

Accepted Manuscript

Improved recursive Green's function formalism for quasi one-dimensional systems with realistic defects

Fabian Teichert, Andreas Zienert, Jörg Schuster, Michael Schreiber

PII: S0021-9991(17)30034-7
DOI: <http://dx.doi.org/10.1016/j.jcp.2017.01.024>
Reference: YJCPH 7086

To appear in: *Journal of Computational Physics*

Received date: 1 December 2015
Revised date: 12 January 2017
Accepted date: 13 January 2017

Please cite this article in press as: F. Teichert et al., Improved recursive Green's function formalism for quasi one-dimensional systems with realistic defects, *J. Comput. Phys.* (2017), <http://dx.doi.org/10.1016/j.jcp.2017.01.024>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Graphical abstract

Improved recursive Green's function formalism for quasi one-dimensional systems with realistic defects

Journal of Computational Physics ••••, •••, •••

Fabian Teichert^{a,c,d,*}, Andreas Zienert^b, Jörg Schuster^{c,d}, Michael Schreiber^{a,d}

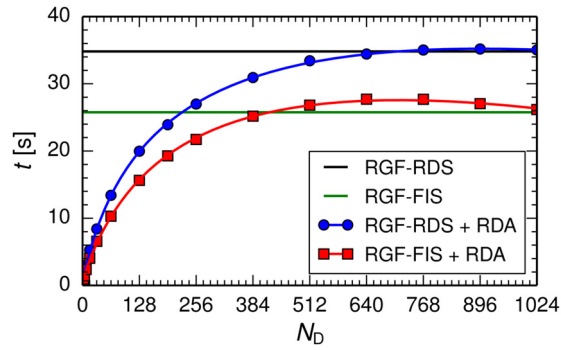
^a Institute of Physics, Faculty of Natural Sciences, Chemnitz University of Technology, 09126 Chemnitz, Germany

^b Center for Microtechnologies, Chemnitz University of Technology, 09126 Chemnitz, Germany

^c Fraunhofer Institute for Electronic Nano Systems (ENAS), 09126 Chemnitz, Germany

^d Dresden Center for Computational Materials Science (DCMS), TU Dresden, 01062 Dresden, Germany

With the improved recursive Green's function formalism + renormalization decimation algorithm (RGF+RDA), the conductance of mesoscopic systems with realistic defects can be computed much faster than with the common RGF. For a constant system length, the calculation time t scales logarithmically with the number of defects N_D . This is true for the recursive decimation scheme (RDS) as well as for the forward iteration scheme (RDS).



Highlights

- An improved version of the recursive Green's function formalism is derived.
- Transport calculations for carbon nanotubes with realistic defects are faster.
- The calculation time is reduced by up to 80% for systems with 1% defects.

Improved recursive Green's function formalism for quasi one-dimensional systems with realistic defects

Fabian Teichert^{a,c,d,*}, Andreas Zienert^b, Jörg Schuster^{c,d}, Michael Schreiber^{a,d}

^a*Institute of Physics, Faculty of Natural Sciences, Chemnitz University of Technology, 09126 Chemnitz, Germany*

^b*Center for Microtechnologies, Chemnitz University of Technology, 09126 Chemnitz, Germany*

^c*Fraunhofer Institute for Electronic Nano Systems (ENAS), 09126 Chemnitz, Germany*

^d*Dresden Center for Computational Materials Science (DCMS), TU Dresden, 01062 Dresden, Germany*

Abstract

We derive an improved version of the recursive Green's function formalism (RGF), which is a standard tool in the quantum transport theory. We consider the case of disordered quasi one-dimensional materials where the disorder is applied in form of randomly distributed realistic defects, leading to partly periodic Hamiltonian matrices. The algorithm accelerates the common RGF in the recursive decimation scheme, using the iteration steps of the renormalization decimation algorithm. This leads to a smaller effective system, which is treated using the common forward iteration scheme. The computational complexity scales linearly with the number of defects, instead of linearly with the total system length for the conventional approach. We show that the scaling of the calculation time of the Green's function depends on the defect density of a random test system. Furthermore, we discuss the calculation time and the memory requirement of the whole transport formalism applied to defective carbon nanotubes.

Keywords: recursive Green's function formalism (RGF), renormalization decimation algorithm (RDA), electronic transport, carbon nanotube (CNT), defect

PACS: 72.10.Bg, 73.23.-b, 73.63.Fg

2010 MSC: 82D77, 82D80, 82B44, 82B80, 81V99

1. Introduction

In the last decades, simulation techniques became more and more important for a huge diversity of topics in theoretical physics. Mechanical, optical, and electronic properties of new materials have been studied, often driven by developments in semiconductor technologies. Simultaneously, the miniaturization process shifts the dimensions of microelectronic devices into the mesoscopic range, which cannot be treated classically. Simulation methods have to deal with a large complexity concerning quantum mechanical effects, different materials and their interplay, and multiple length scales.

Electronic transport properties of arbitrary devices, which are connected to electrodes, can be described by electron scattering and a resulting transmission probability, using the Landauer-Büttiker formalism [1]. In combination with quantum transport theory [2] and an underlying electron structure theory like the tight binding method, density functional theory [3], or hybrid methods, this allows one to calculate currents and conductances through devices precisely via the Green's function of the system using its given Hamiltonian. This approach needs a computationally expensive inversion, whose complexity scales with the third power of the number of involved atoms. However, the recursive Green's function formalism (RGF) [4–7] provides a very efficient, linearly scaling algorithm by dividing the system into layers. It can be applied to arbitrary systems which have a sparse Hamiltonian matrix, as is the case with all short-range interactions in real

*Corresponding author

Email address: fabian.teichert@physik.tu-chemnitz.de (Fabian Teichert)

space. Although this is a common method nowadays, further technical improvements of the theory are necessary, because the numerical complexity still rapidly reaches the limits of today's computer resources for large and complex systems or when different physical effects (i.e. levels of theory) are considered.

It is still difficult to apply the RGF in an optimal way to arbitrary systems. For difficult geometries the question how to arrange the RGF layers or how to get the optimal block-tridiagonal matrix form has to be considered. Different solutions have been found, e.g. a pivoting-like bandwidth minimization method in combination with a block-tridiagonalization algorithm [8], the reverse Cuthill-McKee algorithm for the block-tridiagonalization of connected graphs [9, 10], the Knitting algorithm, which uses the Dyson equation and the resulting decimation scheme to build up the system site by site [11], or accelerations by using singular value decompositions [12]. Besides the RGF, also other methods are often utilized. The inversion can be done directly, using fast LU-decomposition algorithms with nested dissection [13]. The simple basic idea, divide and conquer, leads to optimal scaling in the limit of large systems. But "large" in this context is until now still larger than systems which are treatable within acceptable time. The existing nested dissection method can be further enhanced by using width-one separators instead of width-two separators, which reduces calculation time drastically [14]. LU-decomposition methods also allow to calculate selected elements of an inverse matrix very fast. This is especially interesting for transport problems, because only a few elements of the Green's function are necessary. Based on this, fast recurrence formulas can be found similar to the RGF but based on the LU-decomposition [15, 16]. Also combinations with the RGF like the over-bridging boundary-matching method in combination with the shifted conjugate-orthogonal conjugate-gradient method [17–21] are promising.

In the present paper we describe new methodical improvements to speed up quantum transport calculations of certain quasi one-dimensional systems. In the sections 2 and 3 we shortly explain the most important formulas of (equilibrium) quantum transport theory and common solutions for quasi one-dimensional systems using the RGF. Based on this, in section 4, we discuss an advanced solution for special cases and develop a further improvement of the RGF. We consider the case of disordered quasi one-dimensional systems where the disorder is caused by randomly positioned realistic defects in an otherwise ideal system. In contrast to Anderson-like disorder [22–24], which randomizes the Hamiltonian matrix of the whole system, this system consists of repeated ideal parts which are interrupted by defective parts. That means that the Hamiltonian matrix has many blocks which are equal and only a few ones which are different. For such partly periodic cases the renormalization decimation algorithm (RDA) [25, 26], which calculates the Green's function of the electrodes within an iteration process, allows us to treat the ideal parts much faster than the pure RGF can do. We discuss the calculation time and memory requirement for a simple random test system and especially for carbon nanotubes (CNTs) with divacancy defects.

CNTs are prominent examples for future electronic devices, because of their excellent electronic properties. But many things can negatively affect them. CNTs have to be integrated into conventional devices and thus are contacted to metals. The resulting interplay between these surrounding metals and the CNT leads to unintended reductions of the current and the conductance [27–29]. Also defects have a huge impact. Several publications concerning the influence of vacancies [30–34], substitutional atoms [35], functionalizations [36–38], and Anderson disorder [39] showed that the ballistic transport is driven into the strong localization regime [40]. But mostly this has been done exemplary, as extensive calculations of CNT ensembles are necessary. We contributed to this task with a comprehensive study of the electronic properties of CNTs under the influence of randomly distributed monovacancies and divacancies. We varied the CNT length, the CNT diameter, the defect type, the defect probability, and the temperature, and we discussed the resulting simple analytical dependencies [41].

2. Quantum transport theory

In this section we give a brief introduction into (equilibrium) quantum transport theory and the resulting numerical challenges. Quantum transport theory describes transport properties like transmission and conductance (but also electronic structure properties like density of states and electron density) of open systems [2]. "Open" means that the system of interest (conductor) is connected to comparably large leads (that can be considered as reservoirs where electrons are always at their respective equilibrium), from/to

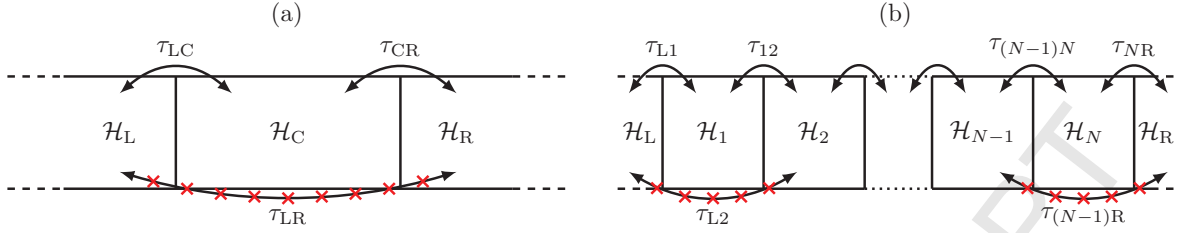


Figure 1: (Color online.) Scheme of the device system used in the transport formalism [41]. The Hamiltonian matrices of the subsystems are denoted by \mathcal{H} and the coupling matrices between the subsystems by τ . Crossed-out couplings are not taken into account in our approach. (a) The two electrodes (L and R) are ideal, semi-infinite parts. The central region (C) includes the defective part. (b) The central region is divided into N subsystems to which the RGF is applied.

which electrons are allowed to enter/leave. For this, the system is described as shown in figure 1a. A finite central region C, which contains the total non-periodic part of the quasi one-dimensional system, is connected to two half-infinite and periodic electrodes (left L and right R), which act as reservoirs providing electrons or holes [2]. We call this a device configuration. The corresponding Schrödinger equation within an orthonormal basis¹ reads

$$\begin{pmatrix} \mathcal{H}_L & \tau_{LC} & 0 \\ \tau_{CL} & \mathcal{H}_C & \tau_{CR} \\ 0 & \tau_{RC} & \mathcal{H}_R \end{pmatrix} \Psi = E \Psi \quad . \quad (1)$$

$\mathcal{H}_{L/C/R}$ denote the Hamiltonian matrices of the parts L, C, and R. τ denote the coupling matrices between two of these parts. If region C is long enough, the direct interaction of L and R can be neglected, $\tau_{LR} = 0$. Note that this is a matrix eigenvalue equation of infinite dimension, which in this form is numerically not treatable. However, it can be reduced to an effective matrix eigenvalue equation of finite dimension by treating the electrodes separately and including their influence as a self energy correction.

We define the advanced Green's function matrix of the central region

$$\mathcal{G}_C(E) = \lim_{\eta \rightarrow 0^+} [(E + i\eta)\mathcal{I} - \mathcal{H}_C - \Sigma_L - \Sigma_R]^{-1} \quad . \quad (2)$$

$\Sigma_L = \tau_{CL}\mathcal{G}_L\tau_{LC}$ and $\Sigma_R = \tau_{CR}\mathcal{G}_R\tau_{RC}$ are self energy matrices, which lead to an energy-dependent shift of the electronic states due to the electrode coupling. $\mathcal{G}_{L/R}$ are the advanced surface Green's functions of the electrodes, which can be calculated with the RDA (see section 3.3). \mathcal{I} is the identity matrix of appropriate dimension. The transmission spectrum $\mathcal{T}(E)$ of the device configuration can be calculated with

$$\mathcal{T}(E) = \text{Tr}(\Gamma_R \mathcal{G}_C \Gamma_L \mathcal{G}_C^\dagger) \quad . \quad (3)$$

Therein, $\Gamma_{L/R} = i(\Sigma_{L/R} - \Sigma_{L/R}^\dagger)$ are broadening matrices, which lead to an energy-dependent broadening of the electronic states due to the electrode coupling. Finally the conductance is given by the Landauer-Büttiker formalism [1]

$$G = -G_0 \int_{-\infty}^{\infty} \mathcal{T}(E) f'(E) dE \quad \text{with} \quad G_0 = \frac{2e^2}{h} \quad \text{and} \quad f(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)} \quad . \quad (4)$$

E_F is the Fermi energy.

¹The Schrödinger equation within a non-orthogonal basis can be obtained by substituting E by ES . S is the overlap matrix, which has the same device block structure as the Hamiltonian matrix. This substitution can be applied to the subsequent equations as well.

In summary, calculating the transmission spectrum involves two main tasks: the electrode calculations (self energies), in which infinite but periodic problems must be solved, and the inversion problem of the central region (which can be very large). In the range of mesoscopic systems with hundreds of thousands of atoms, the latter is a time and memory consuming process (the calculation complexity scales as $\mathcal{O}[(\dim \mathcal{H}_C)^3]$). However, we can take advantage of the block-tridiagonal shape of \mathcal{H}_C . If region C is much longer than the interaction distance, it can be divided into N subsystems with Hamiltonian matrices \mathcal{H}_i and coupling matrices τ_{ij} , as shown in figure 1b. As non-neighboring cells are not interacting, the corresponding coupling matrices are zero and the transmission spectrum can be calculated with the simplified formula

$$\mathcal{T}(E) = \text{Tr} \left(\Gamma'_R \mathcal{G}_{N1} \Gamma'_L \mathcal{G}_{N1}^\dagger \right) \quad . \quad (5)$$

\mathcal{G}_{N1} is the lower left matrix block of \mathcal{G}_C and Γ'_L (Γ'_R) is the upper left (lower right) matrix block of Γ_L (Γ_R).

3. Recursive Green's function formalisms

70 The RGF [6, 7] is a method for calculating \mathcal{G}_{N1} . There exist different implementations of the RGF, which are widely used for mesoscopic systems with hundreds of thousands of atoms. In the following, we discuss the forward iteration scheme (FIS), the recursive decimation scheme (RDS), and the RDA [25, 26]. The latter is a convergent iterative algorithm for calculating the surface Green's functions of periodic systems.

3.1. Forward iteration scheme

The FIS is based on the inversion of a 2×2 block matrix where only the lower left block of the inverse matrix is of interest for transport calculations.

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \Rightarrow B_{21} = -\tilde{A}_{22}^{-1} A_{21} A_{11}^{-1} \quad \text{with} \quad \tilde{A}_{22} = A_{22} - A_{21} A_{11}^{-1} A_{12} \quad (6)$$

This can be repeated by dividing A_{22} into 2×2 blocks and so on, leading to a matrix version of the Gauß-Jordan elimination. The resulting recursion formula can easily be transferred to the calculation of \mathcal{G}_{N1} . Additionally, noticing the self energy corrections in the first and the last subsystem, we obtain

$$\mathcal{G}_1 = \lim_{\eta \rightarrow 0^+} [(E + i\eta)\mathcal{I} - \mathcal{H}_1 - \Sigma_L]^{-1} \quad , \quad (7a)$$

$$\mathcal{G}_i = \lim_{\eta \rightarrow 0^+} [(E + i\eta)\mathcal{I} - \mathcal{H}_i - \tau_{i(i-1)} \mathcal{G}_{i-1} \tau_{(i-1)i}]^{-1} \quad \text{for } 2 \leq i \leq N-1 \quad , \quad (7b)$$

$$\mathcal{G}_N = \lim_{\eta \rightarrow 0^+} [(E + i\eta)\mathcal{I} - \mathcal{H}_N - \tau_{N(N-1)} \mathcal{G}_{N-1} \tau_{(N-1)N} - \Sigma_R]^{-1} \quad , \quad (7c)$$

$$\mathcal{P}_{(i+1)i} = \tau_{(i+1)i} \mathcal{G}_i \quad \text{for } 1 \leq i \leq N-1 \quad , \quad (7d)$$

$$\mathcal{G}_{N1} = \mathcal{G}_N \mathcal{P}_{N(N-1)} \mathcal{P}_{(N-1)(N-2)} \cdots \mathcal{P}_{32} \mathcal{P}_{21} \quad . \quad (7e)$$

75 The computational complexity of this algorithm scales as $\mathcal{O}(N[\dim \mathcal{H}_i]^3)$, which is a factor N^2 better than a direct inversion of the Hamiltonian matrix. A sketch of the RGF-FIS is shown in figure 2a.

3.2. Recursive decimation scheme

The basic idea of the RDS is similar to the previous scheme. The inversion of a 2×2 block matrix where the second row and second column of the inverse matrix are irrelevant, can be reduced to an effective inversion problem of the first block.

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} \mathcal{I} & 0 \\ 0 & \mathcal{I} \end{pmatrix} \Rightarrow \tilde{A}_{11} B_{11} = \mathcal{I} \quad \text{with} \quad \tilde{A}_{11} = A_{11} - A_{12} A_{22}^{-1} A_{21} \quad (8)$$

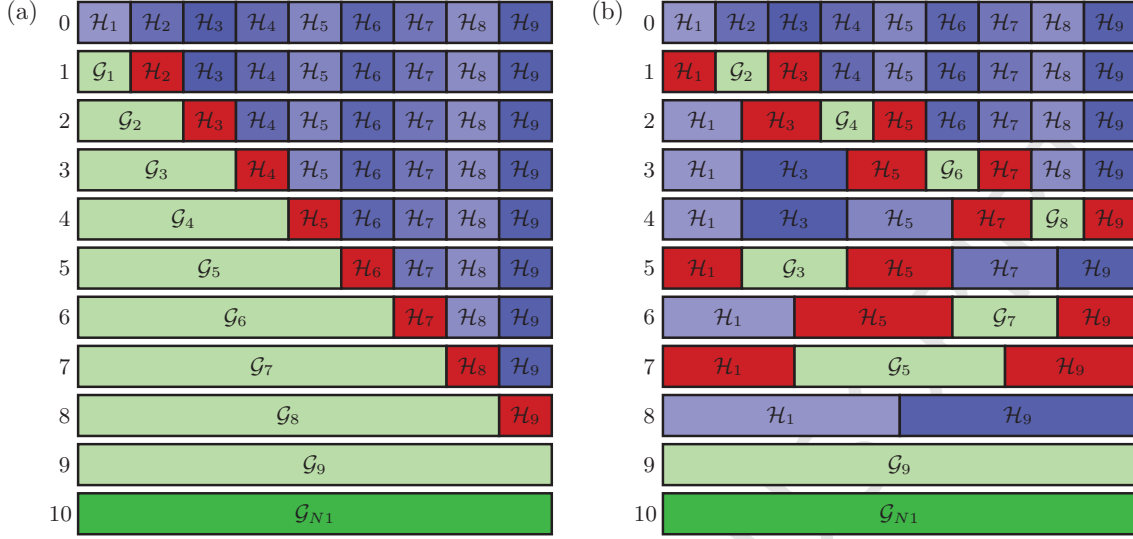


Figure 2: (Color online.) (a) Sketch of the RGF-FIS. (b) Sketch of the RGF-RDS. From top to bottom: Hamiltonian matrices and Green's matrices when applying equations (7) or (9). The numbers at the left of each line denote the iteration step. Different blue blocks denote different Hamiltonian matrices (defect cells). Green blocks denote the calculation of the Green's matrix of the actual iteration step. The dark green block denotes the final step. Red blocks denote Hamiltonian matrices which are coupled to the highlighted Green's matrix. They get a self-energy correction.

Concerning the N parts of the Hamiltonian matrix \mathcal{H}_C and the relevant Green's matrix block \mathcal{G}_{N1} , the subsystems $\mathcal{H}_2, \dots, \mathcal{H}_{N-1}$ can be decimated using this scheme.

For $k \in \{1, 2, \dots, \lceil \log_2(N-1) \rceil\}$: (9a)

For $l \in \left\{0, 1, \dots, \left\lceil \frac{N-1-2^{k-1}}{2^k} - 1 \right\rceil\right\}$: (9b)

$$i := 1 + 2^{k-1} + l \cdot 2^k \quad (9c)$$

$$j_L := 1 + l \cdot 2^k \quad (9d)$$

$$j_R := 1 + (l+1) \cdot 2^k \quad (9e)$$

$$\mathcal{G}_i := \lim_{\eta \rightarrow 0^+} [(E + i\eta)\mathcal{I} - \mathcal{H}_i]^{-1} \quad (9f)$$

$$\mathcal{H}_j := \mathcal{H}_j + \tau_{ji}\mathcal{G}_i\tau_{ij} \quad \forall j \in \{j_L, j_R\} \quad (9g)$$

$$\tau_{j_L j_R} := \tau_{j_L i}\mathcal{G}_i\tau_{i j_R} \quad (9h)$$

$$\tau_{j_R j_L} := \tau_{j_R i}\mathcal{G}_i\tau_{i j_L} \quad (9i)$$

Here, $\lceil \cdot \rceil$ is the ceiling function, i is the actual cell being decimated, and j_L/R is the next left/right cell that has not been decimated so far. In a last step, \mathcal{G}_{N1} can be calculated using equations (7) for a two part system consisting of \mathcal{H}_1 and \mathcal{H}_N . The comparison of equations (7) with equations (9) shows that the FIS is preferable as it needs N inversions and $3N - 3$ multiplications, in contrast to the RDS with N inversions and $6N - 9$ multiplications. A sketch of the RGF-RDS is shown in figure 2b.

3.3. Renormalization decimation algorithm

The RDA is a very fast version of the RGF-RDS for calculating the left (L), right (R), and bulk (B) Green's function of a (half-)infinite periodic system, e.g. the electrodes of a device. Due to the periodicity, this can be done in a convergent iterative process of equations (9) with $N \rightarrow \infty$. For simplicity, let us consider a system with $2^k + 1$ parts. As all subsystems are equal, non-neighboring cells can be decimated

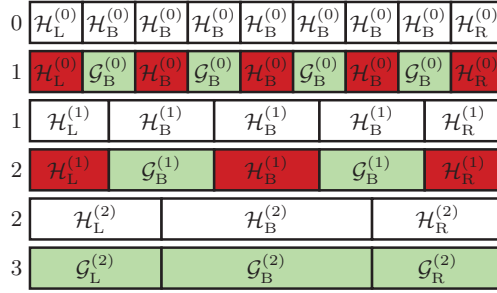


Figure 3: (Color online.) Sketch of the RDA. From top to bottom: Hamiltonian matrices and Green's matrices when applying equations (10). The numbers at the left of each line denote the iteration step. White blocks denote same Hamiltonian matrices (ideal cells). Green blocks denote the calculation of the Green's matrix of the actual iteration step. Red blocks denote Hamiltonian matrices which are coupled to the highlighted Green's matrices. They get a self-energy correction.

in one step. Decimating every second cell, beginning with the second, is most efficient. After this step all remaining subsystems are again equal, except the first and the last one, which will not be decimated. This scheme can be used for all following steps, taking into account that the first and the last cell differ. In the last step, three effective cells remain corresponding to the left, right, and bulk Green's function. The overall RDA reads

$$\mathcal{G}_{L/B/R}^{(i)} = \lim_{\eta \rightarrow 0^+} \left[(E + i\eta)\mathcal{I} - \mathcal{H}_{L/B/R}^{(i)} \right]^{-1} \quad \text{with} \quad \mathcal{H}_{L/B/R}^{(0)} = \mathcal{H} \quad , \quad (10a)$$

$$\alpha^{(i+1)} = \alpha^{(i)} \mathcal{G}_B^{(i)} \alpha^{(i)} \quad \text{with} \quad \alpha^{(0)} = \tau \quad , \quad (10b)$$

$$\beta^{(i+1)} = \beta^{(i)} \mathcal{G}_B^{(i)} \beta^{(i)} \quad \text{with} \quad \beta^{(0)} = \tau^\dagger \quad , \quad (10c)$$

$$\mathcal{H}_B^{(i+1)} = \mathcal{H}_B^{(i)} + \alpha^{(i)} \mathcal{G}_B^{(i)} \beta^{(i)} + \beta^{(i)} \mathcal{G}_B^{(i)} \alpha^{(i)} \quad , \quad (10d)$$

$$\mathcal{H}_L^{(i+1)} = \mathcal{H}_L^{(i)} + \alpha^{(i)} \mathcal{G}_B^{(i)} \beta^{(i)} \quad , \quad (10e)$$

$$\mathcal{H}_R^{(i+1)} = \mathcal{H}_R^{(i)} + \beta^{(i)} \mathcal{G}_B^{(i)} \alpha^{(i)} \quad . \quad (10f)$$

τ is the coupling matrix between the equal subsystems, $\alpha^{(i)}$ and $\beta^{(i)}$ are effective coupling matrices. The iteration can be considered converged if $\|\alpha\| + \|\beta\|$ falls below some threshold. A finite system with $2^k + 1$ parts needs $k - 1$ inversions and $6k - 6$ multiplications. Considering a system with N parts, $\log_2(N + 1) - 1$ inversions and $6\log_2(N + 1) - 6$ multiplications are needed. A sketch of the RDA is shown in figure 3.

4. Improved RGF+RDA

In this section we explain our improved approach, making the RGF, shown in the previous section, faster. We consider a central region with realistic defects like vacancies, substitutional atoms, and functionalisation. In contrast to Anderson disorder, which randomizes the Hamiltonian matrix of the whole system, here, most of the sub-Hamiltonian matrices remain the ideal ones of the periodic system. Only a few sub-Hamiltonian matrices of the defective parts are different. Consequently, the total central region consists of (long) periodic parts which are connected through defective cells. The periodic parts can be treated effectively by decimating all the ideal cells using RDA steps. Especially in cases where the electrodes have the same structure as the defect-free parts of the central region, most of the computations have already been done during the electrode calculation. These decimations lead to corrected Hamiltonian matrices of the defects. Afterwards, the reduced Hamiltonian matrix consists only of corrected defects and can be treated by the original RGF. The computational complexity of the RDA-like part scales as $\mathcal{O}(\log N_D)$, where N_D is the number of defects. The computational complexity of the RGF part scales as $\mathcal{O}(N_D)$.

The RDA-like part is schematically summarized in Figure 4 and can be described as follows. Let us consider one of the periodic parts. Let M be the number of lined up ideal cells. Let l and $r = l + M + 1$ be

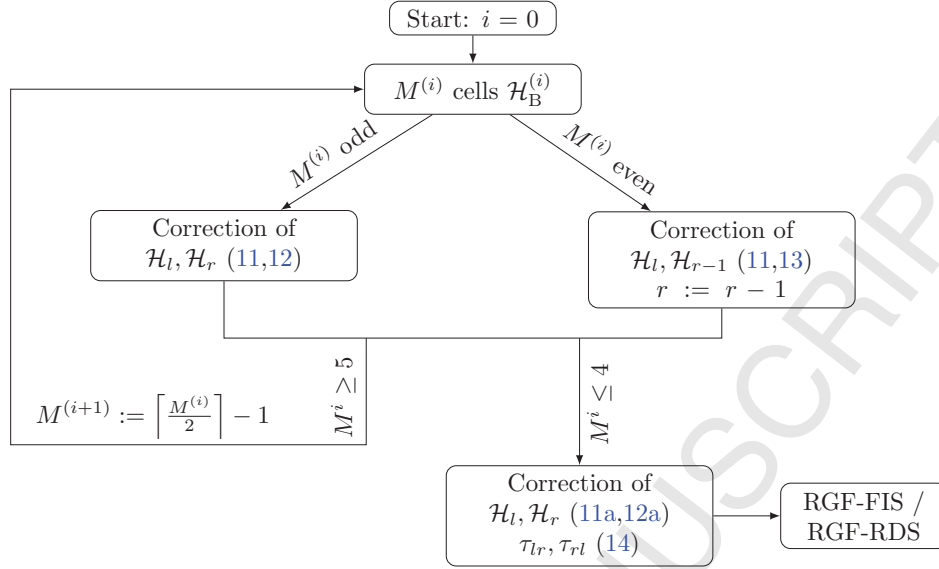


Figure 4: Scheme of the improved RGF+RDA. A periodic part of the system consisting of $M^{(i)}$ ideal cells is treated with RDA-like decimation steps. Different corrections to the neighboring defect cells l and r must be done for odd and even $M^{(i)}$. The iteration is repeated with the half-sized length until all ideal cells have been decimated. After final corrections and the treatment of all other periodic parts, the RGF can be applied.

the indices of the two defect cells at the left and right end. At each step i , one decimates every second cell, starting with index $l + 1$. As defect cell l is connected to cell $l + 1$, the corresponding Hamiltonian matrices and coupling matrices must be corrected in the following way:

$$\mathcal{H}_l^{(i+1)} = \mathcal{H}_l^{(i)} + \tau_{l(l+1)} \mathcal{G}_B^{(i)} \tau_{(l+1)l} \quad , \quad (11a)$$

$$\tau_{l(l+1)}^{(i+1)} = \tau_{l(l+1)}^{(i)} \mathcal{G}_B^{(i)} \alpha^{(i)} \quad , \quad (11b)$$

$$\tau_{(l+1)l}^{(i+1)} = \beta^{(i)} \mathcal{G}_B^{(i)} \tau_{(l+1)l}^{(i)} \quad . \quad (11c)$$

α , β , and \mathcal{G}_B are calculated via equation (10). If $M^{(i)}$ is odd, cell $r - 1$ is decimated and the Hamiltonian matrices and coupling matrices of defect cell r (which is connected to cell $r - 1$) must be corrected in a similar way:

$$\mathcal{H}_r^{(i+1)} = \mathcal{H}_r^{(i)} + \tau_{r(r-1)}^{(i)} \mathcal{G}_B^{(i)} \tau_{(r-1)r} \quad , \quad (12a)$$

$$\tau_{(r-1)r}^{(i+1)} = \alpha^{(i)} \mathcal{G}_B^{(i)} \tau_{(r-1)r}^{(i)} \quad , \quad (12b)$$

$$\tau_{r(r-1)}^{(i+1)} = \tau_{r(r-1)}^{(i)} \mathcal{G}_B^{(i)} \beta^{(i)} \quad . \quad (12c)$$

If $M^{(i)}$ is even, cell $r - 2$ is decimated. In this case, the last ideal cell gets only one correction term instead of two. It cannot be treated any more as an ideal cell within the RDA-like decimation and has to be assigned to the defect cells for further calculation. The corresponding corrections are

$$r := r - 1 \quad , \quad (13a)$$

$$\mathcal{H}_r^{(i+1)} = \mathcal{H}_R^{(i+1)} \quad , \quad (13b)$$

$$\tau_{(r-1)r}^{(i+1)} = \alpha^{(i+1)} \quad , \quad (13c)$$

$$\tau_{r(r-1)}^{(i+1)} = \beta^{(i+1)} \quad . \quad (13d)$$

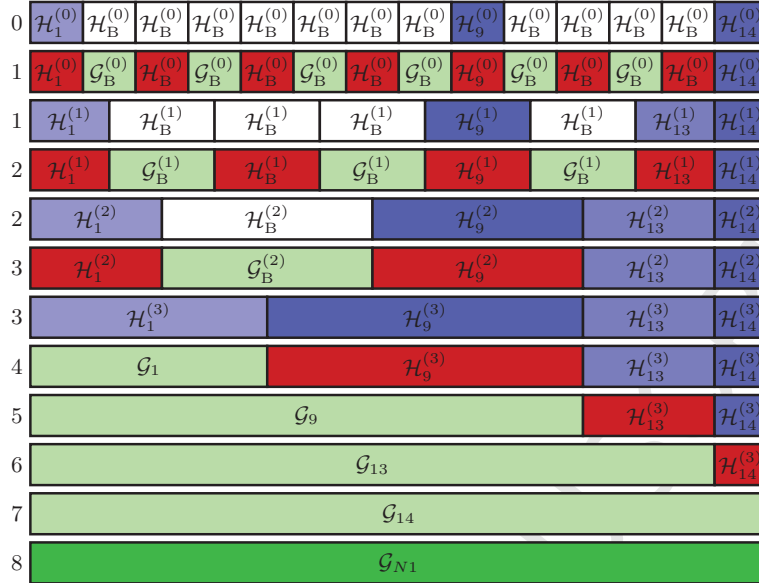


Figure 5: (Color online.) Sketch of the RGF+RDA. From top to bottom: Hamiltonian matrices and Green's matrices when applying equations (11-14). The numbers at the left of each line denote the iteration step. Different blue blocks denote different Hamiltonian matrices (defect cells). White blocks denote same Hamiltonian matrices (ideal cells). Green blocks denote the calculation of the Green's matrix of the actual iteration step. The dark green block denotes the final step. Red blocks denote Hamiltonian matrices which are coupled to the highlighted Green's matrices. They get a self-energy correction. The first seven rows correspond to the RDA, the last six rows correspond to the RGF-FIS (compare figures 3 and 2a).

The last step, which decimates the last cell, yields $l + 1 = r$. Here, we also have to evaluate equations (11a) and (12a), but we get the final effective coupling matrices by means of

$$\tau_{lr}^{(i+1)} = \tau_{l(l+1)}^{(i)} \mathcal{G}_B^{(i)} \tau_{(r-1)r}^{(i)} \quad , \quad (14a)$$

$$\tau_{rl}^{(i+1)} = \tau_{r(r-1)}^{(i)} \mathcal{G}_B^{(i)} \tau_{(l+1)l}^{(i)} \quad , \quad (14b)$$

instead of calculating equations (11b,c;12b,c).

If M is previously decomposed in the basis 2, that means $M = \sum_{i=0}^j M_i 2^i$ with $M_j = 1$, equations (12) of level i (beginning with $i = 0$ and ending with $i = j$) have to be executed if $M_i = 1$ and equations (13) if $M_i = 0$. Equations (11) have always to be executed. For the final step, equations (11a;12a;14) have to be executed.

Up to now, one periodic sequence is described. The other ones can be treated in the same way and independently. Thereafter, the remaining effective defect cells can be used as input for the RGF-FIS or the RGF-RDS. A sketch of the RGF+RDA is shown in figure 5.

The computational complexity of this algorithm cannot be easily obtained exactly and analytically, because the number of additional effective defect cells (and thus the number of matrix multiplications) is determined by the decomposition of M into M_i 's and therefore it strongly depends on M in a non-monotonous way for each sequence. The number of matrix inversions for a sequence of RDA steps, additional effective Hamiltonian matrix calculations, and RGF steps of ideal cells is $\mathcal{N}_{\text{inv}}(M) = 2 \lfloor \log_2 M \rfloor$. The number of corresponding matrix multiplications is in the range $14 \log_2 \frac{M}{3} + 35 \leq \mathcal{N}_{\text{mult}}(M) \leq 19 \log_2 \frac{M+1}{3} + 30$. Considering the whole partly-periodic system with randomly distributed defects, we will get an $\mathcal{O}(\log N_D)$ behavior for these RDA-like calculations (see section 5.1). Afterwards, we are left with N_D effective defect cells, which leads to an $\mathcal{O}(N_D)$ behavior for the RGF. The most important improvement is the logarithmic scaling behavior, stemming from the RDA-like treatment, in contrast to the linear scaling behavior of the RGF. A further advantage is that in cases where the electrodes are identical to the periodic sequences within

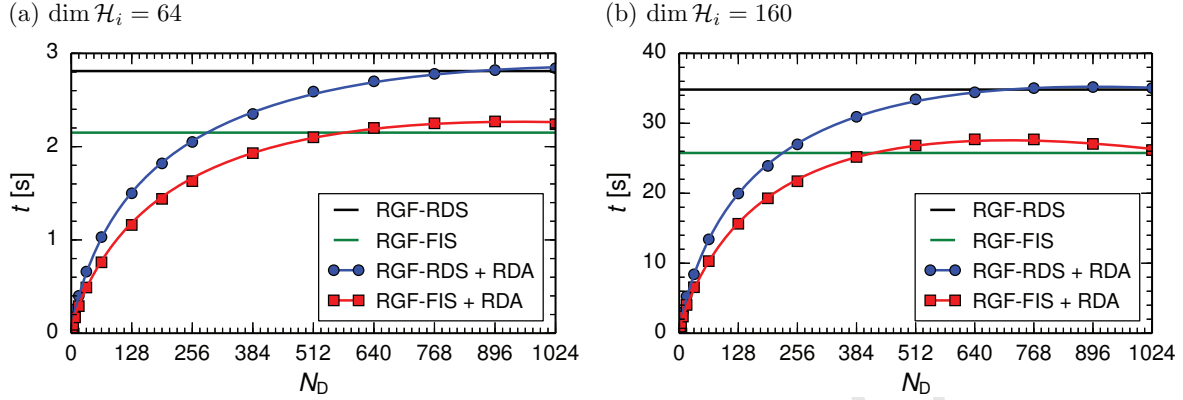


Figure 6: (Color online.) Calculation time for calculating G_{N1} of a random test matrix which consists of 1024 blocks of dimension 64 (a) and 160 (b). $1024 - N_D$ is the number of equal matrix blocks. N_D is the number of deviant matrix blocks (defects). Colors denote the three different algorithms: the RGF-RDS (black), the RGF-FIS (green), and the RGF+RDA (blue and red). Dots and squares are data points. Lines are regressions according to the expected dependence.

the defective bulk part of the system, the RDA steps also occur in the electrode calculation, which further decreases the calculation time.

Finally, we want to make a remark about the generalization to two and three dimensions. For the common RGF this works by treating these systems in a quasi one-dimensional way: A two-dimensional system of $N_1 \times N_2$ cells can be divided into N_1 stripes, which consist of N_2 cells. A three-dimensional system of $N_1 \times N_2 \times N_3$ cells can be divided into N_1 disks, which consist of $N_2 \times N_3$ cells. The RGF can be applied to these N_1 stripes/disks, but their dimensions are then a factor N_2 resp. $N_2 \times N_3$ larger. For the improvement for the case of randomly distributed realistic defects shown in this work, it would in principle be also possible to do so. But this means randomly distributed defective stripes/disks, which are separated by a large two-/three-dimensional area without defects, and randomly distributed defects within these stripes/disks. This does not describe the case of defects which are distributed randomly within the total two-/three-dimensional system. In other words: For a physical two-/three-dimensional system with randomly distributed defects a division into few small stripes/disks with defects cannot work. On the other hand, a cell-wise generalization does not work for the herein discussed real-space RGF, because the decimation of one cell causes coupling elements between all the cells, which are coupled to the decimated one. In one dimension, this transforms the block-tri-diagonal matrix into a similar block-tri-diagonal matrix. But in two/three dimensions, this does not transform the block-penta/hepta-diagonal matrix into a similar block-penta/hepta-diagonal matrix, because for each coupling element, which is removed, 7 resp. 26 coupling elements are added. The resulting algorithm would be worse than direct inversion.

5. Performance test

In this section we evaluate the performance of the algorithm described in section 4, which we have implemented in C++ using the LAPACK routines. We focus on calculation time (wall clock time) and the memory requirement (RAM). We consider two systems: 1) A random test matrix to which we apply only the RGF+RDA. 2) An infinite carbon nanotube with randomly distributed divacancies to which we apply the whole transport formalism.

5.1. Random test matrix

We compare the RGF+RDA steps with the common RGF-FIS and RGF-RDS regarding calculation time and its dependence on the defect fraction. For this purpose, we construct a test matrix, which is block-tridiagonal. The matrix blocks are either ideal ones or defective ones. The ideal matrix blocks are all equal, chosen as a random complex matrix. The defect matrix blocks are all different random complex matrices.

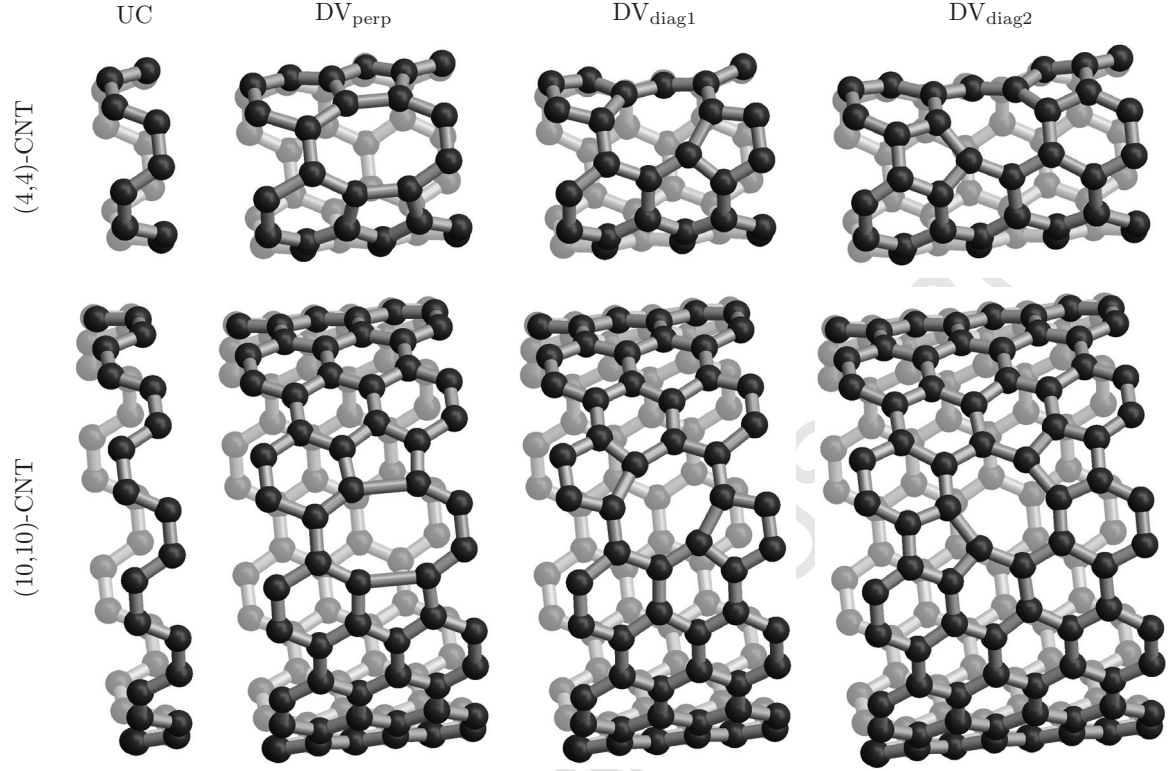


Figure 7: Atomic structure of the used cells of the (4,4)-CNT (upper row) and the (10,10)-CNT (lower row), from left to right: unit cell of the ideal CNT (UC), divacancy with perpendicular orientation (DV_{perp}), and the two types of the divacancy with diagonal orientation (DV_{diag}). Each subsystem of the RGF (figure 1b) is chosen as one of these cells.

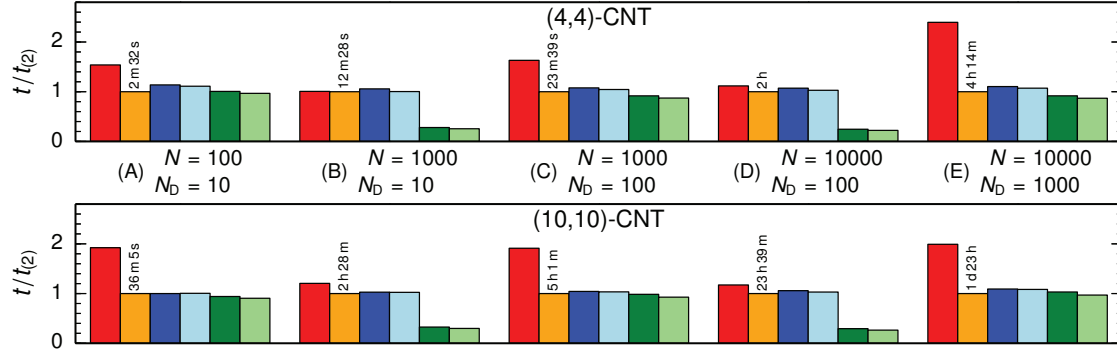
Figure 6 shows the calculation time as a function of the number of defect matrix blocks N_D . The total test matrix consists of $N = 1024$ blocks of dimension $\dim \mathcal{H}_i = 64$ (figure 6a) and $\dim \mathcal{H}_i = 160$ (figure 6b). This is comparable to the respective dimensions of a (4,4)- and a (10,10)-CNT, which we discuss later. The calculation time of the RGF-FIS and the RGF-RDS is independent of the number of matrix blocks. The difference between the FIS and the RDS is the amount of matrix multiplications. The RDS needs $3N - 6$ multiplications more than the FIS.

For low defect fractions, the improved RGF+RDA scales as $\log(N_D)$, according to the RDA part, and as N_D , according to the RGF part. At $N_D = 0$ it results in a pure RDA. Likewise, at $N_D = N$ it results in a pure RGF. Note that for 1024 defect-free cells, a pure RDA needs nearly no time in comparison to a pure RGF. For very high defect fractions (> 0.5), the combined approach requires a bit more calculation time than the RGF alone. This is due to the fact that the RDA-like treatment needs more matrix multiplications for the Hamiltonian matrix corrections than one RGF step. This plays a role, especially, for few and short periodic parts. The calculation time increase, caused by the additional matrix multiplications, is higher than the calculation time reduction, caused by the reduced number of matrix inversions. For fixed N , the overall calculation time of the RGF+RDA can be described by $t = a + bN_D + c \log(d + N_D)$ with specific constants a, b, c, d , in contrast to a constant time within a pure RGF treatment.

5.2. Transport through carbon nanotubes

To study a more realistic system, we choose armchair-(4,4)- and (10,10)-CNTs. We focus on divacancies, which are the most common defects in CNTs [42–44]. Figure 7 shows the atomic structure of their unit cells (UCs) and different orientations of divacancies (DV) within the respective tubes. The unit cell consists of 16

(a) calculation time



(b) memory usage

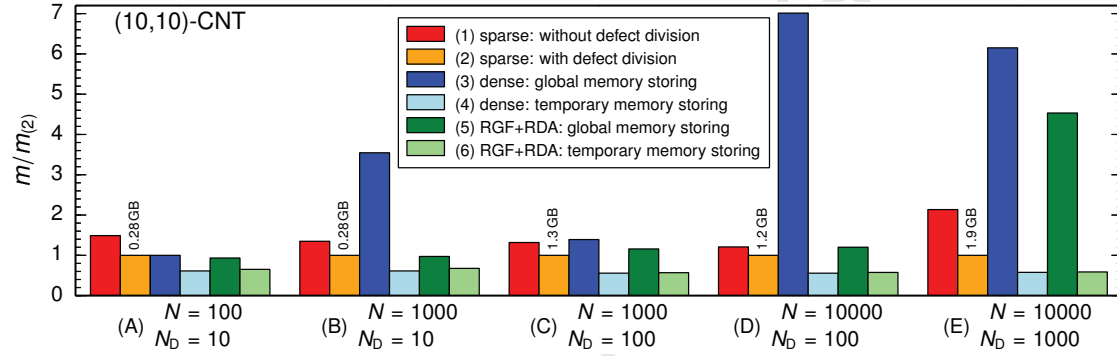


Figure 8: (Color online.) (a) Calculation time for defective (4,4)-CNTs (upper diagram) and (10,10)-CNTs (lower diagram). (b) Memory usage for defective (10,10)-CNTs. Each diagram shows five different cases with different number of CNT cells N and defect cells N_D . Different colors denote different algorithms and matrix treatment (see text for detailed explanation). The values are normalized to case (2), the sparse format with defect division. The absolute values are given above the corresponding bars.

atoms for the (4,4)-CNT and 40 atoms for the (10,10)-CNT. The divacancy defect has the same structure, but with two adjacent atoms removed. There are three different types, labeled DV_{perp} (aligned perpendicular to the tube axis) and $DV_{\text{diag1}}/DV_{\text{diag2}}$ (aligned diagonal). Because of rotational symmetry, each type comprises 8 positions for the (4,4)-CNT and 20 positions for the (10,10)-CNT (which are not shown in the figure). All defect cells are much longer than one unit cell to capture the locally distorted atomic structure, caused by the defect. It results in 46 (DV_{perp} , DV_{diag1}) or 62 (DV_{diag2}) atoms per defect cell for the (4,4)-CNT and 118 (DV_{perp} , DV_{diag1}) or 158 (DV_{diag2}) atoms for the (10,10)-CNT. The geometric structures have been generated by a geometry optimization with density functional theory [3], as implemented in Atomistix ToolKit [45, 46]. We use the local density approximation of Perdew and Zunger [47], norm-conserving Troullier-Martins pseudopotentials [48], and a double zeta plus double polarization (DZDP) basis set of the SIESTA type [49].

The electronic structure (i.e. the Hamiltonian matrices and the coupling matrices) is obtained by a density-functional-based tight-binding (DFTB) model [50, 51]. We use the existing parameter set 3ob [52, 53], which contains onsite energies, distance-dependent hopping energies, and overlap elements for carbon within a four-orbital sp^3 -basis. To reduce the size of the cells, an interaction cutoff distance of twice the carbon-carbon distance was chosen, which is a good compromise between calculation efficiency and accuracy [41]. This leads to a distance-dependent third-nearest-neighbor description, allowing us to use two carbon rings for the unit cell.

Figure 8 shows the calculation time and the memory consumption of the whole transport formalism. Besides the calculation of \mathcal{G}_{N1} , this includes the memory allocation time, some additional calculations to reduce memory requirements, the calculation of the electrodes with the RDA to get $\Sigma_{L/R}$ and $\Gamma_{L/R}$, the calculation of the energy-dependent transmission function using (5), and the calculation of $\mathcal{T}(E)$ after every defect to get the length dependence. We consider five examples with different numbers of cells N and defect cells N_D :

- (A) 100 cells with 10 defects ($p_D = 0.1$),
- (B) 1000 cells with 10 defects ($p_D = 0.01$),
- (C) 1000 cells with 100 defects ($p_D = 0.1$),
- (D) 10 000 cells with 100 defects ($p_D = 0.01$), and
- (E) 10 000 cells with 1000 defects ($p_D = 0.1$).

The defect fraction is either $p_D = 0.1$ (in A, C, E) or $p_D = 0.01$ (in B, D). Furthermore, we distinguish six different cases according to the algorithm and the matrix treatment:

- (1) Defect cells are not divided. The matrices are stored globally in a sparse format. The RGF-FIS (section 3.1) is used.
- (2) Defect cells are divided. The matrices are stored globally in a sparse format. The RGF-FIS is used.
- (3) Defect cells are divided. The matrices are stored globally in a dense format. The RGF-FIS is used.
- (4) Defect cells are divided. The matrices are stored temporarily. The RGF-FIS is used.
- (5) Defect cells are divided. The matrices are stored globally in a dense format. The RGF-FIS+RDA approach (section 4) is used.
- (6) Defect cells are divided. The matrices are stored temporarily. The RGF-FIS+RDA approach is used.

For better comparison the calculation time and the memory consumption shown in figure 8 are normalized to case (2). The corresponding absolute values are given above the respective bars.

First, we want to mention a trivial calculation time reduction, comparing (1) and (2). The cell-wise RGF treatment suggests taking one defect cell in one iteration step (1), but all defects are longer than the ideal UCs due to the extension of the distorted structure. Consequently, the corresponding defect Hamiltonian matrix will also be block-tridiagonal and thus, it can be further divided. Concerning our examples, DV_{perp} and DV_{diag1} are subdivided into two cells and DV_{diag2} into three cells in algorithm (2)². Such a defect division always makes sense as it reduces the calculation time of the corresponding RGF steps in comparison to non-divided defect cells, as can be seen by comparing (1) and (2) in figure 8a.

We implemented two different ways of storing matrices: matrix blocks for different cells types and coupling types are stored once and one of these types is assigned to every cell of the CNT sequence. We call it sparse format. This allows us to do some calculations at the beginning and use the results later, instead of repeating identical computations during the RGF steps. Thus, this should result in a more effective algorithm in comparison to storing the matrix blocks of all cells of the CNT sequence, what we call dense format. In figure 8a we see the reduced calculation time of (2) compared to (3). The reduction lies in the range of 5% to 10% for long systems. The corresponding memory consumption is visualized in figure 8b. The global matrix storage (3) leads to an extensive memory usage, exceeding the acceptable limit of computer resources. But since each iteration step only needs parts of the overall Hamiltonian matrix, they can be loaded temporarily and deleted afterwards. This variant is denoted by (4). The memory usage is reduced drastically and is even better than the one of the sparse format (1, 2). The difference of the matrix storage also affects the calculation time. Since the reservation of memory is lower for the temporary matrix storing (4), also the calculation time is slightly reduced. This reduces the calculation time advantages of using the sparse format (2) to one half.

² Although DV_{perp} and DV_{diag1} have the same length as three UCs, they are subdivided only into two cells, because of the following fact: The interaction cutoff distance was set to twice the carbon-carbon distance, which is equal to the third-nearest-neighbor distance of carbon atoms. At the same time, the UC length is equal to the second-nearest-neighbor distance of carbon atoms, which is allowed, because a coupling between second-nearest-neighbor cells $\tau_{i(i+2)}$ needs an interaction between fourth-nearest-neighbor carbon atoms. Due to the distorted carbon rings in the divacancy structures (see figure 7), a subdivision of DV_{perp} and DV_{diag1} into three cells could cause a second-nearest-neighbor coupling $\tau_{i(i+2)}$, which has to be prevented by dividing DV_{perp} and DV_{diag1} into two cells instead of three. This also applies to the subdivision of DV_{diag2} into three cells instead of four.

The most relevant part of figure 8a is the comparison between (2) and the new RGF+RDA (5) and (6). Note that the improved RGF+RDA cannot be used in combination with the sparse matrix storage because the effective defect matrix blocks are affected differently by the periodic parts in between. For the global matrix storage (5) and the high defect probability $p_D = 0.1$ (cases A, C, E), the calculations take nearly the same time. However, for the low defect probability $p_D = 0.01$ (cases B, D), there is a large reduction of calculation time in the range of 67% to 75%. The calculation time reduction for the (10,10)-CNT is always smaller than for the (4,4)-CNT because of the much larger Hamiltonian matrix blocks. The temporary matrix storage (6) affects the required memory and the calculation time in a similar way as for the dense format. The comparison (6) vs. (5) in figure 8a shows a small decrease of the calculation time. This effect is at most 5%.

In summary, the usage of the RGF+RDA approach with temporary matrix storage (6) is always advantageous compared to the pure RGF (2). It needs less memory (if matrices are only stored temporarily) and it is faster than the common RGF approach. We get calculation time reductions of at least about 5% for large defect probabilities of $p_D = 0.1$, and up to 80% for $p_D = 0.01$. Finally, we want to mention that the results of our calculations, which are not shown in this paper, have been published in [41]. Therein, we presented the conductance of CNTs with monovacancy defects and divacancy defects, and discussed its dependence on various parameters comprehensively.

6. Summary and conclusions

We developed an improved quantum transport algorithm for quasi one-dimensional devices with few realistic defects. The block-tridiagonal Hamiltonian matrix for such devices has long periodic parts, which are interrupted by few defect blocks. The improved RGF+RDA combines the RGF-FIS with the RDA. The periodic parts are treated by the RDA taking advantage of the previously performed electrode calculations. For a fixed number of cells, the scaling behavior of the computational complexity of this new approach has two parts: the RDA part scales logarithmically with the number of defects and the RGF part scales linearly with the number of defects. In contrast to this, the pure RGF would scale linearly with the total number of cells. Overall, this yields a reduction of the computational complexity of the RGF+RDA approach.

The logarithmic scaling behavior is especially advantageous for small defect probabilities, as shown for a test system of random matrices. Applying the algorithm to a more realistic system of practical interest, carbon nanotubes with divacancy defects, the overall calculation time is reduced by up to 80% for $p_D = 0.01$. At that, temporary data loading prevents excessive memory demands (in contrast to global data loading at the beginning), while computation times are nearly unaffected.

This work contributes to the continuing development of numerical implementations in quantum transport theory. Exploring possibilities to unify or combine different approaches, as demonstrated in the present paper, continues to be an important topic for future studies. The RGF+RDA is not limited to CNTs. It can be used for all quasi one-dimensional materials with realistic defects and low defect densities like graphene nanoribbons or nanowires. It is also applicable in a straightforward way to branched systems [54]. Besides that, further improvements towards more arbitrary geometries, e.g. non-periodic edges of graphene nanoribbons or defects in 2D materials would be interesting and promising.

Acknowledgement

This work is funded by the European Union (ERDF) and the Free State of Saxony via the ESF project 100231947 (Young Investigators Group Computer Simulations for Materials Design - CoSiMa).



References

- [1] M. Büttiker, Y. Imry, R. Landauer, S. Pinhas, Generalized many-channel conductance formula with application to small rings, Physical Review B 31 (10) (1985) 6207–6215. doi:10.1103/PhysRevB.31.6207. URL <http://link.aps.org/doi/10.1103/PhysRevB.31.6207>

- [2] S. Datta, Quantum Transport: Atom to Transistor, 1st Edition, Cambridge University Press, Cambridge, 2005.
- [3] K. Capelle, A bird's-eye view of density-functional theory, Brazilian Journal of Physics 36 (4A) (2006) 1318–1343. doi:10.1590/S0103-97332006000700035.
URL http://www.scielo.br/scielo.php?pid=S0103-97332006000700035&script=sci_arttext
- [4] R. Haydock, V. Heine, M. J. Kelly, Electronic structure based on the local atomic environment for tight-binding bands, Journal of Physics C: Solid State Physics 5 (20) (1972) 2845–2858. doi:10.1088/0022-3719/5/20/004.
URL <http://stacks.iop.org/0022-3719/5/i=20/a=004>
- [5] R. Haydock, The recursive solution of the Schrödinger equation, Computer Physics Communications 20 (1) (1980) 11–16. doi:10.1016/0010-4655(80)90101-0.
URL <http://www.sciencedirect.com/science/article/pii/0010465580901010>
- [6] D. J. Thouless, S. Kirkpatrick, Conductivity of the disordered linear chain, Journal of Physics C: Solid State Physics 14 (3) (1981) 235–245. doi:10.1088/0022-3719/14/3/007.
URL <http://stacks.iop.org/0022-3719/14/i=3/a=007>
- [7] A. MacKinnon, The calculation of transport properties and density of states of disordered solids, Zeitschrift für Physik B Condensed Matter 59 (4) (1985) 385–390. doi:10.1007/BF01328846.
URL <http://link.springer.com/article/10.1007/BF01328846>
- [8] M. Wimmer, K. Richter, Optimal block-tridiagonalization of matrices for coherent charge transport, Journal of Computational Physics 228 (23) (2009) 8548–8565. doi:10.1016/j.jcp.2009.08.001.
URL <http://www.sciencedirect.com/science/article/pii/S0021999109004367>
- [9] D. J. Mason, D. Prendergast, J. B. Neaton, E. J. Heller, Algorithm for efficient elastic transport calculations for arbitrary device geometries, Physical Review B 84 (15) (2011) 155401. doi:10.1103/PhysRevB.84.155401.
URL <http://link.aps.org/doi/10.1103/PhysRevB.84.155401>
- [10] E. Cuthill, J. McKee, Reducing the bandwidth of sparse symmetric matrices, in: Proceedings of the 1969 24th National Conference, ACM '69, ACM, New York, NY, USA, 1969, pp. 157–172. doi:10.1145/800195.805928.
URL <http://doi.acm.org/10.1145/800195.805928>
- [11] K. Kazymyrenko, X. Waintal, Knitting algorithm for calculating Green functions in quantum systems, Physical Review B 77 (11) (2008) 115119. doi:10.1103/PhysRevB.77.115119.
URL <http://link.aps.org/doi/10.1103/PhysRevB.77.115119>
- [12] S. Tsukamoto, K. Hirose, S. Blügel, Real-space finite-difference calculation method of generalized Bloch wave functions and complex band structures with reduced computational cost, Physical Review E 90 (1) (2014) 013306. doi:10.1103/PhysRevE.90.013306.
URL <http://link.aps.org/doi/10.1103/PhysRevE.90.013306>
- [13] A. George, Nested dissection of a regular finite element mesh, SIAM Journal on Numerical Analysis 10 (2) (1973) 345–363. doi:10.1137/0710032.
URL <http://epubs.siam.org/doi/abs/10.1137/0710032>
- [14] S. Li, W. Wu, E. Darve, A fast algorithm for sparse matrix computations related to inversion, Journal of Computational Physics 242 (0) (2013) 915–945. doi:10.1016/j.jcp.2013.01.036.
URL <http://www.sciencedirect.com/science/article/pii/S0021999113000892>
- [15] D. E. Petersen, S. Li, K. Stokbro, H. H. B. Sørensen, P. C. Hansen, S. Skelboe, E. Darve, A hybrid method for the parallel computation of Green's functions, Journal of Computational Physics 228 (14) (2009) 5020–5039. doi:10.1016/j.jcp.2009.03.035.
URL <http://www.sciencedirect.com/science/article/pii/S0021999109001648>
- [16] A. Kuzmin, M. Luisier, O. Schenk, Fast methods for computing selected elements of the Green's function in massively parallel nanoelectronic device simulations, in: F. Wolf, B. Mohr, D. an Mey (Eds.), Euro-Par 2013 Parallel Processing, Lecture Notes in Computer Science 8097, Springer-Verlag Berlin, Heidelberg, 2013, pp. 533–544. doi:10.1007/978-3-642-40047-6.
- [17] Y. Fujimoto, K. Hirose, First-principles treatments of electron transport properties for nanoscale junctions, Physical Review B 67 (19) (2003) 195315. doi:10.1103/PhysRevB.67.195315.
URL <http://link.aps.org/doi/10.1103/PhysRevB.67.195315>
- [18] R. Takayama, T. Hoshi, T. Sogabe, S.-L. Zhang, T. Fujiwara, Linear algebraic calculation of the Green's function for large-scale electronic structure theory, Physical Review B 73 (16) (2006) 165108. doi:10.1103/PhysRevB.73.165108.
URL <http://link.aps.org/doi/10.1103/PhysRevB.73.165108>
- [19] T. Ono, Y. Egami, K. Hirose, First-principles transport calculation method based on real-space finite-difference nonequilibrium Green's function scheme, Physical Review B 86 (19) (2012) 195406. doi:10.1103/PhysRevB.86.195406.
URL <http://link.aps.org/doi/10.1103/PhysRevB.86.195406>
- [20] S. Iwase, T. Hoshi, T. Ono, Numerical solver for first-principles transport calculation based on real-space finite-difference method, Physical Review E 91 (6) (2015) 063305. doi:10.1103/PhysRevE.91.063305.
URL <http://link.aps.org/doi/10.1103/PhysRevE.91.063305>
- [21] M. R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, Journal of Research of the National Bureau of Standards 49 (6) (1952) 409–436.
URL http://nvlpubs.nist.gov/nistpubs/jresv049/jresv49n6p409_A1b.pdf
- [22] P. W. Anderson, Absence of diffusion in certain random lattices, Physical Review 109 (5) (1958) 1492–1505. doi:10.1103/PhysRev.109.1492.
URL <http://link.aps.org/doi/10.1103/PhysRev.109.1492>
- [23] A. MacKinnon, B. Kramer, One-parameter scaling of localization length and conductance in disordered systems, Physical Review Letters 47 (21) (1981) 1546–1549. doi:10.1103/PhysRevLett.47.1546.

- URL <http://link.aps.org/doi/10.1103/PhysRevLett.47.1546>
- [24] E. Abrahams, P. W. Anderson, D. C. Licciardello, T. V. Ramakrishnan, *Scaling theory of localization: Absence of quantum diffusion in two dimensions*, Physical Review Letters 42 (10) (1979) 673–676. doi:10.1103/PhysRevLett.42.673. URL <http://link.aps.org/doi/10.1103/PhysRevLett.42.673>
- [25] M. P. López Sancho, J. M. López Sancho, J. Rubio, *Quick iterative scheme for the calculation of transfer matrices: application to Mo (100)*, Journal of Physics F: Metal Physics 14 (5) (1984) 1205–1215. URL <http://stacks.iop.org/0305-4608/14/i=5/a=016>
- [26] M. P. López Sancho, J. M. López Sancho, J. Rubio, *Highly convergent schemes for the calculation of bulk and surface Green functions*, Journal of Physics F: Metal Physics 15 (4) (1985) 851–858. doi:10.1088/0305-4608/15/4/009. URL <http://iopscience.iop.org/0305-4608/15/4/009>
- [27] A. Zienert, J. Schuster, T. Gessner, *Metallic carbon nanotubes with metal contacts: electronic structure and transport*, Nanotechnology 25 (42) (2014) 425203. doi:10.1088/0957-4484/25/42/425203. URL <http://stacks.iop.org/0957-4484/25/i=42/a=425203>
- [28] F. Fuchs, A. Zienert, C. Wagner, J. Schuster, S. E. Schulz, *Interaction between carbon nanotubes and metals: Electronic properties, stability, and sensing*, Microelectronic Engineering 137 (2015) 124–129, Materials for Advanced Metallization 2014. doi:http://dx.doi.org/10.1016/j.mee.2015.02.003. URL <http://www.sciencedirect.com/science/article/pii/S0167931715000556>
- [29] A. Fediai, D. A. Ryndyk, G. Cuniberti, *Electron transport in extended carbon-nanotube/metal contacts: Ab initio based Green function method*, Physical Review B 91 (16) (2015) 165404. doi:10.1103/PhysRevB.91.165404. URL <http://link.aps.org/doi/10.1103/PhysRevB.91.165404>
- [30] B. Biel, F. J. García-Vidal, A. Rubio, F. Flores, *Ab initio study of transport properties in defected carbon nanotubes: an $O(N)$ approach*, Journal of Physics: Condensed Matter 20 (29) (2008) 294214. doi:10.1088/0953-8984/20/29/294214. URL <http://stacks.iop.org/0953-8984/20/i=29/a=294214>
- [31] F. Flores, B. Biel, A. Rubio, F. J. García-Vidal, C. Gómez-Navarro, P. J. de Pablo, J. Gómez-Herrero, *Anderson localization regime in carbon nanotubes: size dependent properties*, Journal of Physics: Condensed Matter 20 (8) (2008) 304211. doi:10.1088/0953-8984/20/30/304211. URL <http://iopscience.iop.org/0953-8984/20/30/304211/>
- [32] B. Biel, F. J. García-Vidal, A. Rubio, F. Flores, *Anderson localization in carbon nanotubes: Defect density and temperature effects*, Physical Review Letters 95 (26) (2005) 266801. doi:10.1103/PhysRevLett.95.266801. URL <http://link.aps.org/doi/10.1103/PhysRevLett.95.266801>
- [33] G. Greene-Diniz, S. L. T. Jones, G. Fagas, M. Haverty, C. M. Lacambra, S. Shankar, J. C. Greer, *Divacancies in carbon nanotubes and their influence on electron scattering*, Journal of Physics: Condensed Matter 26 (4) (2014) 045303. doi:10.1088/0953-8984/26/4/045303. URL <http://stacks.iop.org/0953-8984/26/i=4/a=045303>
- [34] A. T. Lee, Y.-J. Kang, K. J. Chang, *Transport properties of carbon nanotubes: Effects of vacancy clusters and disorder*, The Journal of Physical Chemistry C 116 (1) (2012) 1179–1184. doi:10.1021/jp207943u. URL <http://pubs.acs.org/doi/abs/10.1021/jp207943u>
- [35] F. Khoieini, A. A. Shokri, H. Farman, *Electronic transport through superlattice-like disordered carbon nanotubes*, Solid State Communications 149 (21-22) (2009) 874–879. doi:10.1016/j.ssc.2009.03.005. URL <http://www.sciencedirect.com/science/article/pii/S0038109809001240>
- [36] X. Blase, C. Adessi, B. Biel, A. López-Bezanilla, M.-V. Fernández-Serra, E. R. Margine, F. Triozon, S. Roche, *Conductance of functionalized nanotubes, graphene and nanowires: from ab initio to mesoscopic physics*, Physica Status Solidi B 247 (11-12) (2010) 2962–2967. doi:10.1002/pssb.201000135. URL <http://dx.doi.org/10.1002/pssb.201000135>
- [37] A. López-Bezanilla, X. Blase, S. Roche, *Quantum transport properties of chemically functionalized long semiconducting carbon nanotubes*, Nano Research 3 (4) (2010) 288–295. doi:10.1007/s12274-010-1032-2. URL <http://dx.doi.org/10.1007/s12274-010-1032-2>
- [38] A. López-Bezanilla, F. Triozon, S. Latil, X. Blase, S. Roche, *Effect of the chemical functionalization on charge transport in carbon nanotubes at the mesoscopic scale*, Nano Letters 9 (3) (2009) 940–944. doi:10.1021/nl802798q. URL <http://pubs.acs.org/doi/abs/10.1021/nl802798q>
- [39] M. P. Anantram, T. R. Govindan, *Conductance of carbon nanotubes with disorder: A numerical study*, Physical Review B 58 (6) (1998) 4882–4887. doi:10.1103/PhysRevB.58.4882. URL <http://link.aps.org/doi/10.1103/PhysRevB.58.4882>
- [40] J. Jiang, J. Dong, H. T. Yang, D. Y. Xing, *Universal expression for localization length in metallic carbon nanotubes*, Physical Review B 64 (4) (2001) 045409. doi:10.1103/PhysRevB.64.045409. URL <http://link.aps.org/doi/10.1103/PhysRevB.64.045409>
- [41] F. Teichert, A. Zienert, J. Schuster, M. Schreiber, *Strong localization in defective carbon nanotubes: a recursive Green's function study*, New Journal of Physics 16 (12) (2014) 123026. doi:10.1088/1367-2630/16/12/123026. URL <http://stacks.iop.org/1367-2630/16/i=12/a=123026>
- [42] A. V. Krashenninnikov, K. Nordlund, M. Sirviö, E. Salonen, J. Keinonen, *Formation of ion-irradiation-induced atomic-scale defects on walls of carbon nanotubes*, Physical Review B 63 (24) (2001) 245405. doi:10.1103/PhysRevB.63.245405. URL <http://link.aps.org/doi/10.1103/PhysRevB.63.245405>
- [43] J. A. Rodríguez-Manzo, F. Banhart, *Creation of individual vacancies in carbon nanotubes by using an electron beam of 1 Å diameter*, Nano Letters 9 (6) (2009) 2285–2289, pMID: 19413339. doi:10.1021/nl900463u. URL <http://pubs.acs.org/doi/abs/10.1021/nl900463u>

- [44] S. Kim, H.-J. Kim, H. R. Lee, J.-H. Song, S. N. Yi, D. H. Ha, *Oxygen plasma effects on the electrical conductance of single-walled carbon nanotube bundles*, *Journal of Physics D: Applied Physics* 43 (10) (2010) 305402. doi:10.1088/0022-3727/43/30/305402.
URL <http://iopscience.iop.org/0022-3727/43/30/305402/>
- 410 [45] Atomistix ToolKit version 12.8.2, Quantum Wise A/S.
URL <http://www.quantumwise.com>
- [46] M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, K. Stokbro, *Density-functional method for nonequilibrium electron transport*, *Physical Review B* 65 (16) (2002) 165401. doi:10.1103/PhysRevB.65.165401.
URL <http://link.aps.org/doi/10.1103/PhysRevB.65.165401>
- 415 [47] J. P. Perdew, A. Zunger, *Self-interaction correction to density-functional approximations for many-electron systems*, *Physical Review B* 23 (10) (1981) 5048–5079. doi:10.1103/PhysRevB.23.5048.
URL <http://link.aps.org/doi/10.1103/PhysRevB.23.5048>
- [48] N. Troullier, J. L. Martins, *Efficient pseudopotentials for plane-wave calculations*, *Physical Review B* 43 (3) (1991) 1993–2006. doi:10.1103/PhysRevB.43.1993.
URL <http://link.aps.org/doi/10.1103/PhysRevB.43.1993>
- 420 [49] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, D. Sánchez-Portal, *The SIESTA method for ab initio order-N materials simulation*, *Journal of Physics: Condensed Matter* 14 (11) (2002) 2745–2779. doi:10.1088/0953-8984/14/11/302.
URL <http://iopscience.iop.org/0953-8984/14/11/302>
- 425 [50] D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, R. Kaschner, *Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon*, *Physical Review B* 51 (19) (1995) 12947–12957. doi:10.1103/PhysRevB.51.12947.
URL <http://link.aps.org/doi/10.1103/PhysRevB.51.12947>
- [51] G. Seifert, D. Porezag, T. Frauenheim, *Calculations of molecules, clusters, and solids with a simplified LCAO-DFT-LDA scheme*, *International Journal of Quantum Chemistry* 58 (2) (1996) 185–192. doi:10.1002/(SICI)1097-461X(1996)58:2<185::AID-QUA7>3.0.CO;2-U.
URL [http://dx.doi.org/10.1002/\(SICI\)1097-461X\(1996\)58:2<185::AID-QUA7>3.0.CO;2-U](http://dx.doi.org/10.1002/(SICI)1097-461X(1996)58:2<185::AID-QUA7>3.0.CO;2-U)
- 430 [52] M. Gaus, A. Goez, M. Elstner, *Parametrization and benchmark of DFTB3 for organic molecules*, *Journal of Chemical Theory and Computation* 9 (1) (2013) 338–354. doi:10.1021/ct300849w.
URL <http://pubs.acs.org/doi/abs/10.1021/ct300849w>
- 435 [53] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, G. Seifert, *Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties*, *Physical Review B* 58 (11) (1998) 7260–7268. doi:10.1103/PhysRevB.58.7260.
URL <http://link.aps.org/doi/10.1103/PhysRevB.58.7260>
- 440 [54] G. Thorngilsson, G. Viktorsson, S. I. Erlingsson, *Recursive Green's function method for multi-terminal nanostructures*, *Journal of Computational Physics* 261 (0) (2014) 256–266. doi:10.1016/j.jcp.2013.12.054.
URL <http://www.sciencedirect.com/science/article/pii/S0021999114000096>

Graphical Abstract

With the improved recursive Green's function formalism + renormalization decimation algorithm (RGF+RDA), the conductance of mesoscopic systems with realistic defects can be computed much faster than with the common RGF. For a constant system length, the calculation time t scales logarithmically with the number of defects N_D . This is true for the recursive decimation scheme (RDS) as well as for the forward iteration scheme (RDS).

