Investigation of the Sensitivity of the Monte Carlo Solution for the Barker-Ferry Equation with Sequential and Parallel Pseudo-Random Number Generators*

T.V. Gurov\textsuperscript{1} and P.A. Whitlock\textsuperscript{2}

\textsuperscript{1} CLPP - BAS, Acad. G. Bonchev st, bl. 25 A, 1113 Sofia, Bulgaria, gurov@cern.bas.bg, ivdimov@bas.bg
\textsuperscript{2} Dep. of Comp. and Inf. Sci., Brooklyn College - CUNY 2900 Bedford Ave, Brooklyn, NY 11210, USA whitlock@sci.brooklyn.cuny.edu

Abstract. Random numbers are used to introduce the stochastic dynamics in probabilistic numerical techniques such as the Monte Carlo (MC) method and Langevin dynamics. For computational purposes random numbers are obtained by deterministic rules, implemented in pseudo-random number generators (prng's) which usually rely on simple arithmetic operations. It is important that the prng's have good randomness properties. Generated pseudo-random sequences have to be uniformly distributed, uncorrelated, have a large period of repetition, and can be rapidly calculated using limited computer memory. These properties are crucial for large-scale MC simulations. High-quality prng's are required to ensure correct results. Furthermore the quality of prng's can be tested by using large-scale MC simulation tasks.

In this work a set of several sequential and parallel prng's are tested when we apply a MC approach for solving a quantum-kinetic equation derived from ultra-fast transport in semiconductors. The aim is an optimization of the MC solver for the equation which accounts for quantum effects in the electron-phonon interaction. We consider so-called the Barker-Ferry (B-F) equation written for the case of zero electric field. The original formulation of this equation accounts for the action of the electric field during the process of collision.

The sensitivity of the MC solution for the electron energy distribution is investigated empirically, using prng's under consideration. The results obtained for the computational cost of the MC algorithm, the accuracy and the bias in the MC solution can be used to guide the treatment in the general case.

* Supported by ONR Grant N00014-96-1-1057 and by Center of Excellence BIS-21 grant ICAI-2000-70016, as well as by the NSF of Bulgaria through grant number 1-1201/02.
1 The quantum-kinetic equation

The Barker-Ferry equation [1] was developed as a physical model to describe the femtosecond relaxation process of initially excited electrons by a laser pulse [10]. For zero electrical field, the equation can be written in the following integral form:

\[
    f(k, t) = \int_{0}^{t} dt' \int_{0}^{t'} dt'' \int d^{3}k' S(k', k, t' - t'') f(k', t'') + \phi(k),
\]

where \( S(k', k, t' - t'') = \frac{2V}{(2\pi)^{3} \hbar^{2}} |g_{k' - k}|^{2} \exp(-\Gamma(k', k)(t' - t'')) \)

\[
    \times \{ (n_{q} + 1) \cos(\Omega(k', k)(t' - t'')) + n_{q} \cos(\Omega(k, k')(t' - t'')) \},
\]

and \( g_{k' - k} = -i \left[ \frac{2\pi e^{2} \hbar \omega_{q}}{V} \left( \frac{1}{\epsilon_{\infty}} - \frac{1}{\epsilon_{s}} \right) \frac{1}{|k' - k|^{2}} \right]^{\frac{1}{2}} \)

applies to the Fröhlich interaction, and \((\epsilon_{\infty})\) and \((\epsilon_{s})\) are the optical and static dielectric constants. The damping factor \( \Gamma(k', k) = \Gamma(k') + \Gamma(k) \) is related to the finite carrier lifetime for the scattering process:

\[
    \Gamma(k) = \int d^{3}k' V \frac{2\pi^{2} \hbar}{2V} \sum_{\pm} ||g_{k' - k}||^{2} \delta(\epsilon(k') - \epsilon(k) \pm \hbar \omega_{q})(n_{q} + \frac{1}{2} \pm \frac{1}{2}),
\]

In spherical coordinates \((k, \theta, \varphi)\), with the \( k'_{z} \) axis oriented along \( k \) and zero lattice temperature \((n_{q} = 0)\), the equation (1) can be rewritten as a one-dimensional integral in \( k \) [5]:

\[
    f(k, t) = \int_{0}^{t} dt'' \int_{0}^{t''} d\phi K(k, k') \times
\]

\[
    \int [K_{1}(k, k', t, t'') f(k', t'') + K_{2}(k, k', t, t'') f(k, t'')] + \phi(k),
\]

\[
    K(k, k') = c_{1} \frac{k'}{k} \ln \left( \frac{k + k'}{|k - k'|} \right),
\]

\[
    K_{1}(k, k', t, t'') = \frac{1}{\Omega_{k, k'}^{2} + \Omega_{k'}^{2}} \Gamma_{k, k'} + \frac{1}{\Omega_{k, k'}^{2} + \Omega_{k'}^{2}} \Gamma_{k'} \times
\]

\[
    \exp(-\Gamma_{k', k}(t - t'')) [\Omega_{k', k} \sin(\Omega_{k', k}(t - t'')) - \Omega_{k', k} \cos(\Omega_{k', k}(t - t''))]
\]

and

\[
    c_{1} = e^{2} \omega \left( \frac{1}{\epsilon_{\infty}} - \frac{1}{\epsilon_{s}} \right) / (\pi \hbar).
\]

The functions \( \Gamma(k', k) \) and \( \Omega(k', k) \) depend only on the radial variables \( k \) and \( k' \) and are denoted by \( \Gamma_{k', k} \) and \( \Omega_{k', k} \), respectively where
\[ \Gamma_k = \begin{cases} \epsilon_2 \ln \left( \frac{k + \sqrt{k^2 - \omega_1^2}}{\sqrt{\omega_1}} \right) / k, & \text{if } k^2 \geq \omega_1 \\ 0, & \text{if } k^2 < \omega_1, \end{cases} \]

with \( \omega_1 = 2m \omega_0 / \hbar \), \( \epsilon_2 = (m \omega_0 \omega_1^2 / \hbar^2) |1/\epsilon_\infty - 1/\epsilon_\infty| \).

The Neumann series corresponding to equation (3) converges [5] and a MC approach can be applied to evaluate the electron energy distribution. We note that this approach can be generalized for finite temperatures in a straightforward way.

2 Monte Carlo approach

Define a terminated Markov chain \((\kappa_0, \tau_0) \rightarrow \cdots \rightarrow (\kappa_j, \tau_j) \rightarrow \cdots \rightarrow (\kappa_{l_e}, \tau_{l_e})\), such that every point \((\kappa_j, \tau_j) \in (0, Q) \times (0, \tau_{j-1}), \ j = 1, 2, \ldots, l_e \) (\(\epsilon\) is the truncation parameter) is sampled using an arbitrary transition density function \(r(k', t', t)\) which is tolerant\(^1\) to both kernels in equation (3).

The biased Monte Carlo estimator for the solution of equation (3) at the fixed point \(k = \kappa_0\) at the time \(\tau = \tau_0\) using backward time evolution of the numerical trajectories has the following form:

\[ \xi_{l_e} [\kappa_0, \tau_0] = \phi(\kappa_0) + \sum_{j=1}^{l_e} W^\alpha_j \phi(\kappa_j), \]  

(4)

\[ W^\alpha_j = W^\alpha_{j-1} \frac{K(\kappa_{j-1}, \kappa_j)K_0(\kappa_{j-1}, \kappa_j, \tau_{j-1}, \tau_j)}{p_0 r(\kappa_{j-1}, \kappa_j, \tau_{j-1}, \tau_j)}, \quad W^\alpha_1 = 1, \quad \alpha = 1, 2, \ j = 0, \ldots, l_e. \]

The probabilities \(p_\alpha \ (\alpha = 1, 2)\) are related to the choice of one of the kernels. Now we can define a Monte Carlo method

\[ \frac{1}{N} \sum_{i=1}^{N} (\xi_{l_e} [\kappa_0, \tau_0])_i \rightarrow f(\kappa_0, \tau_0), \]  

(5)

where \(\xi_{l_e} [\kappa_0, \tau_0]_1, \xi_{l_e} [\kappa_0, \tau_0]_2, \ldots, \xi_{l_e} [\kappa_0, \tau_0]_N\) are independent values of the estimator (4) and \(\rightarrow\) means stochastic convergence as \(N \rightarrow \infty\). The relation (5) still does not determine the computation algorithm: we must specify the modeling function (sampling rule) \(\xi_{l_e} [\kappa_0, \tau_0] = g(\beta_1, \ldots, \beta_n)\), where \(\beta_1, \ldots, \beta_n\) are uniformly distributed random numbers in the interval \((0, 1)\). Now both relations (5) and the sampling rule define a Monte Carlo algorithm for (4).

Thus we can say [12] the constructive dimension (c.d.) of the algorithm is \(n\), i.e. \(c.d. = n\). Clearly, the variance of the MC estimator (4) does not depend on the c.d. Nevertheless, the c.d. has suggested a classification of sampling rules and an ordering of tests for pseudo-random numbers.

The transition density function in the Markov chain can be chosen in the following way \(r_\alpha (k, t, k', t') = r(k, k') r(t, t' / k, k')\), \(\alpha = 1, 2\), where

\[ r(k, k') = \frac{1}{k^k} \ln \left( \frac{k + k'}{k - k'} \right) \quad \text{and} \quad r(t, t' / k, k') = \frac{\Gamma_{k, k'} \exp(-\Gamma_{k, k'}(t - t'))}{1 - \exp(-\Gamma_{k, k'} t)}. \]

\(^1\) \(r(x)\) is tolerant of \(g(x)\) if \(r(x) > 0\) when \(g(x) \neq 0\) and \(r(x) \geq 0\) when \(g(x) = 0\).
IV

The normalized density function \( r(k, k') \) can be expressed as an infinite weighted sum of other density functions by expanding \((k'/k)\ln((k + k')/(|k - k'|))\), i.e.

\[
    r(k, k') = \sum_{i=0}^{\infty} C_i r_i(k, k'), \quad C_i \geq 0, \quad \sum_{i=0}^{\infty} C_i = 1,
\]

\[
    r_i(k, k') = \begin{cases} 
        \frac{(2i + 3)(k_i^{2i+2})}{(2i - 1)(Q_{i+1}^{2i-1})}, & \text{when } 0 \leq k' < k \\
        \frac{2}{(Q_{i+1}^{2i-1})(1 - (k_i^{2i-1})^2)}, & \text{when } k < k' \leq Q,
    \end{cases}
\]

\[
    C_i = \frac{(2i+1)(2i+3)}{2i-1} \left( \frac{1}{2k + (Q + k) \ln(\frac{Q_{i+1}^{2i-1}}{k} + 1)} \right)^2,
\]

when \( k < k' \leq Q \).

The decomposition MC approach can be applied to sample \( k' \):

1. Generate \( \beta_1, \beta_2, \beta_3 \) uniform on \([0, 1] \).
2. Define \( C_i \) by \( \beta_1 \) using decomposition MC techniques.
3. Sample \( k' \) with the \( i \)-th density function \( r_i(k, k') \), namely, \( k' = k(\beta_3)^\frac{1}{i+1} \).
   If \( \beta_2 Q < k \). Otherwise, \( k' = k/[1 - \beta_3(1 - (k/Q)^{2i-1})] \).%

Using the normalized conditional probability density function \( r(t, t' \mid k, k') \) we can sample \( t'' \) by \( t'' = \log(\beta_4() \exp(I_k k') - 1) + 1)/(I_k k') \) where \( \beta_4 \in (0, 1) \). Finally, we generate \( \beta_5 \in (0, 1) \) and choose one of the kernels \( K_{\alpha}(k, k', t, t') \), \( \alpha = 1, 2 \) using probabilities \( p_\alpha = [K_{\alpha}(k, k', t, t')] / ([K_1(k, k', t, t')] + [K_2(k, k', t, t')]). Summarizing, we have used 5 uniform random numbers \( \beta_1, \ldots, \beta_5 \) in order to construct the MC estimator (4) for one transition \((k, t) \rightarrow (k', t')\) in the Markov chain.

The computational complexity of the obtained iterative MC algorithm can be measured by the quantity \( F = N \times t_{\text{ran}} \times E(t) \). We note that the number of the random walks, \( N \), and the average number of transitions in the Markov chain, \( E(t) \), are connected with stochastic and systematic errors [5]. However the mean time for modeling one transition, \( t_{\text{ran}} \), \( (n_0 = 5) \) depends on the complexity of the transition density functions and the choice of the random number generator. It is strongly recommended that all simulations be done with two or more different generators, and the results compared to check whether the prng is introducing a bias.

The c.d. of this algorithm can be defined as the average number of uniformly distributed random numbers necessary for carrying out one trial, i.e. c.d. = \( n_0 E(t) \). Thus we can use parallel prng’s that produce \( n_0 = 5 \) independent and non-overlapping random sequences in order to compute every transition in the Markov chain as well as sampling 5 consecutive pseudo-random numbers from a sequential generator.

3 Numerical results and discussions

The simulation results are obtained for GaAs with material parameters taken from [10]. The initial condition is a Gaussian function of the energy. The solution \( f(k, t) \) is estimated by the MC estimator in 60 points of the simulation domain between 0 and \( Q = 66 \times 10^7/m \).

The iterative MC algorithm is realized using the following sequential prng’s
1. CLCG-PL, Combined linear congruential generator with parameters recommended by P. L’Ecuyer [7];
2. EICG, Explicit inverse congruential generator [2];
3. ICG, Inverse congruential generator [3];
4. LCG-F, Linear congruential generator with parameters recommended by Fishman [4];
5. LCG-PM, Linear congruential generator with parameters recommended by Park and Miller [11];
6. MT-MN, Mersenne Twister generator by Matsumoto and Nishimura [9];
as well as the following parallel prng’s

1. SNWS, Shuffled nested Weyl sequences [6] with a multiplier $M = 1234567$. To produce 5 random sequences we use the following seeds:
   $\gamma = \{2^{1/2}, 3^{1/2}, 5^{1/2}, 7^{1/2}\}.$
2. SPRNG, the Scalable Parallel Random Number Generator Library [8,14].

The MC algorithm were implemented in the C language. Numerical tests were performed on a Sun Ultra Enterprise 450 with 4 Ultra-SPARC, 400 MHz CPUs running Solaris.

![Fig. 1. Comparison of the electron energy distribution $kf(k,t)$ versus $k^2$ obtained by using of SNWS and SPRNG.](image)

In all our tests $\varepsilon = 0.0001$. Such a choice of the truncation parameter allow us to ignore the systematic error [5] and to investigate whether any generator under consideration is introducing a bias when different stochastic errors are fixed. The quantity presented on the y-axes in all figures below, $kf(k,t)$, is proportional to the electron energy distribution function multiplied by the density of states. The quantity $k^2$ given on the x-axes in units of $10^{14}/m^2$ is proportional to the electron energy.

Figure 1 compares the solutions for evolution times 100 femtoseconds (fs), 150 fs and 200 fs obtained by using the SNWS and, SPRNG parallel prng’s. The number of realizations of the MC estimator (4) are 1 million (mhn), 5 mhn and 10 mhn, respectively. We see that the solutions coincide. Table 1 shows the
Table 1. Comparison of the accuracy of the solution obtained with the SNWS and SPRNG generators for the 3 points with the biggest variance. The evolution time is 100fs in the case \( N = 1 \) min, 150fs in the case \( N = 5 \) min and 200fs in the case \( N = 10 \) min, respectively.

| \( N \) | \( k \) | \( k_f \) | \( \mu \) | \( k_f \) | \( \mu \) | \( |f_1 - f_2|\) |
|---|---|---|---|---|---|---|
| 48.00 | 9.8192 ±0.0038 | 9.7896 ±0.0038 | 6.2 × 10^{-5} | 48.75 | 10.4557 ±0.0039 | 10.4595 ±0.0039 | 7.8 × 10^{-5} |
| \( \min \) | 49.50 | 10.7340 ±0.0039 | 10.7024 ±0.0039 | 6.4 × 10^{-4} | 48.00 | 14.641 ±0.0041 | 14.7029 ±0.0042 | 1.2 × 10^{-5} |
| \( 5 \) | 48.75 | 15.6877 ±0.0043 | 15.7370 ±0.0043 | 1.0 × 10^{-3} | \( \min \) | 49.50 | 15.4964 ±0.0042 | 15.5394 ±0.0042 | 8.7 × 10^{-4} |
| \( 10 \) | 48.75 | 18.2430 ±0.0066 | 18.0536 ±0.0067 | 3.9 × 10^{-3} | \( \min \) | 49.50 | 17.7436 ±0.0064 | 17.6170 ±0.0064 | 2.6 × 10^{-3} |

mean square error, \( \mu \), and the absolute error for the 3 values of the momentum \( k \) with the biggest variance using the SNWS and SPRNG generators. In this “the worst” case of the variance compared with the variance at the other points, we have \( \mu = O(10^{-3}) \) and absolute errors are in agreement with the mean square error. Let us note that the exact solution of the B-F equation is unknown. Given the excellent agreement and similar variances, we can take any MC solution from Fig. 1 as a “correct” solution. Figures 2 – 5 compare “correct” solutions (using

![Fig. 2. Comparison of the electron energy distribution obtained by using of LCG-F and LCG-PM generators with the "correct" solution on the left and the right pictures, respectively.](image_url)

the results with the SNWS generator) for the evolution times 100fs, 150fs and 200fs with the quantum solutions obtained using all the sequential and parallel
prng’s when the mean square error is $O(10^{-2})$. The number of realizations of the MC estimator are $N = 30000$, $150000$ and $750000$. Results obtained when $k^2 < 2000$ for $k f(k, t)$ using the sequential generators when compared with the “correct” solution show systematic differences. The best case occurs when using the CLG-PM, minimal standard, generator. However, for all times it exhibits small consistent differences.

Systematic differences in the MC solution with increasing evolution time appear when LCG-F, ICG, EICG and MT-MN are used. Random “noise” in the MC solution is observed when the CLCG-PL generator is used, which, however, is unbiased. When $k^2 > 2000$ the results using any prng’s disagree in the first peak of the distribution. This can be explained because the product $k f(k, t)$ for bigger values of $k$ is sensitive to even small errors in the MC solution.

Table 2 shows the computational complexity (CPU time for all 60 points) of the algorithm using all the prng’s. We see that the computational cost is the least when the MT-MN generator is used. The CPU time of the algorithm using the SNWS and SPRNG parallel prng’s is closely to the CLCG-PL sequential generator and they are faster than the EICG and LCG-F sequential generators. The ICG generator delays twice the MC algorithm and therefore it should not be included for solving this problem. Also, the quantity $E(L_e)$ very slowly increases with increasing evolution time.

4 Summary

Statistically, the solution of the B-F equation would be expected to be noisier at $O(10^{-2})$ than at $O(10^{-3})$ mean square error. It is gratifying that the two parallel prng’s used gave the same answer at $O(10^{-3})$ precision. However, even at $O(10^{-2})$ mean square error, if the solution was unbiased, we would expect random fluctuations about the more precise solution. This was only observed.
Fig. 4. Comparison of the electron energy distribution obtained by using of EICG and CLCG-PL generators with the "correct" solution on the left and the right pictures, respectively.

Fig. 5. Comparison of the electron energy distribution obtained by using of MT-MN and SPRNG generators with the "correct" solution on the left and the right pictures, respectively.
with the CLCG-PL prng. All the other sequential generators exhibited systematic rather than random differences. Therefore we conclude that parallel prng’s

<table>
<thead>
<tr>
<th>generator</th>
<th>$150f_s$, $N = 150000$</th>
<th>$200f_s$, $N = 750000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU time</td>
<td>$E(L)$</td>
<td>CPU time</td>
</tr>
<tr>
<td>MT-MN</td>
<td>13m11.80s 15.5084</td>
<td>1h20m51.78s 15.8905</td>
</tr>
<tr>
<td>LCG-PM</td>
<td>20m18.46s 15.5265</td>
<td>1h44m49.75s 15.9065</td>
</tr>
<tr>
<td>CLCG-PL</td>
<td>23m08.77s 15.5162</td>
<td>2h0m45.19s 15.9062</td>
</tr>
<tr>
<td>SNWS</td>
<td>23m17.10s 15.5300</td>
<td>1h50m54.47s 15.9155</td>
</tr>
<tr>
<td>SPRNG</td>
<td>24m50.11s 15.5085</td>
<td>2h6m20.08s 15.8982</td>
</tr>
<tr>
<td>EICG</td>
<td>30m51.42s 15.5265</td>
<td>2h38m47.89s 15.9095</td>
</tr>
<tr>
<td>LCG-F</td>
<td>30m56.13s 15.4899</td>
<td>2h38m58.54s 15.8891</td>
</tr>
<tr>
<td>ICG</td>
<td>47m31.30s 15.5153</td>
<td>3h2m44.73s 15.9023</td>
</tr>
</tbody>
</table>

are preferable to solve this problem as the evolution time increases. In this case, the CPU time of the algorithm become crucial. Thus, to predict the solution we need parallel realizations of the algorithm and/or we have to estimate the solution with coarser stochastic error. In order to obtain a high parallel efficiency in the case of the parallel realization of the algorithm, the random sequences have to be produced with similar CPU times.

References

13. Pseudo-Random Number Generator, (http://statistik.wu-wien.ac.at/prng/).