Failure of complex systems, cascading disasters, and the onset of disease

A.J. Webster[1]

Nuffield Department of Population Health,

Richard Doll Building, University of Oxford,

Old Road Campus, Oxford, OX3 7LF, UK.

(Dated: November 22, 2018)

Abstract

Complex systems can fail through different routes, often progressing through a series of (rate-limiting) steps and modified by environmental exposures. The onset of disease, cancer in particular, is no different. A simple but very general mathematical framework is described for studying the failure of complex systems, or equivalently, the onset of disease. It includes the Armitage-Doll multi-stage cancer model as a particular case, and has potential to provide new insights into how diseases arise and progress. A method described by E.T. Jaynes is developed to provide an analytical solution for the models, and highlights connections between the convolution of Laplace transforms, sums of random samples, and Schwinger/Feynmann parameterisations. Examples include: exact solutions to the Armitage-Doll model, the sum of Gamma-distributed variables with integer-valued shape parameters, a clonal-growth cancer model, and a model for cascading disasters. The approach is sufficiently general to be used in many contexts, such as engineering, project management, disease progression, and disaster risk for example, allowing the estimation of failure rates in complex systems and projects. The intended result is a mathematical toolkit for the study of failure rates in complex systems and the onset of disease, cancer in particular.

I. INTRODUCTION

Complex systems such as a car can fail through many different routes, often requiring a sequence or combination of events for a component to fail. The same can be true for human disease, cancer in particular [1, 2]. For example, cancer can arise through a sequence of steps such as genetic mutations, each of which must occur prior to cancer [3–7]. In addition, there is considerable genetic variation within a single cancer [8], suggesting that similar cancers might arise through a variety of different paths.

Here a simple model is described for the failure of a system through one or more possible routes, each of which consisting of one or more sequential or non-sequential steps. The model is easy to conceptualise and derive, and many specific examples have analytical solutions or approximations, making it ideally suited to the construction of biologically- or physically-motivated models for the incidence of events such as diseases, disasters, or mechanical failures. A method described by E.T. Jaynes [9] generalises to give an exact analytical formula for the sums of random variables needed to evaluate the sequential model. This is evaluated for specific cases. Moolgavkar's exact solution [10] to the Armitage-Doll multistage cancer model is one example that is derived surprisingly easily, and is easily modified. More generally, the mathematical framework is intended to capture more complex cancer models, that might for example include a clonal expansion prior to cancer detection [4–6]. We also consider "cascading disasters" [11], where each disaster can substantially modify the risk of subsequent (possibly different) disasters.

Conventional notation is used [12], with: probability densities $f(t)$, cumulative probability distributions $F(t) = \int_0^t f(t)$, a survival function $S(t) = 1 - F(t)$, hazard function $h(t) = f(t)/S(t)$, and cumulative hazard function $H(t) = \int_0^t h(y)dy$. Noting that $f(t) = -dS/dt$, it is easily seen that $H(t) = \int_0^t f(y)/S(y)dy = -\log S(t)$, $h(t) = -d\log S(t)/dt$, and $S(t) = \exp(-\int_0^t h(y)dy)$.

II. FAILURE BY MULTIPLE POSSIBLE ROUTES

Imagine that we can enumerate all possible routes 1 to n by which a failure can occur. The probability of surviving the ith of these routes after time t is $S_i(t)$, and consequently

the probability of surviving all of these possible routes to failure $S(t)$ is,

$$S(t) = \Pi_{i=1}^{n} S_i(t) \tag{1}$$

or in terms of cumulative hazard functions with $S_i(t) = e^{-H_i(t)}$,

$$S(t) = \exp\left\{-\sum_{i=1}^{n} H_i(t)\right\} \tag{2}$$

The system's hazard rate for failure by any of the routes is,

$$\begin{aligned} h(t) &= -\tfrac{d}{dt}\log\left(S(t)\right) \\ &= -\textstyle\sum_{i=1}^{n} \tfrac{d}{dt}\log\left(S_i(t)\right) \\ &= \textstyle\sum_{i=1}^{n} h_i(t) \end{aligned} \tag{3}$$

and $H(t) = \sum_{i=1}^{n} H_i(t)$. In words, if failure can occur by any of n possible routes, the overall hazard of failure equals the sum of the hazard of failure by all the individual routes.

A few notes on Eq. 2 and its application to cancer modelling. Firstly, if the sth route to failure is much more likely than the others, with $H_s \gg H_j$ for $s \neq j$, then $S(t) = \exp\left\{-H_s(t) + \left(1 + O\left(\sum_{i\neq s} H_i/H_s\right)\right)\right\} \simeq \exp\left\{-H_s(t)\right\}$, which could represent the most likely sequence of mutations in a cancer model for example. Due to different genetic backgrounds, manufacturing processes, or exposures, this most probable route to failure could differ between individuals. Secondly, the stem cell cancer model assumes that cancer can occur through any of n_s equivalent stem cells in a tissue, for which Eq. 2 is modified to, $S = \exp\left\{-n_s \sum_{i=1}^{n} H_i(t)\right\}$. Thirdly, most cancers are sufficiently rare that $S \sim 1$. As a consequence, many cancer models (implicity or explicitly) assume $S \simeq 1 - n_s \sum_{i=1}^{n} H_i(t)$ and $f = -dS/dt \simeq n_s \sum_{i=1}^{n} h_i(t)$, a limit emphasised in the Appendix of Moolgavkar [10].

III. FAILURE REQUIRING m INDEPENDENT EVENTS

Often failure by a particular path will require more than one failure to occur independently. Consider firstly when there are m_i steps to failure, and the order of failure is unimportant. The probability of surviving failure by the ith route, $S_i(t)$ is,

$$\begin{aligned} S_i(t) &= P\,(\text{survive any one or more, necessary step for failure}) \\ &= 1 - P\,(\text{survive all the steps}) \\ &= 1 - \Pi_{j=1}^{m_i} F_{ij}(t) \end{aligned} \tag{4}$$

where $F_{ij}(t)$ is the cumulative probability distribution for failure of the jth step on the ith route within time t. Writing $S_{ij}(t) = 1 - F_{ij}(t)$, this can alternately be written as,

$$S_i(t) = 1 - \Pi_{j=1}^{m_i} \left(1 - S_{ij}(t)\right) \tag{5}$$

IV. RELATION TO RECENT MULTI-STAGE CANCER MODELS

It may be helpful to explain how Eqs. 1 and 4 are used in recently described multi-stage cancer models [13–15]. If we take a rate of mutations μ_j per cell division for each of the rate-limiting mutational steps 1 to j, and d_i divisions of cell i, then the probability of a stem cell surviving without the jth rate limiting mutation is $S_{ij} = (1 - \mu_j)^{d_i}$. Similarly, the probability of a given stem cell having mutation j is $F_{ij} = 1 - (1 - \mu_j)^{d_i}$. This is the solution of Zhang et al. [14] to the recursive formula of Wu et al. [13] (see Appendix of Zhang et al. [14] for details). Using Eq. 4, the survival of the ith stem cell is described by,

$$S_i = 1 - \Pi_{j=1}^{m_i} \left(1 - (1 - \mu_j)^{d_i}\right) \tag{6}$$

Now assuming all n stem cells are equivalent and have equal rates $\mu_i = \mu_j$ for all i, j, and consider only one path to cancer with m mutational steps, then,

$$S_i = 1 - \left(1 - (1 - \mu)^d\right)^m \tag{7}$$

$$\begin{aligned} S &= \Pi_{i=1}^{n} S_i \\ &= \left(1 - \left(1 - (1 - \mu)^d\right)^m\right)^n \end{aligned} \tag{8}$$

The probability of cancer within m divisions, often referred to as "theoretical lifetime intrinsic cancer risk", is,

$$F = 1 - \left(1 - \left(1 - (1 - \mu)^d\right)^m\right)^n \tag{9}$$

This is the equation derived by Calabrese and Shibata [15], and that Zhang found as the solution to the model of Wu et al [13, 14].

Therefore, in addition to the models of Wu and Calabrese being equivalent cancer models needing m mutational steps, the models also assume that the order of the steps is not important. This differs from the original Armitage-Doll model that considered a sequential set of rate-limiting steps, and was exactly solved by Moolgavkar [10]. Eqs. 8 and 9 are equivalent to assuming: (i) equivalent stem cells, (ii) a single path to cancer, (iii) equivalent divisions per stem cell, and, (iv) equivalent mutation rates for all steps.

Despite the differences in modelling assumptions for Eq. 9 and the Armitage-Doll model, their predictions can be quantitatively similar. To see this, use the Armitage-Doll approximation of $\mu d \ll 1$, to expand,

$$(1-\mu)^d = \exp\left(d\log(1-\mu)\right) \simeq \exp\left(\mu d\right) \tag{10}$$

If cell divisions are approximately uniform in time, then we can replace μd with μt, with μ now a rate per unit time. Then expanding $\exp(-\mu t) \simeq 1-\mu t$, gives,

$$F = 1 - \left(1 - \left(1-(1-\mu)^d\right)^m\right)^{n_s} \simeq 1 - \left(1-(\mu t)^m\right)^{n_s} \simeq n_s\left(\mu t\right)^m \tag{11}$$

The incidence rate $h = f/S$ is then $h \simeq n_s \mu^m t^{m-1}$, the same as the original (approximate) Armitage-Doll solution [1]. This approximate solution is expected to become inaccurate at sufficiently long times.

An equivalent expression to Eq. 8 was known to Armitage, Doll, and Pike since at least 1965 [16], as was its limiting behaviour for large n. The authors [16] emphasised that many different forms for the $F_i(t_i)$ could produce approximately the same observed $F(t)$, especially for large n, with the behaviour of $F(t)$ being dominated by the small t behaviour of $F_i(t)$. As a result, for sufficiently small times power-law behaviour for $F(t)$ is likely, and if longer times were observable then an extreme value distribution would be expected [3, 16, 17]. However the power-law approximation can fail for important cases with extra rate-limiting steps such as a clonal expansion [4–6]. It seems likely that a model that includes clonal expansion and cancer detection is needed for cancer modelling, but the power law approximation could be used for all but the penultimate step, for example. A general methodology that includes this approach is described in Section V next, and examples are given in Section VI. The results and examples of the next Sections V and VI are intended to have a broad range of potential applications.

V. FAILURE REQUIRING m SEQUENTIAL STEPS

Some failures require a *sequence* of independent events to occur, each following the one before. A well-known example is the Armitage-Doll multistage cancer model, that requires a sequence of m mutations (failures), that each occur with a different constant rate. The probability density for failure time is the pdf for a sum of the m independent times t_j to

failure at each step in the sequence, each of which may have a different probability density function $f_j(t_j)$. A general method for evaluating the probability density is outlined below, adapting a method described by Jaynes (page 569).

Firstly use marginalisation to write $f\left(t=\sum_{j=1}^{m} t_j\right)$ in terms of $f\left(t=\sum_{j=1}^{m} t_j, t_1, \ldots, t_m\right)$, where (A, B, C) is read as "A and B and C", and expand using the product rule $f(A, B) = f(A|B)f(B)$,

$$\begin{aligned} f\left(t=\sum_{j=1}^{m} t_j\right) &= \int_0^\infty dt_1 \ldots \int_0^\infty dt_m f\left(t=\sum_{j=1}^{m} t_j, t_1, \ldots, t_m\right) \\ &= \int_0^\infty dt_1 \ldots \int_0^\infty dt_m f\left(t=\sum_{j=1}^{m} t_j | t_1, \ldots, t_m\right) f\left(t_j, t_1, \ldots, t_m\right) \end{aligned} \tag{12}$$

Noting that $f\left(t=\sum_{j=1}^{m} t_j, |t_1, \ldots, t_m\right)$ is zero for $t \neq \sum_{j=1}^{m} t_j$ and $1 = \int_0^\infty dt f\left(t=\sum_{j=1}^{m} t_j, |t_1, \ldots, t_m\right)$, indicates that it is identical to a Dirac delta function $\delta\left(t-\sum_{j=1}^{m} t_j\right)$. For independent events $f(t_1, \ldots, t_m) = \Pi_{j=1}^{m} f_j(t_j)$, so we have,

$$f\left(t=\sum_{j=1}^{m} t_j\right) = \int_0^\infty dt_1 \ldots \int_0^\infty dt_m \Pi_{j=1}^{m} f_j\left(t_j\right) \delta\left(t-\sum_{j=1}^{m} t_j\right) \tag{13}$$

To evaluate the integrals, take the Laplace transform with respect to t, to give,

$$\mathcal{L}[f] = \int_0^\infty e^{-st} f(t) dt = \int_0^\infty dt_1 \ldots \int_0^\infty dt_m \Pi_{j=1}^{m} f_j\left(t_j\right) e^{-s(t_1+\cdots+t_m)} \tag{14}$$

This factorises as,

$$\mathcal{L}[f] = \Pi_{j=1}^{m} \int_0^\infty dt_j f_j(t_j) e^{-st_j} \tag{15}$$

Giving a general analytical solution as,

$$f(t) = \mathcal{L}^{-1}\left\{\Pi_{j=1}^{m} \mathcal{L}\left[f_j(t_j)\right]\right\} \tag{16}$$

where $\mathcal{L}^{-1}$ is the inverse Laplace transform. Eq. 15 is similar to the relationship between moment generating functions $M_i(s) = \sum_{t_i=0}^{\infty} e^{st_i} p_i(t_i)$ of discrete probability distributions $p_i(t_i)$, and the moment generating function $M(s)$ for $t = \sum_{i=1}^{m} t_i$, that has,

$$M(s) = \Pi_{i=1}^{m} M_i(s) \tag{17}$$

whose derivation is analogous to Eq. 17 but with integrals replaced by sums. The survival and hazard functions for $f(t)$ can be obtained from Eq. 16 in the usual way. For example,

$$\begin{aligned} S_i(t) &= \int_t^\infty f_i(y) dy \\ &= \int_t^\infty \mathcal{L}^{-1}\left\{\Pi_{j=1}^{m_i} \mathcal{L}\left[f_{ij}(t_{ij})\right]\right\} dy \end{aligned} \tag{18}$$

that can be used in combination with Eq. 1. A number of valuable results are easy to evaluate using Eq. 16, as is illustrated in Section VI.

A useful related result is,

$$f\left(t=\sum_{j=1}^{n} t_j\right)=\mathcal{L}^{-1}\left\{\mathcal{L}\left[f\left(\sum_{j=1}^{n-1} t_j\right)\right]\mathcal{L}\left[f_n(t_n)\right]\right\} \tag{19}$$

that can be inferred from Eq. 16 with $m=2$,

$$f\left(t=t_1+t_2\right)=\mathcal{L}^{-1}\left\{\mathcal{L}\left[f_1(t_1)\right]\mathcal{L}\left[f_2(t_2)\right]\right\} \tag{20}$$

by replacing $f_1(t_1)$ with $f(\sum_{j=1}^{n-1} t_j)$ and $f_2(t_2)$ with $f_n(t_n)$. Eq. 20 can be solved using the convolution theorem for Laplace transforms, that gives,

$$f\left(t=t_1+t_2\right)=\int_0^t f_1(\tau)f_2(t-\tau)d\tau \tag{21}$$

which is sometimes easier to evaluate than two Laplace transforms and their inverse. In general, solutions can be presented in terms of multiple convolutions if it is preferable to do so. Eqs. 19 and 21 are particularly useful for combining a known solution for the sum of $(n-1)$ samples such as for cancer initiation, with a differently distributed nth sample, such as the waiting time to detect a growing cancer. A final remark applies to the sum of random variables whose domain extends from $-\infty$ to ∞, as opposed to the range 0 to ∞ considered so far. In that case an analogous calculation using a Fourier transform with respect to t in Eq. 13 leads to analogous results in terms of Fourier transforms, with $\mathcal{F}[f_i(t_i)]=\int_{-\infty}^{\infty} f_i(t_i)e^{imt_i}dt_i$ in place of Laplace transforms, resulting in,

$$f(t)=\mathcal{F}^{-1}\left\{\Pi_{j=1}^{m}\mathcal{F}\left[f_j(t_j)\right]\right\} \tag{22}$$

Eq. 22 is mentioned for completeness, but is not used here.

A general solution to Eq. 16 can be given in terms of definite integrals, with,

$$\begin{aligned} f(t) &= \mathcal{L}^{-1}\left\{\Pi_{j=1}^{m}\mathcal{L}\left[f_j(t_j)\right]\right\} \\ &= t^{m-1}\int_0^1 dy_1\ldots\int_0^1 dy_{m-1}y_1^0y_2^1\ldots y_{m-1}^{n_m-1} \\ & f_1(ty_1\ldots y_{m-1})f_2(t(1-y_1)y_2\ldots y_{m-1})f_3(t(1-y_2)y_3\ldots y_{m-1})\ldots \\ & f_{m-1}(t(1-y_{m-2})y_{m-1})f_m(t(1-y_{m-1})) \end{aligned} \tag{23}$$

This can sometimes be easier to evaluate or approximate than Eq. 16. A derivation is given in Appendix A1. Eq. 23 allows a generalised Schwinger/Feynmann parameterisation [18] to

be derived. Writing $g_j(s) = \mathcal{L}\left[f_j(t_j)\right]$ and taking the Laplace transform of both sides of Eq. 23, gives,

$$\begin{aligned}\Pi_{j=1}^{m} g_j(s) &= \int_0^1 dy_1 \ldots \int_0^1 dy_{m-1} y_1^0 y_2^1 \ldots y_{m-1}^{n_m-1} \\ &\mathcal{L}\left[t^{m-1}\mathcal{L}^{-1}\left\{g_1(s)\right\}(ty_1\ldots y_{m-1})\mathcal{L}^{-1}\left\{g_2(s)\right\}(t(1-y_1)y_2\ldots y_{m-1})\ldots\right. \\ &\left.\mathcal{L}^{-1}\left\{g_m(s)\right\}(t(1-y_{m-1}))\right]\end{aligned} \tag{24}$$

which includes some well known Schwinger/Feynmann parameterisations as special cases. This is discussed further in Appendix A1.

VI. MODELLING SEQUENTIAL EVENTS - EXAMPLES

In the following examples we consider the time $t = \sum_{i=1}^{m} t_i$ for a sequence of events, with possibly different distributions $f_i(t_i)$ for the time between the $(i-1)$th and ith event. Some of the results are well-known but not usually presented this way, others are new or poorly known. We will use the Laplace transforms (and their inverses), of, $\mathcal{L}^{-1}\mathcal{L}[t^p] = \mathcal{L}^{-1}[\Gamma(p+1)/s^{p+1}] = t^p$, and $\mathcal{L}^{-1}\mathcal{L}[t^p e^{-\mu t}] = \mathcal{L}^{-1}[\Gamma(p+1)/(s+\mu)^{p+1}] = t^p e^{-\mu t}$.

Sums of gamma distributed samples (equal rates): A sum of gamma distributed variables with equal rate parameters can be calculated from,

$$\begin{aligned}\mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\mathcal{L}\left[\mu^{1+p_i}\tfrac{t^{p_i}e^{-\mu t}}{\Gamma(p_i+1)]}\right]\right\} &= \mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\tfrac{\mu^{1+p_i}}{(s+\mu)^{1+p_i}}\right\} \\ &= \mathcal{L}^{-1}\left\{\tfrac{\mu^{m+\sum_{i=1}^{m}p_i}}{(s+\mu)^{m+\sum_{i=1}^{m}p_i}}\right\} \\ &= \mu^{m+\sum_{i=1}^{m}p_i}\tfrac{t^{m-1+\sum_{i=1}^{m}p_i}e^{-\mu t}}{\Gamma\left(m+\sum_{i=1}^{m}p_i\right)}\end{aligned} \tag{25}$$

For a sum of exponentially distributed variables with $\{p_i = 0\}$, then this simplifies to $\mu^m t^{m-1}e^{-\mu t}/\Gamma(m)$, a Gamma distribution.

Power law approximations: For many situations such as most diseases, you are unlikely to get a particular disease during your lifetime. In those cases the probability of survival over a lifetime is close to 1, and the probability density function $f = h/S$, can be approximated by $f \simeq h$, that in turn can often be approximated by a power of time with $f \simeq h \simeq \mu_i t^{p_i}$. Then we have,

$$\begin{aligned}\mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\mathcal{L}\left[\mu_i t^{p_i}\right]\right\} &= \mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\tfrac{\mu_i\Gamma(1+p_i)}{s^{1+p_i}}\right\} \\ &= \Pi_{i=1}^{m}\left(\mu_i\Gamma(1+p_i)\right)\mathcal{L}^{-1}\left\{\tfrac{1}{s^{m+\sum_{i=1}^{m}p_i}}\right\} \\ &= \Pi_{i=1}^{m}\left(\mu_i\Gamma(1+p_i)\right)\tfrac{t^{-1+m+\sum_{i=1}^{m}p_i}}{\Gamma(m+\sum_{i=1}^{m}p_i)}\end{aligned} \tag{26}$$

The Armitage-Doll model: A well known example of this approximation Eq. 26, is (implicitly) in the original approximate solution to the Armitage-Doll multi-stage cancer model. Taking a constant hazard at each step, and approximating $f_i \simeq h_i = \mu_i$, then Eq. 26 gives,

$$\mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\mathcal{L}\left[\mu_i\right]\right\} = \left[\Pi_{i=1}^{m}\mu_i\right]\frac{t^{m-1}}{\Gamma(m)} \tag{27}$$

as was used in the original Armitage-Doll paper. Note that an equivalent distribution can be produced by a different combination of (fewer) hazard functions with $h_i \sim t_i^{p_i}$ provided $t^{\sum_i(p_i+1)} = t^{m-1}$. This flexibility is removed in Moolgavkar's exact solution to the model [10], that is described next.

Moolgavkar's exact solution to the Armitage-Doll model: Moolgavkar's exact solution to the Armitage-Doll model is the solution of,

$$\mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\mathcal{L}\left[\mu_i e^{-\mu_i t}\right]\right\} = \mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\frac{\mu_i}{s+\mu_i}\right] \tag{28}$$

For example, if $n = 3$ then,

$$\mathcal{L}^{-1}\left\{\Pi_{i=1}^{3}\mathcal{L}\left[\mu_i e^{-\mu_i t}\right]\right\} = \mu_1\mu_2\mu_3\mathcal{L}^{-1}\left\{\frac{1}{(s+\mu_1)}\frac{1}{(s+\mu_2)}\frac{1}{(s+\mu_3)}\right\} \tag{29}$$

Using partial fractions, we can write,

$$\frac{1}{s+\mu_1}\frac{1}{s+\mu_2}\frac{1}{s+\mu_3} = \frac{1}{s+\mu_1}\frac{1}{(\mu_1-\mu_2)(\mu_1-\mu_3)} + \frac{1}{s+\mu_2}\frac{1}{(\mu_2-\mu_1)(\mu_2-\mu_3)} + \frac{1}{s+\mu_3}\frac{1}{(\mu_3-\mu_1)(\mu_3-\mu_2)} \tag{30}$$

Allowing the inverse Laplace transforms to be easily evaluated, giving,

$$\begin{aligned}&\mathcal{L}^{-1}\left\{\Pi_{i=1}^{3}\mathcal{L}\left[\mu_i e^{-\mu_i t}\right]\right\}\\ &= \mu_1\mu_2\mu_3\left[\frac{e^{-\mu_1 t}}{(\mu_1-\mu_2)(\mu_1-\mu_3)} + \frac{e^{-\mu_2 t}}{(\mu_2-\mu_1)(\mu_2-\mu_3)} + \frac{e^{-\mu_3 t}}{(\mu_3-\mu_1)(\mu_3-\mu_2)}\right]\end{aligned} \tag{31}$$

Note that the result is independent of the order of sequential events, but unlike the approximate solution to the Armitage Doll model [1], the exact solution allows less variability in the underlying models that can produce it. Also note that the leading order terms of an expansion in t cancel exactly, to give identical leading-order behaviour as for a power-law approximation (with $p = 0$)

A general solution can be formed using a Schwinger/Feynmann parameterisation [18] of,

$$\Pi_{i=1}^{m}\frac{1}{\mu_i} = \Gamma(m)\int_0^1 dy_1\int_0^{y_1} dy_2 \ldots \int_0^{y_{m-2}} dy_{m-1}\frac{1}{(\mu_1 y_{m-1}+\mu_2(y_{m-2}-y_{m-1})+\cdots+\mu_m(1-y_1))^m} \tag{32}$$

Replacing μ_i with $s+\mu_i$ in Eq. 32, then we can write Eq. 28 as,

$$\begin{aligned}
&\mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\tfrac{\mu_i}{s+\mu_i}\right\}\\
&=\left(\Pi_{i=1}^{m}\mu_i\right)\Gamma(m)\int_0^1 dy_1\int_0^{y_1}dy_2\ldots\int_0^{y_{m-2}}dy_{m-1}\mathcal{L}^{-1}\left\{\tfrac{1}{(s+\mu_1 y_{m-1}+\mu_2(y_{m-2}-y_{m-1})+\cdots+\mu_m(1-y_1))^m}\right\}\\
&=\left(\Pi_{i=1}^{m}\mu_i\right)t^{m-1}\int_0^1 dy_1\int_0^{y_1}dy_2\ldots\int_0^{y_{m-2}}dy_{m-1}e^{-(\mu_1 y_{m-1}+\mu_2(y_{m-2}-y_{m-1})+\cdots+\mu_m(1-y_1))t}
\end{aligned} \tag{33}$$

(which is simpler, but equivalent in effect, to repeatedly using the convolution formula). Completing the integrals will generate Moolgavkar's solution for a given value of m. For example, taking $m=3$ and integrating once gives,

$$\mathcal{L}^{-1}\left\{\Pi_{i=1}^{3}\frac{\mu_i}{s+\mu_i}\right\}=\frac{te^{-\mu_3 t}}{(\mu_2-\mu_1)}\int_0^1 dx_1\left(e^{-x_1 t(\mu_1-\mu_3)}-e^{-x_1 t(\mu_2-\mu_3)}\right) \tag{34}$$

Integrating a second time, and simplifying, gives Eq. 31. The relationship between Schwinger/Feynmann parameterisations, Laplace transforms, and the convolution theorem is discussed further in Appendix A.

Moolgavkar [10] used induction to provide an explicit formula for $f(t)$, with,

$$f(t)=\Pi_{i=1}^{m}\mu_i\sum_{i=1}^{m}\chi_i(m)e^{-\mu_i t} \tag{35}$$

where,

$$\chi_i(m)=\frac{1}{(\mu_1-\mu_i)(\mu_2-\mu_i)\ldots(\mu_{i-1}-\mu_i)(\mu_{i+1}-\mu_i)\ldots(\mu_m-\mu_i)} \tag{36}$$

For small times the terms in a Taylor expansion of Eq. 35 cancel exactly, so that $f(t)\simeq\left(\Pi_{i=1}^{m}\mu_i\right)t^{m-1}$, as expected. This feature could be useful for approximating a normalised function when the early-time behaviour approximates a given integer power of time. Further uses of Moolgavkar's solution are discussed next.

Sums of gamma distributed samples (different rates) A useful mathematical result can be found by combining the Laplace transform of Moolgavkar's solution Eq. 35 for $f\left(t=\sum_{i=1}^{m}t_i\right)$ with Eq. 28, to give,

$$\Pi_{i=1}^{m}\frac{1}{s+\mu_i}=\sum_{i=1}^{m}\frac{\chi_i(m)}{s+\mu_i} \tag{37}$$

For example, if the shape parameters p_i of Gamma distributed samples are integer-valued, then we can write,

$$\mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\frac{1}{(s+\mu_i)^{p_i}}\right\}=\mathcal{L}^{-1}\left\{\Pi_{j=1}^{m}(-1)^{p_j-1}\frac{\partial^{p_j-1}}{\partial\mu_j^{p_j-1}}\Pi_{i=1}^{m}\frac{1}{(s+\mu_i)}\right\} \tag{38}$$

where the product of differential operators can be taken outside the product of Laplace transforms because $\partial/\partial\mu_i(1/(s+\mu_j))$ is zero for $i \neq j$. Using Eq. 37 we can replace the product of Laplace transforms with a sum, then only keep derivatives of non-zero terms, giving,

$$\mathcal{L}^{-1}\left\{\Pi_{i=1}^{m}\frac{1}{(s+\mu_i)^{p_i}}\right\} = \mathcal{L}^{-1}\left\{\sum_{i=1}^{m}(-1)^{p_i-1}\frac{\partial^{p_i-1}}{\partial\mu_i^{p_i-1}}\frac{\chi_i(m)}{(s+\mu_i)}\right\} \tag{39}$$

Finally, take the inverse Laplace transform $\mathcal{L}^{-1}[1/(s+\mu_i)] = e^{-\mu_i t}$, to give,

$$f(t) = \left(\Pi_{i=1}^{m}\frac{\mu_i^{p_i}}{\Gamma(p_i)}\right)\sum_{i=1}^{m}(-1)^{p_i-1}\frac{\partial^{p_i-1}}{\partial\mu_i^{p_i-1}}\left(\chi_i(m)e^{-\mu_i t}\right) \tag{40}$$

a general solution for sums of Gamma distributed samples with integer-valued shape parameters p_i (and arbitrary rate parameters μ_i).

For example, if $p_i = 2$ for all i, then evaluation of Eq. 40 can be accomplished by considering the ith term, and noting that for $j \neq i$ there is only one factor involving μ_j in χ_i (Eq. 36). All the derivatives with respect to $j \neq i$ then make $\chi_i e^{-\mu_i t}$ become $(-1)^{m-1}\chi_i(m)^2 e^{-\mu_i t}$. Taking the final derivative with respect to μ_i and cancelling factors of (-1) gives,

$$f(t) = \left(\Pi_{i=1}^{m}\mu_i^2\right)\sum_{i=1}^{m}\chi_i(m)^2 e^{-\mu_i t}\left[t - 2\sum_{j\neq i}\frac{1}{(\mu_j-\mu_i)}\right] \tag{41}$$

for the sum of Gamma distributions with shape parameters $p_i = 2$ and arbitrary rate parameters, and $\chi_i(m)$ as defined in Eq. 36.

Sums of samples with different distributions: An advantage of the method described above, is that it is often easy to calculate pdfs for sums of differently distributed samples. For the first example, consider two samples from the same (or very similar) exponential distribution, and a third from a different exponential distribution. The result can be obtained by writing $\mu_3 = \mu_2 + \epsilon$ in Eq. 31, and letting $\epsilon \to 0$. Considering the terms involving exponents of μ_2 and μ_3,

$$\begin{aligned}\tfrac{e^{-\mu_2 t}}{(\mu_2-\mu_1)(\mu_2-\mu_3)} + \tfrac{e^{-\mu_3 t}}{(\mu_3-\mu_1)(\mu_3-\mu_2)} &= \tfrac{e^{-\mu_2 t}}{\epsilon}\left(-\tfrac{1}{\mu_2-\mu_1} + \tfrac{e^{-\epsilon t}}{\mu_2-\mu_1+\epsilon}\right)\\ &= \tfrac{e^{-\mu_2 t}}{\mu_2-\mu_1}\left(-1 + \left(1 - \epsilon t + O\left(\epsilon^2\right)\right)\right)\\ &= -\tfrac{e^{-\mu_2 t}}{\mu_2-\mu_1}\left(1 + O\left(\epsilon\right)\right)\end{aligned} \tag{42}$$

Using Eq. 31 and letting $\epsilon \to 0$, gives,

$$\begin{aligned}&\mu_1\mu_2\mu_3\left[\tfrac{e^{-\mu_1 t}}{(\mu_1-\mu_2)(\mu_1-\mu_3)} + \tfrac{e^{-\mu_2 t}}{(\mu_2-\mu_1)(\mu_2-\mu_3)} + \tfrac{e^{-\mu_3 t}}{(\mu_3-\mu_1)(\mu_3-\mu_2)}\right]\\ &\qquad\to \mu_1{\mu_2}^2\left[\tfrac{e^{-\mu_1 t}}{(\mu_1-\mu_2)^2} - \tfrac{te^{-\mu_2 t}}{(\mu_2-\mu_1)}\right]\end{aligned} \tag{43}$$

More generally, this argument shows that a sum of exponentially distributed samples with different rates, smoothly approximates a gamma distribution as the rates become increasingly similar. As shown in Eq. 25, the sum of exponentially distributed samples with exactly the same rate, has a gamma distribution

Failure involving a combination of sequential and non-sequential steps: If a path to failure involves a combination of sequential and non-sequential steps, then the necessary set of sequential steps can be considered as one of the non-sequential steps, with overall survival given by Eq. 1 and the survival for any sequential set of steps calculated from Eq. 18.

Clonal-expansion cancer models: Some cancer models have a clonal expansion of cells as a rate-limiting step prior to a cancer's detection [4–6]. For example, Michor et al. [5] modelled clonal expansion of myeloid leukemia as logistic growth, with the likelihood of cancer detection (the hazard function), being proportional to the number of cancer cells. This gives a survival function for cancer detection of,

$$S_i(t) = \exp\left(-a\int_0^t x(y)dy\right) \tag{44}$$

where,

$$x(t) = \frac{1}{1+(N-1)e^{-ct}} \tag{45}$$

a, c, are rate constants, and N is the total number of cells prior to cancer initiation. Noting that $\int_0^t x(y)dy = \log(e^{ct}+(N-1))/c \to t$, as $t \to \infty$ and $x(t) \to 1$, then the tail of the survival curve falls exponentially towards zero with time.

Alternatively, we might expect the likelihood of cancer being diagnosed to continue to increase with time since the cancer is initiated. For example, a hazard function that is linear in time would give a Weibull distribution with $S(t) = e^{-at^2}$. It is unlikely that either this or the logistic model would be an equally good description for the detection of all cancers, although they may both be an improvement on a model without either. Both models have a single peak, but differ in the tail of their distribution, that falls as $\sim e^{-act}$ for the logistic model and $\sim e^{-at^2}$ for the Weibull model. Qualitatively, we might expect a delay between cancer initiation and the possibility of diagnosis, and diagnosis to occur almost inevitably within a reasonable time-period. Therefore a Weibull or Gamma distributed time to diagnosis seems reasonable for many cancers, with the shorter tail of the Weibull distribution making it more suitable approximation for cancers whose diagnosis is almost

inevitable. (The possibility of misdiagnosis or death by another cause is not considered here.)

For example, noting that Moolgavkar's solution is a linear combination of exponential distributions, to combine it with a Weibull distribution for cancer detection $f_1(t_1) = -d/dt_1(e^{-bt_1^2/2})$, we can consider a single exponential term at a time. Taking $f_2(t_2) = ae^{-at_2}$, and using the convolution formula Eq. 21, we get,

$$\begin{aligned} f(t = t_1 + t_2) &= \mathcal{L}^{-1}\left\{L[f_1(t_1)]L[f_2(t_2)]\right\} \\ &= a\int_0^t e^{-a(t-y)}\left(-\tfrac{d}{dy}e^{-by^2/2}\right)dy \\ &= a\left(e^{-at} - e^{-bt^2/2}\right) + a^2e^{-at}e^{a^2/2b}\int_0^t e^{-\frac{b}{2}\left(y-\frac{a}{b}\right)^2}dy \end{aligned} \tag{46}$$

where we integrated by parts to get the last line. This may be written as,

$$\begin{aligned} f(t) &= a\left(e^{-at} - e^{-bt^2/2}\right) \\ &+a^2e^{-at}e^{a^2/2b}\sqrt{\tfrac{\pi}{2b}}\ \mathrm{erf}\left(\sqrt{\tfrac{b}{2}}\tfrac{a}{b}\right) \\ &+a^2e^{-at}e^{a^2/2b}\sqrt{\tfrac{\pi}{2b}}\begin{cases} -\mathrm{erf}\left(\sqrt{\frac{b}{2}}\left(\frac{a}{b}-t\right)\right) & t < \frac{a}{b} \\ +\mathrm{erf}\left(\sqrt{\frac{b}{2}}\left(t-\frac{a}{b}\right)\right) & t \geq \frac{a}{b} \end{cases} \end{aligned} \tag{47}$$

with $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-z^2}dz$. Similarly for a Gamma distribution with $f_1 = b^pt^{p-1}e^{-bt}/\Gamma(p)$ and an exponential, $f_2(t_2) = ae^{-at}$, then assuming $b > a$,

$$\begin{aligned} f(t) &= \tfrac{b^pa}{\Gamma(p)}\int_0^t y^{p-1}e^{-by}e^{-a(t-y)}dy \\ &= b^pa\tfrac{e^{-at}}{(b-a)^p}\tfrac{1}{\Gamma(p)}\int_0^{t(b-a)} u^{p-1}e^{-u}du \\ &= b^pa\tfrac{e^{-at}}{(b-a)^p}\gamma(p, t(b-a)) \end{aligned} \tag{48}$$

where $\gamma(p, t(b-a))$ is the normalised lower incomplete Gamma function, which is available in most computational mathematics and statistics packages. If $a > b$ then f_1 and f_2 must be exchanged and the result is most easily evaluated numerically.

VII. CASCADING FAILURES WITH DEPENDENT SEQUENCES OF EVENTS

Before concluding, we briefly consider how we can model non-independent failures, where the failure of A changes the probability of a failure in B or C. Consider a general situation where m parts must fail for the overall component to fail, but where the failure of each part influences the failure rates of the other parts. Important humanitarian examples are "cascading hazards" [11], where a disaster such as wildfire makes another disaster such

as a mudslide become a realistic possibility. This situation can be considered as a set of $m!$ independent routes to failure, where the probability of a given component failing can be different for each route, and depends on both the component and the particular prior sequence of failures/survivals. For example, if there are three steps A,B,C, then we evaluate the probability of each of the 6 possible routes in turn, and obtain the overall failure probability from Eq. 1. Writing the 6 routes as, 1=ABC, 2=ACB, 3=BAC, 4=BCA, 5=CAB, 6=CBA, and reading e.g. ABC as "A, then B, then C", the survival probability is,

$$S(t) = \Pi_{i=1}^{6} S_i(t) \tag{49}$$

For failure by a particular route ABC we need the probability for the sequence of events, $A\&(\overline{B\&C})$, then $(B\&\bar{C})|A$, then $C|(AB)$. We can calculate this using Eq. 16, for example giving,

$$f_{ABC}(t) = \mathcal{L}^{-1}\left\{\mathcal{L}\left[f_{A\&(\overline{B\&C})}(t_1)\right]\mathcal{L}\left[f_{(B\&\bar{C})|A}(t_2)\right]\mathcal{L}\left[f_{C|(AB)}(t_3)\right]\right\} \tag{50}$$

from which we can construct $S_1(t) = \int_t^\infty f_{ABC}(y)dy$.

Although in principle every term in e.g. Eqs. 49 and 50 need evaluating, there will be situations where results simplify. For example, if one route is much more probable than another - e.g. if it is approximately true that landslides only occur after deforestation, that may be due to fire, then we only need to evaluate the probability distribution for that route. As another example, if all the f_i are exponentially distributed with different rates, then f_{ABC} will be described by Moolgavkar's solution. A more striking example is when there are very many potential routes to failure, as for the Armitage-Doll model where there are numerous stem cells that can cause cancer. In those cases, if the overall failure rate remains low, then the $f_i(t)$ in Eq. 50 must all be small with $S \simeq 1$ and $f \simeq h$, and can often be approximated by power laws. For that situation we have a general result that f_i, F_i, and H_i will be a powers of time, and Eq. 2 gives,

$$S(t) \simeq \exp\left\{-\sum_{i=1}^{n} a_i t^{p_i}\right\} \tag{51}$$

for some $a_i > 0$ and $p_i > 0$. Then $F(t) = 1 - S(t)$, $f(t) = -dS/dt$, and $h(t) \simeq f(t)$, can be approximated by a sum of power series in time. If one route is much more likely than the others then both $f(t)$ and $h(t)$ can be approximated as a single power of time, with the approximation best at early times, and a cross-over to different power-law behaviour at later times.

VIII. SUMMARY

The purpose of this article is to provide a simple mathematical framework to describe existing multi-stage cancer models, that is easily adaptable to model events such as failure of complex systems, cascading disasters, and the onset of disease. The key formulae are given in Eqs. 1, 4, and 16 or equivalently 18. A selection of analytical solutions are given in Sections VI and VII. These examples can be combined in numerous ways to construct a wide range of models. Together the formulae are intended to provide a comprehensive toolkit for developing conceptual models of failure, disaster, and disease, allowing quantitative predictions to be made, and tested against data.

Acknowledgements:

Part of this work was completed with support from Cancer Research UK (grant no. C570/A16491).

---

[1] P. Armitage and R. Doll, "The age distribution of cancer and a multi-stage theory of carcinogenesis," *British Journal of Cancer*, vol. 8, no. 1, pp. 1–12, 1954.

[2] R. Peto, "Epidemiology, multistage models, and short-term mutagenicity tests 1," *International Journal of Epidemiology*, vol. 45, no. 3, pp. 621–637, 2016.

[3] S. H. Moolgavkar, "Commentary: Multistage carcinogenesis and epidemiological studies of cancer," *International Journal of Epidemiology*, vol. 45, no. 3, pp. 645–649, 2016.

[4] E. G. Luebeck and S. H. Moolgavkar, "Multistage carcinogenesis and the incidence of colorectal cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 23, pp. 15095–15100, 2002.

[5] F. Michor, Y. Iwasa, and M. A. Nowak, "The age incidence of chronic myeloid leukemia can be explained by a one-mutation model," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 40, pp. 14931–14934, 2006.

[6] R. Meza, J. Jeon, S. H. Moolgavkar, and E. G. Luebeck, "Age-specific incidence of cancer: Phases, transitions, and biological implications," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 42, pp. 16284–16289, 2008.

[7] M. P. Little, "Cancer models, genomic instability and somatic cellular darwinian evolution,"

Biology Direct, vol. 5, p. 19, 2010.

[8] B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. B. Zhou, L. A. Diaz, and K. W. Kinzler, "Cancer genome landscapes," *Science*, vol. 339, no. 6127, pp. 1546–1558, 2013.

[9] E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.

[10] S. H. Moolgavkar, "Multistage theory of carcinogenesis and age distribution of cancer in man," *Journal of the National Cancer Institute*, vol. 61, no. 1, pp. 49–52, 1978.

[11] A. AghaKouchak, L. S. Huning, O. Mazdiyasni, I. Mallakpour, F. Chiang, M. Sadegh, F. Vahedifard, and H. Moftakhari, "How do natural hazards cascade to cause disasters?," *Nature*, vol. 561, no. 7724, pp. 458–460, 2018.

[12] D. Collett, *Modelling Survival Data in Medical Research*, vol. Third Edition. CRC Press, 2015.

[13] S. Wu, S. Powers, W. Zhu, and Y. A. Hannun, "Substantial contribution of extrinsic risk factors to cancer development," *Nature*, vol. 529, no. 7584, pp. 43–47, 2016.

[14] X. X. Zhang, H. Frohlich, D. Grigoriev, S. Vakulenko, J. Zimmermann, and A. G. Weber, "A simple 3-parameter model for cancer incidences," *Scientific Reports*, vol. 8, 2018.

[15] P. Calabrese and D. Shibata, "A simple algebraic cancer equation: calculating how cancers may arise with normal mutation rates," *BMC Cancer*, vol. 10, 2010.

[16] P. Armitage, R. Doll, and M. C. Pike, "Somatic mutation," *British Medical Journal*, vol. 1, no. 5436, pp. 723–, 1965.

[17] L. Soto-Ortiz and J. P. Brody, "A theory of the cancer age-specific incidence data based on extreme value distributions," *Aip Advances*, vol. 2, no. 1, 2012.

[18] S. Weinberg, *The Quantum Theory of Fields*, vol. 1. Cambridge University Press, 1995.

Appendix A: Derivation of Eq. 23

The solution of Eq. 16 can be written in terms of multiple definite integrals that are sometimes easier to evaluate or approximate than directly evaluating Eq. 16. It is equivalent to expressing the solution as multiple convolutions using Eq. 21, and changing variables appropriately. The equation is obtained by Taylor expanding all functions before taking their Laplace transform, inverting the Laplace transform of the product of all terms (which is easy to do for the powers of time that appear in a Taylor expansion), then using a product

of Beta functions to factorise and re-sum the resulting expression. In mathematical notation,

$$\begin{aligned}\mathcal{L}^{-1}\left\{\Pi_{j=1}^{m}\mathcal{L}\left[f_j(t_j)\right]\right\} &= \mathcal{L}^{-1}\left\{\Pi_{j=1}^{m}\mathcal{L}\left[\sum_{n_j=0}^{\infty} f_{j,n_j}\frac{t^{n_j}}{n_j!}\right]\right\} \\ &= \mathcal{L}^{-1}\left\{\sum_{n_1=0}^{\infty}\cdots\sum_{n_m=0}^{\infty}\frac{f_{1,n_1}}{s^{n_1+1}}\cdots\frac{f_{m,n_m}}{s^{n_m+1}}\right\} \\ &= \sum_{n_1=0}^{\infty}\cdots\sum_{n_m=0}^{\infty} f_{1,n_1}\cdots f_{m,n_m}\frac{t^{-1+\sum_{i=1}^{m}(n_i+1)}}{\Gamma(\sum_{i=1}^{m}(n_i+1))}\end{aligned} \tag{A1}$$

where $f_{i,n_j} = \partial^{n_j} f_i(t_i)/\partial t^{n_j}|_{t_i=0}$. Now noting that the Beta function has,

$$\int_0^1 u^{m-1}(1-u)^{n-1}du = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} \tag{A2}$$

we can write,

$$\frac{1}{\Gamma\left(\sum_{i=1}^{m}(n_i+1)\right)} = \frac{1}{\Gamma\left(\sum_{i=1}^{m-1}(n_i+1)\right)}\frac{1}{\Gamma\left(n_m+1\right)}\int_0^1 y_{m-1}^{-1+\sum_{i=1}^{m-1}(n_i+1)}(1-y_{m-1})^{n_m} \tag{A3}$$

Repeatedly using this gives,

$$\begin{aligned}\frac{1}{\Gamma\left(\sum_{i=1}^{m}(n_i+1)\right)} &= \left(\Pi_{i=1}^{m}\frac{1}{\Gamma(n_i+1)}\right)\int_0^1 dy_1\ldots\int_0^1 dy_{m-1}y_1^0y_2^1y_3^2\ldots y_{m-1}^{m-2}\times \\ &(1-y_1)^{n_2}\ldots(1-y_{m-1})^{n_m}\times(y_1\ldots y_{m-1})^{n_1}(y_2\ldots y_{m-1})^{n_2}\ldots y_{m-1}^{n_m-1}\end{aligned} \tag{A4}$$

Using Eq. A4 to replace $1/\Gamma\left(\sum_{i=1}^{m}(n_i+1)\right)$ in Eq. A1, and grouping terms,

$$\begin{aligned}\mathcal{L}^{-1}\left\{\Pi_{j=1}^{m}\mathcal{L}\left[f_j(t_j)\right]\right\} &= t^{-1+m}\int_0^1 dy_1\ldots\int_0^1 dy_{m-1}y_1^0y_2^1y_3^2\ldots y_{m-1}^{m-2}\times \\ &\left(\sum_{n_1=0}^{\infty} f_{1,n_1}\frac{t^{n_1}(y_1\ldots y_{m-1})^{n_1}}{\Gamma(n_1+1)}\right) \\ &\left(\sum_{n_1=0}^{\infty} f_{2,n_2}\frac{t^{n_2}(1-y_1)^{n_2}(y_2\ldots y_{m-1})^{n_2}}{\Gamma(n_2+1)}\right) \\ &\ldots \\ &\left(\sum_{n_m=0}^{\infty} f_{m,n_m}\frac{t^{n_m}(1-y_{m-1})^{n_m}}{\Gamma(n_m+1)}\right)\end{aligned} \tag{A5}$$

The m Taylor series can now be resummed to give,

$$\begin{gathered}\mathcal{L}^{-1}\left\{\Pi_{j=1}^{m}\mathcal{L}\left[f_j(t_j)\right]\right\} = t^{-1+m}\int_0^1 dy_1\ldots\int_0^1 dy_{m-1}y_1^0y_2^1y_3^2\ldots y_{m-1}^{m-2}\times \\ f_1\left(ty_1\ldots y_{m-1}\right)f_2\left(t(1-y_1)(y_2\ldots y_{m-1})\right)f_3\left(t(1-y_2)(y_3\ldots y_{m-1})\right)\ldots f_m\left(t(1-y_{m-1})\right)\end{gathered} \tag{A6}$$

For example, taking $m = 2$ gives,

$$f(t) = t\int_0^1 dy_1 f_1(ty_1)f_2(t(1-y_1)) \tag{A7}$$

as we could have got from the convolution formula after a simple change of variables.

Eq. A6 might equivalently be regarded as a generalisation of a Schwinger/Feynmann parameterisation, with,

$$\begin{aligned}\Pi_{j=1}^{m} g_j(s) &= \int_0^1 dy_1 \ldots \int_0^1 dy_{m-1} y_1^0 y_2^1 y_3^2 \ldots y_{m-1}^{m-2} \times \\ &\mathcal{L}\left[t^{m-1}\mathcal{L}^{-1}\left\{g_1(s)\right\}(ty_1 \ldots y_{m-1})\right. \\ &\mathcal{L}^{-1}\left\{g_2(s)\right\}(t(1-y_1)y_2 \ldots y_{m-1}) \\ &\ldots \\ &\left.\mathcal{L}^{-1}\left\{g_m(s)\right\}(t(1-y_{m-1}))\right]\end{aligned} \tag{A8}$$

For example, taking $g_j(s) = 1/(s + a_j)^{p_j}$ and noting that $\mathcal{L}^{-1}\left\{1/(s+a_j)^{p_j}\right\} = t^{p_j-1}e^{-a_j t}/\Gamma(p_j)$, then we get,

$$\begin{aligned}\Pi_{j=1}^{m} \tfrac{1}{(s+a_j)^{p_j}} &= \tfrac{\Gamma\left(\sum_{i=1}^{m} p_i\right)}{\Pi_{i=1}^{m}\Gamma(p_i)} \int_0^1 dy_1 \ldots \int_0^1 dy_{m-1} y_1^0 y_2^1 \ldots y_{m-1}^{n_m-1} \\ &(y_1 \ldots y_{m-1})^{p_1-1}\left((1-y_1)y_2 \ldots y_{m-1}\right)^{p_2-1} \ldots \\ &(y_{m-1}(1-y_{m-2}))^{p_{m-1}-1}(1-y_{m-1})^{p_m-1} \\ &\tfrac{1}{[s+(a_1 y_1 \ldots y_{m-1}+a_2(1-y_1)y_2 \ldots y_{m-1}+\cdots+a_m(1-y_{m-1})]^{\sum_{i=1}^{m} p_i}}\end{aligned} \tag{A9}$$

Taking $s = 0$, $m = 2$, and $p_j = 0$ for all j, gives the most well-known form, with [18],

$$\frac{1}{a_1 a_2} = \int_0^1 \frac{dy_1}{(a_2 y_1 + (1-y_1)a_1)^2} \tag{A10}$$

The identity Eq. A9 can be confirmed by writing the denominator as $(A_m)^m$, with,

$$A_m = (a_m(1-y_{m-1}) + y_{m-1}A_{m-1}(y_1, \ldots, y_{m-2})) \text{ and } A_1 = a_1 \tag{A11}$$

and integrating with respect to each of y_{m-1} to y_1 in turn. For example, using the substitution $u = y_{m-1}/(1 + \alpha_{m-1}y_{m-1})$ with $\alpha_{m-1} = (A_{m-1} - a_m)/a_m$ and integrating between $u = 0$ and $u = 1/(1 + \alpha_{m-1})$, the integrand becomes $(1/a_m)(1/{A_{m-1}}^m)$. Repeating this for y_{m-1} to y_1 confirms the identity.