

CP2K — Electron Transport based on Non-Equilibrium-Green's-Functions Method: eCSE 08-09 Technical Report

Sergey K. Chulkov, Matthew B. Watkins, Lev N. Kantorovich, and Iain Bethune

March 29, 2018

Abstract

An efficient Non-equilibrium Green's function method has been implemented in the CP2K atomistic simulation package. A variety of techniques have been employed to reduce associated computational costs and achieve better parallel performance.

1 Introduction

CP2K [1] is an open source quantum chemical software package suitable for atomistic simulation of a wide range of systems including isolated molecules, liquid and solid state systems. It allows the user to tackle systems with an order of magnitude more atoms than standard pure plane-wave codes, while still retaining good accuracy. A large variety of methods is available including standard Kohn-Sham Density Functional Theory (DFT), Linear-Scaling Self Consistent Field (LS-SCF) method, Random Phase Approximation (RPA) and many others.

One area that is currently missing in CP2K is the ability to perform quantum transport calculations. Systems with a length scale of a few nanometres are of particular interest, as the wave nature of the electrons should be explicitly considered for correct description of transport properties. The Non-Equilibrium Green's Function (NEGF) method is a popular and powerful tool to study quantum electron transport phenomena. The aim of this project is to implement the NEGF algorithm within CP2K, so the transport properties of nano-scaled systems could be effectively simulated solely within CP2K without relying on external tools.

As far as we are aware, there are several implementations of the NEGF method already in the literature. Non of them, however, was freely available at the time of initial submission of the proposal. Thus TransSIESTA [2] – an external tool which is a part of the SIESTA software package – was originally implemented in 2002 but remained closed-source until the end of 2016. Alternative closed-source implementations of the NEGF methods can be found in SMEAGOL [3] and ADF [4] software packages. At the end of 2016, NEGF calculations also becomes possible with CP2K through its coupling with an external code OMEN [5].

2 General Theory

2.1 Hamiltonian and basis set

A schematic illustration of a typical NEGF setup is shown in Figure 1. An extended system is formed by connecting a molecular device to a number of defect-free semi-infinite electrodes, which serve as electron reservoirs to provide and remove electrons from the central region. Each electrode c has an associated temperature T_c as well as chemical potential μ_c , the latter is defined as a sum of the contact's Fermi energy $E_{F,c}$ and the external potential eV_c applied to the contact, where e is the charge of the electron.

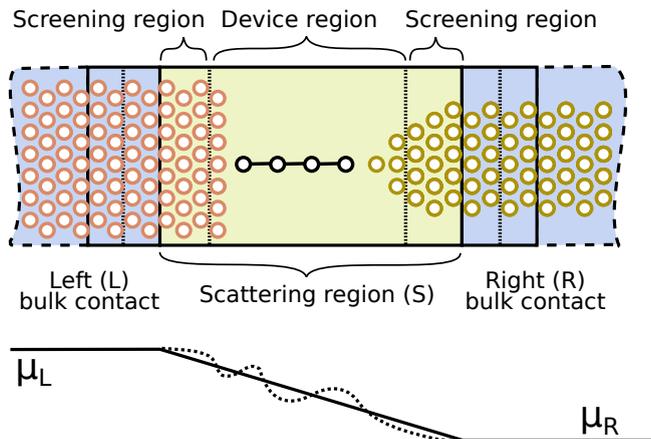


Figure 1: Schematic illustration of a basic NEGF setup.

From the quantum mechanical point of view, the electronic properties of such extended system (e.g. electron density $\rho(\mathbf{r})$) can be obtained by constructing the associated Hamiltonian operator \hat{H} . By choosing a basis set, this Hamiltonian operator can be written in a block-matrix form:

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_d & \mathbf{H}_{dc^1} & \cdots & \mathbf{H}_{dc^n} \\ \mathbf{H}_{dc^1}^\dagger & \mathbf{H}_{c^1} & \cdots & \mathbf{H}_{c^1c^n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{dc^n}^\dagger & \mathbf{H}_{c^1c^n}^\dagger & \cdots & \mathbf{H}_{c^n} \end{pmatrix}, \quad (1)$$

where indices d and c^i refer to the molecular device and the i -th electrode. Note that all contacts' matrix blocks (\mathbf{H}_{c^i} and \mathbf{H}_{dc^i}) have an infinite rank. Fortunately, the electrodes some distance away from the molecular device have a rather regular periodic structure, which can be exploited to make the problem tractable. Thus, away from the molecular device the electrodes have bulk-like properties and can be taken into account via the corresponding self-energy terms. By using an atomic-centred Gaussian basis set and choosing a unit cell of appropriate size along its semi-infinite direction, it is always possible to ensure near-neighbour interaction. In other words we assume that electrons centred on atoms from a given unit cell of each electrode interact exclusively with nuclei and other electrons from this cell and its immediate neighbours only. It follows that both the overlap (\mathbf{S}_c) and Hamiltonian (\mathbf{H}_c) matrices of the c -th electrode in the given basis set representation have the following block-tridiagonal form:

$$\mathbf{H}_c = \begin{pmatrix} \mathbf{H}_{c_0} & \mathbf{H}_{c_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{c_1}^\dagger & \mathbf{H}_{c_0} & \mathbf{H}_{c_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{c_1}^\dagger & \mathbf{H}_{c_0} & \ddots \\ \mathbf{0} & \mathbf{0} & \ddots & \ddots \end{pmatrix}, \quad (2)$$

where indices c_0 and c_1 denote diagonal and off-diagonal matrix blocks respectively.

As we approach the molecular device, the local atomic structure of the electrode near the surface boundary becomes different from that in the bulk. Moreover, reconstruction of the contact's surface can also take place due to chemical interaction with the connected molecular device. Because of this, we should consider such atoms explicitly by including them into the device region. For proper description of the system it is also important that the electrostatic potential near the bulk region boundary should be identical to that in its bulk. For this reason, few layers of the bulk contact at the boundaries are assigned to a special screening region – one for every electrode – which forms the final NEGF scattering region alongside with the device region itself.

By choosing the scattering region of the appropriate size it is also possible to avoid direct interaction between different electrodes. The infinite-rank overlap (\mathbf{S}) and Hamiltonian (\mathbf{H}) matrices from Eq. 1

become substantially sparse with zero blocks located everywhere except the first row, the first column, and the main diagonal:

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_S & \mathbf{H}_{Sc^1} & \cdots & \mathbf{H}_{Sc^n} \\ \mathbf{H}_{Sc^1}^\dagger & \mathbf{H}_{c^1} & \mathbf{0} & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{H}_{Sc^n}^\dagger & \mathbf{0} & \mathbf{0} & \mathbf{H}_{c^n} \end{pmatrix}, \quad (3)$$

where indices S and c^i denote the scattering region and the i -th electrode respectively. If only the nearest unit cells interact, the off-diagonal matrix block \mathbf{H}_{Sc^i} couples the scattering region with a nearest unit cell c_0^i of the i -th semi-infinite electrode and can be written as:

$$\mathbf{H}_{Sc^i} = \begin{pmatrix} \mathbf{H}_{Sc_0^i} & \mathbf{0} & \mathbf{0} & \cdots \end{pmatrix}. \quad (4)$$

The following identity:

$$[E\mathbf{S} - \mathbf{H}] \mathbf{G}(E) = \mathbf{I}, \quad (5)$$

defines the Green's function matrix for the given Hamiltonian with \mathbf{I} being an identity matrix. Applying it separately to the scattering S and i -electrode Hamiltonian, one can compute the retarded Green's functions of the scattering region \mathbf{G}_S and the i -th electrode \mathbf{G}_c as:

$$\mathbf{G}_S(E) = [(E + i\eta)\mathbf{S}_S - \mathbf{H}_S]^{-1}, \quad (6)$$

$$\mathbf{G}_{c^i}(E) = [(E + i\eta)\mathbf{S}_{c^i} - \mathbf{H}_{c^i}]^{-1}, \quad (7)$$

where η is an infinitesimal real number to make the inversion numerically well-defined for any given energy E .

2.2 Surface Green's function

Within the NEGF method the presence of the semi-infinite electrodes is taken into account via the corresponding self-energy term:

$$\Sigma_c(E) = (E\mathbf{S}_{Sc} - \mathbf{H}_{Sc})\mathbf{G}_c(E)(E\mathbf{S}_{Sc}^\dagger - \mathbf{H}_{Sc}^\dagger), \quad (8)$$

where the matrices \mathbf{G}_c , \mathbf{S}_{Sc} , and \mathbf{H}_{Sc} are defined via Eqs. 3 and 7. Note that the self-energy matrix has a finite rank due to sparsity of the underlying overlap and Hamiltonian matrices:

$$\begin{aligned} \Sigma_c(E) &= (E\mathbf{S}_{Sc_0} - \mathbf{H}_{Sc_0} \quad \mathbf{0} \quad \mathbf{0} \quad \cdots) \begin{pmatrix} \mathbf{G}_{c_{00}} & \mathbf{G}_{c_{01}} & \mathbf{G}_{c_{02}} & \cdots \\ \mathbf{G}_{c_{10}} & \mathbf{G}_{c_{11}} & \mathbf{G}_{c_{12}} & \cdots \\ \mathbf{G}_{c_{20}} & \mathbf{G}_{c_{21}} & \mathbf{G}_{c_{22}} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} E\mathbf{S}_{Sc_0}^\dagger - \mathbf{H}_{Sc_0}^\dagger \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \end{pmatrix} \\ &= (E\mathbf{S}_{Sc} - \mathbf{H}_{Sc})\mathbf{G}_{c_{00}}(E)(E\mathbf{S}_{Sc}^\dagger - \mathbf{H}_{Sc}^\dagger). \end{aligned} \quad (9)$$

For this reason, despite the fact that the retarded Green's function matrix is rather dense, only a single matrix block $\mathbf{G}_{c_{00}}$ associated with the boundary electrode's cell only needs to be known. This matrix block is called the retarded surface Green's function, and can be computed using a López Sancho recursive algorithm [6].

The foundation of this algorithm is rather straightforward. We start from the identity given in Eq. 5 for the retarded Green's function:

$$\begin{pmatrix} \mathbf{A}_{c_0} & \mathbf{A}_{c_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{c_1}^\dagger & \mathbf{A}_{c_0} & \mathbf{A}_{c_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{c_1}^\dagger & \mathbf{A}_{c_0} & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \mathbf{G}_{c_{00}} & \mathbf{G}_{c_{10}} & \mathbf{G}_{c_{20}} & \mathbf{G}_{c_{i,0}} \\ \mathbf{G}_{c_{10}} & \mathbf{G}_{c_{11}} & \mathbf{G}_{c_{21}} & \mathbf{G}_{c_{i,1}} \\ \mathbf{G}_{c_{20}} & \mathbf{G}_{c_{21}} & \mathbf{G}_{c_{22}} & \cdots \\ \mathbf{G}_{c_{0,i}} & \mathbf{G}_{c_{1,i}} & \cdots & \cdots \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots \end{pmatrix}, \quad (10)$$

where \mathbf{G}_{c_0} is the target surface Green's function and the individual matrix blocks are defined as:

$$\mathbf{A}_{c_0} = E\mathbf{S}_{c_0} - \mathbf{H}_{c_0}, \quad \mathbf{A}_{c_1} = E\mathbf{S}_{c_1} - \mathbf{H}_{c_1}, \quad \mathbf{G}_{c_i,j}(E) = \mathbf{G}_{c_j,i}(E). \quad (11)$$

Multiplication of the first row of the matrix \mathbf{A}_c by the first column of the matrix \mathbf{G}_c gives us the first chain equation:

$$(E\mathbf{S}_{c_0} - \mathbf{H}_{c_0})\mathbf{G}_{c_0} + (E\mathbf{S}_{c_1} - \mathbf{H}_{c_1})\mathbf{G}_{c_1} = \mathbf{I}. \quad (12)$$

The second chain equation can be derived by multiplying the $i+1$ -th row of the matrix \mathbf{A} by the first column of the matrix \mathbf{G} :

$$(E\mathbf{S}_{c_0} - \mathbf{H}_{c_0})\mathbf{G}_{c_i,0} + (E\mathbf{S}_{c_1} - \mathbf{H}_{c_1})^T\mathbf{G}_{c_{i-1},0} + (E\mathbf{S}_{c_1} - \mathbf{H}_{c_1})\mathbf{G}_{c_{i+1},0} = \mathbf{0}. \quad (13)$$

Comparison of the last two equations with Eq. 3 from Ref. [6] allows us to extend an iterative López Sancho method for the case of a non-orthogonal basis set.

Thus, in order to obtain the retarded surface Green's function of the semi-infinite electrode c , initial values of four matrices (α , β , ϵ , and $\epsilon^{\text{surf.}}$) need to be computed:

$$\epsilon_0^{\text{surf.}} = \epsilon_0 = E\mathbf{S}_{c_0} - \mathbf{H}_{c_0}, \quad \alpha_0 = E\mathbf{S}_{c_1} - \mathbf{H}_{c_1}, \quad \beta_0 = (E\mathbf{S}_{c_1} - \mathbf{H}_{c_1})^T. \quad (14)$$

These matrices are then iteratively refined using the following relations:

$$\begin{aligned} g_i &= \epsilon_i^{-1}, & \alpha_{i+1} &= \alpha_i g_i \alpha_i, & \beta_{i+1} &= \beta_i g_i \beta_i, & \zeta_{i+1} &= \alpha_i g_i \beta_i, & \eta_{i+1} &= \beta_i g_i \alpha_i, \\ \epsilon_{i+1} &= \epsilon_i - \zeta_{i+1} - \eta_{i+1}, & \epsilon_{i+1}^{\text{surf.}} &= \epsilon_i^{\text{surf.}} - \zeta_{i+1}. \end{aligned} \quad (15)$$

until convergence of the matrices α and β . Finally, the surface Green's function is computed by simple matrix inversion:

$$\mathbf{G}_{c_0}(E) = \left(\epsilon_{\text{last}}^{\text{surf.}} \right)^{-1}. \quad (16)$$

2.3 NEGF method

Having self-energy matrices Σ_c for every contact c , we can compute the retarded Green's function matrix of the scattering region which is the central quantity of the NEGF method:

$$\mathbf{G}_S(E) = \left[(E + i\eta)\mathbf{S}_S - \mathbf{H}_S^{\text{consistent}} - \sum_c \Sigma_c(E - V_c) \right]^{-1}. \quad (17)$$

Another important quantity is the broadening matrix:

$$\Gamma_c(E) = i \left(\Sigma_c(E) - \Sigma_c^\dagger(E) \right), \quad (18)$$

where the advanced self-energy $\Sigma_c^\dagger(E)$ can be obtained by conjugate transposition of the corresponding retarded self-energy matrix $\Sigma_c(E)$.

The knowledge of these quantities gives access to various properties of the extended molecule. Among them are density of states (DOS):

$$\text{DOS}(E) = -\frac{1}{\pi} \text{Tr}[\mathbf{G}_S(E)\mathbf{S}_S], \quad (19)$$

transmission from the c -th electrode to the c' -th electrode [7]:

$$T_{cc'}(E) = \text{Tr}[\mathbf{G}_S(E)\Gamma_c(E)\mathbf{G}_S^\dagger(E)\Gamma_{c'}(E)], \quad (20)$$

and the steady-state current through the c -th electrode:

$$I_c = \frac{1}{2\pi} \int \sum_{c'} T_{cc'}(E) [f_{T_c}(E - \mu_c) - f_{T_{c'}}(E - \mu_{c'})] dE, \quad (21)$$

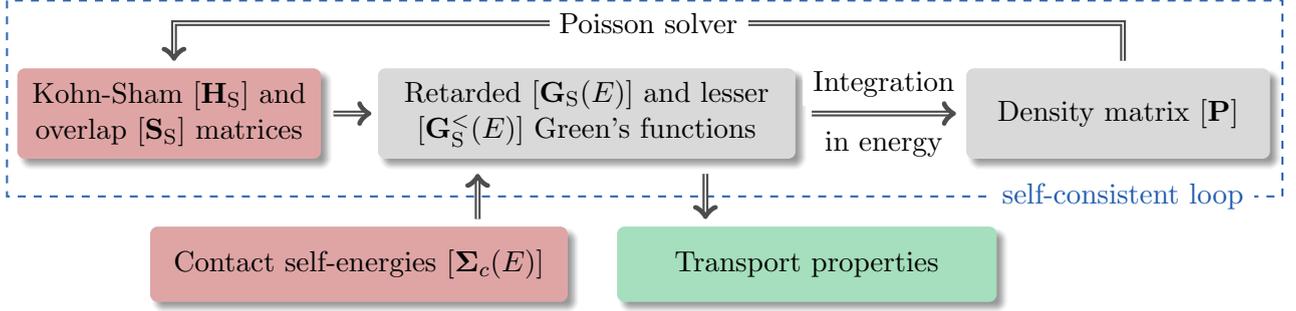


Figure 2: Block diagram of the NEGF algorithm.

where f_T denotes the Fermi function at the temperature T :

$$f_T(E) = \left[\exp\left(\frac{E}{k_B T}\right) + 1 \right]^{-1}, \quad (22)$$

and k_B is the Boltzmann constant.

In the zero-bias case the Hamiltonian matrix obtained by solving the electronic Schrödinger equation – which can be done for example, by using Kohn-Sham DFT – is already consistent with the electron density of the molecular system. Because of this, the term $\mathbf{H}_S^{\text{consistent}}$ in Eq. 17 becomes simply the Hamiltonian matrix of the scattering region \mathbf{H}_S (see Eq. 3). On the contrary, when an external potential is applied to electrodes, current flows which affects the electron density within the scattering region. It means that the initial Hamiltonian matrix \mathbf{H}_S is no longer consistent with the actual density and a self-consistent procedure needs to be introduced to refine the Hamiltonian matrix.

The main idea behind this self-consistent procedure is the relation between the density matrix \mathbf{P} and the lesser Green's function $\mathbf{G}_S^<(E)$:

$$\mathbf{P} = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \mathbf{G}_S^<(E) dE. \quad (23)$$

As it is shown in Ref. [3], in NEGF formalism the lesser Green's function can be expressed in terms of the retarded Green's function and the corresponding broadening matrices:

$$\mathbf{G}_S^<(E) = i\mathbf{G}_S(E) \sum_c [\Gamma_c(E - V_c) f_{T_c}(E - \mu_c)] \mathbf{G}_S^\dagger(E), \quad (24)$$

The density matrix from Eq. 23 can be used to refine the Hamiltonian matrix – for example, by solving Kohn-Sham equations – which in turn gives new retarded and lesser Green's function as well as an updated density matrix. The block-scheme of the typical NEGF algorithm is shown in Figure 2 and it will be discussed in details in the next section.

3 Implementation

The current implementation of the NEGF method takes advantage of the heterogeneous nature of the electronic device. On the one hand, it contains the scattering region where atomic structure needs to be explicitly considered. On the other hand, it also contains a number of semi-infinite electrodes, which are taken into account via the corresponding self-energy terms. Such composite system can be naturally split across a number of `FORCE_ENV` sections. By convention, atoms belonging to the scattering region should be assigned to the section with index 0, while the remained sections correspond to semi-infinite contacts. In particular, the section with index c should contains atoms which belong to a single bulk unit cell of the c -th electrode. If needed, however, the `FORCE_ENV` sections can be given in arbitrary order and then rearranged with help of the `MULTIPLE_FORCE_EVALS` section.

One advantage of using CP2K multiple force environments is the fact that this approach allows us to reduce the number of active atoms in actual calculations by using appropriate boundary conditions

for different parts of the system. Thus, Dirichlet boundary conditions can be applied to the scattering region by using the relaxation Poisson solver [8] or the implicit Poisson solver [9] during NEGF self-consistent iterations. On the other hand, a single contact's unit cell with periodic boundary conditions is sufficient to properly describe bulk properties of the corresponding semi-infinite lead. By solving the Kohn-Sham equation at a set of \mathbf{k} -points aligned along the semi-infinite direction, it is then possible to obtain the contact's matrix blocks \mathbf{H}_{c_0} and \mathbf{H}_{c_1} needed for NEGF calculation. Thus, having Kohn-Sham matrices $\mathbf{H}_{c,\mathbf{k}}$ for every \mathbf{k} -point \mathbf{k} , the required matrix blocks are computed as:

$$\mathbf{H}_{c_0} = \sum_{\mathbf{k}} w_{\mathbf{k}} \mathbf{H}_{c,\mathbf{k}}, \quad \mathbf{H}_{c_1} = \sum_{\mathbf{k}} w_{\mathbf{k}} \mathbf{H}_{c,\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}}, \quad (25)$$

where $w_{\mathbf{k}}$ is the weight assigned to the \mathbf{k} -th \mathbf{k} -point, and \mathbf{R} is the lattice vector of the contact's unit cell along the semi-infinite direction.

Despite the fact that we do need \mathbf{k} -points to be set up for all contact `FORCE_ENV` sections, they are not currently supported for the device `FORCE_ENV` section itself. Whereas it is relatively easy to implement \mathbf{k} -points for the scattering region as well (see Ref [7] for more details), such feature seems to be rarely needed. The reason is that the device in question is usually large enough to be well-described by using Γ -point only. Moreover, it is more common to simulate transport properties of a single instance of the device rather than the series of identical devices arranged in close proximity to each other and repeated in transverse directions. Because of this, even if the size of the molecular device itself is relatively small, the size of the simulation box for the device region is still appropriate for the Γ -point calculation.

The workflow of the NEGF algorithm is the following.

1. By iteratively solving Kohn-Sham equations within DFT obtain diagonal and off-diagonal Kohn-Sham (\mathbf{H}_{c_0} , \mathbf{H}_{c_1}) and overlap (\mathbf{S}_{c_0} , \mathbf{S}_{c_1}) matrix blocks for each contact c :
2. Solve Kohn-Sham equations for the scattering region and obtain the following matrix blocks: \mathbf{H}_S , $\mathbf{H}_{S_{c_0}}$, \mathbf{S}_S , and $\mathbf{S}_{S_{c_0}}$.
3. If not provided explicitly via the input file, refine the Fermi energy of each bulk contact c using the NEGF algorithm by considering the semi-infinite electrode as a two-terminal device with left (l) and right (r) contacts and under zero bias. Taking into account the nearest-neighbour interaction, the relevant blocks of both overlap (\mathbf{S}) and Kohn-Sham (\mathbf{H}) matrices satisfy the following relations:

$$\begin{aligned} \mathbf{H}_S &= \mathbf{H}_{l_0} = \mathbf{H}_{r_0} = \mathbf{H}_{c_0}, \\ \mathbf{H}_{S_{r_0}} &= \mathbf{H}_{r_0} = \mathbf{H}_{S_{l_0}}^\dagger = \mathbf{H}_{l_0}^\dagger = \mathbf{H}_{c_1}. \end{aligned} \quad (26)$$

Using a linear interpolation technique we then try to locate the optimal Fermi energy that converges the NEGF algorithm in one iteration.

4. Using a López Sancho algorithm compute the surface Green's functions and self-energy matrices for every semi-infinite electrode. These quantities remain constant during NEGF iterations, so self-energy matrices are precomputed once and then cached into the main memory.
5. Using the self-consistent procedure shown in Figure 2, obtain the Kohn-Sham matrix block H_S which is consistent with the corresponding density matrix \mathbf{P} . The main bottleneck here is the energy integration step due to the fact that the lesser Green's function has poles on the real axis, so a fine grid mesh (typically 100 points per 1 eV) is required. Following the common practice [2], we add and subtract the term $\sum_{c' \neq c} \Gamma_{c'}(E - V_{c'}) f_{T_{c'}}(E - \mu_{c'})$ from the Eq. 24 which

gives us the equilibrium ($\mathbf{P}_{\text{eq.}}$) and non-equilibrium (\mathbf{P}_V) components of the density matrix \mathbf{P} :

$$\mathbf{P} = \mathbf{P}_{\text{eq.}}^c + \sum_{c' \neq c} \mathbf{P}_V^{c,c'}, \quad (27)$$

$$\mathbf{P}_{\text{eq.}}^c = -\frac{1}{\pi} \int_{-\infty}^{\infty} \text{Im}(\mathbf{G}_S(E)) f_{T_c}(E - \mu_c) dE, \quad (28)$$

$$\mathbf{P}_V^{c,c'} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{G}_S(E) \mathbf{\Gamma}_{c'}(E - V_{c'}) \mathbf{G}_S^\dagger(E) (f_{T_{c'}}(E - \mu_{c'}) - f_{T_c}(E - \mu_c)) dE. \quad (29)$$

In the case of the equilibrium density component, the expression under the integral sign is an analytic function, so the corresponding integral can be easily calculated along a closed-contour path in the complex plane. In particular, the function along the integration path is sampled at a discrete set of points which form Chebyshev roots. Then the integral is computed using an efficient adaptive Clenshaw-Curtis algorithm which allows us to reduce the number of grid points up to 3 times compared to the standard approach based on Simpson rules. In contrast, the non-equilibrium part of the density matrix needs to be computed along the real axis within the narrow interval $[\mu_c + k_B T_c \ln \epsilon, \mu_c + k_B T_c \ln \epsilon]$, where ϵ is the target accuracy for the density matrix. As the lesser Green's function has a substantial amount of poles on the real axis, the adaptive integration method based on Simpson rules typically outperforms more advanced approaches, such as adaptive Clenshaw-Curtis method.

6. Upon convergence of the density matrix, various properties for the extended molecule are computed. Individual properties are requested via their respective PRINT subsections.

Calculation of Green's function matrices requires a substantial amount of matrix inversions, which are handled by using ScaLAPACK routines. In CP2K dense matrices are stored in BLACS format and by default distributed across all available MPI processes in a block-cyclic fashion. Being in general very reasonable, this distribution scheme, however, is not optimal in case of a small matrices as it leads to non-uniform distribution of the matrix blocks. Moreover, this distribution scheme can be non-optimal even for a moderately sized matrices when the number of parallel processes is rather large. Thus, substantial amount of CPU time is spent in matrix-matrix multiplication and matrix inversion routines which asymptotically scales with the number of processor as \sqrt{p} . In order to increase CPU utilisation, all MPI processes are arranged into a number of parallel groups; each group computes Green's functions at their respective set of energy points. In a typical NEGF calculation the number of points where Green's functions need to be evaluated is several hundreds. It has to be noted, however, that the actual position of these grid points is not known in advance, but is determined iteratively by the adaptive integration algorithms used. In order to get most of the parallelisation over energy approach, the number of new grid points at every iteration should be proportional to the number of available parallel groups. Thus, at every iteration of adaptive algorithm, we first identify intervals where the integral needs to be refined. Then we subdivide only those intervals which have largest error estimate and allows us to satisfy the constraint on the number of new points at the same time.

4 Results

The results presented in this report were obtained using CP2K 6.0 (revision 18380) on ARCHER – a Cray XC30 supercomputer with standard nodes containing two 2.7 GHz 12-core Intel Xeon E5-2697 v2 series processors and 64 GB of random access memory (RAM).

We start from a very basic test case – an infinite chain of the metallic atoms. In particular we focus on the lithium which is the first element in periodic system with strong metallic properties. The structure of the molecule was optimised in CP2K using the LDA exchange-correlation functional based on Padé approximant [10] and SZV-MOLOPT-SR-GTH basis set [11], while core electrons were taken into account using Goedecker-Hutter-Tetter (GTH) pseudo-potentials [10]. A rectangular unit cell with size of 10 Å in transverse direction and plane-wave cutoff of 500 Ry was employed with periodic

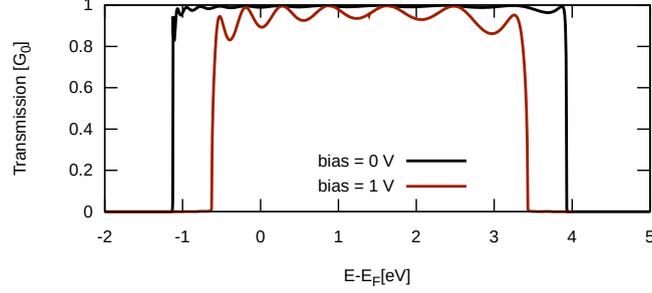


Figure 3: Transmission coefficients of Li-chain.

boundary conditions applied in all three dimensions. The optimal distance between lithium atoms was found to be 2.99 \AA . The contact's unit cell was composed of 4 atoms, which is found to be sufficient to satisfy the nearest-neighbour assumption. To be more specific, the absolute value of any element of the Hamiltonian matrix belonging to the first and the ninth lithium atoms in the chain does not exceed 10^{-17} . The scattering region was constructed using 12 atoms, four of them were explicitly included into the device region while the remaining 8 atoms – 4 atoms per contact – form the screening region. Transmission coefficients for zero bias case and with the applied external bias are shown in Figure 3 which is in agreement with reference calculations available in the literature [4].

As the second example we consider a di-thiol-benzene molecule connected to two identical semi-infinite Au(111) electrodes with 3×3 gold atoms per electrodes' plane. At the beginning the unit cell of bulk gold was optimised using the Padé LDA exchange-correlation functional, 500 Ry plane wave cutoff, and SZV-MOLOPT-SR-GTH basis set. The equilibrium distance between gold atoms in bulk was found to be equal to 2.92 \AA . The di-thiol-benzene molecule was deposited on the hollow site between the gold electrodes separated by 9.2 \AA . The geometry optimisation was then performed using the same exchange-correlation functional and TZVP-MOLOPT-GTH [11] basis set for all atoms except gold, for which SZV-MOLOPT-SR-GTH basis set was utilised. All gold atoms were frozen except the ones from the surface plane which were allowed to relax. The Au-S and C-S interatomic distances are found to be equal to 2.45 \AA and 1.75 \AA which are in good agreement with the reference calculation [12]. The obtained structure was then split into several parts to set up a NEGF calculation. Each contact's unit cell was constructed using 3 gold layers (27 atoms) from the bulk region, while the di-thiol-benzene molecule itself and the nearest 6 gold layers from each electrode (120 atoms in total) were assigned to the scattering region.

Computed zero-bias transmission coefficient and steady-state current as a function of bias are shown in Figure 4 and found to be in good agreement with reference calculations available in the

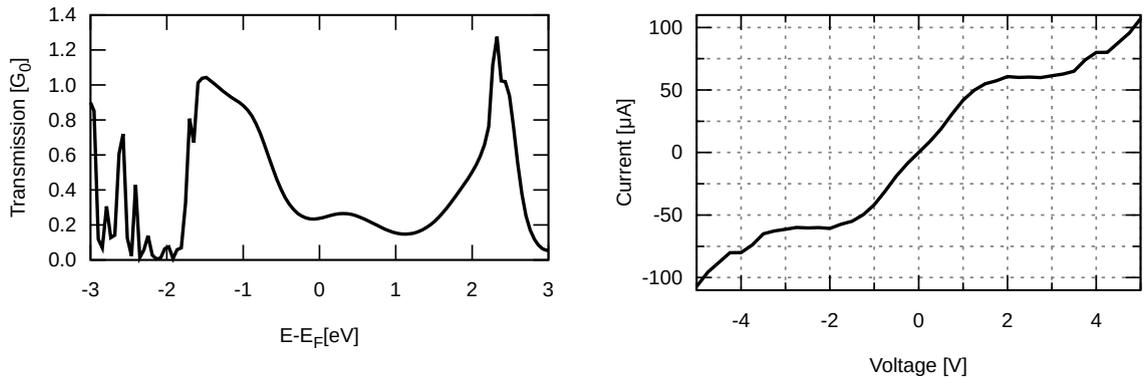


Figure 4: Zero-bias transmission coefficient (left) and current as a function of bias (right) of di-thiol-benzene molecule connected to two Au(111) electrodes.

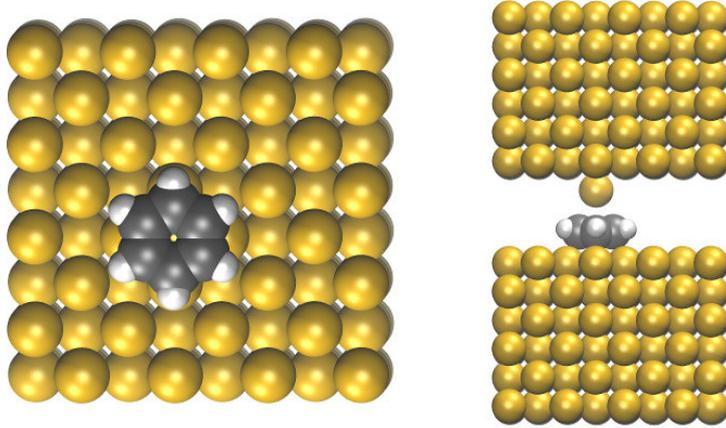


Figure 5: Structure of benzene molecule adsorbed on a Au(001) surface and a gold tip. Top and side view.

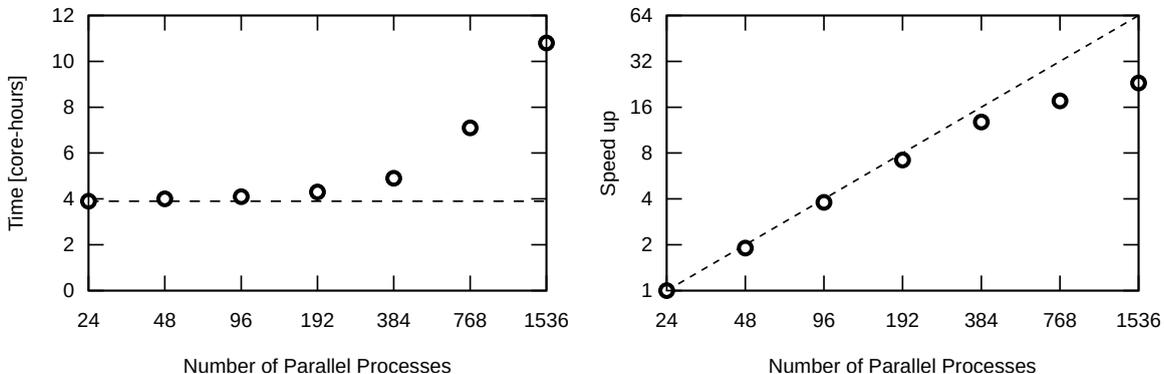


Figure 6: Computational performance of NEGF calculation with respect to the number of MPI processes. The left plot shows the cumulative computational time spent to perform a single NEGF iteration of a benzene molecule adsorbed on a Au(001) surface and the gold tip (397 atoms, 3603 spherical basis functions). The right plot visualises the performance gain achieved by using parallelisation over energy points. Each parallel group contains 24 MPI processes. Dash line corresponds to ideal scaling.

literature [4, 12]. In particular, the transmission coefficient has a recognisable shape with two broad peaks located approximately at -1 eV and 2.5 eV relatively the Fermi level. The (I-V) characteristic curve is antisymmetric about origin and it has three regions on the interval between 0 and 5 V. When the applied bias is rather small (up to 1.7 Å) the current increases at rate of $40\mu\text{A}/\text{V}$. In the interval between 1.7 Å and 3 Å the current flow becomes constant and equal to $61\mu\text{A}$, and starts to grows at average rate of $20\mu\text{A}/\text{V}$ with further increase in applied voltage.

NEGF method heavily relies on matrix inversion, which scales as $O(N^3)$ with the number of basis functions in the scattering region. Because of this, even in the case of a moderately-sized system such calculations need to be run in a parallel environment. We demonstrate scalability of the code in case when the number of basis functions remain constant by using a benzene molecule adsorbed on the hollow site of the Au(001) surface with a gold tip with a tetrahedral asperity placed 3.0 Å above the centre of the molecule. As previously, the calculation was carried out by using LDA exchange-correlation functional and SZV-MOLOPT-SR-GTH and TZVP-MOLOPT-GTH basis sets for gold and all other atoms respectively. The contacts' unit cells were composed of 4 gold planes with 32 gold atoms per plane. We found that the size of such unit cell is large enough to satisfy the nearest-neighbour interaction: all matrix elements between the contact's cell and the next but one cell do not exceed 10^{-15} which is the value of the EPS_DEFAULT threshold. The scattering region was composed of

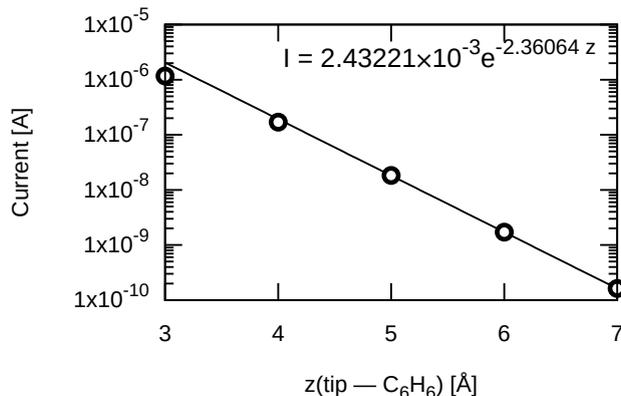


Figure 7: Tunnelling current as a function of distance between a benzene molecule adsorbed on a Au(001) surface and a gold tip. The applied bias is 0.1 V.

397 atoms. Among them are 12 atoms belonging to the benzene molecule, 192 gold atom from each electrode and an extra gold atom which acts as the tip. The optimal molecular structure is shown in Figure 5.

A single NEGF iteration with no applied bias was then performed using different number of MPI processes. All processes were organised in parallel groups with 24 processes in each group. Computational time spent by all the processes in order to calculate every element of the density matrix with 10^{-6} accuracy is shown in the left plot of Figure 6 while the performance gain is visualised in the right plot. It can be seen that the cumulative computational time remains constant in case of the moderate number of parallel groups (from 1 to 8). In contrast, when the number of groups becomes large (≥ 32) load-balance issues come into play as the parallelisation over energy approach requires us to distribute energy points evenly across the groups. However, this is not always possible due to the fact that we rely on an adaptive integrator which helps to evaluate Green's functions only at points which contribute most to the density matrix. Thus, at some refinement iteration the number of new energy points can become smaller than the actual number of parallel groups available, so some processors may become idle. It has been found that the ratio of four energy points per parallel group gives the optimal parallel performance.

As the final test case, we computed the current as a function of distance between a benzene molecule adsorbed on a Au(001) surface and a gold tip under the constant bias of 0.1 V. The obtained results are plotted in Figure 7. It is well known from the scanning tunnelling microscopy theory that the current decreases exponentially with the distance between the tip and the sample surface:

$$I \sim \exp(-2\kappa z), \quad (30)$$

where κ is the decay constant. By using this relation we can estimate the decay constant which is found to be equal to 1.18 \AA^{-1} . This is in perfect agreement with experimental decay constant for gold atoms which is equal to 1.19 \AA^{-1} [13].

5 Conclusions

We demonstrate that our implementation of Non-Equilibrium Green's Function method within the CP2K package is working correctly and shows good efficiency. It opens possibility to investigate a number of research and application areas – from quantum transport properties of nano-devices containing thousands of atoms till *ab-initio* scanning tunnelling microscopy images – by solely using CP2K without relying on external tools.

Acknowledgments

This work was funded under the embedded CSE programme of the ARCHER UK National Supercomputing Service (<http://www.archer.ac.uk>). We also gratefully thank the UK's HEC Materials Chemistry Consortium, which is funded by EPSRC (EP/L000202), for computational resources on the ARCHER UK National Supercomputing Service.

References

- [1] *CP2K developers home page*, <http://cp2k.org/>.
- [2] M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, *Phys. Rev. B* **65**, 165401 (2002).
- [3] A. R. Rocha et al., *Phys. Rev. B* **73**, 085414 (2006).
- [4] C. J. O. Verzijl and T. J. M., *J. Phys. Chem. C* **116**, 24393 (2012).
- [5] M. Luisier, A. Schenk, W. Fichtner, and G. Klimeck, *Phys. Rev. B* **74**, 205323 (2006).
- [6] M. P. López Sancho, J. M. López Sancho, and J. Rubio, *J. Phys. F: Met. Phys.* **15**, 851 (1985).
- [7] N. Papior, N. Lorente, T. Frederiksen, A. García, and M. Brandbyge, *Computer Physics Communications* **212**, 8 (2017).
- [8] S.-C. Chang, *J. Comput. Phys.* **67**, 91 (1986).
- [9] M. H. Bani-Hashemian, S. Brüch, M. Luisier, and J. VandeVondele, *J. Chem. Phys.* **144**, 044113 (2016).
- [10] S. Goedecker, M. Teter, and J. Hutter, *Phys. Rev. B* **54**, 1703 (1996).
- [11] J. VandeVondele and J. Hutter, *J. Chem. Phys.* **127**, 114105 (2007).
- [12] K. Stokbro, J. Taylor, M. Brandbyge, J.-L. Mozos, and P. Ordejon, *Comput. Materials Science* **27**, 151 (2003).
- [13] C. J. Chen, *Introduction to Scanning Tunneling Microscopy*, Oxford University Press, 2016.