



# HHS Public Access

Author manuscript

Phys Rev E. Author manuscript; available in PMC 2022 January 01.

Published in final edited form as:

Phys Rev E. 2021 January ; 103(1-1): 012412. doi:10.1103/PhysRevE.103.012412.

## Generalized stochastic microdosimetric model: The main formulation

**F. Cordoni,**

Department of Computer Science, University of Verona, Verona, Italy and TIFPA-INFN, Trento, Italy

**M. Missiaggia,**

Department of Physics, University of Trento, Trento, Italy and TIFPA-INFN, Trento, Italy

**A. Attili,**

INFN, Roma 3, Italy

**S. M. Welford,**

Department of Radiation Oncology, University of Miami, Miller School of Medicine, Miami, Florida 33136, USA

**E. Scifoni,**

TIFPA-INFN, Trento, Italy

**C. La Tessa**

Department of Physics, University of Trento, Trento, Italy and TIFPA - INFN, Trento, Italy

### Abstract

The present work introduces a rigorous stochastic model, called the generalized stochastic microdosimetric model (GSM<sup>2</sup>), to describe biological damage induced by ionizing radiation. Starting from the microdosimetric spectra of energy deposition in tissue, we derive a master equation describing the time evolution of the probability density function of lethal and potentially lethal DNA damage induced by a given radiation to a cell nucleus. The resulting probability distribution is not required to satisfy any *a priori* conditions. After the initial assumption of instantaneous irradiation, we generalized the master equation to consider damage induced by a continuous dose delivery. In addition, spatial features and damage movement inside the nucleus have been taken into account. In doing so, we provide a general mathematical setting to fully describe the spatiotemporal damage formation and evolution in a cell nucleus. Finally, we provide numerical solutions of the master equation exploiting Monte Carlo simulations to validate the accuracy of GSM<sup>2</sup>. Development of GSM<sup>2</sup> can lead to improved modeling of radiation damage to both tumor and normal tissues, and thereby impact treatment regimens for better tumor control and reduced normal tissue toxicities.

### I. INTRODUCTION

Currently, around 50% of all patients with localized malignancies undergo treatment including ionizing radiation, mostly in combination with tumor resection and/or chemotherapy [1,2]. Conventional therapy with high-energy photons is by far the most

common approach, but the use of accelerated particles has grown exponentially, especially in the past decade. The well-defined, energy-dependent, range with sharp distal fall-off and the limited lateral beam spread, typical of ions when penetrating a medium, translate into a dose profile delivered with millimeter precision. In addition, charged particles, especially for larger charge, present enhanced biological effectiveness compared to photons, resulting in reduced cellular repair [3–5]. Thus the field of radiation oncology is evolving toward broader application of radiotherapy with ions, while still several key physical and biological questions remain to be fully unraveled. In particular, the need to account for a biologically effective dose, beyond a purely physical energy deposition, imposes an advanced characterization of a beam [6]. The calculation of the effective dose distribution delivered to the patient during a treatment indeed requires detailed knowledge of the radiation field composition at the tumor site and surrounding tissue. The beam quality across its propagation in the medium is, in fact, modified by nuclear and electromagnetic interactions of the primary ions with the patient's body nuclei, atoms, and molecules, creating a mixed radiation field composed of primary as well as secondary nuclear fragments of different charge and kinetic energy [3,7].

In such a complex radiation field, many different mechanisms deliver a variety of nanoscopic damages to the biological target molecules, mainly mediated by their secondary electrons distribution [8,9] and by the radicals generated from the scattered electrons [10,11], although there are lesser processes [12,13]. Clearly, such a nanoscopic pattern of energy deposition produces a *complexity* of molecular damage [14], which correlates with a different reparability and thus a different biological response. An accurate approach for characterizing the complex radiation field produced by an ion beam is microdosimetry [15]. There are two main points of strength in using this methodology: (i) the energy deposited by radiation is measured in an area with dimensions comparable to a cell nucleus; and (ii) stochastic fluctuations of energy deposition, e.g., from cell to cell, are taken into account. Microdosimetry is considered to be a link between the physical characteristics and the biological effectiveness of radiation with the advantage of an experimentally measurable physical quantity, and it has been used in radiobiological models to describe radiation quality. One of the most relevant examples is the microdosimetric kinetic model (MKM), which was formulated in its original version in [16,17] as an elaboration of the theory of dual radiation action (TDRA) [18,19] and of the repair-misrepair model (RMR) [20,21]. The MKM exploits microdosimetric spectra to calculate the energy deposited by radiation and predicts cell survival modeling the DNA-damage repair kinetics. Today, it is one of two radiobiological models employed clinically in particle therapy, together with the local effect model (LEM) [22,23].

Although based on microdosimetry, the MKM is a purely deterministic model as only the average number of lethal lesions induced by radiation to the DNA is considered. The model aims to provide a mathematical formulation of the kinetic evolution of double-strand breaks (DSB) in the DNA in order to calculate the cell survival fraction. Mathematically, the temporal evolution of a DSB is described by a system of two ordinary differential equations representing the average number of lethal and potentially lethal damages as a function of time. This description is accurate only as long as the lethal and potentially lethal damage distributions are Poissonian, and results in a cell survival curve that follows a linear-

quadratic behavior [16]. However, it has been widely shown in the literature that the DNA damage distribution deviates significantly from a Poisson function under several irradiation conditions, such as high-dose or high-LET (linear energy transfer) [24]. For this reason, several recent studies focused on implementing corrections to the original MKM formulation to account for non-Poissonian behaviors of the DNA damage distribution [17,24–28]. An extensive collection of the original MKM formulation and its subsequent generalizations can be found in [29]. Nonetheless, all MKM versions are based on the original deterministic formulation described by Hawkins [16].

The main goal of the present work is to develop a fully probabilistic model of DNA damage formation and its kinetic evolution based on microdosimetry. The new model, called the Generalized Stochastic Microdosimetric Model ( $GSM^2$ ), will provide a rigorous and general mathematical description of DNA damage time-evolution without using any *a priori* assumption on the lesion distribution (e.g., a Poisson). The model accuracy will be tested for different irradiation conditions (beam quality, dose, and dose rate) and compared with MKM predictions to prove both  $GSM^2$  validity and advances compared to the current standard.

In our model formulation, the potential damage induced by radiation to the proteins of the repairing machinery was not considered. This approximation can have a non-negligible impact on the cell response, but normally only at a second-order level compared to DNA damage, as it is also assumed in similar radiation damage models such as the MKM and the LEM [30].

The classical approach for mathematically modeling a complex physical system, such as the one resulting from the interaction between cells and ionizing radiation that leads to the formation of DNA lesions, is achieved with deterministic models. In these approaches, given an initial condition the system time-evolution can be completely characterized at each state. Recent studies [31] have shown that this approach fails mainly for three reasons: (i) a precise and accurate estimation of the parameters is often not feasible; (ii) such models become unrealistic in accounting for all relevant interactions as the system complexity increases, [32]; and (iii) certain systems can be oversensitive to some input parameters, typically the initial values. The inclusion of stochasticity in the modeling of complex systems using suitable random variables is natural for many systems, and lets such models avoid much of the difficulties of deterministic models.

To model complex physical processes, such as lesions formation following a radiation exposure, the standard method is to consider the *macroscopic* system, so that the main focus is on the system as a whole; this approach typically shows that the principal relationship governing the physical or biological processes is deterministic, and its predictions are assumed to represent average values. In a *microscopic* (or often nanoscopic) approach, instead, each element of the system is usually modeled using Brownian dynamics [33–35]. However, the complexity of lesion formation and time-evolution makes a full Brownian dynamics-representation intractable.

To obtain a more general and accurate description of DNA lesion formation and evolution than those provided by a *macroscopic* approach, and yet to maintain suitable mathematical

tractability of the main equations, which is often missing in a *microscopic* approach, a hybrid methodology, known as *mesoscopic*, can be considered. This approach takes into account the stochastic nature of a system while remaining manageable from both the analytical and numerical points of view. The *mesoscopic* method is based on the assumption that the process driving the system evolution is a Markov jump process [36]. The equations of motion are described via the so-called *master equation*, which contains a DNA damage probability density function of the whole system [33,36,37].

### A. Rationale for a new model

In GSM<sup>2</sup>, we will introduce an equation, referred to as the microdosimetric master equation (MME), that governs the time evolution of the joint probability density function for lethal and sublethal damages inside the cell nucleus and is based on just three rate parameters  $a$ ,  $b$ , and  $r$ , defined below. The main innovation in comparison to the existing approaches is that in the proposed MME, account is taken of the variations in both lesions formation and evolution caused by the randomness of these processes. In particular, we will use microdosimetry spectra for describing radiation quality and include the stochastic nature of energy deposition.

To provide a rigorous mathematical formulation of the DNA damage kinetics, we will consider lethal and sublethal lesions inside a single cell nucleus. Potentially lethal lesions can either be repaired or not, in which case they become lethal lesions. A cell in which at least one lethal lesion has been formed is considered inactivated. A potentially lethal damage induced by radiation can undergo three main processes: (i) it can be repaired at a rate  $r$ ; (ii) it can become a lethal damage at a rate  $a$ ; or (iii) it can combine with another potentially lethal lesion to form a lethal lesion at rate  $b$ .

Starting from some probabilistic assumptions on the lesions formation, we will derive a master equation that describes the time evolution for the joint probability density function of DNA lesions for both lethal and potentially lethal damage. The density function solution will be shown to have a first moment in agreement with the standard MKM driving equations. One important goal of this study is to overcome the Poissonian assumption on lethal lesions.

In the present work, we will further generalize the MME in two main directions. In particular, besides the damage kinetic mechanisms (i), (ii), and (iii) introduced above, we will additionally consider that (iv) either a lethal or sublethal damage can be formed randomly due to the effect of the ionizing radiation at a rate  $\dot{d}$ ; and (v) lethal lesions can move inside the cell nucleus. Case (iv) can represent DNA damage formation resulting from a continuous irradiation field. In fact, together with standard lesion interactions, we will also take into account random jumps in the number of lethal and sublethal lesions caused by the stochastic nature of radiation-energy deposition.

Case (v) accounts for the fact that we also allow lesions to move between adjacent domains. Because the GSM<sup>2</sup> model considers pairwise interactions of potentially lethal lesions, the domain size plays a crucial role. In fact, a domain that is too big might imply that lesions created far away from each other can interact to form a lethal lesion. On the other hand, a domain that is too small results in a lower number of lesions per domain so that the

probability of double events can be underestimated. In the limiting case in which the domain size approaches zero, lesion interactions across neighboring domains also approach zero [38,39]. To minimize the model dependence on the domain size, we will allow interactions between lesions both belonging to the same domain and to different domains [31].

In summary, we will introduce a general master equation that models the joint probability distribution of DNA lethal and potentially lethal lesion inside a cell nucleus. The derived master equation will encompass a number of effects. Besides the effect of potentially lethal lesion repair and death due to either spontaneous death or pairwise interaction, the model also has the stochastic effect of energy deposition due to ionizing radiation and lesions movements between adjacent domains, providing a description of the cell nucleus as a whole. Not yet included is a time-delayed active repair of double-strand breaks, nor is cell apoptosis, although a spontaneous repair rate term can simulate much of the effect of repair mechanisms when applied to protracted dosages. Damage due to radicals produced by the ionizing radiation begins near the time of exposure, and so their effects are represented.

To validate GSM<sup>2</sup>, we will consider microdosimetric energy spectra obtained from GEANT4 simulations [40]. We will show how different assumptions related to the probability distribution of damages number, as well as model parameters, show significant deviation from the Poisson distribution assumed by most existing models, including the MKM. We will further compute the survival probability and compare it to the classical *linear-quadratic* (LQ) model [41,42].

The innovations presented in this work are several. We will develop a fully probabilistic description of the DNA damage kinetic. In particular, the joint probability distribution of the number of sublethal and lethal lesions will be modeled. We will further generalize the model including interdomain movements and continuous damage formation due to protracted dose. The resulting *master equation* solution will provide the real probability distribution without any *a priori* assumption on the density function, allowing the computation of several biological end points. The proposed approach will be able to fully describe the stochastic nature of energy deposition both in time and space, improving the existing models where the energy deposition is averaged over both the whole cell nucleus and cell population. In doing so, we will be able to reproduce several behaviors referred to in the literature as *non-Poissonian effects*, which cannot be predicted by the MKM and its variants and are typically included in the models with ad hoc corrections [17,24,25,28].

Because of GSM<sup>2</sup> flexibility and generality, analytical solutions both on the probability density function and on the resulting survival curve are not of easy derivation. Therefore, the present study is intended as a first step of a systematic investigation of the stochastic nature of energy deposition and how it influences lesion formation. In particular, a further investigation will focus on the long-time behavior of the *master equation* and the resulting survival curve. Furthermore, the principles used in the current approach will be used to develop a fully stochastic model of intercellular damage formation optimized to improve radiation field characterization via a novel hybrid detector for microdosimetry, [43].

With GSM<sup>2</sup> and its future developments, we try to shed new light on non-Poissonian effects in order to obtain a deeper mechanistic understanding, which will allow us to model them more accurately.

## B. Structure of the paper

The present paper is structured as follows: Section II recalls basic assumptions and formulation of the MKM and its variants, [17,24,26–28]. Then Sec. III introduces the main *master equation* describing the probability distributions of lethal lesions. Connections of the current model to the standard MKM are presented in Sec. III A. Section III B shows in detail how microdosimetric spectra can be used to extract the energy deposition. Sections III C and III D introduce the above-mentioned generalization of the *master equation* to consider split dose and domain interconnection. Connections of the current model to the standard MKM are presented in Sec. III A. Further, long-time behavior and survival probability resulting from the GSM<sup>2</sup> are presented in Sec. III E. Finally, Sec. IV presents some numerical examples aiming at highlighting specific aspects resulting from the governing *master equation*.

## II. FUNDAMENTALS ON THE MICRODOSIMETRIC KINETIC MODEL AND RELATED NON-POISSONIAN GENERALIZATIONS

The microdosimetric kinetic model (MKM) is based on the following assumptions:

- (i) The cell nucleus can be divided into  $N_d$  independent domains.
- (ii) Radiation can create two different kinds of DNA damage, referred to as sublethal and lethal.
- (iii) Lethal lesions cannot be repaired. On the contrary, sublethal lesions can either be repaired or evolve into a lethal lesions either by spontaneous death or by interaction with another sublethal lesion.
- (iv) The number of sublethal and lethal lesions in a single domain  $d$  is proportional to the specific energy  $z$  delivered by radiation to the site.
- (v) Cell death occurs if at least one domain suffers at least one lethal lesion.

In the described setting, lethal lesions represent clustered double-strand breaks that cannot be repaired, whereas sublethal lesions are double-strand breaks that can be repaired.

We denote by  $\bar{x}_{g, z_d}$  and  $\bar{y}_{g, z_d}$  the average number of sublethal and lethal lesions, respectively, induced in the domain  $d$  that received a specific energy  $z_d$ . For ease of notation, we will omit the subscript  $(d, z)$  and indicate  $\bar{x}_{d, z} := \bar{x}$  and  $\bar{y}_{d, z} := \bar{y}$ . The following set of coupled ordinary differential equations (ODEs) is satisfied:

$$\begin{aligned}\frac{d}{dt}\bar{y}(t) &= a\bar{x} + b\bar{x}^2, \\ \frac{d}{dt}\bar{x}(t) &= -(a+r)\bar{x} - 2b\bar{x}^2.\end{aligned}\quad (1)$$

The term proportional to  $\bar{x}_{g, z_d}$  in the first equation comes about because the number of lethal lesions increases, in part, from the conversion of single nonlethal lesions that are present, and the term proportional to  $\bar{x}_{g, z_d}^2$  is due to the interaction of pairs of nonlethal lesions producing a lethal one. In the second equation, the  $r$  term determines the rate of repaired nonlethal lesions, while the  $2b$  term accounts for the loss of nonlethal lesions when two make one lethal one.

If further  $(a+r)\bar{x}_{d, z_d} \gg 2b\bar{x}_{d, z_d}^2$ , then Eq. (1) can be simplified as

$$\begin{aligned}\frac{d}{dt}\bar{y}(t) &= a\bar{x} + b\bar{x}^2, \\ \frac{d}{dt}\bar{x}(t) &= -(a+r)\bar{x},\end{aligned}\quad (2)$$

making solutions expressible in terms of dropping exponentials in time.

One of the main goals of the MKM is to predict the survival probability of cell nuclei when exposed to ionizing radiation, whose quality is described with a microdosimetry approach. To achieve this result, an additional assumption to those listed above must be made:

(vi) Lethal lesions follows a Poissonian distribution.

Under the latter assumption, the probability  $S_{d, z_d}$  that a domain  $d$  survives as  $t \rightarrow \infty$  when receiving the specific energy  $z_d$  can be computed as the probability that the random outcome of a Poisson random variable is null. Therefore,  $S_{d, z_d}$  is given by

$$S_{d, z_d} = e^{-\lim_{t \rightarrow \infty} \bar{y}_{d, z_d}(t)}. \quad (3)$$

An explicit computation [16,44,45] shows that the number of lethal lesions as  $t \rightarrow \infty$  can be expressed as

$$\lim_{t \rightarrow \infty} \bar{y}_{d, z_d}(t) = \left( \lambda + \frac{a\kappa}{a+r} \right) z_d + \frac{b\kappa^2}{2(a+r)} z_d^2. \quad (4)$$

Combining Eqs. (4) and (3), we obtain

$$S_{d, z_d} = e^{-Az_d - Bz_d^2},$$

with  $A$  and  $B$  some suitable constants independent of  $d$  and  $z_d$ .

The survival probability (3) can be extended to the whole cell nucleus ( $S_n$ ) by averaging it on all domains as

$$S_{n, z_n} := \exp\left(-\sum_{d=1}^{N_d} \left\langle \lim_{t \rightarrow \infty} \bar{y}_{d, z_d}(t) \right\rangle\right). \quad (5)$$

Finally, by averaging  $S_{n, z_n}$  over the entire cell population, the overall cell survival can be calculated as

$$S = \exp(-\alpha D - \beta D^2), \quad (6)$$

where  $D$  is the macroscopic dose delivered to the entire cell population. Details on how the survival function  $S$  was derived can be found in [16,17,24,25].

Several generalizations [24–28,46,47] have been proposed to take into account effects due to a deviation of the lethal lesion behavior from a Poissonian distribution. All models try to correct the survival probability (6) by introducing some correction term based on the overkilling effects. An overkilling effect may come about because a single particle deposits much more energy than is required to kill a cell [48], resulting in fewer cells killed per absorbed dose. The typical survival correction is of the form [24]

$$S = \exp(-[\alpha_0 + f(\bar{z}_d, \bar{z}_n)\beta]D - \beta D^2),$$

where  $f(\bar{z}_d, \bar{z}_n)$  is a suitable correction term that depends on both energy deposition on the single domain ( $\bar{z}_d$ ) and on the cell nucleus ( $\bar{z}_n$ ). An alternative form is given by [25]

$$S = \exp[-(\alpha_0 + \bar{z}_d^* \beta)D - \beta D^2],$$

where  $\bar{z}_d^*$  is a term that accounts for the overkilling effects.

We refer to [29] for a comprehensive review of the biophysical models of DNA damage based on microdosimetric quantities.

It is worth highlighting that all corrections so far proposed for non-Poissonian effects rely on ad hoc terms derived from empirical considerations. The final goal of this study, instead, is to obtain analogous corrections based on physical considerations stemming from the stochastic nature of energy deposition [49].

### III. THE GENERALIZED STOCHASTIC MICRODOSIMETRIC MODEL GSM<sup>2</sup>

As part of this study, we investigated how the models described in Sec. II could be developed to rely on the whole probability distribution for  $x$  and  $y$  rather than simply on its mean value. In fact, all proposed generalizations of the MKM always consider deterministic driving equations for predicting the number of lethal and sublethal lesions. Non-Poissonian



effects are often proposed as correction terms added to the survival fraction predicted by the MKM with no formal mathematical derivation and mainly based on empirical evaluations.

The MKM formulation is based on the probability distribution of inducing a damage when a specific energy  $z$  is deposited. Once the survival for a given  $z$  is computed, the specific energy is averaged over the whole cell population to yield the overall expected survival probability. To the best of our knowledge, there is no systematic investigation that aims at capturing the true stochasticity of both energy deposition and lesion formation.

The main goal of the present work is thus to generalize microdosimetric-based models in order to describe the full probability distribution of lethal and sublethal lesions. We will take advantage of assumptions (i)–(v) described in Sec. II. Regarding assumption (iv), the MKM assumes that the initial distribution of the lethal lesions, given an energy deposition  $z$ , follows a Poisson law. We will generalize this assumption assuming a general initial distribution, allowing to fully describe the stochastic nature of energy deposition. This point will be treated in detail in Sec. III B.

An additional remark on the importance of the initial distribution is necessary to fully understand the implication of the generalization we will carry out in this study. The stochasticity of energy deposition in a microscopic volume is the basic foundation of microdosimetry, and assuming every probability distribution to be Poissonian is a restrictive assumption that limits the model application.

To capture the real stochastic nature of energy deposition and related DNA damage formation, we will provide a probabilistic reformulation of Eq. (1). We denote by  $(Y(t), X(t))$  the system state at time  $t$ , where  $X$  and  $Y$  are two  $\mathbb{N}_0$ -valued random variables representing the number of lethal and sublethal lesions, respectively. We will consider a standard complete filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  that satisfies the usual assumptions of right-continuity and saturation by  $\mathbb{P}$ -null sets.

Let us consider two different sets  $\mathcal{X}$  and  $\mathcal{Y}$  containing the possible values for the number of sublethal and lethal lesions, respectively. As we have indicated after Eq. (1), the heuristic interpretation of the coefficients in Eq. (1) is that  $a$  is the rate at which a sublethal lesion becomes a lethal lesion,  $r$  is the rate at which a sublethal lesion recovers and goes to the set  $\emptyset$  (i.e., that of the healthy cells), whereas  $b$  is the rate at which two sublethal lesions interact to become a single lethal lesion. These considerations can be mathematically expressed as



At a given time  $t$ , the probability to observe  $x$  sublethal lesions and  $y$  lethal lesions is expressed as

$$p(t, y, x) = \mathbb{P}[(Y(t), X(t)) = (y, x)].$$

Also,

$$p_{t_0, y_0, x_0}(t, y, x) := p(t, y, x | t_0, y_0, x_0) = \mathbb{P}[(Y(t), X(t)) = (y, x) | (Y(t_0), X(t_0)) = (y_0, x_0)]$$

is the probability conditioned to the fact that at  $t = t_0$  there were  $x_0$  and  $y_0$  sublethal and lethal lesions, respectively.

To determine the governing master equation for the above probability density  $p(t, y, x)$ , we need to account for all possible system changes in the infinitesimal time interval  $dt$ .

Thus, the following scenarios may happen:

(i) At time  $t$  we have exactly  $(y, x)$  lesions and they remain equal with a rate  $[1 - (a + r)x - bx(x - 1)]$ , namely

$$\mathbb{P}[(Y(t + dt), X(t + dt)) = (y, x) | (Y(t), X(t)) = (y, x)] = 1 - [(a + r)x - bx(x - 1)]dt + O(dt^2).$$

(ii) At time  $t$  we have exactly  $(y, x + 1)$  lesions, and one sublethal lesion recovers with rate  $(x + 1)r$ , namely

$$\mathbb{P}[(Y(t + dt), X(t + dt)) = (y, x) | (Y(t), X(t)) = (y, x + 1)] = (x + 1)rdt + O(dt^2).$$

(iii) At time  $t$  we have exactly  $(y - 1, x + 1)$  lesions, and one sublethal lesion becomes a lethal lesion with a rate  $(x + 1)a$ , namely

$$\mathbb{P}[(Y(t + dt), X(t + dt)) = (y, x) | (Y(t), X(t)) = (y - 1, x + 1)] = (x + 1)adt + O(dt^2).$$

(iv) At time  $t$  we have exactly  $(y - 1, x + 2)$  lesions, and two sublethal lesions become one lethal lesion with a rate  $(x + 2)(x + 1)b$ , namely

$$\mathbb{P}[(Y(t + dt), X(t + dt)) = (y, x) | (Y(t), X(t)) = (y - 1, x + 2)] = (x + 2)(x + 1)bdt + O(dt^2).$$

Grouping the equations derived in Sec. III, we obtain

$$p(t + dt, y, x) = p(t, y, x) \left( 1 - [(a + r)x - bx(x - 1)]dt + O(dt^2) \right) + p(t, y, x + 1) \left( (x + 1)rdt + O(dt^2) \right) + p(t, y - 1, x + 1) \left( (x + 1)adt + O(dt^2) \right) + p(t, y - 1, x + 2) \left( (x + 2)(x + 1)bdt + O(dt^2) \right).$$

Taking the limit as  $dt \rightarrow 0$ , we arrive at the *microdosimetric master equation* (MME)

$$\begin{aligned} \partial_t p(t, y, x) = & -[(a+r)x - bx(x-1)]p(t, y, x) + (x+1)rp(t, y, x+1) \\ & + (x+1)ap(t, y-1, x+1) + (x+2)(x+1)bp(t, y-1, x+2), \end{aligned} \quad (8)$$

where  $\partial_t$  denotes the partial derivative with respect to the time variable. Equation (8) must be equipped with a suitable initial condition  $p(0, y, x) = p_0(y, x)$ .

We remark that the above-derived MME arises solely from the probabilistic assumptions regarding lesion formation.

The MME (8) can be written for short as

$$\begin{aligned} \partial_t p(t, y, x) = & (E^{-1,2} - 1)[x(x-1)bp(t, y, x)] + (E^{-1,1} - 1)[xap(t, y, x)] \\ & + (E^{0,1} - 1)[xrp(t, y, x)] \\ = & \mathcal{E}^{-1,2}[x(x-1)bp(t, y, x)] + \mathcal{E}^{-1,1}[xap(t, y, x)] + \mathcal{E}^{0,1}[xrp(t, y, x)], \end{aligned} \quad (9)$$

where above we have denoted the creation operators defined as

$$\mathcal{E}^{i,j}[f(t, y, x)] := (E^{i,j} - 1)[f(t, y, x)] := f(t, y+i, x+j) - f(t, y, x).$$

## A. Connection with the MKM

The present section aims at showing that the mean value of the master equation does satisfy, under certain assumptions, the kinetic equations (1). In what follows,  $\mathbb{E}$  denotes the mean value of a random variable defined as

$$\bar{x}(t) := \mathbb{E}[X(t)] = \sum_{x, y \geq 0} xp(t, y, x),$$

$$\bar{y}(t) := \mathbb{E}[Y(t)] = \sum_{x, y \geq 0} yp(t, y, x).$$

Note that, for a general function  $f$ , the following holds true:

$$\begin{aligned} \sum_{x, y \geq 0} x \mathcal{E}^{i,j}[f(y, x)p(t, y, x)] &= -\mathbb{E}jf(Y, X), \\ \sum_{x, y \geq 0} y \mathcal{E}^{i,j}[f(y, x)p(t, y, x)] &= -\mathbb{E}if(Y, X). \end{aligned} \quad (10)$$

Therefore, multiplying the MME (9) by  $x$  and then by  $y$ , we obtain using (10)

$$\begin{aligned} \frac{d}{dt}\mathbb{E}[Y(t)] &= b\mathbb{E}\{X(t)[X(t)-1]\} + a\mathbb{E}[X(t)], \\ \frac{d}{dt}\mathbb{E}[X(t)] &= -2b\mathbb{E}\{X(t)[X(t)-1]\} - (a+r)\mathbb{E}[X(t)]. \end{aligned} \quad (11)$$

Equations (11) are still not of the form of Eqs. (2); in particular, they depend on a second-order moment  $\mathbb{E}\{X(t)[X(t) - 1]\}$ . Nonetheless, explicit computation will show that, if we tried to compute a kinetic equation for the second-order moment  $\mathbb{E}\{X(t)[X(t) - 1]\}$ , we would obtain a dependence on higher moments, and thus an infinite set of coupled ODEs. To solve the impasse, we shall make what is called a *mean-field* assumption, that is, we assume that

$$\mathbb{E}\{X(t)[X(t) - 1]\} \sim \mathbb{E}[X(t)]^2.$$

Under the above *mean-field assumption*, Eqs. (11) become

$$\begin{aligned} \frac{d}{dt} \bar{y}(t) &= b\bar{x}^2(t) + a\bar{x}(t), \\ \frac{d}{dt} \bar{x}(t) &= -2b\bar{x}^2(t) - (a+r)\bar{x}(t), \end{aligned} \quad (12)$$

and the original kinetic equations are in turn recovered.

A quick remark on the *mean-field assumption* is needed. In the case of  $x$  being large enough, we have that the following approximation holds true:  $\mathbb{E}\{X(t)[X(t) - 1]\} \sim \mathbb{E}[X^2(t)]$ ; therefore, the *mean-field assumption* means that  $\mathbb{E}[X^2(t)] - \mathbb{E}[X(t)]^2 \sim 0$ . Noticing that the last term is merely the variance, and recalling that the variance for a random variable is null if and only if the random variable is in fact deterministic, if the *mean-field assumption* is reasonable, then the realized number of lesions does not differ much from the mean value so that all we need to know is contained in the mean value. On the contrary, if there is evidence that the mean value is not a reasonable approximation for the realized number of lesions, the *mean-field assumption* must be considered unacceptable, in which case more detailed knowledge of the probability distribution is essential to have a good understanding of the system.

## B. Initial distribution for the number of lethal and sublethal lesions

One of the main advantages of the proposed model is that the distribution of DNA damages induced by an ionizing radiation  $z$  does not need to be chosen as Poissonian. In the present section, we will show how the number of induced lesions can be evaluated starting from microdosimetric spectra.

Let  $f_{1;d}(z)$  be the single-event distribution of energy deposition on a domain  $d$ ; see [15]. The single-event distribution  $f_{1;d}$  can be either computed numerically via the Monte Carlo toolkit supplied by GEANT4 [40], or by experimental microdosimetric measurements.

The full probability distribution of an energy deposition thus depends on the number of events that deposit energy on the cell nucleus. Given a cell nucleus, composed by  $N_d$  domains, the probability that  $\nu$  events deposit an energy  $z$  obeys a Poissonian distribution of mean  $\lambda_n := \frac{z_n}{z_F}$ , where  $z_n$  is the mean energy deposition on the nucleus, i.e.,

$$z_n = \int_0^\infty z f(z|z_n) dz,$$

and  $z_F$  is the first moment of the single event distribution  $f_{1;d}$  defined as

$$z_F := \int_0^\infty z f_{1;d}(z) dz. \quad (13)$$

Then, assuming a Poissonian probability that a domain registers  $\nu$  events, the energy deposition distribution is given by

$$f(z|z_n) := \sum_{\nu=0}^{\infty} \frac{e^{-\frac{z_n}{z_F}} \left(\frac{z_n}{z_F}\right)^\nu}{\nu!} f_{\nu;d}(z),$$

where  $f_{\nu;d}(z)$  is the energy deposition distribution resulting from  $\nu$  depositions.

In particular, given a domain  $d$  that suffers  $\nu$  energy deposition events, the distribution resulting from  $\nu$  events can be computed convolving  $\nu$  times the single event distribution; see [15,28]. Therefore, the imparted energy  $z$  has distribution  $f_{\nu;d}$  computed iteratively as

$$\begin{aligned} f_{2;d}(z) &:= \int_0^\infty f_{1;d}(\bar{z}) f_{1;d}(z-\bar{z}) d\bar{z}, \\ &\dots \\ f_{\nu;d}(z) &:= \int_0^\infty f_{1;d}(\bar{z}) f_{\nu-1;d}(z-\bar{z}) d\bar{z}. \end{aligned}$$

For a certain energy deposition  $z$ , the induced number of lesions is a random variable. The standard assumption is that the distribution of  $X$  given  $z$  is a Poisson random variable of mean value  $\kappa_z$ . Analogous reasoning holds for  $Y$ , being the number of induced lethal lesions, a Poisson random variable of mean  $\lambda_z$ . Given the high flexibility of the proposed approach, the number of induced lesions given an energy deposition  $z$  can be any random variable. It is worth stressing that the chosen distribution may vary with LET.

In the following general treatment we will denote by  $p_z^X(x|\kappa_z) [p_z^Y(y|\lambda_z)]$  the initial random distribution for the number of sublethal (lethal) lesions given an energy deposition  $z$ . We remark again that both  $p_z^X(x|\kappa_z)$  and  $p_z^Y(y|\lambda_z)$  can be any probability distribution. Specific relevant examples will be considered in the numerical implementation.

Putting all the above reasoning together, the MME (9) reads

$$\partial_t p(t, y, x) = \mathcal{E}^{-1,2}_{[x(x-1)bp(t, y, x)]} + \mathcal{E}^{-1,1}_{[xap(t, y, x)]} + \mathcal{E}^{0,1}_{[xrp(t, y, x)]},$$

$$p(0, y, x) = p_0^X(x) p_0^Y(y), \quad (14)$$

where the initial distribution is obtained as

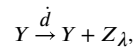
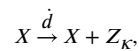
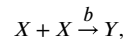
$$\begin{aligned}
 p_0^X(x) &= \int_0^\infty p_z^X(x|\kappa_z) f(z|z_n) dz, \\
 p_0^Y(y) &= \int_0^\infty p_z^Y(y|\kappa_z) f(z|z_n) dz.
 \end{aligned}
 \tag{15}$$

### C. The protracted dose case for the generalized stochastic microdosimetric model

The MME can be further generalized to consider *protracted dose* irradiation. We refer to a *protracted dose* as a continuous dose delivered over a range of time. In contrast, a dose delivered over a short time is called *acute dose irradiation*, whereas a series of such acute irradiations at prescribed time steps is referred to as *split dose irradiation*. Existing models fail at properly describing protracted dose, being unable to fully capture the stochasticity inherent to energy deposition. Usually, strong assumptions are used to treat protracted dose, [24], or a split dose is used to approximate a continuous dose delivery [26]. Nonetheless, models cannot fully predict experimental data [26].

The generalization of the GSM<sup>2</sup> master Eq. (9) to consider a continuous dose irradiation is not trivial. In fact, at random time  $t$  the number of lesions, either lethal or sublethal, exhibits a jump upward of a random quantity that depends on the energy deposition  $z$ , which we recall is a random variable.

More formally, the possible interactions now become



where  $Z_\lambda$  and  $Z_\kappa$  are two random variables with integer-valued distributions  $p_0^X$  and  $p_0^Y$ , respectively, defined as in Eq. (14). The parameter  $\dot{d}$  represents the dose rate (see [46,47]), and it is given by  $\dot{d} := \frac{z_n}{T_{\text{irr}} z_F}$ , with  $z_F$  given in Eq. (13) and  $T_{\text{irr}}$  is the total irradiation time.

We have the following:

(i)

$$\begin{aligned} & \mathbb{P}[(Y(t+dt), X(t+dt)) = (y, x) | (Y(t), X(t)) = (y, x)] \\ & = 1 - \left\{ (a+r)x + bx(x-1) + d \left[ 1 - p_0^X(0) \right] \left[ 1 - p_0^Y(0) \right] \right\} dt + O(dt^2), \end{aligned}$$

(ii)

$$\begin{aligned} & \mathbb{P}[(Y(t+dt), X(t+dt)) = (y, x) | (Y(t), X(t)) = (y - i_y, x - i_x)] \\ & = d p_0^X(i_x) p_0^Y(i_y) dt + O(dt^2), \quad i_x = 1, \dots, x, i_y = 1, \dots, y. \end{aligned}$$

Furthermore, reactions (ii), (iii), and (iv) in Sec. III remain valid.

Therefore, a similar analysis to the one carried out in Sec. III leads to the following MME:

$$\begin{aligned} \partial_t p(t, y, x) &= (E^{-1,2} - 1)[x(x-1)bp(t, y, x)] + (E^{-1,1} - 1)[xap(t, y, x)] \\ &+ (E^{0,1} - 1)[xrp(t, y, x)] + \left( \sum_{i_x=1}^x \sum_{i_y=1}^y E_d^{-i_y, -i_x} - [1 - p_0^X(0)][1 - p_0^Y(0)] \right) [dp(t, \\ &y, x)] \\ &\mathcal{E}^{-1,2}[x(x-1)bp(t, y, x)] + \mathcal{E}^{-1,1}[xap(t, y, x)] + \mathcal{E}^{0,1}[xrp(t, y, x)] + \mathcal{E}_d^{-y, -x} \\ &[dp(t, y, x)]. \end{aligned} \quad (16)$$

The operator in the last line of Eq. (16), right-hand side, has been defined as

$$\begin{aligned} \mathcal{E}_d^{-y, -x} f(t, y, x) &:= \left( \sum_{i_x=1}^x \sum_{i_y=1}^y E_d^{-i_y, i_x} - [1 - p_0^X(0)][1 - p_0^Y(0)] \right) f(t, y, x) \\ &= \sum_{i_x=1}^x \sum_{i_y=1}^y p_0^X(i_x) p_0^Y(i_y) f(t, y - i_y, x - i_x) - [1 - p_0^X(0)][1 - p_0^Y(0)] f(t, y, x). \end{aligned}$$

The protracted dose is assumed to be delivered up to a finite time  $T_{irr} < \infty$ , beyond which no irradiation is considered and the system evolves according to (9).

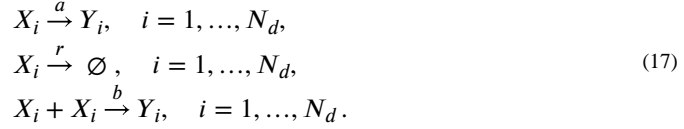
#### D. The diffusive cell nucleus model for GSM<sup>2</sup>

At the beginning of Sec. III, we investigated the time evolution for lethal and sublethal lesions in the cell nucleus. As we discussed above, one of the major weaknesses of the standard MKM and its extensions is the choice of the cell domains [31]. In fact, too small domains translate into a null probability of double events, whereas too big domains imply that distant lesions may combine to produce a lethal lesion. To overcome this problem, the cell nucleus is split into several domains so that the time evolution in each domain can be considered independently. Further, in the following treatment we ameliorate the above limitations by allowing domains interaction and variability in shape and dimension.

In the current section, we will show how the MME (16) can be extended to include interactions between the domains. To keep the treatment as clear as possible, no protracted

dose will be considered. The general case of a continuous irradiation can easily be included in the following treatment via arguments analogous to the ones used in Sec. III C.

Let us consider  $N_d$  domains (referred to also as voxels) that can undergo one of the following possible reactions:



A reasoning similar to the one carried out at the beginning of Sec. III leads to the following MME:

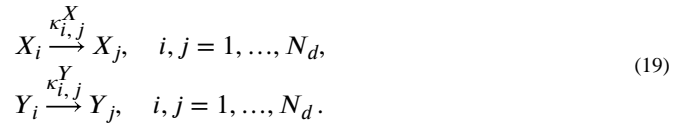
$$\begin{aligned} \partial_t p(t, y, x) &= \sum_{i=1}^{N_d} \mathcal{E}_i^{-1,2} [x_i(x_i - 1)bp(t, y, x)] \\ &+ \sum_{i=1}^{N_d} \left( \mathcal{E}_i^{-1,1} [x_i ap(t, y, x)] + \mathcal{E}_i^{0,1} [x_i rp(t, y, x)] \right). \end{aligned} \quad (18)$$

In Eq. (18), the variables  $x$  and  $y$  are  $N$ -dimensional vectors with the  $i$ th components given by  $x_i$  and  $y_i$ , representing the number of sublethal or lethal lesions, respectively, within the  $i$ th domain ( $i = 1, \dots, N$ ).

**Remark III.1.**—To keep the notation as simple as possible, in with the  $i$ th components given by  $x_i$  and  $y_i$ , representing the Eq. (17) we chose the rates  $a$ ,  $b$ , and  $r$  independent of the donumber of sublethal or lethal lesions, respectively, within the main. Similar results would be obtained with voxel-dependent rates  $a_i$ ,  $b_i$ , and  $r_i$ ,  $i = 1, \dots, N_d$ .

Empirical evidence shows that the lesions, together with interacting within the same voxel, may also move to a different voxel. In fact, lesion spatial movement inside a cell has been demonstrated to be significantly higher than the typical voxel size [50]. To account for this behavior, we will add an additional term to the MME (18).

Besides reactions considered in Eq. (17), we now assume further the following:



**Remark III.2.**—We assumed possible interactions also between nonadjacent domains. If the reactions described by Eq. (19) are to be intended as lesion movements inside the cell nucleus, the most reasonable choice for the interaction rates is to set

$$\kappa_{i,j}^X = \kappa_{i,j}^Y = 0$$



for  $j \notin \Gamma_i$ , where  $\Gamma_i$  is the set of adjacent domains to  $i$ .

Following the same process described in Sec. III, we obtain the MME,

$$\begin{aligned} \partial_t p(t, y, x) &= \sum_{i=1}^N \mathcal{E}_i^{-1,2} [x_i(x_i - 1)bp(t, y, x)] \\ &+ \sum_{i=1}^N \left( \mathcal{E}_i^{-1,1} [x_i ap(t, y, x)] + \mathcal{E}_i^{0,1} [x_i r p(t, y, x)] \right) \\ &+ \sum_{i,j=1}^N X \mathcal{E}_{i,j}^{-1,1} [x_i \kappa_{i,j}^X p(t, y, x)] + \sum_{i,j=1}^N Y \mathcal{E}_{i,j}^{Y,-1,1} [y_i \kappa_{i,j}^Y p(t, y, x)], \end{aligned} \quad (20)$$

where the operators are defined as

$$X \mathcal{E}_{i,j}^{-1,1} f(t, y, x) = \left( E_i^{0,1} E_j^{0,-1} - 1 \right) f(t, y, x),$$

$$Y \mathcal{E}_{i,j}^{-1,1} f(t, y, x) = \left( E_i^{1,0} E_j^{-1,0} - 1 \right) f(t, y, x).$$

The first two lines of Eq. (20) account for reactions within the same voxel, whereas the last line arises from movements between adjacent domains.

Using the same approach for modeling the initial damage distribution (Sec. III B), the resulting MME reads

$$\begin{aligned} \partial_t p(t, y, x) &= \sum_{i=1}^{N_d} \mathcal{E}_i^{-1,2} [x_i(x_i - 1)bp(t, y, x)] + \sum_{i=1}^{N_d} \left( \mathcal{E}_i^{-1,1} [x_i ap(t, y, x)] + \mathcal{E}_i^{0,1} [x_i r p(t, y, x)] \right) \\ &+ \sum_{i,j=1}^{N_d} X \mathcal{E}_{i,j}^{-1,1} [x_i \kappa_{i,j}^X p(t, y, x)] + \sum_{i,j=1}^{N_d} Y \mathcal{E}_{i,j}^{Y,-1,1} [y_i \kappa_{i,j}^Y p(t, y, x)], \end{aligned}$$

$$p(0, y, x) = \prod_{i=1}^N p_{0,i}^X(x_i) p_{0,i}^Y(y_i), \quad (21)$$

where  $p_{0,i}^X(x_i) p_{0,i}^Y(y_i)$  denotes the initial distribution for the voxel  $i$  as computed in Eqs. (14) and (15).

## E. Survival probability

Cell survival is one of the most relevant biological end points in radiobiology, and it is defined as the probability for a cell to survive radiation exposure, mostly measured by its ability to form clonogens, i.e., to retain its reproductive potential. Taking into account assumption (v), no lethal lesions must be present in the cell nucleus after a sufficiently large time has passed from the irradiation. An estimate of cell survival can be obtained from the

solution to the MME (9). In this study, we will focus on a single domain, because the calculations for the entire cell are completely analogous.

The survival probability for the domain  $d$  under the assumptions (i)–(v) introduced above is defined as

$$S^Y := \mathbb{P}\left(\lim_{t \rightarrow \infty} Y(t) = 0\right). \quad (22)$$

To assess the survival probability, the limiting long-time distribution for the MME (9) must be studied.

From an heuristic perspective, since the number of sublethal lesions can only decrease, the points  $\{(y, 0) : y \in \mathbb{N}_0\}$  are absorbing states. Furthermore, the system reaches an absorbing state in a finite time with probability 1, converging toward a limiting stationary distribution. By *absorbing state* we mean that once the system reaches the point  $(y, 0)$ , it stays there and future evolutions are no longer considered.

Due to the high generality of the GSM<sup>2</sup> model, especially because no detailed balance is satisfied and no explicitly conserved quantities can be obtained, the closed form for the limiting distribution is not easily computable. For this reason, in the present work the survival probability will be computed from the corresponding master equation numerical solution as

$$S^Y = \lim_{t \rightarrow \infty} p(t, 0, 0).$$

In forthcoming developments, we will study in more detail the survival probability resulting from the proposed GSM<sup>2</sup> model and its explicit form. In general, it is worth mentioning that besides the numerical approach, such as the one used here, and the analytical approach in which the survival probability is explicitly computed, an efficient approach is to introduce suitable approximations in the driving equation so that a formal expansion of the survival probability can be computed [36].

#### IV. NUMERICAL IMPLEMENTATION

To calculate a numerical solution to the MME (9), the following steps are performed:

- (i) We choose the number  $N_d$  of domains in which the cell nucleus is divided. As GSM<sup>2</sup> does not rely on any specific assumption for the probability distribution, the domains do not need to be assumed to be of equal size. For each domain, the *single event* energy deposition distribution  $f_{1;d}(z)$  is obtained with GEANT4 [40] simulations.
- (ii) The number of lethal and sublethal lesions are sampled from the distributions  $p_0^X(x)$  and  $p_0^Y(x)$  as derived in Eq. (15). The standard assumption is that  $p_z^X(p_z^Y)$  is a Poisson distribution of mean  $\kappa_z(\lambda_z)$ . Given the general setting, we will compare the results with an initial Gaussian distribution of different possible variances.

(iii) Given the initial number of lesions, the evolution paths are simulated via the *stochastic simulation algorithms* (SSA) (Ref. [51], Chap. 13).

(iv) Steps (i)–(iii) are repeated to obtain the Monte Carlo empirical distribution of lethal and sublethal lesions over the cell nucleus.

(v) The survival probability in the single domain as well as the cell nucleus are calculated from the empirical distribution obtained in step (iv).

Previous steps can be computed independently for each domain if no interaction between domains is assumed or the paths for the whole nucleus can be estimated simultaneously, in the case of a dependent-voxel model. The computational effort for the latter is substantially higher. It should be noted here that developing an efficient simulation algorithm is beyond the scope of the present work, and we refer to [52] for a review of possible simulation algorithms.

### A. The numerical solution

The present section is devoted to finding and discussing the numerical solution of MME derived in Sec. III. In particular, the full master equation (9) is solved via the *stochastic simulation algorithms* (SSA) (Ref. [51], Chap. 13), so that the density is estimated with a Monte Carlo simulation. We simulate  $10^6$  events, and the density function is thus reconstructed empirically.

The goal of this section is also to highlight how a different setting affects the density distribution of the lesions. In particular, it will emerge how the density distribution resulting from the corresponding master equation changes for different lesion evolution parameters, initial probabilistic conditions, and irradiation conditions.

To assess the energy deposited on the domain, we used the microdosimetry approach as discussed in Sec. III B. With GEANT4, we simulated microdosimetric spectra of a 20 MeV/u carbon ion beam traversing a 1.26-cm-diam sphere filled with pure propane gas with a low density ( $1.08 \times 10^{-4} \text{ g/cm}^3$ ), such that the energy depositions are equivalent to those in  $2 \mu\text{m}$  of tissue. This geometry reproduces a standard tissue equivalent proportional counter (TEPC) as used, for example, in [53]. Specific energies acquired with the TEPC are then converted to the domain size of interest as reported in Ref. [29] (Sec. 2). The choice to simulate a microdosimeter has been made with the aim of remaining as consistent as possible with real experiments. In addition, carbon ions have been chosen since the existing model fails at predicting relevant radiobiological end points under high-LET regimes.

In the calculations, we consider high doses, so that multievent distributions as described in Sec. III B are computed for  $z_n \gg 1$ . This choice is due to the fact that the plotted distributions refer to a single-cell nucleus domain, and thus to highlight differences at such a small scale, a high dose needs to be considered. At lower doses, differences between the MME solution for a single nucleus domain for different parameters are more difficult to appreciate. Nonetheless, small differences at the domain level can translate into relevant dissimilarities at the macroscopic level.

Figure 1 reports different path realizations for the lethal and sublethal evolution; the stochastic paths are also compared to the mean value, which evolves according to the MKM kinetic equations (1). The plots indicate that the mean value cannot be representative of the whole path realization distribution.

Figure 2 shows the master equation solution at different times. The left panels show the contour plots of the joint probability distributions of lethal and sublethal damages, together with their marginal distributions depicted along the axis. The right panels are 3D representations of the density function solutions. Yellow regions represent values with higher probability, whereas purple regions denote values for which the probability is close to 0. At a starting time  $t_1$ , there is a high variability in the number of reparable lesions while small fluctuations are present in the number of lethal lesions. As time increases to  $t_2$  it can be seen that variability in the number of lethal lesions increases. At a later time  $t_3$ , instead, the situation is the exact opposite, with a greater variability in the number of lethal lesions against small fluctuations in the number of sublethal lesions.

Figure 3 compares lethal and sublethal lesion distributions for different types of irradiation conditions, namely acute dose delivery at initial time, split dose at uniform time steps, and protracted dose according to Eq. (16). A split dose at uniform times yields a rather similar lesion distribution to that of a fully stochastic protracted dose irradiation, while the solution differs significantly for the acute dose case. This result is caused by the nonlinear effect that double events have on the probability distribution of the lesions.

The long-time distribution of lethal lesions is compared with a Poisson distribution for different parameters and doses in Fig. 4. At lower doses and for  $b$  negligible with respect to  $r$ , the MME solution is in fact Poissonian (top panel). As the dose increases, the MME solution can be non-Poissonian even if  $r$  dominates  $b$  (middle panel). Finally, for higher doses and higher  $b$ , the MME solution differs significantly from a Poisson distribution (bottom panel).

## B. Effect of the initial law on the lethal lesions distribution and cell survival

The goal of the present section is to emphasize the dependence on the initial law of the long-time lethal lesion distribution, showing that the marginal distribution of the lethal lesions might differ from the Poisson distribution that is typically assumed.

We considered different initial conditions for Eq. (15). In particular, the following initial distributions were selected for  $p_z^X(x|\kappa_z)$  and  $p_z^Y(y|\lambda_z)$ : (i) a Poisson random variable with mean value  $\mu$ , and (ii) a Gaussian with mean value  $\mu$  and variance between  $0.5\mu$  and  $1.5\mu$ . The mean value  $\mu$  has been set to  $\lambda_z$  for sublethal lesions and  $\kappa_z$  for lethal lesions. The results are plotted in Fig. 5 and indicate that a more peaked initial distribution corresponds to a more peaked long-time distribution, meaning that the initial value can sharpen or broaden lethal and sublethal lesion distributions. This effect has a straightforward consequence on the resulting survival probability shown in Fig. 6.

We test both the typically used Poisson initial distribution and a Gaussian random variable with different variance. Figure 5 shows the comparison of lethal and sublethal lesion

distributions for different initial conditions. In particular, an initial datum has been taken to be a Poisson random variable with mean value  $\mu$ . Additionally, the case of an initial distribution to be Gaussian and with mean value  $\mu$  and variance  $0.5\mu$  and  $1.5\mu$  has been considered. The mean value  $\mu$  has been set to  $\lambda_z$  for sublethal lesions and  $\kappa_z$  for lethal lesions. It can be seen how the initial value can sharpen or broaden lethal and sublethal lesion distributions, with a straightforward consequence on the resulting survival probability; see Fig. 6.

Survival probability is one of the most used and relevant radiobiological observables. Figure 6 highlights how a different initial condition affects the resulting survival curve. In particular, it is important to notice that the probability of survival rises or falls in the high-dose region. One of the major flaws in classical models, with particular reference to the linear-quadratic model, is the fact that it significantly underestimates the probability of survival for high doses. In particular, it can be seen how the resulting survival probability differs from the classical linear quadratic survival. For low-dose irradiation, it emerges how the survival probability exhibits a linear-quadratic behavior. On the contrary, for higher doses the survival curve maintains linear patterns as shown by experimental data. Given the deep importance of the survival probability, this topic is currently being investigated in more detail, and it will be the subject of future research.

## V. CONCLUSION

The present work represents a step toward an advanced and systematic investigation of the stochastic nature of energy deposition by particle beams, with a particular focus on how it affects DNA damage. Starting from basic probabilistic assumptions, a *master equation* for the probability distribution of the number of lethal and sublethal lesions induced by radiation of a cell nucleus has been derived. The model, called the Generalized Stochastic Microdosimetric Model (GSM<sup>2</sup>), provides a simple and yet fundamental generalization of all existing models for DNA-damage prediction, being able to truly describe the stochastic nature of energy deposition. This advance results in a more general description of DNA-damage formation and time-evolution in a cell nucleus for different irradiation scenarios, from which radiobiological outcomes can be assessed.

Most of the existing models assume a Poissonian distribution of lethal damage, ignoring the true space-time stochastic nature of energy deposition. To overcome the limits of this assumption, *ad hoc* corrections have been introduced, called non-Poissonian corrections in the literature, but to the best of our knowledge an extensive survey on the complete stochasticity of biophysical processes has never been carried out.

This work aims at highlighting how the stochastic nature of energy deposition can lead to different cell survival estimations, and how non-Poissonian effects emerge naturally in the general setting developed. Remarkably, in particular, GSM<sup>2</sup> does not require any *ad hoc* corrections for taking into account overkill effects, in contrast to prior models.

In a separate work, we will focus on verification and optimization of the prediction of the survival curves for different systems, i.e., radiation type, irradiation conditions, and cell line.

In addition, given the general nature of the proposed model, closed-form solutions for lesion distribution and survival curve are typically difficult to obtain. However, it is fair to say that, due to the several processes involved, approximation methods provide powerful tools to estimate several quantities of interests. Among the most important approximation methods, we mention system size expansions [33,36,54,55] and the related small-noise asymptotic expansions [36]. Both approaches will be investigated in future research to provide accurate estimates of several biological end points, such as cell survival.

Further investigation will also be devoted to developing a more efficient numerical implementation of the driving *master equation*.

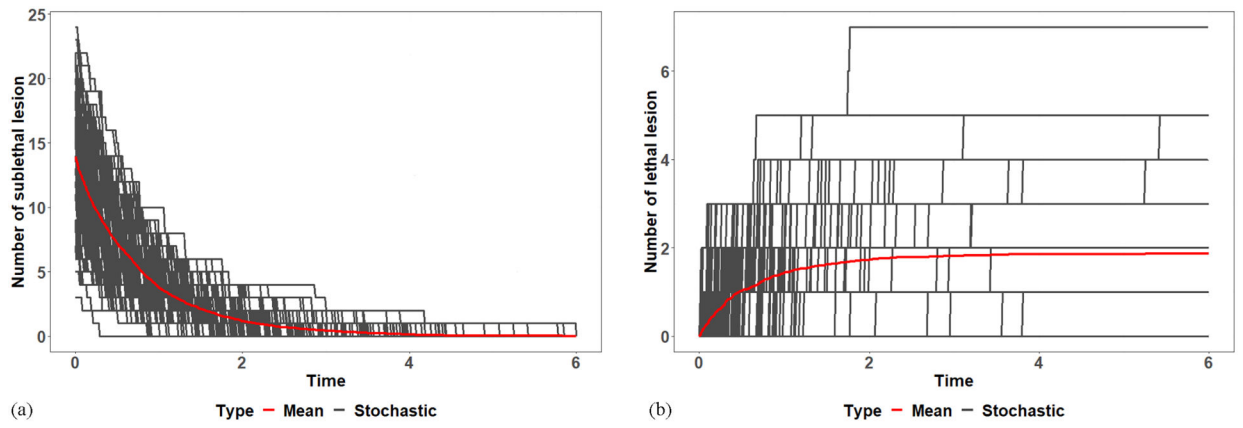
## ACKNOWLEDGMENTS

This work was partially supported by the INFN CSNV projects MoVe-IT and NEPTune.

## References

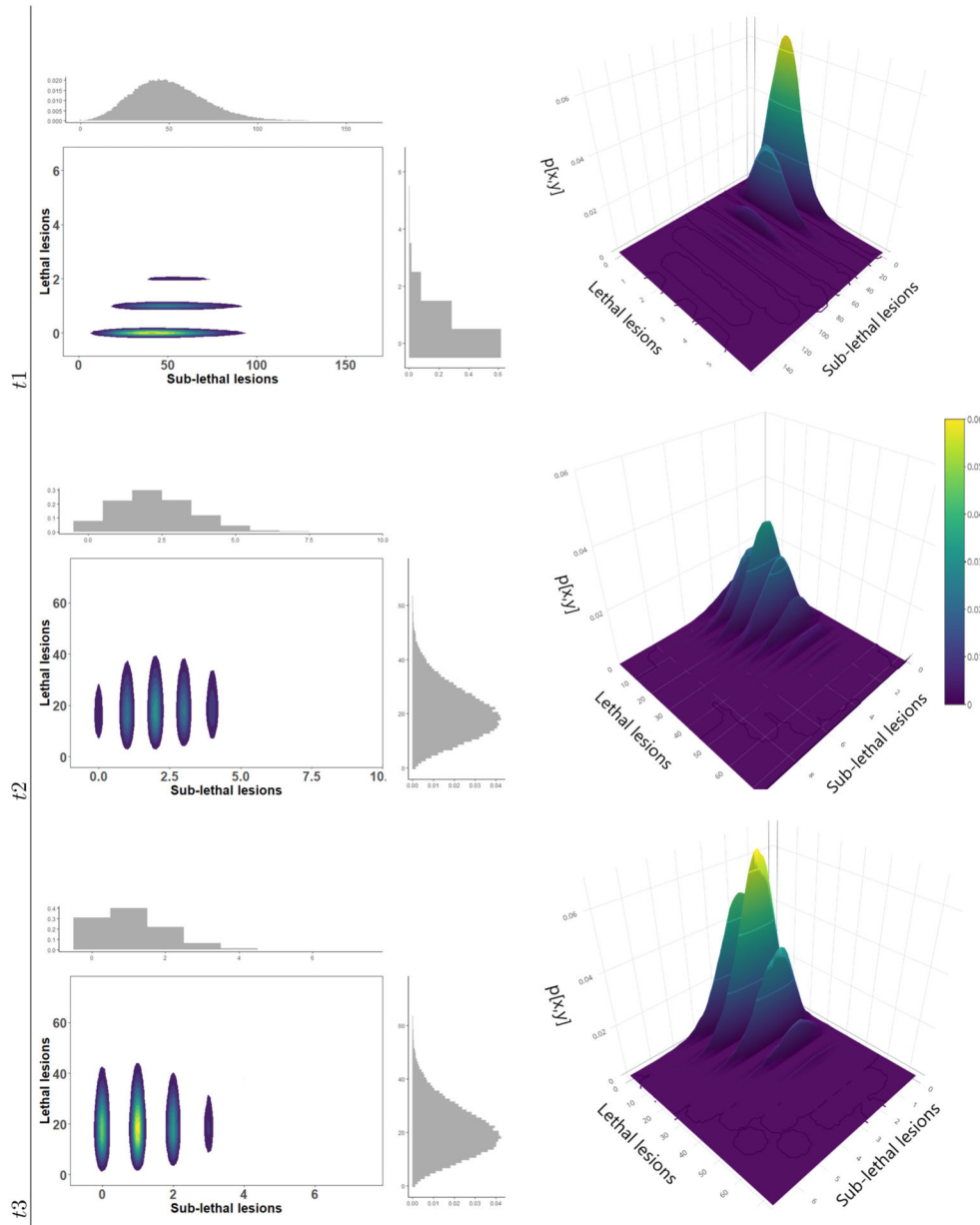
- [1]. Durante M and Loeffler JS, Nat. Rev. Clin. Oncol 7, 37 (2010). [PubMed: 19949433]
- [2]. PTCOG, “Ptcog website” (2020), <https://www.ptcog.ch/index.php/ptcog-patient-statistics>.
- [3]. Durante M and Paganetti H, Rep. Prog. Phys 79, 096702 (2016). [PubMed: 27540827]
- [4]. Schardt D, Elsässer T, and Schulz-Ertner D, Rev. Mod. Phys 82, 383 (2010).
- [5]. Tinganelli W and Durante M, Cancers 12, 3022 (2020).
- [6]. Krämer M, Jäkel O, Haberer T, Kraft G, Schardt D, and Weber U, Phys. Med. Biol 45, 3299 (2000). [PubMed: 11098905]
- [7]. Kraft G, Prog. Part. Nucl. Phys 45, S473 (2000).
- [8]. Scifoni E, Mod. Phys. Lett. A 30, 1540019 (2015).
- [9]. Plante I, Ponomarev A, and Cucinotta FA, Radiat. Prot. Dosim 143, 156 (2011).
- [10]. Plante I, Radiat. Environ. Biophys 50, 389 (2011). [PubMed: 21562854]
- [11]. Plante I, Radiat. Environ. Biophys 50, 405 (2011). [PubMed: 21594646]
- [12]. Surdutovich E and Solov'yov AV, Phys. Rev. E 82, 051915 (2010).
- [13]. Toulemonde M, Surdutovich E, and Solov'yov AV, Phys. Rev. E 80, 031913 (2009).
- [14]. Surdutovich E, Gallagher DC, and Solov'yov AV, Phys. Rev. E 84, 051918 (2011).
- [15]. Zaider M, Rossi BHH, and Zaider M, Microdosimetry and Its Applications (Springer, Berlin Heidelberg, 1996).
- [16]. Hawkins RB, Radiat. Res 140, 366 (1994). [PubMed: 7972689]
- [17]. Hawkins RB, Radiat. Res 189, 104 (2018). [PubMed: 29045193]
- [18]. Kellerer AM and Rossi HH, Radiat. Res 75, 471 (1978).
- [19]. Zaider M and Rossi H, Radiat. Res 83, 732 (1980). [PubMed: 7413933]
- [20]. Curtis SB, Radiat. Res 106, 252 (1986). [PubMed: 3704115]
- [21]. Tobias CA, Radiat. Res 104, S77 (1985).
- [22]. Elsässer T, Krämer M, and Scholz M, Int. J. Radiat. Oncol. Biol. Phys 71, 866 (2008). [PubMed: 18430521]
- [23]. Pfuhl T, Friedrich T, and Scholz M, Radiat. Res 193, 130 (2020). [PubMed: 31804150]
- [24]. Hawkins RB, Radiat. Res 160, 61 (2003). [PubMed: 12816524]
- [25]. Kase Y, Kanai T, Matsumoto Y, Furusawa Y, Okamoto H, Asaba T, Sakama M, and Shinoda H, Radiat. Res 166, 629 (2006). [PubMed: 17007551]
- [26]. Inaniwa T, Suzuki M, Furukawa T, Kase Y, Kanematsu N, Shirai T, and Hawkins RB, Radiat. Res 180, 44 (2013). [PubMed: 23768075]
- [27]. Inaniwa T, Furukawa T, Kase Y, Matsufuji N, Toshito T, Matsumoto Y, Furusawa Y, and Noda K, Phys. Med. Biol 55, 6721 (2010). [PubMed: 21030747]

- [28]. Sato T and Furusawa Y, *Radiat. Res* 178, 341 (2012). [PubMed: 22880622]
- [29]. Bellinzona VE, Attili A, Cordoni F, Missiaggia M, Tommasino F, Scifoni E, and La Tessa C, arXiv:2007.07673 [Frontiers in Physics (to be published)].
- [30]. Radman M, *DNA Repair* 44, 186 (2016). [PubMed: 27264559]
- [31]. Smith S and Grima R, *Bull. Math. Biol* 81, 2960 (2019). [PubMed: 29785521]
- [32]. Chatzipapas KP, Papadimitroulas P, Emfietzoglou D, Kalospyros SA, Hada M, Georgakilas AG, and Kagadis GC, *Cancers* 12, 799 (2020).
- [33]. Van Kampen NG, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992), Vol. 1.
- [34]. Solov'yov AV, Surdutovich E, Scifoni E, Mishustin I, and Greiner W, *Phys. Rev. E* 79, 011909 (2009).
- [35]. Plante I, Devroye L, and Cucinotta FA, *Comput. Phys. Commun* 185, 697 (2014).
- [36]. Gardiner CW et al., *Handbook of Stochastic Methods*, 3rd ed. (Springer, Berlin, 1985).
- [37]. Weber MF and Frey E, *Rep. Prog. Phys* 80, 046601 (2017). [PubMed: 28306551]
- [38]. Isaacson SA, *SIAM J Appl. Math* 70, 77 (2009).
- [39]. Hellander S, Hellander A, and Petzold L, *Phys. Rev. E* 85, 042901 (2012).
- [40]. Agostinelli S, Allison J, Amako KA, Apostolakis J, Araujo H, Arce P, Asai M, Axen D, Banerjee S, Barrand G et al., *Nucl. Instrum. Methods Phys. Res., Sect. A* 506, 250 (2003).
- [41]. Bodgi L, Canet A, Pujo-Menjouet L, Lesne A, Victor J-M, and Foray N, *J. Theor. Biol* 394, 93 (2016). [PubMed: 26807808]
- [42]. McMahon SJ, *Phys. Med. Biol* 64, 01TR01 (2018).
- [43]. Missiaggia M, Pierobon E, Castelluzzo M, Perinelli A, Cordoni F, Vignali MC, Borghi G, Bellinzona V, Scifoni E, Tommasino F et al., *Front. Phys* (to be published), arXiv:2007.06276.
- [44]. Manganaro L, Russo G, Cirio R, Dalmasso F, Giordanengo S, Monaco V, Muraro S, Sacchi R, Vignati A, and Attili A, *Med. Phys* 44, 1577 (2017). [PubMed: 28130821]
- [45]. Parke W, *Phys. Rev. E* 56, 5819 (1997).
- [46]. Hawkins RB and Inaniwa T, *Radiat. Res* 182, 72 (2014). [PubMed: 24937780]
- [47]. Hawkins RB and Inaniwa T, *Radiat. Res* 180, 584 (2013). [PubMed: 24191898]
- [48]. Chang DS, Lasley FD, Das IJ, Mendonca MS, and Dynlacht JR, in *Basic Radiotherapy Physics and Biology* (Springer, New York, 2014), pp. 235–240.
- [49]. Loan M, Alameen M, Bhat A, and Tantary M, arXiv:2009.06802
- [50]. Schettino G, Ghita M, Richard D, and Prise K, *Radiat. Prot. Dosim* 143, 340 (2011).
- [51]. Weinan E, Li T, and Vanden-Eijnden E, *Applied Stochastic Analysis* (American Mathematical Society, Providence, Rhode Island, 2019), Vol. 199.
- [52]. Simoni G, Reali F, Priami C, and Marchetti L, *Wiley Interdisc. Rev.: Syst. Biol. Med* 11, e1459 (2019).
- [53]. Missiaggia M, Cartechini G, Scifoni E, Rovituso M, Tommasino F, Verroi E, Durante M, and La Tessa C, *Phys. Med. Biol* 65, 245024 (2020). [PubMed: 32554886]
- [54]. Kurtz TG, in *Stochastic Systems: Modeling, Identification and Optimization, I* (Springer, Amsterdam: North-Holland, 1976), pp. 67–78.
- [55]. Meleard S and Bansaye V, *Stochastic Models for Structured Populations: Scaling Limits and Long Time Behavior* (Springer, Cham, 2015), Vol. 1.

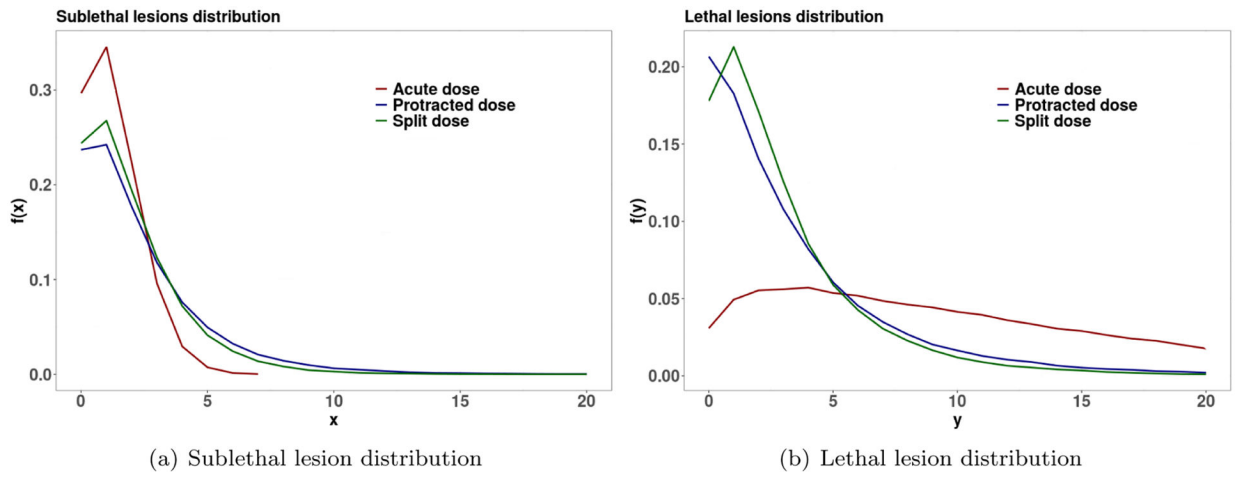


**FIG. 1.** Sublethal lesions (a) and lethal lesions (b) evolution. GSM<sup>2</sup> parameters were set to  $r = 1$ ,  $a = 0.1$ , and  $b = 0.01$ . The red line represents the average value.

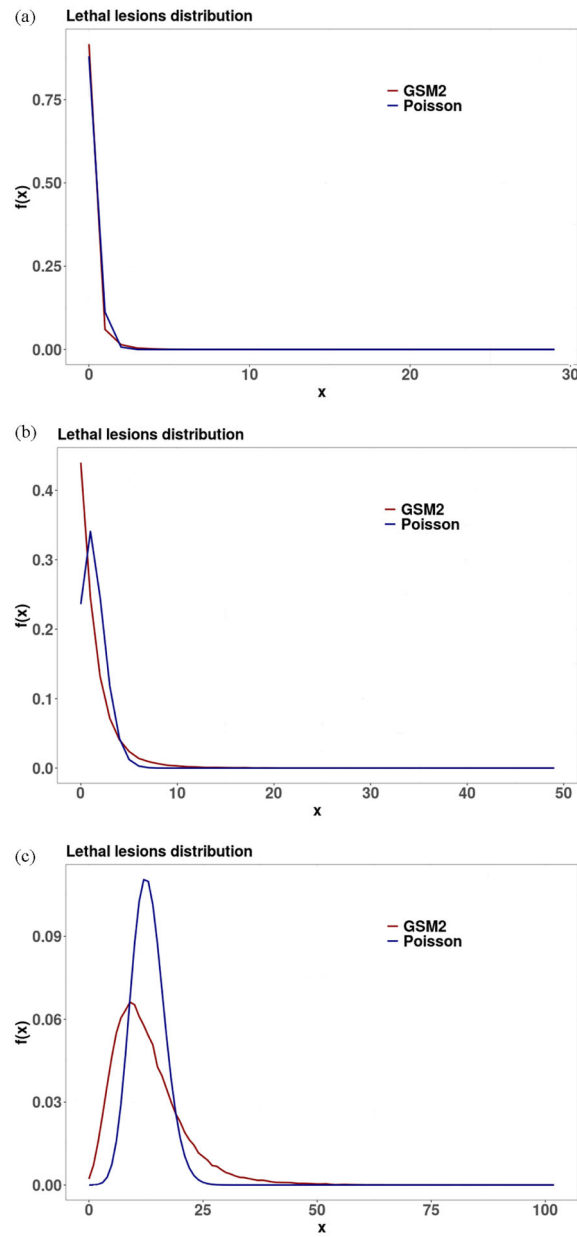




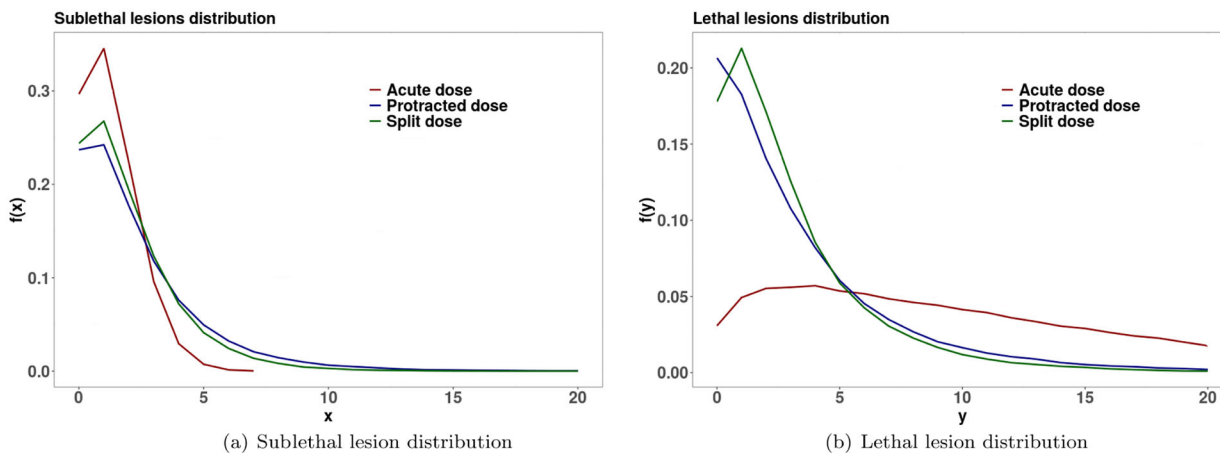
**FIG. 2.** Master equation solution at time  $t = 1$  a.u. (top panel),  $t = 100$  a.u. (middle panel), and  $t = 150$  a.u. GSM<sup>2</sup> parameters were set to  $r = 1$ ,  $a = 0.2$ , and  $b = 0.1$ . The left panels report the contour plots of the joint probability distributions of lethal and sublethal damages, with the marginal distributions depicted along the axis. The right panels are 3D plots for the density function. Yellow regions represent values with higher probability, whereas purple regions denote values for which the probability is close to 0.



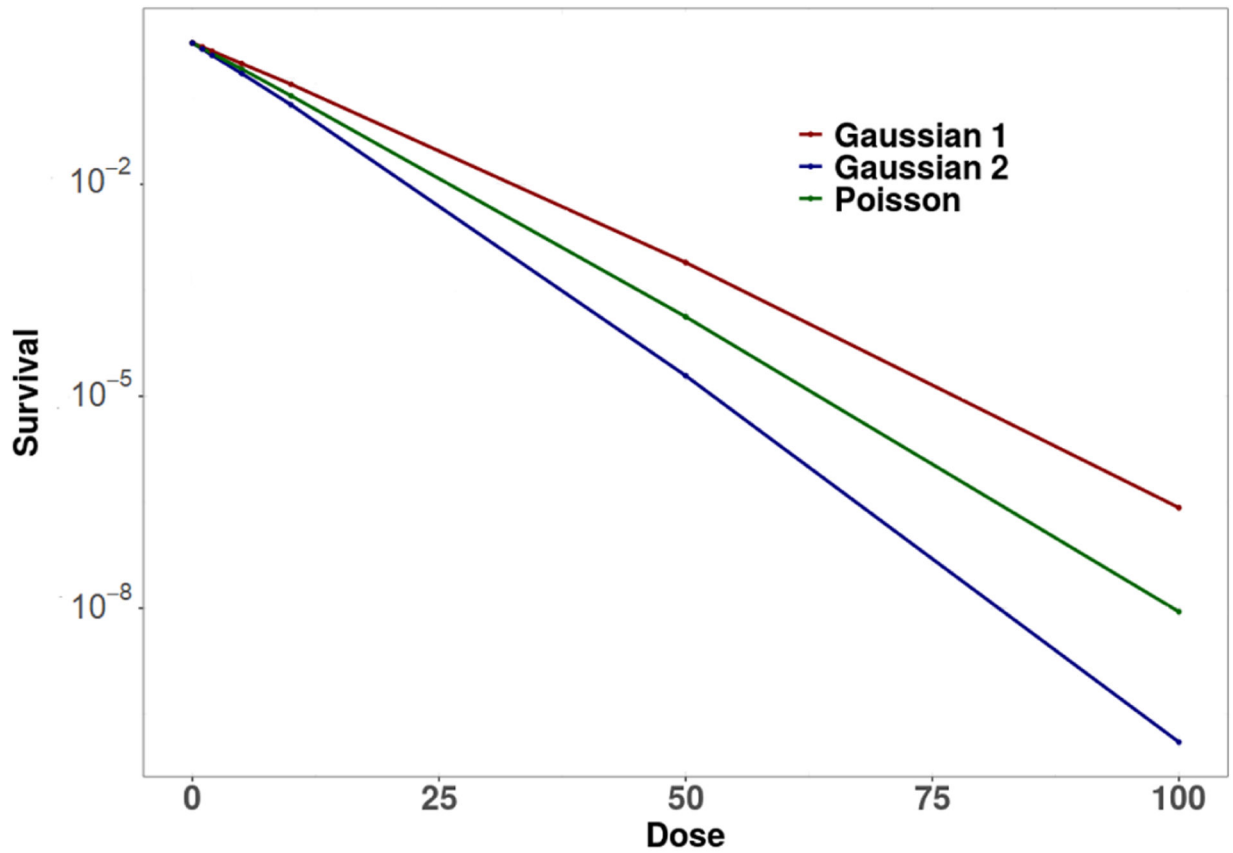
**FIG. 3.** Master equation solution for acute, split, and protracted doses of 100 Gy. GSM<sup>2</sup> parameters were set to  $r = 1$ ,  $a = 0.2$ , and  $b = 0.1$ .



**FIG. 4.** Comparison of long-time lethal lesion distributions and Poisson distributions. (a) dose = 5 Gy,  $r = 1$ ,  $a = 0.1$ , and  $b = 0.01$ . (b) dose = 100 Gy,  $r = 5$ ,  $a = 0.1$ , and  $b = 0.01$ . (c) dose = 150 Gy,  $r = 5$ ,  $a = 0.2$ , and  $b = 0.1$ .



**FIG. 5.** Lethal and sublethal lesion distribution depending on the chosen initial distribution at time  $t1 = 1$  a.u. and  $\beta = 150$  a.u. The initial distributions  $p_z^X$  and  $p_z^Y$  have been chosen as a Poisson distribution of mean  $\mu = \{\lambda_z, \kappa_z\}$  or as a Gaussian distribution with mean  $\mu = \{\lambda_z, \kappa_z\}$  and variance  $\sigma^2 \in \{0.5\mu, \dots, 1.5\mu\}$ . The MME parameters were set to  $r = 1$ ,  $a = 0.2$ , and  $b = 0.1$ .



**FIG. 6.**

Cell survival function calculated for different initial conditions. The initial distributions  $p_z^X$  and  $p_z^Y$  have been chosen as a Poisson distribution of mean  $\mu = \{\lambda_z, \kappa_z\}$ , or as a Gaussian distribution with mean  $\mu = \{\lambda_z, \kappa_z\}$  and variance  $\sigma^2 \in \{0.5\mu, \dots, 1.5\mu\}$ . The MME parameters were set to  $r = 1$ ,  $a = 0.2$ , and  $b = 0.1$ .