Available online at www.sciencedirect.com
ScienceDirect
Polymer 48 (2007) 1770–1784
polymer
www.elsevier.com/locate/polymer

Conditional Monte Carlo sampling to find branching architectures of polymers from radical polymerizations with transfer to polymer and recombination termination

Piet D. Iedema [a,*], Michael Wulkow [b], Huub C.J. Hoefsloot [a]

[a] *Universiteit van Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, Netherlands*
[b] *Computing in Technology GmbH, Rastede, Germany*

Received 28 November 2006; received in revised form 19 January 2007; accepted 22 January 2007
Available online 25 January 2007

Abstract

A model is developed that predicts branching architectures of polymers from radical polymerization with transfer to polymer and termination by disproportionation and recombination, in a continuously stirred tank reactor (CSTR). It is a so-called conditional Monte Carlo (MC) method generating architectures of molecules of specified dimensions. The relevant dimensions in the present case are the number of branch points, n_p, and the number of combined parts a molecule consists of, n_c. These branch points and combination points together are decisive for the connectivity inside molecules. The modeling strategy is based on backtracking of the molecular growth history in terms of the chemical events determining connectivity, transfer to polymer and recombination termination. The recombination termination mechanism requires the model to develop parts of the architecture following several paths back to the initial primary polymers that form the starting points for the molecules. The algorithm requires the construction of probability density functions being evaluated using a fast Galerkin-FEM method. The architectures generated by the conditional Monte Carlo method are compared to those from a full MC method using several qualifiers. One of these is the number of initial primary polymers in a molecule as well as their lengths, another is the radius of gyration contraction factor. Perfect agreement is found between the architectures found by the conditional and full MC methods.

Keywords: Radical polymerization; Branching architecture; Monte Carlo simulation

1. Introduction

The issue of branching architectures in terms of connectivity patterns between linear chain parts was already addressed a long time ago, for instance by Zimm and Stockmeyer [1]. Even until today – 50 years later – the relationship derived by these authors is still in use in characterization of branched polymers. The since long recognized importance of this issue is now also seen in recent developments in rheological modeling. A range of new rheological models have been proposed being directly based on branching architectures, like the 'pom-pom' model from the group of McLeish [2,3]. Both radical polymerization of ethylene, where branching is caused by transfer to polymer, and catalyzed ethylene polymerization, where long branches are formed by insertion of chains with unsaturated chain ends, have received considerable attention [4–9]. This study will be devoted to radical polymerization in a continuously stirred tank reactor (CSTR) with disproportionation and recombination as the termination mechanisms.

The main objective of our work is finding the architectures in relation to kinetics. Moreover, our task of finding architectures for molecules of *specified dimensions* has led to a special view of the radical polymerization. A similar phenomenon took place in the case of termination by disproportionation only [10], where we noticed some interesting phenomena concerning lengths of primary polymers (linear elements) and sequence order of growth in one molecule. Length turns out to decrease with increasing sequence number. This

* Corresponding author.
E-mail address: piet@science.uva.nl (P.D. Iedema).

0032-3861/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.polymer.2007.01.046

fact – hitherto unrealized – became fully understandable by analyzing the problem at the level required to construct the model. Similarly, in the present paper we observe some interesting phenomena particularly due to the occurrence of recombination termination. It has forced us to realize by which chemical events – in this particular kinetic case – the connectivity between the linear elements (primary polymers) is determined. These events are *transfer to polymer* and *recombination termination* leading to two different types of connection points: branch points and combination points. Now, one of the factors determining the architecture turned out to be the particular *sequence* of termination by either disproportionation or recombination. Finding the sequences is one of the ingredients of the model that is being realized by backtracking the growth history in inversed chronological order (Fig. 2). In this paper we also show that the occurrence of one or the other event in a growing molecule of certain dimensions (chain length, number of branch points) is directly related to concentrations of molecules of such dimensions. Concentration distributions of both living and dead chains, being different due to the recombination reaction, are involved.

As an alternative to this new method, molecular architectures can also be found by a different Monte Carlo method, developed by Tobita [11–13], although that procedure does not allow to do this for a molecule of given dimensions. In contrast to our new method, Tobita's method – in this article hereafter called the *full* Monte Carlo method – generates samples of molecules of varying dimensions, while in addition it provides information by the most probable architectures belonging to each of the molecules can be found. This then serves as a reference to our conditional Monte Carlo procedure – for smaller molecules only. Finding statistically representative samples of large molecules is a bottleneck in full MC, while it is not in conditional MC, which is one of its advantages. To illustrate this point, we here draw the attention to a result discussed at the end of this paper, and presented in Fig. 17. This compares the radius of gyration contraction factor as calculated from two samples of 100,000 molecules of *exactly identical chain length and number of branch points*, but one made without and the other with recombination termination under the same kinetic conditions. The question was to find out whether one kind of molecules was more compact than the other. These sets were easily generated by conditional Monte Carlo, while it is evident that creating sets of molecules with exactly identical dimensions is practically impossible with full Monte Carlo.

It is obvious that the conditional method requires the a priori knowledge of the concentration distribution of chain length, number of primary polymers (number of branch points) and number of combination points. Several methods to find this 2D-solution have been discussed by us in previous work [14] and will here be referred to. In addition, the probability density functions used in the history backtracking algorithm have to be constructed and we will show that this leads to a considerable combinatorial problem, requiring a specific solution, based on a Galerkin-FEM method (cf. PREDICI® [14]).

A note of caution must be made concerning the fact that in this work we will assume that the rate coefficients of termination are independent of the length of the macro-radicals involved. In radical polymerization science it is, however, commonly accepted that the rate of termination is often controlled by diffusion and hence does depend on chain length [15]. This will influence the distributions of chain length, number of branch points and number of combination points. After having explained the structure of the algorithms, in the course of this article we will provide the arguments indicating that an extension to chain length dependent termination is straightforward. For the sake of simplicity and also to be able to compare to full Monte Carlo sampling, we have chosen to employ termination coefficients that are independent of chain length.

This article is structured as follows. After an explanation of the terminology defining the dimensions of molecules the history backtracking algorithm is introduced, together with the required probability density functions. This is followed by a description of the algorithm that constructs the architectures from the molecular growth histories. This is completed by a section where the probability density functions are derived. The first part of Section 8 reports on the comparison of the probability density functions as calculated from Galerkin FEM and as they may be inferred from full Monte Carlo simulations. Finally, the architectures from the full and conditional Monte Carlo methods will be compared and discussed using various architectural qualifiers, among which the radius of gyration contraction factor [10].

2. Architectures generation by conditional Monte Carlo sampling

The conditional Monte Carlo algorithm is based on the chemistry of the radical polymerization system as briefly summarized in the reaction and population balance equations in Table 1. The

Table 1

Initiator dissociation	$I_2 \xrightarrow{k_d} 2I$
Initiation	$I + M \xrightarrow{k_i} R_{1,0,0}$
Propagation	$R_{n,i,k} + M \xrightarrow{k_p} R_{n+1,i,k}$
Termination by disproportionation	$R_{n,i,k} + R_{m,j,l} \xrightarrow{k_{td}} P_{n,i,k} + P_{m,j,l}$
Termination by recombination	$R_{n,i,k} + R_{m,j,l} \xrightarrow{k_{tc}} P_{n+m,i+j,k+l+1}$
Transfer to polymer	$R_{n,i,k} + P_{m,j,l} \xrightarrow{k_{tp}m} P_{n,i,k} + R_{m,j+1,l}$
Population balance equations	$dR_{n,i,k}/dt = k_p M(-R_{n,i,k} + R_{n-1,i,k}) - (k_{td} + k_{tc})\lambda_0 R_{n,i,k} + k_{tp}(-\mu_1 R_{n,i,k} + \lambda_0 n P_{n,i-1,k}) - 1/\tau R_{n,i,k}$
	$dP_{n,i,k}/dt = k_{td}\lambda_0 R_{n,i,k} + 1/2k_{tc}\sum_{m=1}^{n-1}\sum_{j=0}^{i}\sum_{l=0}^{k-1} R_{m,j,l}R_{n-m,i-j,k-l-1} + k_{tp}(\mu_1 R_{n,i,k} - \lambda_0 n P_{n,i,k}) - 1/\tau P_{n,i,k}$

m, n: chain length; i, j: number of branch points per chain; k, l: number of combination sites per chain.

solution of the balance equations, for instance by the direct solution method as described in previous work [14], is supposed to have provided us with the 3D chain length, n, number of branch points, i, number of combination points, k, distribution for dead and living chains, $P(n,i,k)$ and $R(n,i,k)$, respectively. Now it should be noted that in the algorithm we do use length, n, number of primary polymers, n_p, and number of combined part, n_c, instead of branch points and combination points. A primary polymer in this work is defined as a single linear element created by initiation or transfer to polymer, propagation and termination. Hence, according to this terminology a linear chain created by initiation at two sites, propagation and recombination termination counts as two (connected) primary polymers. This definition differs from that in the full Monte Carlo approach [11–13], where such an element counts for one. The relations between the number of combination points, k, and branch points, i, as appearing in the population balances, and the number of combined parts, n_c, and primary polymers, n_p, are, respectively:

$$n_c = k + 1 \tag{1}$$

$$n_p = i + k + 1 \tag{2}$$

We prefer working with n_p and n_c over i and k, since simple additivity holds for the former. The 3D-distributions $P(n,n_p,n_c)$ and $R(n,n_p,n_c)$ are inferred from $P(n,i,k)$ and $R(n,i,k)$ in a straightforward way.

3. The architecture generation algorithm

In order to construct the architecture of a branched molecule, like in the case of disproportionation only, we firstly need to find the correct sequences and lengths of the primary polymers, and then secondly build up the architectures. To perform the first task, we had to develop a new algorithm, the *history backtracking algorithm*. This algorithm will now be described. Later we show how to construct the architectures from the resulting primary polymer sequences and lengths. It should be realized that this algorithm rigorously and fully reflects the chemistry that leads to a particular structure.

4. History backtracking algorithm

4.1. Growth without or with recombination

This algorithm is similar to sequence/length finding in the case without recombination termination in the sense that it starts with a molecule of specified length, n, number of primary polymers, n_p, and number of combined parts, n_c. The basic idea that all the chemical events ultimately leading to a complete molecule are backtracked, one-by-one, is the same. In the case of disproportionation termination only the chemical events leading to a new primary polymer, transfer to polymer, subsequent propagation and termination, could mathematically be captured in one step, employing just one probability density function to determine the lengths of the primary polymer and the structure to which it is attached. Backtracking of the growth history then turned out to be relatively easy.

In the present case we have to deal with such a growth process, leading to a new primary polymer as well, but here it is intertwined with a combination process. This complicates the history backtracking.

Growth of a molecule may be initiated at different sites and different instants in time, leading to several *initial primary polymers*. Thus, several independently growing structures exist with separate growth histories that on some point in time merge with others. Only from these points on the structures have their growth process in common. Consequently, the algorithm in backtracking the growth history starts with the common history, but soon has to branch into different paths to backtrack the individual partial growth history until the initial primary polymers. We will now present the algorithm, of which a flowchart is shown in Fig. 1, and subsequently illustrate it with an exemplary molecule. The probability density functions employed by it will be defined along this explanation, but their derivation will be reported in a later section.

- *Start of the algorithm: recombination test*. The history backtracking algorithm starts with a dead molecule of given dimensions, n, n_p, n_c. As in most radical polymerization systems, we are mostly interested in dead molecules. For such molecules the final chemical event before leaving the reactor is always the reaction from its living predecessor(s) to the dead form either by termination or by transfer

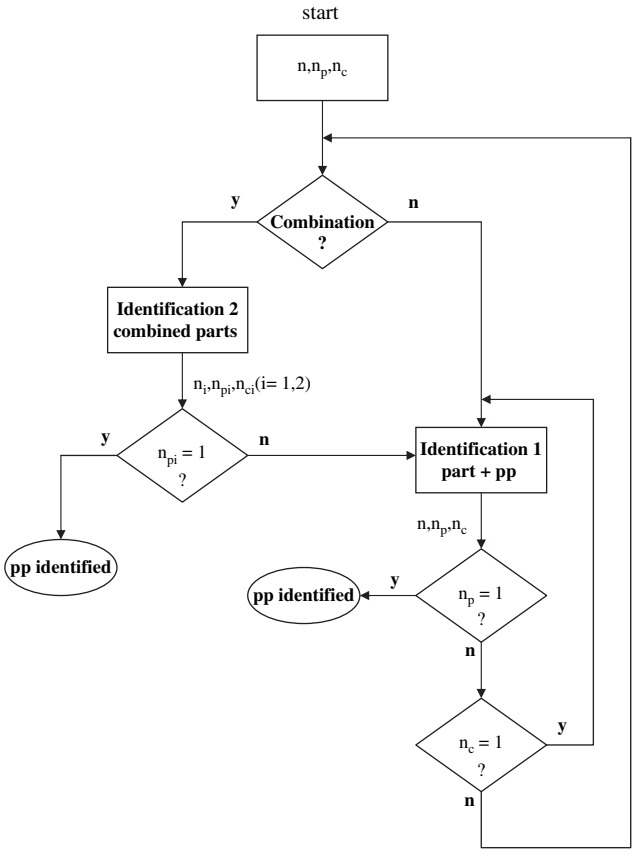

Fig. 1. Flowchart of the growth history backtracking algorithm.

to polymer. In the case of disproportionation termination and transfer the dead molecule's dimensions are identical to those of the living one, in case of recombination termination it is the sum of the two combining molecules. The probability that the last event is a recombination step is a function of n, n_p and n_c of the specified molecule, and we define it as $\Re_c(n, n_p, n_c)$, which is a number between 0 and 1. By choosing a random number between 0 and 1, r, we have recombination as the last event, if $r < \Re_c$, and disproportionation/transfer otherwise. This step takes place in the upper diamond block of the flowchart, Fig. 1.

- *Last step is no recombination*. We proceed now with explaining the case of no recombination, indicated as the path on right-hand side in the flowchart. Here, it should be realized that *living* molecules of n, n_p and n_c are created from smaller *dead* molecules with $n_p - 1$ primary polymers by a *transfer to polymer step* and a consecutive *growth step*. This part of the algorithm is exactly identical to that without recombination [10]. The specifications of the newly grown primary polymer, part 1 with length n_2, and the molecular structure on which it grows, part 2 with length n_1, number of primary polymers $n_{p1} = n_p - 1$, follow the probability density function $\Re_{t1}(n_1|n, n_p, n_c)$. This pdf defines the probability that given the specifications of the resulting molecule, n, n_p and n_c, the length of part 1 equals n_1 (and part 2 the complementary length $n_2 = n - n_1$). As shown in the flowchart (rectangular block on RHS) this step results in the identification of a dead molecular structure of specified dimensions and a primary polymer of known length (marked with the oval banner) attached to it at a yet unknown position. The algorithm then proceeds with testing the dimensions of the larger dead structure, as indicated in the upper diamond block on RHS. If it turns out that the molecule consists of only one primary polymer, $n_p = 1$, this path terminates with the identification of another primary polymer, which is in this case an initial primary polymer. If this is not the case, then the algorithm proceeds with another test (lower diamond block on RHS). If the molecule has more than one combination part, then the algorithm returns to the recombination test (upper diamond block) and a new round is started. If $n_c = 1$, then no combination points are left in the molecule and the combination test can be omitted. In that case the algorithm directly returns to the routine where the lengths are identified to the last grown primary polymer and the structure on which it grows (upper rectangular block on RHS).
- *Last step is recombination*. If the recombination test reveals that the last step was a combination step, the algorithm takes the path on left-hand side of the flowchart. The task carried out here (rectangular block on LHS) is to find the dimensions of the two parts from all the possible combinations that add up to the prevailing dimensions of the molecule, n, n_p and n_c. Hence, a 3D sampling problem is at hand. We decided to perform this using three 1D probability density functions in a (arbitrary) sequence. The first pdf determines the number of combination points, n_{c1} and n_{c2}, of the two parts: $\Re(n_{c1}|n, n_p, n_c)$. This function defines the probability that given n, n_p and n_c, the number of combination points on part 1 equals n_{c1} (and the complementary number $n_{c2} = n_c - n_{c1}$ on part 2). This yields n_{c1} and n_{c2}. A second pdf is employed to find the number of primary polymers, n_{p1} and n_{p2}: $\Re(n_{p1}|n, n_p, n_{c1}, n_{c2})$, which defines the probability that given n and n_p of the complete molecule and n_{c1} and n_{c2} of the parts the number of primary polymers of part 1 equals n_{p1} (and the complementary number $n_{p2} = n_p - n_{p1}$ on part 2). Finally, the lengths n_1 and n_2 are determined using $\Re(n_1|n, n_{p1}, n_{p2}, n_{c1}, n_{c2})$, the probability that part 1 has length n_1 ($n_2 = n - n_1$ for part 2) given total length n and n_{c1}, n_{c2}, n_{p1}, n_{p2} of the parts. After this step we end up with two parts being living molecules with specifications n_1, n_{c1}, n_{p1} and n_2, n_{c2}, n_{p2}.

The algorithm continues by testing the dimensions of the two parts (diamond block at LHS). If *only one* part consists of one primary polymer, $n_{pi} = 1$, then this primary polymer has been fully identified. In that case the primary polymer is not an initial one. If *both* parts consist of one primary polymer, both are fully identified and in that case two initial primary polymers are found; then this path of the algorithm is terminated. The algorithm continues for the one or two parts that have $n_{pi} > 1$. For each part the last event before recombination was a transfer to polymer step followed by propagation. Hence, the lengths of the primary polymer last grown and the structure on which it grows should be determined in the manner already described. Consequently, the algorithm – for one or two parts – enters the routine denoted in the flowchart by the rectangular block on RHS. If there are two parts to be further unraveled, and then it is here that the algorithm branches into two independent paths.

Thus, all the paths of the algorithm are worked through for this molecule, yielding all the parts with their lengths, number of primary polymers and combination points specified.

4.2. Backtracking example

The growth history backtracking algorithm is now demonstrated on a molecular example as shown in Fig. 2, cartoon 1. Its dimensions are: n (we will not specify it in this demonstration), $n_p = 15$, $n_c = 7$. The recombination test for this molecule, using $\Re_c(n, n_p, n_c)$, revealed that the last event was a recombination step. The dimensions of the two parts are then identified using $\Re(n_{c1}|n, n_p, n_c)$, $\Re(n_{p1}|n, n_p, n_{c1}, n_{c2})$ and $\Re(n_1|n, n_{p1}, n_{p2}, n_{c1}, n_{c2})$. The result is shown in cartoon 2: one part is a single primary polymer with length n_1 that is now fully identified, the other part is a larger structure of dimensions (n_2,14,6). Next, as shown in cartoon 3, the lengths of the last grown primary polymer, n_1, and that of the structure on which it grows, n_2, are determined, using $\Re_{t1}(n_1|n, n_p, n_c)$. This yields a fully identified primary polymer again and a structure of dimensions (n_2,13,6). At this point the algorithm returns to the recombination test and a new round is entered. As shown in cartoon 4, again the last step turns out to be

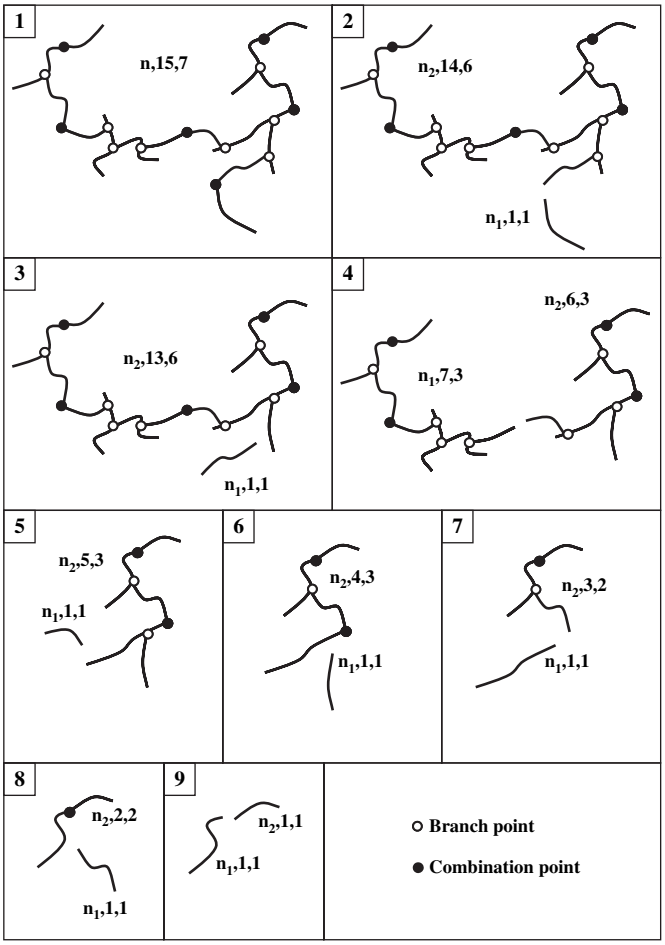

Fig. 2. Illustration of the algorithm. (1) Example molecule of length n, 15 primary polymers, 7 combination parts. (2) First step (last chemical event in the molecule) is recombination; identification of parts, one of these is a single pp of length n_1. (3) Identification of pp grown on larger structure prior to combination event of 2. (4) Recombination step; identification of two larger structures. (5) Identification of pp grown on one of the larger structures in 4. (6) This step is a disproportionation/transfer event. (7) Recombination step. (8) Disproportionation/transfer step. (9) Recombination step; this path of the algorithm ends up with two initial pp's. The algorithm continues with unraveling the other structure identified in step 4 (not shown).

recombination, this time resulting in two larger structures of dimensions (n_1,7,3) and (n_2,6,3). Now, the algorithm branches into two paths that are worked through one after another. Fig. 2 shows only one of these parts, that of structure (n_2,6,3). In cartoon 5, for this part, the result is shown of the transfer to polymer and growth step immediately preceding the recombination step of cartoon 4. Using $\Re_{t1}(n_1|n, n_p, n_c)$ the lengths of the primary polymer (n_1) and the larger structure (n_2,5,3) involved are now identified. At this point once again the algorithm returns to the recombination test. This time the last event is identified as a disproportionation/transfer step (cartoon 6), yielding a primary polymer of n_1 and a larger structure (n_2,4,3). The algorithm then enters a new round again and the recombination test finds a recombination event as the last step, yielding two parts: (n_1,1,1) (single primary polymer) and (n_2,3,2), see cartoon 7. The lengths of the primary polymer last grown and the structure on which it grows are determined, see cartoon 8. At this point the algorithm enters the last round with a structure (n,2,2). For a molecule of this dimensions the last event must be a recombination step (or $\Re_c(n, 2, 2) = 1$, for $n > 1$). Thus, the algorithm terminates this path by identifying two (initial) primary polymers as shown in cartoon 9. Subsequently, the algorithm picks up a new path, in this example for the structure (n_1,7,3) of cartoon 4. After having unraveled this part, the algorithm stops, since for this example all primary polymer lengths and their sequences per path are known.

5. Generating architectures

The construction of architectures from the primary polymers is basically the same as the one described in our previous work [10] on the case of disproportionation only, except for the fact that now sequences *per path* and *connectivity* between pairs of primary polymers have to be included. To this end the history backtracking algorithm stores as the necessary identifiers per primary polymer: a (arbitrary) row number ranging from 1 to n_p; a growth sequence number per path; (eventually) the row number of primary polymer to which it is connected by a combination point – the connection identifier.

The construction algorithm is schematically shown in Fig. 3. N counts the number of primary polymers incorporated in the structure; obviously the algorithm stops when $N = n_p$, the total number of primary polymers in the molecule. The algorithm starts by arbitrarily selecting one of the initial primary polymers. A test then is performed, using the connection identifier, whether it is connected to another primary polymer. If so, the connected structure of two primary polymers is formed, and N increases by one. If $N < n_p$, the algorithm continues by selecting the primary polymer next in sequence; it is attached on an arbitrary monomer unit of the structure already formed. If $N < n_p$, a test follows, to check whether the primary polymer is connected to another primary polymer. If not, then a new primary polymer is selected. A subsequent test checks whether this connecting primary polymer is a single one, in which case it is connected (N increases by one) and – if $N < n_p$ – the algorithm returns to selecting a new primary polymer next in sequence. The test may also reveal that the connecting primary polymer is not single but instead has grown on an intermediate structure that was started from another initial polymer. In this case a test follows, whether this structure already has been developed in a previous round of the algorithm. If not, the algorithm stores the structure found so far as an intermediate structure and returns to selecting a new initial primary polymer. If instead the intermediate structure to be connected to was already known, the coupling actually takes place and the algorithm – if $N < n_p$ – returns to selecting a new primary polymer next in sequence. The algorithm continues until, at some point, N becomes equal to n_p and then it terminates.

6. Probability density functions

As regards transfer to polymer we realize that *living* molecules of n, n_p and n_c, are created from smaller *dead* molecules with $n_p - 1$ primary polymers by a *transfer to polymer step*

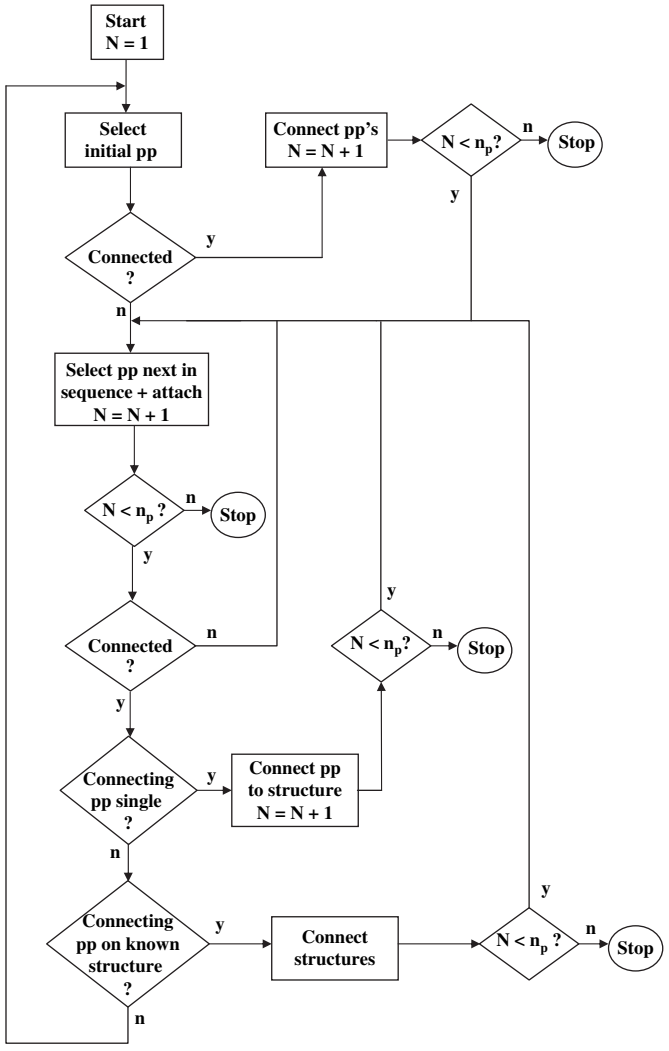

Fig. 3. Flowchart of the construction algorithm.

and a consecutive *growth step* producing a new primary polymer of length n_1. In the first step from a dead chain of $n-n_1$, n_p-1, n_c, a living chain of the same length and number of combination parts and n_p primary polymers is created with probability p_1:

$$p_1\left(n-n_1,n_p-1,n_c\right)\sim(n-n_1)P\left(n-n_1,n_p-1,n_c\right) \tag{3}$$

Here, P is the 3D-distribution of length, number of primary polymers and number of combination parts of *dead* molecules as calculated from the Galerkin-FEM procedure. In p_1, the multiplication by $n-n_1$ arises from the fact that any of the $n-n_1$ units of the molecule can undergo this step. The second step is growth of a new arm by propagation at the new branch point to length n_1 until termination. The probability, p_2, of creating an arm length n_1 obeys the Flory distribution (the same as the one sampled from in the full MC algorithm):

$$p_2(n_1)\sim p_F(n_1)=\frac{(1-1/\overline{n})^{(n-1)}}{\overline{n}} \tag{4}$$

where the average primary polymer length $\overline{n}$ is given by [2]:

$$\overline{n}=\frac{k_pMI+k_{tp}\mu_1\lambda_0+k_pM\lambda_0}{\lambda_0\left(k_{tp}\mu_1+k_{td}\lambda_0+k_{tc}\lambda_0+1/\tau\right).} \tag{5}$$

The probability of thus creating a dead molecule of total length n and n_p primary polymers is simply the product of p_1 and p_2, since length n_1 is independent of length $n-n_1$. Hence, for the conditional probability that a dead chain is created from a dead chain of $n-n_1$, n_p-1, n_c and subsequent growth of an arm of length n_1 we can write:

$$p\left(n_1\middle|n,n_p,n_c\right)\sim(n-n_1)P\left(n-n_1,n_p-1,n_c\right)p_F(n_1). \tag{6}$$

Now, following this argument we state that a molecule of n, n_p, n_c can be created from *any* molecule with n_p-1 primary polymers of (shorter) length $n-n_1$ undergoing transfer to polymer and subsequent growth of an arm with the complementary length n_1. According to Eq. (6) the probability of the various combinations is proportional to $(n-n_1)P(n-n_1,n_p-1,n_c)$ $p_F(n_1)$. Hence, given the length n of the original molecule with n_p-1 primary polymers, the probability distribution of arm length n_1, $\Re_{t1}(n_1|n,n_p,n_c)$, obeys the following probability distribution function (pdf):

$$\Re_{t1}\left(n_1\middle|n,n_p,n_c\right)=\frac{p_F(n_1)P\left(n-n_1,n_p-1,n_c\right)(n-n_1)}{\sum_{n_1=1}^{n-1}p_F(n_1)P\left(n-n_1,n_p-1,n_c\right)(n-n_1).} \tag{7}$$

This result is exactly identical to that for the case without recombination termination [10].

In order to determine, whether the last chemical event having occurred in a dead molecule before leaving the reactor was a termination by recombination or another termination step, the reaction rates of both termination possibilities have to be known. According to the reaction mechanisms in Table 1, living chains may terminate by disproportionation/recombination or by transfer to polymer. The reaction rate of recombination termination for molecules of dimensions n, n_p and n_c is given by the following sum:

$$r_c\left(n,n_p,n_c\right)=k_{tc}\sum_{n_1=1}^{n-1}\sum_{n_{p1}=1}^{n_p}\sum_{n_{c1}=1}^{n_c}R\left(n_1,n_{p1},n_{c1}\right)R\left(n_2,n_{p2},n_{c2}\right);$$
$$n_2=n-n_1,n_{p2}=n_p-n_{p1},n_{c2}=n_c-n_{c1} \tag{8}$$

Here, R is the 3D-distribution of length, number of primary polymers and number of combination parts of *living* molecules as calculated from the Galerkin-FEM procedure. The total rate of reaction for termination by disproportionation and transfer to polymer for molecules of the same dimensions is as follows as:

$$r_{dt}\left(n,n_p,n_c\right)=\left(k_{td}\lambda_0+k_{tp}\mu_1\right)R\left(n,n_p,n_c\right) \tag{9}$$

The probability that a dead molecule of these dimensions is created by recombination follows from the ratio between these two rates:

$$\Re_c(n,n_p,n_c) = 1/\{1 + r_c(n,n_p,n_c)/r_{dt}(n,n_p,n_c)\}. \quad (10)$$

Next follows the derivation of the probability density function describing the characteristics of the pairs of molecular parts involved in a combination step. The probability density function describing the probability that a dead molecule of dimensions n, n_p and n_c is the product of the reaction between two living molecules, of which one has n_{c1} combination parts (and the other $n_c - n_{c1}$) is given by:

$$\Re(n_{c1}|n,n_p,n_c) = \frac{\sum_{n_1=1}^{n-1}\sum_{n_{p1}=1}^{n_p} R(n_1,n_{p1},n_{c1})R(n_2,n_{p2},n_{c2})}{\sum_{n_{c1}}^{n_c}\sum_{n_1=1}^{n-1}\sum_{n_{p1}=1}^{n_p} R(n_1,n_{p1},n_{c1})R(n_2,n_{p2},n_{c2})};$$
$$n_2 = n - n_1, n_{p2} = n_p - n_{p1}, n_{c2} = n_c - n_{c1}. \quad (11)$$

The following probability density function defines the probability that given n and n_p of the complete molecule and n_{c1} and n_{c2} of the parts the number of primary polymers of part 1 equals n_{p1} (and the complementary number $n_{p2} = n_p - n_{p1}$ on part 2):

$$\Re(n_{p1}|n,n_p,n_{c1},n_{c2}) = \frac{\sum_{n_1=1}^{n-1} R(n_1,n_{p1},n_{c1})R(n_2,n_{p2},n_{c2})}{\sum_{n_{p1}=1}^{n_p}\sum_{n_1=1}^{n-1} R(n_1,n_{p1},n_{c1})R(n_2,n_{p2},n_{c2})};$$
$$n_2 = n - n_1, n_{p2} = n_p - n_{p1}, n_{c2} = n_c - n_{c1}. \quad (12)$$

Finally, the probability that part 1 has length n_1 ($n_2 = n - n_1$ for part 2) given total length n and n_{c1}, n_{c2}, n_{p1}, n_{p2} of the parts is expressed by the following probability density function:

$$\Re(n_1|n,n_{p1},n_{p2},n_{c1},n_{c2}) = \frac{R(n_1,n_{p1},n_{c1})R(n_2,n_{p2},n_{c2})}{\sum_{n_1=1}^{n-1} R(n_1,n_{p1},n_{c1})R(n_2,n_{p2},n_{c2})};$$
$$n_2 = n - n_1, n_{p2} = n_p - n_{p1}, n_{c2} = n_c - n_{c1}. \quad (13)$$

Now, to find the rate $r_c(n,n_p,n_c)$ and the pdf $\Re(n_{c1}|n,n_p,n_c)$, Eqs. (8) and (11) poses us a huge combinatorial problem, since in principle all possible combinations of chain length, number of primary polymers and number of combined parts have to be computed in order to find the convolution sums. We have developed an efficient method that completes this task at acceptable computational effort. It employs the leading moments of the individual distributions $R(n,n_p,n_c)$ in combination with a global representation of the convolution sums. The details of this method are presented in Appendix.

Now, having explained the new algorithms, we briefly draw the attention to the impact of the fact that termination is diffusion controlled leading to a chain length dependent termination rate [15]. We can now see that under such circumstances the reaction steps that lead to the ultimate branched structure, transfer to polymer and recombination termination, are the same as in the case without chain length dependent termination. It is the relative probabilities of these steps – as expressed in the various probability density functions – that indeed are changing. Hence, the algorithms fully cover the chain length dependent termination case, but the pdf's have to be reconstructed on the basis of available relationships between chain length of macro-radicals and rate coefficients.

7. How to analyze primary polymer sequences and lengths

As in the disproportionation only case, we dispense with the explicit chronology of growth times of primary polymers (or residence times), again assuming that the *sequence* of events is the decisive factor. The existence of just one unique sequence per molecule of given number of primary polymers in the case of disproportionation only allowed us to compare primary polymer lengths as varying (decreasing) with sequence order between full and conditional Monte Carlo simulations. Now, the previous explanation of the algorithms applicable to the recombination case may have revealed that molecules of given dimensions may have different sequences. This complicates the comparison to the full Monte Carlo sampling method, since lengths as a function of sequence order cannot be determined in a straightforward manner. On the other hand, the particular sequencing structure of a molecule also provides information about its architecture. We will use this information in the sense that we will record the number of initial primary polymers in a molecule, which in fact equals the number of sequences. As regards lengths of primary polymers, it is still possible to compare the lengths of the primary polymer formed in the last event, as well as the lengths of the initial primary polymers. We will indeed give a report on these quantities in the following section.

8. Results

Reaction equations, population balances, kinetic and simulation data for both full and conditional MC algorithms are listed in Tables 1–3. All results have been obtained using this data set. The essential part of the full MC algorithm was implemented according to the original description published by Tobita [11–13]. However, in order to obtain the explicit graph theoretical description of the generated molecules

Table 2
Kinetic and simulation data

	Symbol	Value	Unit
Dissociation	k_d	0.5	s^{-1}
Propagation	k_p	5000	$m^3\,kmol^{-1}\,s^{-1}$
Disproportionation termination	k_{td}	5×10^6	
Recombination termination	k_{tc}	5×10^6	
Transfer to polymer	k_{tp}	1.5	
Average residence time CSTR	τ	30	s
Feed monomer concentration	M_f	16.75	$kmol\,m^{-3}$
Feed initiator concentration	$I_{2,f}$	5×10^{-3}	
Monomer concentration	M	9.1067022	
Macro-radical concentration	λ_0	5.58850×10^{-6}	
Incorporated monomer concentration	μ_1	7.638741575	

Table 3
Full Monte Carlo simulation parameters

MC simulation parameter	Formula	Value
Number average chain length	$\overline{n}$ (Eq. (3))	676.807243
Average branching density	$\overline{\rho} = k_{tp}\mu_1\tau$	2.5148257×10^{-4}
Branching probability	$P_b = k_{tp}\lambda_0/\{(k_{tp}+k_{td}+k_{tc})\lambda_0 + k_{tp}\mu_1\}$	0.170145237
Recombination probability	$P_c = k_{tc}\lambda_0^2/\{(k_{tc}+k_{td})\lambda_0^2 + k_{tp}\lambda_0\mu_1\}$	0.41492738

we had to add new features to the algorithm. We also observed that a fair amount of extra computation time was associated with these extra features.

The overall chain length and number of branch points distributions for the conditions listed in Tables 1–3 have already been presented before [9,10]. Since the specific chain length and number of primary polymer distributions for molecules consisting of 11 combination parts, $n_c = 11$, will serve as a test case for the conditional Monte Carlo simulations, we present these as resulting from the Galerkin-FEM calculations and full Monte Carlo simulations in Figs. 4 and 5. Note that the chain length distribution is the weighted concentration distribution of dead chains, $\sum_{n_p=20}^{\infty} nP(n, n_p, n_c = 11)$, which is directly comparable to the distribution of the sample from full Monte Carlo simulations with $n_c = 11$, in view of the weight fraction sampling procedure in those simulations. Note that the lowest possible number of primary polymers, n_p, for $n_c = 11$ equals 20. Similarly, the number of primary polymers' distribution shown is a weighted distribution, $\sum_{n=20}^{\infty} nP(n, n_p, n_c = 11)$. The value 20 as the lower chain length limit in the summation is a theoretical one, the practical one being much higher. Apart from scatter due to the limited full Monte Carlo sample size (45,563 molecules), both distributions coincide. All quantities to be calculated for the full

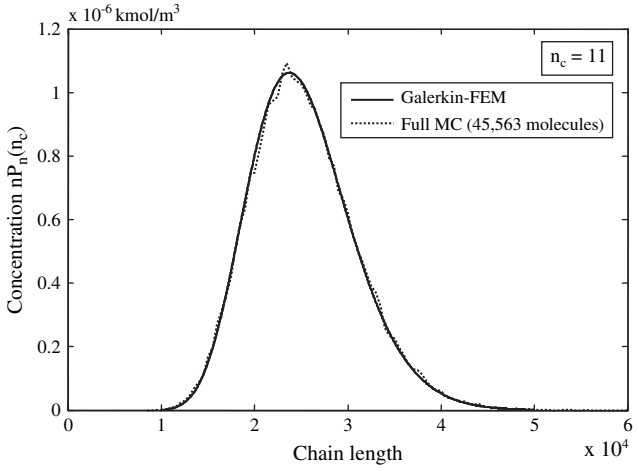

Fig. 4. Weighted chain length distribution, $nP_n(n_c = 11) = \sum_{n_p=20}^{\infty} nP(n, n_p, n_c = 11)$, of dead molecules consisting of 11 parts from Galerkin-FEM method and full Monte Carlo simulations for conditions in Tables 1–3.

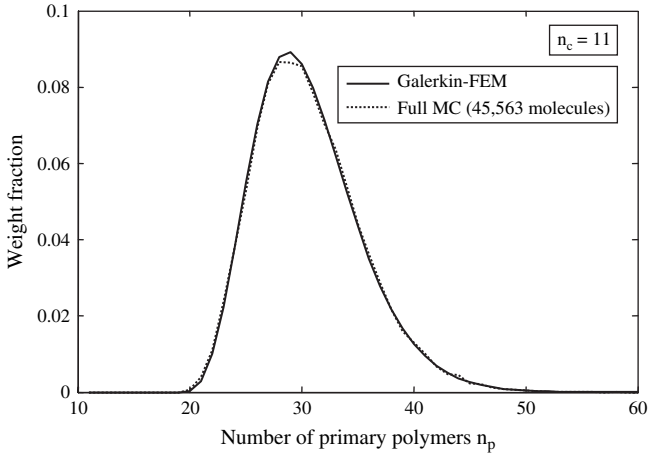

Fig. 5. Number of primary polymers' distribution, $\sum_{n=20}^{\infty} nP(n, n_p, n_c = 11)$, of molecules consisting of 11 parts (corresponding to chain length distribution of Fig. 4) from Galerkin-FEM method and full Monte Carlo simulations for conditions in Tables 1–3. The minimum number of primary polymers is 20.

Monte Carlo case with $n_c = 11$ are inferred from this sample. Chain lengths and number of primary polymers as input data for the conditional Monte Carlo architecture simulations are sampled from the distributions shown in Figs. 4 and 5. We also have used the full Monte Carlo generated chain lengths and number of primary polymers as input, but the results from this could not be distinguished from the ones we actually are going to present. Only for smaller sample size, as will be shown for larger molecules, it is recommendable to use the full Monte Carlo data as input for conditional Monte Carlo to better compare architecture samples. It must be realized, however, that the variability in architectures even for one fixed set of n, n_p and n_c is considerable (see e.g., Figs. 12 and 16).

Results will now be discussed in two parts. First, values and trends of probability density functions as presented in the previous section will be presented and compared to full Monte Carlo simulation data. Secondly, samples of molecules generated with both conditional and full Monte Carlo simulations will be analyzed using several architectural indicators.

8.1. Values and trends of probability density functions

These are calculated using the definitions presented before from the dead and living chain distribution Galerkin-FEM representations. It is now interesting to calculate such values from full Monte Carlo simulations. This required some extra effort, since obviously such features are not normally employed in the algorithm. Computational means had to be implemented to analyze the two pieces of a molecule at either side of a combination point. To this end the growth history of molecules generated by full Monte Carlo simulation had to be reconstructed from their architectures.

The probability density function describing the probability that a molecule of given dimensions has been created by a termination by recombination reaction, as defined by Eq. (10), is plotted in Fig. 6 as a function of chain length for the cases of a smaller molecule of $n_c = 3$ combination parts with $n_p = 3$

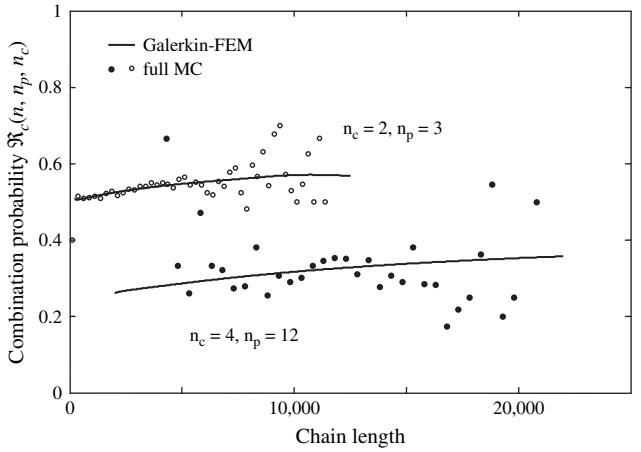

Fig. 6. Combination probability $\Re_c(n, n_p, n_c)$ as defined by Eq. (10) for 3 combination parts/3 primary polymers and 4 combination parts/12 primary polymers as a function of chain length from Galerkin FEM and full Monte Carlo simulations.

primary polymers and a larger molecule of $n_c = 4$ combination parts with $n_p = 12$ primary polymers. The values from the full Monte Carlo simulations feature considerable scatter due to the limited sample size (some thousands of molecules) – more for the smaller sample of the larger molecule – but still we observe good agreement. These pdf's turn out to be only weak functions of chain length.

Fig. 7 shows the probability that a molecule of $n_c = 4$ combination parts has been formed from a molecule with $n_{c1} = 1$ and a molecule with $n_{c2} = 3$ combination parts as a function of chain length, according to the definition of Eq. (11). Note that this one of the two options possible for such molecules, the other being $n_{c1} = n_{c2} = 2$. To create a sufficiently large full Monte Carlo sample the summation of the probabilities

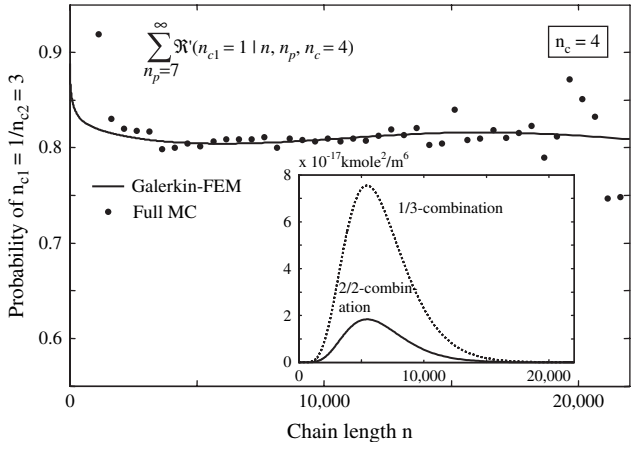

Fig. 7. Probability that a molecule of $n_c = 4$ combination parts has been formed from a molecule with $n_{c1} = 1$ and a molecule with $n_{c2} = 3$ combination parts as a function of chain length, n, according to Eq. (11), calculated by Galerkin FEM and simulated by full MC. To create a sufficiently large full MC sample the summation of the probabilities over all possible total number of primary polymers, n_p (minimum 7) has been taken: $\sum_{n_p=7}^{\infty} \Re'(n_{c1} = 1|n, n_p, n_c = 4)$ – the quote implying that the probability density function has been re-normalized. The inset shows the magnitude of the corresponding convolution sums (Eq. (11), Galerkin FEM) versus chain length for both combination options: 1/3 and 2/2.

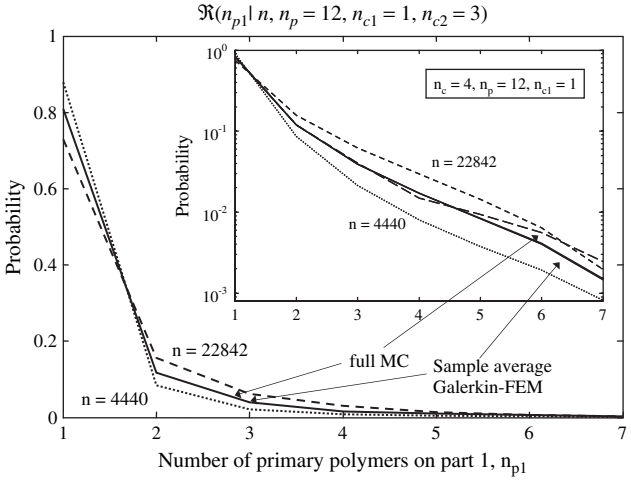

Fig. 8. Probability density function for number of primary polymers on combination part 1, n_{p1}, given total chain length, n, number of primary polymers, $n_p = 12$, number of combination parts, $n_c = 4$, and number of combination parts on part 1, $n_{c1} = 1$: $\Re(n_{p1}|n, n_p = 12, n_{c1} = 1, n_{c2} = 3)$ (Eq. (12)). By full MC a sample of such molecules is generated having various lengths. For these lengths the pdf's according to Eq. (12) are computed with Galerkin FEM. Differences at higher n_{p1} are due to small sample size of full MC. Also pdf's from Galerkin-FEM for extreme chain lengths in this sample are shown.

over all possible total number of primary polymers, n_p (minimum 7) has been taken: $\sum_{n_p=7}^{\infty} \Re'(n_{c1} = 1|n, n_p, n_c = 4)$ – the quote implying that the probability density function has been re-normalized. The inset of Fig. 7 shows the magnitude of the corresponding convolution sums (Eq. (11), Galerkin FEM) versus chain length for both combination options: 1/3 and 2/2. Despite some scatter we again conclude that good agreement exists between full Monte Carlo and the Galerkin-FEM results, as well as a weak variation of the pdf with chain length.

In Fig. 8 a complete probability density function is shown. It represents the probability distribution of number of primary polymers on one of the two combination parts, n_{p1}, given total chain length, n, total number of primary polymers, $n_p = 12$, total number of combination parts, $n_c = 4$, and number of combination parts on part 1, $n_{c1} = 1$: $\Re(n_{p1}|n, n_p = 12, n_{c1} = 1, n_{c2} = 3)$ as defined by Eq. (12). The sample of such molecules generated by full Monte Carlo contains various lengths. The comparison with Galerkin FEM, for which we can compute the pdf for a specified chain length, has been made on the basis of exactly the same various lengths. In addition two pdf's are shown for specified chain lengths: $n = 4440$ and 22,842. Differences at higher n_{p1} are only visible in the semi-logarithmic inset of Fig. 8, being caused by the small sample size of the full Monte Carlo results.

8.2. Architectural results

One interesting source of information on the architecture of a branched molecule formed by recombination reactions arises when regarding the number of *connecting points* on the parts of the molecules. Larger molecules may possess numerous different number of connecting point distributions, whereas the

smallest just have one. For instance, a molecule with one combination point *always* consists of two parts with each one connecting point, and a molecule with two combination points *always* consists of two parts with one connecting point and one part with two connecting points. However, for a molecule with three combination points two options exist. Either it has three parts with one connecting point and one with three connecting points or it has two with one and two with two connecting points. Hence, a molecule of $n_c = 4$ combination parts may have two different architectures as its way of being composed from the parts is concerned. Note that this is just a manner of analyzing architectures rather than describing the way the molecule has been growing, which is an intermingled process of growing of the parts and coupling of them.

In Fig. 9 the frequency distribution of the number of connecting points per parts is shown for samples of molecules with $n_c = 11$ combination points. The plot has been obtained by counting the parts with a certain number of connecting points among all the molecules of the sample. By definition, each molecule contains a number of parts with one connecting point, hence the relative abundance of such parts is large. At the other side of the spectrum, only very few molecules contain one part with 10 connecting points (plus 10 parts with one connecting point). Obviously, the molecules of the samples may differ in number of primary polymers, n_p, and chain length, n. The samples underlying Fig. 9 are the same as those in Figs. 4 and 5 showing the n and n_p distributions, respectively. In the semi-logarithmic plot of Fig. 9 no difference is visible between the results from the conditional and full Monte Carlo architectures. Hence, we conclude that from the perspective of number of connecting points per part distribution both methods yield identical results. Fig. 10 shows similar distributions for molecules with 40 combination parts, $n_c = 40$. Since such molecules are very rare it took a few days of full Monte Carlo simulations to generate a sample of 268 molecules, obviously of different lengths and number of primary

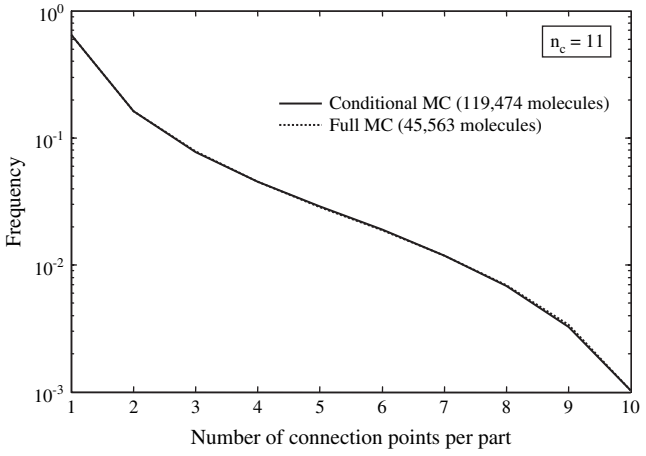

Fig. 9. In a branched molecule with 10 combination points consisting of 11 combined parts, each part has between 1 and 10 connection points. The frequency distribution of the number of connection points per parts is strongly dependent on architecture. This figure shows these distributions in samples generated by conditional and full MC. Molecules in this sample have lengths, n, and numbers of primary polymers, n_p, distributed as shown in Figs. 4 and 5.

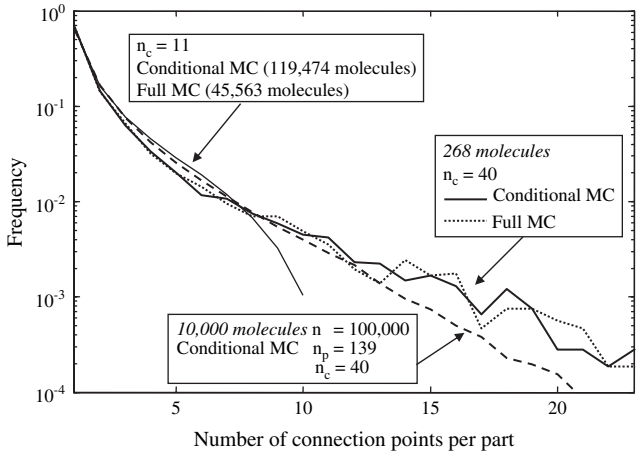

Fig. 10. Distribution of number of connection points for samples of a large molecule, $n_c = 40$. The varying chain lengths, n, and number of primary polymers, n_p, for the conditional MC sample are taken identical to those generated by full MC. These necessarily small samples show considerable scatter, but globally agree. For comparison, the distribution of a larger sample of large molecules from conditional MC for a single n, n_p, n_c-combination is shown.

polymers. The comparison has been made with architectures from conditional Monte Carlo simulation for the same set of chain lengths and number of primary polymers. In the figure the results feature equal scatter, but on average they overlap. The curve for a sample of 10,000 molecules of specified dimensions ($n_c = 40$, $n_p = 139$, $n = 100{,}000$) using the conditional method is considerably smoother and clearly deviates from that for the smaller samples of non-specified molecules.

During the introduction to the history backtracking algorithm the concept of initial primary polymers was already described; their existence is caused by the fact that growth of a molecule may be initiated at different sites and different instants in time. The initial primary polymers form the starting point for an equal number of growth sequences that ultimately all merge together. All molecules formed by recombination termination possess at least one initial primary polymer. A molecule consisting of n_c-combination parts may have an equal number of initial primary polymers at maximum. Fig. 11 shows the frequency distribution of number of initial primary polymers for the samples from full and conditional Monte Carlo sampling of $n_c = 11$ introduced before. About 10% has just one initial primary polymer, roughly one-third has 3 and very few have 9 initial primary polymers; none out of these samples have the theoretical maximum of 11 initial primary polymers. The curves from the conditional and full Monte Carlo sampling exactly overlap for the largest parts of the curves, except for the rare molecules in the higher number of initial primary polymer region, only visible in the semi-logarithmic inset of Fig. 11. Hence we conclude that also from the perspective of this particular architectural property both the Monte Carlo methods lead to identical results. Fig. 12 confirms this for samples of larger molecules ($n_c = 40$) – the same as used to construct Fig. 10 – be it with a larger scatter due to the much smaller sample sizes.

Fig. 13 compares the lengths of initial primary polymers as a function of the number of initial primary polymers in the

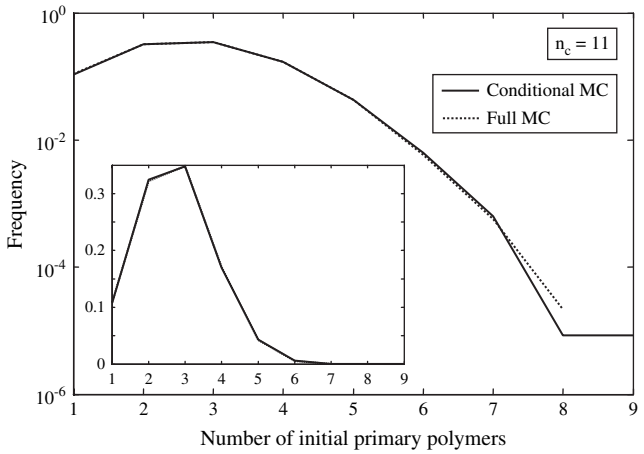

Fig. 11. A branched molecule with 10 combination points has grown from at least 1 and at maximum 11 initial primary polymers, strongly dependent on architecture. This figure shows the frequency distribution for the same samples as in Fig. 9 generated by conditional and full MC. Maximum number found from the former is 9 (2 molecules) and for the latter is 8 (1 molecule).

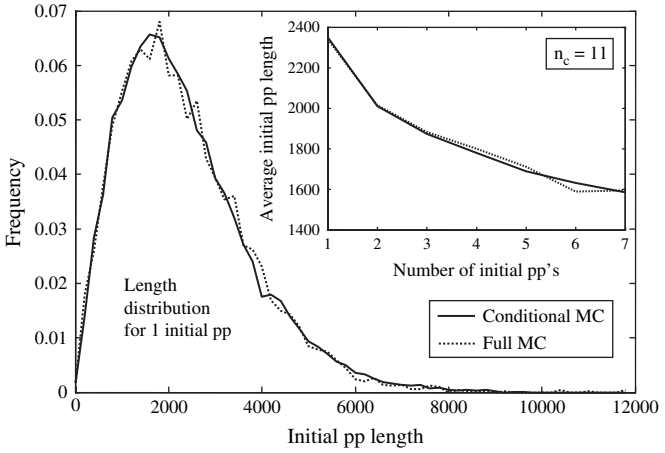

Fig. 13. The inset shows the average length of initial primary polymers as a function of the number of initial molecules resulting from conditional and full MC, same samples as in Fig. 9, overall length distribution in Fig. 4. Such primary polymers turn out to be smaller when their number per molecule is larger. The length distribution of the initial primary polymer for molecules containing a single one from both MC methods shows good agreement.

molecule. The average length (inset) is seen to decrease with the number of initial primary polymers; the curves from the two Monte Carlo methods practically overlap. The length distributions of the initial primary polymer for molecules with just one such primary polymer are also observed to agree very well. Fig. 14 shows that the length distributions of the last grown primary polymer of the molecules agree equally well. We hence conclude that also with respect to the details of these special primary polymers the conditional and full Monte Carlo methods yield identical results.

Finally, we show the results of comparisons on a third architectural qualifier: the radius of gyration contraction factor in Figs. 15–17. The contraction factor distributions for the samples with $n_c = 11$ combination parts (same sample as above) from both methods are fairly smooth and feature

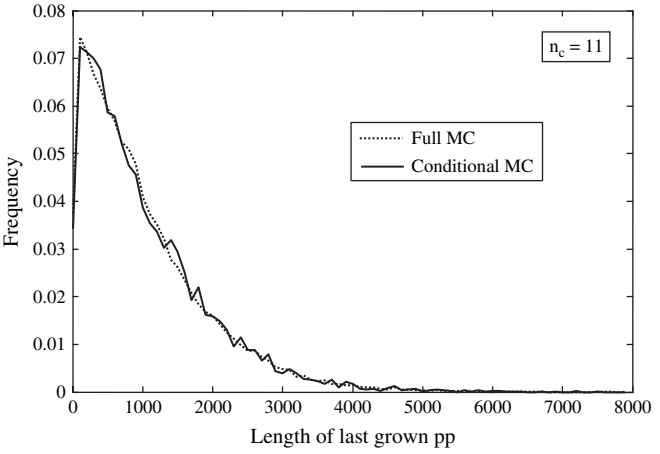

Fig. 14. Length distribution of last grown primary polymer for the samples with $n_c = 11$ of Fig. 9 from conditional and full MC, showing good agreement.

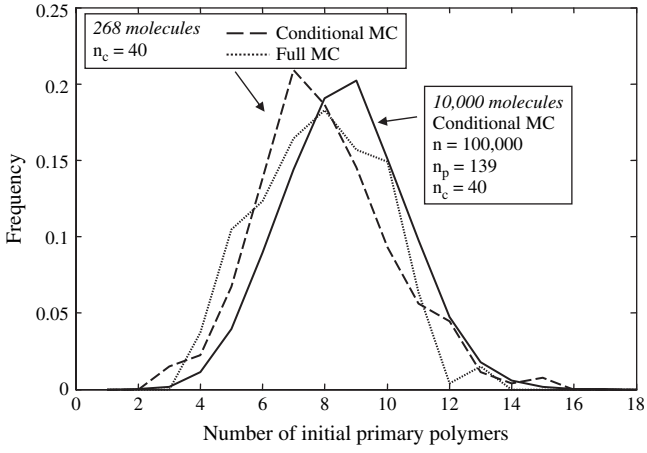

Fig. 12. Frequency distribution number of initial primary polymers for the same samples of large molecules, $n_c = 40$, as in Fig. 10 generated by conditional and full MC. Although the theoretical lower and upper limits are 1 and 40, in these samples they turn out to be 3 and 15. Small samples from both MC methods globally agree, while the larger conditional MC sample for one n, n_p, n_c-combination is more regular and lightly shifted to higher numbers.

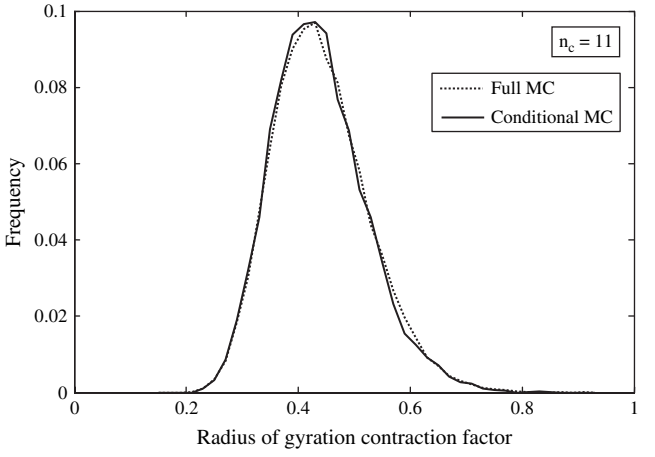

Fig. 15. Radius of gyration contraction factor distribution as calculated from the architectures [10]. Samples used, for $n_c = 11$, are the same as in Fig. 9. Results from conditional and full MC show excellent agreement.

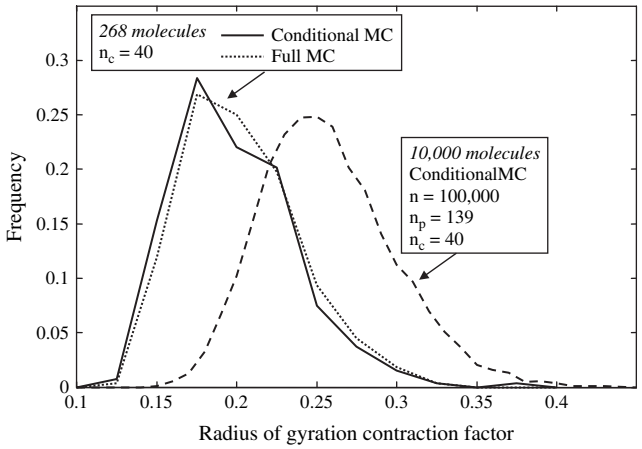

Fig. 16. Radius of gyration contraction factor distribution as calculated from the architectures of large molecules, $n_c = 40$ [10]. Samples used are the same as in Fig. 10. The smaller samples from conditional and full MC show, despite fair scatter, good agreement. These larger molecules, more strongly branched, feature more important contraction than the smaller ones in Fig. 15. Again, the conditional MC results for one single n, n_p, n_c-combination are more regular, but they show considerably weaker contraction.

excellent overlap. Again, the smaller samples of larger molecules (268) lead to more scatter, but still the similarity between the distributions is strong. This is especially true when comparing these distributions to the one obtained for a much larger sample (10,000) of specified molecules, which is considerably shifted towards the weaker contraction region.

To illustrate the usefulness of the conditional MC method as regards finding the influence of kinetics or other circumstances on architectures, the relation between termination mechanism and architecture is now discussed. We have posed ourselves the following question. Suppose, we have molecules

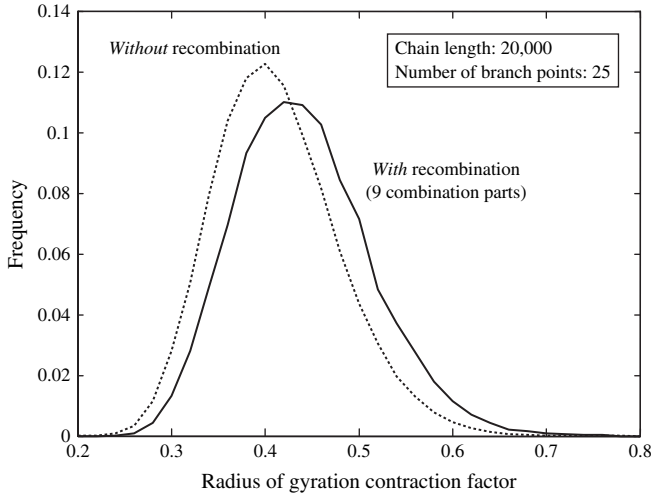

Fig. 17. Radius of gyration contraction factor distribution for molecules made without recombination termination ($k_{td} = 10^7\ m^3\,kmol^{-1}\,s^{-1}$) and with recombination ($k_{tc} = k_{td} = 5 \times 10^6\ m^3\,kmol^{-1}\,s^{-1}$) and further identical kinetic conditions. All molecules have chain length $n = 20{,}000$ and $N = 25$ branch points. The recombined molecules consist of 9 recombination parts. The plot indicates that molecules made by disproportionation only are more compact than those *of the same dimensions (n,N)* made by both recombination and recombination. Sample sizes: 100,000 molecules.

of identical dimensions in terms of chain length and number of branch points, but different in termination mechanism. One kind has been made without recombination, so disproportionation only, while the other did undergo recombination termination indeed. Which of these should be expected to be more compact as measured from the radius of gyration contraction factor? Before comparing the architectures, it should be noted that the averages in chain length and number of branch points of molecules made in the presence of recombination are higher than those made without it. Taking molecules with identical dimensions therefore implies sampling from the *lower* end of the whole population of chain lengths and number of branch points for the first category and from the *higher* end for the second. The answer on the question on compactness is given in Fig. 17 for the example of chain length $n = 20{,}000$ and number of branch points, $N = 25$. The kinetic conditions are identical, in both cases the overall termination rate coefficient equals $10^7\ m^3\,kmol^{-1}\,s^{-1}$, but in one case $k_{td} = 10^7\ m^3\,kmol^{-1}\,s^{-1}$, while in the other $k_{tc} = k_{td} = 5.10^6\ m^3\,kmol^{-1}\,s^{-1}$. Architectures were calculated for samples of 100,000 molecules in both cases, a matter of few hours computation time. The conclusion is that the molecules made by disproportionation turn out to be slightly, but clearly more compact than those having experienced recombination. Note that this does *not* mean that molecules having undergone recombination in general are less compact than molecules of the other category. Partly, the result of Fig. 17 must be attributed to what was observed about the sampling of molecules of identical dimensions, namely that this happens from opposite parts of the complete populations. It should also be noted that generating samples of 100,000 molecules of exactly these dimensions is practically impossible with the full MC method. In view of the small difference between the two kinds of molecules, such a conclusion could not possibly have been drawn from full MC.

9. Conclusions and discussion

Our primary goal in this article is to develop a model to find the architectures of polymer molecules of given chain length and number of branch points, polymerized by a radical mechanism with transfer to polymer and recombination termination. It was precisely this conditionality that the molecules are of specified dimensions – and fixed but as yet unknown architecture – that forced us to think of a method that could find the history of chemical events leading to the architecture in *anti*-chronological order. In this particular kinetic case these events are condensed to *two* types both involving branch point formation or chain initiation. In one type this reaction is followed by termination without recombination, resulting in addition of an arm. In the other type simultaneous branch formation/initiation steps at two different molecules are followed by a recombination termination step connecting these molecules. Considering the well-known radical polymerization process in this manner offers a nice view on the essential parts of the growth of a branched structure, as shown in Fig. 2. The probability, that one or the other event is happening follows

from probability density functions that directly are inferred from concentration distributions of living and dead molecules of certain dimensions. As has been argued in previous work [10], apart from the chain length dimensions and number of branch points, the number of combination points as a third dimension has to be taken into account in an explicit manner. The importance of the number of combination points for the architecture is obvious, if one realizes its impact on the connectivity possibilities in a molecule. The 3D-concentration distribution can be obtained in several ways, by solving 3D-population balances or by a full Monte Carlo scheme [11–13]. We choose to solve the population balances using an improved version of the earlier published method [10] (see Appendix), which turned out to be the most efficient manner finding the concentrations in a sufficiently accurate way. Finally, the growth history backtracking algorithm (see Fig. 3) constructs possible growth paths using the concept of the two architecture determining events and the proper probability density functions and thus finds statistically representative samples of architectures.

To test the validity of our approach we performed several comparisons to the full Monte Carlo scheme. Although the latter does not naturally provide the probability density functions required for the conditional MC method, it still contains the information to infer these functions. By developing means to infer this information from the full MC method we could compare our conditional method to the full MC method on the level of probability density functions. However, this could only be realized for small molecules, since the full MC method does not generate sufficiently large samples of larger molecules in reasonable computation time to warrant statistically sound conclusions – which precisely constitutes a major incentive to invent the conditional MC method after all. Comparisons to full MC have also been performed on the level of resulting architectures, again and for the same reason for small molecules only. In all cases we found sufficient agreement to conclude that both methods represent the same architectural growth process – be it in quite different ways.

Thus, the validity of the conditional MC method has been proven for the case of a CSTR at steady state with transfer to polymer and recombination termination. As regards the extendibility of the conditional MC method to other reactor and kinetic systems in general it is required that the proper concentration distributions are available. A very easy extension is including transfer to small molecules reactions. Adding such reactions does not lead to qualitatively different architectures, but rather have a quantitative effect similar to increasing the disproportionation termination rate. Another possible extension is that to batch reactors. In the present case we need 3D chain length/number of branch points/number of combination point distributions. For a batch reactor with the same kinetics this distribution should be available as a function of conversion. Especially interesting are extensions to those systems that cannot be dealt with using the full MC method, like radical polymerization with chain length dependent termination or accounting for limited accessibility of the inner part of larger molecules to the transfer to polymer reaction. In this case the termination depends on the overall size of macro-radicals, which is a quantity not explicitly featuring in the full MC method, for which this mechanism therefore poses a serious fundamental problem. In contrast, in the conditional MC method the implementation of chain length dependent termination is straightforward, since it does not require a modification at all, once the proper probability density functions are constructed. A limited accessibility of the inner part of molecules to transfer to polymer, not accounted for in the present work, again gives rise to fundamental problems in the full MC method, since the growing architecture is not explicitly featuring in its concept. The architectural information required to distinguish between 'inside' and 'outside' of a molecule indeed is available in the conditional MC method, which in principle would offer a key to solving this complex problem. Summarizing, we conclude that several options exist to extend the conditional MC to a wider scope of applications.

Appendix A. Modified discrete Laguerre polynomials for fast evaluation of convolution sums

Finding the rate $r_c(n, n_p, n_c)$ and the pdf $\Re(n_{c1}|n, n_p, n_c)$, Eqs. (8) and (11) of the main text, forms a huge combinatorial problem, since in principle all possible combinations of chain length, number of primary polymers and number of combined parts have to be computed in order to evaluate the convolution sums. We have developed an efficient method that completes this task at acceptable computational effort. It employs the leading moments of the individual distributions $R(n,n_p,n_c)$ in combination with a global representation of the convolution sums based on the Galerkin h-p method, also used in PREDICI® [14].

Discrete Galerkin methods

A discrete Galerkin method for the representation of chain length distributions has been introduced by Wulkow [14] in a weighted finite element h-p-description. This method has already found many applications as it is implemented in the PREDICI® software package. The earlier studied weighted method (Refs. [8,9] in Ref. [6]) makes use of a *global* representation, based on discrete orthogonal Laguerre polynomials for the whole chain length axis, and *local* representations on chain length intervals, based on discrete *Chebyshev* polynomials. The global representation allows fast handling and has some interesting properties, as for instance a direct relation to the leading moments of a distribution. However, its success depends on the form of the distributions to be approximated. The locally adapted h-p-representation can be highly accurate by choosing chain length intervals sufficiently small and the order sufficiently high. Here no weight function is employed. In this paper we will make use of both approaches and review the weighed method first.

The basic form of the global representation is, in one (chain length) dimension:

$$\Psi_{\rho,\alpha}(n) = (1-\rho)^{1+\alpha}\binom{n-1+\alpha}{n-1}\rho^{n-1,} \tag{A1}$$

which for $\alpha = 0$ becomes a Flory distribution. The associated orthogonal polynomials are modified discrete Laguerre polynomials, as follows:

$$l_s(n;\rho,\alpha) = \sum_{t=0}^{s}\rho^{s-t}(\rho-1)^t\binom{s+\alpha}{s-t}\binom{n-1}{t}, \tag{A2}$$

where s is the order of the polynomial. Distributions are represented as:

$$P_{\rho,\alpha}(n) = \Psi_{\rho,\alpha}(n)\sum_{t=0}^{s}a_t(n)l_t(n;\rho,\alpha), \tag{A3}$$

where the α_t are the expansion coefficients, to be found in varying ways, depending from the problem at hand. The parameters ρ and α are related to the leading moments by:

$$\rho = \frac{\mu_0\mu_2 - \mu_1^2 - \mu_1\mu_0 + \mu_0^2}{\mu_0\mu_2 - \mu_1^2}$$

$$\alpha = \frac{2\mu_1^2 - \mu_1\mu_0 - \mu_0\mu_2}{\mu_0\mu_2 - \mu_1^2 - \mu_1\mu_0 + \mu_0^2,} \tag{A4}$$

while the higher coefficients are related to the moments by:

$$\mu_s = \sum_{t=0}^{s}b_{st}a_t\gamma_t^{\rho,\alpha}, \tag{A5}$$

where the b_{st} follows as:

$$n^t = \sum_{t=0}^{s}b_{st}l_t(n) \tag{A6}$$

and $\gamma_t^{\rho,\alpha}$ comes from the orthogonality relation:

$$\sum_{n=1}^{\infty}l_s(n)l_t(n)\Psi_{\rho,\alpha}(n) = \delta_{st}\gamma_s^{\rho,\alpha}, \quad \gamma_s^{\rho,\alpha} = \rho^s\binom{s+\alpha}{s}. \tag{A7}$$

The local approximation is based on Chebyshev polynomials of the first kind, $l_s(x)$, on x-intervals $[-1,1]$, corresponding to chain length intervals $[n_1,n_2]$. The expansion coefficients a_t follow from the distribution values $P(x_u)$ by:

$$\begin{aligned} a_t &= \frac{2}{s}\sum_{u=1}^{s}P(x_u)l_t(x_u) \\ &= \frac{2}{s}\sum_{u=1}^{s}P\left[\cos\left\{\frac{\pi(u-\frac{1}{2})}{s}\right\}\right]\cos\left\{\frac{\pi t(u-\frac{1}{2})}{s}\right\} \end{aligned} \tag{A8}$$

The distribution is then approximated by the h-p representation:

$$P(x) = \sum_{t=0}^{s-1}a_t l_t(x) - \frac{1}{2}a_0. \tag{A9}$$

A fast and accurate manner to evaluate the numerous convolution sums mentioned above we found to proceed is as follows. An accurate solution of the 3D-problem was found using the Galerkin h-p method to solve, for each number of branch points, a 1D (chain length) problem. This proceeds in a recursive manner, since the balance equations for i branch points can be solved using the solutions for those $<i$. It must be realized that these solutions should be accurate, since even small errors may propagate from lower through higher i, in our case up to $i = 100$. For this reason, the Galerkin h-p method is employed here. This yields a 2D-solution (n,i). To solve the complete 3D-problem, two additional number of combination point moments is computed in a manner explained elsewhere [9]. Further following this work on the basis of the three combination point moments and using another type of orthogonal polynomials, the *Krawtchouk* polynomials, we arrive at a full solution of the 3D-set. This solution consists of h-p representations for each individual combination of number of branch points and combination points, around 5000 in total. Typically, one h-p representation in this case consists of 15 coefficients (order 15) at around 20 intervals. It includes the interval boundaries as well as the leading moments of the distributions.

In order to find a convolution sum $P_c(n)$ from two distributions, P_1 and P_2 as a distribution in n, $P_c(n) = \sum_{m=1}^{n-1}P_1(m)P_2(n-m)$, we calculate the moments of P_c from those of P_1 and P_2 (stored in their h-p representations): $m_1(0\ldots k)$ and $m_2(0\ldots k)$. For this we use the expression:

$$m_c(j) = \sum_{i=0}^{k}a(n,i)m_1(i)m_2(j-i), \tag{A10}$$

where $a(n,i)$ are the elements of a coefficient matrix:

$$\mathbf{a} = \begin{bmatrix} 1 & & & & & \\ 1 & 1 & & & & \\ 1 & 2 & 1 & & & \\ 1 & 3 & 3 & 1 & & \ldots \\ 1 & 4 & 6 & 4 & 1 & \\ 1 & 5 & 10 & 10 & 5 & 1 \\ \ldots & & & & & \end{bmatrix}. \tag{A11}$$

Using the relations between moments and the expansion coefficients in the global representation given above, we obtain these coefficients and store them. In the architectures algorithm they are easily retrieved for point- or distribution-wise evaluation of the convolution sum of any two distributions P_1 and P_2.

Performance

The accuracy and computation speed are now demonstrated for two known distributions to be convoluted:

$$\begin{aligned} &P_1 = (1-\rho_1)\rho_1^{(n-1)}n^2;\ P_2 = (1-\rho_2)\rho_2^{(n-1)}n^3; \\ &\rho_1 = 1 - 1/1000;\ \rho_2 = 1 - 1/3000. \end{aligned} \tag{A12}$$

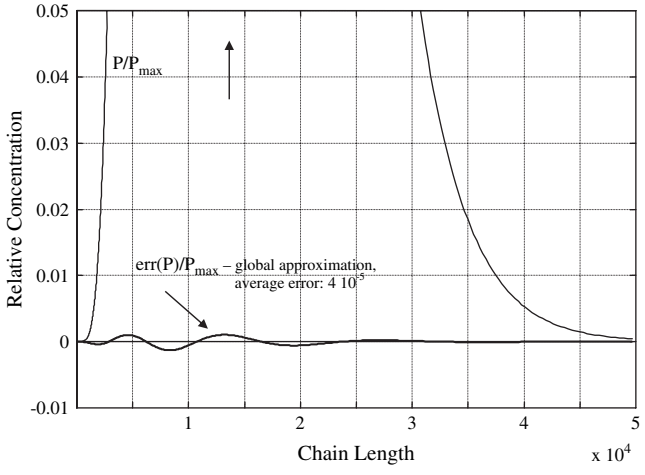

Fig. A1. Convolution distribution P_c/P_c^{max} and errors of global and local approximations. Average error local approximation 3.7×10^{-9} – too small to be visible on this figure's scale.

P_1 and P_2 are relatively narrow distributions being representative for the $R_{n,i,k}$ (main text, Eq. (3)) in the underlying problem. Obviously, in this example, the convolution distribution $P_c(n)$ can be calculated directly for each n. Now first, we approximate the convolution distribution $P_c(n)$ using the accurate local h-p representation by first finding the h-p presentations of P_1 and P_2 themselves and then evaluating it for P_c using an algorithm similar to that used in PREDICI®. Interval length and order of the Chebyshev polynomials were *not* optimized as happens in the package. Instead, we employed 20 intervals (logarithmically equidistant) and order 15. This was a slow procedure, but when using interval and order optimization as in PREDICI® typically 0.5–1 s CPU-time is required for one complete convolution operation with simultaneous grid adaptation. The relative error $\Delta P_c^{rel} = \sum_{n=1}^{10^5} |P_c - P_c^{app}| / \sum_{n=1}^{10^5} P_c$ turned out to be 3.7×10^{-9}.

Secondly, we applied the global approximation, computing m_c from m_1 and m_2 and subsequent finding ρ, α d the coefficients a_c. This turned out to be very fast, taking 0.05 s CPU-time, while the relative error was still small, 4.1×10^{-5}. In this case taking up to 21 Laguerre polynomials with the equal number of expansion coefficients turned out to be optimal.

In a normal distribution plot both approximations to the eye coincide with the exact solution. Hence, to illustrate the difference, the error ΔP_c as relative to the maximum of P_c was plotted in Fig. A1, together with the (lower part of) distribution itself. The error of the global approximation here appears as the wavy curve, but in fact that of the local solution is still invisible at this scale.

We conclude that with the global approximation a fast and sufficiently accurate procedure has been found to evaluate the numerous convolution sums, Eq. (3) of main text, as a part of the underlying architectures problem.

References

[1] Zimm BH, Stockmeyer WH. J Chem Phys 1949;17:1301–14.
[2] Bick DK, McLeish TCB. Phys Rev Lett 1996;76(14):2587–90.
[3] Read DJ, McLeish TCB. Macromolecules 2001;34:1928–45.
[4] Iedema PD, Slot JJM, Kim D-M, Hoefsloot HCJ. Macromol Theory Simul 2004;13:400–18.
[5] Costeux S, Wood-Adams P, Beigzadeh D. Macromolecules 2002;35:2514–28.
[6] Hoefsloot HCJ, Iedema PD. Macromol Theory Simul 2003;12:484–98.
[7] Iedema PD, Hoefsloot HCJ. Macromol Symp 2004;206:93–106.
[8] Iedema PD, Hoefsloot HCJ. Polymer 2004;17:6071–82.
[9] Iedema PD, Hoefsloot HCJ. Macromol Theory Simul 2005;14:505–18.
[10] Iedema PD, Hoefsloot HCJ. Macromolecules 2006;39:3081–8.
[11] Tobita H. Polym React Eng 1993;3:379.
[12] Tobita H. J Polym Sci B Polym Phys 1994;32:911.
[13] Tobita H. e-Polymers, http://www.e-polymers.org, 2004;032.
[14] Wulkow M. Macromol Theory Simul 1996;5:393–416.
[15] Buback M, Muller E, Russell GT. J Phys Chem A 2006;110:3222–30.