Evaluation of Spin Lifetime in Thin-Body FETs: A High Performance Computing Approach

Joydeep Ghosh(✉), Dmitry Osintsev, Viktor Sverdlov, Josef Weinbub, and Siegfried Selberherr

Institute for Microelectronics, TU Wien, Vienna, Austria
{ghosh,osintsev,sverdlov,weinbub,selberherr}@iue.tuwien.ac.at

Abstract. Silicon, the prominent material of microelectronics, is perfectly suited for spin-driven applications because of the weak spin-orbit interaction resulting in long spin lifetime. However, additional spin relaxation on rough interfaces and acoustic phonons may strongly decrease the spin lifetime in modern silicon-on-insulator and trigate transistors. Because of the need to perform numerical calculation and appropriate averaging of the strongly scattering momenta depending spin relaxation rates, an evaluation of the spin lifetime in thin silicon films becomes prohibitively computationally expensive. We use a highly parallelized approach to calculate the spin lifetime in silicon films. Our scheme is based on a hybrid parallelization approach, using the message passing interface MPI and OpenMP. The algorithm precalculates wave functions and energies, and temporarily stores the results in a file-based cache to reduce memory consumption. Using the precalculated data for the spin relaxation rate calculations drastically reduces the demand on computational time. We show that our approach offers an excellent parallel speedup, and we demonstrate that the spin lifetime in strained silicon films is enhanced by several orders of magnitude.

1 Introduction

For almost half a century Moore's law has successfully predicted the persistent miniaturization of semiconductor devices, such as the transistor feature size in microprocessors and unit cell in magnetic storage disks and random access memories. However, as devices are scaled down to the nano-scale, fundamental physical limitations will hinder further improvements in device performances in the upcoming years. Employing spin as an additional degree of freedom is promising for boosting the efficiency of future low-power integrated electronic circuits. Silicon is characterized by a weak spin-orbit interaction and long spin lifetime. It is therefore an attractive material for spin-driven applications. A long spin transfer distance of conduction electrons has been shown experimentally [1]. However, a large experimentally observed spin relaxation in electrically-gated silicon films could become an obstacle in realizing spin-driven devices [2]. Henceforth, a deeper understanding of the fundamental spin relaxation mechanism in silicon MOSFETs is required. We consider the surface roughness (SR) and the

© Springer International Publishing Switzerland 2015
I. Lirkov et al. (Eds.): LSSC 2015, LNCS 9374, pp. 285–292, 2015.
DOI: 10.1007/978-3-319-26520-9_31

longitudinal (LA) and transversal (TA) acoustic phonons to cause the prominent spin relaxation mechanisms. In (001) silicon films the conduction electrons are positioned close to the minima of the pair of valleys near the edges of the Brillouin zone along the axis. Each state is described by the subband index, the in-plane wave vector $\mathbf{k}$, and the spin orientation (spin-up and spin-down) on a chosen axis. The subband wave functions and the rates are obtained by the perturbative $\mathbf{k}\cdot\mathbf{p}$ model [3]. The spin relaxation time is calculated from the obtained rates by thermal averaging [4,5]:

$$\frac{1}{\tau} = \frac{\int \frac{1}{\tau(\mathbf{K_1})}\cdot f(E)(1-f(E))d\mathbf{K_1}}{\int f(E)d\mathbf{K_1}}, \tag{1}$$

$$f(E) = \frac{1}{1+\exp(\frac{E-E_F}{K_B T})}, \int d\mathbf{K_1} = \int_0^{2\pi} d\phi \cdot \int_0^{\infty} \frac{|\mathbf{K_1}(\phi,E)|}{|\frac{\partial E(\mathbf{K_1})}{\partial \mathbf{K_1}}|}dE \tag{2}$$

E is the electron energy, $\mathbf{K_1}$ is the in-plane after-scattering wave vector, K_B is the Boltzmann constant, T is the temperature, and E_F is the Fermi level.

The SR-limited spin relaxation rate is expressed as:

$$\frac{1}{\tau_{SR}(\mathbf{K_1})} = \frac{4\pi}{\hbar(2\pi)^2}\sum_{i,j=1,2}\int_0^{2\pi}\pi\triangle^2 L^2\cdot\frac{1}{\epsilon_{ij}^2(\mathbf{K_2}-\mathbf{K_1})}\cdot\frac{\hbar^4}{4m_l^2}\cdot\frac{|\mathbf{K_2}|}{|\frac{\partial E(\mathbf{K_2})}{\partial \mathbf{K_2}}|}\cdot [(\frac{d\psi_{i\mathbf{K}_{1\sigma}}}{dz})^*(\frac{d\psi_{j\mathbf{K}_{2-\sigma}}}{dz})]^2_{z=\pm\frac{t}{2}}\cdot\exp(\frac{-(\mathbf{K_2}-\mathbf{K_1})^2L^2}{4})d\phi \tag{3}$$

$\mathbf{K}_2$ and $\mathbf{K}_1$ are the in-plane wave vectors before and after scattering, ϕ is the angle between $\mathbf{K}_1$ and $\mathbf{K}_2$, ϵ is the dielectric permittivity, L is the autocorrelation length, $\triangle$ is the mean square value of the SR-fluctuations, $\psi_{i\mathbf{K}_{1\sigma}}$ and $\psi_{j\mathbf{K}_{2\sigma}}$ are the wave functions, $\sigma = \pm 1$ is the spin projection to the [001] axis, and m_l is the longitudinal effective mass. The TA-phonon induced intravalley spin relaxation rate can be written as:

$$\frac{1}{\tau_{TA}(\mathbf{K_1})} = \frac{\pi K_B T}{\hbar\rho\nu_{TA}^2}\sum\int_0^{2\pi} d\phi\cdot\frac{|\mathbf{K}_2|}{|\frac{\partial E(\mathbf{K}_2)}{\partial \mathbf{K}_2}|}\left[1-\frac{|\frac{\partial E(\mathbf{K}_2)}{\partial \mathbf{K}_2}|f(E(\mathbf{K}_2))}{|\frac{\partial E(\mathbf{K}_1)}{\partial \mathbf{K}_1}|f(E(\mathbf{K}_1))}\right]\cdot \int_0^t\int_0^t \exp(-\sqrt{q_x^2+q_y^2}|z-z'|)[\psi^{\dagger}_{\mathbf{K}_{2\sigma}}(z)M\psi_{\mathbf{K}_{1-\sigma}}(z)]^*[\psi^{\dagger}_{\mathbf{K}_{2\sigma}}(z')M\psi_{\mathbf{K}_{1-\sigma}}(z')]\cdot [\sqrt{q_x^2+q_y^2}-\frac{8q_x^2q_y^2-(q_x^2+q_y^2)^2}{q_x^2+q_y^2}|z-z'|]dzdz' \tag{4}$$

$\rho = 2329\,\mathrm{kg/m^3}$ is the silicon density, $\nu_{TA} = 5300\,\mathrm{m/s}$ is , t is the film thickness, $(q_x,\ q_y)$=$\mathbf{K_1}-\mathbf{K_2}$, and M written in the two valley plus two spin projection basis is ($D = 14\,\mathrm{eV}$: shear deformation potential) $\begin{bmatrix} 0 & 0 & \frac{D}{2} & 0 \\ 0 & 0 & 0 & \frac{D}{2} \\ \frac{D}{2} & 0 & 0 & 0 \\ 0 & \frac{D}{2} & 0 & 0 \end{bmatrix}$.

The intravalley spin relaxation rate due to LA-phonons is:

$$\frac{1}{\tau_{LA}(\mathbf{K_1})} = \frac{\pi K_B T}{\hbar \rho \nu_{LA}^2} \sum \int_0^{2\pi} d\phi \cdot \frac{|\mathbf{K_2}|}{|\frac{\partial E(\mathbf{K_2})}{\partial \mathbf{K_2}}|} \left[1 - \frac{|\frac{\partial E(\mathbf{K_2})}{\partial \mathbf{K_2}}| f(E(\mathbf{K_2}))}{|\frac{\partial E(\mathbf{K_1})}{\partial \mathbf{K_1}}| f(E(\mathbf{K_1}))}\right]$$
$$\cdot \int_0^t \int_0^t \exp(-\sqrt{q_x^2+q_y^2}|z-z'|)[\psi^{\dagger}_{\mathbf{K}_{2\sigma}}(z) M \psi_{\mathbf{K}_{1-\sigma}}(z)]^* [\psi^{\dagger}_{\mathbf{K}_{2\sigma}}(z') M \psi_{\mathbf{K}_{1-\sigma}}(z')] \cdot$$
$$\frac{4 q_x^2 q_y^2}{(q_x^2+q_y^2)^{3/2}} [\sqrt{q_x^2+q_y^2}|z-z'|+1] dz dz' \tag{5}$$

$\nu_{LA} = 8700\,\mathrm{m/s}$ is the longitudinal phonon velocity.

The intervalley spin relaxation rate due to acoustic phonons contains Elliot and Yafet contributions and is calculated as:

$$\frac{1}{\tau_{LA}(\mathbf{K_1})} = \frac{\pi K_B T}{\hbar \rho \nu_{LA}^2} \sum \int_0^{2\pi} d\phi \cdot \frac{|\mathbf{K_2}|}{|\frac{\partial E(\mathbf{K_2})}{\partial \mathbf{K_2}}|} \left[1 - \frac{|\frac{\partial E(\mathbf{K_2})}{\partial \mathbf{K_2}}| f(E(\mathbf{K_2}))}{|\frac{\partial E(\mathbf{K_1})}{\partial \mathbf{K_1}}| f(E(\mathbf{K_1}))}\right]$$
$$\cdot \int_0^t [\psi^{\dagger}_{\mathbf{K}_{2\sigma}}(z) M' \psi_{\mathbf{K}_{1-\sigma}}(z)]^* [\psi^{\dagger}_{\mathbf{K}_{2\sigma}}(z) M' \psi_{\mathbf{K}_{1-\sigma}}(z)] dz \tag{6}$$

M' is $\begin{bmatrix} D' & 0 & 0 & D_{SO}(q_y - iq_x) \\ 0 & D' & D_{SO}(-q_y - iq_x) & 0 \\ 0 & D_{SO}(-q_y + iq_x) & D' & 0 \\ D_{SO}(q_y + iq_x) & 0 & 0 & D' \end{bmatrix}$ with $D' = 12\,\mathrm{eV}$, $D_{SO} = 15\,\mathrm{meV/k_0}$ (k_0=$0.15\frac{2\pi}{a}$, a being silicon lattice constant).

2 Simulations

In order to evaluate the spin relaxation time, a multi-dimensional integral (2) over the energy E and angle ϕ must be computed. The spin relaxation matrix elements are characterized by very narrow and sharp peaks (so-called spin hot spots [6]). To resolve these sharp features, the mesh in the $\mathbf{K}$ space has to be precise. In our application we have determined that the energy step value $\triangle E$ should be upper-bounded by 0.5meV. The lower limit of the integral over E is zero, and we have also identified that it is sufficient to set the corresponding upper limit to be 0.7eV. This particular simulation setup requires almost 1400 points. The lower and the upper limits of the integral over ϕ are 0° and 360° respectively. The step value for ϕ, or $\triangle\phi$, is set to be smaller than 0.5°. Hence, the inner integral over ϕ on before- and after-scattering directions at fixed energies require almost 1000 points each. Thus, the scattering matrix elements and the Jacobians (the derivative of the dispersion energy over the wave vector) must be calculated numerically for almost 1,400,000 times. To compute the matrix elements, the eigenfunction problems for the 4×4 Hamiltonian matrix must be solved for the two wave vectors before and after scattering for a broad range of parameters, which makes the numerical spin relaxation time calculation prohibitively expensive: When utilizing a standard adaptive integration technique,

we found that a month of calculations on 20 cores, or 15000 core-hours total, was required to evaluate a single data point of τ as a function of stress $\varepsilon_{\mathrm{xy}}$. In order to improve the calculation time, we have divided the entire computation into two levels. The first level calculates and archives all static wave functions and energy data to a binary file (a file-based cache technique), and the second level performs the spin lifetime calculations by loading those data in memory. Both of these levels individually perform the calculations in parallel by using the message passing interface MPI and OpenMP.

2.1 Parallelization Algorithm for Spin Relaxation Rate Calculations

In the following our two-level computation algorithm is outlined.
Level 1

- (**1**) Divide the range of angle ϕ into sub-domains for each MPI process.
- (**1.1**) Divide the range of energy E into sub-domains for each OpenMP thread.
- (**1.1.1**) Calculate the derivatives at the interface $(\frac{d\psi}{dz})_{z=\pm\frac{t}{2}}$, and $\frac{|\mathbf{K}|}{|\frac{\partial E(\mathbf{K})}{\partial \mathbf{K}}|}$ in parallel (MPI, OpenMP).
- (**2**) Collect all the cached values at the master MPI process.
- (**3**) Archive the cache to a binary file.

Level 2

- (**4**) Load archived cache by the master MPI process.
- (**5**) Divide the range of ϕ into sub-domains for each MPI process.
- (**5.1**) Divide the range of E into sub-domains for each OpenMP thread.
- (**5.1.1**) Calculate (1) for a given range of values in parallel (MPI, OpenMP).
- (**6**) Collect all calculated relaxation rates into the final relaxation rate.

The performance is measured on the Vienna Scientific Cluster (VSC-2) [7], which consists of 1285 nodes with 2 processors each (AMD Opteron 6132 HE, 2.2 GHz and 8 cores) and 32 GB main memory on each node. We investigate the calculation time, memory consumption, and utilization of computational resources for each level.

In the first level we scrutinize the limitations of our file-based cache calculation technique. We have examined different configurations of MPI and OpenMP with a fixed number of cores 96 (i.e. number of nodes is 6). Using our discretization scheme, we have found that the memory consumption increases by a factor of around three by using a pure MPI approach, compared to a hybrid MPI-OpenMP (i.e. 6 MPI processes, each using 16 threads) configuration. However, even very accurate calculations with a pure MPI approach require less than 10 GB of memory per node, hence memory limitations are not an issue considering a modern supercomputer and this particular simulation setup. In contrast, we find that the total calculation time is reduced by 30 % with a pure MPI configuration, compared to a hybrid MPI-OpenMP scheme. The performance decrease of the hybrid approach is due to data locality issues arising in shared-memory

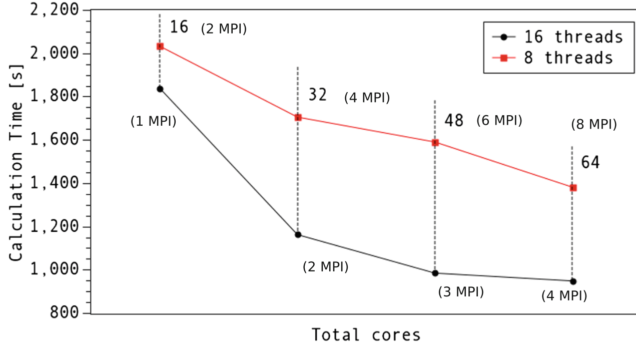

Fig. 1. Dependence of the total calculation time on total cores for a fixed number of threads. 1, 2, 3, and 4 nodes are used respectively. For 16(8)-threaded application, we use one(two) MPI job(s)/node.

techniques. Therefore, we conclude that our cache approach is most efficient for a pure MPI configuration.

In the second level, we compute τ by using the archived data in parallel. For our discretization parameters, the size of the archived cache is up to 3 GB. The even smaller energy steps ($\triangle E = 0.2\,\text{meV}$) improve the computational accuracy but increase the required cache size to more than 7.5 GB. This archived data is to be loaded into the memory while calculating τ by each MPI job, and thus the number of parallel MPI jobs on a single node becomes strictly limited. Theoretically only three processes can work together on a single node on our supercomputer, thus leading to a significant loss of computational resources. Under these conditions, it becomes inevitable to use a hybrid MPI-OpenMP configuration, albeit its execution performance limitations. Figure 1 shows the dependence of the total calculation time on the number of cores and the number of threads. This confirms that increasing the total number of cores at a fixed number of threads decreases the demand on computing time, which is further reduced when the number of threads is increased. This approach was tested with 416 cores and requires only around 40 min for a single relaxation time data point (around 280 core-hours). Hence, we conclude that our suggested two-level computation technique tremendously reduces the overall computational time.

3 Results

First we assume that the spin is injected along the perpendicular OZ-direction, and we investigate the dependence of τ on shear strain ε_{xy}. It is observed in Fig. 2 that the spin relaxation rate with ε_{xy} is dramatically reduced, and the corresponding τ is increased by orders of magnitude. The figure confirms that at higher temperature the phonon scattering rate increases to reduce τ as compared to that at lower temperature, for all values of ε_{xy}. Figure 2 also describes the inter- and intravalley scattering spin relaxation components, and we notice

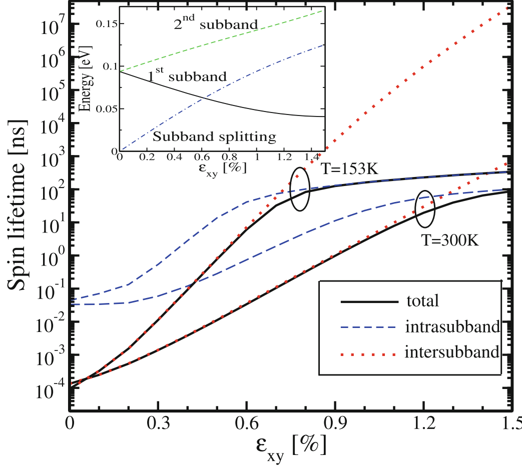

Fig. 2. Total spin relaxation time with components at two distinct temperatures. $t = 1.36\,\text{nm}$, $N_S = 10^{12}\,\text{cm}^{-2}$.

that at low to intermediate strain the major contribution to spin relaxation comes from the intersubband processes, whereas at higher stress the intrasubband component becomes significant. The physical reason for the spin lifetime enhancement by shear strain lies in the ability to completely remove the degeneracy between the two [001] valleys in a confined electron systems by ε_{xy} (inset of Fig. 2). The enhanced valley splitting makes the intervalley spin relaxation much less pronounced which results in a giant spin lifetime enhancement.

Next we study the spin lifetime dependence on the spin injection orientation. The surface roughness scattering matrix elements (M_{SR}), taken to be proportional to the product of the subband wave function derivatives at the interface, are shown in Fig. 3 for several injection orientations, when the additional valley splitting in unstrained films [6] is included. The spin injection orientation is described by the polar angle θ measured from the perpendicular OZ-axis towards the XY-plane. We find that, when θ increases, M_{SR} decreases (Fig. 3 inset). The dependence of M_{SR} on θ can be expressed as:

$$M_{SR}^2(\theta) \propto 1 + \cos^2\theta \cdot (\frac{p_x}{p_y})^2 \tag{7}$$

(p_x, p_y) is the in-plane momentum.

Figure 4 describes the variation of τ with ε_{xy}. Contrary to Fig. 2 we find that the increase of τ with ε_{xy} becomes less pronounced, but even in this case τ is boosted by almost two orders of magnitude. In accordance with Fig. 3 we also find in Fig. 4 that the spin relaxation rate (time) decreases (increases), when θ increases, thus τ reaches its minimum (maximum) value for an in-plane spin injection. The inset of Fig. 4 highlights the variation of τ with θ at a fixed stress point ($\varepsilon_{\text{xy}} = 0.5\,\%$). We note that the ratio of τ computed for two different

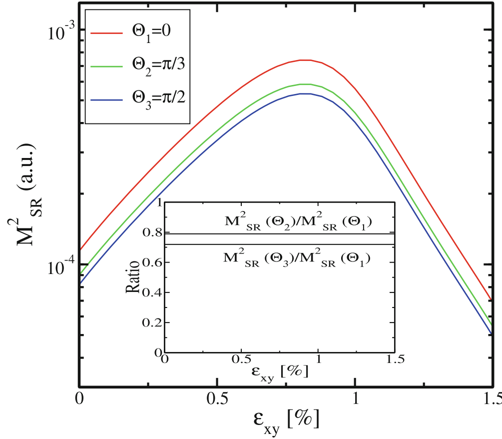

Fig. 3. Variation of intersubband spin relaxation matrix elements with ε_{xy}, $t = 1.36\,\text{nm}$, $k_x = 0.5\,\text{nm}^{-1}$, $k_y = 0.8\,\text{nm}^{-1}$. The inset shows the ratio.

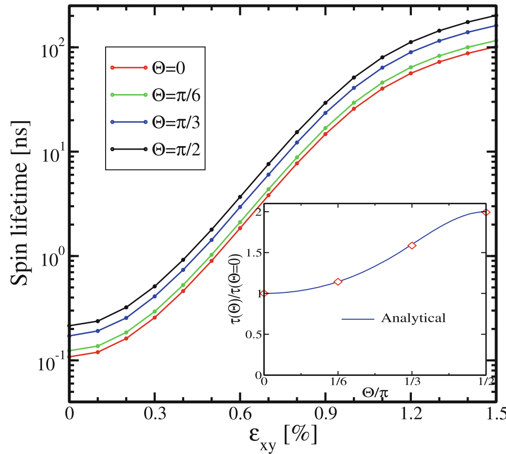

Fig. 4. Variation of τ with ε_{xy} with the spin injection orientation (θ) as parameter. $t = 1.36\,\text{nm}$, $N_S = 10^{12}\,\text{cm}^{-2}$, $T = 300\,\text{K}$. The inset shows the ratio, dots are simulation points.

injection directions at the same stress value does not depend on shear strain. An analytical expression describing the variation of τ with θ for a fixed ε_{xy} can be deduced from (7) by averaging M_{SR}^2 over the in-plane momentum direction, and can be expressed as, $\frac{1}{\tau(\theta)} \propto 1 + \cos^2\theta$. The simulated results and the analytical expression show a perfect agreement. We point out that a similar dependence of spin lifetime on the injection direction was also reported in bulk silicon [8], indicating that the spin lifetime only depends on the spin injection orientation relative to the valley orientation it is injected to. We conclude that the spin lifetime can be further increased, when spin is injected in-plane.

4 Summary

We have described a two-level parallelization scheme to calculate the spin lifetime in silicon. The proposed algorithm precalculates wave functions and energies (first level), and computes the spin relaxation rate by using the precalculated data (second level). These calculations are performed in parallel. We have analyzed the memory and computation time requirements for different parallelization configurations (pure MPI, hybrid MPI-OpenMP), and found that the precalculation step is best performed through a pure MPI scheme, whereas the spin relaxation calculations are efficiently performed by a hybrid approach due to memory demands. Finally, we have depicted that shear strain can boost the spin lifetime by orders of magnitude in thin silicon films. The spin lifetime is further enhanced, once spin is injected in-plane.

Acknowledgements. This work is supported by the European Research Council through the grant #247056 MOSILSPIN. The computational results presented have been partly achieved using the Vienna Scientific Cluster (VSC).

References

1. Huang, B., Monsma, D.J., Appelbaum, I.: Coherent spin transport through a 350 micron thick silicon wafer. Phys. Rev. Lett. **99**, 177209 (2007)
2. Li, J., Appelbaum, I.: Modeling spin transport in electrostatically-gated lateral-channel silicon devices: role of interfacial spin relaxation. Phys. Rev. B **84**, 165318 (2011)
3. Li, P., Dery, H.: Spin-orbit symmetries of conduction electrons in silicon. Phys. Rev. Lett. **107**, 107203 (2011)
4. Fischetti, M.V., Ren, Z., Solomon, P.M., Yang, M., Rim, K.: Six-band $\mathbf{k} \cdot \mathbf{p}$ calculation of the hole mobility in silicon inversion layers: dependence on surface orientation, strain, and silicon thickness. Phys. Rev. Lett. **94**, 1079 (2003)
5. Song, Y., Dery, H.: Analysis of phonon-induced spin relaxation processes in silicon. Phys. Rev. B **86**, 085201 (2012)
6. Osintsev, D., Sverdlov, V., Neophytou, N., Selberherr, S.: Valley splitting and spin lifetime enhancement in strained thin silicon films. In: Proceedings IWCE (2014) ISBN:9781479954346
7. Vienna Scientific Cluster: http://www.vsc.ac.at/systems/vsc-2/
8. Dery, H., Song, Y., Li, P., Zutic, I.: Silicon spin communication. Appl. Phys. Lett. **99**, 082502 (2011)