Theoretical prediction of improved figure-of-merit in Si/Ge quantum dot superlattices

Gregor Fiedler and Peter Kratzer
Faculty of Physics and Center for Nanointegration (CENIDE), University Duisburg-Essen, Lotharstrasse 1, D-47048 Duisburg, Germany
E-mail: Peter.Kratzer@uni-duisburg-essen.de

New Journal of Physics 15 (2013) 125010 (15pp)
Received 29 June 2013
Published 6 December 2013
Online at http://www.njp.org/
doi:10.1088/1367-2630/15/12/125010

Abstract. A detailed theoretical model for thermoelectric transport perpendicular to the multilayers of a Si–Ge heterostructure is presented. The electronic structure of a three-dimensional superlattice, consisting of a regular array of Ge quantum dots in each layer, capped by Si layers, is calculated using an atomistic tight-binding approach. The Seebeck coefficient, the electric conductivity and the contribution of the electrons to the thermal conductivity for n-doped samples are worked out within Boltzmann transport theory. Using experimental literature data for the lattice thermal conductivity, we determine the temperature dependence of the figure of merit $ZT$. A nonlinear increase of $ZT$ with temperature is found, with $ZT > 2$ at $T = 1000$ K in highly doped samples. Moreover, we find an enhanced thermoelectric power factor already at room temperature and below, which is due to highly mobile electrons in strain-induced conductive channels.

1 Author to whom any correspondence should be addressed.
1. Introduction

In their search for novel thermoelectric materials, researchers have been guided by the paradigm of the ‘thermal glass–electron crystal’ [1], in other words, by the vision of a material displaying high electrical but low thermal conductivity. With the advent of nanotechnology, the deliberate manufacturing of nanostructured materials has opened up different routes to approach this goal. Most frequently, researchers have pursued the aim of reducing the thermal conductivity of a material by nanostructuring, exploiting the idea that phase boundaries at the nanoscale act as scattering centers for phonons. In this work, we follow the alternative route of enhancing the electrical properties (Seebeck coefficient and electrical conductivity) that has received less attention so far. One possible reason may be found in the observation that the electronic contribution to the thermal conductivity, at least in many materials, increases roughly proportionally to the electrical conductivity, thus canceling any expected gain in thermoelectric efficiency in highly conductive materials. For three-dimensional (3D) metals or semi-metals, the proportionality factor is known under the notion of the Lorenz number. In nanostructured materials, however, both the electronic density of states (DOS) and the electron scattering rates are significantly altered, and thus a strict proportionality is no longer to be expected. Therefore, nanostructured semiconductors, in addition to having low thermal conductivity, may also hold the promise of an enhanced electrical power factor. In this paper, we show by a theoretical analysis that a Si/Ge quantum dot (QD) superlattice is one such candidate system.

It has long been known that SiGe alloys have low thermal conductivity, compared to pure Si, and hence random alloys of these elements have been used to fabricate radioisotope thermoelectric generators that typically operate at high temperatures. On the other hand, alloying reduces not only the thermal, but also the electrical conductivity due to disorder scattering. As a consequence, the figure of merit $ZT$ reached in such materials did not exceed values of about 1.5, even at temperatures as high as 1000 K [2]. For further improvement, samples with inhomogeneous Si/Ge composition need to be examined. Recently, nanoparticle composites of Si and Ge and their possible use for thermoelectrics have been investigated [3–5]. Moreover, it has been demonstrated that QD superlattices that can be grown with high structural perfection allow one to tune the thermal conductivity [6, 7], and to achieve values as low as 1 W K$^{-1}$ m$^{-1}$. The present paper aims at making a theoretical prediction for the thermoelectric figure of merit of precisely the samples used in [6, 7]. Thus, we are addressing the issue of
Figure 1. Sketch of the QD superlattice. Multiple layers of QDs are grown on each other. The temperature gradient is applied perpendicular to the QD layers. On the right is a magnification of two such layers: eight monolayer (ML) high pure Ge QD’s (black) stand on a six ML Si$_{0.5}$Ge$_{0.5}$ wetting layer (gray) and are covered by 22 ML of Si. $L_x$ and $L_z$ are the dimensions of the periodic supercell and $w$ is the base distance between neighboring QD.

enhancing the figure of merit by tailoring the electrical transport properties of such samples, rather than just thermal transport properties.

In semiconductor physics, it is well known that the strain introduced in the host matrix around QDs affects its electronic structure. In this paper, we aim at exploiting electronic states that are lowered by strain effects below the conduction band edge in order to improve the carrier mobility in nanostructured Si/Ge samples even beyond the values for pure Si. We show that this gives access to a regime of lower operating temperatures for SiGe thermoelectric generators: the limitation of SiGe alloys to the use in high-temperature thermoelectric generators is a consequence of the almost linear $T$-dependence of the figure of merit when using bulk alloys (see [2, 8, 9]), which allows efficient operation only at the highest temperatures compatible with the stability of the device. Employing QD superlattices epitaxially embedded in the Si host crystal, we show theoretically that a significant enhancement of $ZT$ already at low $T$ values can be realized due to the contribution of transport channels spatially confined between the QDs.

2. General model and transport theory

In this work, we study electrical transport in a Ge QD superlattice embedded in a Si host crystal, as illustrated schematically in figure 1. A temperature gradient is applied in the direction perpendicular to the Si(100) lattice planes on which the Ge QDs are arranged. Generally speaking, the presence of a QD superlattice may affect the electric transport in several ways. One possible scenario (not studied here) involves transport through one-dimensional channels caused by charge carriers tunneling between the QDs along a QD stack [10, 11]. While this type of systems shows enhanced thermoelectric properties for a suitable choice of doping and QD geometry [11], the Si/Ge system does not lend itself to the realization of such a one-dimensional (1D) transport regime. This is because SiGe inclusions in a Si matrix form type-II QDs: the SiGe QDs act as obstacles for the electron transport in n-type material, while the hole carriers in p-type material form deeply confined states in the SiGe QDs [12, 13]. For both carrier types,
the 1D transport along a QD stack is therefore hindered by large barriers. In the strained regions of the Si matrix around the QDs, however, special low-lying states for the electrons may form. In the present paper, we focus on conductive channels appearing inside the band gap due to strain-induced energy shifts. Since we are aiming at high-mobility transport, there is another reason to restrict ourselves to electrons (rather than holes) as carriers: The electrons in Si have much lower effective mass than the holes, and hence are preferable for achieving an enhanced conductivity in strained heterostructures [14, 15].

We are interested in the electronic structure of specific samples that can be obtained experimentally in a self-assembled growth process using molecular beam epitaxy. Embedded in a Si host crystal, there are Si$_x$Ge$_{1-x}$ QDs with the shape of shallow, truncated pyramids (so-called 'huts'), stacked vertically on top of each other. The misfit strain is accommodated elastically; the samples are practically free of dislocations or other extended defects that could adversely affect electronic transport. Experimentally it is possible to stack a large number (up to 20 layers [16]) of such QDs. Since the spacings, both vertically and horizontally, between the QDs are in the range of a few nanometers up to tens of nanometers, elaborate techniques are required to calculate the electronic structure. Due to the nanoscale dimensions of the QDs, the effective-mass model of the carriers in a homogeneous semiconductor reaches its limits in such systems. Here, we use a tight-binding approach to calculate selected electronic bands. The superlattice is built up from orthorhombic supercells with dimensions $L_x$, $L_y$ and $L_z$ that contain one QD each. Periodic boundary conditions for the wavefunctions are applied in all three spatial directions. From our calculations, we obtain the band structure as a function of the crystal momentum $k_z$, along the direction of stacking, see figure 2(b).

In bulk Si, the conduction band consists of six valleys placed along the lines of $\Delta_1$ symmetry in the Brillouin zone. In material under uniaxial strain, their energetic degeneracy is partly lifted, and they split up into valleys with the remaining degeneracy factor $g_T = 4$ and $g_L = 2$. In the strained matrix of a QD superlattice, the strain-induced electronic states have the dominant character of one of these valleys, as demonstrated by effective-mass-type calculations of the electronic structure [13]. Guided by these considerations, we introduce a model for the electron transport in a Si host with embedded QD superlattice, consisting of a bulk continuum.
plus several (say, \(N\)) isolated bands that correspond to strain-induced, thread-like or sheet-like electronic wavefunctions. These model assumptions are supported by our tight-binding calculations (see below). As can be seen from the results displayed in figure 2(b), there are two such bands (with some hidden degeneracy) split off from the bottom of the conduction band for the specific system studied by us.

In order to describe the transport, we employ the semiclassical Boltzmann transport theory using an energy-dependent relaxation time for the carriers. The DOS, the carrier velocity and the relaxation rates are obtained from the electronic structure. Moreover, the chemical potential entering via the equilibrium Fermi distribution function is calculated self-consistently within the model, starting from a given concentration \(n_D\) of dopants.

Specifically, we calculate the electrical conductivity \(\sigma\), the Seebeck coefficient \(S\) and the electronic contribution to the thermal conductivity, \(\kappa_{el}\). All three quantities can be expressed by the moments, equation (1), of the ‘transport density of states’ (TDOS). This is the DOS \(g(\varepsilon)\) multiplied by the electron velocity \(v(\varepsilon)\) squared, the relaxation time \(\tau(\varepsilon)\) and the derivative of the Fermi function \(f_0^\mu(\varepsilon)\). For the model considered here, consisting of a continuum with DOS \(g_c(\varepsilon)\) and \(n = 1, \ldots, N\) bands, each with DOS \(g_n(\varepsilon)\), the moments are calculated from the integral

\[
L^{(\gamma)} = \int_{-\infty}^{\infty} d\varepsilon (\varepsilon - \mu)^\gamma \left( -\frac{\partial f_0^\mu}{\partial \varepsilon} \right) \left( g_c(\varepsilon)\tau_c(\varepsilon)v_c^2(\varepsilon) + \sum_{n=1}^{N} g_n(\varepsilon)\tau_n(\varepsilon)v_n^2(\varepsilon) \right) .
\]

Then, the Seebeck coefficient can be expressed as

\[
S = -\frac{1}{eT} \frac{L^{(1)}}{L^{(0)}}
\]

with the electron charge \(-e\) and the temperature \(T\). For the power factor \(\sigma S^2\) we also require the conductivity, given by \(\sigma = e^2 L^{(0)}\). For the figure of merit \(ZT\), we need, in addition to \(S\) and \(\sigma\), the thermal conductivity. We assume that the lattice contribution is constant in temperature and can be as low as \(\kappa_{ph} = 1\ \text{W K}^{-1}\ \text{m}^{-1}\), as observed experimentally by Pernot et al \[6\], and the electronic contribution is

\[
\kappa_{el} = \frac{1}{T} \left( L^{(2)} - \frac{(L^{(1)})^2}{L^{(0)}} \right) .
\]

Then, the figure of merit is given by

\[
ZT = \frac{\sigma S^2}{\kappa_{el} + \kappa_{ph}} T .
\]

2.1. Electronic structure

In order to obtain the parameters of our model, namely, the energy positions of the bands and their effective masses in transport direction, as well as the energy position of the continuum, we consider an atomistic model of the QD superlattice and carry out a tight-binding calculation of its electronic structure. A supercell containing one pure Ge QD, a Si\(_{0.5}\)Ge\(_{0.5}\) alloy wetting layer and a Si capping layer, with all together 103 544 atoms, is employed. The details of our atomistic modeling will be published elsewhere. In brief, we start by optimizing the height of the supercell and by relaxing the positions of the atoms within the supercell in order to minimize the elastic energy. Using a classical force field \[17\], this is achieved by the software
LAMMPS [18]. The height of the supercell is $L_z = 38.56 \text{Å}$, and the length and width are $L_x = L_y = 43a_{\text{Si}} = 233.49 \text{Å}$, with $a_{\text{Si}} = 5.43 \text{Å}$ being the silicon lattice constant. Next, we set up a tight-binding Hamiltonian using an $sp^3$ orbital basis with distance-dependent first and second-neighbor interactions [19]. Single eigenvalues and the associated eigenvectors (wavefunctions) are calculated numerically using an iterative linear algebra solver.

The band structure obtained by solving the eigenvalue problem repeatedly for various values of $k_z$ is displayed in figure 2(b). As already pointed out, two bands split off from the bulk Si continuum can be identified; thus we set $N = 2$. An analysis of the wavefunctions of these bands shows that they are located in thin sheets between the QDs. Each band is two-fold degenerate, with corresponding symmetry-related wavefunctions that are either localized in the $yz$-plane (figure 2(a)) or in the $xz$-plane (not shown). The inhomogeneous strain on the Si matrix between the Ge QDs defines the confining potential for these states. The lowest band, shown in figure 2(a), corresponds to the lowest bound state in this potential, while the second-lowest band corresponds to an excited state that has an extra node along the confining dimension. At $\varepsilon_c = 0.01 \text{eV}$ above the conduction band edge in unstrained bulk Si, the DOS increases, and hence we treat the states at these higher energies as a 3D continuum.

Thus, the electronic structure of the QD superlattice can be modeled as follows: there are $N = 2$ two-dimensional (2D) bands with constant DOS. The DOS, summed over both spin orientations, and the velocity read

$$g_n(\varepsilon) = g_L \frac{m_n}{\pi^2 \hbar^2 w} \Theta(\varepsilon - \varepsilon_n), \quad v_n = \sqrt{2(\varepsilon - \varepsilon_n)/m_n}$$

with $w = 6a_{\text{Si}}$ being the effective width of the transport channel, as marked in figure 1. The quantities $\varepsilon_n$ and $m_n$ are the energy of the band minimum and the corresponding effective mass, respectively. The band energies are given relative to the position of the conduction band in unstrained silicon. From a fit to the calculated band structure (figure 2(b)), the values $\varepsilon_1 = -0.023 \text{eV}$, $\varepsilon_2 = -0.008 \text{eV}$ and $m_1 = 0.171m_e$, $m_2 = 0.147m_e$ are obtained. The start of the 3D continuum is shifted to a slightly higher energy, $\varepsilon_c = 0.01 \text{eV}$, compared to unstrained bulk Si. Its DOS, summed over both spin orientations, and velocity are described by

$$g_c(\varepsilon) = (g_T + g_L) \frac{m^{3/2}}{\pi^2 \hbar^3} \sqrt{\varepsilon - \varepsilon_c} \Theta(\varepsilon - \varepsilon_c),$$

$$v_c = \sqrt{\frac{2(\varepsilon - \varepsilon_c)}{g_T + g_L} \left( \frac{g_T}{m_T} + \frac{g_L}{m_L} \right)}$$

with $m^* = (m^2_T m_L)^{1/3}$, $m_T$ and $m_L$ being the transverse (small) and the longitudinal (large) effective mass of electrons in bulk Si. The values $m_T = 0.19m_e$ and $m_L = 0.98m_e$ have been adopted from the literature [20]. In the expression for the velocity $v_c$, a weighted average of all six valleys is used. In both expressions above, $\Theta(\varepsilon)$ denotes the Heaviside function.

### 2.2. Doping model

For studying the temperature dependence of the transport coefficients, a consistent calculation of the position of the chemical potential of the electrons $\mu(T)$ is important when evaluating the moments in equation (1). At any given temperature, overall charge neutrality requires the

New Journal of Physics 15 (2013) 125010 (http://www.njp.org/)
number of thermally ionized donors and the number of carriers occupying the bands to be equal. The latter quantity is calculated from our model

\[ n(\mu) = \int_{-\infty}^{\infty} d\varepsilon f_{\mu}^0(\varepsilon) \left( g_c(\varepsilon) + \sum_{n=1}^{N} g_n(\varepsilon) \right), \]

keeping in mind that \( f_{\mu}^0(\varepsilon) \) depends both on \( T \) and \( \mu \). On the other hand, \( n(\mu) \) can be determined by

\[ n(\mu) = n_D \frac{1}{2 \exp \left( \frac{\mu - \varepsilon_D}{k_B T} \right) + 1}, \]

where \( n_D \) is the concentration of donor atoms and \( \varepsilon_D \) is the donor charge transfer level [20]. To be specific, we selected phosphorus as a shallow donor, using the value \( \varepsilon_D = -0.054 \) eV adopted from the literature [21]. Both equations (8) and (9) together define an implicit relation for \( \mu(T) \). For any given \( T \), \( \mu(T) \) is obtained by solving the two implicit equations numerically.

2.3. Relaxation rates

The energy dependence of the transport DOS is crucial for calculating the transport properties at elevated temperatures, in particular for \( S \) and \( \kappa_{el} \) that are determined by the first and second moment of the TDOS, respectively. In the TDOS itself, the relaxation time enters as an energy-dependent factor. Therefore, a realistic and accurate treatment of the carrier relaxation is required. We define separate energy-dependent relaxation rates for each of the isolated bands and for the continuum. The following scattering processes are included in the spirit of Matthiesen’s rule: for the continuum, scattering to other continuum states is considered, either by acoustic phonons, by intervalley scattering or by ionized impurities

\[ \frac{1}{\tau_c} = \frac{1}{\tau_{acu}^c} + \frac{1}{\tau_{interv}^c} + \frac{1}{\tau_{imp}^c}. \]

The impurity scattering rate, which we calculate using the Conwell–Weisskopf approach [21], is proportional to the concentration of ionized donors, which is equal to the number of free carriers \( n(\mu) \). For calculating \( \tau_{acu}^c \), we use the deformation potential model of Herring and Voigt [22]. The intervalley scattering is treated as suggested by Ferry [23]. More details are given in the appendix.

For each of the 2D bands, we consider intra-band scattering by acoustic phonons (\( \tau_{acu}^n \)), as well as acoustic phonon scattering into the continuum states of the same valley (\( \tau_{acu,c}^n \)) that becomes accessible at high electron energies. Moreover, intervalley scattering (\( \tau_{interv}^n \)) is taken into account within the same band, and to the other (degenerate or non-degenerate) band(s) if energetically possible. For scattering by charged impurities (\( \tau_{imp}^n \)), we consider the dielectric screening of remote impurities due to the electron gas in the 2D bands [24]

\[ \frac{1}{\tau_n} = \frac{1}{\tau_{acu}^n} + \frac{1}{\tau_{acu,c}^n} + \frac{1}{\tau_{interv}^n} + \frac{1}{\tau_{imp}^n}. \]

We discuss in somewhat more detail the behavior of the relaxation rates in the strain-induced 2D bands as a function of energy. It is interesting to note that, in a 2D band, the acoustic phonon scattering rate increases linearly with temperature, but is independent of the electron energy, since the DOS \( g_n(\varepsilon) \) is constant. For calculating thermoelectric properties, however,
Figure 3. Total relaxation rates for the 2D bands of the QD superlattice at 100 K and dopant concentration \( N_D = 10^{17} \text{cm}^{-3} \) as function of electron energy. The structuring is due to contributions from different scattering mechanisms: (a) lower 2D band and (b) upper 2D band.

care must be taken to describe the energy relaxation of carriers with higher energy, several \( k_B T \) above the band edge. For these relaxation processes, the emission of optical phonons becomes important, and hence the energy dependence of the associated matrix elements is of interest. According to Ferry [23], the intervalley scattering is best described by an incoherent superposition of two contributions that are first order and zeroth order in the phonon energy. Therefore, each intervalley scattering rate is built up from two contributions \( 1/\tau^{(0)} \) and \( 1/\tau^{(1)} \). Moreover, summation over different final state bands \( m \) must be carried out. Thus, we obtain

\[
\frac{1}{\tau_{\text{interv}}} = \sum_{m=1}^{N} \frac{1}{\tau^{(0)}_{nm}} + \frac{1}{\tau^{(1)}_{nm}}.
\]

The detailed evaluation of the scattering rates is described in the appendix. In brief, one has to distinguish between g-processes that mediate scattering between valleys at opposite \( \vec{k} \), i.e. from \( \vec{k} \) to \(-\vec{k}\), and f-processes that account for off-axis scattering to the four other valleys. In the present case, two wavefunctions whose orbital character is derived from different valleys of the bulk band structure are spatially separated and located at different edges of the QD. Therefore, the f-process intervalley scattering is strongly suppressed in the low-lying 2D bands. Moreover, since the degeneracy of band edges has been lifted by the strain, different energy thresholds for the various scattering channels need to be taken into account: the energy difference between the final and initial state, e.g. \( \varepsilon_2(\vec{k}') \) and \( \varepsilon_1(\vec{k}) \), must equal the phonon energy. The conservation laws of both energy and crystal momentum determine the onset and the available phase space for a certain intervalley scattering process. For both these reasons, the scattering rates in the 2D bands are reduced compared to the scattering rate in the 3D continuum.

The energy dependence of the scattering rates can be seen in figure 3. Due to the thresholds of the various scattering processes, \( \tau^{-1}(\varepsilon) \) is a piecewise defined function. It starts at the band minimum with a constant defined by acoustic phonon scattering within the band. This is followed by a square-root-like increase. From this energy onward, scattering by acoustic phonons into the continuum states becomes possible, and the shape of the curve is given by the continuum DOS, see equation (6). The abrupt jumps visible in figure 3(a) mark the onset of intervalley scattering. The scattering rates in the second band, displayed in figure 3(b), show similar, but shifted structures compared to the lowest band. Both figures 3(a) and (b) clearly show that \( \tau(\varepsilon) \) in the 2D bands is strongly energy-dependent, and thus cannot be taken out of the
integral in equation (1), as it is sometimes done in oversimplified treatments of thermoelectric transport. Depending on the position of the chemical potential \( \mu \), the skewness of the TDOS around \( \mu \) will determine the size the Seebeck coefficient. Hence, this property is affected by the energy dependence of \( \tau(\varepsilon) \).

3. Results

Using equations (1)–(4), we calculate the electrical conductivity \( \sigma \), the mobility, the Seebeck coefficient and the power factor \( \sigma S^2 \), the electrical contribution \( \kappa_{el} \) to the thermal conductivity and the figure of merit \( ZT \). The results are shown in figures 4(a)–(e) as function of temperature for three different concentrations of dopants, \( n_D = 10^{17} \), \( 10^{18} \) and \( 10^{19} \) cm\(^{-3} \).

As expected for a semiconducting sample, the electrical conductivity \( \sigma \) (figure 4(a)) goes through a maximum when raising the temperature. At very low temperatures, \( \sigma \) increases with the number of free carriers. At high temperatures, the trend is reversed by the increased scattering rate due to thermally excited phonons. Comparing to the calculated conductivity for bulk Si for \( n_D = 10^{19} \) cm\(^{-3} \) (dot-dashed line in figure 4(a)), one can clearly see that the unusually sharp peak originates from the contribution of the 2D bands. This peak appears at low temperatures (below room temperature), when the chemical potential \( \mu(T) \) passes through the low-lying 2D bands. The increase of the peak height as function of the dopant concentration \( n_D \) is sublinear. This is a consequence of the thermal activation of the carriers: As the doping level is increased at low temperatures, the chemical potential rises, and thus the absolute number of carriers increases, but a smaller fraction of the dopants is ionized.

As shown in figure 4(f), we obtain carrier mobilities at \( T = 250 \) K in the range of 2000–2500 cm\(^2\) V\(^{-1}\) s\(^{-1}\) for samples with high and low dopant concentration, respectively. At room temperature, the mobility is about 1500 cm\(^2\) V\(^{-1}\) s\(^{-1}\) and rather independent of doping, i.e. still as high as for perfect Si bulk crystals. The reason is the much-reduced scattering in the 2D bands: impurity scattering, which is most important for low-energy carriers, is reduced due to dielectric screening of remote impurities by the 2D electron gas. At higher carrier energies, where electron–phonon scattering dominates, the suppression of f-process intervalley scattering, originating from the modified wavefunctions between the Ge QDs in the Si host, is the main reason for the high mobility in the 2D bands.

As seen in figure 4(b), the power factor \( \sigma S^2 \), which is the numerator of the figure of merit, has a clear maximum at about 110 K. Its position in temperature seems to be unaffected by the doping concentration, but its magnitude clearly is. The coincidence of the peak position in the power factor and in the electrical conductivity indicates that the latter is responsible for the maximum in the power factor. The Seebeck coefficient \( S(T) \) (figure 4(c)) for the highest dopant concentrations ranges between \(-0.24 \) and \(-0.42 \) mV K\(^{-1}\), and shows little temperature variability between 200 and 1000 K. Compared to the results for bulk Si, the absolute magnitude of \( S \) is actually lower in the nanostructured samples. Thus, the increased power factor is solely due to the high carrier mobility.

Our results for the electronic contribution \( \kappa_{el} \) to the thermal conductivity, figure 4(d), show that the electrons make a minor contribution, less than 3 mW K\(^{-1}\) cm\(^{-1}\) for the highest temperature and doping level, compared to the lattice contribution \( \kappa_{ph} \), which is known from experiment [6] to be 10 mW K\(^{-1}\) cm\(^{-1}\). Different from bulk materials, \( \kappa_{el} \) and \( \sigma \) are not proportional, and hence there is little risk of these two factor compensating each other in the figure of merit \( ZT \).

New Journal of Physics 15 (2013) 125010 (http://www.njp.org/)
Figure 4. Electrical conductivity (a), mobility (b), power factor (c), Seebeck coefficient (d), electronic contribution to the thermal conductivity (e) and figure of merit ZT (c) for doping concentrations of $10^{17}$ cm$^{-3}$ (dashed, orange), $10^{18}$ cm$^{-3}$ (solid, red) and $10^{19}$ cm$^{-3}$ (dotted, blue). For comparison, the calculated values of bulk silicon are shown for the doping concentration of $10^{19}$ cm$^{-3}$ (dot-dashed, gray) in (a)–(d). (f) Mobility averaged over all bands as function of doping concentration $n_D$ for $T = 250$ K (dashed), 500 K (solid) and 1000 K (dotted line).

At the temperature where the power factor is maximum, we observe a steep rise in $ZT$ (see figure 4(e)). For doping concentrations of $10^{17}$ and $10^{18}$ cm$^{-3}$, this rise in the figure of merit tapers off and passes through a maximum at about 200 K. For $n_D = 10^{18}$ cm$^{-3}$, $ZT \approx 1$ is reached already below room temperature. In materials systems with similar composition, such as sintered SiGe nanopowders [2], values of $ZT$ above unity are not achieved until the
temperature reaches about 800 K. Moreover, our calculations show a different behavior of $ZT$ in the highly doped regime, which starts in our case at $10^{19}$ cm$^{-3}$. Whereas for low doping concentrations $ZT$ eventually decreases slightly with temperature, it keeps rising linearly with $T$ for the highest doping value, only the slope changes and becomes smaller above $T \approx 200$ K. At these high carrier concentrations, not only the 2D bands, but also the 3D continuum become populated by carriers at high temperatures. While the carriers in the 3D host matrix now also contribute to the power factor, the contribution of the highly mobile carriers in the 2D bands persists, resulting in an enhancement of $ZT$.

We note that the numerical value of $ZT$ that can be reached with the QD superlattice depends on some assumptions entering our calculations. For instance, we have used a very low, yet realistic value of the lattice thermal conductivity $\kappa_{\text{ph}} = 10 \text{ mW K}^{-1} \text{ cm}^{-1}$ as input, that has actually been measured in QD superlattices of similar geometry. [6]. For samples fabricated from SiGe nanoparticles via current-induced sintering, $\kappa_{\text{ph}}$ was found to be a factor of 2 [3] to 2.5 [2] bigger, showing little temperature dependence. At room temperature and below, the enhanced $ZT$ value is mostly due to the high carrier mobility in the fully crystalline QD superlattices.

It is clear that the geometry of the QD superlattice investigated by us is somewhat idealized. In real samples, segregation of Ge will lead to a composition grading near the QDs, rather than a sharp interface. For the same reason, the Ge concentration in the QD will not be 100%, but rather be in the 80% range. Due to both effects, the Si host material around the QD will be less strained, and the 2D-like bands will be less clearly separated from the conduction band continuum. Therefore, we would expect that the peak in $\sigma$ and the shoulder in $ZT$ will be less pronounced and occur at even lower temperatures in more realistic samples. However, the effect of an enhanced thermoelectric performance due to high-mobility transport channels in the QD superlattice will persist, and the present calculations provide an upper limit for the size of the effects to be expected in real samples.

4. Conclusions

In summary, our transport model on the basis of tight-binding calculations provides evidence for an enhanced figure of merit in fully crystalline, epitaxially grown Si/Ge QD superlattices. The values $ZT > 1$ have their origin in the 2D electronic states introduced below the conduction band edge of silicon due to the strain around the QDs. The design principle of the materials system proposed here, the Si/Ge QD superlattice, is complementary to the typical highly doped thermoelectric materials, demonstrating that high $ZT$ values can be achieved with lower carrier concentrations if samples are fully crystalline and support transport channels for highly mobile carriers. In the present case, the reduced intervalley scattering and the low impurity scattering due to dopants help to realize such high electron mobility. Our findings suggest the possibility of extending the range of efficient operation of SiGe nanostructured compounds down to lower temperatures, and even below room temperature.

Acknowledgments

We gratefully acknowledge useful discussions about the theory with Igor Bejenari and about the experimental realization of quantum dot superlattices with Pei-Xuan Chen and
Appendix. Calculation of the intervalley scattering rate

In order to describe the intervalley scattering in Si, we follow the approach of Ferry [23] that has been successfully applied to transport calculations in Si quantum wells [25]. In this approach, two effective phonon energies and corresponding deformation potentials are used. The latter ones are specified as $D_0 = 9.0 \times 10^8$ eV cm$^{-1}$ and $D_1 = 5.6$ eV for the zeroth-order and first-order contribution, respectively. The phonon frequencies are characterized by effective temperatures according to the relation $\omega_i = k_B T_i / \hbar$. For the $g$-process, the values are $T_0 = 750$ and $T_1 = 134$ K; for the $f$-process $T_0 = 630$ K and $T_1 = 230$ K are used. The absolute square of the matrix element can be obtained from a deformable-ion model [26]. For the zeroth order one obtains

$$|\langle \tilde{k} | H | \tilde{k} \pm \tilde{q} \rangle|^2 = \frac{D_0^2 \hbar}{2V \rho_m \omega_0} |I|^2 \left( n(\omega_0) + \frac{1}{2} \pm \frac{1}{2} \right)$$

(A.1)

and for the first order

$$|\langle \tilde{k} | H | \tilde{k} \pm \tilde{q} \rangle|^2 = \frac{D_1^2 \hbar q^2}{2V \rho_m \omega_1} |I|^2 \left( n(\omega_1) + \frac{1}{2} \pm \frac{1}{2} \right).$$

(A.2)

Here, $n(\omega)$ is the Bose–Einstein distribution describing the thermal occupation of a phonon with energy $\hbar \omega$ at temperature $T$, and the $+$ or $-$ sign in the factor $n(\omega) + \frac{1}{2} \pm \frac{1}{2}$ stands for phonon emission or absorption, respectively. $\omega_0$ and $\omega_1$ are the frequencies of the phonon involved in the zeroth order or first order process, respectively. $\rho_m$ is the crystal mass density and $V$ the crystal volume, and $|I|^2$ the overlap integral of electronic states with $\tilde{k}$ and $\tilde{k}'$, while $\tilde{k}' = \tilde{k} \pm \tilde{q}$. Our treatment relies on qualitative similarity of the phonon spectrum of the QD superlattice and bulk Si, in particular, it relies on the persisting presence of a quasi-continuum of phonon modes, which is justified by comparison to explicit calculations [27]. Only if spatially confined phonon modes become involved in the scattering, their locally enhanced amplitude, which is observable e.g. in Raman scattering [28], is missed in the present treatment of electron–phonon coupling.

To obtain the scattering rate, the sum over all wavevectors $\tilde{q}$ of the phonon must be carried out. Following Ferry [25] and considering that the states are 2D, we obtain for the scattering rate from the subband $n$ to $m$ from the zeroth order term

$$\frac{1}{\tau_{nm}^{(0)}(\varepsilon)} = \frac{D_0^2 m_T}{\hbar^2 \rho_m \omega_0} |I_{nm}|^2 \left( n(\omega_0) \Theta(\varepsilon - \varepsilon_m + \hbar \omega_0) + (n(\omega_0) + 1) \Theta(\varepsilon - \varepsilon_m - \hbar \omega_0) \right),$$

and from the first order term

$$\frac{1}{\tau_{nm}^{(1)}(\varepsilon)} = \frac{D_1^2 m_T^2}{\hbar^4 \rho_m \omega_1} |I_{nm}|^2 \left( n(\omega_1) (2 \varepsilon - \varepsilon_m - \varepsilon_n + \hbar \omega_1) \Theta(\varepsilon - \varepsilon_m + \hbar \omega_1) ight.$$

$$\left. + (n(\omega_1) + 1)(2 \varepsilon - \varepsilon_m - \varepsilon_n - \hbar \omega_1) \Theta(\varepsilon - \varepsilon_m - \hbar \omega_1) \right).$$

These formulas are adopted from [25] with slightly modified notation. Since our model implies different band edge energies for each band, and the kinetic energy $E$ must be counted from the band edge

$$E := \frac{\hbar k^2}{2m} + \varepsilon_x, \quad x = 1, 2 \text{ or } c.$$  

(A.3)

We made this explicit in the formulas by substituting $E \mapsto \varepsilon - \varepsilon_x$. 

New Journal of Physics 15 (2013) 125010 (http://www.njp.org/)
Table A.1. Matrix elements \( |I_{nm}|^2 \) in \( 10^5 \text{ cm}^{-1} \) calculated from tight-binding wavefunctions of the lowest 2D bands according to equation (A.4).

|          | \( I_{00} \) | \( I_{01} \) | \( I_{11} \) |
|----------|--------------|--------------|--------------|
| g-process| 17.39        | 12.65        | 18.75        |
| f-process| 2.61         | 6.45         | 11.70        |

Moreover, instead of using the width \( w \) of the quantum well, as in Ferry’s work, we calculate the matrix element explicitly from the tight-binding wavefunctions, equation (A.4).

Apart from the matrix elements \( I_{nm} \) for the g-process, where \( \vec{k} \) and \( \vec{k}' \) are antiparallel, there are f-process matrix elements for scattering into states \( m' \), where \( m' \) denotes the degenerate partner to state \( m \), i.e. a state with \( \varepsilon_{m'} = \varepsilon_m \), but with a wavefunction rotated by 90° around the \( z \)-axis through the QD center. The corresponding matrix elements \( I_{nm'} \) are smaller than \( I_{nm} \), because the wavefunctions are located on different sides of the QD, and thus hardly overlap. In effect, f-process intervalley scattering is suppressed in the low-lying bands, as seen from table A.1.

In order to calculate the matrix elements \( |I_{nm}|^2 \) that enter the scattering rates, we follow standard approximation procedures in the literature. The matrix elements can be reduced to the overlap of probability densities (see e.g. the appendix in [29]), if the scattering occurs by acoustic phonons or by optical phonons with negligible dispersion. In the present case of 2D electronic states, e.g. in the \( yz \)-plane, the wavefunctions are plane waves in these directions. We therefore normalize the probability density to the area of the periodic supercell in these two directions

\[
\int_0^{L_x} \int_0^{L_y} \int_0^{L_z} \Psi_n(x, y, z) \Psi_n(x, y, z)^* \, dx \, dy \, dz = L_x L_y L_z.
\]

To calculate the scattering matrix element with this choice of the normalization, one has to average over the extended directions \( y \) and \( z \), and to integrate over the direction of confinement, \( x \)

\[
|I_{nm}|^2 = \int_0^{L_x} \int_0^{L_y} \int_0^{L_z} \frac{L_y L_z}{(\hbar^2 \rho_m v_s^2)} |\Psi_n(x, y, z)|^2 |\Psi_m(x, y, z)|^2
\]

Using tight-binding wavefunctions, the 3D integral is evaluated in the on-site approximation, as described in [30].

For acoustic phonon scattering within a 2D band, we use the expression

\[
\frac{1}{\tau_{acu}^n} = \frac{m_T \Sigma_u^2}{\hbar^2 \rho_m v_s^2} |I_{nm}|^2 k_B T
\]  

(A.5)

with \( v_s = 5840 \text{ m s}^{-1} \) being the speed of sound in silicon, and the acoustic deformation potential \( \Sigma_u = 8.77 \text{ eV} \).

The formalism for the intervalley scattering within the continuum states can also be taken from Ferry’s work [23] on bulk Si. Again, both zeroth-order and first-order contributions are considered,

\[
\frac{1}{\tau_{interv}^c} = \frac{1}{\tau_{c,g}^{(0)}} + \frac{g_T}{\tau_{c,g}^{(1)}} + \frac{1}{\tau_{c,f}^{(0)}} + \frac{g_T}{\tau_{c,f}^{(1)}}.
\]

(A.6)

New Journal of Physics 15 (2013) 125010 (http://www.njp.org/)
The formula for the zeroth-order term reads

$$\frac{1}{\tau_{c,0}^{(0)}(\epsilon)} = \frac{\sqrt{2m^*} D_0^2}{\hbar^2 \rho_m \omega_{0,x}} (n(\omega_{0,0}) \sqrt{\epsilon + \hbar \omega_{0,0} \Theta(\epsilon - \epsilon_c + \hbar \omega_{0,0})} + (n(\omega_{0,0}) + 1) \sqrt{\epsilon - \hbar \omega_{0,0} \Theta(\epsilon - \epsilon_c - \hbar \omega_{0,0}))},$$

where $x \in \{f, g\}$, and for the first-order term

$$\frac{1}{\tau_{c,0}^{(1)}(\epsilon)} = \frac{\sqrt{2m^*} D_0^2}{\hbar^2 \pi \rho_m \omega_{1,x}} (n(\omega_{1,0}) \sqrt{\epsilon + \hbar \omega_{1,0} (2(\epsilon - \epsilon_c) + \hbar \omega_{1,0}) \Theta(\epsilon - \epsilon_c + \hbar \omega_{1,0})} + (n(\omega_{1,0}) + 1) \sqrt{\epsilon - \hbar \omega_{1,0} (2(\epsilon - \epsilon_c) - \hbar \omega_{1,0}) \Theta(\epsilon - \epsilon_c - \hbar \omega_{1,0}))}.$$
[18] Plimpton S 1995 Fast parallel algorithm for short-range molecular dynamics J. Comput. Phys. 117 1–19
[19] Grosso G and Piermarocchi C 1995 Tight-binding model and interactions scaling laws for silicon and germanium Phys. Rev. B 51 16772–7
[20] Ashcroft N W and Mermin D N 1976 Solid State Physics (Philadelphia, PA: Saunders)
[21] Yu P Y and Cardona M 2003 Fundamentals of Semiconductors Physics and Materials Properties 3rd edn (Berlin: Springer)
[22] Herring C and Voigt E 1956 Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering Phys. Rev. 101 944–61
[23] Ferry D K 1976 First-order optical and intervalley scattering in semiconductors Phys. Rev. B 14 1605–9
[24] Ando T, Fowler A B and Stern F 1982 Electronic properties of two-dimensional systems Rev. Mod. Phys. 53 437
[25] Ferry D K 1976 Hot-electron effects in silicon quantized inversion layers Phys. Rev. B 14 5364–71
[26] Harrison W A 1956 Scattering of electrons by lattice vibrations in nonpolar crystals Phys. Rev. 104 1281
[27] Gillet J-N, Chalopin Y and Volz S 2009 Atomic-scale three-dimensional photonic crystals with a very low thermal conductivity to design crystalline thermoelectric devices J. Heat Transfer 131 043206
[28] Lazarenkova O L and Balandin A A 2003 Raman scattering from three-dimensionally regimented quantum dot superlattices Superlatt. Microstruct. 33 95–101
[29] Knezevic I, Ramayya E B, Vasileska D and Goodnick S M 2009 Diffusive transport in quasi-2D and quasi-1D electron systems J. Comput. Theor. Nanosci. 6 1725–53
[30] Neophytou N and Kosina H 2011 Effects of confinement and orientation on the thermoelectric power factor of silicon nanowires Phys. Rev. B 83 245305