Relative frequencies of constrained events in stochastic processes: An analytical approach

S. Rusconi,1,* E. Akhmatskaya,1,2 D. Sokolovski,2,3 N. Ballard,4 and J. C. de la Cañ1
1BCAM - Basque Center for Applied Mathematics, Alameda de Mazarredo 14, 48009 Bilbao, Bizkaia, Spain
2Ikerbasque, Basque Foundation for Science, E-48013 Bilbao, Spain
3Departamento de Química-Física, Universidad del País Vasco, UPV/EHU, Leioa, Spain
4POLYMAT and Grupo de Ingeniería Química, Dpto. de Química Aplicada, University of the Basque Country UPV/EHU,
Joxe Mari Korta Zentroa, Tolosa Etorbidea 72, 20018 Donostia/San Sebastián, Spain

(Received 5 June 2015; published 19 October 2015)

The stochastic simulation algorithm (SSA) and the corresponding Monte Carlo (MC) method are among the most common approaches for studying stochastic processes. They rely on knowledge of interevent probability density functions (PDFs) and on information about dependencies between all possible events. Analytical representations of PDFs are difficult to specify in advance, in many real life applications. Knowing the shapes of PDFs, and using experimental data, different optimization schemes can be applied in order to evaluate probability density functions and, therefore, the properties of the studied system. Such methods, however, are computationally demanding, and often not feasible. We show that, in the case where experimentally accessed properties are directly related to the frequencies of events involved, it may be possible to replace the heavy Monte Carlo core of optimization schemes with an analytical solution. Such a replacement not only provides a more accurate estimation of the properties of the process, but also reduces the simulation time by a factor of order of the sample size (at least \( \approx 10^4 \)). The proposed analytical approach is valid for any choice of PDF. The accuracy, computational efficiency, and advantages of the method over MC procedures are demonstrated in the exactly solvable case and in the evaluation of branching fractions in controlled radical polymerization (CRP) of acrylic monomers. This polymerization can be modeled by a constrained stochastic process. Constrained systems are quite common, and this makes the method useful for various applications.

DOI: 10.1103/PhysRevE.92.043306

PACS number(s): 02.70.-c, 05.40.-a, 82.20.Fd, 82.35.—x

I. INTRODUCTION

Simulation of stochastic processes is a powerful tool for modeling and describing the evolution of various phenomena in natural and human-made systems. A well known example of such modeling is the stochastic simulation algorithm (SSA) developed by Gillespie [1]. In the physics community, it is also known as the \( n \)-fold way, introduced by Bortz et al. [2]. SSA is a Monte Carlo (MC) based method: it draws multiple realizations of the process and then computes statistics on them. From now on, we will refer to SSA as the MC method. This approach is based on the assumption that the studied system is well mixed, and also memoryless. These assumptions lead to independent exponentially distributed interevent-time probability density functions [3]. As intuition suggests, this set of hypotheses does not hold for all phenomena of practical interest. One such example is a constrained stochastic process, where the occurrence of some events may depend on the previous history of the process (see [4] for examples of constrained stochastic processes). In this cases, the dependencies can be realized either explicitly by introducing constraints in the SSA algorithm or implicitly through a modification of interevent-time probability density functions (PDFs).

Either way, well posed formulations of stochastic processes require the knowledge of interevent-time probability density functions. One can assume a particular functional shape for these PDFs, but meaningful values of the PDF parameters are also needed, in order to complete the description of the model and run the corresponding Monte Carlo simulation. In practice, appropriate PDF parameters are difficult to identify.

One possible way to estimate the unknown parameters is to employ an optimization scheme, which uses available experimental data. The idea behind a fitting scheme is to build a cost function \( J \), which measures the disagreement between the experimental data and the data obtained by the proposed MC method. A fitting scheme seeks to minimize the cost function \( J \), in order to find the set of PDF parameters which gives the best agreement with the experimental data.

Since the stochastic simulation is a part of the cost function evaluation, multiple runs are needed until a good fitting is obtained. Thus, regardless of the choice of the optimization route and of the particular cost function, any MC based fitting scheme is computationally expensive.

In this paper we derive an analytical approach for evaluation of \( J \), as a function of PDF parameters, without running the corresponding MC simulation. The method is free of the statistical errors affecting Monte Carlo based simulations. It can be incorporated into fitting schemes used to study various phenomena for which the experimentally observed properties are directly related to frequencies of the events involved. Examples of such phenomena include branching rate reduction in controlled radical polymerization [5], high-frequency pulsed laser polymerization of acrylates [6], reduction in rate of polymerization in reversible addition-fragmentation chain transfer radical polymerizations [7,8], and the absence of chain transfer to poly(vinyl alcohol) in single electron transfer-degenerative chain transfer living radical polymerization of poly(vinyl chloride) in aqueous media [9].

Our goal is to provide analytical expressions for the asymptotic frequencies of events involved in constrained stochastic

\[\text{*srusconi@bcamath.org}\]
processes, thus avoiding costly Monte Carlo evaluations, used in a fitting scheme. A natural application for the new method is the computation of branching fractions in controlled radical polymerization (CRP) of acrylic monomers [10].

The paper is organized as follows. In Sec. II we propose a general approach for calculation of asymptotic relative frequencies of events in constrained stochastic processes. It is solved analytically in Sec. III for a process with a single constraint, and the results are compared with the ones earlier provided in [11]. CRP and the formation of branches in CRP of acrylics is briefly described in Sec. IV. In Sec. IV A the MC fitting method earlier proposed for evaluation of the branching fraction in the CRP [12] is summarized. Our alternative analytical approach is developed in Sec. IV B. The results of comparison between two different methodologies are provided in [11]. CRP and the formation of branches in CRP is solved analytically in Sec.III for a process with a single constraint, and the results are compared with the ones earlier [12]. Our objective is to compute ratios between asymptotic frequencies of events into 2

\[ \text{II. RELATIVE ASYMPTOTIC FREQUENCIES OF EVENTS IN CONSTRAINED STOCHASTIC PROCESSES} \]

Consider a stochastic process with a given total number of events \( n_T \gg 1 \). These events correspond to the realizations of the random variables building the stochastic process. Each event may have \( N \) possible outcomes. A particular outcome is given by a possible value assumed by the random variables. The outcomes are constrained: i.e., the outcome \( i = 1, \ldots, N \) can occur if and only if at least \( c_{ij} \) outcomes \( j = 1, \ldots, N \) have already occurred after the previous occurrence of the \( i \)th outcome.

Our objective is to compute ratios between asymptotic numbers of occurred outcomes, \( n_i \) and \( n_j \), corresponding to two events of different types, \( i \) and \( j \).

It is worth noting that there are the conditions the system needs to satisfy in order to be able to evolve its state. First, each outcome must not be constrained by itself, or

\[ c_{ii} = 0 \quad \forall i = \{1, \ldots, N\}, \quad (1) \]

\[ c_{ij} = 0 \quad \forall i = \{1, \ldots, N\}, \quad (2) \]

(s.t. is a shorthand for “such that”). Otherwise, if \( c_{ij} \neq 0 \) for all \( i, j \), none of the events are possible.

The \( n_T \) events may have \( 2^N \) nonoverlapping subsets labeled \( (j_1, j_2, \ldots, j_N) \), where \( j_k = 0, 1 \) and \( k = 1, 2, \ldots, N \). \( j_k \) means that the \( k \)th event is possible, whereas if \( j_k = 0 \) it is not possible due to constraints. Let \( n(j_1, j_2, \ldots, j_N) \) be the number of events in a subset; e.g.,

\[ n(1,0,\ldots,0) \]

is the number of events for which the outcome \( i = 1 \) is possible, but the rest of them are not.

Obviously

\[ \sum_{j_1,\ldots,j_N=0,1} n(j_1, j_2, \ldots, j_N) = n_T. \quad (3) \]

Let \( n_i(j_1, j_2, \ldots, j_N) \) be the number of outcomes of a kind \( i \) in the subset \((j_1, j_2, \ldots, j_N)\). The assumption is that the events belonging to each subset \((j_1, j_2, \ldots, j_N)\) are independent and identically distributed (i.i.d.). Such an assumption does not lead to loss of generality because the constraints used in building different subsets can contain the information about the dependencies between events [13].

Then if \( n(j_1, j_2, \ldots, j_N) \rightarrow +\infty \), the probability \( P_i(j_1, j_2, \ldots, j_N) \) for the \( i \)th outcome to occur in the set \((j_1, j_2, \ldots, j_N)\) corresponds to the limit of frequency of the \( i \)th outcome, i.e.,

\[ P_i(j_1, j_2, \ldots, j_N) = \frac{n_i(j_1, j_2, \ldots, j_N)}{n(j_1, j_2, \ldots, j_N)} \quad \forall i \in \{1, \ldots, N\}. \quad (4) \]

A recent formulation of the Gillespie stochastic simulation algorithm (SSA) [3] can be used to define the probability \( P_i(j_1, j_2, \ldots, j_N) \). Let independent random variables \( T_1, T_2, \ldots, T_N \) be the times required for the next occurrence of the respective outcome. The SSA suggests picking the outcome that realizes the minimal occurrence time among the possible ones. In other words,

\[ \text{if } j_k = 0, \quad \text{if } j_k = 1 \quad \forall k \neq i \quad \text{s.t. } j_k = 1; \left\{ \begin{array}{l} j_k = 1, \quad \text{if } j_k = 0 \quad \forall k \neq i \end{array} \right. \]

\[ \text{if } j_k = 0, \quad \text{if } j_k = 1 \quad \forall k \neq i \quad \text{s.t. } j_k = 1; \left\{ \begin{array}{l} j_k = 1, \quad \text{if } j_k = 0 \quad \forall k \neq i \end{array} \right. \]

possible dependencies between different events. In general, complex dependencies between different subsets may lead to nonconverged asymptotic behavior of the ratios of interest. However, when the constraints are limited and well defined, as in the examples presented in the following sections, the ratios of interest can be evaluated exactly for any value of \( N \).

\[ \text{III. A PROCESS WITH A SINGLE CONSTRAINT} \]

Before applying the suggested methodology to the fitting procedure outlined in Sec. I we test it on the simple model introduced in [11]. We consider the case of a stochastic process with only two possible outcomes, 1 and 2. The first one is
free to occur with the occurrence rate $c_1$, but the second one must wait till at least $n_0$ occurrences of kind 1 after its own previous occurrence. Its occurrence rate is $c_2$. There are only two possible subsets in this case (the order of events is preserved as described above):

$$(1,0), (1,1).$$

(8)

The corresponding probabilities are given by

$$
P_1(1,0) = 1, \quad P_1(1,1) = P(T_1 < T_2),
$$

$$
P_2(1,0) = 0, \quad P_2(1,1) = P(T_2 < T_1).
$$

(9)

The total number of events $n(1,0)$ in the subset $(1,0)$ is given by $n_0$ outcomes 1 for each occurrence of outcome 2:

$$
n(1,0) = n_2 n_0.
$$

(10)

where $n_2$ is the asymptotic total number of outcomes 2. The total number of events in the complementary subset $n(1,1)$ can be computed by subtraction:

$$
n(1,1) = n_T - n_2 n_0.
$$

(11)

Equation (7) gives

$$
n_1 = n_2 n_0 + P(T_1 < T_2)(n_T - n_2 n_0),
$$

(12)

$$
n_2 = P(T_2 < T_1)(n_T - n_2 n_0),
$$

(13)

where $T_1$ and $T_2$ are the times required to fire the next occurrences of kind 1 and 2 respectively.

Equations (12) and (13) can be rewritten as

$$
n_1 = \frac{P(T_1 < T_2) + n_0 P(T_2 < T_1)}{1 + n_0 P(T_2 < T_1)} n_T,
$$

(14)

$$
n_2 = \frac{P(T_2 < T_1)}{1 + n_0 P(T_2 < T_1)} n_T.
$$

(15)

Equations (14) and (15) show that as $n_T \to \infty$ the asymptotic fraction $n_2/n_1$ reaches a fixed asymptotic value. In particular

$$
\frac{n_2}{n_1} = \frac{P(T_2 < T_1)}{P(T_1 < T_2) + n_0 P(T_2 < T_1)}.
$$

(16)

Following [3], we assign independent exponentially distributed probability density functions to the random variables $T_1$ and $T_2$:

$$
T_1 \overset{\text{ind}}{\sim} \text{Exp}(c_1), \quad T_2 \overset{\text{ind}}{\sim} \text{Exp}(c_2).
$$

(17)

The probabilities in (16) can be computed as suggested in the AppendixA, which gives

$$
P(T_1 < T_2) = \frac{c_1}{c_1 + c_2}, \quad P(T_2 < T_1) = \frac{c_2}{c_1 + c_2}.
$$

(18)

Thus, the asymptotic ratio between the number of outcomes 2 and the number of outcomes 1 becomes

$$
\frac{n_2}{n_1} = \frac{c_2}{c_1 + n_0 c_2}.
$$

(19)

This result confirms the validity of the analytical approach presented for this model in [11]. It is also in good agreement with the data obtained using the Monte Carlo method proposed in [11], as is shown in Fig. 1.

\section*{IV. CONTROLLED RADICAL POLYMERIZATION}

In this section we introduce a real phenomenon that can be studied by the proposed analytical method. In particular, we are interested in a polymerization process known as controlled radical polymerization (CRP) of acrylic monomers [10].

Radical polymerization is a method by which high molecular weight polymer molecules can be formed by successive addition of individual monomer units. In radical polymerization of vinyl monomers, each addition, or propagation, outcome regenerates the active radical species at the chain end, and the chain continues to grow until it terminates, typically by reaction with another radical.

Alternatively, the reactive radical at the end of the chain can be transferred to a carbon atom within the chain, thus generating a so-called mid-chain radical. In radical polymerization of acrylic monomers this is particularly important and it occurs via a process (outcome) known as backbiting. Due to the specific molecular orientation required for the backbiting outcome to occur, at least three propagation
outcomes must take place before the next backbiting outcome occurs.

Addition of monomer units to the mid-chain radical, which is formed by backbiting, results in the formation of a branched structure. Obtaining information about the kinetics of the process, and the relative rate of reactions between backbiting and propagation, are important for understanding polymerization of acrylic monomers. The ratio of the two competitive outcomes has a strong impact on polymer microstructure and the mechanical properties of the resulting polymer. It can be measured by evaluating the branching fraction, which is determined experimentally as the number of branches compared to the number of propagation outcomes that have occurred.

Controlled radical polymerization (CRP) is a special type of radical polymerization which is conducted in the presence of an additional chemical reagent known as a control agent. In CRP the reactive radical chain end is subjected to frequent deactivation and reactivation steps, allowing one to control the molecular weight of the polymer chain. Reversible deactivation of the reactive chain end is an outcome which occurs in competition with propagation and backbiting. We refer to it as the deactivation outcome.

Although classical chemical reaction kinetics dictate that the imposition of the additional competitive process of chain deactivation should not impact the ratio of backbitings to propagations, experimental evidence has shown that there is a strong reduction in branching fraction under CRP conditions [5,12,14]. In order to explain this reduction, the existence of a nonexponential probability density function has been proposed [12]. In [12], a Monte Carlo fitting scheme was suggested for evaluation of the appropriate parameters of the nonexponential PDF in order to fit the experimental branching fractions. The data provided in [12], shown in Fig. 2, will be used to compare our analytical approach with the MC based procedure. Two sets of data with corresponding uncertainty intervals were obtained by two different experimental procedures, known as bulk and solution polymerization. Each polymerization was conducted at different control agent concentrations, thus giving a range of data points for the fitting procedure.

A. The Monte Carlo fitting scheme


The MC method follows the SSA [3], where independent linear exponential PDFs, Linexp($b_i, \tau_i$), are assumed for required times for each outcome. This means that the time required for a propagation is $T_p^{\text{ind}} \sim \text{Linexp}(b_p, \tau_p)$, for a backbiting it is $T_b^{\text{ind}} \sim \text{Linexp}(b_r, \tau_r)$, and for a deactivation it is $T_d^{\text{ind}} \sim \text{Linexp}(b_d, \tau_d)$. The linear exponential PDF is defined as

$$T_i \sim \text{Linexp}(b_i, \tau_i) \Leftrightarrow f_{LE}(t; b_i, \tau_i)$$

$$= \begin{cases} \frac{2}{b_i^2 + 2b_i \tau_i} t, & \text{if } 0 \leq t < b_i, \\ \frac{2}{2b_i} e^{-\left(\frac{t}{b_i^2 + 2b_i \tau_i}\right)}, & \text{if } t \geq b_i. \end{cases} \tag{20}$$

Different concentrations of monomers and control agents are considered as proposed in the supporting information of [12]. The MC method is summarized in Algorithm 1.

The optimization routine can be described as follows:

(1) Choose the values of the parameters $\{b_p, \tau_p, b_r, \tau_r, b_d, \tau_d\}$ for a particular attempt of the chosen optimization algorithm.

(2) Run the MC method proposed in Algorithm 1 and calculate the mean ratio between numbers of backbitings and propagations, $r$, corresponding to the current choice of parameters.

(3) Choose and compute the cost function $J$. The middle point of each uncertainty interval is used.

(4) Restart from step 1 until $J$ is minimized.

We have to stress that for the purposes of this paper the choice of an optimization algorithm is not important. However, in practice one should go for the most efficient one available, since this would help to reduce the number of Monte Carlo runs required for finding the optimal parameters. Next we show how the MC algorithm in the optimization scheme described above can be replaced with a significantly more effective analytical approach.

The reason to invoke Monte Carlo procedure is because it is one of the most commonly used methods to study such processes. Thus, we began by defining the kinetics of the system in terms of Monte Carlo rules, or by defining the explicitly constrained PDF proposed in Algorithm 1. Next we show how to directly solve the dynamics of the system without having to resort to the MC approach.
The following are the probabilities for each outcome in each controlled radical polymerization carried out in the presence of Algorithm 1. The MC method for simulation of the evolution of relative frequencies of constrained events in . . . PHYSICAL REVIEW E 92 backbiting needs at least propagation and deactivation are always possible, whereas backbiting is not possible: backbiting, and by the number of deactivations occurred when the branching fraction can be calculated as a ratio between 

\[ n(1,1,0) = \frac{n_0 \rho_r}{P(T_p < T_d)}. \] (24)

Equation (23) can be rewritten as

\[ n(1,1,0) = n_T - n(1,1,0). \] (25)

Apparently,

\[ n(1,1,1) = n_T - n(1,1,0). \] (26)

Then, Eq. (7) yields

\[ n_p = \frac{P(T_p < T_d) n(1,1,0) + P(T_p < T_d,T_r)[n_T - n(1,1,0)]}{P(T_p < T_d)} \] (27)

\[ n_d = \frac{P(T_d < T_p)n(1,1,0) + P(T_d < T_p,T_r)[n_T - n(1,1,0)]}{P(T_d < T_p)} \] (28)

From Eqs. (26) and (28) we have

\[ n_r = \frac{P(T_r < T_p,T_d) + n_0 P(T_r < T_p,T_d)}{P(T_p < T_d)} \] (29)

\[ n_p = \frac{P(T_p < T_d) + n_0 P(T_r < T_p,T_d)}{P(T_p < T_d)} \] (30)

Equations (29) and (30) show that as \( n_T \to \infty \), the branching fraction \( r = n_r/n_p \) reaches the fixed asymptotic value. In particular, the ratio \( r \) is a function of probabilities involving the random variables \( T_p, T_d, \) and \( T_r \) only.

This implies two important consequences: (i) our solution holds for any choice of interevent-time PDF and (ii) it is possible to express the branching fraction as a function of the PDF parameters, once the probabilities of interest are computed. We explain how to compute the required probabilities for a generic case in the Appendix. This method is still valid for the linear exponential case (20) considered in this study.

**B. Analytical expressions for asymptotic relative frequencies of Events in CRP**

We will follow the ideas described in Sec. II. In particular, let \( n_p, n_d, \) and \( n_r \) respectively be the asymptotic mean number of occurred propagations, deactivations, and backbitings. Then the branching fraction can be calculated as a ratio between \( n_1 \) and \( n_0 \) using Eqs. (6) and (7). As stated in Sec. IV, propagation and deactivation are always possible, whereas backbiting needs at least \( n_0 = 3 \) previous propagations to occur. Hence, we have two subsets [the order is propagation \((p)\), deactivation \((d)\), backbiting \((r)\)]:

\[ (1,1,0), (1,1,1). \] (21)

The following are the probabilities for each outcome in each subset:

\[ P_p(1,1,0) = P(T_p < T_d), \quad P_p(1,1,1) = P(T_p < T_d,T_r), \]

\[ P_d(1,1,0) = P(T_d < T_p), \quad P_d(1,1,1) = P(T_d < T_p,T_r), \]

\[ P_r(1,1,0) = 0, \quad P_r(1,1,1) = P(T_r < T_p,T_d). \] (22)

The total number of events that cannot be a backbiting, \( n(1,1,0) \), is given by \( n_0 \) propagations for each occurred backbiting, and by the number of deactivations occurred when backbiting is not possible:

\[ n(1,1,0) = n_0 \rho_r + n_d(1,1,0) = n_0 \rho_r + P(T_d < T_p)n(1,1,0). \] (23)

\[ FIG. 3. Fitted data obtained by the analytical approach (crosses) and by the MC method (open circles) (MC sample size \( G = 10^4 \)) are presented for two polymerization reactions, bulk and solution. Both approaches use the linear exponential interevent-time PDF (20). \]
evaluates the same cost function \( J \) explained in Sec.IV A. The second method (crosses) follows shown by open circles) corresponds to the optimization routine the experimental observations. The first approach (the data are approaches for calculation of the branching fractions using the same optimization route (Nelder-Mead method \([15]\)) and tine (Nelder-Mead method \([ 15]\)) performed with an increasing factor of 104 compared with the MC based optimization method.

The proposed analytical approach is based on the assumption that the asymptotic limit is reached. Thus, it describes the statistics of the asymptotic states of a stochastic process. Reaching the asymptotic limit via Gillespie type simulations is costly, since also the transient dynamics need to be performed. Availability of analytical shortcuts avoids having to deal with fluctuations present in finite systems. In particular, in the considered application (CRP), the reaction continues until all the reactants are used up. For this reason, our method agrees with the experimental data for a reaction which may be considered to have reached its asymptotic state.

### C. Numerical results

We present the results of numerical tests run for validation of the proposed approach. The experimental data used in the tests were described in Sec. IV.

In Fig. 3, we assess the accuracy provided by two fitting approaches for calculation of the branching fractions using the experimental observations. The first approach (the data are shown by open circles) corresponds to the optimization routine explained in Sec. IV A. The second method (crosses) follows the same optimization route (Nelder-Mead method \([15]\)) and evaluates the same cost function \( J \), but the Monte Carlo evaluation of the branching fractions is replaced with the analytical expressions given in (29) and (30). Both the MC method and the analytical approach use a linear exponential interevent-time PDF \([20]\). As follows from Fig. 3, the sample size \( G = 10^4 \) in the Monte Carlo approach guarantees the same level of accuracy provided by the analytical method. Table I shows the optimized parameters for the linear exponential interevent-time PDF \([20]\).

Although both methods can offer comparable accuracies, this is not the case for the computational cost. The optimization routine performed with the analytical approach is up to 104 times faster than the one using the MC method of the same level of accuracy (MC sample size \( G = 10^4 \)). Comparative computational times are shown in Fig. 4.

It is clear that the degree of speedup provided by the analytical approach over the MC method is determined by the MC sample size \( G \). Indeed, the computational complexity of the analytical method is \( O(1) \) whereas it is \( O(G) \) in the case of Monte Carlo. This is confirmed by the numerical tests.

Different optimization algorithms and interevent-time PDFs were tested. In Figs. 5 and 6 and in Table II we present the results obtained with the fitting scheme in which both the optimization routine and the interevent-time PDFs differ from those applied in the previous test. In particular,

### Table I. Optimized parameters of linear exponential interevent-time PDFs \([20]\) obtained by analytical solution fitting (AS) and Monte Carlo fitting (MC). All results assume unitary concentrations of monomers and control agents (supporting information can be found in Ref. \([12]\)).

<table>
<thead>
<tr>
<th>Polymerization</th>
<th>Fitting</th>
<th>( b_p )</th>
<th>( b_r )</th>
<th>( b_d )</th>
<th>( \tau_p )</th>
<th>( \tau_r )</th>
<th>( \tau_d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution</td>
<td>AS</td>
<td>( 1.74 \times 10^{-1} )</td>
<td>6.53</td>
<td>( 2.28 \times 10^{-4} )</td>
<td>( 9.1 \times 10^{-1} )</td>
<td>1.31</td>
<td>3.58 \times 10^{-2}</td>
</tr>
<tr>
<td>Solution</td>
<td>MC</td>
<td>( 1.74 \times 10^{-1} )</td>
<td>6.53</td>
<td>( 2.28 \times 10^{-4} )</td>
<td>( 9.1 \times 10^{-1} )</td>
<td>1.31</td>
<td>3.58 \times 10^{-2}</td>
</tr>
<tr>
<td>Bulk</td>
<td>AS</td>
<td>( 2.8 \times 10^{-1} )</td>
<td>1.58 \times 10^{-1}</td>
<td>( 1.57 \times 10^{-2} )</td>
<td>( 8.53 \times 10^{-1} )</td>
<td>11.54</td>
<td>3.43 \times 10^{-2}</td>
</tr>
<tr>
<td>Bulk</td>
<td>MC</td>
<td>( 1.64 \times 10^{-1} )</td>
<td>1.40 \times 10^{-1}</td>
<td>( 3.557 \times 10^{-2} )</td>
<td>( 9.16 \times 10^{-1} )</td>
<td>12.01</td>
<td>2.56 \times 10^{-2}</td>
</tr>
</tbody>
</table>

![FIG. 4. Computational times required for the optimization routine (Nelder-Mead method \([15]\)) performed with an increasing number of iterations for bulk and solution polymerization. The analytical approach (crosses and squares) speeds up the procedure by the factor of 10^4 compared with the MC based optimization method (open circles and triangles) of the same level of accuracy.](043306-6)

![FIG. 5. Fitted data obtained by the analytical approach (crosses) and by the MC method (open circles) (MC sample size \( G = 10^4 \)) are presented for two polymerization reactions, bulk and solution. Both approaches use exponential PDF’s for propagation and deactivation, and a linear exponential PDF \([20]\) for backbiting. Genetic algorithm is employed for optimization.](043306)
Figure 6. Computational times required for the optimization routine (genetic algorithm [16]) performed with an increasing number of iterations for bulk and solution polymerization. Here exponential PDF’s are chosen for propagation and deactivation, whereas a linear exponential PDF was assigned for backbiting. The analytical approach (crosses and squares) speeds up the procedure by a factor of $10^3$ compared with the MC based optimization method (open circles and triangles) of the same level of accuracy.

A genetic algorithm [16] has been selected for optimization. A pure exponential PDF was assigned for propagation and deactivation, whereas a linear exponential PDF (20) was chosen for backbiting. The rationale behind the choice of the PDFs for propagation and deactivation is confirmed by our previous test, described above, which results in very small optimal parameters $b_p$ and $b_d$. This suggests that the optimal PDF choices for propagation and deactivation are very close to an exponential function. The results shown in Fig. 5 justify this choice. Also, these results confirm that the proposed analytical approach is valid for various choices of optimization routines and interevent-time PDFs.

V. CONCLUSIONS

Stochastic simulation algorithms produce multiple realizations of the full evolution of studied processes. In principle, this information allows a systematic study of all properties of the simulated system. The price for this detailed description includes, however, excessive computational time, needed to perform the Monte Carlo simulations, as well as the presence of unavoidable statistical errors. If the focus is on extracting particular quantities of interest, it may be possible to solve the problem analytically. The proposed analytical approach is one example of this strategy. We have demonstrated the superiority of this analytical approach over a traditional Monte Carlo simulation in the computation of the branching fraction in controlled radical polymerization. The approach is free of statistical errors, and thus guarantees more accurate estimations, than those provided by a Monte Carlo simulation. In addition, the method is significantly (an order of the sample size) faster than the Monte Carlo approach. The performed tests show that the choice of optimization algorithm is not important and the analytical approach works with any choice of the interevent-time PDF. The proposed method is also general enough to be used in other applications, especially in those involving constrained stochastic processes which are difficult and often impossible to simulate using a conventional statistical simulation algorithm. Finally, the proposed approach can be used as an efficient tool for finding the optimal set of parameters for interevent-time PDFs, to be further utilized in detailed Monte Carlo simulations.

ACKNOWLEDGMENTS

This research is supported by the Severo Ochoa Grant No. SEV-2013-0323, the Predoctorales Severo Ochoa 2014 Grant No. SVP-2014-068451. One of us (E.A.) acknowledges the support of the Project No. MTM2013-46553-C3-1-P. D.S. acknowledges support of the Project Grupos Consolidados UPV/EHU del Gobierno Vasco (IT-472-10) and the MINECO Grant No. FIS2012-36673-C03-01. The authors thank J.M. Asua (POLYMAT, Spain) for valuable discussions.

APPENDIX: COMPUTATION OF THE PROBABILITIES

Below we explain how to compute the probability $P(T_k < T_j \forall j \neq k)$, with $k,j \in \{1, \ldots, N\}$. $T_1, \ldots, T_N$ are independent random variables with the generic PDF $f_i(t)$, $i \in \{1, \ldots, N\}$:

$$T_i \sim f_i(t), \quad \forall i \in \{1, \ldots, N\}. \quad (A1)$$

The cumulative density function $F_i(t)$ of the random variable $T_i$ is

$$F_i(t) := P(T_i \leq t) = \int_0^t f_i(\tau)d\tau, \quad \forall i \in \{1, \ldots, N\}. \quad (A2)$$

The probability of interest is then computed as

$$P(T_k < T_j \forall j \neq k)$$

$$= \int_0^{+\infty} P(T_k < T_j \forall j \neq k, T_k = \tau)d\tau$$

$$= \int_0^{+\infty} P(T_k = \tau) \prod_{j \neq k} P(T_j > \tau)d\tau$$

$$= \int_0^{+\infty} f_k(\tau) \prod_{j \neq k}(1 - F_j(\tau))d\tau, \quad \forall k \in \{1, \ldots, N\}. \quad (A3)$$

TABLE II. Optimized parameters corresponding to the model and interevent-time PDFs.

<table>
<thead>
<tr>
<th>Polymerization</th>
<th>$b_p$</th>
<th>$b_d$</th>
<th>$\tau_p$</th>
<th>$\tau_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution</td>
<td>5.94</td>
<td>1</td>
<td>5.97</td>
<td>4.78 x 10^{-2}</td>
</tr>
<tr>
<td>Bulk</td>
<td>3.82 x 10^{-1}</td>
<td>1</td>
<td>8.80</td>
<td>1.33</td>
</tr>
</tbody>
</table>