. Lett. **93**, 095505 (2004).
[10] F. Kun, H. A. Carmona, J. S. Andrade Jr., and H. J. Herrmann, Phys. Rev. lett. **100**, 094301 (2008).
[11] N. Yoshioka, F. Kun, and N. Ito, Phys. Rev. Lett. **101**, 145502 (2008).
[12] S. Lennartz-Sassinek, I. G. Main, Z. Danku, and F. Kun, Phys. Rev. E **88**, 032802 (2013).
[13] L. Viitanen, M. Ovaska, S. K. Ram, M. J. Alava, and P. Karppinen, Phys. Rev. Appl. **11**, 024014 (2019).
[14] X. G. Peng and G. B. McKenna, Phys. Rev. E **93**, 042603 (2016).
[15] T. Jonsson, J. Mattsson, C. Djurberg, F. A. Khan, P. Nordblad, and P. Svedlindh, Phys. Rev. Lett. **75**, 4138 (1995).
[16] L. Laloux and P. LeDoussal, Phys. Rev. E **57**, 6296 (1998).

[17] A. Dechant, E. Lutz, D. A. Kessler, and E. Barkai, Phys. Rev. X **4**, 011022 (2014).
[18] P. Lunkenheimer, R. Wehn, U. Schneider, and A. Loidl, Phys. Rev. Lett. **95**, 055702 (2005).
[19] S. Boettcher, D. M. Robe, and P. Sibani, Phys. Rev. E **98**, 020602(R) (2018).
[20] Y. T. Lou, J. F. Xia, W. Tang, and Y. Chen, Phys. Rev. E **96**, 062418 (2017).
[21] H. Pham and C. D. Lai, IEEE T. Reliab. **56**, 454 (2007).
[22] D. V. Lindley, J. R. Stat. Soc.: Ser. B **20**, 102 (1958).
[23] Z. Ahmad, G. G. Hamedani, and N. S. Butt, Pak. J. Stat. Oper. Res. **15**, 87 (2019).
[24] C. D. Lai, M. Xie, and D. N. P. Murthy, Handbook Stat. **20**, 69 (2001).
[25] R. Jiang, Reliab. Eng. Syst. Safe. **119**, 44 (2013).
[26] R. Jiang, Int. J. Perform. Eng. **9**, 569 (2013).
[27] N. Ebrahimi, Sankhyā Ser. A **58**, 48 (1996).
[28] N. Ebrahimi, Sankhyā Ser. A **62**, 236 (2000).
[29] M. Asadi and N. Ebrahimi, Stat. Probab. Lett. **49**, 263 (2000).
[30] M. Asadi, N. Ebrahimi, G. G. Hamedani, and E. S. Soofi, J. Appl. Prob. **41**, 379 (2004).
[31] M. Asadi, N. Ebrahimi, E. S. Soofi, and S. Zarezadeh, Nav. Res. Log. **61**, 427 (2015).
[32] A. D. Crescenzo and M. Longobard, J. Appl. Prob. **39**, 434 (2002).
[33] X. Guan, R. Jha, and Y. Liu, J. Intell. Manufact. **23**, 163 (2012).
[34] X. Guan, R. Jha, and Y. Liu, Probab. Eng. Mech. **29**, 157 (2012).
[35] P. Rocchi and G. Capacci, Entropy **17**, 502 (2015).
[36] P. Rocchi, *Reliability is a New Science* (Springer International Publishing, Berlin, 2017).
[37] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[38] E. T. Jaynes, Phys. Rev. **106**, 620 (1956).
[39] E. T. Jaynes, Phys. Rev. **108**, 171 (1957).
[40] E. T. Jaynes, Proc. IEEE **70**, 939 (1982).
[41] See, for example, N. W. Ashcroft and N. David Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976), Chap. 26.
[42] See, for example, C. A. Fuchs and J. V. D. Graaf, IEEE Trans. Inf. Theory **45**, 1216 (1999).
[43] J. R. Carey, P. Liedo, and J. W. Vaupel, Experiment. Gerontol. **30**, 605 (1995).