ELSEVIER

Chemical Engineering Science 63 (2008) 111–121

Chemical Engineering Science

www.elsevier.com/locate/ces

Fast Monte Carlo methodology for multivariate particulate systems-II: τ-PEMC

Roberto Irizarry*

DuPont Electronic Technologies, 14 T.W. Alexander Drive, P.O. Box 13999, Research Triangle Park, NC 27709-3999, USA

Received 12 December 2006; received in revised form 5 September 2007; accepted 6 September 2007
Available online 15 September 2007

Abstract

The point ensemble Monte Carlo (PEMC) was introduced [Irizarry, R., 2007. Fast Monte Carlo methodology for multivariate particulate systems-I: point ensemble Monte Carlo. Chemical Engineering Science, in press, doi:10.1016/j.ces.2007.09.007.] as a method to accelerate the simulation speed of particulate processes solved by Monte Carlo methods. The PEMC method is a "constructed" jump Markov model that approximates the dynamics of the original particulate process without losing a detailed description of individual particles. The PEMC method is integrated using the stochastic simulation algorithm, which is exact in time. A natural extension of the PEMC algorithm is to consider a coarse-graining strategy for the time to further accelerate the MC simulation. In this work, the τ-leap method is adapted to the PEMC. It is shown that when the τ-parameter is selected properly, the τ-PEMC can also give accurate results with faster computational speed than the PEMC method. Furthermore, similar to the PEMC, the dynamic of complex intra-particle phenomena can be represented accurately. Numerical experiments show that this algorithm can improve the computational load of the exact method by orders of magnitude without sacrificing computational accuracy. The methodology is useful especially in stochastic optimization applications where many function calls (simulations) are required.

Keywords: Population balance; Multivariate population balance; Monte Carlo simulation; Aggregate coagulation; Aggregate sintering; Nonlinear dynamics; τ-Leap method; Optimization

1. Introduction

Monte Carlo solutions of population balance equation were reviewed in Part I (Irizarry, 2007). These methods (Gillespie, 1975; Garcia et al., 1987; Liffman, 1992; Kruis et al., 2000; Maisels et al., 2004; Smith and Matsoukas, 1998; Lin et al., 2002; Ramkrishna, 2000) consist of an artificial realization of the population dynamics of a finite system (finite number of particles in a very small volume) to estimate the properties of the whole system. The estimate becomes exact as the number of particles approaches infinity. In many cases the simulation speed of the MC simulation is important. One example is the optimization and dynamic optimization of processes described by PBE (Irizarry, 2005, 2006).

The point ensemble Monte Carlo method (PEMC) was introduced by Irizarry (2007) to accelerate the simulation speed of population balances solved by the MC method. The PEMC method consists of replacing the simulation box with a set of sub-ensembles (representing characteristic sizes) and construction of a network of reactions called random product channels (RPC), which approximates the mechanisms of particle formation (i.e., aggregation, breakage, nucleation and growth). When a given RPC is fired, particles from existing sub-ensembles are selected to form newborn particles ("product" particles). Each product particle is then allocated into one of the existing sub-ensembles using probability functions that are property conserving. A very important aspect of the RPC structure is that a coarser view of the process is taken without losing the integrity of individual particles. In Irizarry (2007) it was demonstrated that this coarsening of the MC simulation can accelerate the speed by orders of magnitude with minimal loss of accuracy.

In this work, the τ-leap algorithm is embedded into the PEMC framework to further accelerate the simulation speed of the PEMC method. The modified PEMC strategy is called

* Tel.: +1 919 248 5246.
E-mail address: Roberto.Irizarry@usa.dupont.com.

0009-2509/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ces.2007.09.006

the τ-PEMC method. In the PEMC the only source of error was the coarsening of the original particle population into sub-ensembles and the number of particles used in the simulation. In the τ-PEMC method, a new possible source of error comes from the approximations of the τ-leap method. In this work it is shown that τ-PEMC can be $10\times - 50\times$ faster than PEMC with accurate results.

In Section 2 the stochastic simulation algorithm (SSA) and the τ-leap method are briefly reviewed. The PEMC framework is reviewed in Section 3. The τ-PEMC is described in Section 4. In Section 5 the performance of the approximated method is studied with complex coagulation kernels. In Section 6 competitive reactions in droplets are studied. Finally, in Section 7, simultaneous coagulation-breakup problems are considered.

2. Exact stochastic simulation and τ-method

The exact SSA of a well mixed chemical reaction system has been developed by Gillespie (1976). Let $S=(S_1,\ldots,S_M)$ be the vector of species and $N=(N_1,\ldots,N_M)$ the vector of number of molecules for each specie in the simulation volume. M is the total number of species in the system and T the total number of reactions. For example consider the parallel reaction system: $NaOH + HCl \rightarrow NaCl + H_2O$ and $NaOH + CH_2ClCOOC_2H_5 \rightarrow CH_2ClCOONa + C_2H_5OH$. This system has six species (M=6) NaOH (S_1), HCl (S_2), $CH_2ClCOOC_2H_5$ (S_3), NaCl (S_4), $CH_2ClCOONa$ (S_5), C_2H_5OH (S_6) with the following two reactions ($T=2$).

$$E_1 : S_1 + S_2 \rightarrow S_4, \tag{1}$$

$$E_2 : S_1 + S_3 \rightarrow S_5 + S_6. \tag{2}$$

Each reaction is considered an event E_i that can change the system from the current state, N, to a new state. The method consists of answering the following question: for a system in a given state what reaction occurs next, E_f, and when does it occur, τ_{micro}? The simulation generates random number pairs (τ_{micro}, E_f) from conditional probabilities which are functions of the propensity function for each reaction, $R(E_s)$, defined as

$R(E_s)\,\mathrm{d}t \equiv$ the probability that reaction E_s occurs in the time interval $[t, t+\mathrm{d}t)$.

In each iteration of the SSA algorithm, two random numbers r_1, r_2 are selected from the uniform distribution (0, 1). The time for the next reaction is calculated as

$$\tau_{\text{micro}} = \frac{-\log(r_1)}{\sum_{j=1}^{T} R(E_j)}. \tag{3}$$

The time is moved to $t+\tau_{\text{micro}}$. The reaction fired at $(t+\tau_{\text{micro}})$ is selected by solving the following equation:

$$\sum_{s=1}^{f} R(E_s) < r_2 \sum_{s=1}^{T} R(E_s) \leqslant \sum_{s=1}^{f+1} R(E_s). \tag{4}$$

The selected event is E_f. These steps are repeated until a final time is reached. As one reaction is fired at a time in a microscopic time interval, τ_{micro}, the method can be slow in some cases, for example, if the concentrations are high or if the number of reactions is very large.

The τ-leap is an approximated method introduced by Gillespie (2001) to accelerate the simulation speed of the SSA. Differently from the SSA method, in which one reaction is "fired" in a microscopic time (also determined as part of the algorithm), in the τ-leap method, all reactions are fired many times during a coarser time, τ ($\tau \gg \tau_{\text{micro}}$). The method is based on the assumption that τ is small enough that the propensity functions will not change appreciably during τ. When this condition is satisfied, all reaction channels can be considered as independent events and the number of firings for each reaction, E_j, is a Poisson random variable with distribution, $P_{PD}(k_j; R(E_j), \tau)$, where

$$P_{PD}(k; a, \tau) = \frac{\mathrm{e}^{-a\tau}}{k!}(a\tau)^k. \tag{5}$$

The algorithm simply consists of selecting τ and then for each reaction channel E_j generating a sample k_j from the corresponding Poisson random variable. The system is updated by firing each reaction E_j, k_j times and setting t to $t+\tau$. The accuracy and speed of the method depends on the selection of time τ during the simulation. A method for selecting τ was proposed by Gillespie and Petzold (2003). As the Poisson distribution is not bounded, it could generate negative values of the concentration.

Another improvement to the method is to replace the Poisson distribution with a binomial distribution (Tian and Burrage, 2004;
Chatterjee et al., 2005).

$$P_{BD}(k; p, k_{\max}) = \frac{k_{\max}}{k!(k_{\max}-k)!} p^k (1-p)^{k_{\max}-k}. \tag{6}$$

In this case the number of firings for each reaction is a sample, k_j, from a binomial random variable with distribution, $P_{PD}(k_j; p_j, k^j_{\max})$, where $k^j_{\max}$ is the maximum number of times reaction j can be fired (after consuming the limiting component). The firing probability for E_j is $p_j = R(E_j)\tau/k^j_{\max}$. The binomial distribution eliminates the problem of negative concentrations and is more robust with respect to larger τ values.

The τ-leap method will be embedded into the PEMC algorithm to further accelerate the PEMC algorithm. In theory, any version of the τ-leap method can be utilized for this purpose. In this work the version proposed by Chatterjee et al. (2005) is implemented. Also, from the different strategies for selecting τ, the one employed in Chatterjee et al. (2005) is used due to its simplicity: $\tau = f\langle\tau_{\text{micro}}\rangle$, where f is a coarse-graining factor ($f>1$).

3. Point ensemble Monte Carlo

The PEMC method was introduced in Irizarry (2007). The method consists in the SSA solution of the Jump Markov model called particle ensemble random product (PERP) model. This model is based on three ideas. First, the total population is