Adaptive Energy Integration of Non-Equilibrium Green's Functions

O. Baumgartner, M. Karner, S. Holzer, M. Pourfath, T. Grasser, and H. Kosina

Institute for Microelectronics, TU Vienna,
Gusshausstrasse 27–29/E360, 1040 Wien, Austria
Baumgartner@iue.tuwien.ac.at

ABSTRACT

To obtain the physical quantities of interest within the non-equilibrium Green's function formalism, numerical integration over energy space is essential. Several adaptive methods have been implemented and tested for their applicability. The number of energy grid points needed and the convergence behavior of the Schrödinger-Poisson iteration have been evaluated. An adaptive algorithm based on a global error criterion proved to be more efficient than a local adaptive algorithm.

Keywords: NEGF, numerical integration, adaptive grid refinement

1   INTRODUCTION

For nanoscaled devices, numerical simulations based on the non-equilibrium Green's functions (NEGF) formalism are commonly performed [1]–[4]. A very efficient implementation of this method has been achieved by means of a recursive algorithm [5]. Proper numerical integration methods are vital for the stability and accuracy of NEGF simulations.

2   THE NEGF FORMALISM

The retarded and advanced Green's functions are determined by the equation

$$G^{\mathrm{R}}(\mathbf{r},\mathbf{r}',\mathcal{E}) = G^{\mathrm{A}\dagger}(\mathbf{r},\mathbf{r}',\mathcal{E})$$
$$= [\mathcal{E}I - H(\mathbf{r},\mathbf{r}',\mathcal{E}) - \Sigma^{\mathrm{R}}(\mathbf{r},\mathbf{r}',\mathcal{E})]^{-1},$$

where $H(\mathbf{r},\mathbf{r}',\mathcal{E})$ is the Hamiltonian of the system and $\Sigma^{\mathrm{R}}(\mathbf{r},\mathbf{r}',\mathcal{E})$ is the retarded self-energy. The less-than Green's function is calculated as

$$G^{<}(\mathbf{r},\mathbf{r}',\mathcal{E}) = G^{\mathrm{R}}(\mathbf{r},\mathbf{r}',\mathcal{E})\Sigma^{<}(\mathbf{r},\mathbf{r}',\mathcal{E})G^{\mathrm{A}}(\mathbf{r},\mathbf{r}',\mathcal{E}).$$

The lesser self energy of the left and right contact is given by $\Sigma^{<}_{\mathrm{l,r}}(\mathcal{E}) = \imath\Im\left\{\Sigma^{\mathrm{R}}_{\mathrm{l,r}}(\mathcal{E})\right\}f_{\mathrm{l,r}}(\mathcal{E})$ with the occupation function $f_{\mathrm{l,r}}(\mathcal{E})$ of the left and right lead, respectively. The Green's functions allow the calculation of physical quantities of interest such as the local density of states, $N(\mathbf{r},\mathbf{r},\mathcal{E}) = -\frac{1}{\pi}\Im\left\{G^{\mathrm{R}}(\mathbf{r},\mathbf{r},\mathcal{E})\right\}$, and the electron and current density

$$n(\mathbf{r}) = -2\imath \int G^{<}(\mathbf{r},\mathbf{r},\mathcal{E})\frac{d\mathcal{E}}{2\pi},$$
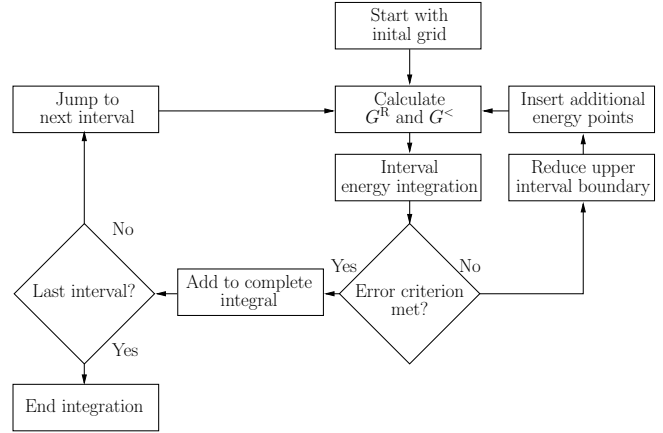

Figure 1: Illustration of the adaptive integration algorithm with grid refinement. Grid points are added as long as the error criterion is not met.

$$j(\mathbf{r}) = -\frac{\hbar q}{m^*} \int \left[(\nabla - \nabla')G^{<}(\mathbf{r},\mathbf{r}',\mathcal{E})\right]\Big|_{\mathbf{r}'=\mathbf{r}} \frac{d\mathcal{E}}{2\pi}.$$

For these quantities integration over energy is required.

3   INTEGRATION METHODS

The numerical evaluation of these quantities require a discretization of the energy space. A simple approach using an equidistant energy grid suffers from two problems. A small number of grid points will not correctly resolve narrow resonances, whereas a vast number can lead to an unpredictable summation of numerical errors and to intractable memory requirements. These effects can yield instability or poor convergency of a self-consistent iteration loop [6]. Therefore adaptive energy integration (AEI) on a non-equidistant grid is required to increase accuracy, numerical stability, and memory efficiency. The following section outlines the different approaches that were implemented and tested for the applicability within the NEGF formalism.

3.1   Simpson's Rule

Simpson's rule is a closed Newton-Cotes rule of second order. The integral of a function $f(x)$ over an in-

terval $[a, b]$ is given by

$$I_1 = \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right].$$

One strategy to decrease the interpolation error is to subdivide the interval into two equal parts and to apply Simpson's rule on each subinterval. This leads to the composite Simpson rule which, for five grid points, writes as

$$I_2 = \frac{b-a}{12}\left[f(a) + 4f\left(\frac{a+b}{4}\right) + 2f\left(\frac{a+b}{2}\right)\right.$$
$$\left. + 4f\left(3\frac{a+b}{4}\right) + f(b)\right].$$

To obtain an error criterion for the adaptive integration algorithm (Fig. 1), the electron concentration within the current integration interval is calculated using the 3-point and the 5-point composite Simpson rule. This leads to a local error which is compared to the desired error tolerance τ

$$\left|\frac{I_1 - I_2}{I_2}\right| < \tau.$$

If this condition is satisfied, the integral is considered accurate enough and the grid in the given interval is not further refined.

3.2 Polynomial Interpolation

Simpson's rule is based on equidistant grid points and an interpolation polynomial of second order. As an alternative, a more general approach with non-equidistant grid points and polynomials of arbitrary degree can be considered. For a monomial power basis the interpolation polynomial on N nodes takes the form

$$p(x) = \sum_{i=1}^{N} a_i x^{i-1}.$$

To obtain the coefficient vector $\mathbf{a} = [a_1, a_2, a_3, \ldots, a_N]^T$ an equation system of rank N needs to be solved

$$\underbrace{\begin{bmatrix} 1 & x_1 & \cdots & x_1^{N-1} \\ 1 & x_2 & \cdots & x_2^{N-1} \\ \vdots & & \ddots & \vdots \\ 1 & x_N & \cdots & x_N^{N-1} \end{bmatrix}}_{\underline{\mathbf{V}}} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix},$$

where $\underline{\mathbf{V}}$ is called the Vandermonde matrix. Unfortunately this system is often ill-conditioned and its solution may become numerically unstable. Björck and Pereyra [7] developed an algorithm that is able to calculate the coefficient vector $\mathbf{a}$ in a fast and stable manner.

After the coefficients of the polynomial are obtained the integral of the interpolation function in the interval $[x_1, x_N]$ can be calculated. For an arbitrary odd number N of grid points, a subset of $(N+1)/2$ grid points may be used to obtain a second polynomial and consequentially a second approximation of the integral. These two results are then compared to yield the error criterion for the adaptive integration algorithm. Unfortunately, polynomial interpolation functions on equidistant points suffer from Runge's phenomenon for a higher degree. This can be avoided by using non-equidistant grid points as done by the Clenshaw-Curtis Rule described in the succeeding section.

3.3 Clenshaw-Curtis Integration

Fejér [9] proposed to use the zeros of the Chebyshev polynomial $T_n = \cos(n\arccos x)$ in the interval $]-1, 1[$ as quadrature points of the integral of $f(x)$,

$$\int_{-1}^{1} f(x)\mathrm{d}x = \sum_{k=0}^{n} w_k f(x_k). \tag{1}$$

For Fejér's second rule, the $n-1$ extreme points of T_n are used. Clenshaw and Curtis [10] extended this open rule to a closed form which includes the boundary points $x_0 = -1$ and $x_n = 1$ of the interval. The $n+1$ quadrature points are

$$x_k := \cos(\vartheta_k), \quad \vartheta_k := k\frac{\pi}{n}, \quad k = 0, 1, \ldots, n.$$

The weights w_k in equation (1) are to be obtained by an explicit expression or by means of discrete Fourier transforms [8]. The explicit expressions of the Clenshaw-Curtis weights are:

$$w_k = \frac{c_k}{n}\left(1 - \sum_{j=1}^{\lfloor n/2 \rfloor} \frac{b_j}{4j^2 - 1}\cos(2j\vartheta_k)\right).$$

The coefficients b_j and c_k are given by

$$b_j = \begin{cases} 1, & \text{if } j = n/2 \\ 2, & \text{if } j < n/2 \end{cases},$$

$$c_k = \begin{cases} 1, & \text{if } k = 0 \mod n \\ 2, & \text{otherwise.} \end{cases}$$

A useful property of the Clenshaw-Curtis rule is the option to create subsets of the quadrature nodes. To move from $n+1$ to $2n+1$ points only n new function values need to be evaluated.

3.4 Extended Doubly Adaptive Quadrature Routine

So far the presented methods used a local error criterion for adaptive energy integration. A different approach, which comprises two refinement strategies, has
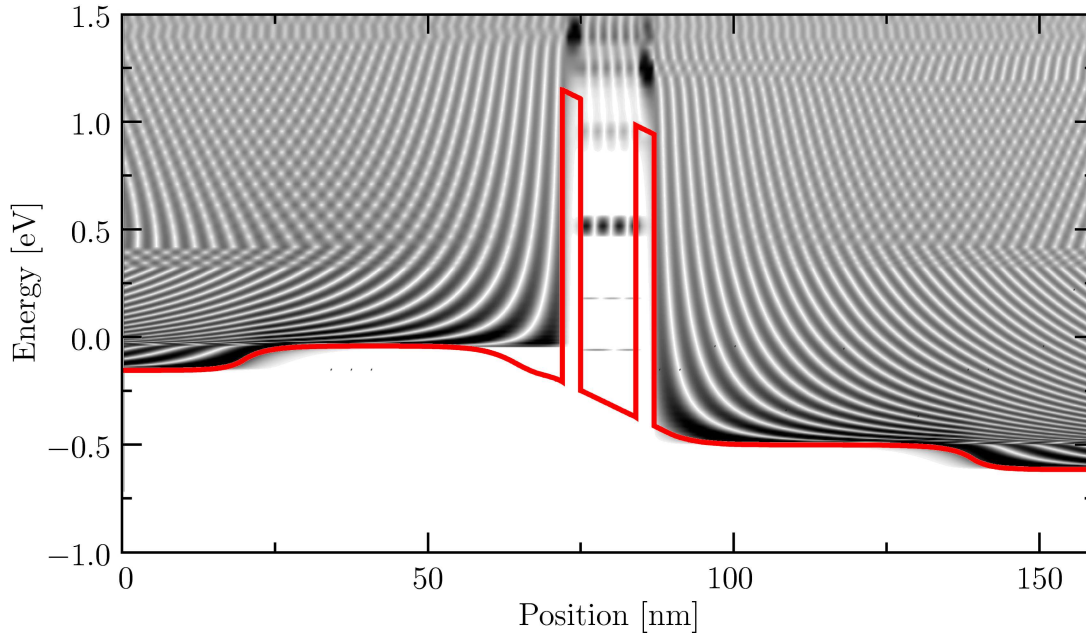
Figure 2: Local density of states of the resonant tunneling diode. The penetration into classically forbidden regions, the reflections at the barriers, as well as the resonant tunneling states within the quantum well can be seen clearly.

been presented by Espelid [11]. A global error criterion is used to find the most erroneous subinterval. This interval is then treated locally either by subdivision and applying a smaller order Newton-Cotes rule, or by inserting additional energy grid points and using a higher order rule, depending on the estimated error. The local integral and error of the superior method for a given subinterval are then added to the global values. This procedure is repeated until the global error is below a given tolerance as depicted in Fig. 3.

4   RESULTS

The implemented energy integration algorithms were applied to an unbiased double barrier structure. The number of energy grid points needed to meet a given relative error τ is plotted in Figure 4. For the polynomial interpolation and the Clenshaw-Curtis method a relative error $\tau < 10^{-4}$ was needed to correctly resolve the resonance in the quantum well. Using Simpson's rule even $\tau < 10^{-5}$ is required to set enough grid points. Comparing the local adaptive procedures, the polynomial interpolation performs best considering the number of grid points.

To evaluate the convergence behavior of a self-consistent band edge calculation as seen in Figure 2, the square of the potential update norm is plotted over the iteration number (Fig. 5). For the polynomial and the Clenshaw-Curtis method a local error criterion of $\tau = 10^{-6}$ has been chosen. For the global adaptive algorithm the relative error has been set to $\tau = 10^{-3}$ and
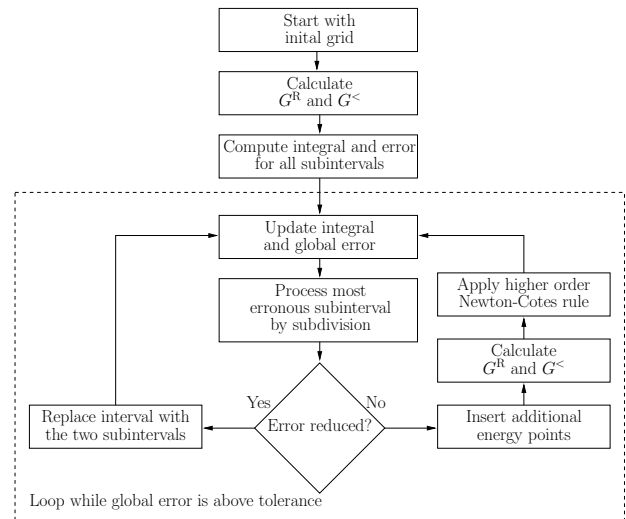

Figure 3: Illustration of the extended doubly adaptive quadrature routine.

$\tau = 10^{-5}$, respectively. All methods show similar good convergence, whereas the number of energy grid points differs considerably. The global adaptive method requires about half of the points of the polynomial interpolation and a third of the Clenshaw-Curtis integration.

Figures 6 and 7 show the distribution of the grid points versus energy for the self-consistent calculation of the bandedge of a resonant tunneling diode under bias. The histograms give the number of grid points in energy intervals of 1 meV width. For both applied
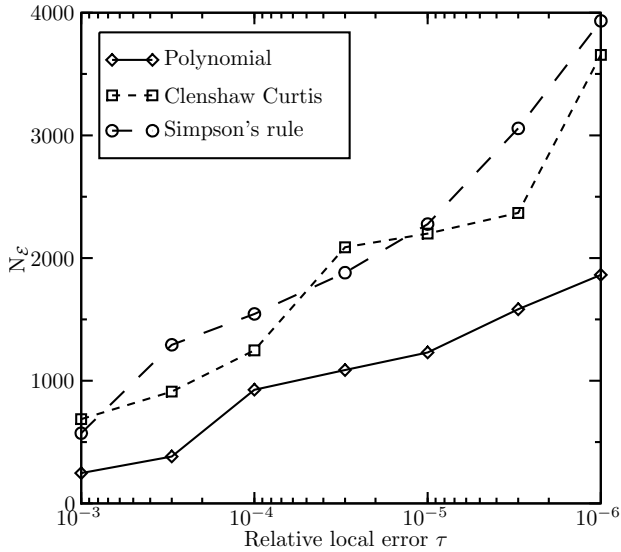
Figure 4: Number of energy grid points needed for a given relative error tolerance.

methods, the significant energies at resonant levels or the contact chemical potentials can be distinguished. At these energies, many more grid points are placed by the algorithms. The global adaptive procedure requires approximately half of the grid points to properly resolve a resonance as compared to the polynomial interpolation.

5 CONCLUSION

Local as well as a global adaptive integration strategies have been used in NEGF simulations. The Simpson rule does not suffice the demands of the diverse energy spectrum of a nano-electronic device. Although the polynomial interpolation and the Clenshaw-Curtis method combined with a local error criterion prove suitable for the numerical energy integration, a global adaptive approach is superior due to less grid points.

REFERENCES

[1] R. Lake et al., J. Appl. Phys. 81, 7845, 1997.
[2] R. C. Bowen et al., J. Appl. Phys. 81, 3207, 1997.
[3] A. Cresti et al., Phys. Rev. B 68, 075306, 2003.
[4] M. Pourfath et al., J. Comp. Electron. 5, 155, 2006.
[5] A. Svizhenko et al., J. Appl. Phys. 91, 2343, 2002.
[6] D. John et al., in Nanotech, 3, 65–68, 2004.
[7] A. Björck and V. Pereyra, Math. Comp. 24, 893, 1970.
[8] J. Waldvogel, BIT 46, 195, 2006.
[9] L. Fejér, Math. Z. 37, 287-309, 1933.
[10] C. W. Clenshaw and A. R. Curtis, Num. Math. 2, 197-205, 1960.
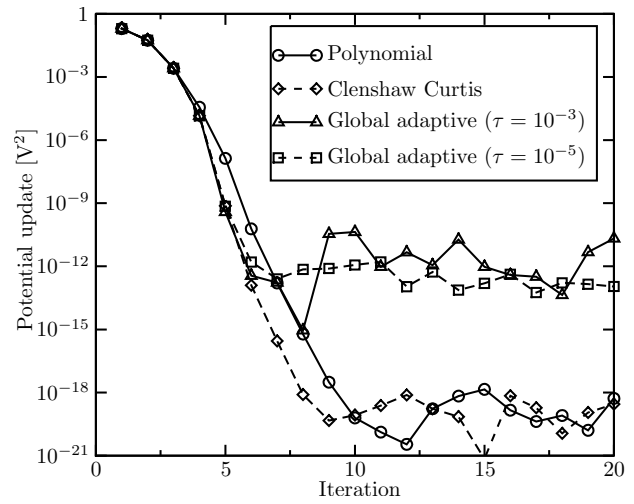[11] T. O. Espelid, BIT 43, 319-337, 2003.

Figure 5: The evolution of the squared potential update norm is given for a self-consistent bandedge calculation of a double barrier structure.
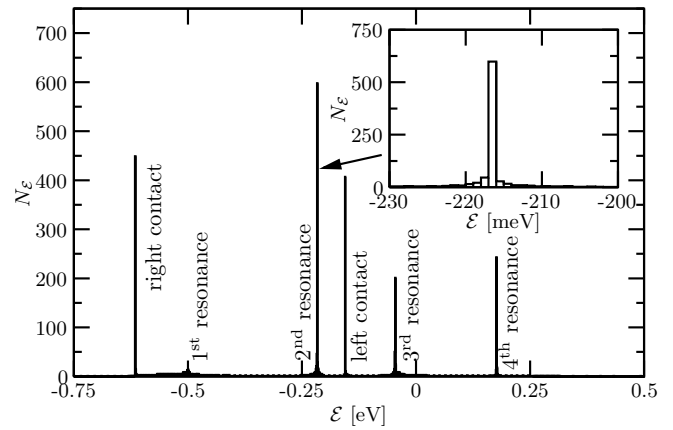

Figure 6: Histogram of energy grid points for the polynomial interpolation.
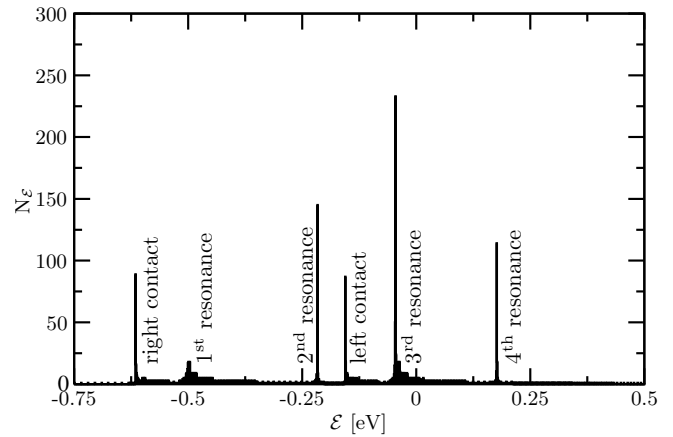

Figure 7: Histogram of energy grid points for the doubly global adaptive integration.