Computer Physics Communications 185 (2014) 28–39

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

# Efficient linear-scaling quantum transport calculations on graphics processing units and applications on electron transport in graphene

Zheyong Fan *, Andreas Uppstu, Topi Siro, Ari Harju

*COMP Centre of Excellence and Helsinki Institute of Physics, Department of Applied Physics, Aalto University, Helsinki, Finland*

ARTICLE INFO

*Article history:*
Received 28 June 2013
Received in revised form
12 August 2013
Accepted 14 August 2013
Available online 22 August 2013

*Keywords:*
Quantum transport
Kubo–Greenwood formula
Chebyshev polynomial expansion
Graphics processing unit
CUDA
Graphene

ABSTRACT

We implement, optimize, and validate the linear-scaling Kubo–Greenwood quantum transport simulation on graphics processing units by examining resonant scattering in graphene. We consider two practical representations of the Kubo–Greenwood formula: a Green–Kubo formula based on the velocity auto-correlation and an Einstein formula based on the mean square displacement. The code is fully implemented on graphics processing units with a speedup factor of up to 16 (using double-precision) relative to our CPU implementation. We compare the kernel polynomial method and the Fourier transform method for the approximation of the Dirac delta function and conclude that the former is more efficient. In the ballistic regime, the Einstein formula can produce the correct quantized conductance of one-dimensional graphene nanoribbons except for an overshoot near the band edges. In the diffusive regime, the Green–Kubo and the Einstein formalisms are demonstrated to be equivalent. A comparison of the length-dependence of the conductance in the localization regime obtained by the Einstein formula with that obtained by the non-equilibrium Green's function method reveals the challenges in defining the length in the Kubo–Greenwood formalism at the strongly localized regime.



## 1. Introduction

Quantum simulations are very important tools to study transport phenomena in the nanoscale, both for electrons and phonons. There are mainly two numerical approaches for quantum transport simulations: one is the widely used non-equilibrium Green's function (NEGF) method [1] and the other is the Kubo–Greenwood method [2,3]. Both methods have been widely used to study the electronic transport properties of graphene, a two-dimensional sheet of carbon atoms [4,5]. Despite this, the field of electronic transport in graphene has remained very actively debated.

So far, the NEGF method has been mostly used to simulate relatively small systems, due to the cubic scaling of the computational effort associated with matrix inversion. Although an efficient iterative method [6] enables the simulation of very long systems, this method is still restricted to studying quasi-one-dimensional (1D) systems, such as carbon nanotubes and graphene nanoribbons (GNRs). However, many realistic systems exhibit structural length scales up to several micrometers. Examples of these include large-scale two-dimensional (2D) graphene antidot lattices [7] and polycrystalline graphene produced by chemical vapor deposition [8]. Furthermore, in many applications, graphene-based materials have to be fabricated or assembled into three-dimensional (3D) architectures, which require even larger simulation sizes. The application of the NEGF method to these realistically sized 2D and 3D systems is still not feasible.

In contrast, for the Kubo–Greenwood method, a real-space linear-scaling method has been developed [9–12] and used to study transport properties of both quasi-1D systems [13–15] and 2D graphene sheets [16–23]. Moreover, this method has been generalized to studying thermal conductivity [24]. Besides the real-space Kubo method [9–12], which expresses the conductivity as a time-derivative of the mean square displacement, another seemingly different approach [25], which expresses the conductivity as a time-integration of the velocity auto-correlation function, has also been used to study the electronic transport properties of large-scale single-layer [25,26] and multi-layer [27] graphene sheets, and disordered graphene antidot lattices [28].

Although both of the above methods are based on the Kubo–Greenwood formula, no connection has been made between them. One of our purposes is to identify the time-derivative approach and the time-integration approach as an Einstein relation and the corresponding Green–Kubo relation and demonstrate their equivalence numerically. Furthermore, a thorough validation of Kubo–Greenwood formula-based quantum transport methods for all the transport regimes is also absent. We thus aim to perform a comprehensive evaluation of the applicability of the linear-scaling Kubo–Greenwood quantum transport simulation method for all three transport regimes: the ballistic, diffusive, and localized regimes.

* Corresponding author. Tel.: +358 0465378411.
*E-mail addresses:* zheyong.fan@aalto.fi, brucenju@gmail.com (Z. Fan).

0010-4655/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cpc.2013.08.009

To achieve the above, we find that an efficient implementation is very desirable. Despite the linear-scaling nature of these numerical methods, they are still computationally demanding in most cases. Nowadays, the use of graphics processing units (GPUs) has played a more and more important role in computational physics; finding the solutions to many problems in computational physics has become impressively accelerated by using a single or multiple GPUs [29]. In this work, we consider the implementation of the Kubo–Greenwood quantum transport simulation on the GPU, with a unified treatment of the various involved theoretical formalisms and numerical techniques. We will evaluate the performance and correctness of our implementation, as well as the applicability of the method itself.

This paper is organized as follows. In Section 2, we present the theoretical background of the Kubo–Greenwood formula and the Green–Kubo and Einstein relations which are both derived. In Section 3, we give a detailed discussion of the involved numerical techniques and their GPU implementations. After making a performance evaluation in Section 4, we thoroughly evaluate the computational method in different transport regimes in Section 5. Section 6 concludes.

## 2. Theoretical formalism

The Kubo–Greenwood formula [3] for DC conductivity $\sigma^{\mathrm{KG}}_{\mu\nu}(E)$ as a function of the energy $E$ at zero temperature is

$$\sigma^{\mathrm{KG}}_{\mu\nu}(E) = \frac{2\pi\hbar e^2}{\Omega}\mathrm{Tr}\left[V_\mu\delta(E-H)V_\nu\delta(E-H)\right] \tag{1}$$

where $\hbar$ is the reduced Planck constant, $e$ is the electron charge, $\Omega$ is the system volume, $V_\mu$ is the velocity operator in the $\mu$-direction, $H$ is the Hamiltonian of the system, and Tr denotes the trace. The factor of 2 results from spin degeneracy. For simplicity, we only consider transport along one direction. Then, the above formula can be simplified to be

$$\sigma^{\mathrm{KG}}(E) = \frac{2\pi\hbar e^2}{\Omega}\mathrm{Tr}\left[V\delta(E-H)V\delta(E-H)\right]. \tag{2}$$

By Fourier transforming one of the $\delta$ functions in the above formula,

$$\delta(E-H) = \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} dt e^{i(E-H)t/\hbar}, \tag{3}$$

we have

$$\sigma(E) = \frac{e^2}{\Omega}\int_{-\infty}^{+\infty} dt\mathrm{Tr}\left[e^{iEt/\hbar}Ve^{-iHt/\hbar}V\delta(E-H)\right], \tag{4}$$

or equivalently,

$$\sigma(E) = \frac{e^2}{\Omega}\int_{-\infty}^{+\infty} dt\mathrm{Tr}\left[e^{iHt/\hbar}Ve^{-iHt/\hbar}V\delta(E-H)\right] \tag{5}$$

due to the remaining $\delta$ function. Through a change of variables, $t \to -t$, we get the following Green–Kubo formula [30,2], which expresses the running electrical conductivity (REC) as a time-integration of the velocity auto-correlation (VAC) $C_{vv}(E,t)$:

$$\sigma^{\mathrm{GK}}(E,t) = e^2\rho(E)\int_0^t C_{vv}(E,t)dt, \tag{6}$$

$$C_{vv}(E,t) = \frac{\mathrm{Tr}\left[\frac{2}{\Omega}\delta(E-H)\left(V(t)V+VV(t)\right)/2\right]}{\mathrm{Tr}\left[\frac{2}{\Omega}\delta(E-H)\right]}, \tag{7}$$

$$\rho(E) = \mathrm{Tr}\left[\frac{2}{\Omega}\delta(E-H)\right], \tag{8}$$

where $V(t) = U^\dagger(t)VU(t) = e^{iHt/\hbar}Ve^{-iHt/\hbar}$ is the velocity operator in the Heisenberg representation, and $\rho(E)$ the density of states (DOS). The Green–Kubo relation constitutes essentially the formalism used by Yuan et al. [25,27].

For a specific Green–Kubo formula, there is generally a corresponding Einstein formula. By integrating the Green–Kubo formula, we obtain the following Einstein formula, which expresses the REC as a time-derivative of the mean square displacement (MSD) $\Delta X^2(E,t)$:

$$\sigma^{\mathrm{E1}}(E,t) = e^2\rho(E)\frac{d}{2dt}\Delta X^2(E,t), \tag{9}$$

$$\Delta X^2(E,t) = \frac{\mathrm{Tr}\left[\frac{2}{\Omega}\delta(E-H)\left(X(t)-X\right)^2\right]}{\mathrm{Tr}\left[\frac{2}{\Omega}\delta(E-H)\right]}, \tag{10}$$

where $X(t) = U^\dagger(t)XU(t)$ is the position operator in the Heisenberg representation. An alternative definition, in which the derivative in the above equation is replaced by a division,

$$\sigma^{\mathrm{E2}}(E,t) = e^2\rho(E)\frac{\Delta X^2(E,t)}{2t}, \tag{11}$$

is frequently used, since it gives smoother curves for the REC than $\sigma^{\mathrm{E1}}(E,t)$ does. The above Einstein relation is exactly the real-space Kubo method [9–12].

We will demonstrate the equivalence of the Green–Kubo formalism and the Einstein formalism numerically. Specifically, we will show that $\sigma^{\mathrm{E1}}(E,t)$ and $\sigma^{\mathrm{GK}}(E,t)$ are equivalent, while $\sigma^{\mathrm{E2}}(E,t)$ deviates from the other two to some degree.

By going from the Kubo–Greenwood formalism to the Green–Kubo or the Einstein formalism, the conductivity becomes a function of not only the energy $E$, but also the correlation time $t$. Usually, one takes the following large time limit:

$$\sigma^{\mathrm{KG}}(E) = \lim_{t\to\infty}\sigma^{\mathrm{GK}}(E,t) = \lim_{t\to\infty}\sigma^{\mathrm{E1}}(E,t). \tag{12}$$

However, the convergence of this limit is only ensured for diffusive transport, in which case the VAC decays to zero and the MSD becomes proportional to $t$, resulting in a converged REC. For ballistic transport, the VAC oscillates around a fixed value and the MSD increases quadratically with increasing $t$, resulting in a divergent REC. In the localized regime, the VAC develops negative values and the slope of the MSD decreases, resulting in a decaying REC.

In this paper, we take graphene as our test system. We use $N_x$ to represent the number of dimer lines located along the zigzag edge and $N_y$ to represent the number of zigzag-shaped chains across the armchair edge. Thus, an $N_x \times N_y$ graphene sample has $N = N_xN_y$ carbon atoms, and the lengths in the zigzag and armchair directions are $L_x = \sqrt{3}N_xa/2$ and $L_y = 3N_ya/2$, respectively, where $a = 0.142$ nm is the carbon–carbon bond length used. For 2D graphene, periodic boundary conditions are applied in both directions; for quasi-1D armchair graphene nanoribbon (AGNR) and zigzag graphene nanoribbon (ZGNR), we use periodic boundary conditions along the transport (longitudinal) direction and non-periodic boundary conditions along the perpendicular direction.

We use a nearest-neighbor $p_z$ orbit tight-binding Hamiltonian for pristine systems:

$$H = \sum_{\langle mn\rangle}H_{mn}|m\rangle\langle n| = -\sum_{\langle mn\rangle}\gamma_0|m\rangle\langle n|, \tag{13}$$

where the hopping parameter $\gamma_0$ is chosen to be 2.7 eV. With this notation, the position and velocity operators can be expressed as

$$X = \sum_m X_m|m\rangle\langle m|, \tag{14}$$