# On the Validity of Momentum Relaxation Time in Low-Dimensional Carrier Gases

Z. Stanojević*, O. Baumgartner*, M. Karner†, L[a]. Filipović*, C. Kernstock†, and H. Kosina*
*Institute for Microelectronics, TU Wien, Gußhausstraße 27-29/E360, 1040 Vienna, Austria
†Global TCAD Solutions GmbH., Landhausgasse 4/1a, 1010 Vienna, Austria
Email: {stanojevic|baumgartner|lidijafilipovic|kosina}@iue.tuwien.ac.at, {m.karner|c.kernstock}@globaltcad.com

***Abstract*—The momentum relaxation time (MRT) is widely used to simplify low-field mobility calculations including anisotropic scattering processes. Although not always fully justified, it has been very practical in simulating transport in bulk and in low-dimensional carrier gases alike. We review the assumptions behind the MRT, quantify the error introduced by its usage for low-dimensional carrier gases, and point out its weakness in accounting for inter-subband interaction, occurring specifically at low inversion densities.**

## I. Introduction

Semi-classical carrier transport at low driving fields is described by the linearized Boltzmann equation (LBTE), which, excluding inelastic scattering processes, reads

$$\sum_{n',\mathbf{k}'} S_{n,n'}(\mathbf{k},\mathbf{k}')\left[f_n^1(\mathbf{k})-f_{n'}^1(\mathbf{k}')\right] = -\mathbf{F}\cdot\mathbf{v}_n(\mathbf{k})\frac{\mathrm{d}f^0}{\mathrm{d}E}. \quad (1)$$

The carrier distribution response $f^1$ can be expressed in terms of a *microscopic relaxation time tensor* $\underset{\sim}{\tau}_n(\mathbf{k})$ [1], which results in an equivalent representation of Eq. (1),

$$\sum_{n',\mathbf{k}'} S_{n,n'}(\mathbf{k},\mathbf{k}')\left[\underset{\sim}{\tau}_n(\mathbf{k})\cdot\mathbf{v}_n(\mathbf{k}) - \underset{\sim}{\tau}_{n'}(\mathbf{k}')\cdot\mathbf{v}_{n'}(\mathbf{k}')\right] = \mathbf{v}_n(\mathbf{k}). \quad (2)$$

The exact knowledge of $\underset{\sim}{\tau}_n(\mathbf{k})$ would require a solution of above integral equation, and two approximations can be introduced to simplify the integral equation to a simple integration in $\mathbf{k}$-space.

The first approximation is to replace the tensor $\underset{\sim}{\tau}_n(\mathbf{k})$ by a scalar $\tau_n(\mathbf{k})$ and multiply Eq. (2) by $\mathbf{v}_n(\mathbf{k})$, so it can be rewritten as

$$\sum_{n',\mathbf{k}'} S_{n,n'}(\mathbf{k},\mathbf{k}')\left[\tau_n(\mathbf{k})v_n^2(\mathbf{k}) - \tau_{n'}(\mathbf{k}')\mathbf{v}_n(\mathbf{k})\cdot\mathbf{v}_{n'}(\mathbf{k}')\right] = v_n^2(\mathbf{k}). \quad (3)$$

Factoring out $v_n^2(\mathbf{k})$ and $\tau_n(\mathbf{k})$, one obtains

$$\frac{1}{\tau_n(\mathbf{k})} = \sum_{n',\mathbf{k}'} S_{n,n'}(\mathbf{k},\mathbf{k}')\left[1-\frac{\tau_{n'}(\mathbf{k}')}{\tau_n(\mathbf{k})}\frac{\mathbf{v}_n(\mathbf{k})\cdot\mathbf{v}_{n'}(\mathbf{k}')}{v_n^2(\mathbf{k})}\right], \quad (4)$$

which is the *implicit definition* [2] of the momentum relaxation time (MRT).

The second approximation makes the assumption that $\tau_{n'}(\mathbf{k}')/\tau_n(\mathbf{k}) \approx 1$, thus neglecting inter-subband coupling in the LBTE's scattering operator and giving the explicit momentum relaxation time [1, 2],

$$\frac{1}{\tau_n^{\mathrm{m}}(\mathbf{k})} = \sum_{n',\mathbf{k}'} S_{n,n'}(\mathbf{k},\mathbf{k}')\left[1-\frac{\mathbf{v}_n(\mathbf{k})\cdot\mathbf{v}_{n'}(\mathbf{k}')}{v_n^2(\mathbf{k})}\right]. \quad (5)$$

Generally, one can write the MRT as an integral of the transition rate $S_{n,n'}(\mathbf{k},\mathbf{k}')$ multiplied by a weighting factor $\Theta_{n,n';\mathbf{k},\mathbf{k}'}$,

$$\frac{1}{\tau_n^{\mathrm{m}}(\mathbf{k})} = \sum_{n',\mathbf{k}'} S_{n,n'}(\mathbf{k},\mathbf{k}')\Theta_{n,n';\mathbf{k},\mathbf{k}'}. \quad (6)$$

Other expressions for $\Theta_{n,n';\mathbf{k},\mathbf{k}'}$ than the one derived in Eq. (5),

$$\Theta_{n,n';\mathbf{k},\mathbf{k}'} = 1-\frac{\mathbf{v}(\mathbf{k})\cdot\mathbf{v}(\mathbf{k}')}{v^2(\mathbf{k})}, \quad (7)$$

can be found throughout literature. Reference [3] for instance suggests

$$\Theta_{n,n';\mathbf{k},\mathbf{k}'} = 1-\frac{k'}{k}\cos(\vartheta_{\mathbf{k}}-\vartheta_{\mathbf{k}'}) = 1-\frac{\mathbf{k}\cdot\mathbf{k}'}{k^2}. \quad (8)$$

A systematic overview is given in Table I. Here we distinguish two classes of $\Theta_{n,n';\mathbf{k},\mathbf{k}'}$ expressions: normalized (2 and 5) and non-normalized (3 and 4). The value of the normalized expressions is guaranteed to lie within $0 \le \Theta_{n,n';\mathbf{k},\mathbf{k}'} \le 2$, while this is not the case for the non-normalized ones.

TABLE I. Systematic overview of the different expressions for $\Theta_{n,n';\mathbf{k},\mathbf{k}'}$ in the momentum relaxation time formula; the numbering scheme was chosen to be consistent with [4].

| direction quantity | $\mathbf{k}$ | $\mathbf{v}_n(\mathbf{k})$ |
|---|---|---|
| normalized | [2] $1-\frac{\mathbf{k}\cdot\mathbf{k}'}{kk'}$ | [5] $1-\frac{\mathbf{v}_n(\mathbf{k})\cdot\mathbf{v}_{n'}(\mathbf{k}')}{v_n(\mathbf{k})v_{n'}(\mathbf{k}')}$ |
| non-normalized | [3] $1-\frac{\mathbf{k}\cdot\mathbf{k}'}{k^2}$ | [4] $1-\frac{\mathbf{v}_n(\mathbf{k})\cdot\mathbf{v}_{n'}(\mathbf{k}')}{v_n^2(\mathbf{k})}$ |

The usage of the non-normalized expressions for confined systems can be problematic as they can result in negative momentum relaxation times. This unphysical behavior typically affects Coulomb scattering, for which $S_{n,n'}(\mathbf{k},\mathbf{k}')$ rapidly decreases with $\|\mathbf{k}-\mathbf{k}'\|$. For $n \neq n'$, $\mathbf{k}$ and $\mathbf{k}'$ may well be parallel but have different length; especially, $k'$ might be larger than $k$, in which case $\Theta_{n,n';\mathbf{k},\mathbf{k}'} = 1-\mathbf{k}\cdot\mathbf{k}'/k^2$ becomes negative. Combined with the aforementioned property of the function $S_{n,n'}(\mathbf{k},\mathbf{k}')$, this renders the non-normalized expressions for

978-1-4799-5288-5/14/$31.00 

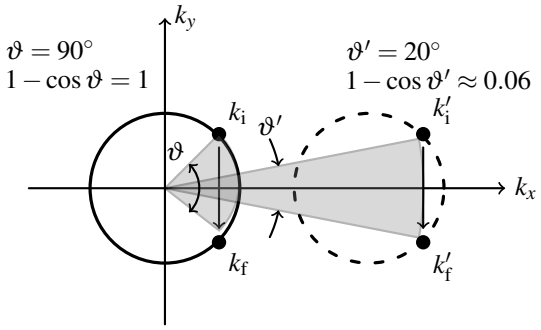


Fig. 1. Assume a valley centered around $\mathbf{k} = \mathbf{0}$; a transition from initial state $k_\mathrm{i}$ to the final one $k_\mathrm{f}$ would be weighted with a $1-\cos\vartheta = 1$ to obtain the momentum relaxation time. If we choose a different reference point in $\mathbf{k}$-space, effectively shifting the valley to a position other than $\mathbf{k} = \mathbf{0}$, the momentum relaxation time factor changes dramatically. Physically, however, mobility should not depend in the reference $\mathbf{k}$-point.

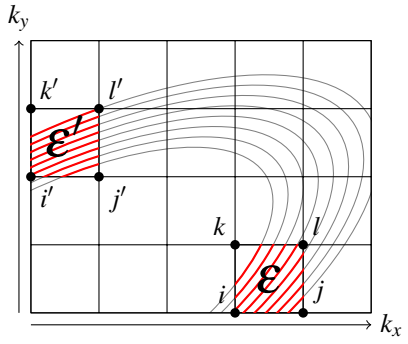


Fig. 2. Energy contours that pass trough $\mathbf{k}$-grid elements $\varepsilon$ and $\varepsilon'$ couple the elements' vertices, $i$, $j$, $k$, $l$, $i'$, $j'$, $k'$, and $l'$.

$\Theta_{n,n';\mathbf{k},\mathbf{k}'}$ extremely sensitive to the accuracy of the numerical quadrature used to compute Eq. (5).

The usability of expressions 2 and 3 in Table I, which are based on initial and final wave vectors rather than group velocities, are also questionable. Let us consider band structures with multiple degenerate valleys, which is typical of conduction bands of indirect semiconductors, where one example would be the two-band $\mathbf{k}\cdot\mathbf{p}$ model for electrons in Si. Here, two valleys lie on either side of the X-point and the choice of the reference $\mathbf{k}$-point greatly influences the resulting MRT, as Fig. 1 illustrates.

Ruling out non-normalized and $\mathbf{k}$-based expressions for $\Theta_{n,n';\mathbf{k},\mathbf{k}'}$ leaves only

$$\Theta_{n,n';\mathbf{k},\mathbf{k}'} = 1 - \frac{\mathbf{v}_n(\mathbf{k})\cdot\mathbf{v}_{n'}(\mathbf{k}')}{v_n(\mathbf{k})v_{n'}(\mathbf{k}')}. \tag{9}$$

This expression can also be derived from Eq. (4) assuming $\tau_{n'}(\mathbf{k}')\mathbf{v}_{n'}(\mathbf{k}')/\tau_n(\mathbf{k})\mathbf{v}_n(\mathbf{k}) \approx 1$ rather than $\tau_{n'}(\mathbf{k}')/\tau_n(\mathbf{k}) \approx 1$, but has not been used in any of the publications known to us. Although it appears to be a reliable choice it still neglects the anisotropy of the microscopic relaxation time as well as inter-subband coupling.

## II. Numerical Solution of the LBTE

To be able to assess the error introduced by using the MRT we need a way to obtain macroscopic quantities without the use of MRT. To this end we need to go back to the Boltzmann transport equation, Eq. (1). This time we will not introduce a relaxation at all but rather solve for the carrier distribution response $f^1$ directly.

It should be noted that Eq. (1) is linear in both carrier response $f^1$ and driving field $\mathbf{F}$. This allows us to factor out the driving field modulus, leaving the reduced LBTE

$$\sum_{n',\mathbf{k}'} S_{n,n'}(\mathbf{k},\mathbf{k}')\left[\tilde{f}^1_n(\mathbf{k}) - \tilde{f}^1_{n'}(\mathbf{k}')\right] = -\mathbf{e}_\mathbf{F}\cdot\mathbf{v}_n(\mathbf{k})\frac{\mathrm{d}f^0}{\mathrm{d}E}, \tag{10}$$

where $\mathbf{e}_\mathbf{F}$ is a unit vector pointing in the direction of the driving field and $\tilde{f}^1 = f^1/F$ is the reduced distribution response.

Before Eq. (10) can be solved it needs to be discretized. This can be done using the same $\mathbf{k}$-space grid that is employed to obtain the subband structure in the first place, e.g. by diagonalizing a $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian. The $\mathbf{k}$-grid imposes the concept of discrete $\mathbf{k}$-cells that are coupled by probability fluxes due to scattering. The general form of the discretized equation will thus be

$$\sum_{\nu'} \hat{S}_{\nu,\nu'} w_{\nu,\nu'}\left[\tilde{f}^1_\nu - \tilde{f}^1_{\nu'}\right] = -\mathbf{e}_\mathbf{F}\cdot\mathbf{v}_\nu\frac{\mathrm{d}f^0}{\mathrm{d}E}V_\mathbf{k}, \tag{11}$$

where $\nu = (n,\mathbf{k})$ is a global index that denotes the index of each $\mathbf{k}$-grid cell on in each subband and $V_\mathbf{k}$ is the volume, area, or length of a $\mathbf{k}$-grid cell depending on the dimensionality of the carrier gas. The discrete states are coupled through $\hat{S}_{\nu,\nu'} w_{\nu,\nu'}$, where

$$\hat{S}_{\nu,\nu'} = \frac{2\pi}{\hbar}\langle|H_{n,n';\mathbf{k},\mathbf{k}'}|^2\rangle \tag{12}$$

is the transition rate due to Fermi's golden rule without the energy-conserving $\delta(E-E')$ and $w_{\nu,\nu'}$ are weights that arise from the discretization of the scattering operator.

To compute the coupling weights $w_{\nu,\nu'}$ we need to consider the total probability flux from one cell $\nu$ to another $\nu'$. When an equi-energy contour passes through two elements on the $\mathbf{k}$-grid it couples each element's vertices, as shown in Fig. 2. A numerical integration $\int \mathrm{d}E$ is performed, where the contribution of each contour is accumulated

$$\mathrm{d}w_{\nu,\nu'} = g_\nu(E)g_{\nu'}(E)\mathrm{d}E, \tag{13}$$

with $g_\nu(E) = 1/\hbar\|\mathbf{v}_\nu\|$ being the local density of states. Fig. 3 shows such an integration for a one-dimensional subband structure.

Energy conservation in the scattering operator makes the resulting $w_{\nu,\nu'}$ sparse. Ordering the discrete states by their absolute energy will result in a dense skyline-type arrangement of the non-zero elements in the matrix as shown in Fig. 4. This is important since the number of non-zero elements in a realistic device can easily reach tens of millions, and dense storage has practically no memory overhead and results in fast access times, since the elements don't have to be *searched for* in maps.

The sparsity of $w_{\nu,\nu'}$ also means that $\hat{S}_{\nu,\nu'}$ only needs to be computed for the non-zero elements. This is crucial since the computation of the transition rates in confined systems is

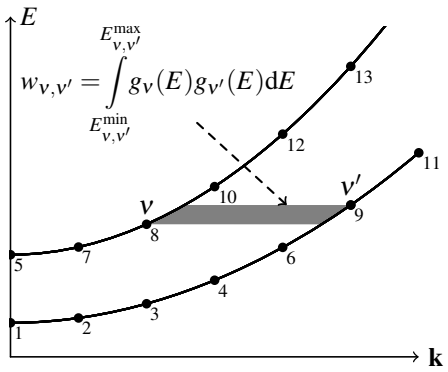


Fig. 3. Calculation of the coupling weights for an elastic scattering operator; $w_{\nu,\nu'}$ is obtained by integrating the product of the density of states of state $\nu$ and $\nu'$ over the energy interval where $\nu$ and $\nu'$ overlap. Multiplied by a transition rate it gives the probability flux between $\nu$ and $\nu'$.

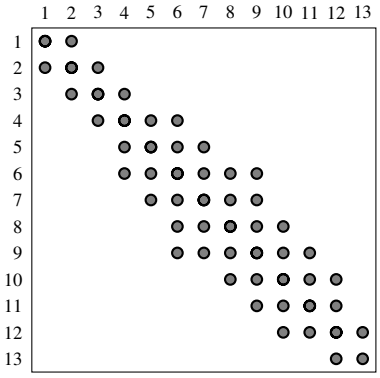

Fig. 4. The resulting non-zero pattern of the discretized scattering operator for the example subband structure from Fig. 3; sorting all states by absolute energy produces dense symmetric skyline matrix, thus eliminating storage overhead.

by far the most time demanding task and the effort must be kept to a minimum.

Having discretized Eq. (10) and computed the rates from Eq. (10), we can now solve the system using a linear solver. Due to system size, sparsity positive semi-definiteness, an iterative solver can be used to efficiently calculate the $\tilde{f}^1$ solution. The effort to obtain the solution from an assembled matrix is negligible compared to the effort needed to calculate the transition rates. This makes our approach computationally competitive with MRT, since the rates need to be computed regardless whether one uses the MRT or the $f^1$ approach.

## III. Results

We will now quantitatively compare the MRT approximation with the $f^1$ approach. To do so we will compute the mobility of planar Si MOS channels and nanowires using both approaches and compare the results. Both n and p-type channels were investigated with the channel doping set to $N_\mathrm{A} = 3\times10^{17}\,\mathrm{cm}^{-3}$ for n-type and $N_\mathrm{D} = 3\times10^{17}\,\mathrm{cm}^{-3}$ for p-type channels. Different crystal orientations were used in order to thoroughly test MRT against the $f^1$ approach. For the MOS the substrate orientations $\{100\}$, $\{110\}$, $\{111\}$, and $\{112\}$ were used while keeping the transport direction at $\langle110\rangle$.

In the first step, the subband structure of the devices was

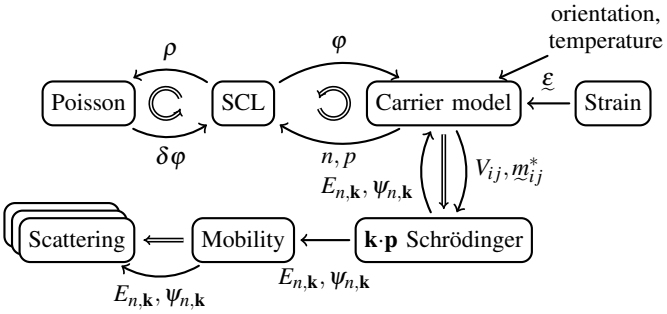


Fig. 5. A self-consistent loop (SCL) is used to obtain potential and carrier concentration. The subbands and wavefunctions from the converged results are then used in the mobility model and the scattering models to compute transition rates, probability fluxes, and finally the channel mobility.

TABLE II. Scattering model parameters used in the mobility calculations

| Parameter | Value | |
|---|---|---|
| Acoustic deformation potential (electrons) | 14.6 | eV |
| Acoustic deformation potential (holes) | 10.2 | eV |
| Roughness RMS amplitude | 5 | Å |
| Roughness autocorrelation length | 10 | Å |

obtained using a self-consistently $\mathbf{k}\cdot\mathbf{p}$-Schrödinger-Poisson setup shown in Fig. 5. A two-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian [5] was used for modeling the conduction band and a six-band Hamiltonian [6, 7] including spin-orbit coupling for the valence band.

In the second step, the scattering operators were assembled and transition rates computed as shown in the previous section. The scattering processes involved were acoustic phonon scattering, ionized-impurity scattering and surface-roughness scattering [8, 9] with an exponential surface roughness autocorrelation function. The parameters used in the scattering models are summarized in Table II.

In the third and final step, $\tilde{f}^1$ is either computed by solving Eq. (11), or approximated using one of the MRT formulas. Subband conductivity is obtained from

$$\sigma_n = -\frac{q_0}{(2\pi)^d}\int_{\mathbb{R}^d} \mathbf{v}_n(\mathbf{k})\tilde{f}^1_n(\mathbf{k})\mathrm{d}^d k, \tag{14}$$

and from it the channel mobility.

All models involved in the calculations were implemented as part of the Vienna Schrödinger-Poisson simulator [10, 11].

Fig. 6 shows the electron mobility curves of the n-type MOS channels. The $\mathbf{v}(\mathbf{k})$-based MRT models (4 and 5) approximate the $f^1$ solution very well at hight inversion densities while they systematically underestimate electron mobility at low densities. The $\mathbf{k}$-based MRT models completely fail for electrons due to the reasons discussed in Fig. 1.

In Fig. 7 the hole mobility for p-type channels is plotted. Again, MRT 4 and 5 underestimate mobility at low inversion densities while approximating the mobility curve well at high densities, except for the $\{112\}$ channel where the error persists throughout the curve. Also, numerical instabilities can be observed in the MRT 4 curve. $\mathbf{k}$-based MRT models (2 and 3) generally overestimate hole mobility. The accuracy of all four MRT models is clearly *orientation-dependent*.

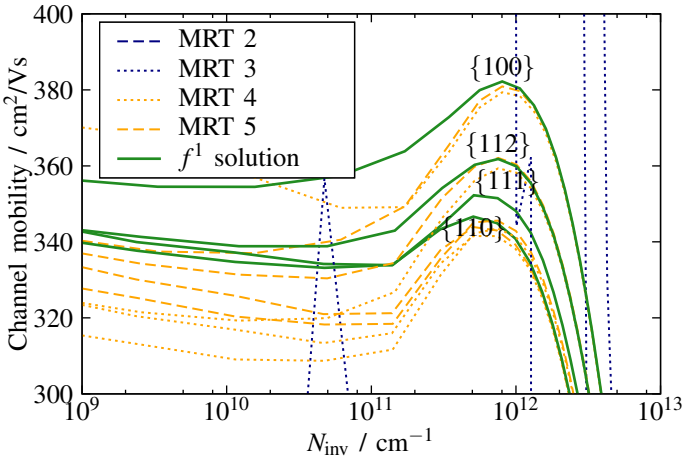


Fig. 6. Electron mobility calculated using MRT (dashed: normalized, dotted: non-normalized, blue: **k**-based, orange: **v**(**k**)-based) and solving $f^1$ (solid green) for n-type MOS channels at different substrate orientations. Transport direction if ⟨110⟩. MRT 4 and 5 deviate from $f^1$ at low inversion densities but are a good approximation at high densities. MRT 2 overestimates mobility by a factor of ten (not shown) while MRT 3 diverges.

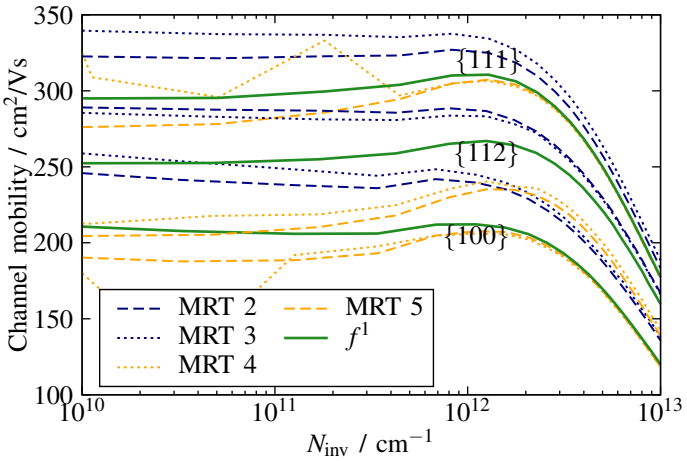


Fig. 7. Same as Fig. 6 for holes in p-type MOS channels; in general, **k**-based MRT models (2 and 3) tend to overestimate hole mobility while **v**(**k**)-based ones tend to underestimate it. Clearly, the discrepancy between the MRT and $f^1$ solutions is orientation dependent.

Finally, Fig. 8 shows the hole distribution response $f^1$ in the top subband of the p-type {100} and {110} channels approximated by MRT 5 and obtained by numerical solution of Eq. (11). The numerical $f^1$ solution shows a redistribution of carriers due to inter-subband coupling in the scattering operator, an effect not seen in the MRT picture.

## IV. Conclusion

We conclude that the MRT can give remarkably accurate results regarding its simplifications, although one needs to be careful about the choice of the weighting factor $\Theta_{n,n';\mathbf{k},\mathbf{k}'}$. The normalized, group velocity-based expression in Eq. (9) gives the most accurate and stable results. However, systematic errors still occur in multi-subband systems and accuracy is orientation-dependent. Simultaneously, we developed a computationally efficient method to directly compute the carrier distribution response and, from it, the exact channel mobility,

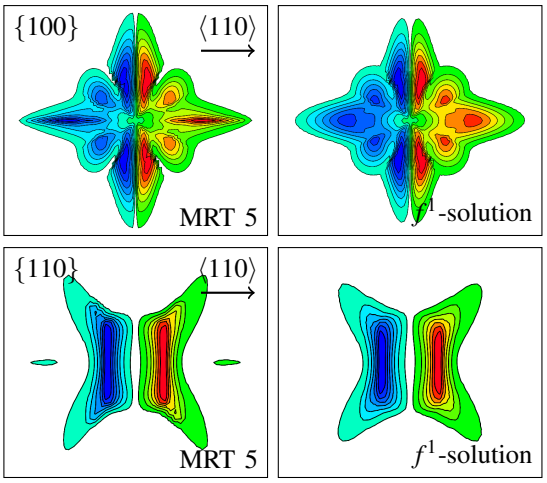


Fig. 8. **k**-space plots of the carrier distribution response for {100} and {110}-oriented p-type MOS channels in inversion; the left pictures show the response approximated by MRT 5, the right ones the $f^1$ response. The coupling between subbands causes a redistribution of carriers not seen in the MRT picture.

thus avoiding the MRT and its problems while remaining competitive in terms of computational effort.

## Acknowledgment

This work has been supported by the Austrian Science Fund through contracts F2514 and I841-N16.

## References

[1] C. Jungemann and B. Meinerzhagen, *Hierarchical device simulation: The monte-carlo perspective*. Springer, 2003.

[2] M. V. Fischetti, Z. Ren, P. M. Solomon, M. Yang, and K. Rim, "Six-band $k\cdot p$ calculation of the hole mobility in silicon inversion layers: Dependence on surface orientation, strain, and silicon thickness," *Journal of Applied Physics*, vol. 94, no. 2, pp. 1079–1095, 2003.

[3] M. V. Fischetti, "Long-range Coulomb interactions in small Si devices. Part II. Effective electron mobility in thin-oxide structures," *Journal of Applied Physics*, vol. 89, no. 2, pp. 1232–1250, 2001.

[4] *Sentaurus Device Monte Carlo User Guide, Part III: Band Structure and Mobility Calculation*, Synopsys Inc.

[5] J. C. Hensel, H. Hasegawa, and M. Nakayama, "Cyclotron Resonance in Uniaxially Stressed Silicon. II. Nature of the Covalent Bond," *Phys. Rev.*, vol. 138, no. 1A, pp. A225–A238, Apr. 1965.

[6] G. Dresselhaus, A. F. Kip, and C. Kittel, "Cyclotron Resonance of Electrons and Holes in Silicon and Germanium Crystals," *Phys. Rev.*, vol. 98, pp. 368–384, Apr 1955.

[7] T. Manku and A. Nathan, "Valence energy-band structure for strained group-IV semiconductors," *J. Appl. Phys.*, vol. 73, no. 3, pp. 1205–1213, 1993.

[8] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Rev. Mod. Phys.*, vol. 54, pp. 437–672, Apr 1982.

[9] Z. Stanojevic and H. Kosina, "Surface-Roughness-Scattering in Non-Planar Channels – the Role of Band Anisotropy," in *Intl. Conf. on Simulation of Semiconductor Processes and Devices*, 2013, pp. 352–355.

[10] O. Baumgartner, Z. Stanojevic, K. Schnass, M. Karner, and H. Kosina, "VSP–a quantum-electronic simulation framework," *J. Comput. Electron.*, 2013.

[11] "http://www.globaltcad.com/en/products/vsp.html."