Copyright
by
Dax Michael Crum
2016

The Dissertation Committee for Dax Michael Crum
certifies that this is the approved version of the following dissertation:

Advanced modeling for end-of-the-roadmap CMOS and potential beyond-CMOS applications

Committee:

Leonard F. Register, Supervisor

Sanjay K. Banerjee

Emanuel Tutuc

Jack C. Lee

Allan H. MacDonald

Advanced modeling for end-of-the-roadmap CMOS and potential beyond-CMOS applications

by

Dax Michael Crum, B.S.E.E.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2016

For my parents, the first teachers I ever had

Acknowledgments

My journey as a student of physics and engineering required the help and guidance of many people.

First and foremost I acknowledge my thesis adviser L. Frank Register. Dr. Register ignited my love of the quantum with a pen and paper sitting many, many hours at my desk. Frank believed in me as a scientist before I knew what the Schrödinger equation even *was*. I have the deepest respect and admiration for Frank Register.

I thank my parents Robert H. and Kimberly W. Crum for their immense sacrifices on my behalf. I acknowledge my older brother Conan C. Crum from whom I constantly find a steady role model. I cherish my sisters Cassidy C. Cavin, Mariko C. Rajamand, and Chevon C. Herbert for their love and support. I am proud to mention my family Derek T. Carr, Amy B. Crum, Sina Rajamand, and Charles C. Herbert. I note my little nieces and nephews C. Christopher, Bryce M., and Audrey E. Crum, Jack R. Cavin, and Layla and Roya Rajamand, who all bring me much joy.

I thank Samir Lounis for giving me the opportunity to work with him

in the prestigious Forschungszentrum Jülich (FZJ), and teaching me density-functional theory. Samir is among the finest teachers I've known. I also thank Stefan Blügel who devoted many hours to our research on skyrmions. My time spent abroad was a turning point in my career, and I am grateful for the tutelage of Samir and Stefan who made this fantastic collaboration possible.

I thank each of my committee members, beginning with Sanjay K. Banerjee, who financially supported me upon entrance to UT graduate school. This generosity gave my Ph.D. its start and we never looked back. I thank Allan H. MacDonald for all the lessons on physics—and life—that he taught me over four courses and four years. Allan is an inspiration to me and many others. I thank Jack C. Lee and Emanuel Tutuc for teaching me the fundamentals of MOSFETs and for many insightful conversations.

I thank other teachers who impacted my life. I mention Bill Dillon from Horizon High School who, when I was 15 years old, first suggested I consider being an engineer. From the University of Wyoming, I mention Steven Barrett, Stanislaw Legowski, and Sadrul Ula. I must thank John Spitler who taught me calculus and helped me muster the courage to pursue graduate school. From UT, I acknowledge Seth Bank, Ed Yu, Willy Fischler, and Steven Weinberg. From FZJ, I thank Nikolai Kiselev, Manuel dos Santos Dias, Gustav Bihlmayer, Robert Jones, H. Long Nguyen, Bo Persson, and, especially, Phivos Mavropoulos. I thank Ansgar Liebsch who generously lent me his guitar while I was living in Jülich and working at FZJ.

I must mention some of my Ph.D. colleagues. First must be John K.

David, my predecessor upon whose Monte Carlo work and software I continued. I thank with sincere gratitude Dharmendar Reddy Palle, who taught me more about physics and programming than almost anyone. Dharmendar even programmed some solvers for our quantum corrections and some software to generate our FinFET visualizations (Figs. 2.1 and 3.1). I thank Dipanjan Basu for professional and research-related advice. I thank theorists Amith Valsaraj, Nitin Prasad, Christopher Reilly, Jiwon Chang, Priyamvada Jadaun, Urmimala Roy, Xuehao Mou, Xian Wu, Rik Dey, and Tanmoy Pramanik from UT. I thank experimentalists William Hsu, Andreas Hsieh, Hema Chandra Prakash Movva, Chris Corbet, Dave C. Dillon, Babak Fallahazad, Stefano Larentis, Kayoung Lee, Tanuj Trivedi, Scott D. Sifferman, Sangwoo Kang, Seyoung Kim, and Sushant Sonde. From FZJ, I thank Bendikt Schweflinghaus for his mentorship concerning DFT and magnetism. I also thank Boris Lorentz, Mohammed Bouhassoune, Juba Bouaziz, Bernd Zimmermann, Julen Ibanez Azpiroz, Flaviano J. Dos Santos, Felipe Guimaraes, Ashis Nandy, and Vasile Caciuc.

I am grateful for being the recipient of four highly distinguished fellowships. I thank the US National Science Foundation for my Graduate Research Fellowship Program (GRFP) and UT for my Microelectronics and Computer Development (MCD) and Cockrell School of Engineering (CSE) Fellowships. I also thank the German Academic Exchange Service (DAAD) for my *RISE*-Professional Fellowship, which allowed me to study at FZJ for nearly 9 months.

I thank Bhagawan Sahu and other *GLOBALFOUNDRIES INC., USA*

colleagues for our fruitful collaboration.

I must acknowledge Jean Toll, Melanie Gulick, and Ute Winkler, whose administrative work provided the organization to enable my research. I thank Kim Praderas for personal guidance.

I thank Steve Cea and Mark Stettler of Intel Corporation for the opportunity to further my career in science.

I give a special thanks to Roza Kotlyar, who told me to plot the as-calculated quantum-mechanical charge distribution of my FinFETs and compare them with my semi-classical charge distributions (Fig. 2.5 of Section 2.4.2).

I give one more special thanks to Philipp Rüßmann who helped me program the Python code to generate Fig. 4.11, my favorite visualization I have ever created as a scientist so far.

The utmost possible distinction, however, must be paid to my close friend and confidante Tiffany C. Sims, who edited even some of these very pages.

Dax Michael Crum

May, 2016

Austin, Texas USA

Advanced modeling for end-of-the-roadmap CMOS and potential beyond-CMOS applications

Publication No. ____________________

Dax Michael Crum, Ph.D.
The University of Texas at Austin, 2016

Supervisor: Leonard F. Register

End-of-the-roadmap CMOS devices are explored via particle-based ensemble semi-classical Monte Carlo (MC) methods employing quantum corrections (QCs) to address quantum confinement and degenerate carrier populations. The significance of such QCs is illustrated through simulation of n-channel III-V and Si FinFETs. Original contributions include our treatment of far-from-equilibrium degenerate statistics and QC-based modeling of surface-roughness scattering, as well as considering quantum-confined phonon and ionized-impurity scattering in 3D. Typical MC simulations approximate degenerate carrier populations as Fermi distributions to model the Pauli-blocking (PB) of scattering to occupied final states. To allow for increasingly far-from-equilibrium non-Fermi carrier distributions in ultra-scaled and III-V devices, we instead generate the final-state occupation probabilities used for PB by sampling the local carrier populations as a function of energy and energy valley. This process is aided by the use of fractional carriers or *sub-carriers*,

which minimizes classical carrier-carrier scattering. Quantum confinement effects are addressed through quantum-correction potentials (QCPs) generated from coupled Schrödinger-Poisson solvers, as commonly done. However, we use our valley- and orientation-dependent QCPs not just to redistribute carriers in real space, or even among energy valleys, but also to calculate confinement-dependent phonon, ionized-impurity, and surface-roughness scattering rates. Collectively, these quantum effects can substantially reduce and even eliminate otherwise expected benefits of considered $In_{0.53}Ga_{0.47}As$ FinFETs over otherwise identical Si FinFETs, despite higher thermal velocities in $In_{0.53}Ga_{0.47}As$.

Beyond-CMOS device concepts are also being considered for future applications. Thin-film sub-5 nm magnetic skyrmions constitute an ultimate scaling alternative for beyond-CMOS data storage technologies. These robust non-collinear spin-textures can be moved and manipulated by spin-polarized or *non*-spin-polarized electrical currents, which is extremely attractive for integration with current memory technologies. An innovative technique to detect isolated nano-skyrmions with a current-perpendicular-to-plane is shown, which has immediate implications for device concepts. Such a mechanism is explored by studying the atomistic electronic structure of the magnetic quasiparticles. The tunneling conductance is quite sensitive to spatial variations in the electronic structure, as a large atomistic conductance anisotropy up to ~20% is found for magnetic skyrmions in Pd/Fe/Ir(111) magnetic thin-films. This spin-mixing magnetoresistance effect possibly could be incorporated in future magnetic storage technologies.

Table of Contents

Acknowledgments **v**

Abstract **ix**

List of Tables **xiv**

List of Figures **xv**

Chapter 1. Introduction **1**

1.1 Goals of 21st century electronics 1

1.2 High-performance simulation of nano-scale materials, devices, and device physics 3

1.3 Dissertation overview 4

Chapter 2. Ensemble Monte Carlo for III-V and Si n-channel FinFETs considering non-equilibrium degenerate statistics and quantum-confined scattering **6**

2.1 Modeling ultra-scaled FinFETs using particle-based Monte Carlo methods 7

2.2 3D FinFET device and uncorrected MC simulator 12

2.2.1 Device structure 13

2.2.2 Baseline purely semi-classical Monte Carlo simulator . . 15

2.2.3 Silicon and III-V material models 17

2.3 Quantum-corrections for non-equilibrium degenerate statistics 19

2.3.1 The Pauli-blocking of scattering 19

2.3.2 Effects of degenerate statistics under equilibrium conditions 23
2.3.3 Effects of degenerate statistics under far-from-equilibrium conditions . . . 27
2.4 Quantum-corrections for electron quantum confinement 30
2.4.1 Calculating the quantum-correction potential 31
2.4.2 Real-space redistribution of charge under quantum confinement . . . 33
2.4.3 Modeling quantum-confinement-dependent phonon and ionized-impurity scattering . . . 36
2.4.4 Modeling quantum-confinement-dependent surface-roughness scattering . . . 41
2.5 Model comparisons . . . 44
2.5.1 Illustrative impact of included quantum corrections on III-V and Si FinFET performance . . . 44
2.5.2 Effects of satellite valleys on III-V electron transport . . 49
2.6 Discussion . . . 53

Chapter 3. Device scaling studies 55
3.1 Dimensional scaling study . . . 55
3.1.1 Device structure and results . . . 55
3.1.2 Discussion . . . 59
3.2 Dielectric scaling study and the impact of the gate oxide effective mass on channel quantum confinement . . . 60
3.2.1 Alternate gate oxide materials for scaling CMOS 61
3.2.2 Gate oxide dielectric constant effects . . . 63
3.2.3 Gate oxide effective mass effects . . . 65
3.2.4 Gate oxide trade-off: alumina vs. hafnia . . . 67
3.2.5 Combined gate oxide stacks . . . 68

Chapter 4. Perpendicular Reading of Single Confined Magnetic Skyrmions 71
4.1 Single magnetic skyrmions: the ultimate bit . . . 72
4.2 Material systems and computational procedure . . . 76
4.2.1 Charge density and angular momentum of ferromagnetic thin-films . . . 78

4.2.2 Non-collinear inhomogeneity in nano-skyrmions 79
4.3 Electronic structure of isolated confined skyrmions 81
4.3.1 Angular dependence of the local density-of-states in nano-skyrmions 83
4.3.2 Two-atom extended Alexander-Anderson model 87
4.3.3 Electronic structure in 1.7 nm Pd/Fe/Ir(111) skyrmions 93
4.3.4 Electronic structure in 2.2 nm Pd/Fe/Ir(111) skyrmions 95
4.3.5 Electronic structure in Pd/Pd/Fe/Ir(111) skyrmions . . 98
4.4 Tunneling spin-mixing magnetoresistance (TXMR) 102
4.4.1 TXMR in larger 2.2 nm skyrmions 104
4.4.2 All-electrical skyrmion detection 106
4.5 Discussion 107
4.5.1 Skyrmion racetracks for dense magnetic memories . . . 109
4.6 Methods 113
4.6.1 Computational details 113
4.6.2 Thin-film slab configurations 113
4.6.3 Calculating whole skyrmions within density-functional theory 114

Chapter 5. Conclusion **115**
5.1 Dissertation recap 115
5.2 Future work 118

Appendix **122**

Bibliography **125**

Vita **142**

List of Tables

2.1 **ON-state performance by model.** The transconductance g_{M} and the drain current I_{ON}, sampled at $V_{\mathrm{DS}} = 600$ mV and $V_{\mathrm{ON}} = V_{\mathrm{GS}} - V_{\mathrm{T}} = 350$ mV. Both have been normalized by the fin perimeter, $(2T_{\mathrm{fin}} + W_{\mathrm{fin}})$. 47

3.1 **Effects of the variation of the gate oxide dielectric constant on FinFET performance.** A larger gate oxide dielectric constant improves ON-state transconductance and subthreshold behavior. 63

3.2 **Effects of the variation of the gate oxide conduction band effective mass on FinFET performance.** A larger gate oxide conduction band effective mass improves ON-state performance by reducing the quantum confinement experienced by electrons in the FinFET channel. 65

3.3 **Effects of the variation of the gate oxide material stack on FinFET performance.** Al_2O_3, with a heavier effective mass, is shown to compete effectively against HfO_2 in performance despite its lower dielectric permittivity. 68

4.1 **FM state of Pd/Fe/Ir(111).** The atomic number (Z), the total electronic charge (n), and the charge transfer ($\Delta n = n - Z$) are given in units of electrons. The spin-moment ($\mathrm{M_S}$) and the orbital-moment ($\mathrm{M_L}$) are given in units of μ_{B}. 78

4.2 **FM state of Pd/Pd/Fe/Ir(111).** The atomic number (Z), the total electronic charge (n), and the charge transfer ($\Delta n = n - Z$) are given in units of electrons. The spin-moment ($\mathrm{M_S}$) and the orbital-moment ($\mathrm{M_L}$) are given in units of μ_{B}. 79

List of Figures

2.1 **FinFET structure.** **(a)** Side view of the simulated device with relevant dimensions. Outer materials are shown transparent to visualize the semiconductor fin. **(b)** Edge view with cross-sectional clip to show different material regions. We vary the work functions in the $\delta\Phi_{\rm m}$ metal contacts to set the injection boundary conditions in the S/D such that there are flat-band conditions at the interface. 14

2.2 **Effects of degeneracy in $In_{0.53}Ga_{0.47}As$.** MC simulated electron distributions (symbols) with the PB of scattering ($n_{\rm PB}^{\rm MC}$) and without ($n_{\rm CL}^{\rm MC}$) in the Γ-valley of $In_{0.53}Ga_{0.47}As$. The results are shown under flat-band and equilibrium conditions, with reference distributions (solid lines) calculated directly from Fermi-Dirac ($n_{\rm PB}$) and Boltzmann ($n_{\rm CL}$) statistics, respectively. The distributions have been normalized to the peak electron density in the classical limit as a reference. There is 40% less Γ-valley charge (area under the curve) in the degenerate case. 24

2.3 **Impact of sub-carriers on classical carrier-carrier interactions.** MC simulated electron distributions in the Γ-valley of $In_{0.53}Ga_{0.47}As$ under equilibrium with different sub-carrier factors. The product of the 300 K Fermi-Dirac distribution and DOS is shown for reference ($n_{\rm FD}$). With $N_{\rm sub} = 1$, the unwanted thermalizing effects of classical carrier-carrier scattering drive the MC electron distribution to a high-temperature-like shape. Increasing the sub-carrier factor to $N_{\rm sub} = 10$ and then $N_{\rm sub} = 100$, the device distributions approach the theoretical expectation in shape and average kinetic energy $\langle E \rangle$. These distributions are normalized to have the same area under each curve. (The actual electron concentration in the Γ-valley is not conserved due to the occupation of peripheral valleys as a function of sub-carrier factor.) 26

2.4 **Non-equilibrium electron distributions.** **(a)** FinFET edge view showing the device slice containing the potential energy barrier-top. **(b)** Potential profile showing the location of the barrier-top. **(c)** Normalized MC simulated electron distributions sampled at the potential energy barrier-top to the channel in the Γ-valley of $In_{0.53}Ga_{0.47}As$ under non-equilibrium. The much larger forward-going contribution (+) versus backward-going (−) is consistent with a high injection efficiency into the channel. **(d)** FinFET edge view showing near the drain end of the channel. **(e)** Potential profile showing the location of the drain end. **(f)** Normalized MC simulated electron distributions again but sampled near the drain. The forward-going distribution (+) is the superposition of the drain near-equilibrium charge distribution and the quasi-ballistic population injected from the source. 29

2.5 **Real-space redistribution of charge under quantum confinement.** **(a)** $In_{0.53}Ga_{0.47}As$ FinFET with the channel slice containing the charge distributions of **b** and **c**. **(b)** At the overdrive gate voltage of 0.35 V above threshold, the classical electron charge density is attracted to the channel surface (top left, ρ_{CL}). Including QCPs repels carriers from the interface (top middle, ρ_{QC}) although degeneracy effects smear the spatial distribution relative to the as-calculated equilibrium Boltzmann-weighted quantum-mechanical charge density (top right, ρ_{QM}). **(c)** Here ρ_{QC} (bottom middle) more closely resembles ρ_{QM} (bottom right) at threshold, but effects of degeneracy and non-equilibrium behavior are still evident. 35

2.6 **Modeling quantum-confined scattering using adjusted bulk rates.** The scattering rate for short range nominally randomizing elastic scattering within the first sub-band for 1D quantum confinement in **(a)** infinite square wells of varying width w, and **(b)** infinite one-sided triangular wells of varying fixed interface-normal field $\mathbf{E}$. The curves are labeled by how many sub-bands are picked up within the given energy range for each respective confinement condition. The scattering rates oscillate about the bulk rate (bulk) when the carrier energy is referenced to the expectation value of the electrostatic potential, more closely approaching the bulk limit with reduced confinement. For increased confinement, the scattering rate for low-energy first sub-band carriers is increased substantially. By using the bulk scattering rate but starting at the quantum-corrected band-edge V_{QC} (**a**, quantum-corrected rate), our approximation captures the overall increase in the basic scattering rate, if missing the oscillations and being somewhat conservative for low-energy carriers. By comparison, simply shifting the zero of the bulk rate by the quantum-confined band-edge energy (**a**, shifted bulk rate) results in a much larger error. 38

2.7 **Selecting the choice of final state for quantum-confined scattering.** Quantum-confined intervalley scattering from the Γ-valley to an L-valley via absorption of an optical phonon of energy $\delta E = \hbar\omega$ in $\mathrm{In_{0.53}Ga_{0.47}As}$, whereby the choice of final state scattering rate becomes a function of the total energy $E^{\Gamma}_{\mathrm{QC}} = \varepsilon^{\Gamma}_{\mathbf{k}} + V^{\Gamma}_{\mathrm{QC}}$, where the QCP V^{Γ}_{QC} raises the initial and, thus, final state energies relative to the uncorrected valley edges alike. 40

2.8 **Enhanced scattering rates due to quantum confinement.** Calculated $\mathrm{In_{0.53}Ga_{0.47}As}$ Γ-valley scattering rates comparing the bulk rate against those including quantum-confined scattering (+QCS) and additional surface-roughness scattering (+SR) in the device channel center of a 6 nm wide fin under equilibrium conditions at threshold. Not only are the rates increased due to quantum confinement, but intervalley scattering, and thus the intervalley transfer (IVT) of electrons, occurs at lower kinetic energies versus the bulk case. 42

2.9 **(100)/⟨100⟩ FinFET performance comparison of $In_{0.53}Ga_{0.47}As$ versus Si with varying quantum-corrected models.** **(a)** As the Pauli-blocking (PB) of scattering, and then quantum-correction potentials (+QCP) are added, as well as QC-dependent phonon (+QCS) and surface-roughness scattering (+SR), the performance of the $In_{0.53}Ga_{0.47}As$ FinFET device (left) is substantially degraded relative to classical (CL) expectations. **(b)** Si is shown to be more robust against carrier degeneracy and quantum-confinement effects (right), although its performance is also moderated compared to classical expectations. 46

2.10 **The role of satellite valleys in $In_{0.53}Ga_{0.47}As$ transport.** The individual curves are labeled by their doping density and are $In_{0.53}Ga_{0.47}As$ channels unless noted as Si. **(a)** With a larger injection velocity, the $In_{0.53}Ga_{0.47}As$ device having no intervalley scattering (no IV) outperforms the device including the peripheral valleys, although it has a smaller quantum capacitance. It also outperforms the $In_{0.53}Ga_{0.47}As$ device having a larger doping density of 5×10^{19} cm^{-3} and even crosses-over Si at larger gate biases. **(b)** Si, however, still has the largest ON-current at the given drive voltage determined by ITRS predictions for future CMOS. 51

3.1 **Edge view of simulated FinFET transistor showing each material region.** The strain relaxed buffer provides a defect-free fin-channel, while the shallow trench isolation electrically isolates adjacent FinFETs. The source and drain reservoirs are doped with a grading falling off as 1 decade/nm while the fin-channel is otherwise undoped. Including realistic spacer and metal regions effectively captures the relevant electrostatics. All dimensions are shown in nm. 56

3.2 **Comparing ON-state performance.** Transfer characteristic $I_{\rm DS}$ vs. $V_{\rm GS} - V_{\rm T}$ for FinFETs from Fig. 3.1 with various channels. The right figure is a zoomed view of the overdrive voltage $V_{\rm ON} = V_{\rm GS} - V_{\rm T} = 350$ mV (with end-of-the-ITRS[65] supply $V_{\rm DD}$ and threshold $V_{\rm T}$ voltages 0.6 V and 0.25 V). Si ⟨100⟩ exhibits the largest ON-state current $I_{\rm ON}$, while Si ⟨110⟩ and $In_{0.53}Ga_{0.47}As$ have nearly identical signals. The $I_{\rm ON}$ advantage for Si ⟨100⟩ and disadvantage for $In_{0.70}Ga_{0.30}As$ are a result of more and less abrupt turn-on, respectively, beyond the assumed threshold current of $I_{\rm DS}\big|_{V_{\rm T}} = 0.01$ mA/μm. 57

3.3 **Dependence of channel peak transconductance on source/drain doping.** $In_{0.53}Ga_{0.47}As$ exhibits higher peak g_M than Si ⟨100⟩ does at the same doping. However, the advantage decreases with increasing doping, and greater doping is available in Si devices. 58

3.4 **Channel length scaling.** Holding all parameters constant including the 6 nm fin width, the device of Fig. 3.1 is scaled via L_{CH} = 21, 18, and 15 nm respectively. Si (100)/⟨100⟩ outperforms $In_{0.53}Ga_{0.47}As$ in all metrics. Setting a benchmark for electrostatic control of $S < 70$ mV/dec and DIBL< 70 mV/V, a design rule of $L_{CH} = 3W_{fin}$ is more than sufficient for Si, while $In_{0.53}Ga_{0.47}As$ requires somewhat longer channels, approximately $L_{CH} = 3(W_{fin} + 1)$. 59

3.5 **FinFET device scaling.** The device from Fig. 3.1 is then scaled via $L_{CH} = 3W_{fin}$ according to possible device nodes at 2016 (L_{CH} = 15 nm, W_{fin} = 5 nm), 2019 (L_{CH} = 12 nm, W_{fin} = 4 nm), and 2022 (L_{CH} = 9 nm, W_{fin} = 3 nm). ON-state drive current remains comparable between Si (100)/⟨110⟩ versus $In_{0.53}Ga_{0.47}As$, while the Si (100)/⟨100⟩ device boasts a roughly 3× greater drive current. The Si devices continue exercising somewhat better short channel control in terms of subthreshold swing S and DIBL, with both metrics becoming worse for shorter channels. In terms of I_{ON} it might prove more feasible to move to Si ⟨100⟩ channels rather than integrate $In_{0.53}Ga_{0.47}As$ material processes. Currently, nMOS Si ⟨110⟩ channels are incorporated to match pMOS technologies, which for which Si ⟨110⟩ orientation is preferred. 60

3.6 **Varying oxide dielectric constant. (a)** I_{DS} vs. $V_{GS} - V_T$ for otherwise identical FinFETs having gate oxide dielectric ε_R = 7.8 (gray) compared to ε_R = 22.3 (black). **(b)** Increasing the dielectric constant (from gray to black) induces a larger channel inversion charge density Q_{ch} (right-axis) and capacitance $C_{ch} = dQ_{ch}/dV_{GS}$ (left-axis). 64

3.7 **Varying oxide effective mass. (a)** Transfer curves comparing gate oxide effective mass m^*_{ox} = 0.4 (gray) to m^*_{ox} = 0.08 (black). **(b)** Decreasing the gate dielectric's effective mass increases the average quantum-confinement energy (left-axis) and occupation of heavy-mass peripheral valley states (right-axis). 66

3.8 **Trade-offs between Al_2O_3 and HfO_2 gate dielectrics.** Al_2O_3 (black), with its smaller dielectric constant than HfO_2 (gray), generates a smaller inversion charge density Q_{ch} (right-axis). However, with its heavier effective mass and smaller quantum confinement, Al_2O_3 retains more Γ-valley carriers than HfO_2 (left-axis). 69

4.1 **Out-of-plane detection of nano-skyrmions.** **(a)** Illustrative heterostructure cross-section for the perpendicular reading of single nano-skyrmions. Due to energy-dependent spin-mixing perturbations to the atomistic electronic structure as a function of position within skyrmions, the electric current relation $i_{\mathrm{FM}} \neq i_{\mathrm{Sk}}$ holds, and therefore the magnetic data information can be sensed in a CPP-geometry. **(b)** Illustrative STM-spectroscopy experiment of fcc-Pd/Fe overlayer on single-crystal fcc-Ir(111) bulk substrate. The tunneling conductance is modified by the combined effects of local magnetic non-collinearity and substrate-induced spin-orbit interaction. For similar physical reasons as (**a**), the tunneling conductance at position x_0 is different from that at position x_1. 77

4.2 **Real-space relaxation of nano-skyrmions with increasing size.** **(a-c)** Plots of axisymmetric cycloidal spin-whirls inside a magnetically active Fe-layer centered about increasingly larger skyrmionic defects in fcc-Pd/Fe overlayer on fcc-Ir(111) bulk substrate. Confining spins in the FM-background are shown transparent. We define θ as the typical polar angle with the vertical and dθ as the difference in polar angle between adjacent pairwise atoms. **(d-f)** Again but in fcc-Pd/Pd/Fe overlayer on fcc-Ir(111) bulk substrate. The color bar in (**a**) represents the magnitude of the z-component of the magnetization for each spin in (**a**-**f**). 81

4.3 **Alexander-Anderson prediction for *d*-*d* hybridization.** Beginning from dθ=0°, ferromagnetic coupling imposes a hybridization between the two d_{z^2} orbitals localized at the same energy but on neighboring atoms. This produces a splitting into a bonding- and antibonding-state (green curve). Upon rotation, the atoms eventually become antiferromagnetically coupled (black curve). The change in the LDOS can be qualitatively estimated with a cos(dθ) fitting parameter. 89

4.4 **Energy- and spatially-dependent ΔLDOS in skyrmions.** **(a)** Fe-layer atom-by-atom *ab initio* results (blue dots) of the total change in LDOS as a function of θ in a 1.7 nm diameter skyrmion in Pd/Pd/Fe/Ir. ΔLDOS is given relative to the background-FM and is measured at the Fermi energy while including SOI. These results are then fitted to $f(\theta)$ (solid curve). The $f(\theta)$ in (**a,b**) refers to Eq. 4.17, while the numbers labeling each dot respresent the corresponding atom in the inset extending radially from the skyrmion core. **(b)** Again, Fe-layer first-principles results (blue dots) of the atomistic change in LDOS in Pd/Pd/Fe/Ir and fitted against $f(\theta)$ (solid curve), but for a larger 2.2 nm skyrmion and measured at the resonant energy $E_{\mathrm{R}} = 0.5$ eV. **(c)** We plot the ΔLDOS as a function of dθ given by our two-atom Alexander-Anderson model (black curve) to isolate the non-collinear component of the ΔLDOS. While the main contribution comes from cos(dθ) (red-dashed, referring to Eq. 4.16), a higher-order term is required to complete the fit (blue-dashed). **(d)** Rotation parameters as found from the first-principles calculations, referring to (**a**) and (**b**). . 92

4.5 **Electronic structure of a realistic $D_{\mathrm{Sk}} \approx 1.7$ nm skyrmion in Pd/Fe/Ir.** **(a)** LDOS in the magnetically-active Fe-layer resolved into minority (solid) and majority (dashed) spin-channels. The resonance-peak near $E \approx 0.5$ eV in the FM-background (green) shifts in energy when approaching the center of the skyrmion (black). **(b)** The modification of the electronic structure in Fe contributes to a strong resonance in the LDOS in vacuum via hybridization through surface Pd-states. Arbitrary units are used so as to include in the same plot the nature of the Pd-d_{z^2} surface-state (black-dashed), whose resonance-peak near 0.5 eV survives in the vacuum. **(c)** Illustrative legend for (**a**, **b**) where the numbered spheres represent a line extending radially from the skyrmion's center. The vacuum domains are represented by empty spheres. FM-Fe, FM-Pd, and FM-Vac represent the unperturbed background ferromagnet. 94

4.6 **LDOS of a $D_{\mathbf{Sk}} \approx$ 2.2 nm skyrmion in Pd/Fe/Ir.** (**a**) Electronic structure in the magnetically-active Fe-layer resolved into minority (solid) and majority (dashed) spin-channels. The splitted-structure of the FM-background (green) is modified due to quasi-AFM interactions when approaching the center of the skyrmion (black). (**b**) The alteration of the LDOS in Fe contributes to a strong resonance in vacuum via hybridization through surface Pd-states. Arbitrary units are used so as to include in the same plot the nature of the Pd-d_{z^2} surface-state (black-dashed), whose resonance-peak near 0.5 eV survives in the vacuum. (**c**) Illustrative legend for (**a**, **b**) where the numbered spheres represent a line extending radially from the skyrmion's center. The vacuum domains are represented by empty spheres. FM-Fe, FM-Pd, and FM-Vac represent the unperturbed background ferromagnet. 96

4.7 **LDOS of a $D_{\mathbf{Sk}} \approx$ 1.7 nm skyrmion in Pd/Pd/Fe/Ir.** (**a**) LDOS in the magnetically-active Fe-layer resolved into minority (solid) and majority (dashed) spin-channels. The resonance-peak near $E \approx 0.5$ eV in the FM-background (green) shifts in energy when approaching the center of the skyrmion (black). (**b**) The modification of the electronic structure in Fe impacts the vacuum LDOS via hybridization through surface Pd-states. Arbitrary units are used so as to include in the same plot the nature of the Pd-d_{z^2} surface-state (black-dashed), whose features at negative energies survive in the vacuum. (**c**) Illustrative legend for (**a**, **b**) where the numbered spheres represent a line extending radially from the skyrmion's center. The vacuum domains are represented by empty spheres. FM-Fe, FM-Pd, and FM-Vac represent the unperturbed background ferromagnet. 99

4.8 **LDOS of a $D_{\mathbf{Sk}} \approx$ 2.2 nm skyrmion in Pd/Pd/Fe/Ir.** (**a**) Electronic structure in the magnetically-active Fe-layer resolved into minority (solid) and majority (dashed) spin-channels. The splitted-structure of the FM-background (green) is modified due to quasi-AFM interactions when approaching the center of the skyrmion (black). (**b**) The alteration of the LDOS in Fe impacts the vacuum structure via hybridization through surface Pd-states. Arbitrary units are used so as to include in the same plot the nature of the Pd-d_{z^2} surface-state (black-dashed). (**c**) Illustrative legend for (**a**, **b**) where the numbered spheres represent a line extending radially from the skyrmion's center. The vacuum domains are represented by empty spheres. FM-Fe, FM-Pd, and FM-Vac represent the unperturbed background ferromagnet. 101

4.9 **Tunneling spin-mixing magnetoresistance.** **(a)** Energy-resolved TXMR signals measured at the skyrmion's core comparing the effects of SOI in a $D_{\mathrm{Sk}} \approx 1.7$ nm skyrmion in Pd/Fe/Ir. **(b)** TXMR signals again, but for a line of atoms extending radially from the center of the skyrmion, illustrating the spatial dependence of the effect. **(c)** Illustrative legend for (**a**, **b**) where the spheres are numbered and color-coded to identify the appropriate curves to the corresponding vacuum domains above the skyrmion. **(d-f)** Plots corresponding to (**a**-**c**) but in the Pd/Pd/Fe/Ir system. 105

4.10 **Spacially-varying TXMR in 2.2 nm skyrmions.** **(a)** We plot the TXMR in vacuum for Pd/Fe/Ir beginning in the core of the skyrmion (Vac-0, black) and then moving radially outwards (Vac-1, red; Vac-2, blue; Vac-3, navy) towards the confining FM environment. The vacuum resonance near 0.5 eV (see Fig. 4.6b) contributes a strong peak in the TXMR near the same energy. **(b)** Again but in Pd/Pd/Fe/Ir, where the peak-signal strength appears near -0.8 eV, as predicted by Fig. 4.8b. 106

4.11 **Perpendicular reading of single confined magnetic skyrmions.** Expected STS-signal when approaching a skyrmionic defect in the single-Pd case. The electrical contrast has been projected onto the plane below the skyrmion. Near the injection energy $eV_{\mathrm{bias}} \approx 0.5$ eV, there are about $\sim 20\%$ fewer tunneling states in the skyrmion's core compared to the unperturbed FM-environment. This increases the local magnetoresistance, allowing for the reliable spin-averaged electrical detection of skyrmions in a CPP-geometry. 107

4.12 **Concept spin-mixing magneto-memory versus DRAM.** **(a)** 2×2 1-T 1-C DRAM unit. In a random access memory, any memory cell can be accessed in roughly the same time since the WLs and BLs are connected to each bit in parallel. The ideal minimum packing footprint goes as $4f^2_{\mathrm{DRAM}}$/bit in units of area, where f_{DRAM} is the node generation. **(b)** Spin-mixing racetracks are not random access but acquire a latency associated with moving the bits sequentially out of the way to reach the requested address. However, with realistic skyrmion velocities, the access time could be roughly the same as in DRAM, while dramatically increasing the packing density. 111

Chapter 1

Introduction

1.1 Goals of 21st century electronics

Imagine a (not-too-distant) future where, e.g., (i) self-driving cars buzz you to-and-from work with aggregate accident-related fatalities approaching 0.000%, (ii) cell phones and smart devices need only be charged once every *other* year, and (iii) genetic birth defects are repaired inside the embryonic DNA on a gene-protein level *within the womb*. Could such astounding ideas ever really become a reality? Just like magnetic resonance imaging (MRI) and WiFi-enabled debit cards, such possibilities become foregone conclusions—provided that we have sufficiently advanced transistors.

Transistors are the building blocks of electronics. Improve the single transistor, and you can create a better electronic "building". In the 1960s, we had a transistor building block allowing us to build what would now seem analogous to small shanties, like the first 8-bit adder. What kind of transistor building block is required to build the electronic "skyscrapers" of (i–iii) and

more?

- *Ultra-fast speed.* Self-driving cars must process information not only related to their own environments, but also relay and communicate their status to other nearby cars, which could easily number in the hundreds within a few crowded city blocks. Faster devices come by minimizing transistor switching times, charging delay times, and acquired interconnect latency. Transistors employing high-mobility channels (such as III-Vs) possibly could provide faster operation relative to those employing Si due to faster light effective mass carriers, and associated higher mobilities and higher thermal velocities.[1]

- *Ultra-low power.* Cell phones and other smart devices consume electrical power. Reducing the power consumed by the single transistor building block vastly reduces the power consumed by the entire aggregate device considered. Transistor power consumption is reduced when, e.g., the transistor supply voltage is reduced in concert. Unfortunately, however, voltage scaling of Si CMOS has come to a near stand-still due to fundamental scaling physics[2] (e.g., no engineer can scale the thermal voltage $k_{\mathrm{B}}T$ below 25.9 mV at 300 K). To achieve nearly lossless devices, alternate phases of matter such as, e.g., superconducting condensates[3] or tunnel-devices[4] may need to be incorporated.

- *Ultra-small size.* Of all the *Gedankenexperimente* above, (iii) may be the most challenging to realize. Each double-helixed DNA strand contains

some 3 billion base pairs for protein encoding, but has a diameter of roughly 20 Å.[5] To have any hope of gene-level repair, the nanoelectronic devices not only must be able to search the DNA strands for irregularities, but the devices themselves must have at least *some* minimum feature of the order of the DNA diameter to enable protein-level manipulation of the genetic material. Materials and structures alien to Si CMOS, for instance topologically-protected magnetic skyrmions of diameter 1-2 nm, might prove useful in such an environment.[6]

To achieve tomorrow's devices will require investigation of multiple avenues of research, including improving now ubiquitous CMOS logic, as well as developing beyond CMOS capabilities.

1.2 High-performance simulation of nano-scale materials, devices, and device physics

Ultimately, technology moves forward when new devices are fabricated, tested, and reliably manufactured in industrial laboratories. To this end, theoretical work and experimental work are the perfect counterbalances of engineering and scientific exploration. A single great experimental result can change our entire physical understanding and lead to a new theoretical framework (e.g., the photo-electric effect). Certainly the reverse is true also. One or a few theorists can provide intuition for innumerable experimentalists to follow by developing new theory (e.g., relativity or quantum mechanics) or, on a much less grand scale, creating a new or improved simulation tool.

Theoretical modeling is imperative to the advancement of the state-of-the-art. Not only does simulation provide a fast and effective manner by which to explore complex experimental design spaces, it also provides important insights into device and material physics which can only be probed through theoretical means. It is toward the area of high-performance theoretical modeling and simulation that the work of my thesis is aimed.

1.3 Dissertation overview

For my thesis, I have followed a two-pronged approach, one directed toward improvement of current field-effect effects transistors until the exhaustion of end-of-the-roadmap CMOS (Chapters 2 and 3), and one directed toward potential beyond CMOS materials and operating principles (Chapter 4):

- Chapter 2 develops advanced quantum corrections applicable to semi-classical ensemble Monte Carlo methods to predictively model ultra-scaled Si and III-V transistors. The main reference for this chapter is Ref. 7.

- Chapter 3 describes device modeling studies of dimensional scaling and also gate oxide scaling in high-performance FinFETs, using the methods described in Chapter 2. The main reference for this chapter is Ref. 8.

- Chapter 4 focuses on the electronic properties of single confined magnetic skyrmions and how they possibly can be incorporated in future memory technologies. The main reference is Ref. 6.

Finally, Chapter 5 summarizes the work of my dissertation and provides some ideas for future related work.

Chapter 2

Ensemble Monte Carlo for III-V and Si n-channel FinFETs considering non-equilibrium degenerate statistics and quantum-confined scattering

Particle-based ensemble semi-classical Monte Carlo (MC) methods employ quantum corrections (QCs) to address quantum confinement and degenerate carrier populations to model tomorrow's ultra-scaled MOSFETs. Here a simulator is presented with the most complete treatment of quantum confinement and carrier degeneracy effects in a three-dimensional (3D) MC device simulator to date, and the significance of these effects is illustrated through simulation of n-channel Si and III-V FinFETs. Original contributions include our treatment of far-from-equilibrium degenerate statistics and QC-based modeling of surface-roughness scattering, as well as considering quantum-confined phonon and impurity scattering in 3D. Typical MC simulations approximate degenerate carrier populations as Fermi distributions to model the Pauli-blocking

(PB) of scattering to occupied final states. To allow for increasingly far-from-equilibrium non-Fermi carrier distributions in ultra-scaled and III-V devices, we instead generate the final-state occupation probabilities used for PB by sampling the local carrier populations as a function of energy and energy valley. This process is aided by the use of fractional carriers or sub-carriers, which minimizes classical carrier-carrier scattering. Quantum confinement effects are addressed through quantum-correction potentials (QCPs) generated from coupled Schrödinger-Poisson solvers, as commonly done. However, we use our valley- and orientation-dependent QCPs not just to redistribute carriers in real space, or even among energy valleys, but also to calculate confinement-dependent phonon, impurity, and surface-roughness scattering rates. FinFET simulations are used to illustrate the contributions of each of these QCs. Collectively, these quantum effects can substantially reduce and even eliminate otherwise expected benefits of considered $In_{0.53}Ga_{0.47}As$ FinFETs over otherwise identical Si FinFETs, despite higher thermal velocities in $In_{0.53}Ga_{0.47}As$.

2.1 Modeling ultra-scaled FinFETs using particle-based Monte Carlo methods

Multi-gate metal-oxide-semiconductor-field-effect-transistors (MOSFETs) have supplanted planar MOSFETs as the clear device choice for future integrated circuit technology. The three-dimensional (3D) fin-shaped MOSFET, or FinFET,[9] is electrostatically superior[10,11] to its planar relatives and already directing current technologies and future complimentary MOS

(CMOS) scaling.[12,13] In addition, high mobility III-Vs are being considered as channel replacements for Si.[1,14] $In_{0.53}Ga_{0.47}As$, which is lattice-matched to fabrication-friendly InP,[15] is being considered to provide a drive-current boost via light-effective mass carriers with associated large thermal injection velocities.[16,17]

Modeling such devices presents challenges for predictive device simulators, which are needed to optimize the large design space and estimate future scaling benefits. Alternate channel materials and associated transport physics require a microscopic description of their behavior. Quasi-ballistic transport cannot be completely described by continuum diffusive models in these devices,[18–20] yet scattering remains crucial, even as channel lengths are scaled well below 22 nm.[21–23] Certainly fully coherent methods, such as non-equilibrium Green function (NEGF) techniques,[24] have demonstrated their value to studying such systems[25–28] and represent the reference standard in the ballistic limit. However, upon the inclusion of scattering in realistic device geometries, pure quantum methods can become computationally impractical for many applications. Non-randomizing polar optical phonon scattering, which dominates Γ-valley transport in III-V channels as considered here, still has not been achieved in 3D NEGF simulations. Therefore it still remains important to extend the validity of semi-classical methods via so-called quantum corrections (QCs) to model these nanoscale devices while maintaining reasonable computational efficiency.

For these reasons, particle-based ensemble semi-classical Monte Carlo

(MC) remains a benchmark in semiconductor device research. It allows modeling of various distinct scattering mechanisms (including non-randomizing processes) and consideration of complex device geometries. MC is known to predictively model diffusive through ballistic transport including non-local field effects such as velocity overshoot. MC, however, suffers from its own drawbacks. Traditional particle-based MC is rigorous only in large systems where the carrier distributions are well-approximated by the bulk energy dispersion relations and scattering rates. Cutting-edge electron devices, however, often go well beyond these limits. In today's maximally-doped source and drain (S/D) transistor reservoirs and above threshold in the channel, degenerate carrier populations must be considered, along with associated Pauli-blocking (PB) of scattering. FinFET fin widths of a few nm (already 8 nm in 22 nm node devices[12]) modify not only the carrier distributions in real-space but also the band structure and even scattering rates. Each of these quantum effects is exacerbated in materials with very light effective masses m^* (e.g., $m^* = 0.044\ m_{\mathrm{e}}$ for Γ-valley electrons in $\mathrm{In_{0.53}Ga_{0.47}As}$), now being considered for MOSFETs. The continued use of particle-based MC under these conditions requires modification to the semi-classical methodology.

In this chapter, we present an ensemble 3D semi-classical MC simulator for n-channel devices whose treatment of electron degeneracy and quantum confinement institutes new approaches for particle-based simulations. We focus on our original contributions to the state-of-the-art, including our treatment of (i) far-from-equilibrium degenerate statistics, (ii) QC-based modeling

of surface-roughness scattering, and (iii) extending our group's previously introduced treatment of quantum-confined phonon and impurity scattering to 3D. In doing so, we expand upon our techniques, verify our methodologies, and refine and extend results introduced in a short preliminary study.[29]

After a brief description of the underlying purely semi-classical simulator, we detail our treatment of far-from-equilibrium degenerate carrier statistics. To consider the Pauli exclusion principle in MC simulation, scattering processes for electrons are either accepted or rejected according to the probability that the final scattering state is already occupied. The distributions of final states are typically approximated as being Fermi distributions, even if hot, dictated by the average local electron density and energy.[30–36] However, this approximation cannot be justified under strong non-equilibrium conditions approaching the ballistic limit of performance. In this work, we avoid *a priori* assumptions about the shape of the electronic distribution functions. Instead, we sample the electron populations locally in energy, energy valley, and propagation direction to generate the occupation numbers for the PB of scattering to states which are already occupied. Such approaches have been executed in **k**-space for bulk calculations,[37,38] but now we extend this method to include real-space variations in the distribution function for device simulation. This process is aided by the use of fractional carriers or sub-carriers, which not only improves statistics but, as the principle motivation, minimizes classical carrier-carrier scattering otherwise introduced via the time-dependent solution of Poisson's equation, which is incompatible with degenerate statistics.

Next, QCs for various quantum-confinement effects are provided through multiple uses of valley-, space-, orientation-, and time-dependent quantum-correction potentials (QCPs). Here, we calculate the set of QCPs based on the solutions of effective mass Schrödinger's equations defined in each channel slice normal to the transport direction[39–47] on a valley-by-valley basis[42–45] considering 2D confinement,[45–47] a first-principles strategy requiring no adjustable parameters. However, it is our uses of the QCPs, not their method of calculation, which is the focus here. Indeed, it may be possible to extend such uses of QCPs in these ways, however calculated, to still more computationally efficient drift diffusion and hydrodynamic simulations. The QCPs redistribute the MC electrons in real-space (e.g., away from interfaces) to reflect the quantum-mechanical spatial density. In addition, our QCPs naturally alter energy separations between energy valley minima, leading to degeneracy-splitting and redistribution of charge among energy valleys through scattering. Further, we use the QCPs to adjust 3D phonon and ionized-impurity scattering rates self-consistently on-the-fly, an extension of our previous 2D strategy.[42–44] Lastly, for the first time in MC simulation, we model surface-roughness (SR) scattering rates as a function of our QCPs. This SR method is quite general and allows for arbitrary potential-well shapes and confining geometries, moving beyond typical triangular-well assumptions for SR rate calculations.

Our QCs capture the main qualitative effects of quantum confinement and electron degeneracy within MC simulation. Using $In_{0.53}Ga_{0.47}As$ and Si n-channel FinFETs as examples, we illustrate the importance of each quantum

effect by analyzing simulation results with and without QCs. While both III-V and Si devices suffer these quantum effects, the scale is decidedly greater for III-V devices. In $In_{0.53}Ga_{0.47}As$ Γ-valleys, the light effective masses, low densities of states (DOS), and encountered highly-degenerate carrier populations lead to undesirable low quantum capacitances and high SR scattering rates. Confinement-reduced intervalley energy separations lead to sizable transfer of Γ-electrons to peripheral L- and X-states, reducing channel injection velocities, although also beneficially increasing the quantum capacitance. A measure of the significance of the here-modeled quantum effects is that the simulated ultra-scaled III-V devices exhibit worse ON-state transconductance than otherwise identical Si devices.

In Section 2.2, we introduce our device simulator before discussing the details of our QCs in Sections 2.3 and 2.4. In Section 2.5, we present a detailed comparison of devices with different levels of quantum-corrected modeling. Finally, we summarize our MC study in Section 2.6.

2.2 3D FinFET device and uncorrected MC simulator

The device structure and the baseline purely semi-classical MC simulator used in this work are intended as vehicles for illustrating the QCs that are the focus of this study. Indeed, both the structure and baseline MC simulator are somewhat idealized for this purpose.

2.2.1 Device structure

The device used as a test bed in this work is shown in Fig. 2.1. It has a fin-shaped semiconducting channel, connecting two heavily-doped electron reservoirs. The n-type channel materials we study here are industry-standard Si and $In_{0.53}Ga_{0.47}As$, the latter being lattice-matched to fabrication-friendly InP[15] and a candidate for future CMOS. The S/D reservoirs are doped to $N_{\mathrm{D}} = 5 \times 10^{19}$ cm^{-3}, a realistic activated dopant density which can be reached in III-V materials with current *in situ* growth technology.[48] Certainly Si devices are doped much more heavily than this in practice, as in our separate ongoing simulation study focused on device scaling and short-channel performance. Here, however, the focus is on our simulation methods, so we choose equal doping concentrations between the materials as a control. This allows a more fair comparison with regard to the essential transport physics. We also consider a lower dopant density of 1×10^{19} cm^{-3} in $In_{0.53}Ga_{0.47}As$. The correspondingly lower chemical potential in the S/D avoids contact injection directly into the peripheral valleys in $In_{0.53}Ga_{0.47}As$, isolating the role of intervalley scattering within the device simulation region. We model all devices as having perfectly injecting and absorbing boundary conditions, an idealization for both systems but more so for $In_{0.53}Ga_{0.47}As$, who requires more careful materials processing than Si to develop ohmic contacts.[49] Perfectly injecting and absorbing boundary conditions give a more fair comparison between the channel materials by decoupling the channel performance from current experimental and technological constraints regarding the metal contacts. The

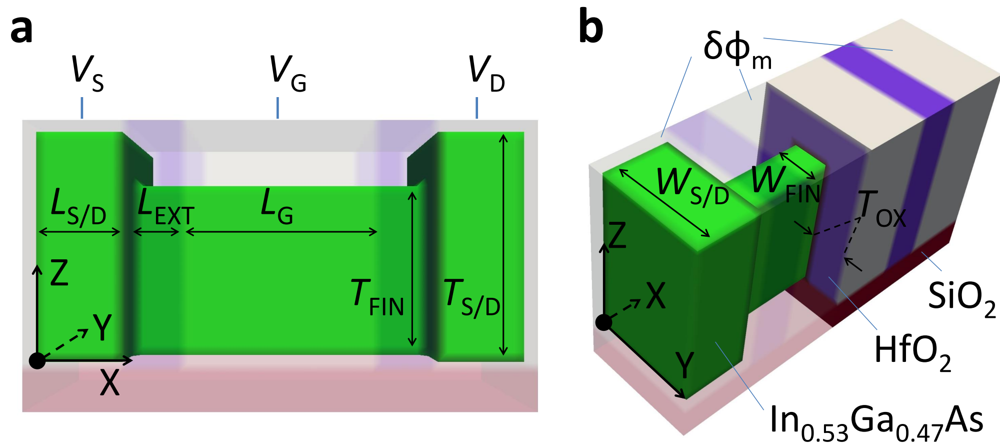


Figure 2.1 – FinFET structure. (a) Side view of the simulated device with relevant dimensions. Outer materials are shown transparent to visualize the semiconductor fin. **(b)** Edge view with cross-sectional clip to show different material regions. We vary the work functions in the $\delta\Phi_\mathrm{m}$ metal contacts to set the injection boundary conditions in the S/D such that there are flat-band conditions at the interface.

undoped fin-channel sits atop a 5 nm insulating layer of SiO_2 ($\varepsilon_\mathrm{r} = 3.9$) in a semiconductor-on-insulator (SOI) fully-depleted configuration. The gate oxide is a 5 nm insulating layer of HfO_2 ($\varepsilon_\mathrm{r} = 22.3$) and is wrapped around a 6 nm wide fin-channel. The channel is 20 nm long beneath the gate with 6 nm extensions. The channel sidewall orientation is (surface)/$\langle$channel$\rangle$ = (100)/$\langle$100$\rangle$, which is optimized for n-type transport in Si, although not for Si CMOS as a whole. This orientation is more interesting in terms of the quantum effects due to the nature of the valley degeneracy-breaking, as will be discussed later.

2.2.2 Baseline purely semi-classical Monte Carlo simulator

Our baseline MC simulator follows the basic methods described in Refs. 37, 38. In later sections and plots, this purely semi-classical model (CL) provides a reference. It contains no considerations for the Pauli exclusion principle or quantum-confinement effects. We summarize important details of the implementation here. Specifically, we generate a uniform 3D real-space cubic mesh of 1 nm resolution. Each MC loop has a 1.2 fs time step such that, e.g., a very fast carrier moving 8×10^7 cm/sec can almost cross one grid site per time step. This time length is chosen large enough to minimize the computational burden while being small enough to converge the numerical data. Poisson's equation is solved every time step consistent with the updated electron density. The electrostatics are modeled for each material based on their static dielectric permittivity. Simulated electrons couple to the mesh via a nearest grid-point assignment of charge. This approach is simple and eliminates self-forces while any electron remains in the cubic nm cell centered about its grid site. However, when carriers do cross a cell boundary between grid points during a time step, the forces and scattering rates are adjusted instantaneously.

At this point a significant self-force would result from an electron in the new cell being repelled by its own contribution to the charge in the old cell until the next update of Poisson's equation. The energy gain due to self-forces is exacerbated in the $In_{0.53}Ga_{0.47}As$ Γ-valley as compared to Si. The ratio of energy gain between these two materials should roughly follow the ratio of the respective effective masses, which is in proportion to the product

of the probability that a carrier will leave a grid site before Poisson's equation is updated, and how far it will likely travel in that remaining time period (where change in energy equals force times change in distance), both of which vary inversely with the square root of effective mass. Normally a self-force correction would be required to alleviate this artifact.[50,51] However, with our use of sub-carriers, as detailed subsequently and whose impact is quantified below, the issue is nearly moot. The energy provided by the remaning effects of self-forces are small compared to the thermalizing effects on the carrier population by phonon scattering and the device contacts, as will be shown.

The contacts are modeled by coupling the semiconductor S/D reservoirs to equilibrium electron distributions in the metal leads. To realize perfectly injecting and absorbing contacts, we adjust the work function of the metal $q\Phi_{\mathrm{m}}$ to set the Fermi level E_{F} with respect to the conduction band edge E_{C} at the interface to provide a free electron carrier density corresponding to the doping density N_{D} within the S/D. Within the electron affinity rule, this means $q\Phi_{\mathrm{m}} = q\chi - (E_{\mathrm{F}} - E_{\mathrm{C}})$, where $q\chi$ is the electron affinity of the semiconductor. This creates flat-band conditions at the contact interface. (The common value of $E_{\mathrm{F}} - E_{\mathrm{C}}$ is altered between simulations modeling classical and quantum statistics, however, for a given doping concentration.) Warm-up periods of 2.4 picoseconds have proven sufficient to remove initial simulation transients. Final statistics were averaged over 18 picosecond intervals per gate bias.

2.2.3 Silicon and III-V material models

We generally follow the material parameters assembled in Refs. 38, 52 for valley-specific effective masses, non-parabolicity constants, and deformation potentials. We choose an analytic non-parabolic description of the band structure that accurately reproduces the DOS in Si up to 1.5 eV in the conduction band compared to full-band calculations,[53] far larger than carrier energies produced by applied voltages of interest in our scaled devices. For Si, we model 6 ellipsoidal Δ-valleys, while for $In_{0.53}Ga_{0.47}As$ we include 1 Γ-, 4 L-, and 3 X-valleys. The Γ-valley is modeled as spherical while the L- and X-valleys are modeled as ellipsoidal.

We employ a virtual crystal approximation (VCA) to model $A_{\mathrm{x}}B_{1-\mathrm{x}}C$ ternary alloys such as $In_{0.53}Ga_{0.47}As$ considered here. We specify bowing parameters for the intervalley separations between the Γ-, L-, and X-valleys, while all other parameters within the VCA are linearly interpolated. We typically model the intervalley separation $E_{\Gamma\mathrm{L}}$ between the light-mass Γ-valley and heavier-mass peripheral L-valleys as $E_{\Gamma\mathrm{L}} = 487$ meV, determined by a set of bowing parameters recommended by Vurgaftman and colleagues in their comprehensive review article.[54] This value is a compromise between a commonly cited tight-binding calculation[55] ($E_{\Gamma\mathrm{L}} = 460$ meV) and the only experimental determination[56] ($E_{\Gamma\mathrm{L}} = 550$ meV) to date. However, within the literature there is significant uncertainty in $E_{\Gamma\mathrm{L}}$.[57] Recent density-functional calculations have estimated $E_{\Gamma\mathrm{L}}$ to be as large as 1.31 eV.[58] Given such uncertainty, later we will analyze a fictitious $In_{0.53}Ga_{0.47}As$ device having no satellite valleys

whatsoever ($E_{\Gamma L} \rightarrow \infty$). We will show that the main impacts of the peripheral valleys are to (i) increase the quantum capacitance via enhanced DOS, (ii) reduce the injection velocity due to heavier masses, and (iii) reduce the injection efficiency due to larger back-scattering. These effects are competing and it is not clear from the outset whether heavy occupation of the peripheral valleys will enhance or degrade device performance in III-V channels.

For scattering, our simulator includes intravalley acoustic phonons within an elastic equipartition approximation and inelastic intra/intervalley deformation potential optical phonon scattering with a constant phonon energy.[37,38] Umklapp f- and g-type intervalley scattering processes are included for Si[59] and polar optical intravalley phonon scattering is considered for III-Vs.[37,38] Degenerate ionized impurity scattering rates are calculated using a Brooks-Herring approach[60] employing a Thomas-Fermi screening model.[61] We found this model to more readily reproduce low-field mobilities consistent with experiments in the degenerate limit compared to a Debye screening model. Alloy scattering is modeled with a crystal disorder deformation potential.[62] (SR scattering is included via QCs as discussed subsequently). We reproduced known bulk scattering rates as a function of energy for each scattering process individually to confirm our approach in each material.

We verified bulk transport by reproducing experimental carrier drift-velocity versus electric field curves for Si[37] and $In_{0.53}Ga_{0.47}As$[63,64] including the temperature dependence of the phonon bath at 300 K and 77 K. We not only matched the low-field mobilities but also the peak velocities in both

materials to experimental data. Further, we verified that the onset of negative differential behavior in $In_{0.53}Ga_{0.47}As$, which denotes intervalley transfer of Γ-valley electrons to peripheral valley L-states, occurred at the correct electric field strength. Reproducing the velocity-field curves required small tunings of various deformation potentials, which is commonplace to MC simulation where deformation potentials are viewed as adjustable parameters.[52] All our final simulation parameters and their references are listed in Appendix A.

2.3 Quantum-corrections for non-equilibrium degenerate statistics

2.3.1 The Pauli-blocking of scattering

Modern MOSFET devices employ large carrier concentrations throughout the device. With effective oxide thicknesses (EOTs) below 1 nm, and multi-gate geometries, large carrier concentrations can be obtained in the channel under gating in the ON-state. Activated S/D doping densities approaching solid-solubility limits then are used to improve performance by making the semiconductor reservoirs more metallic, reducing parasitic series S/D resistance, and increasing the overall device transconductance $g_{\mathrm{M}} = (\mathrm{d}I_{\mathrm{DS}}/\mathrm{d}V_{\mathrm{GS}})$. However, such carrier concentrations also can far exceed the conduction band effective DOS N_{C}, raising the chemical potential well above the conduction band edge, the more so for lower DOS. Such degenerate carrier populations invalidate classical statistics (Boltzmann statistics in the equilibrium limit), and quantum statistics (Fermi statistics in the equilibrium limit) must be con-

sidered.

These quantum statistics are self-consistently produced by the consideration of what can simply be referred to as the PB of scattering. That is, the scattering rate $S(\mathbf{k}_i, \mathbf{k}_f)$ from any initial state $\mathbf{k}_i$ of occupation probability $f(\mathbf{k}_i)$ to some final state of occupation probability $f(\mathbf{k}_f)$ will be reduced in proportion to $1 - f(\mathbf{k}_f)$ compared to what otherwise would be expected,

$$S(\mathbf{k}_i, \mathbf{k}_f) = P(\mathbf{k}_i, \mathbf{k}_f) f(\mathbf{k}_i)\big(1 - f(\mathbf{k}_f)\big), \tag{2.1}$$

to accommodate the Pauli exclusion principle. Here $P(\mathbf{k}_i, \mathbf{k}_f)$ is the scattering probability per unit time from a full state to an empty state. To address quantum statistics in otherwise semi-classical MC, with initial states intrinsically fully occupied in the MC method ($f(\mathbf{k}_i) = 1$), the PB of scattering typically is treated stochastically. Scattering events are first selected consistent with $P(\mathbf{k}_i, \mathbf{k}_f)$ pre-calculated by Fermi's Golden Rule. Then the scattering events are stochastically rejected with a probability $1 - f(\mathbf{k}_f)$ according to the likelihood that the final state is already occupied. The question becomes what to use for the distribution function $f(\mathbf{k}_f)$, and how to determine it.

In a commonly employed approximation,[30–36] the distribution function of final scattering states $f(\mathbf{k}_f)$ at position $\mathbf{r}$ is assumed to be a Fermi-Dirac distribution $f_{\mathrm{FD}}\big(\mathbf{r}, E(\mathbf{k}_f)\big)$ for the purposes of PB in Eq. (2.1). In this strategy, the shape of f_{FD} is determined by the local quasi-Fermi level and temperature consistent with the local carrier concentration and average energy. This approximation represents a great improvement over neglecting the PB

of scattering when assuming classical statistics, being rigorous in the equilibrium limit. However, under strong non-equilibrium conditions including quasi-ballistic transport, actual distribution functions can become locally non-Fermi-like throughout the considered nanoscale device, including the channel.

In this work, we make no *a priori* assumptions about the shape of the distribution function. Instead, we calculate the distribution function to be used for PB directly by sampling the local carrier population $N(\mathbf{r}, E, g, \pm)$ as a function of position $\mathbf{r}$, valley g, energy E relative to its respective valley-edge, and propagation directions forward toward the drain end (+) or backward toward the source end (−). The corresponding distribution function $f(\mathbf{r}, E, g, \pm)$ then is obtained from

$$f(\mathbf{r}, g, E, \pm) = \frac{N(\mathbf{r}, g, E, \pm)}{D(g, E)/2} \;, \tag{2.2}$$

where $D(g, E)/2$ is the position independent DOS per energy valley reduced by a factor of two for these half space (±) distributions. (In the case of quantum confinement considered below, E is the energy referenced to, specifically, the non-quantum-corrected valley edge.) The resolution of $f(\mathbf{r}, E, g, \pm)$ in energy, ΔE, is chosen depending on the equilibrium $E_{\mathrm{F}} - E_{\mathrm{C}}$ value in the S/D electron reservoirs. For light-mass $\mathrm{In_{0.53}Ga_{0.47}As}$ with an activated doping density of $N_{\mathrm{D}} = 5 \times 10^{19}$ $\mathrm{cm^{-3}}$ and a corresponding $E_{\mathrm{F}} - E_{\mathrm{C}} = 510$ meV, we choose an energy discretization of $\Delta E = k_{\mathrm{B}}T = 25.9$ meV at 300 K. For Si with a much greater DOS, at the same donor density we choose $\Delta E = k_{\mathrm{B}}T/4$ corresponding to a much smaller $E_{\mathrm{F}} - E_{\mathrm{C}} = 24.5$ meV. Si exhibits about 20×

greater quantum capacitance $C_{\mathrm{Q}} = \mathrm{d}Q/\mathrm{d}(E_{\mathrm{F}}/q)$ than $\mathrm{In}_{0.53}\mathrm{Ga}_{0.47}\mathrm{As}$ when considering degenerate statistics.

Such approaches have been executed in **k**-space for bulk calculations,[37,38] but extension to device simulations has so far been prohibitive. This is because of large random-access memory demands and limited sample sizes. We use three basic methods to increase our sample size: (i) averaging over short time periods, (ii) averaging over small regions in space and (iii) the use of fractional electrons or *sub-carriers*. Considering (i), we average $N(\mathbf{r}, E, g, \pm)$ over a time period of 120 fs or 100 time steps, which is still an order of magnitude shorter that the switching time for even a THz transistor. With electrons moving only on the scale of Angstroms/time step, this time averaging effectively increases the sample size by roughly an order of magnitude. For (ii), we average $N(\mathbf{r}, E, g, \pm)$ over the central nearest neighbor grid sites, increasing the sample size on average by a factor of 27 except at the device boundaries, at any given point in time. (Arguably, although not our motivation, averaging over nearest neighbor grid sites or beyond is perhaps more physically realistic than not doing so given the actual quantum-mechanical nature of the particles.) Finally considering (iii), we represent each real electron with 100 sub-carriers each carrying $1/100^{\mathrm{th}}$ the fundamental charge, increasing our sample size by another two orders of magnitude. All told, the sample size from which we calculate $N(\mathbf{r}, E, g, \pm)$ and, thus, $f(\mathbf{r}, E, g, \pm)$, is effectively over four orders of magnitude larger than the physical number of carriers that would be expected at any grid point at any point in time!

2.3.2 Effects of degenerate statistics under equilibrium conditions

Fig. 2.2 illustrates convergence of our PB method to the known equilibrium results in Γ-valley $In_{0.53}Ga_{0.47}As$ electrons with a uniform device carrier concentration of 5×10^{19} cm^{-3} at 300 K. (Since quantum effects are more pronounced in III-V materials, within this section and the upcoming Section III we illustrate our QCs in the $In_{0.53}Ga_{0.47}As$ system, before returning to consider Si devices as well.) Here we closed the device boundaries and enforced flat-band conditions, considering only scattering and allowing the simulation to come to equilibrium. We then plotted the average local charge density distribution versus carrier kinetic energy, sampled over a single quasi-instantaneous 120 fs time interval. The carrier densities are normalized to the peak theoretical density in the classical limit. Our device populations both with ($n_{\mathrm{PB}}^{\mathrm{MC}}$) and without ($n_{\mathrm{CL}}^{\mathrm{MC}}$) the PB of scattering show excellent agreement compared to the reference equilibrium Fermi-Dirac (n_{PB}) and Boltzmann (n_{CL}) statistics. In $In_{0.53}Ga_{0.47}As$, the Fermi level rises nearly 400 meV into the conduction band upon considering degeneracy to accommodate the modeled 5×10^{19} cm^{-3} carrier concentration. This is in stark contrast to Si, where the Fermi energy only moves up 13 meV in the conduction band when considering PB at the given doping level. For $In_{0.53}Ga_{0.47}As$, the large change in the chemical potential occurs despite partial pinning of the Fermi level by the introduction of charge carriers into the peripheral L-valleys. L-valley electrons are not shown in Fig. 2.2 for clarity. However, their occupation probabilities converge to the proper Fermi distribution as well, in concert with the Γ-valley. At the consid-

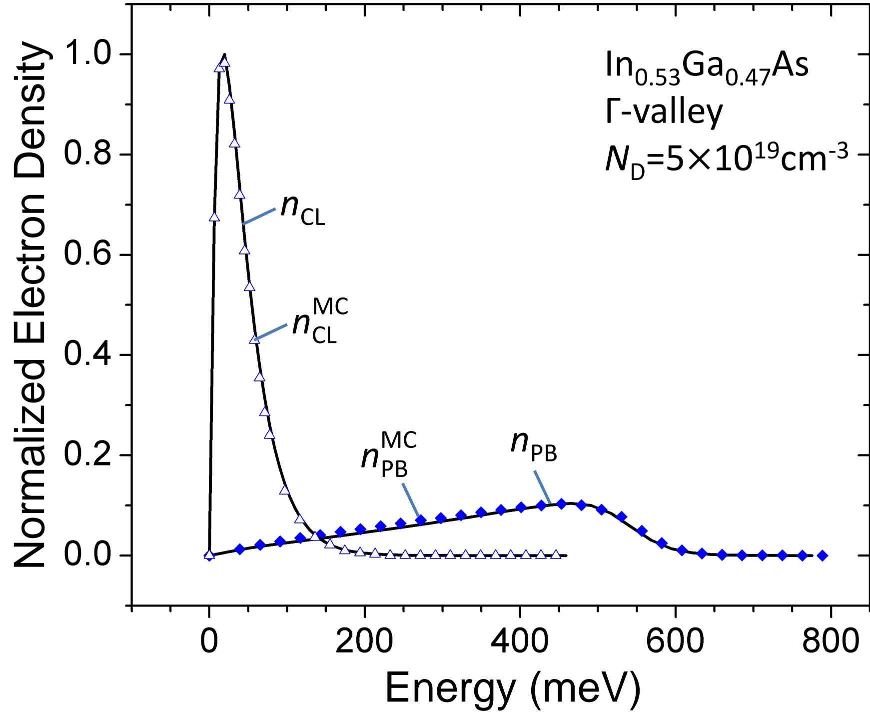


Figure 2.2 – Effects of degeneracy in $In_{0.53}Ga_{0.47}As$. MC simulated electron distributions (symbols) with the PB of scattering ($n_{\mathrm{PB}}^{\mathrm{MC}}$) and without ($n_{\mathrm{CL}}^{\mathrm{MC}}$) in the Γ-valley of $In_{0.53}Ga_{0.47}As$. The results are shown under flat-band and equilibrium conditions, with reference distributions (solid lines) calculated directly from Fermi-Dirac (n_{PB}) and Boltzmann (n_{CL}) statistics, respectively. The distributions have been normalized to the peak electron density in the classical limit as a reference. There is 40% less Γ-valley charge (area under the curve) in the degenerate case.

ered doping, the PB Γ-valley population is reduced to just 60% of the total equilibrium density, with 40% being in the L-valleys.

While our sub-carriers significantly enhance our simulation statistics, our primary motivation for their use was to minimize classical carrier-carrier scattering which results from charge interacting via the time-dependent solution of Poisson's equation. This is in contrast to usual ensemble MC simu-

lators, which typically embrace treating carrier-carrier scattering classically. The ability to model carrier-carrier scattering via the Poisson solution is commonly thought to be a benefit of ensemble MC simulation versus full-quantum methods due to its simplicity and speed. However, classical molecular dynamics carrier-carrier scattering intrinsically neglects the PB of final-state pairs. Therefore, although these interactions serve to thermalize the carrier population, they do so towards a high-temperature Boltzmann distribution which is incompatible with non-equilibrium degenerate statistics. Moreover, the Coulomb force between two electrons at 2.5 nm apart (roughly the average separation for a carrier density of 5×10^{19} cm^{-3}) is quite strong at over 20 mV/nm, maximizing not only this classical thermalization effect, but also fictitious self-forces. However, with our introduction of $N_{\rm sub}$ sub-carriers per real electron, each of which contributes only $q/N_{\rm sub}$ to the charge density, the Coulomb force among sub-carriers is reduced by $N_{\rm sub}$ (but not $N_{\rm sub}^2$) to

$$\left|\mathbf{F}_{\rm e-e}\right| \simeq \frac{1}{N_{\rm sub}} \frac{q^2}{4\pi\varepsilon|\mathbf{r}_1 - \mathbf{r}_2|} \; . \tag{2.3}$$

This force still must be taken as proportional to the full charge on a real electron considering the local electric field to properly model the effects of the applied source, drain, and gate voltages.

Using the Golden Rule scattering rate as a measure, while the number of carriers to scatter off increases by $N_{\rm sub}$, the scattering rate between any two sub-carriers decreases by $N_{\rm sub}^2$, for a net reduction in the scattering rate by $N_{\rm sub}$. (Indeed, a future goal would be to introduce carrier-carrier scattering

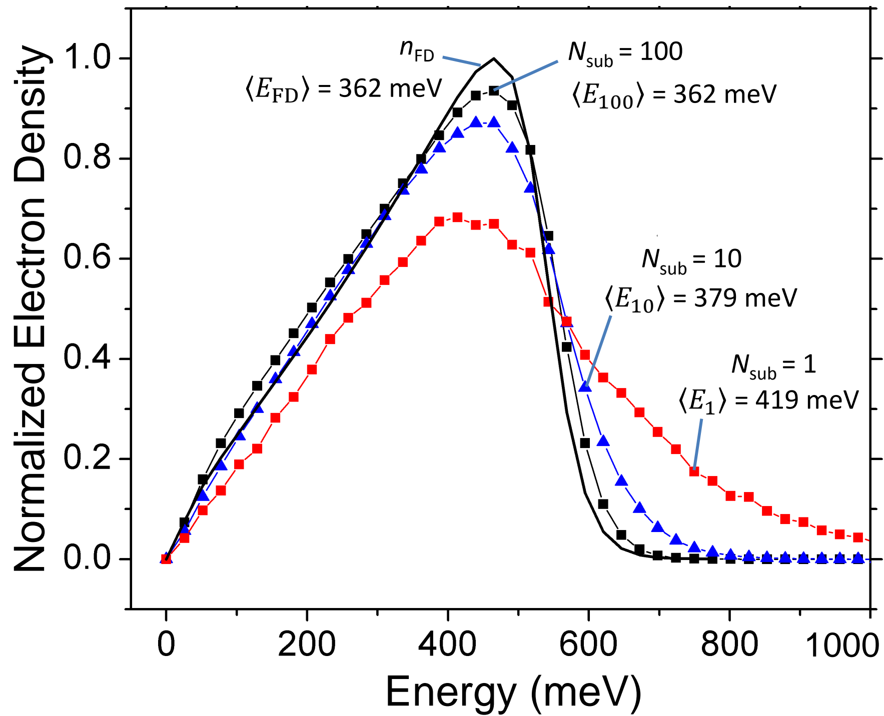


Figure 2.3 – Impact of sub-carriers on classical carrier-carrier interactions. MC simulated electron distributions in the Γ-valley of $In_{0.53}Ga_{0.47}As$ under equilibrium with different sub-carrier factors. The product of the 300 K Fermi-Dirac distribution and DOS is shown for reference ($n_{\rm FD}$). With $N_{\rm sub} = 1$, the unwanted thermalizing effects of classical carrier-carrier scattering drive the MC electron distribution to a high-temperature-like shape. Increasing the sub-carrier factor to $N_{\rm sub} = 10$ and then $N_{\rm sub} = 100$, the device distributions approach the theoretical expectation in shape and average kinetic energy $\langle E \rangle$. These distributions are normalized to have the same area under each curve. (The actual electron concentration in the Γ-valley is not conserved due to the occupation of peripheral valleys as a function of sub-carrier factor.)

within a practical framework for which PB still can be considered.) Thus we not only reduce the classical Coulomb force between sub-carriers but also their effective carrier-carrier scattering rates as well.

In Fig. 2.3, we turn on the self-consistent Poisson equation, open the

S/D reservoirs to the metal contacts, and re-run the previous simulation study of Fig. 2.2, which was considered under flat-band conditions. The average local Γ-valley $In_{0.53}Ga_{0.47}As$ charge density distribution versus carrier kinetic energy is shown, along with the average carrier kinetic energy per distribution, for differing sub-carrier factors N_{sub} to quantify their impact. With $N_{\mathrm{sub}} = 1$, classical carrier-carrier scattering thermalizes the electron population to an undesirable high-temperature-like distribution in energy, while self-forces raise the average kinetic energy about 60 meV above the theoretical expectation, despite the coupling to the boundaries and phonon scattering driving the electron distribution toward the Fermi-Dirac limit. However, as the sub-carrier factor is increased to 10, and then 100, the shape and average energy of the MC electron distributions approach those of the product of the DOS and the 300 K Fermi-Dirac distribution, n_{FD}, and $\langle E_{\mathrm{FD}} \rangle$. This agreement is evidence that our sub-carrier strategy both mitigates the non-PB thermalization effects of classical carrier-carrier scattering, and relegates energy gains due to self-forces to negligible levels.

2.3.3 Effects of degenerate statistics under far-from-equilibrium conditions

The strength of our method, however, is that no *a priori* assumption of an equilibrium—or any—specific distribution is made, in contrast to the use of a Fermi approximation for PB. We therefore conclude this section by illustrating our PB method under far-from-equilibrium conditions by sampling the $In_{0.53}Ga_{0.47}As$ Γ-valley charge distributions under bias in Fig. 2.4. The bias

conditions are source-to-drain voltage $V_{\mathrm{DS}} = 0.6$ V and gate overdrive above threshold $V_{\mathrm{ON}} = V_{\mathrm{GS}} - V_{\mathrm{T}} = 0.35$ V in accordance with ITRS predictions for future scaled MOSFETs.[65] The carrier distributions are sampled in the plane normal to the transport direction at the top of the channel potential-energy barrier-top (Fig. 2.4a) and also at the drain end of the channel (Fig. 2.4d). The electrostatic potentials, which are plotted along the dotted-white lines in Fig. 2.4a,d for reference, are visualized in Fig. 2.4b,e, where x_{b} represents the location of the barrier-top, and x_{d} the location of the beginning of the drain. The forward-going (+) and backward-going (−) carrier distributions necessarily differ greatly at the top of the channel barrier (Fig. 2.4c) consistent with a high injection efficiency. The forward-going distribution at the drain end (Fig. 2.4f) shows two peaks, the lower energy peak consistent with a nearly equilibrium distribution of charge carriers in the drain reservoir, and the higher energy peak consistent with quasi-ballistic electrons injected from the source.

It is clear from Fig. 2.4 that the non-equilibrium electron occupation probability distribution is not consistent with any single Fermi distribution even as a local function of position. However, we cannot provide a direct comparison between the effects of the Fermi approximation versus our self-consistently obtained non-equilibrium distributions on device performance, such as on the drain current. The use of the Fermi approximation presents its own set of programming and computational challenges. Either sophisticated non-linear solvers must be used to determine the quasi-Fermi levels and temperatures, or large two-dimensional (E_{F}, T) reverse look-up tables, e.g.,

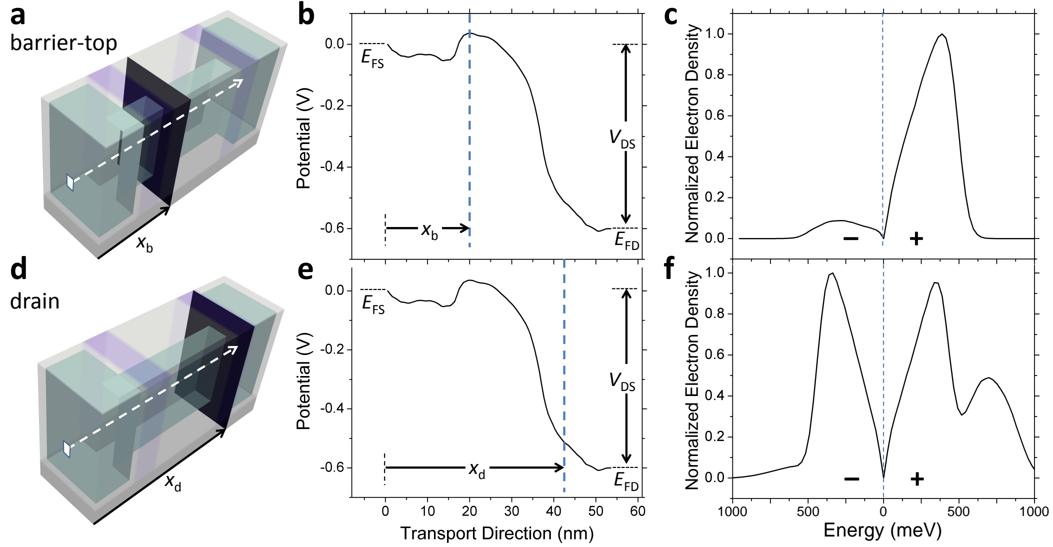


Figure 2.4 – Non-equilibrium electron distributions. (a) FinFET edge view showing the device slice containing the potential energy barrier-top. **(b)** Potential profile showing the location of the barrier-top. **(c)** Normalized MC simulated electron distributions sampled at the potential energy barrier-top to the channel in the Γ-valley of $In_{0.53}Ga_{0.47}As$ under non-equilibrium. The much larger forward-going contribution $(+)$ versus backward-going $(-)$ is consistent with a high injection efficiency into the channel. **(d)** FinFET edge view showing near the drain end of the channel. **(e)** Potential profile showing the location of the drain end. **(f)** Normalized MC simulated electron distributions again but sampled near the drain. The forward-going distribution $(+)$ is the superposition of the drain near-equilibrium charge distribution and the quasi-ballistic population injected from the source.

$(n, \langle E \rangle) \rightarrow (E_{\mathrm{F}}, T)$ would be required. Both strategies become increasingly difficult to use for highly degenerate statistics as considered in this work, where the average energy $\langle E \rangle$ becomes an increasingly weak function of T, and the numerical stability of the non-linear solvers becomes a concern. Thus, the inclusion of the Fermi approximation approach to degenerate statistics in our quantum-corrected MC simulator for the purpose of comparing methods is impractical.

2.4 Quantum-corrections for electron quantum confinement

Electrostatic quantum-correction potentials (QCPs) are widely used in MC simulation for the purpose of modeling quantum confinement. Each of our QCs for confinement described in this work are designed to employ the same set of QCPs. We employ first-principles QCPs which inherently require no fitting parameters as opposed to other methods like effective quantum potentials,[66–68] perturbative approaches,[69,70] or density-gradient models[71,72] which require calibration. Specifically, we provide a valley-by-valley treatment of the space-, orientation-, and time-dependent QCPs based on the solutions of 2D effective mass Schrödinger's equations solved in each transport slice.[39–47] The valley and orientation dependence is provided by including the reciprocal effective mass tensor in the model Hamiltonian.[42–45] This point is necessary to capture the self-consistent modification of intervalley separations and degeneracy splitting of otherwise equivalent valleys.

Our uses of the QCPs include altering energy separations between energy valley minima and calculating quantum-confinement-dependent phonon and surface-roughness scattering rates, in addition to redistributing charge carriers in real space and modifying source-to-channel potential barriers. (They are, however, not designed nor used to model quantum-mechanical tunneling-related leakage currents along the channel or through the gate.) Although it is our uses of the QCPs that are the focus of this work, we still describe their method of calculation here for clarity and completeness.

2.4.1 Calculating the quantum-correction potential

In general for this approach, for each valley g (but not set of equivalent valleys) at position $\mathbf{r}$, the QCPs are defined by the relation

$$\rho^g_{\mathrm{CL}}\big(\,\bar{V}(\mathbf{r}) + V^g_{\mathrm{QC}}(\mathbf{r})\,\big) = \rho^g_{\mathrm{QC}}\big(\,\bar{V}(\mathbf{r})\,\big)\ . \tag{2.4}$$

Here $V^g_{\mathrm{QC}}(\mathbf{r})$ is defined as the effective potential which, upon addition to the electrostatic potential $\bar{V}(\mathbf{r})$, will produce a classical device space-charge distribution ρ_{CL} equal to the quantum-mechanical one ρ_{QM}. To smooth granularities in the potential found in the instantaneous device solutions, $\bar{V}(\mathbf{r})$ is a time-average over 100 time steps (120 fs total) of the potential $V(\mathbf{r})$ obtained from the self-consistent solution of Poisson's equation within the particle MC simulation, and includes the band and valley offsets in its definition. In equilibrium and for the 2D confinement considered here, within the $y-z$ plane of confinement, for each value of x along the transport direction, Eq. 2.4 can be written as

$$\int_{\bar{V}(\mathbf{r})+V^g_{\mathrm{QC}}(\mathbf{r})}^{+\infty} \mathrm{d}E\, D^g_{\mathrm{3D}}\big(E-\bar{V}(\mathbf{r})-V^g_{\mathrm{QC}}(\mathbf{r})\big)\, f_{\mathrm{FD}}(E) = \sum_{i=1}^{M} \int_{\bar{V}(\mathbf{r})}^{+\infty} \mathrm{d}E\, D^g_{\mathrm{1D}}\big(E-\bar{V}(\mathbf{r})\big)\, \big|\Psi^g_i(y,z)\big|^2\Big|_x f_{\mathrm{FD}}(E)\ . \tag{2.5}$$

We include $M = 20$ modes i in the summation for the simulations of this work. D^g_{1D} and D^g_{3D} are the valley-wise 1D and 3D DOS, respectively. The $\Psi^g_i(y,z)$ are the 2D eigenvectors that diagonalize the Hamiltonian matrix H within the 2D effective mass Schrödinger equation,

$$H\,\Psi^g_i(y,z) = \big[-\frac{\hbar^2}{2}\,\nabla_\perp \cdot \frac{1}{\boldsymbol{m}^*(y,z)} \cdot \nabla_\perp + \bar{V}(\mathbf{r})\big]\,\Psi^g_i(y,z) = E^g_i\,\Psi^g_i(y,z)\ , \tag{2.6}$$

where $1/\boldsymbol{m}^*$ is the reciprocal effective mass tensor, $\nabla_\perp = \hat{y}(\partial/\partial y) + \hat{z}(\partial/\partial z)$ is the transverse momentum operator, and E_i^g are the valley- and sub-band-dependent eigenenergies. We solve the eigenvalue problem using a finite difference scheme that preserves the continuity of the probability current across the semiconductor interface. Here, the integrations of Eq. 2.5 are completed within the parabolic limit inside a Boltzmann approximation for the Fermi-Dirac distribution,[39–47]

$$N_{\text{3D}}^g \, e^{\frac{E_{\text{F}} - \bar{V}(\mathbf{r}) - V_{\text{QC}}^g(\mathbf{r})}{k_{\text{B}}T}} = N_{\text{1D}}^g \, e^{\frac{E_{\text{F}} - \bar{V}(\mathbf{r}) - E_1^g}{k_{\text{B}}T}} \sum_{i=2}^{M} \left|\Psi_i^g(y,z)\right|^2 e^{\frac{-E_i^g + E_1^g}{k_{\text{B}}T}} \, . \tag{2.7}$$

Here, N_{1D}^g and N_{3D}^g are the valley-wise 1D and 3D effective DOS along the transport direction, respectively. The QCPs are revealed via logarithmic inversion as

$$V_{\text{QC}}^g(\mathbf{r}) = E_1^g - k_{\text{B}}T \ln\left(\frac{N_{\text{1D}}^g}{N_{\text{3D}}^g} \sum_{i=2}^{M} \left|\Psi_i^g(y,z)\right|^2 e^{\frac{-E_i^g + E_1^g}{k_{\text{B}}T}}\right) . \tag{2.8}$$

Finally, we then make a non-parabolicity correction to the QCPs by the reassignment

$$V_{\text{QC}}^g(\mathbf{r}) \Leftarrow \frac{1}{\alpha_g}\left(\sqrt{\frac{1}{4} + \alpha_g V_{\text{QC}}^g(\mathbf{r})} - \frac{1}{2}\right) , \tag{2.9}$$

where α_g is the valley non-parabolicity constant. This correction is consistent with the reduction of the electron energy relative to the band edge due to non-parabolicity for a carrier of fixed wavelength, here defined by the quantum confinement. Electrons, or sub-carriers here, are moved within the total potential according to the equations of motion

$$\mathbf{F}_g(\mathbf{r}) = \frac{d}{dt}\big(\hbar\mathbf{k}\big) = -q\nabla \cdot V_{\text{tot}}(\mathbf{r}) = -q\nabla \cdot \big(V(\mathbf{r}) + V_{\text{QC}}^g(\mathbf{r})\big), \tag{2.10}$$

which govern the evolution of their crystal momentum. Consistent with our PB statistical updates and the time-averaging of $\bar{V}(\mathbf{r})$, our QCPs are updated every 120 fs in this work.

There is no quantum-confinement in the S/D semiconductor regions since the electron wave functions can escape into the metal contacts. Thus there arises a question of how to approximate the 3D effects at the beginning and end of the conduction channel, where the quantum-confinement gradually turns-on and off moving from the unconfined S/D regions into and out of the restricted fin-channel. To estimate this transition, we first generated a Γ-valley QCP from a 2D slice *along* the transport direction in the horizontal $x-y$ plane (not transverse). We observed an approximately linear turn-on in this QCP approaching the channel that was rougly equal to the physical width of the fin. (The effective channel width allowing for barrier penetration is actually larger, which reduces the QCPs from what otherwise would be obtained.[8]) To appoximate this effect, we linearly ramp-up the QCPs $V^{g}_{\mathrm{QC}}(\mathbf{r})$ from the outer edges of the drain and source extentions toward the gated channel over a distance W_{fin}, the physical width of the channel in the horizontal plane.

2.4.2 Real-space redistribution of charge under quantum confinement

We see from Eq. 2.10 that the first effect of the QCPs is their application of classical forces on particles to redistribute them in real-space according to the as-calculated quantum-mechanical thermal charge distribution. This is

provided to accurately model the capacitance of the gate, where it is known that the channel wave function is actually repelled from the oxide interface under quantum confinement.

This spatial effect of the QCPs on the charge distribution is illustrated in Fig. 2.5 under 0.6 V drain bias for the channel cross-section of Fig. 2.5a located near the beginning of the channel at the location of the potential energy barrier-top. We analyze the distribution at gate voltages of 0.35 V above threshold (Fig. 2.5b) and at threshold (Fig. 2.5c) in an $In_{0.53}Ga_{0.47}As$ FinFET. Under each condition, we compare the purely classical MC device charge distribution (ρ_{CL}), the MC device charge distribution including PB and quantum confinement (ρ_{QC}), and the as-calculated equilibrium quantum-mechanical charge distribution in the Boltzmann limit from the preceding subsection (ρ_{QM}, the right-hand side of Eq. 2.7).

Above threshold, a strong interface potential well attracts electrons to the surface, as expected classically (ρ_{CL}, top left). Upon the inclusion of the QCPs, however, we see a device distribution with the population focused in the center of the channel and repelled from the interface as expected quantum-mechanically (ρ_{QC}, top middle). Yet this corrected shift of the carrier population is not as strong as for the reference Boltzmann equilibrium calculation (ρ_{QM}, top right), nor should it be. Unlike the calculations from which the QCPs were obtained, the quantum-corrected MC device simulations, corresponding to ρ_{QC}, are also subject to degenerate statistics. Carrier degeneracy, combined with effectively reduced DOS for predominantly down-

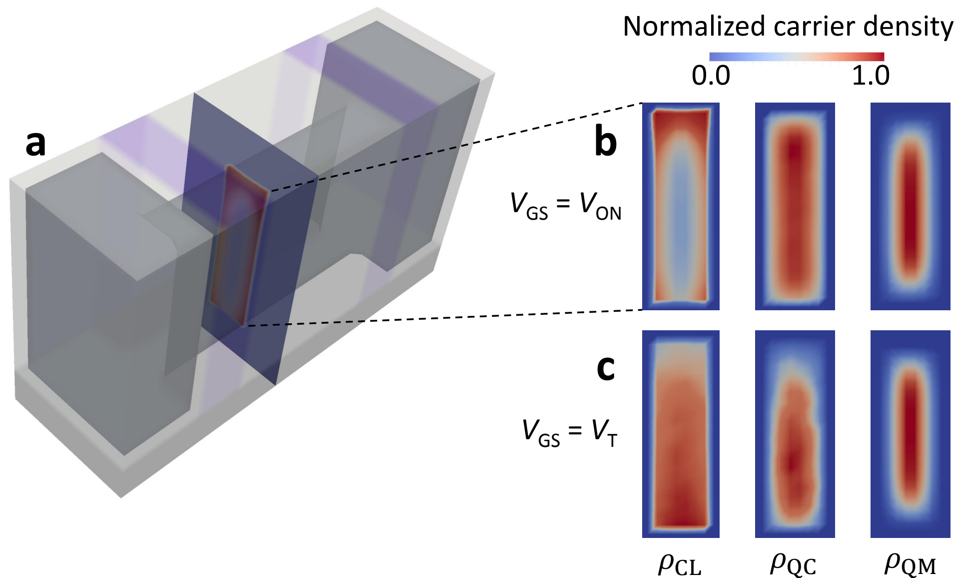


Figure 2.5 – Real-space redistribution of charge under quantum confinement. **(a)** $In_{0.53}Ga_{0.47}As$ FinFET with the channel slice containing the charge distributions of **b** and **c**. **(b)** At the overdrive gate voltage of 0.35 V above threshold, the classical electron charge density is attracted to the channel surface (top left, ρ_{CL}). Including QCPs repels carriers from the interface (top middle, ρ_{QC}) although degeneracy effects smear the spatial distribution relative to the as-calculated equilibrium Boltzmann-weighted quantum-mechanical charge density (top right, ρ_{QM}). **(c)** Here ρ_{QC} (bottom middle) more closely resembles ρ_{QM} (bottom right) at threshold, but effects of degeneracy and non-equilibrium behavior are still evident.

channel directed carriers (reduced by a factor of two in the ballistic limit), pushes carriers up in energy in the Γ-valley, as well as significantly into the peripheral valleys. This intervalley transfer of charge is due to a reduced Γ-L valley-splitting $E_{\Gamma L}$. The satellite valleys have much larger effective masses, weaker quantum effects, and correspondingly smaller QCPs, as captured by our valley-by-valley treatment of the QCPs. In $In_{0.53}Ga_{0.47}As$ quantum wells, we typically see the quantum-corrected $E_{\Gamma L}$ reduced by 200−300 meV in the

channel of the considered FinFET depending on the voltage conditions.

At threshold, but still under non-equilibrium degenerate conditions if less so, the quantum-corrected device distribution (ρ_{QC}, bottom middle) looks more like the reference quantum-mechanical distribution (ρ_{QM}, bottom right), but differences remain clear. In this way, our use of the equilibrium Boltzmann approximation for the purposes of calculating the QCPs does not prevent us from using these same QCPs to address quantum confinement as applied to non-equilibrium degenerate carrier populations. Moreover, our QCPs are most accurate but also most important for E_{F} near the quantum-corrected valley edges. For example, the QCPs change the energy barrier heights for electrons to enter the constricted FinFET channel and, thus, in particular, the threshold condition. A good rule of thumb for estimating the resulting shift in threshold voltage ΔV_{T} in the center of the channel due to quantum confinement is simply $\Delta V_{\mathrm{T}} \approx V_{\mathrm{QC}}^{\Gamma}$ for III-V materials (with a similar relation for Si considering V_{QC}^{Δ}). For more energetic and more degenerate carrier populations, however, it is easier for electrons to reach the interface than would be expected in a fully quantum-mechanical calculation, so our QCPs remain somewhat conservative.

2.4.3 Modeling quantum-confinement-dependent phonon and ionized-impurity scattering

Strong quantum-confinement-enhanced phonon scattering long has been recognized, sufficient enough to more than halve the electron mobility in Si conduction channels with a few tenths of a MV/cm effective interface

normal field absent even surface-roughness scattering.[73] This effect also has been seen in quantum transport calculations considering phonon scattering (as well as collision broadening thereof).[27,28] Scattering rates as a function of energy for electrons under quantum confinement oscillate about the bulk electron scattering rates with the introduction of each new final state sub-band when the confined carrier's energy is referenced to the expectation value of the uncorrected position-dependent valley edge. This behavior is illustrated via analytic Golden-Rule-based calculations for nominally randomizing short-range quasi-elastic (e.g., deformation potential acoustic phonon) scattering for one-dimensional confinement in an infinite square well and in a (one-sided) perfect triangular well in Fig. 2.6a and b, respectively. (Initial and final state occupation probabilities have been neglected in these rates.) This illustrative calculation considers not only the modification of the final-state DOS but also the overlap between initial and final states which leads to a preference for intra-sub-band scattering while remaining otherwise randomizing. In the limit of wide wells and low interface fields, these results converge to the bulk limit as they must. For narrow wells and high interface fields, however, the scattering rate for the lowest energy carriers, which are substantially above the bottom of the square well or the expected potential for the triangular well due to quantum confinement, increases continuously as the confinement increases.

To model these effects, we simply adjust the MC scattering rate by shifting the energy of the argument of the scattering rate for a given kinetic energy $\varepsilon^i_{\mathbf{k}}$ in the initial valley by the initial valley quantum correction V^i_{QC},

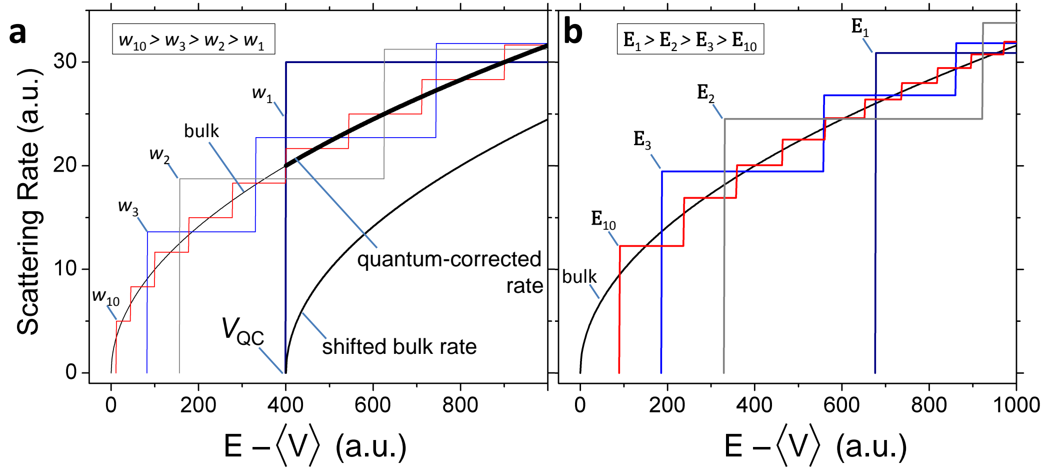


Figure 2.6 – Modeling quantum-confined scattering using adjusted bulk rates. The scattering rate for short range nominally randomizing elastic scattering within the first sub-band for 1D quantum confinement in **(a)** infinite square wells of varying width w, and **(b)** infinite one-sided triangular wells of varying fixed interface-normal field $\mathbf{E}$. The curves are labeled by how many sub-bands are picked up within the given energy range for each respective confinement condition. The scattering rates oscillate about the bulk rate (bulk) when the carrier energy is referenced to the expectation value of the electrostatic potential, more closely approaching the bulk limit with reduced confinement. For increased confinement, the scattering rate for low-energy first sub-band carriers is increased substantially. By using the bulk scattering rate but starting at the quantum-corrected band-edge V_{QC} (**a**, quantum-corrected rate), our approximation captures the overall increase in the basic scattering rate, if missing the oscillations and being somewhat conservative for low-energy carriers. By comparison, simply shifting the zero of the bulk rate by the quantum-confined band-edge energy (**a**, shifted bulk rate) results in a much larger error.

as also illustrated (Fig. 2.6a, quantum-corrected rate). That is, the quantum-corrected scattering rate $R^{\mathrm{QC}}_{i\to f}$ from some initial state i to some final state f is given in terms of the uncorrected (classical) scattering rate $R^{\mathrm{CL}}_{i\to f}$ as

$$\begin{aligned} R^{\mathrm{QC}}_{i\to f}(\varepsilon^i_{\mathbf{k}}) &= R^{\mathrm{CL}}_{i\to f}(\varepsilon^i_{\mathbf{k}} + V^i_{\mathrm{QC}}) \quad &;\quad \varepsilon^f_{\mathbf{k}} = (\varepsilon^i_{\mathbf{k}} + V^i_{\mathrm{QC}} - V^f_{\mathrm{QC}} - \Delta_{i,f} + \delta E) > 0 \\ R^{\mathrm{QC}}_{i\to f}(\varepsilon^i_{\mathbf{k}}) &= 0 \quad &;\quad \varepsilon^f_{\mathbf{k}} = (\varepsilon^i_{\mathbf{k}} + V^i_{\mathrm{QC}} - V^f_{\mathrm{QC}} - \Delta_{i,f} + \delta E) < 0 \end{aligned} \tag{2.11}$$

for any assumed energy conserving scattering processes, where $\Delta_{i,f}$ is the uncorrected energy separation from the initial valley edge to the final valley edge, and δE is the energy gained from (positive) or lost to (negative) the environment in the scattering process. We make the same adjustment for intravalley and intervalley scattering, quasi-elastic and inelastic, and (nominally) randomizing and non-randomizing scattering alike. (For intervalley scattering, the correction V^i_{QC} to the total energy of a carrier with kinetic energy $\varepsilon^i_{\mathbf{k}}$ raises the total energy with respect to the uncorrected band edge of the final valley just as for the initial valley.) Once the scattering rate has been selected, a specific final state subject to $\varepsilon^f_{\mathbf{k}} = \varepsilon^i_{\mathbf{k}} + V^i_{\mathrm{QC}} - V^f_{\mathrm{QC}} - \Delta_{i,f} + \delta E$ is chosen consistent with the relative probability determined by the difference in the initial and allowed final state $\mathbf{k}$ values in the usual way as appropriate for each scattering process.

While this approach misses the oscillations with energy and, in that way, is somewhat conservative for the lowest energy carriers, it captures the larger overall shift in the scattering rates with quantum confinement with no requirement to actually recalculate the scattering rates themselves from first-principles. Certainly this approach is more effective than simply shifting the zero of the bulk scattering rate itself by the QCP (Fig. 2.6a, shifted bulk rate). However we note that the effect of quantum confinement (and of this quantum correction in particular), can be relatively small or even reduce the scattering rates for non-randomizing long-range scattering interactions, consistent with their bulk energy dependence. An illustrative example of selecting the correct

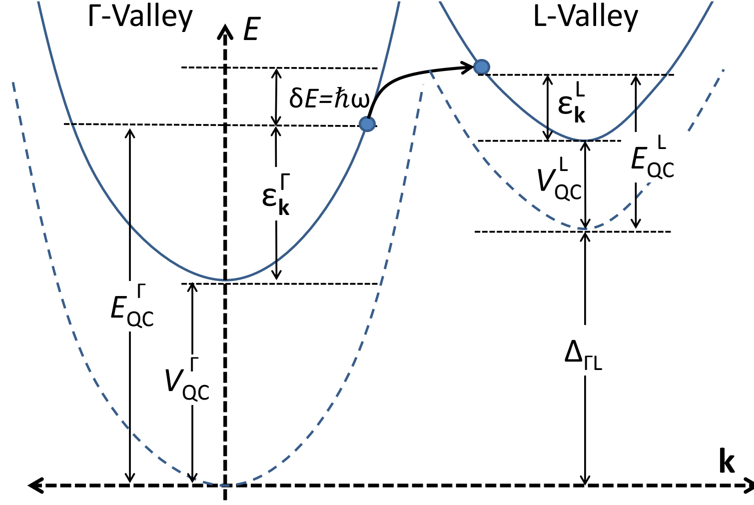


Figure 2.7 – Selecting the choice of final state for quantum-confined scattering. Quantum-confined intervalley scattering from the Γ-valley to an L-valley via absorption of an optical phonon of energy $\delta E = \hbar\omega$ in $In_{0.53}Ga_{0.47}As$, whereby the choice of final state scattering rate becomes a function of the total energy $E_{\text{QC}}^{\Gamma} = \varepsilon_{\mathbf{k}}^{\Gamma} + V_{\text{QC}}^{\Gamma}$, where the QCP V_{QC}^{Γ} raises the initial and, thus, final state energies relative to the uncorrected valley edges alike.

final quantum-corrected state energy for intervalley optical phonon absorption, and identifying the energies needed for the quantum-corrected scattering rate, is shown in Fig. 2.7. Consistent with the underlying localized-particle MC method, this correction is implemented in terms of the final-state valley QCPs as a local function of $\mathbf{r}$, $V_{\text{QC}}^{f}(\mathbf{r})$. This approach has been used previously by our group[43,44] for 1D confinement in 2D MC simulations, but it is extended here in this study to 2D confinement in 3D simulations.

In Fig. 2.8 we illustrate our quantum-corrected scattering model applied to phonon scattering rates (+QCS), as well as the contribution from SR scattering (+SR) as discussed in the next section. The corrected rates

are compared to the reference bulk rate (bulk) for electrons in the Γ-valley of $In_{0.53}Ga_{0.47}As$. The employed QCPs correspond to the channel center of a 6 nm wide fin geometry at equlibrium threshold conditions leading to moderately high values of $V_{\rm QC}^{\Gamma} = 375$ meV, and smaller values of $V_{\rm QC}^{\rm L}$ and $V_{\rm QC}^{\rm X}$, 150 meV and 100 meV, respectively. Not only are the rates of allowed scattering processes enhanced conistent with the $V_{\rm QC}^{\Gamma}$ shift in the energy argument, but the onset of intervalley scattering, and thus the intervalley transfer (IVT) of electrons, is reduced in energy by the difference in valley QCPs.

2.4.4 Modeling quantum-confinement-dependent surface-roughness scattering

Surface-roughness (SR) scattering is calculated formally as a function of the effects of variation in the surface location and the quantum-confined energies, which makes it a candidate for approximation via QCPs. We motivate our final use of our position-, valley-, and orientation-dependent QCPs to model SR scattering by considering some limiting behaviors of SR scattering rates in well-known systems. In quantum wells defined by triangular confining potentials, such as in inversion layers in planar MOSFETs, the SR scattering rate varies as $S_{\rm SR} \propto E_{\rm eff}^2$ where $E_{\rm eff}$ is the interface-normal effective electric field defining the triangular well.[74–76] Since the eigenenergies of the confined states E_i have an $E_{\rm eff}^{3/2}$ dependence on the effective electric field, we note that the SR scattering rate varies as E_i^3. In narrow infinite square wells of width w, SR scattering rates have been observed[77,78] to obey $S_{\rm SR} \propto w^{-6}$, while the eigenenergies E_i have a w^{-2} dependence. Again we note that SR scattering

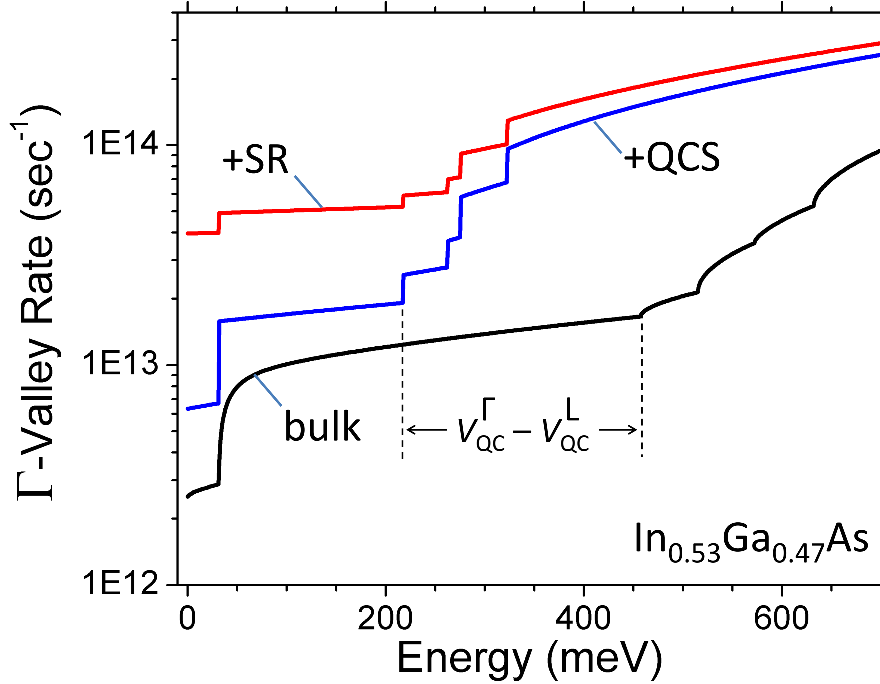


Figure 2.8 – Enhanced scattering rates due to quantum confinement. Calculated $In_{0.53}Ga_{0.47}As$ Γ-valley scattering rates comparing the bulk rate against those including quantum-confined scattering (+QCS) and additional surface-roughness scattering (+SR) in the device channel center of a 6 nm wide fin under equilibrium conditions at threshold. Not only are the rates increased due to quantum confinement, but intervalley scattering, and thus the intervalley transfer (IVT) of electrons, occurs at lower kinetic energies versus the bulk case.

varies as E_i^3. We therefore postulate, at least as a first *ansatz*, a generalized approximate SR scattering rate for an electron at position $\mathbf{r}$ in valley g as

$$S_{\mathrm{SR}}^{g}(\mathbf{r}) \equiv C_{\mathrm{SR}} \times \left(V_{\mathrm{QC}}^{g}(\mathbf{r})\right)^{3} \tag{2.12}$$

for our quantum-corrected MC simulations. In addition to reproducing the basic confinement dependence of SR scattering in these two limits, this *ansatz* also scales to the required result for very wide wells approaching the classical

limit. There the SR is only significant for carriers near the surface and where our QCPs also would remain nonzero, modeling quantum-mechanical surface repulsion which remains no matter how wide the well, but is otherwise width independent.

In Eq. 2.12, while $\left(V^g_{\mathrm{QC}}(\mathbf{r})\right)^3$ represents the effects of quantum confinement on SR scattering, the lead coefficient C_{SR} represents the amount of actual SR. SR is a function of interface quality and, thus, will vary with both channel and dielectric material and even strain and growth conditions, particularly in nascent technologies. However, the lack of universal experimental results and rapid evolution in both FinFET and III-V technologies makes a calibration of C_{SR} for these systems problematic. For this work, we calibrated C_{SR} to reproduce known experimental SR scattering rates obtained for a planar Si/thermal SiO_2 interface channel, having also considered confined phonon scattering in both cases.[73] This approach likely is optimistic for Si FinFETs and, more so, for $In_{0.53}Ga_{0.47}As$ FinFETs. However, it also provides another control for our simulations, allowing us to focus on the intrinsic properties of the two host material systems. However, although the assumed SR represented by C_{SR} is held constant, the behavior of the quantum confinement—represented here via the QCPs—and, thus, the actual SR *scattering* will vary with material as well as energy valley and valley orientation.

Here, since the interface roughness interaction with the channel wave function is non-local along the channel, we have chosen an elastic long-range non-randomizing polar optical phonon-like selection procedure for the final

state after SR scattering,[37,38] calibrated as noted above, which should be sufficient for the purposes of this work. We have plotted the thus-calculated SR scattering rate contribution for the electrons in the Γ-valley of $In_{0.53}Ga_{0.47}As$ of the same FinFET channel in Fig. 2.8. Initially the SR scattering more than doubles the scattering rate despite the likely underestimated amount of SR. However, with increasing energy, the randomizing intervalley phonon scattering processes soon become the dominant scattering mechanisms again.

2.5 Model comparisons

2.5.1 Illustrative impact of included quantum corrections on III-V and Si FinFET performance

We illustrate the QCs, and the effects of degenerate populations and quantum confinement, by benchmarking $In_{0.53}Ga_{0.47}As$ devices against industry standard Si devices, and by isolating the impact of each quantum correction. The FinFET sidewall orientation is $(100)/\langle 100 \rangle$ and the reference FinFET geometry is described in detail in Section 2.2. We compare the two materials' respective transfer curves $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ in Fig. 2.9 and their transconductances and ON-state currents in Table 2.1 sampled at the overdrive gate voltage of $V_{\mathrm{ON}} = V_{\mathrm{GS}} - V_{\mathrm{T}} = 350$ mV, while adding modeled quantum effects one at a time. We also analyze these results in terms of a quasi-ballistic representation of the current motivated by Lundstrom[21,22]

$$I_{\mathrm{DS}} = q\, n_{\mathrm{b}}\, v_{\mathrm{inj}}\, \gamma \;, \tag{2.13}$$

where q is the fundamental charge and, by definition, n_b is the total cross-sectional charge density at the top of the source-to-channel potential barrier. The injection velocity v_inj is the average velocity along the channel of incident charge carriers (those moving toward the drain) at the barrier-top, and γ is the injection efficiency. The injection efficiency is

$$\gamma = \frac{1 - R_j}{1 + R_n} \,, \tag{2.14}$$

with a distinction made here between reflection in terms of current j reflection in the numerator and charge n reflection in the denominator. The ratio of reflected ($-$) current to incident ($+$) current is $R_j \equiv \langle n_- \rangle \langle v_- \rangle \big/ \langle n_+ \rangle \langle v_+ \rangle$, the ratio of reflected charge to incident charge is $R_n = \langle n_- \rangle \big/ \langle n_+ \rangle$, $\langle v_+ \rangle \equiv v_\mathrm{inj}$, and $\langle n_+ \rangle + \langle n_- \rangle \equiv n_\mathrm{b}$, with all parameters once again measured at the barrier-top. Meanwhile, n_b depends on the gate overdrive voltage, controlled by the series combination of the dielectric capacitance and the channel quantum (DOS) capacitance.

$In_{0.53}Ga_{0.47}As$ strongly outperforms Si under classical MC simulation assumptions (Fig. 2.9a and b, CL), consistent with a smaller transport effective mass in its Γ-valley than in the six equivalent Si Δ-valley carriers. Since degeneracy is not considered in our CL model, both devices exhibit similar gate capacitances, and the 3$\times$ difference in I_ON can be explained by the roughly 3$\times$ larger injection velocity of $In_{0.53}Ga_{0.47}As$ versus Si in the semi-classical limit. The observed injection velocity of $In_{0.53}Ga_{0.47}As$ is curbed, however, by the large non-parabolicity constant α_Γ of its Γ-band, especially at larger kinetic energies, where α_γ can decrease incident velocities as much as 50%.

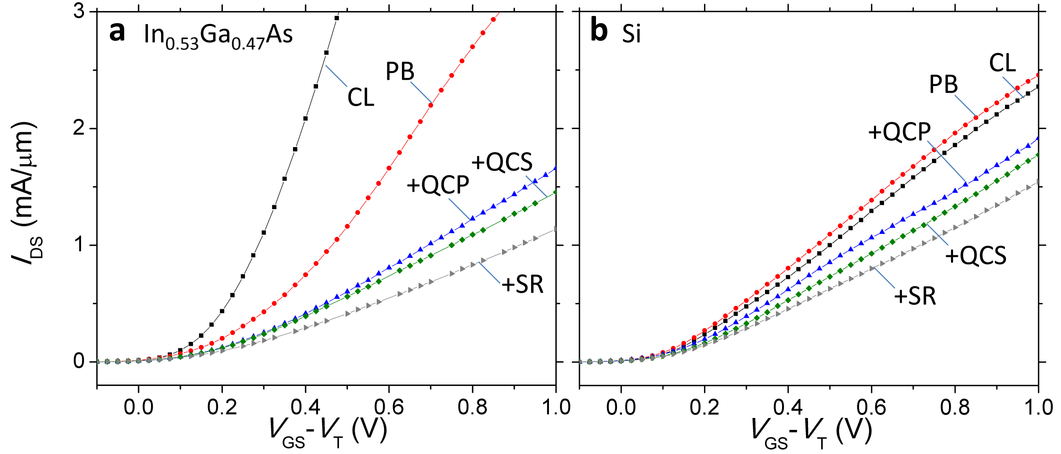


Figure 2.9 – $(100)/\langle 100 \rangle$ FinFET performance comparison of $In_{0.53}Ga_{0.47}As$ versus Si with varying quantum-corrected models. **(a)** As the Pauli-blocking (PB) of scattering, and then quantum-correction potentials (+QCP) are added, as well as QC-dependent phonon (+QCS) and surface-roughness scattering (+SR), the performance of the $In_{0.53}Ga_{0.47}As$ FinFET device (left) is substantially degraded relative to classical (CL) expectations. **(b)** Si is shown to be more robust against carrier degeneracy and quantum-confinement effects (right), although its performance is also moderated compared to classical expectations.

The inclusion of the PB of scattering and resulting degenerate statistics then greatly moderates the performance of $In_{0.53}Ga_{0.47}As$ relative to Si (Fig. 2.9a and b, PB). Indeed, PB-Si actually performs better than in the CL-Si case due to increased thermal velocities of degenerate electrons injected from the source versus the non-degenerate CL case. Although the same is true of $In_{0.53}Ga_{0.47}As$, whose injection velocity increases from 3.2 to 4.3×10^7 cm/sec considering degeneracy, its cross-sectional charge density is reduced 85% in the PB case due to severely reduced quantum (DOS) capacitance in this light-effective mass material.

In addition to degeneracy, we then employ the quantum-correction po-

tentials (Fig. 2.9a and b, +QCP) to model our first level of quantum confinement. The performance of $In_{0.53}Ga_{0.47}As$ is diminished as the charge density is redistributed among the energy valleys through scattering via reduced intervalley separations. In the PB case, 99% of the sampled carriers at the barrier-top occupy Γ-states, which is reduced to only 26% in the +QCP case. This corresponds to a reduction in the injection efficiency from $\gamma_{\mathrm{PB}} = 90\%$ down to $\gamma_{\mathrm{QCP}} = 43\%$, coinciding with the greater occupation of heavier-mass satellite valleys, which experience worse backscattering. Notably the thermal velocity also dwindles from 4.3 down to 2.5×10^7 cm/sec. However, there is a competition here, as occupation of the L- and X-valleys rapidly increases the differential quantum capacitance of $In_{0.53}Ga_{0.47}As$, and gate control over n_{b}, correspondingly. Si is more robust to quantum confinement in contrast, however. In this $(100)/\langle 100\rangle$ fin configuration, the confinement-induced

$In_{0.53}Ga_{0.47}As$

	CL	PB	+QCP	+QCS	+SR
g_{M} (mA/μm/V)	9.90	3.17	1.62	1.50	1.09
I_{ON} (mA/μm)	1.57	0.57	0.33	0.31	0.23

Si

	CL	PB	+QCP	+QCS	+SR
g_{M} (mA/μm/V)	2.51	2.71	2.30	2.04	1.63
I_{ON} (mA/μm)	0.60	0.67	0.50	0.43	0.37

Table 2.1 – ON-state performance by model. The transconductance g_{M} and the drain current I_{ON}, sampled at $V_{\mathrm{DS}} = 600$ mV and $V_{\mathrm{ON}} = V_{\mathrm{GS}} - V_{\mathrm{T}} = 350$ mV. Both have been normalized by the fin perimeter, $(2T_{\mathrm{fin}} + W_{\mathrm{fin}})$.

degeneracy-splitting of the Δ-bands results in larger occupation of the Δ_2-valleys oriented normal to the sides of the fin, which have the lightest transport effective mass and higher thermal velocity. This partially offsets worse quantum (DOS) capacitance, leading to an overall smaller relative reduction in current for Si compared to $In_{0.53}Ga_{0.47}As$. These results in both materials highlight the importance of a valley-by-valley treatment of quantum corrections.

We next add the quantum-confined scattering using the same QCPs (Fig. 2.9a and b, +QCS). Performance degrades in both material systems due to significantly enhanced scattering rates, although this quantum correction had the smallest effect on the overall drive current performance in the considered device structure.

Finally, we add the surface-roughness scattering (Fig. 2.9a and b, +SR). I_{ON} and g_{M} suffer in both materials, albeit worse for $In_{0.53}Ga_{0.47}As$. The light-mass Γ-valley carriers, having the largest QCPs, experience the worst SR scattering, reducing the $In_{0.53}Ga_{0.47}As$ overall injection efficiency to $\gamma_{\mathrm{SR}} = 33\%$ down from $\gamma_{\mathrm{QCS}} = 43\%$. Nevertheless, the effect of SR scattering is smaller than one might otherwise expect, even allowing for the proximity to the ballistic limit. Previously our group has observed in both MC[43,44] and quantum transport simulations[27] that SR and phonon scattering rates are not simply additive concerning the effects on channel transport, and that the whole is less than the sum of the parts. (E.g., phonon emission reducing a carrier's energy to below the barrier-top can prevent subsequent SR induced

back-scattering.) Also, because of calibration to the amount of SR to the more ideal planar Si-SiO_2 interface, the estimates of SR and, thus, SR scattering, therefore, are likely conservative.

2.5.2 Effects of satellite valleys on III-V electron transport

In terms of the effects of occupation of the peripheral valleys in $In_{0.53}Ga_{0.47}As$, there is ambiguity for two reasons, however. First, at the considered valley offset of $E_{\Gamma L} = 487$ meV and doping density of $N_D = 5 \times 10^{19}$ cm^{-3}, the Fermi level is sufficiently high in the source and drain that we inject directly into the bottom of the satellite L-valleys in the simulations above. Thus the mechanism by which carriers reach the peripheral valleys within the channel, whether by this injection from the boundaries or by scattering within the simulation region, is inconclusive. Therefore the role of intervalley scattering on transport is unclear. Moreover, there is significant uncertainty in reported intervalley separations, with some models suggesting[58] sufficiently large separations that carriers would remain almost entirely localized to the Γ-valley. We address both of these issues by reducing the assumed source and drain doping to 1×10^{19} cm^{-3}, below that achievable through post-crystal growth doping, such that the Fermi level at $E_F - E_C = 290$ meV remains well below the peripheral valley edges and eliminates injection into the L-valleys from the contacts. We then perform simulations with and without intervalley scattering to isolate its impact. Including intervalley scattering provides understanding of the role of intervalley scattering within the simulation

region on transport provided the peripheral valleys are available for occupation. Excluding intervalley scattering allows analysis of device behavior absent the availability of peripheral valleys for occupation.

With the intervalley separations as considered previously, intervalley scattering turned on, but the S/D doping reduced to 1×10^{19} cm^{-3}, we still find large occupation of the satellite valleys in the channel in the ON-state, where the gate and not the S/D doping controls the carrier density and the intervalley separation is reduced by quantum confinement. Roughly 58% of the charge in the channel occupies L- and X-states, although all the charge from the source and drain is injected into the Γ-valley. This result points to the continuing need to model intervalley scattering within the simulation region—absent contact injection, and perhaps even with it, ballistic treatments of transport cannot model the role of energetically available peripheral valleys.

However, the absence of energetically available peripheral valleys—or ignoring scattering to available ones—does have a significant impact on device performance. Transfer curves are shown in Fig. 2.10 for devices with full QCs comparing $In_{0.53}Ga_{0.47}As$ FinFETs with S/D doped to $N_D = 1 \times 10^{19}$ cm^{-3} with intervalley scattering and without (no IV). Transfer curves for Si and $In_{0.53}Ga_{0.47}As$ devices from Fig. 2.9 doped to $N_D = 5 \times 10^{19}$ cm^{-3} also are shown for reference. As stated previously, there are competing effects of intervalley scattering on device performance. With intervalley scattering turned on, overall channel injection efficiency is poor ($\gamma = 36\%$), having significant occupation of the high-scattering rate peripheral valleys, as well as sub-

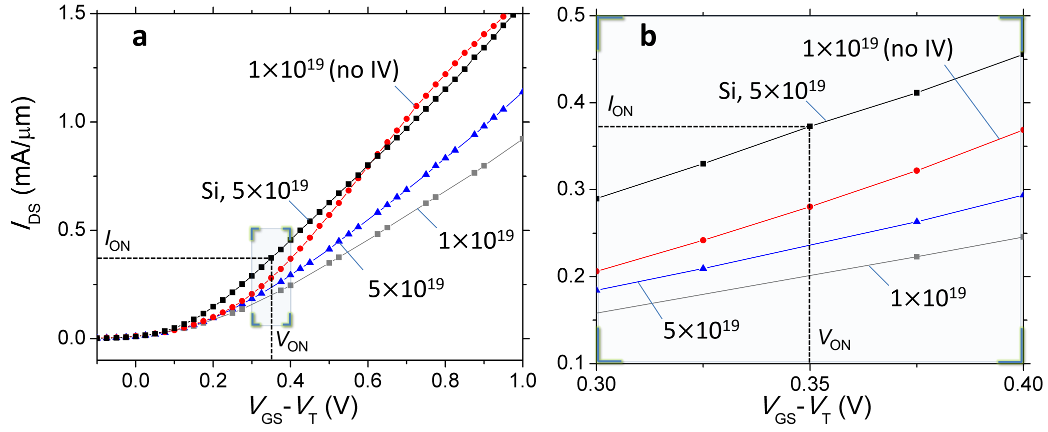


Figure 2.10 – The role of satellite valleys in $In_{0.53}Ga_{0.47}As$ transport. The individual curves are labeled by their doping density and are $In_{0.53}Ga_{0.47}As$ channels unless noted as Si. **(a)** With a larger injection velocity, the $In_{0.53}Ga_{0.47}As$ device having no intervalley scattering (no IV) outperforms the device including the peripheral valleys, although it has a smaller quantum capacitance. It also outperforms the $In_{0.53}Ga_{0.47}As$ device having a larger doping density of 5×10^{19} cm^{-3} and even crosses-over Si at larger gate biases. **(b)** Si, however, still has the largest ON-current at the given drive voltage determined by ITRS predictions for future CMOS.

stantial intervalley scattering that, itself, produces back-scattering. In contrast, the quantum capacitance increases with the occupation of the satellite valleys, leading to an overall larger carrier population in the channel for a given gate voltage than otherwise would be found. However, the larger gate control of the device including intervalley scattering is not enough to offset the greater benefits enjoyed by the device where the peripheral valleys are no longer available (Fig. 2.10, no IV), whose large injection velocity of 4.3×10^7 cm/sec is great enough to outperform even the $In_{0.53}Ga_{0.47}As$ device doped to $N_D = 5 \times 10^{19}$ cm^{-3} and included intervalley scattering. This intervalley-scattering-free device even outperforms Si at larger gate biases well

above threshold, as its DOS and, thus, quantum capacitance grows with increasing energy. However, at the considered ITRS voltages where the quantum capacitance is lower, it still exhibits a smaller I_{ON} where there is a more rapid turn-on of the drain current in the Si device just above threshold. Notably, with much higher available doping (e.g., $N_{\mathrm{D}} > 5 \times 10^{20}$ cm^{-3} in Ref. 79), Si devices would be expected to have still better performance relative to III-V devices due to reduced S/D resistances and better contacts.

Our device results with full quantum corrections can be compared qualitatively to previous studies. However, quantitative comparison to Si-FinFET experimental results is difficult due to, e.g., different geometries,[79,80] strain considerations, and uncertainty in surface roughness. The same can be said regarding existing III-V FinFET experimental devices and technologies.[11,81] It is also difficult to make quantitative comparisons among the experimental results for similar reasons. In addition, our simulations excluded realistic modeling of contact resistance, which was done to isolate the respective channel behavior across materials to study the relevant transport physics considered. For comparison to others simulations, there are no other MC simulators currently that treat the array of quantum effects modeled in this work. However, ballistic quantum transport simulations have exhibited the same qualitative trend exhibited here, namely that Si devices may continue to outperform III-V devices moving forward.[82,83]

2.6 Discussion

In this chapter, we provided an ensemble MC methodology with the most complete set of quantum corrections in terms of the number of quantum mechanical effects addressed: far-from-equilibrium degenerate statistics and associated PB of scattering and limited quantum (DOS) capacitance, and confinement effects including altered energy separations between energy valley minima and quantum-confinement-dependent phonon and surface-roughness scattering, in addition to electron redistribution in real space and modified source-to-channel potential barriers. We developed each of our methods individually within this article with a focus on our new contributions and discussed their relevance in terms of nanoscale n-channel FinFET device performance, illustrated through application to example $In_{0.53}Ga_{0.47}As$ and Si devices.

For the treatment of the PB of scattering, we avoid the common use of Fermi-Dirac equilibrium electron distributions to approximate the final state occupation probabilities. Instead, our method directly samples even far-from-equilibrium forward-going and backward-going local electron populations as a function of energy valley and energy, and uses those occupation probabilities self-consistently to model PB. We also introduced sub-carriers (fractional carriers) to suppress classical molecular dynamics carrier-carrier interactions that inherently do not consider the Pauli exclusion principle, with the added benefits of enhancing simulation statistics and minimizing self-forces. Our method of calculating degenerate carrier populations was shown to limit to Fermi-Dirac statistics under equilibrium conditions, while flexibly adapting to

more complex distributions under bias.

We modeled the above-noted quantum-confinement effects via space, valley, and orientation-dependent quantum-correction potentials. In doing so, we extended to 3D a treatment of quantum-confined phonon and ionized-impurity scattering developed previously in-house, and found a versatile method for modeling surface-roughness scattering with these potentials that extends to arbitrary potential-well shapes, giving material, valley, and orientation-dependent SR scattering in various device geometries.

We showed that collectively these modeled quantum effects can substantially degrade or even eliminate otherwise expected benefits of considered $In_{0.53}Ga_{0.47}As$ devices over industry-standard Si devices, despite lower bulk electron masses, higher mobilities, and higher thermal velocities found in III-V materials, even while neglecting non-ideal contacts and reduced interface quality that are likely to be worse for III-Vs.

We note that it also may be possible to use quantum corrections in the ways described herein within simpler drift-diffusion or hydrodynamic simulators, albeit using more computationally efficient methods for calculating the potentials themselves.

Chapter 3

Device scaling studies

3.1 Dimensional scaling study

In an analysis of possible future logic device scaling trends, high performance n-channel Si and InGaAs FinFETs are simulated down to channel lengths $L_{\mathrm{ch}} = 9$ nm using ensemble Monte Carlo with state-of-the-art quantum corrections. Si-channel devices exhibit somewhat superior short channel behavior as evidenced by lower subthreshold swing $S = (\ln 10)\mathrm{d}V_{\mathrm{GS}}/\mathrm{d}(\ln I_{\mathrm{DS}})$ and drain-induced barrier lowering DIBL$= \mathrm{d}\Phi_{\mathrm{b}}/\mathrm{d}V_{\mathrm{DS}}$, where Φ_{b} is the channel potential barrier. In addition, InGaAs devices exhibit just comparable drive currents to Si. The less-than-hoped-for InGaAs ON-state performance is a result of more severe adverse quantum degeneracy and confinement effects.

3.1.1 Device structure and results

The simulations of this Chapter employ a slightly different device as compared to the FinFET of Chapter 2. This device was developed in collab-

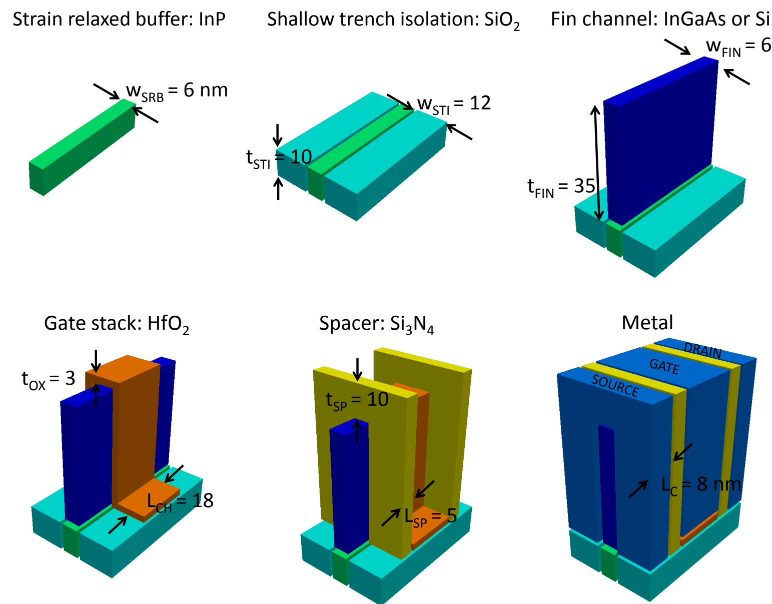


Figure 3.1 – Edge view of simulated FinFET transistor showing each material region. The strain relaxed buffer provides a defect-free fin-channel, while the shallow trench isolation electrically isolates adjacent FinFETs. The source and drain reservoirs are doped with a grading falling off as 1 decade/nm while the fin-channel is otherwise undoped. Including realistic spacer and metal regions effectively captures the relevant electrostatics. All dimensions are shown in nm.

oration with *GLOBALFOUNDRIES USA*, Inc. and uses a flat narrow fin as opposed to a fin which opens up into the S/D reserviors (compare Fig. 3.1 to Fig. 2.1 of Chapter 2).

We first simulated the device of Fig. 3.1 with various channels: Si (100) surface/⟨100⟩ channel, Si (100)/⟨110⟩, $\mathrm{In_{0.53}Ga_{0.47}As}$, and $\mathrm{In_{0.70}Ga_{0.30}As}$ (Fig. 3.2, left), where the InGaAs devices have shown little channel orientation dependence. All channel materials lead to similar peak transconductances g_{M}, but Si (100)/⟨100⟩ provides a superior ON-sate current I_{ON} (Fig. 3.2, right),

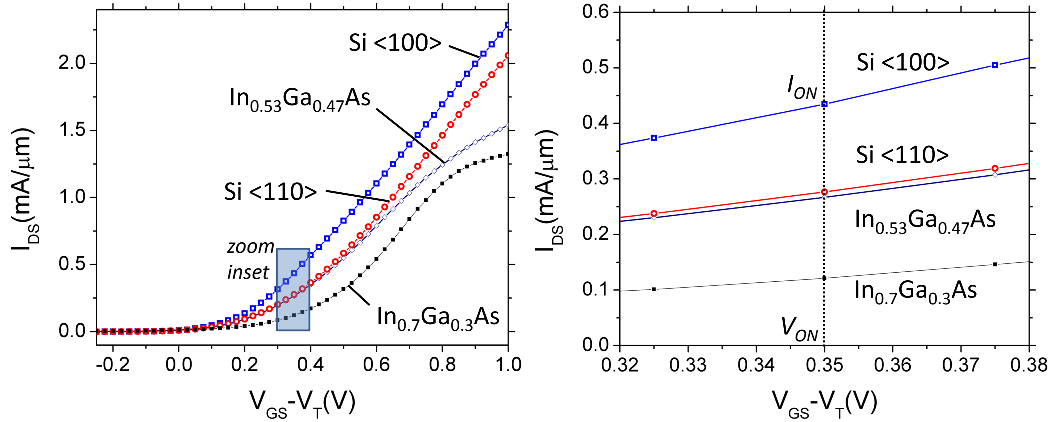


Figure 3.2 – Comparing ON-state performance. Transfer characteristic $I_{\rm DS}$ vs. $V_{\rm GS}-V_{\rm T}$ for FinFETs from Fig. 3.1 with various channels. The right figure is a zoomed view of the overdrive voltage $V_{\rm ON}=V_{\rm GS}-V_{\rm T}=350$ mV (with end-of-the-ITRS[65] supply $V_{\rm DD}$ and threshold $V_{\rm T}$ voltages 0.6 V and 0.25 V). Si $\langle 100\rangle$ exhibits the largest ON-state current $I_{\rm ON}$, while Si $\langle 110\rangle$ and $In_{0.53}Ga_{0.47}As$ have nearly identical signals. The $I_{\rm ON}$ advantage for Si $\langle 100\rangle$ and disadvantage for $In_{0.70}Ga_{0.30}As$ are a result of more and less abrupt turn-on, respectively, beyond the assumed threshold current of $I_{\rm DS}|_{V_{\rm T}}=0.01$ mA/μm.

substantially due to more abrupt turn-on after reaching the assumed threshold current of 0.01 mA/μm.

Si devices have the advantage of greater source and drain (S/D) doping than the InGaAs devices ($N_{\rm D}=2\times10^{20}$ cm^{-3} vs. $N_{\rm D}=1\times10^{19}$ cm^{-3} in the simulations of Fig. 3.2). However, there is limited benefit to increasing the doping further in the InGaAs devices, at least in terms of peak $g_{\rm M}$, while Si devices exhibit greater benefit to complement their greater available dopant concentrations (Fig. 3.3). In InGaAs, as we learned from Chapter 2, increasing the S/D doping has diminishing returns once the Fermi level moves into the satellite L-valleys. Conversely, in Si, increasing the doping continues to

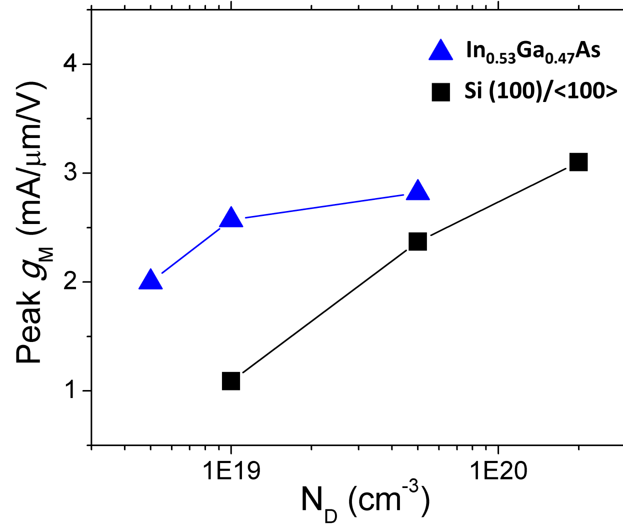


Figure 3.3 – Dependence of channel peak transconductance on source/drain doping. $In_{0.53}Ga_{0.47}As$ exhibits higher peak g_M than Si $\langle 100 \rangle$ does at the same doping. However, the advantage decreases with increasing doping, and greater doping is available in Si devices.

enhance the average thermal velocity, which gets a bonus contribution as the S/D become more and more degenerate.

Moreover, on this scale, the Si devices more than meet short channel metrics of 70 mV/decade subthreshold swing S and 70 mV/V DIBL with a $L_{CH} = 3W_{fin}$ design rule, while $In_{0.53}Ga_{0.47}As$ devices require slightly longer channels to meet these metrics (Fig. 3.4). In general, III-V devices exhibit poor SCE compared to Si due to their larger dielectric permittivity which effectively couples the S/D to the channel potential more strongly.

We then used the $L_{CH} = 3W_{fin}$ design rule to scale the FinFET structure of Fig. 3.1 according to ITRS[65] predictions for future nodes. Comparing the common Si (100)/$\langle 110 \rangle$ option to $In_{0.53}Ga_{0.47}As$ (Fig. 3.5), Si devices main-

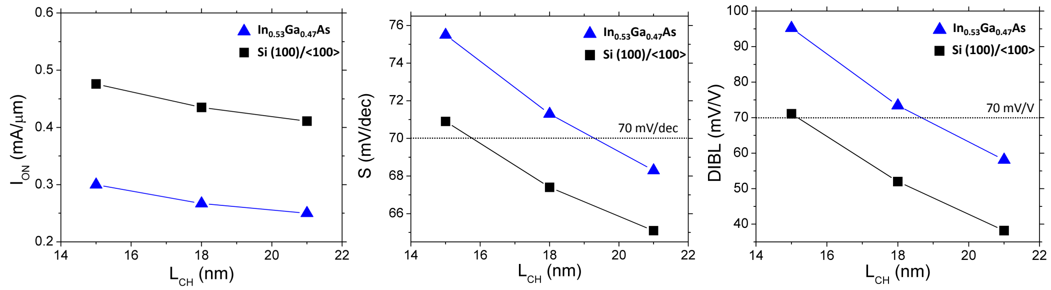


Figure 3.4 – Channel length scaling. Holding all parameters constant including the 6 nm fin width, the device of Fig. 3.1 is scaled via $L_{\mathrm{CH}} = 21$, 18, and 15 nm respectively. Si $(100)/\langle 100 \rangle$ outperforms $\mathrm{In}_{0.53}\mathrm{Ga}_{0.47}\mathrm{As}$ in all metrics. Setting a benchmark for electrostatic control of $S < 70$ mV/dec and DIBL< 70 mV/V, a design rule of $L_{\mathrm{CH}} = 3W_{\mathrm{fin}}$ is more than sufficient for Si, while $\mathrm{In}_{0.53}\mathrm{Ga}_{0.47}\mathrm{As}$ requires somewhat longer channels, approximately $L_{\mathrm{CH}} = 3(W_{\mathrm{fin}} + 1)$.

tain better electrostatic control, although diminishing in both systems. Both channels have comparable I_{ON} through 2022. The excellent Si $(100)/\langle 100 \rangle$ performance and robust scaling trends even up to 2022 could possibly motivate multi-orientation CMOS in future nodes if necessay (i.e., n-channel $(100)/\langle 100 \rangle$ and p-channel $(100)/\langle 110 \rangle$ on the same chip).

3.1.2 Discussion

For future FinFETs, the Monte Carlo simulations of this section suggest that proposed InGaAs n-channel devices would exhibit roughly comparable drive currents and somewhat inferior short-channel electrostatic control as compared to Si FinFETs. A combination of negative quantum effects substantially moderates otherwise expected advantages of the InGaAs devices based on bulk non-degenerate thermal velocities.

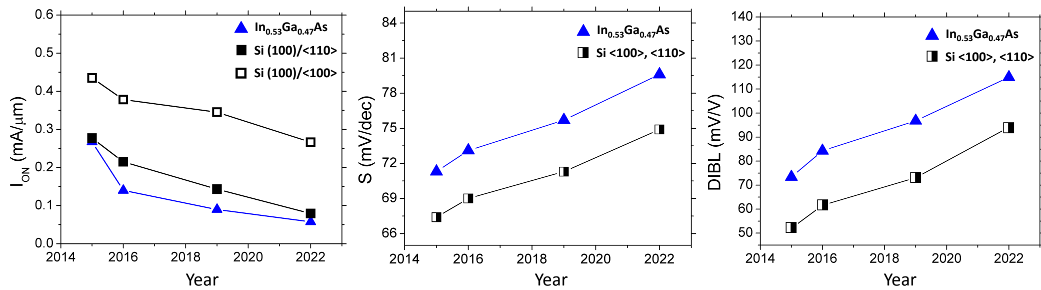


Figure 3.5 – FinFET device scaling. The device from Fig. 3.1 is then scaled via $L_{\mathrm{CH}} = 3W_{\mathrm{fin}}$ according to possible device nodes at 2016 ($L_{\mathrm{CH}} = 15$ nm, $W_{\mathrm{fin}} = 5$ nm), 2019 ($L_{\mathrm{CH}} = 12$ nm, $W_{\mathrm{fin}} = 4$ nm), and 2022 ($L_{\mathrm{CH}} = 9$ nm, $W_{\mathrm{fin}} = 3$ nm). ON-state drive current remains comparable between Si $(100)/\langle 110 \rangle$ versus $\mathrm{In_{0.53}Ga_{0.47}As}$, while the Si $(100)/\langle 100 \rangle$ device boasts a roughly $3\times$ greater drive current. The Si devices continue exercising somewhat better short channel control in terms of subthreshold swing S and DIBL, with both metrics becoming worse for shorter channels. In terms of I_{ON} it might prove more feasible to move to Si $\langle 100 \rangle$ channels rather than integrate $\mathrm{In_{0.53}Ga_{0.47}As}$ material processes. Currently, nMOS Si $\langle 110 \rangle$ channels are incorporated to match pMOS technologies, which for which Si $\langle 110 \rangle$ orientation is preferred.

3.2 Dielectric scaling study and the impact of the gate oxide effective mass on channel quantum confinement

FinFET geometries have been developed for the sub-22 nm regime to extend Si-CMOS scaling via improved electrostatics compared to planar technology. Moreover, engineers have incorporated high-k oxide gate stacks. Beyond leakage current, less discussed is the impact of the gate oxide's complex band structure on the device performance. However, it defines the boundary condition for the channel wave function at the interface, which, in turn, affects the quantum-confinement energy for channel electrons. Here we show that the ON-state performance of n-channel FinFETs may be sensitive to the oxide's complex band structure, especially with light-mass III-V channel materials,

such as $In_{0.53}Ga_{0.47}As$. We study this effect using the ensemble semi-classical Monte Carlo device simulator of Chapter 2 which includes advanced quantum corrections for degeneracy and confinement effects. Our simulations suggest that using a surface oxide with a heavy effective mass may lower the channel carrier confinement energies, mitigating unwanted quantum side-effects that hinder device performance. Ultimately, future high-k stacks may benefit from oxide gate stack heterostructures balancing effective mass and dielectric permittivity considerations.

3.2.1 Alternate gate oxide materials for scaling CMOS

Physical limits[84] have become the pressing challenge for Si planar metal-oxide-semiconductor-field-effect-transistors (MOSFETs) when considering future device nodes. Simply reducing device dimensions is no longer sufficient, as nano-scale planar CMOS devices suffer from debilitating short-channel effects (SCE).[10] Poor electrostatic control can lead to substantial drain-induced barrier lowering (DIBL) and degraded device subthreshold swing (S), as well as limiting transconductance (g_{m}), contributing to poor ON/OFF ratios.

Incorporation of new three-dimensional (3D) geometries, especially that of the 3D fin-shaped MOSFET (FinFET),[9] has extended the scaling life of MOSFETs via improved gate control, as well as reduced on-chip surface area. FinFETs are likely to drive scaling in future device nodes.[12,13] Moreover, high dielectric constant (high-k) oxide stacks have improved gate-to-channel capacitive coupling, allowing for reduced effective electrostatic oxide thicknesses

(EOT), despite increased physical thicknesses to prevent tunneling.

Alternate channel materials also are being considered widely for future device nodes, such as III-Vs for n-channel devices in particular.[16] III-Vs may provide a performance boost versus Si in the ON-state via higher bulk mobilities and higher thermal injection velocities associated with lighter masses, the latter being more important approaching the ballistic limit. In particular, $In_{0.53}Ga_{0.44}As$, which is lattice-matched to fabrication-friendly InP,[15] provides Γ-valley electrons with an effective mass of $m^*_\Gamma = 0.044\ m_e$. However, for deeply-scaled devices a host of quantum-mechanical effects may diminish the otherwise expected advantages. Considerations of the Pauli exclusion principle lead to reduced quantum/density-of-states (DOS) capacitance (C_q), while quantum confinement reduces intervalley separations, increases phonon scattering rates, and decreases electrostatic capacitance as the carriers are shifted further from the interface.

In this section, we show that the choice of gate oxide may affect device performance not just through the electrostatics but also through these quantum-confinement effects, as it establishes the boundary conditions on the channel wave function. To this end, we use the ensemble semi-classical Monte Carlo (MC) simulation tool with advanced quantum corrections that was described in Chapter 2 to model nano-scale $In_{0.53}Ga_{0.47}As$ FinFETs. We find that use of lighter gate oxide effective masses (m^*_ox) actually leads to increased quantum-confinement effects, including increased scattering and intervalley transfer, and, thus, reduced ON-state performance. In this way, we find that

the otherwise expected advantage of higher-k materials may be minimized by often-associated lighter masses, and that oxide stacks combining higher-mass surface dielectrics with higher-k dielectric over-layers could prove optimal.

3.2.2 Gate oxide dielectric constant effects

$In_{0.53}Ga_{0.47}As$ transistors with the geometry of Fig. 3.1 are simulated with varying gate oxide dielectric constants. Fig. 3.6 shows the result of increasing the gate oxide dielectric constant from an Al_2O_3-like $\varepsilon_r = 7.8$ (gray curves) to a HfO_2-like $\varepsilon_r = 22.3$ (black curves), but otherwise using Al_2O_3-like parameters[85] in both cases.

As expected, a larger ε_r delivers better performance (Fig. 3.6a), emphasized by the larger peak transconductance $g_M = \text{MAX}\{dI_{DS}/dV_{GS}\}$, lowered OFF-state S, and DIBL= $d\Phi_b/dV_{DS}$, where Φ_b is the channel potential barrier (Table 3.1). The better performance can be understand in terms of the increased capacitive coupling of the gate to the channel (Fig. 3.6b).

Dielectric constant	$\varepsilon_r = 7.8$	$\varepsilon_r = 22.3$
EOT (nm)	1.0	0.35
g_M (mA/μm/V)	2.4	3.8
S (mV/dec)	76	67
DIBL (mV/V)	94	55

Table 3.1 – Effects of the variation of the gate oxide dielectric constant on FinFET performance. A larger gate oxide dielectric constant improves ON-state transconductance and sub-threshold behavior.

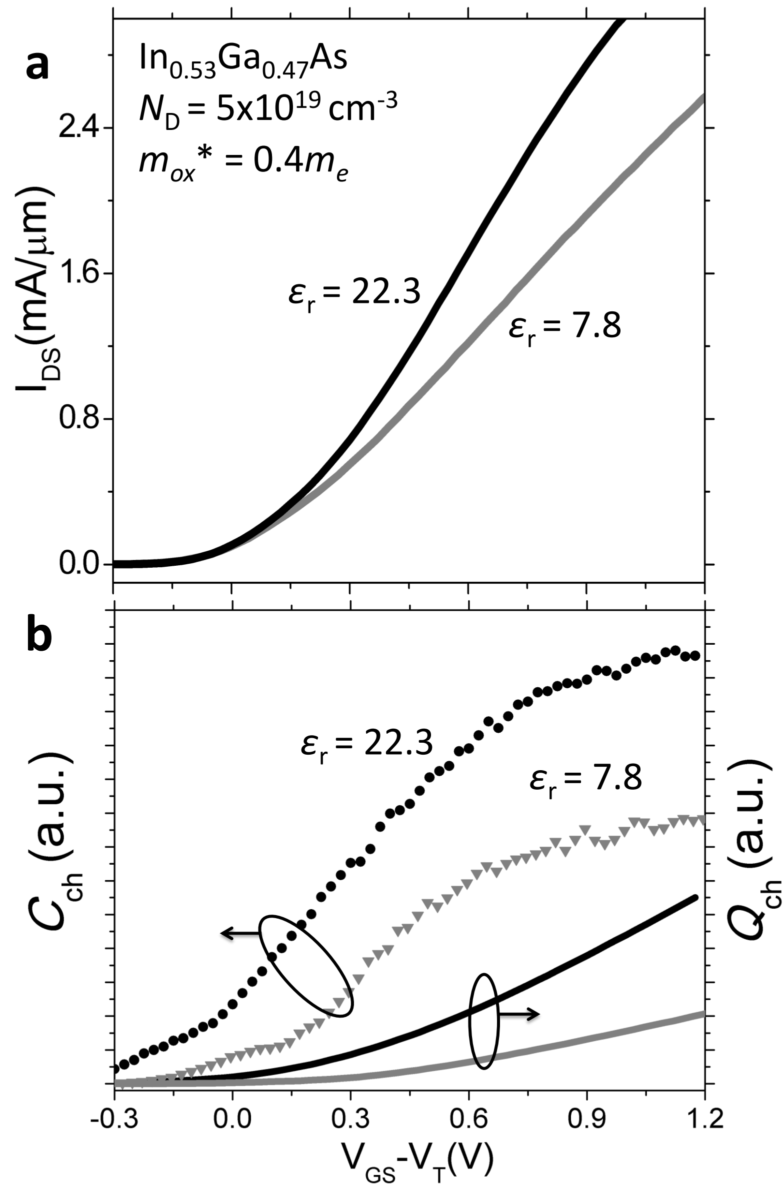


Figure 3.6 – Varying oxide dielectric constant. (a) $I_{\rm DS}$ vs. $V_{\rm GS} - V_{\rm T}$ for otherwise identical FinFETs having gate oxide dielectric $\varepsilon_{\rm R} = 7.8$ (gray) compared to $\varepsilon_{\rm R} = 22.3$ (black). **(b)** Increasing the dielectric constant (from gray to black) induces a larger channel inversion charge density $Q_{\rm ch}$ (right-axis) and capacitance $C_{\rm ch} = {\rm d}Q_{\rm ch}/{\rm d}V_{\rm GS}$ (left-axis).

Effective mass	$m^*_{\text{ox}} = 0.4m_{\text{e}}$	$m^*_{\text{ox}} = 0.08m_{\text{e}}$
EOT (nm)	1.0	1.0
g_{M} (mA/μm/V)	2.4	1.73
S (mV/dec)	76	75
DIBL (mV/V)	94	94

Table 3.2 – Effects of the variation of the gate oxide conduction band effective mass on FinFET performance. A larger gate oxide conduction band effective mass improves ON-state performance by reducing the quantum confinement experienced by electrons in the FinFET channel.

3.2.3 Gate oxide effective mass effects

Fig. 3.7 shows the result of decreasing the gate oxide effective mass from an Al_2O_3-like $m^*_{\text{ox}} = 0.4m_{\text{e}}$ from Ref. 85 (gray curves) to a HfO_2-like $m^*_{\text{ox}} = 0.08m_{\text{e}}$ from Ref. 86 (black curves), but otherwise using Al_2O_3-like parameters in both cases.

Fig. 3.7a shows diminished drive current when lowering the oxide effective mass. Table 3.2 shows that the g_{M} is reduced by almost 30% relative to the device with higher m^*_{ox}. The culprit is increased quantum confinement for the device with lighter m^*_{ox}, as shown by the larger ensemble average quantum correction for carriers injected over the channel potential barrier-top (Fig. 3.2b, left-axis). In these simulations, the oxide's complex band structure is parameterized by its m^*_{ox} along with its barrier height. An increase in calculated quantum-confined bound state energies with reduction in m^*_{ox} can be seen as a by-product of the required mass-difference-induced discontinuity in the wave function derivative at the interface,[87] despite also increased barrier

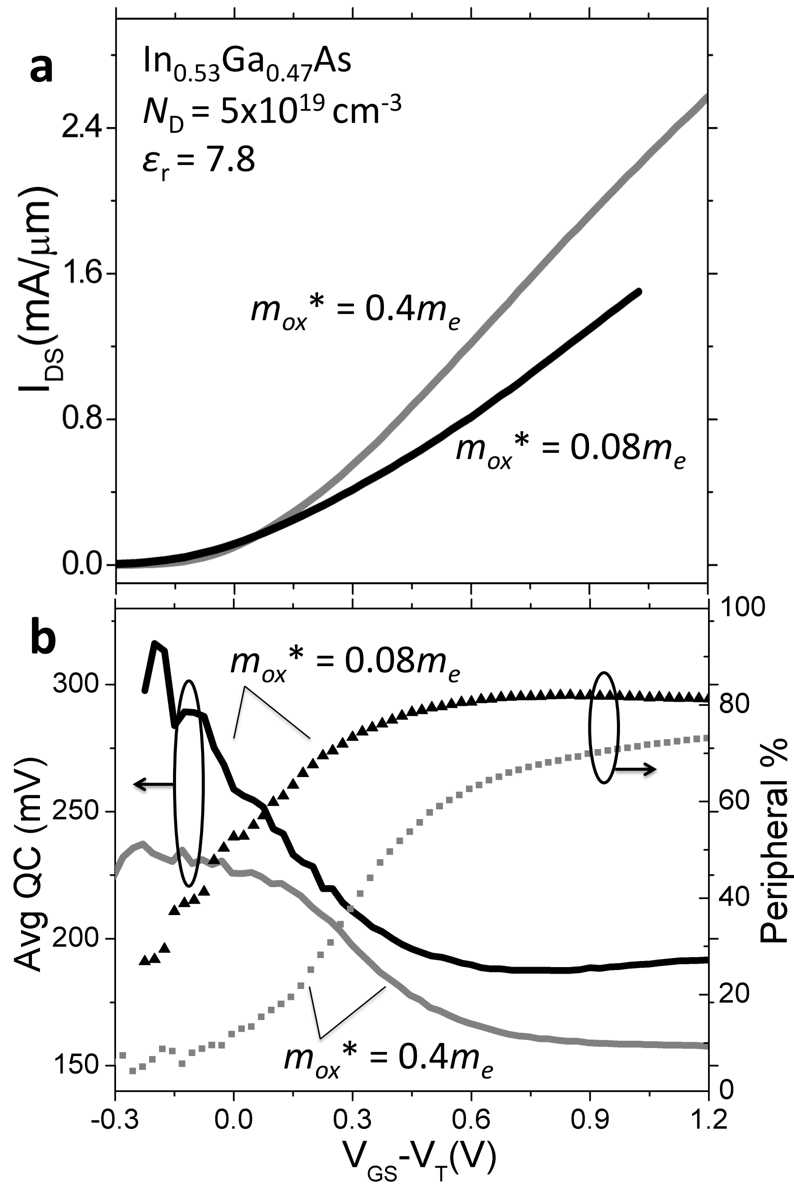


Figure 3.7 – Varying oxide effective mass. (a) Transfer curves comparing gate oxide effective mass $m^*_{\text{ox}} = 0.4$ (gray) to $m^*_{\text{ox}} = 0.08$ (black). **(b)** Decreasing the gate dielectric's effective mass increases the average quantum-confinement energy (left-axis) and occupation of heavy-mass peripheral valley states (right-axis).

penetration depths. (We note that our Schrödinger equation and wave function are discretized in our calculations, with the mass change reflected only in (Hermitian) changes in the inter-site hopping potentials. However, this approach preserves this basic effect of the impact of the mass change across the interface.)

The resulting increase in quantum corrections bolsters an already strong transfer of electrons to the heavy-mass peripheral valleys (Fig. 3.2b, right-axis) and degrades the current density. However, SCE are shown to be insensitive to confinement, as the S and DIBL are relatively unaffected (Table 3.2).

3.2.4 Gate oxide trade-off: alumina vs. hafnia

Finally, we compare two otherwise identical FinFET structures, changing only the gate oxide material from Al_2O_3 ($m^*_{\rm ox} = 0.4m_{\rm e}$, $\varepsilon_r = 7.8$, and an electron affinity[85] producing a 2.56 eV band offset within the electron-affinity rule) to HfO_2 ($m^*_{\rm ox} = 0.08m_{\rm e}$, $\varepsilon_r = 22.3$, and an electron affinity[86] producing a 2.46 eV band offset). The *physical* thickness of the gate oxide is kept at $t_{\rm ox} = 2$ nm in both cases as a control, although in practice a greater physical thickness would be expected for the higher-k material.

HfO_2, with a larger ε_r and therefore thinner EOT for a given physical gate oxide thickness, exhibits superior short channel control in terms of S and DIBL (Table 3.3). A more interesting comparison is the ON-state performance. Use of Al_2O_3 leads to weaker electrostatic coupling, and generates a smaller inversion charge density than the use of HfO_2 (Fig. 3.8, right-axis).

Gate oxide	Al_2O_3	HfO_2
ε_r (ε_0)	7.8	22.3
$m^*_{\mathrm{ox}}(m_{\mathrm{e}})$	0.4	0.08
EOT (nm)	1.0	0.35
g_{M} (mA/μm/V)	2.4	2.5
S (mV/dec)	76	68
DIBL (mV/V)	94	55

Table 3.3 – Effects of the variation of the gate oxide material stack on FinFET performance. Al_2O_3, with a heavier effective mass, is shown to compete effectively against HfO_2 in performance despite its lower dielectric permittivity.

However, with a heavier m^*_{ox}, use of Al_2O_3 produces smaller quantum corrections in our model, and thus relatively more Γ-valley carriers in the channel and larger carrier velocities as compared to the use of HfO_2 (Fig. 3.8, left-axis). The counterbalance between degraded electrostatics but moderated quantum-confinement effects leads to nearly equal channel g_{M} between the materials (Table 3.3).

3.2.5 Combined gate oxide stacks

Performing MC simulations with state-of-the-art quantum corrections, we have shown that FinFET devices with light-mass III-V channels may be sensitive to the gate oxide's complex band structure via its impact on channel confinement. For the common device structure considered here, we have found that heavier-mass Al_2O_3, even with a smaller dielectric constant and larger EOT for the same physical oxide thickness, compares well against HfO_2

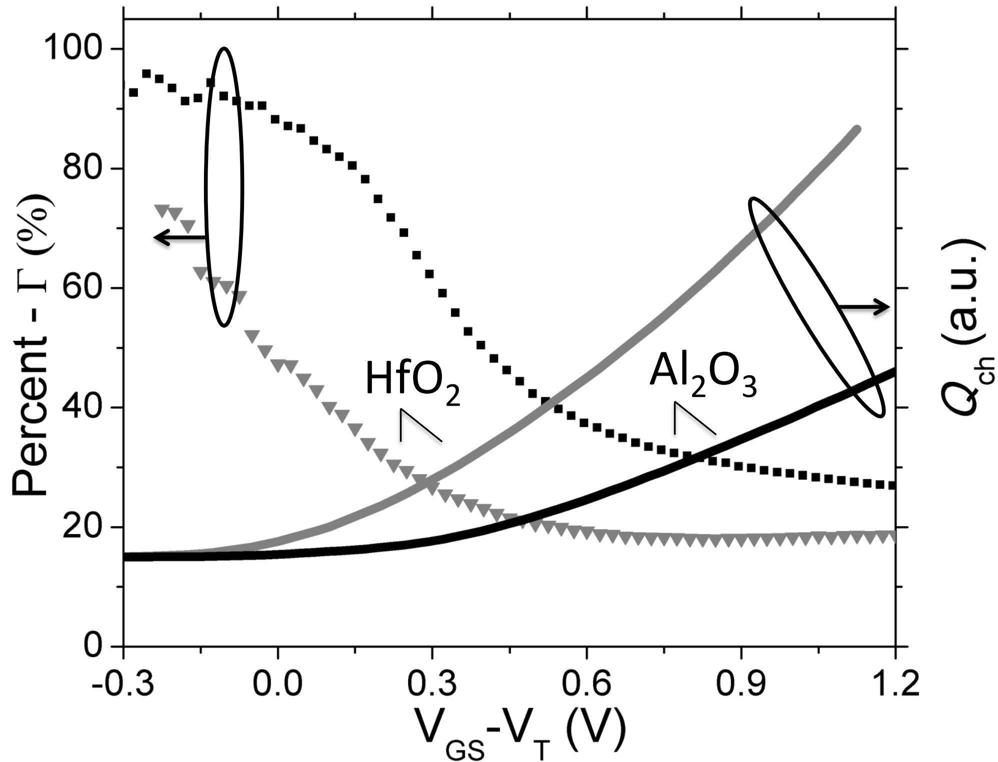


Figure 3.8 – Trade-offs between Al_2O_3 and HfO_2 gate dielectrics. Al_2O_3 (black), with its smaller dielectric constant than HfO_2 (gray), generates a smaller inversion charge density Q_{ch} (right-axis). However, with its heavier effective mass and smaller quantum confinement, Al_2O_3 retains more Γ-valley carriers than HfO_2 (left-axis).

in the ON-state due to reduced carrier confinement conditions. Qualitatively, these results suggest that Al_2O_3 may be preferred to HfO_2 at the same EOT. A larger oxide dielectric constant provides increased gate capactitance. However, a larger oxide effective mass reduces quantum confinement while also reducing the wave function penetration depth in the oxide. Therefore, combining a heavy-mass interface oxide with an over-layer oxide with a large dielectric constant could provide both benefits simultaneously. (Notably, such

a combination already exists in Si-channel devices with high-k gate stacks incorporating heavy-mass native oxide SiO_2 surface layers. However, less and less deleterious—and some even beneficial—quantum confinement occurs in Si devices[8,43]). Finally, we focused on the effect of differing oxide m^*_{ox} on device performance (here with the band offsets/barrier heights expected to be much the same for HfO_2 and Al_2O_3[85,86]). More generally, this work suggests that anything that significantly affects the boundary conditions on the quantum-mechanical wave function at the channel-gate dielectric interface may impact device performance, especially for strong quantum confinement as expected in scaled III-V based n-channel MOSFETs.

Chapter 4

Perpendicular Reading of Single Confined Magnetic Skyrmions

Thin-film sub-5 nm magnetic skyrmions constitute an ultimate scaling alternative for future digital data storage. Skyrmions are robust non-collinear spin-textures that can be moved and manipulated by small electrical currents. A technique is shown here to detect isolated nano-skyrmions with a current-perpendicular-to-plane geometry, which has immediate implications for device concepts. The physics behind such a mechanism are explored by studying the atomistic electronic structure of the magnetic quasiparticles. Investigation from first-principles showed how the isolated skyrmion local-density-of-states, from which electrons tunnel into the vacuum, when compared to that of the ferromagnetic background, is modified by the site-dependent spin-mixing of electronic states with different relative canting angles. Local transport properties are sensitive to this effect, as an atomistic conductance anisotropy of up to $\sim$20% is discovered for magnetic skyrmions in Pd/Fe/Ir(111) thin-films. In single skyrmions, engineering this spin-mixing magnetoresistance could possi-

bly be incorporated in future magnetic storage technologies.

4.1 Single magnetic skyrmions: the ultimate bit

Si complimentary-metal-oxide-semiconductor[10] (CMOS) compatible magnetic devices represent the current state-of-the-art in information data storage circuits.[88] In such devices, the information is encoded by manipulation of different spatial magnetic domains, and the data is read by sensing the variation in the magnetoresistance as a function of the magnetization direction.[89] Typically, the different spatial magnetic domains are homogeneous ferromagnetic domains separated by domain walls (DWs), but more exotic magnetic configurations, such as magnetic skyrmions, may be ultimately preferred in device applications.

Magnetic skyrmions, for topological reasons,[90] are relatively robust particle-like field configurations,[91,92] stable even up to near-room temperature,[93] and could possibly be information carriers or memory storage units in future information technology nodes. Incorporation of magnetic skyrmions instead of DWs may improve device performance and scaling possibilities.[94] DWs are sensitive to defect-pinning[95–97] while skyrmions typically are not.[98–100] This explains reports of skyrmions being moved by electric current densities orders of magnitude smaller than DWs[101,102] while achieving smaller yet sizeable magnitudes of velocity.[100,103–105] Smaller current densities mean that skyrmion-based devices would dissipate less power than DW-based devices, and possibly meet stringent power requirements for future deeply-scaled technologies.[65]

Recent work has shown more subtle defect-pinning dynamics in skyrmions, however. In certain spin-polarized electric current regimes, skyrmions may be trapped by point magnetic defects,[106] and drift mobilities of skyrmions in some realistic films more closely resemble those of DWs.[107]

Regardless of defect-pinning (and therefore power considerations), moving from DWs to skyrmions is attractive from a dimensional scaling point of view, as single skyrmions can be confined at will, and their shape and size controlled with an external magnetic field down to diameters smaller than 5 nm.[108]

With this in mind, an interesting device scaling route is an extension[94] of the racetrack memory[109] to incorporate single[110,111] confined magnetic skyrmions instead of DWs. Such a concept constitutes an ultimate scaling alternative in terms of packing density, speed, and power consumption.[94] Consider a metallic thin-film magnetic heterostructure[111,112] where sub-5 nm chiral skyrmionic quasiparticles are generated (via materials engineering and external magnetic fields) and moved laterally along a magnetic racetrack by in-plane currents, which has been shown experimentally.[101,102,113] The bit-wise data would be encoded by out-of-plane currents which create or annihilate the individual skyrmions, thus setting or resetting the bit-state. This has also been shown experimentally;[111] thus, two important ingredients to a viable skyrmion racetrack memory system, lateral bit-wise movement and set/reset of each bit-state, have been established experimentally.

But how can one read each bit-state? Current-in-plane (CIP) detection

of skyrmions has been shown experimentally[113–115] and understood theoretically as a topological Hall effect,[116,117] but may be costly in terms of power consumption and difficult to fabricate in terms of device geometries. A better option would be the direct detection of the nano-skyrmions via current-perpendicular-to-plane (CPP)[118,119] geometries.

Near a skyrmion, the microscopic magnetoresistance varies as a function of the atomic magnetization direction due to the spin-mixing of electronic states on neighboring sites. Considering a skyrmion's central spin-flipped atom, its local magnetic environment is inequivalent from those atoms in the ferromagnetic background. Thus, in the presence of non-collinearity, one could expect already a conductance anisotropy in the center of the skyrmion due to the change in the electronic structure relative to the ferromagnetic (FM) state. In addition, the heavy-metal substrate induces a large spin-orbit interaction (SOI), coupling the local magnetization to the real-space direction, further modifying the electronic structure as a function of magnetization rotation relative to the substrate plane. This tunneling anisotropic magnetoresistance (TAMR)[120] effect has been studied in detail but typically for homogeneous magnetic domains (either different ferromagnetic domains[120,121] or homogeneous spin-spirals[122]), not inhomogeneous spin-textures such as skyrmions.

In this chapter, we demonstrate a spin-averaged electrical detection mechanism for single skyrmions in a CPP-geometry. This all-electrical signature is a departure from typical experiments which rely on spin-polarized injection to detect magnetic structures, and thus inherently an improvement from

a device application perspective. Our mechanism is viable due to the reliable modification of the local tunneling magnetoresistance into the sample near the vicinity of a skyrmion. Using single confined magnetic skyrmions as examples, we explain this combined (non-collinear and SOI-induced) spin-mixing magnetoresistance in terms of the first fully self-consistent calculations of entire skyrmions altogether from first-principles, rather than employing extrapolated models as commonly done (see for example Refs. 112, 123–125). There we have direct access to the electronic structure of not just each skyrmion as a whole, but even of the states decaying into the vacuum, which are essential to the tunneling conductance. We find a rather large atomistic conductance anisotropy of up to ∼20% (∼10%) for magnetic skyrmions in Pd/Fe/Ir (Pd/Pd/Fe/Ir) magnetic thin-films, which potentially could be detected in a realistic device exploiting a CPP-geometry. Developing the physics of this generalized tunneling spin-mixing magnetoresistance (TXMR) could possibly inspire the design of future nano-magnetic devices based on such a mechanism.

In Section 4.2, we introduce the studied material systems before discussing the details of the electronic structure of single magnetic skyrmions in Section 4.3. In Section 4.4, we discuss how a robust magnetoresistance emerges in the vicinity of localized magnetic skyrmions. Finally, we summarize this *ab initio* analysis in Section 4.5.

4.2 Material systems and computational procedure

Thin-film systems containing single nano-scopic skyrmions make a dynamic combination from an engineering perspective, due to the robustness of the skyrmion data-carriers and the ultimate scalability of future fabricated devices. Such an auspicious system is schematically illustrated in Fig. 4.1a. In the upcoming discussions, we will show how the tunneling current between a suspended metal contact through vacuum depends on the non-collinear magnetic state-of-phase below. Suspended metal contacts are possible with state-of-the-art fabrication techniques,[126] but one could also imagine tunneling through a weakly-interacting two-dimensional (2D) insulator, such as hexagonal boron nitride (hBN) or molybdenum disulfide (MoS_2). Such a process can be intimately understood in a non-spin-polarized scanning tunneling microscopy (STM) experiment, Fig. 4.1b.

In this study, we consider two magnetic thin-film heterostructures similar to Fig. 4.1b purely from *ab initio*: fcc-overlayers of Pd/Fe and Pd/Pd/Fe on single-crystal bulk fcc-Ir(111). These systems are attractive for a number of reasons. First, they generate large Dzyaloshinskii-Moriya interactions (DMIs),[127–129] whose competition with the isotropic exchange interaction J determines the size and chirality of the skyrmions.[94] DMIs are large here because of the strength and nature of the inversion-symmetry-breaking in the heterostructures. At the Fe/Ir(111) interface, a large spin-orbit interaction in the underlying heavy-metal substrate, here Ir(111), is relatively uncompensated by the overlayer Pd/Fe or Pd/Pd/Fe interface, leading to a large

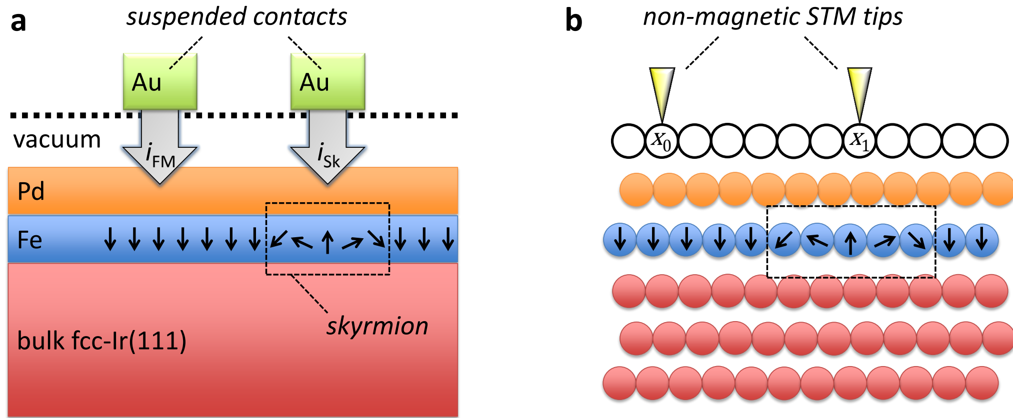


Figure 4.1 – Out-of-plane detection of nano-skyrmions. (a) Illustrative heterostructure cross-section for the perpendicular reading of single nano-skyrmions. Due to energy-dependent spin-mixing perturbations to the atomistic electronic structure as a function of position within skyrmions, the electric current relation $i_{\rm FM} \neq i_{\rm Sk}$ holds, and therefore the magnetic data information can be sensed in a CPP-geometry. **(b)** Illustrative STM-spectroscopy experiment of fcc-Pd/Fe overlayer on single-crystal fcc-Ir(111) bulk substrate. The tunneling conductance is modified by the combined effects of local magnetic non-collinearity and substrate-induced spin-orbit interaction. For similar physical reasons as (**a**), the tunneling conductance at position x_0 is different from that at position x_1.

DMI vector preferentially in the plane of Fe, denoted by $\mathbf{D}$. The ratio of $|\mathbf{D}|/J$, along with an external magnetic field, can stabilize isolated skyrmions with diameter $D_{\rm Sk}$ ≈1-5 nm in size, and has been shown experimentally.[108,111] Second, by choosing a double-Pd overlayer (Pd/Pd/Fe/Ir) versus a single-Pd overlayer (Pd/Fe/Ir), one can alter the exchange interactions in Fe due to the modified nature of the interface hybridization and electronic charge transfer (see Section 4.2.1 and Tables 4.1 and 4.2 below). We investigate this effect to illuminate conceptual studies where other overlayer combinations and materials are used to engineer the size, shape, and stability of the isolated

skyrmions.[112,124,125]

We focus specifically on single skyrmions and do not investigate networks or lattices of skyrmions. We perform self-consistent density functional theory (DFT) calculations based on a full-potential Green function formalism including SOI,[130] which allows a perfect embedding of real-space defects, such as isolated skyrmions, into the ferromagnetic background system. Additional specifics of our computational scheme are given in the Methods Section 4.6.

4.2.1 Charge density and angular momentum of ferromagnetic thin-films

In Table 4.1, we give information regarding the charge density, charge transfer, and spin and orbital magnetic moments of the near-surface atomic layers in the ferromagnetic (FM) state in the single-Pd system (Pd/Fe/Ir).

We see that the spin magnetic moment of Fe is rather large (2.70 μ_B) and the orbital moment is not negligible (0.10 μ_B). Fe induces a sizable spin-moment in Pd of 0.31 μ_B which stiffens the isotropic exchange constant for first-nearest neighbors J compared to the system[124,131] having no Pd-overlayer

	Z	n	Δn	M_S	M_L
Pd	46	45.72	−0.28	0.31	0.02
Fe	26	26.18	0.18	2.70	0.10
Ir	77	76.84	−0.16	0.02	−0.01

Table 4.1 – FM state of Pd/Fe/Ir(111). The atomic number (Z), the total electronic charge (n), and the charge transfer ($\Delta n = n - Z$) are given in units of electrons. The spin-moment (M_S) and the orbital-moment (M_L) are given in units of μ_B.

($J^{\mathrm{Pd/Fe/Ir}} = 14.7$ meV vs. $J^{\mathrm{Fe/Ir}} = 5.7$ meV). Pd is known to be a good Stoner system, i.e. it can easily develop a spin magnetic moment.

We also provide a similar table for the double-Pd case (Pd/Pd/Fe/Ir) in Table 4.2. Interestingly in this case, the spin magnetic moments of the Fe-layer and its nearest-neighbor Pd-layer did not change greatly compared to the values obtained for Pd/Fe/Ir. However, the charge transfer of this inner Pd-layer is now much smaller. Since the surface Pd-layer interacts with Fe indirectly through the inner Pd-layer, the induced spin magnetic moment is rather small (0.08 μ_{B}) on the surface.

4.2.2 Non-collinear inhomogeneity in nano-skyrmions

Before coming to the essential physics of the tunneling spin-mixing magnetoresistance (TXMR) effect, we first self-consistently relax different-sized nano-skyrmions in otherwise ferromagnetic backgrounds (see Fig. 4.2), in both single- and double-Pd overlayer material stacks. We control the size of the skyrmionic defects by allowing different finite numbers of atoms to re-

	Z	n	Δn	$\mathrm{M_S}$	$\mathrm{M_L}$
Pd	46	45.77	−0.23	0.08	0.01
Pd	46	45.94	−0.06	0.28	0.02
Fe	26	26.20	0.20	2.63	0.09
Ir	77	76.83	−0.17	0.02	−0.01

Table 4.2 – FM state of Pd/Pd/Fe/Ir(111). The atomic number (Z), the total electronic charge (n), and the charge transfer ($\Delta n = n - Z$) are given in units of electrons. The spin-moment ($\mathrm{M_S}$) and the orbital-moment ($\mathrm{M_L}$) are given in units of μ_{B}.

lax their magnetic moments in size and direction after the central atom has been spin-flipped as an initial condition. We investigate three different realistic skyrmion sizes: $D_{\mathrm{Sk}} \approx 1.7$, 2.2, and 2.7 nm in diameter. The spin-textures exhibit a fixed and unique rotational sense as demanded by the DMI, which seeks energy gain by torquing the moments to rotate with respect to their neighbors. These structures are cycloidal and radial in nature as expected for magnetic thin-films. Thus our theoretical calculations are consistent in generating realistic nano-skyrmions which have been experimentally detected using magnetic spin-polarized currents.[108, 111]

We illustrate the spin-moment global rotation versus the vertical (polar angle θ) of each atom and the pairwise difference between adjacent polar angles (dθ). We will show that the spin-mixing perturbations to the local-density-of-states (LDOS) are a function of these angular parameters because the relative canting between different pairwise atomic sites varies as a function of space, in addition to the absolute canting relative to the substrate. While traversing across the diameter of any of the nanoscopic skyrmions shown in Fig. 4.2, we mention that dθ itself is not constant between different nearest-neighbor atom pairs, such that there exists an inhomogeneity on the atomic-scale in the rotation of the magnetization direction with respect to the substrate plane. Furthermore, these inhomogeneities themselves are a function of diameter when comparing skyrmions of different sizes (Fig. 4.2a,b,c).

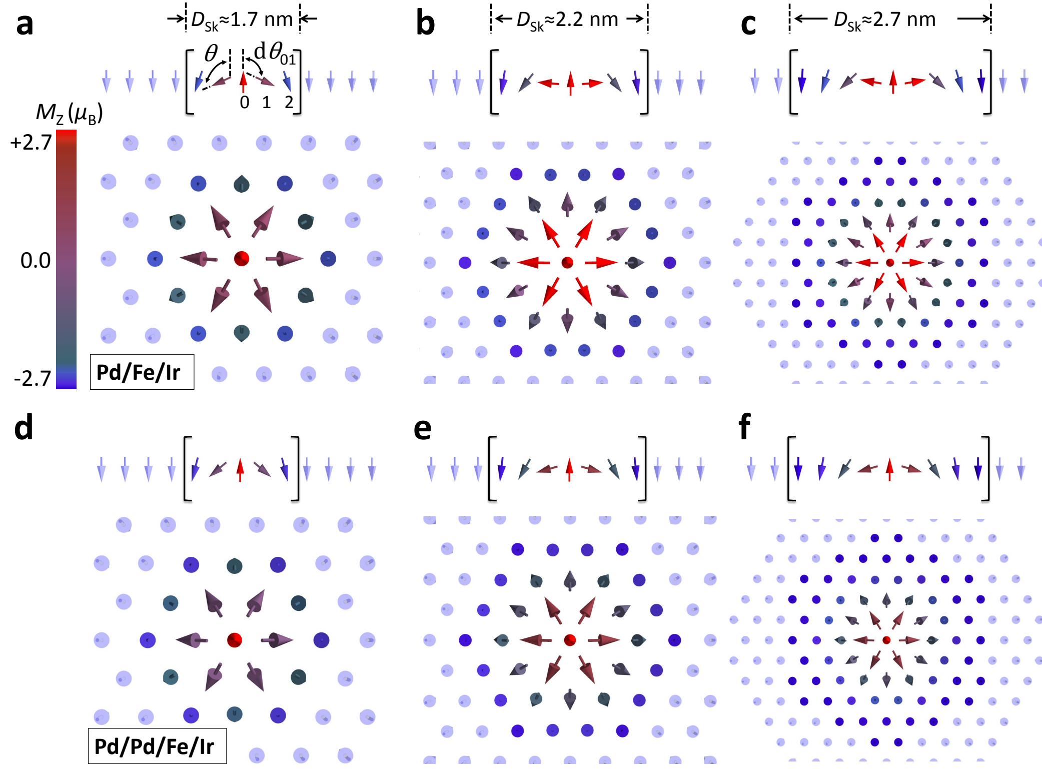


Figure 4.2 – Real-space relaxation of nano-skyrmions with increasing size. **(a-c)** Plots of axisymmetric cycloidal spin-whirls inside a magnetically active Fe-layer centered about increasingly larger skyrmionic defects in fcc-Pd/Fe overlayer on fcc-Ir(111) bulk substrate. Confining spins in the FM-background are shown transparent. We define θ as the typical polar angle with the vertical and dθ as the difference in polar angle between adjacent pairwise atoms. **(d-f)** Again but in fcc-Pd/Pd/Fe overlayer on fcc-Ir(111) bulk substrate. The color bar in (**a**) represents the magnitude of the z-component of the magnetization for each spin in (**a**-**f**).

4.3 Electronic structure of isolated confined skyrmions

We now move to establish the physics behind the TXMR effect within a scanning tunneling microscopy/spectroscopy (STM or STS) experiment employing a non-spin-polarized tip, for which according to the Tersoff-Hamann model,[132] the differential conductance dI/dV is proportional to the LDOS of

the sample, calculated at the tip position, $\mathbf{R}_{\text{tip}}$, and the given bias energy E_{bias}:

$$\frac{\mathrm{d}I}{\mathrm{d}V} \propto \text{LDOS}\,(\,\mathbf{R}_{\text{tip}}; E_{\text{bias}}, \{\mathbf{s}_i\}) \ . \tag{4.1}$$

The LDOS depends on the configuration $\{\mathbf{s}_i\}$ of atomic spins of the sample relative to each other, e.g. in terms of $\mathrm{d}\theta_{ij} = \theta_j - \theta_i$ for all atom-pairs (i, j), and relative to the lattice in terms of the absolute polar angle θ_i. The transport phenomenon related to the latter is known as the tunneling anisotropic magnetoresistance (TAMR),[120] an effect related to spin-mixing due to the SOI. The former results from the spin-mixing hybridization of majority- and minority-states due to non-collinearity. Both can be subsumed as tunneling spin-mixing magnetoresistance (TXMR). Common to both is that the probability of tunneling into majority- and minority-states depends on angles. The difference is that the impact on the electronic structure due to non-collinearity can be of first-order and thus larger than that due to the SOI, which is typically of second-order. While the SOI is nominally of the order of tens of meV, non-collinear interactions are mediated by the exchange splitting of the electronic states ($\sim$few eVs) and the strength of their hybridization, which here is of the order of hundreds of meV (see Sections 4.3.1 and 4.3.2 and Figs. 4.3 and 4.4 below for more details). Also the appearance of both are different. For example, the TXMR due to non-collinearity in a homogeneous magnetic spiral is the same across the spiral, because $\mathrm{d}\theta_{ij} = \mathrm{d}\theta$ for all atom-pairs (i, j), but different for spirals of different pitches under the transformation $\mathrm{d}\theta \to \mathrm{d}\theta'$. In contrast, the TAMR is modulated across the spiral[122] as θ_i changes from atom-

to-atom. Thus, the TXMR is used to measure conductance differences between two different magnetic states such as the difference between a skyrmion and the FM-state, but can also be used to resolve magnetization inhomogenieties inside complex spin-textures such as skyrmions or domain walls.

4.3.1 Angular dependence of the local density-of-states in nano-skyrmions

We now analytically derive the change in the LDOS at site r, denoted $\Delta\text{LDOS}(r)$, inside an axially symmetric skyrmion measured from the origin at $r = 0$ as a function of energy and magnetization rotation direction defined by the unit vector of the magnetic moment $\hat{\mathbf{s}}(r) = \big(\sin\theta(r)\cos\phi(r), \sin\theta(r)\sin\phi(r), \cos\theta(r)\big)$. Examples of rotation parameters θ and $\mathrm{d}\theta$ for given skyrmion magnetic moments can be found in Fig. 4.4d. When compared to the ferromagnetic state (with all moments pointing out-of-plane), there are two contributions to $\Delta\text{LDOS}(r)$, one due to spin-orbit interaction (SOI) and one due to non-collinearity (NC): $\Delta\text{LDOS}(r) = \Delta\text{LDOS}^{SOI}(r) + \Delta\text{LDOS}^{NC}(r)$. The contribution from SOI, the so-called anisotropic magnetoresistance, is well known:[120–122]

$$\Delta\text{LDOS}^{SOI}(r, E, \theta) \propto A(r, E) \cdot \left[1 - \hat{s}_z^2(r)\right] \quad , \tag{4.2}$$

where $A(r, E)$ is a coefficient depending on the site r and energy E, and $\hat{s}_z(r)$ is the z-component of the spin-moment at site r. Thus upon including SOI, we expect, for example, a $\sin^2\theta(r)$ dependence, which contributes to $\Delta\text{LDOS}(r)$ in second-order.

The contribution from NC, intuitively, comes from the change in the electronic structure upon rotation of the magnetic moments at consecutive sites $(i,\ j)$. For homogeneous magnetic spirals, a constant deviation in the LDOS from the ferromagnetic (FM) state will be observed for each atom in the spiral (ignoring SOI). In such a spiral, the smooth rotation of moments $\theta(i) \rightarrow \theta(j) = \theta(i) + \mathrm{d}\theta$ for each atom pair is a symmetry operation commuting with a Hamiltonian having translational invariance, making each atom equivalent, and the electronic structure the same for each atom in the spiral. However, upon transforming the spiral's pitch $\mathrm{d}\theta \rightarrow \mathrm{d}\theta'$, one would find a different constant deviation of the electronic structure from the FM-state, such that spirals of different pitch can be identified by their different magnitudes in $\Delta\mathrm{LDOS}^{NC}$.

In skyrmions, however, the rotation of the magnetic moments is not homogeneous, i.e. $\mathrm{d}\theta$ is not constant for all nearest-neighbor atom pairs while traversing the skyrmion's diameter. Thus there will be a site-dependent deviation in the LDOS among the atoms inside a skyrmion with respect to each other. In what follows, we demonstrate that these deviations are a complex function of the rotation angles, which will depend on the details of the electronic structure, the energy probed, and even the size of the skyrmions.

To derive $\Delta\mathrm{LDOS}^{NC}(r)$, we utilize multiple scattering theory: G is the Green function describing the whole system upon rotation of the magnetic moments, and g is the Green function describing the initial FM-state. The Green function will be used to evaluate the change in the LDOS induced by

the rotation of the magnetic moments:

$$\Delta\text{LDOS}^{NC}(r,E) = -\frac{1}{\pi}\,\Im\Big\{\text{Tr}_{\text{LS}}\big[\Delta G^{rr}(E)\big]\Big\}\ , \tag{4.3}$$

as given in a matrix notation where a trace over orbital and spin angular momenta has to be performed. G can be evaluated via the Dyson equation connecting the non-collinear state to the ferromagnetic one:

$$G = g + g\Delta V G = g + g\Delta V g + g\Delta V g\Delta V g + ... \tag{4.4}$$

where ΔV describes the change of the potential upon rotation of the magnetic moments. It can be expressed as:

$$\Delta V(r) = V_{\text{diff}}(r)\,(\boldsymbol{\sigma}\cdot\hat{\mathbf{s}}(r) - \sigma_z)\ \ , \tag{4.5}$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices, and V_{diff} is the difference of the two spin components of the ferromagnetic potential $(V_0^{\uparrow} - V_0^{\downarrow})/2$.

We execute a similar expansion for the ferromagnetic initial Green function matrix g:

$$g = g_{\text{sum}}1_2 + g_{\text{diff}}\sigma_z\ \ \ , \tag{4.6}$$

where 1_2 is the 2×2 identity matrix.

Let us evaluate the first-order and second-order terms contributing to the Dyson equation (Eq. 4.4):

$$\Delta G^{rr} = \sum_i g^{ri}\Delta V^i g^{ir} + \sum_{ij} g^{ri}\Delta V^i g^{ij}\Delta V^j g^{jr} + ... \tag{4.7}$$

where i and j are sites surrounding site r, or can be the site r itself. Since the trace over spin has to be performed, we will focus only on the terms that in

the end will contribute to Eq. 4.3. We use a pair of useful properties of the Pauli matrices:

$$\mathrm{Tr_S}\big[\sigma_x\big] = \mathrm{Tr_S}\big[\sigma_y\big] = \mathrm{Tr_S}\big[\sigma_z\big] = 0 \quad , \tag{4.8}$$

and

$$(\boldsymbol{\sigma}\cdot\hat{\mathrm{s}})(\boldsymbol{\sigma}\cdot\hat{\mathrm{s}}') = \hat{\mathrm{s}}\cdot\hat{\mathrm{s}}' + \mathrm{i}\,\boldsymbol{\sigma}\cdot(\hat{\mathrm{s}}\times\hat{\mathrm{s}}') \quad , \tag{4.9}$$

where i is the imaginary unit. After simplifying, we find the following result:

$$-\frac{1}{\pi}\Im\Big\{\mathrm{Tr_{LS}}\big[g^{ri}\Delta V^i g^{ir}\big]\Big\} \quad = \quad B^{rir}(E)\,[1-\hat{s}_z(i)] \tag{4.10}$$

where the coefficient B is given by

$$B^{rir} = \frac{2}{\pi}\Im\Big\{\mathrm{Tr_L}\big[g^{ri}_{\mathrm{sum}}V^i_{\mathrm{diff}}g^{ir}_{\mathrm{diff}} + g^{ri}_{\mathrm{diff}}V^i_{\mathrm{diff}}g^{ir}_{\mathrm{sum}}\big]\Big\} \quad . \tag{4.11}$$

In other words, the first sum in Eq. 4.7 leads to a behavior like $(1-\cos\theta_i)$.

The second-order term is given by

$$-\frac{1}{\pi}\Im\Big\{\mathrm{Tr_{LS}}\big[g^{ri}\Delta V^i g^{ij}\Delta V^j g^{jr}\big]\Big\} = C^{rijr}\big[\hat{\mathbf{s}}(i)\cdot\hat{\mathbf{s}}(j) - \big(\hat{s}_z(i)+\hat{s}_z(j)\big) + 1\big], \tag{4.12}$$

where the coefficient C is related to the Green functions and V_{diff} by:

$$\begin{aligned} C^{rijr} = -\frac{2}{\pi}\Im\Big\{\mathrm{Tr_L}\big[&g^{ri}_{\mathrm{sum}}V^i_{\mathrm{diff}}g^{ij}_{\mathrm{sum}}V^j_{\mathrm{diff}}g^{jr}_{\mathrm{sum}} \;+\; g^{ri}_{\mathrm{diff}}V^i_{\mathrm{diff}}g^{ij}_{\mathrm{diff}}V^j_{\mathrm{diff}}g^{jr}_{\mathrm{sum}} \\ &+ g^{ri}_{\mathrm{sum}}V^i_{\mathrm{diff}}g^{ij}_{\mathrm{diff}}V^j_{\mathrm{diff}}g^{jr}_{\mathrm{diff}} \;+\; g^{ri}_{\mathrm{diff}}V^i_{\mathrm{diff}}g^{ij}_{\mathrm{sum}}V^j_{\mathrm{diff}}g^{jr}_{\mathrm{diff}}\big]\Big\} \quad . \end{aligned} \tag{4.13}$$

Thus we obtain a dependence on the dot product of the unit vectors of the magnetic moments $(1-\cos\mathrm{d}\theta\cos\mathrm{d}\phi)$ and a contribution depending only on the z-components of the unit vectors of the magnetic moments.

We have thus demonstrated that due to NC, the dependence of the change in the LDOS with respect to the ferromagnetic state upon rotation of the magnetic moments is not trivial, and will have terms depending on the dot product between magnetic moments, contrary to the contribution coming from SOI. The non-collinear contribution is then

$$\Delta\mathrm{LDOS}^{NC}(r, E, \{\mathbf{s}\}) =$$
$$\sum_i B^{rir}(E)\big(1 - \hat{s}_z(i)\big) + \sum_{ij} C^{rijr}(E)\big[\hat{\mathbf{s}}(i)\cdot\hat{\mathbf{s}}(j) - \big(\hat{s}_z(i) + \hat{s}_z(j)\big) + 1\big], \quad (4.14)$$

where $\{\mathbf{s}\}$ is the spin configuration. Of course, depending on the details of the electronic structure and strength of perturbation related to the non-collinearity, higher-order terms can be important and have to be included in Eq. 4.14.

Combining $\Delta\mathrm{LDOS}^{NC}$ and $\Delta\mathrm{LDOS}^{SOI}$, in the next section we will fit the change in the LDOS in terms of trigonometrical functions that depend on the rotation angles of the magnetic moments. We will apply these fits to our *ab initio* results as well as to an extended Alexander-Anderson model used to interpret the variation of the LDOS resonance-splitting upon rotation of the magnetic moments on neighboring sites.

4.3.2 Two-atom extended Alexander-Anderson model

We wish to estimate the change in the LDOS and qualitatively understand the shifting in energy of resonant d-states in Fe as a function of the rotation angle between adjacent moments. To this end, we consider for simplification two magnetic atoms $(i, j) = (1, 2)$ each having one localized orbital

d_{z^2} whose single-particle eigenenergy is centered about $E = \epsilon$. The initial Hamiltonian describing this model is diagonal in spin-space. We could also consider an orbital of the type d_{xz} in order to address the coupling induced by SOI between the d_{z^2} and d_{xz}, as done by Caffrey *et al.*[133] However, since the impact of SOI on the LDOS has already been discussed by others, we focus here on the impact of NC on the LDOS. We study the ΔLDOS as we vary $\mathrm{d}\theta = \theta_1 - \theta_2$ between the two atoms. We restrict the hopping from atom-to-atom to non-spin-flip processes, characterized by the interaction parameter V_{hop}.

In terms of Green functions, the following equation gives the LDOS for site 1:

$$\mathrm{LDOS}(1; E, \{\mathbf{s}\}) = -\frac{1}{\pi}\Im\Big\{\mathrm{Tr_S}\big[G_{11}(E)\big]\Big\} = -\frac{1}{\pi}\Im\Big\{\mathrm{Tr_S}\left[E - H + i\Gamma\right]^{-1}\Big\}_{11}, \tag{4.15}$$

where Γ takes care of the broadening of the states. Instead of solving exactly the previous equation, one could also use perturbation theory, as described in the previous Section 4.3.1, simplifying Eq. 4.14 to:

$$\Delta\mathrm{LDOS}^{NC}(1; E, \{\mathbf{s}\}) = D(E) \cdot (1 - \cos \mathrm{d}\theta)\,, \tag{4.16}$$

where $D = B^{121} + C^{1221}$.

The energy of the resonant d-states, their width, and splittings come from our first-principles calculations, e.g. Fig. 4.6a (atom Fe-3, navy curve). To obtain the proper splitting we choose a hopping parameter $V_{\mathrm{hop}} \approx 300 -$

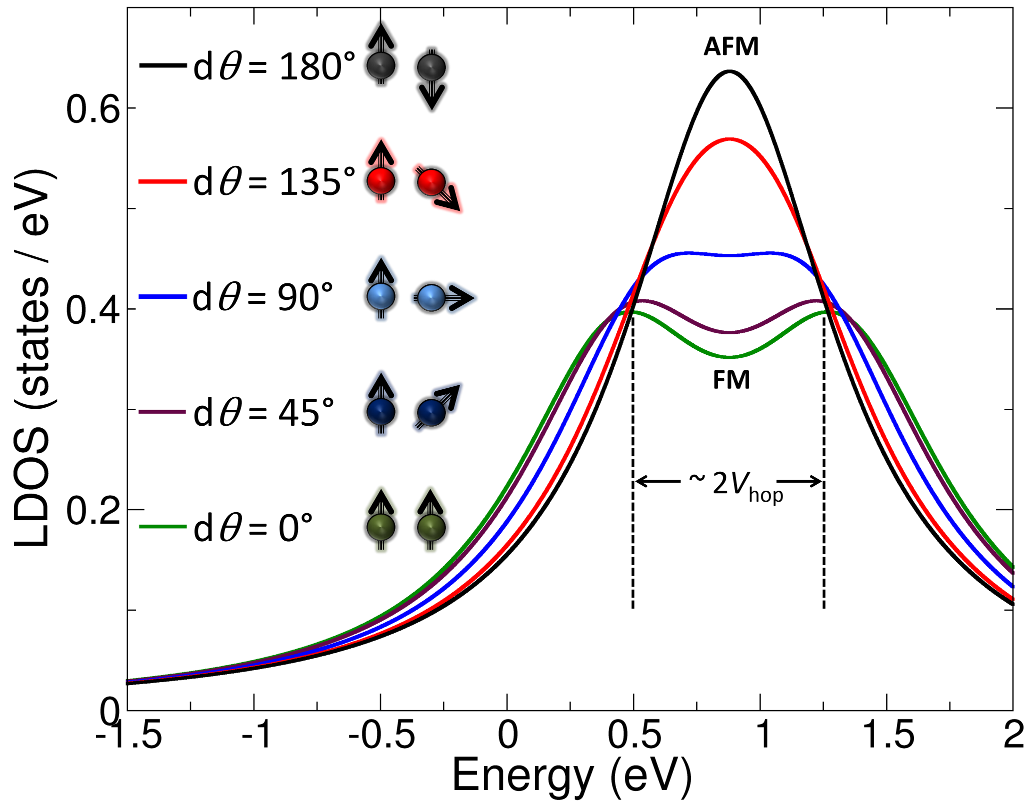


Figure 4.3 – Alexander-Anderson prediction for d-d hybridization. Beginning from dθ=0°, ferromagnetic coupling imposes a hybridization between the two d_{z^2} orbitals localized at the same energy but on neighboring atoms. This produces a splitting into a bonding- and antibonding-state (green curve). Upon rotation, the atoms eventually become antiferromagnetically coupled (black curve). The change in the LDOS can be qualitatively estimated with a $\cos(\mathrm{d}\theta)$ fitting parameter.

400 meV. We show the resulting LDOS in Fig. 4.3 for five different rotation angles dθ. There we reproduce the splitting of the resonance-peaks that we have seen in our first-principles calculations, where d-d hybridization is important, as seen in the upcoming figures, e.g. Figs. 4.5a, 4.6a, 4.7a, and 4.8a.

For the limiting cases of a ferromagnetic state and an antiferromagnetic

state, we recover the expectations of the Alexander-Anderson model.[134,135] Indeed, when the two atoms are in the ferromagnetic state ($\mathrm{d}\theta = 0°$), the *d*-*d* hybridization leads to a splitting of the original single orbital into bonding- and antibonding-states, seen as broad resonances near 0.5 eV and 1.25 eV, respectively (Fig. 4.3, green curve). The splitting is then given by $2V_{\mathrm{hop}}$.

If the magnetic state is antiferromagnetic ($\mathrm{d}\theta = 180°$), there is repulsion between the minority and majority spin-states leading to a shift given by $V_{\mathrm{hop}}/(E^{\downarrow}_{i,dz^2} - E^{\uparrow}_{i,dz^2})$. In our simple model, $E^{\downarrow}_{i,dz^2} - E^{\uparrow}_{i,dz^2}$ is extremely large, thus the shift is not observed in Fig. 4.3. It is interesting to see how the splitting between the resonance-peaks decreases upon rotation of the magnetic moments until they merge to a single resonance in the antiferromagnetic case (Fig. 4.3, black curve). This is in accordance with the behavior of the LDOS calculated from first-principles for different sized skyrmions in the two systems Pd/Fe/Ir (Fig. 4.5a and Fig. 4.6a) and Pd/Pd/Fe/Ir (Figs. 4.7a and 4.8a).

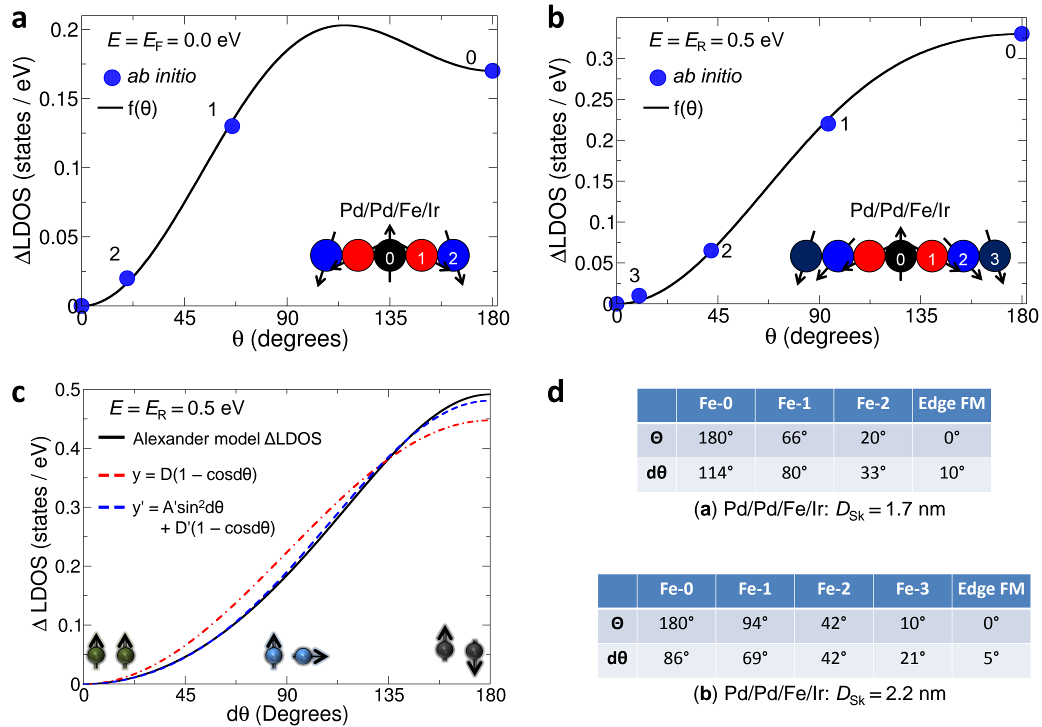
Next, we wish to estimate the change in the LDOS as a function of rotation as previously discussed, which leads to the TXMR signal. In Fig. 4.4, we plot the change in LDOS compared against the background-FM for the model along with our *ab initio* results in Pd/Pd/Fe/Ir for skyrmions 1.7 nm and 2.2 nm in diameter, respectively. In addition we show fitted functions against θ and $\mathrm{d}\theta$. Interestingly, a good fit to the change in LDOS shown in Fig. 4.4a,b is found by considering

$$\Delta\mathrm{LDOS}(r; E, \{\mathbf{s}\}) \approx A(E) \cdot \sin^2\theta(r) + D(E) \cdot \big[1 - \cos\theta(r)\big] \tag{4.17}$$

instead of strictly employing the terms given by Eq. 4.14. This result is similar to what we found in the simple two-orbital Alexander-Anderson model (Eq. 4.16). In such systems, the contribution from all spin-moment dot products (see Eq. 4.12) behaves on average like $\cos(\theta)$. This is naturally satisfied in the two-orbital model.

In Fig. 4.4a the change in LDOS at the Fermi energy $E = E_\mathrm{F} = 0$ is depicted. The blue dots represent the total ΔLDOS given from our *ab initio* calculations for each atom extending radially from the skyrmion core. The fitting-function $f(\theta)$ is Eq. 4.17 and shown to be effective in fitting the first-principles data. We note that the change in the LDOS at the Fermi energy was not large. This is different than the situations in Fig. 4.4b,c where we probed the energy resonances near 0.5 eV, and a larger change in the LDOS was induced. In Fig. 4.4b, the fitting function $f(\theta)$ is again Eq. 4.17 and shown to be effective in fitting the *ab initio* data.

In Fig. 4.4c, we plot the ΔLDOS as given by the two-atom Alexander-Anderson model (black curve) with two fitting-functions. The first-order fit (red-dashed curve) based on Eq. 4.16 is shown to be slightly inaccurate when fitting the model ΔLDOS. Instead a higher-order term is needed to fit the data (blue-dashed curve). Thus we learned from the Alexander-Anderson model that depending on the probed energy, additional terms can be needed to improve the fit. A term proportional to $\sin^2(\theta)$, similar to the one expected when SOI is included, improves considerably the fit to the non-collinear contribution to ΔLDOS. For example, near energy resonances, especially if they



	Fe-0	Fe-1	Fe-2	Edge FM
Θ	180°	66°	20°	0°
dθ	114°	80°	33°	10°

(**a**) Pd/Pd/Fe/Ir: $D_{Sk} = 1.7$ nm

	Fe-0	Fe-1	Fe-2	Fe-3	Edge FM
Θ	180°	94°	42°	10°	0°
dθ	86°	69°	42°	21°	5°

(**b**) Pd/Pd/Fe/Ir: $D_{Sk} = 2.2$ nm

Figure 4.4 – Energy- and spatially-dependent ΔLDOS in skyrmions. (a) Fe-layer atom-by-atom *ab initio* results (blue dots) of the total change in LDOS as a function of θ in a 1.7 nm diameter skyrmion in Pd/Pd/Fe/Ir. ΔLDOS is given relative to the background-FM and is measured at the Fermi energy while including SOI. These results are then fitted to $f(\theta)$ (solid curve). The $f(\theta)$ in (**a**,**b**) refers to Eq. 4.17, while the numbers labeling each dot respresent the corresponding atom in the inset extending radially from the skyrmion core. **(b)** Again, Fe-layer first-principles results (blue dots) of the atomistic change in LDOS in Pd/Pd/Fe/Ir and fitted against $f(\theta)$ (solid curve), but for a larger 2.2 nm skyrmion and measured at the resonant energy $E_{\mathrm{R}} = 0.5\ \mathrm{eV}$. **(c)** We plot the ΔLDOS as a function of $\mathrm{d}\theta$ given by our two-atom Alexander-Anderson model (black curve) to isolate the non-collinear component of the ΔLDOS. While the main contribution comes from $\cos(\mathrm{d}\theta)$ (red-dashed, referring to Eq. 4.16), a higher-order term is required to complete the fit (blue-dashed). **(d)** Rotation parameters as found from the first-principles calculations, referring to (**a**) and (**b**).

are sharp, the perturbative expansion up to second-order from Section 4.3.1 begins to break down and higher-order terms are necessary to improve the

fitting-functions.

4.3.3 Electronic structure in 1.7 nm Pd/Fe/Ir(111) skyrmions

Since the spin-mixing perturbations due to non-collinearity and SOI are magnetic in nature, we show in Fig. 4.5a the spin-dependent LDOS in the magnetically-active Fe-layer as a function of the atomic position for the $D_{\mathrm{Sk}} \approx 1.7$ nm skyrmion. For brevity we plot only the Pd/Fe/Ir case (see the upcoming Section 4.3.4 and Fig. 4.6 for the larger $D_{\mathrm{Sk}} \approx 2.2$ nm case and Section 4.3.5 and Figs. 4.7 and 4.8 for the double-Pd cases). We note that the majority and minority spin-channels are given in the local spin-frame of reference for each Fe-atom. The color-coding of the plot, which corresponds to different atoms extending radially from the skyrmion's center, is explained in Fig. 4.5c. The energy-zero is the Fermi energy, $E_{\mathrm{F}} = 0$.

The resonant-states between 0.5 and 1.0 eV above E_{F} are of Fe d-band minority-spin character and consist of d_{xy/x^2-y^2}, $d_{xz/yz}$, and d_{z^2} states. These states hybridize with the sp states in the Pd-overlayer and give rise to Fe-Pd-spd_{z^2} hybrid states, named in short Pd-d_{z^2} states localized in the Pd overlayer-film around $E \approx 0.5$ eV, as shown in Fig. 4.5b (black-dashed curve). It is clear that the surface-layer Pd-d_{z^2} state (shown only for the background-FM Pd-surface film), which has the proper orbital symmetry to decay slowly transverse to the substrate, controls the electronic structure in vacuum as a function of energy, characterized by a strong resonance in the vacuum-LDOS. An all-electrical STS measurement will be sensitive to this vacuum resonance-

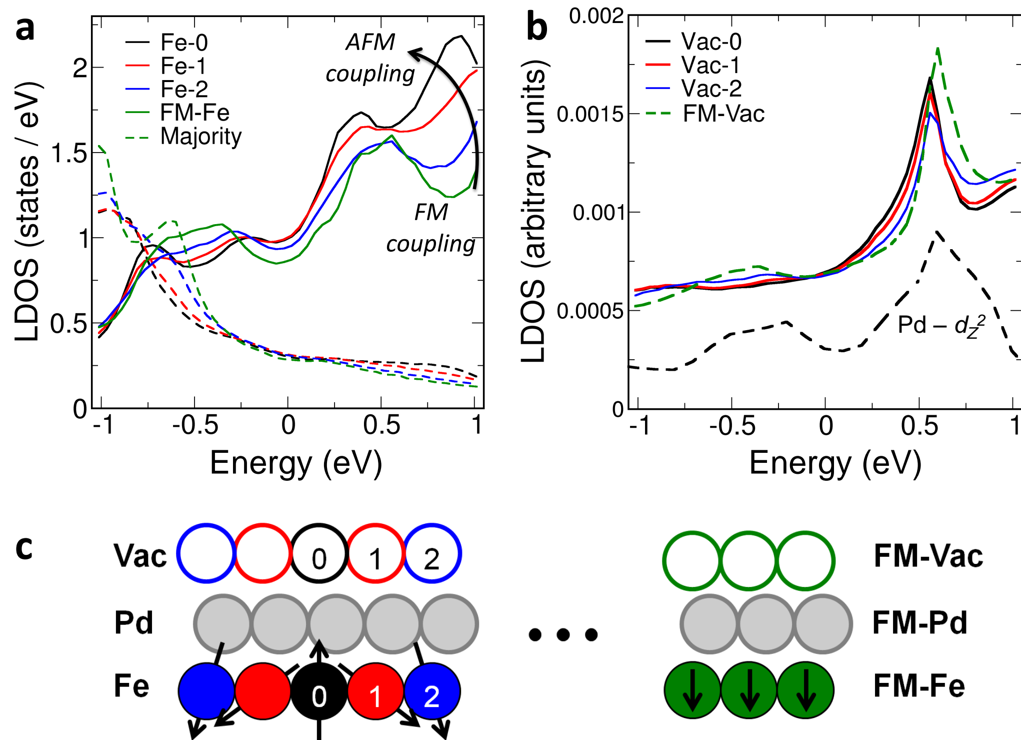


Figure 4.5 – Electronic structure of a realistic $D_{\mathrm{Sk}} \approx 1.7$ nm skyrmion in Pd/Fe/Ir. (a) LDOS in the magnetically-active Fe-layer resolved into minority (solid) and majority (dashed) spin-channels. The resonance-peak near $E \approx 0.5$ eV in the FM-background (green) shifts in energy when approaching the center of the skyrmion (black). **(b)** The modification of the electronic structure in Fe contributes to a strong resonance in the LDOS in vacuum via hybridization through surface Pd-states. Arbitrary units are used so as to include in the same plot the nature of the Pd-d_{z^2} surface-state (black-dashed), whose resonance-peak near 0.5 eV survives in the vacuum. **(c)** Illustrative legend for (**a**, **b**) where the numbered spheres represent a line extending radially from the skyrmion's center. The vacuum domains are represented by empty spheres. FM-Fe, FM-Pd, and FM-Vac represent the unperturbed background ferromagnet.

peak.

The origin of this resonance and its behavior upon rotation can be un-

derstood by analyzing the energy window near $E \approx 0.5$ eV in Fig. 4.5a, where one can see the resonance-peaks shifting in energy in the Fe skyrmion-LDOS as a function of position. The green curve, which represents the ferromagnetic state of the background Fe-film, shows an electronic structure consistent with Fe-minority *d*-*d* hybridization when adjacent atoms couple ferromagnetically (see Section 4.3.4 below). Conversely, moving towards the center of the skyrmion, the quantization axes between two neighboring atoms becomes different, and majority-states of one atom can hybridize with minority-states of the second. This effect is especially pronounced at the central spin-flipped atom (black curve), where the resonance-peak has shifted lower in energy – as expected for antiferromagnetic (AFM) coupling (see Section 4.3.4 below). We reproduce these effects within the context of a simple model, where we can qualitatively predict the change in LDOS as a function of the non-collinear magnetization rotation parameter $\mathrm{d}\theta$ as defined in Fig. 4.2a (see previous Sections 4.3.1 and 4.3.2).

The energy-dependent disturbance to the LDOS resonance-peaks as a function of position moving radially along the skyrmion will manifest as a perturbation to the local electrical conductivity, and is the physical basis for the space-dependent TXMR effect.

4.3.4 Electronic structure in 2.2 nm Pd/Fe/Ir(111) skyrmions

Let us now turn to the case of a single-Pd system but for a larger skyrmion, $D_{\mathrm{Sk}} \approx 2.2$ nm, compared to the one shown in Section 4.3 ($D_{\mathrm{Sk}} \approx$

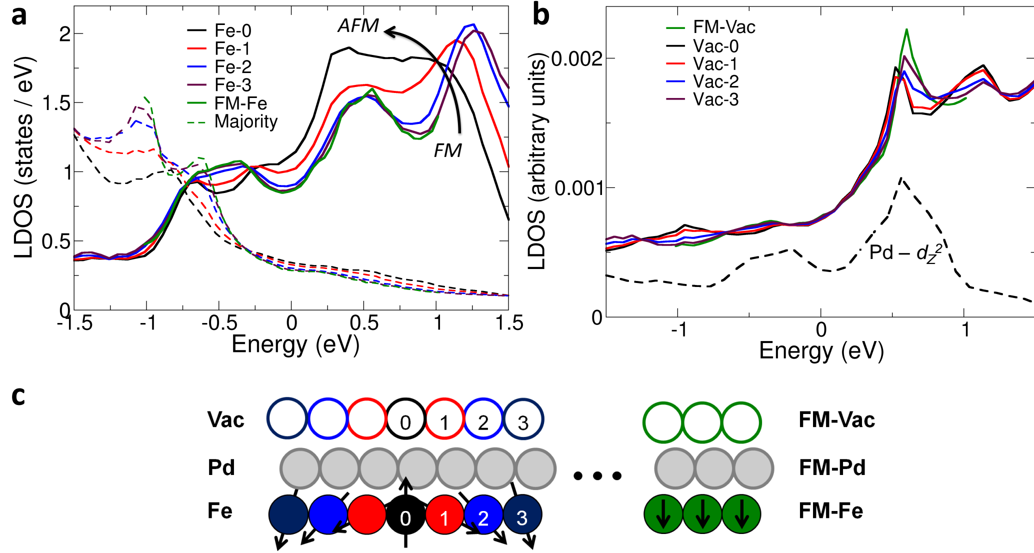


Figure 4.6 – LDOS of a $D_{\mathrm{Sk}} \approx 2.2$ nm skyrmion in Pd/Fe/Ir. (a) Electronic structure in the magnetically-active Fe-layer resolved into minority (solid) and majority (dashed) spin-channels. The splitted-structure of the FM-background (green) is modified due to quasi-AFM interactions when approaching the center of the skyrmion (black). **(b)** The alteration of the LDOS in Fe contributes to a strong resonance in vacuum via hybridization through surface Pd-states. Arbitrary units are used so as to include in the same plot the nature of the Pd-d_{z^2} surface-state (black-dashed), whose resonance-peak near 0.5 eV survives in the vacuum. **(c)** Illustrative legend for (**a**, **b**) where the numbered spheres represent a line extending radially from the skyrmion's center. The vacuum domains are represented by empty spheres. FM-Fe, FM-Pd, and FM-Vac represent the unperturbed background ferromagnet.

1.7 nm, Fig. 4.5). In Fig. 4.6, we plot the LDOS in a realistic nano-skyrmion in Pd/Fe/Ir(111) for the magnetically-active Fe-layer (**a**) and its corresponding vacuum sites (**b**), but with a slightly larger energy window $[-1.5, +1.5]$ eV as compared to the Section 4.3, where a window of $[-1.0, +1.0]$ eV was analyzed.

Similar to the case of the smaller skyrmion in the single-Pd system as shown in Section 4.3 (Fig. 4.5a, FM-Fe, green curve), a large resonance peak

is observed in the background-FM Fe-film at 0.5 eV (Fig. 4.6a, FM-Fe, green curve). But by increasing the energy window beyond 1.0 eV, we clearly see a second resonance appearing at 1.25 eV in the nearly-FM atom Fe-3 at the edge of the skyrmion (Fig. 4.6a, Fe-3, navy curve). The appearance of these two resonance peaks in tandem suggests a strong interaction due to ferromagnetic coupling. In other words, the resonant-peaks at 0.5 eV and 1.25 eV are actually the result of FM-interactions which have split a single resonance in two. An analogy can be made between these FM-splitted resonance-peaks and localized degenerate atomic orbitals at different sites which upon interaction split in energy between bonding- and antibonding-states. The essential physics of this process were reproduced within an extended Alexander-Anderson model for *d*-*d* hybridization between FM-Fe *d*-states[134, 135] in the previous Section 4.3.2.

The splitting in energy between the two resonance-peaks reduces when the rotation angle of the magnetic moment increases (Fig. 4.6a, Fe-2, blue curve, and Fe-1, red curve). Finally, in the central spin-flipped atom of the skyrmion (Fig. 4.6a, Fe-0, black curve), the splitting almost disappears as the two resonance-peaks quasi-merge. This is evidence of antiferromagnetic coupling between the central atom and its nearest-neighbors.[134, 135] This coupling was also reproduced within our Alexander-Anderson model in the antiferromagnetic case in the previous Section 4.3.2.

An important difference with respect to the center of the skyrmion (Fig. 4.6a, Fe-0, black curve) is observed when comparing to the smaller $D_{\mathrm{Sk}} \approx$ 1.7 nm case (Section 4.3 Fig. 4.5a, Fe-0, black curve). In the smaller skyrmion,

the quasi-merged resonance still shows two separate peaks, while in the larger skyrmion there is only a single, broad resonance peak. Thus we expect a difference in the TXMR signal in the larger skyrmion compared to the $D_{\mathrm{Sk}} \approx$ 1.7 nm case. In general, there will be a weak size dependence of the spin-mixing signal, which should slowly decay as the diameter of the skyrmion is increased. In addition, from Fig. 4.6b we can see that the vacuum resonance near 0.5 eV in the smaller defect (Section 4.3 Fig. 4.5b, solid curves) survives in the larger skyrmion (Fig. 4.6b, solid curves), and the TXMR remains detectible, as seen in Fig. 4.10a.

To complete our analysis of the LDOS shown in Fig. 4.6, we mention that with regards to magnifying the strength of the TXMR effect, we need not restrict ourselves only to the energy window near 0.5 eV, but also look for other energy windows where a large change in the LDOS is observed as a function of position. We see this in the majority-states near -1.0 eV (Fig. 4.6a, dashed-curves), where the spin-mixing effect is perhaps even stronger than in the 0.5 eV window composed of minority-states. Indeed, the TXMR signal is shown to be large for negative bias energies near -1.0 eV in this system (see Section 4.3 Fig. 4.9a,b and Fig. 4.10a).

4.3.5 Electronic structure in Pd/Pd/Fe/Ir(111) skyrmions

In Fig. 4.7, we plot the LDOS in a skyrmion in Pd/Pd/Fe/Ir(111) with diameter $D_{\mathrm{Sk}} \approx 1.7$ nm for the magnetically-active Fe-layer (**a**) and its corresponding vacuum sites (**b**).

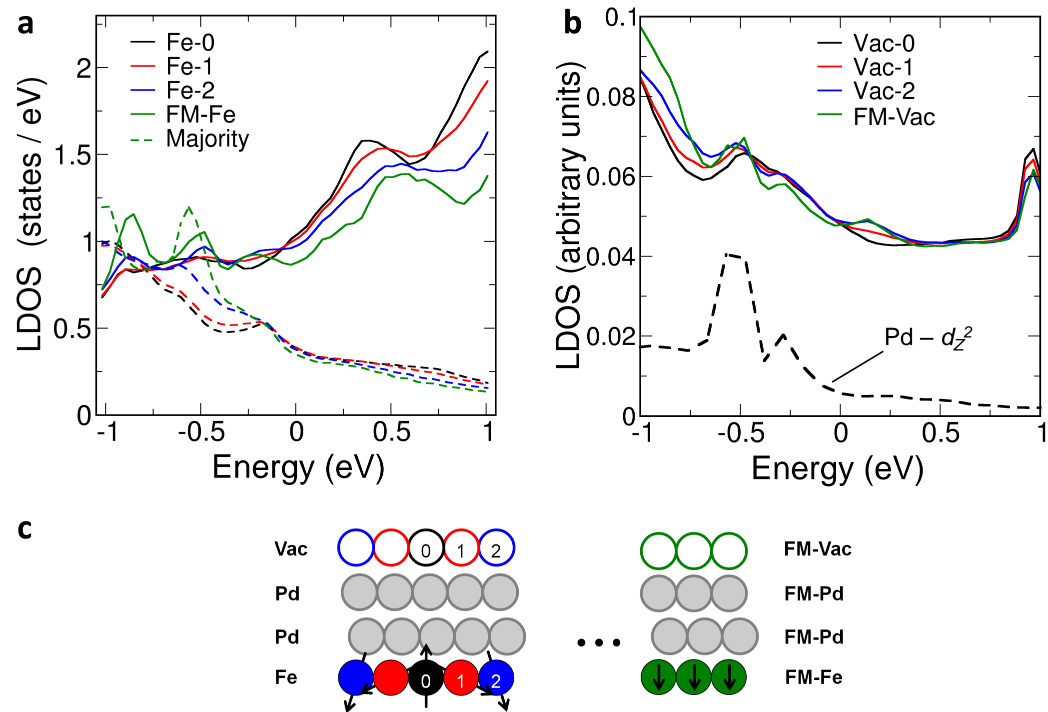


Figure 4.7 – LDOS of a $D_{\text{Sk}} \approx 1.7$ nm skyrmion in Pd/Pd/Fe/Ir. (a) LDOS in the magnetically-active Fe-layer resolved into minority (solid) and majority (dashed) spin-channels. The resonance-peak near $E \approx 0.5$ eV in the FM-background (green) shifts in energy when approaching the center of the skyrmion (black). **(b)** The modification of the electronic structure in Fe impacts the vacuum LDOS via hybridization through surface Pd-states. Arbitrary units are used so as to include in the same plot the nature of the Pd-d_{z^2} surface-state (black-dashed), whose features at negative energies survive in the vacuum. **(c)** Illustrative legend for (**a**, **b**) where the numbered spheres represent a line extending radially from the skyrmion's center. The vacuum domains are represented by empty spheres. FM-Fe, FM-Pd, and FM-Vac represent the unperturbed background ferromagnet.

Similar to the case of the smaller skyrmion in the single-Pd system as shown in Section 4.3 (Fig. 4.5a, FM-Fe, green curve), a strong resonance-peak appears near 0.5 eV composed of minority-spin channel d-states (Fig. 4.7a, FM-Fe, green curve). This resonance-peak shifts in energy upon increased

rotation of the magnetic moments moving towards the center of the skyrmion (Fig. 4.7a, Fe-2, blue curve; Fe-1, red curve; and finally Fe-0, black curve). This observed change in the electronic structure as a function of position due to the spin-mixing of majority- and minority-states suggests there will be a TXMR signal in this system as shown previously for the single-Pd system.

Contrary to the single-Pd case, however, the resonance in the hybrid Fe-Pd-spd_{z^2} state (Pd-d_{z^2} for short) observed in the Pd-overlayer around 0.5 eV vanishes, and instead appears near -0.5 eV (Fig. 4.7b, black-dashed curve). The steepness of this resonance leads to a disturbance in the vacuum-LDOS as a function of position in slightly more negative energies, near -0.8 eV (Fig. 4.7b, solid curves). The character of the hybrid Pd-d_{z^2} surface-state leads to a flat region in the vacuum-LDOS at positive energies. We then expect here a smaller TXMR signal in comparison to the single Pd-system for positive bias voltages. Therefore, in the double-Pd system, we suggest probing negative bias energies near -0.8 eV, where the spin-mixing effect is more significant, as exemplified in Section 4.3 Fig. 4.9d,e.

In Fig. 4.8, we plot the LDOS for a larger skyrmion, $D_{\mathrm{Sk}} \approx 2.2$ nm, in Pd/Pd/Fe/Ir(111). We use a similar labeling convention when decomposing the LDOS. Interestingly, the d-resonances are broader (Fig. 4.8a, solid curves) than for the smaller skyrmion (Fig. 4.7a, solid curves). In the center of the skyrmion, only one broad d-resonance is observed (Fig. 4.8a, Fe-0, black curve). To reiterate, this indicates that the TXMR signal will be different upon increasing the diameter of any skyrmion since its electronic structure is modified

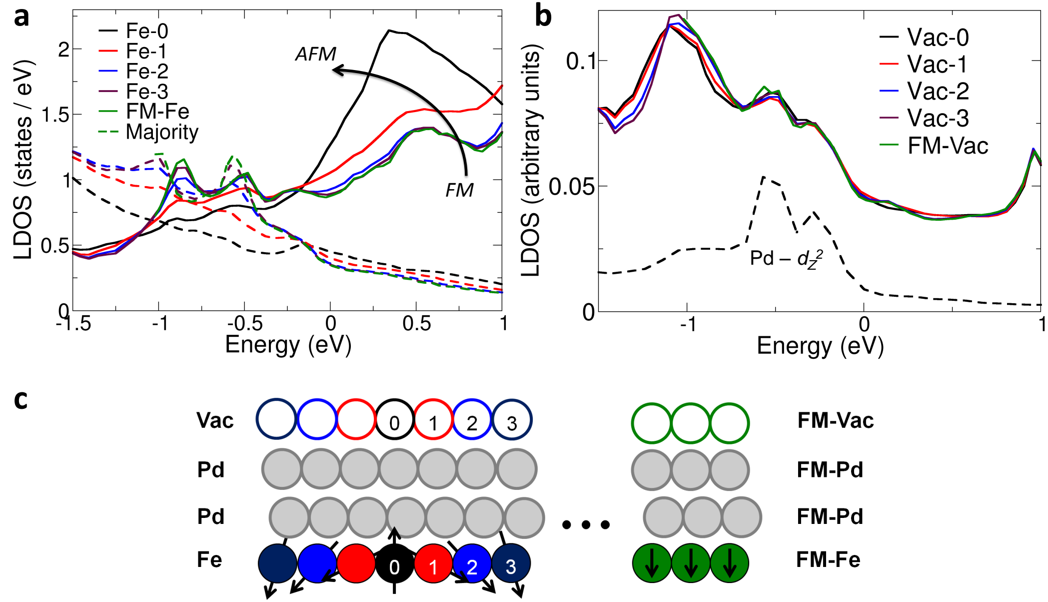


Figure 4.8 – LDOS of a $D_{\mathrm{Sk}} \approx 2.2$ nm skyrmion in Pd/Pd/Fe/Ir. (a) Electronic structure in the magnetically-active Fe-layer resolved into minority (solid) and majority (dashed) spin-channels. The splitted-structure of the FM-background (green) is modified due to quasi-AFM interactions when approaching the center of the skyrmion (black). **(b)** The alteration of the LDOS in Fe impacts the vacuum structure via hybridization through surface Pd-states. Arbitrary units are used so as to include in the same plot the nature of the Pd-d_{z^2} surface-state (black-dashed). **(c)** Illustrative legend for (**a**, **b**) where the numbered spheres represent a line extending radially from the skyrmion's center. The vacuum domains are represented by empty spheres. FM-Fe, FM-Pd, and FM-Vac represent the unperturbed background ferromagnet.

upon increasing D_{Sk}. From the shape of the vacuum-LDOS (Fig. 4.8b, solid curves), the TXMR is nicely detectable at an energy range around -0.8 eV, as confirmed in Fig. 4.10b, and in the same energy window as the smaller $D_{\mathrm{Sk}} \approx 1.7$ nm Pd/Pd/Fe/Ir(111) case.

From Sections 4.3.4 and 4.3.5 we can conclude that for each skyrmion in a Pd/Fe/Ir(111) or Pd/Pd/Fe/Ir(111) system, we can select an energy region

in which a large tunneling spin-mixing magnetoresistance (TXMR) will be found. Due to the relevant energy of detection, one can distinguish skyrmions of different sizes, and discriminate between skyrmions in Pd-films of different thicknesses. Finally, and perhaps most importantly, the spatial-variation of the spin-mixing signal means that the internal atomic structure of each individual skyrmion can be resolved. This means the TXMR can be used to visualize the size, shape, and structure of individaul defects. In devices based on spin-mixing, this also means that skyrmions can be used as nano-scopic information carriers, where the TXMR would be used to read the magnetic state-of-bit.

4.4 Tunneling spin-mixing magnetoresistance (TXMR)

We now define the TXMR and make predictions for future experimental observation of the effect. The TXMR is the percent deviation of the local conductance from a reference conductance due to the spin-mixing from non-collinearity and SOI. As long as the magnetic state under consideration has a different non-collinearity than the reference state, there will be a TXMR. If one is interested in the spatial resolution of a complex spin texture (ignoring SOI), however, then an additional inhomogeneity within the non-collinearity is required, as is the case for nano-skyrmions.

The TXMR is by definition measured in vacuum. Here, we choose the reference to be somewhere far from the skyrmion in the FM-background. Then

the normalized TXMR measured at site r is

$$\mathrm{TXMR}(r) = \frac{\mathrm{LDOS}^{\mathrm{vac}}_{\mathrm{FM}} - \mathrm{LDOS}^{\mathrm{vac}}_{\{\mathbf{S}\}}(r)}{\mathrm{LDOS}^{\mathrm{vac}}_{\mathrm{FM}}} \times 100\% \ , \tag{4.18}$$

where $\mathrm{LDOS}^{\mathrm{vac}}_{\mathrm{FM}}$ is the LDOS in the vacuum just above the FM, and $\mathrm{LDOS}^{\mathrm{vac}}_{\{\mathbf{S}\}}(r)$ is the LDOS of the complex spin-texture in the vacuum just above site r.

Integrating the TXMR over the entire device injection boundary, over all energies up to the bias energy eV_{bias}, would give a measure of the total change in conductance, and would be the state-of-bit detection mechanism in a CPP-TXMR device like discussed in Fig. 4.1a. In an STS experiment, however, the effect could be amplified by selecting specific energy windows where the TXMR were largest as a function of position.

In Fig. 4.9a,d we show the energy-resolved TXMR of the $D_{\mathrm{Sk}} \approx 1.7$ nm skyrmion's central spin-flipped vacuum-site, with and without SOI, for the single- and double-Pd cases, respectively. We notice a sizeable TXMR effect for both systems. This holds true for all skyrmions that we studied, noting a small size-dependence of the effect which varies weakly as a function of D_{Sk} (see the upcoming Section 4.4.1 and Fig. 4.10).

TXMR signals in the different single-Pd and double-Pd material systems vary from skyrmion-to-skyrmion, however. Since the SOI arises from the Fe-Ir interface, the impact of SOI is much more pronounced in the single-Pd system (Fig. 4.9a), where the TXMR can peak out at an impressive $\sim$40% when ignoring spin-orbit coupling, but decreases down to $\sim$ 20% when SOI

is included. Interestingly, the spin-mixing due to SOI is to compete with the effects due to inhomogeneous non-collinearity, reducing the overall TXMR signal. In the double-Pd system, most of the TXMR signal is due to inhomogeneous non-collinearity, with a small contribution coming from the SOI (Fig. 4.9d).

In Fig. 4.9b,e we plot the spacial variation of the TXMR signal, which differs significantly from atom-to-atom within the same skyrmion. We see that the TXMR effect is reduced when approaching the edge of the skyrmion (Fig. 4.9b,e blue curves), since the effective non-collinearity is reduced as the complex spin-texture fades into the ferromagnetic background.

4.4.1 TXMR in larger 2.2 nm skyrmions

In Fig. 4.10 we plot the TXMR signal spatially-resolved for the 2.2 nm skyrmions in both systems. The spin-mixing signals retain the same general features and shapes as in the smaller defects (Section 4.3 Fig. 4.9a-d). The important peak near 0.5 eV in the Pd/Fe/Ir case remains robust (Fig. 4.10a), along with another strong peak near -1.0 eV. In the case of Pd/Pd/Fe/Ir (Fig. 4.10b), the TXMR has the strongest signal near -0.8 eV, as before. Therefore one could infer from Fig. 4.10 that as the diameter of skyrmionic quasiparticles is increased, the spin-mixing effect not only survives, but also that specific locations of strong TXMR signals remain in similar energy windows as in smaller structures.

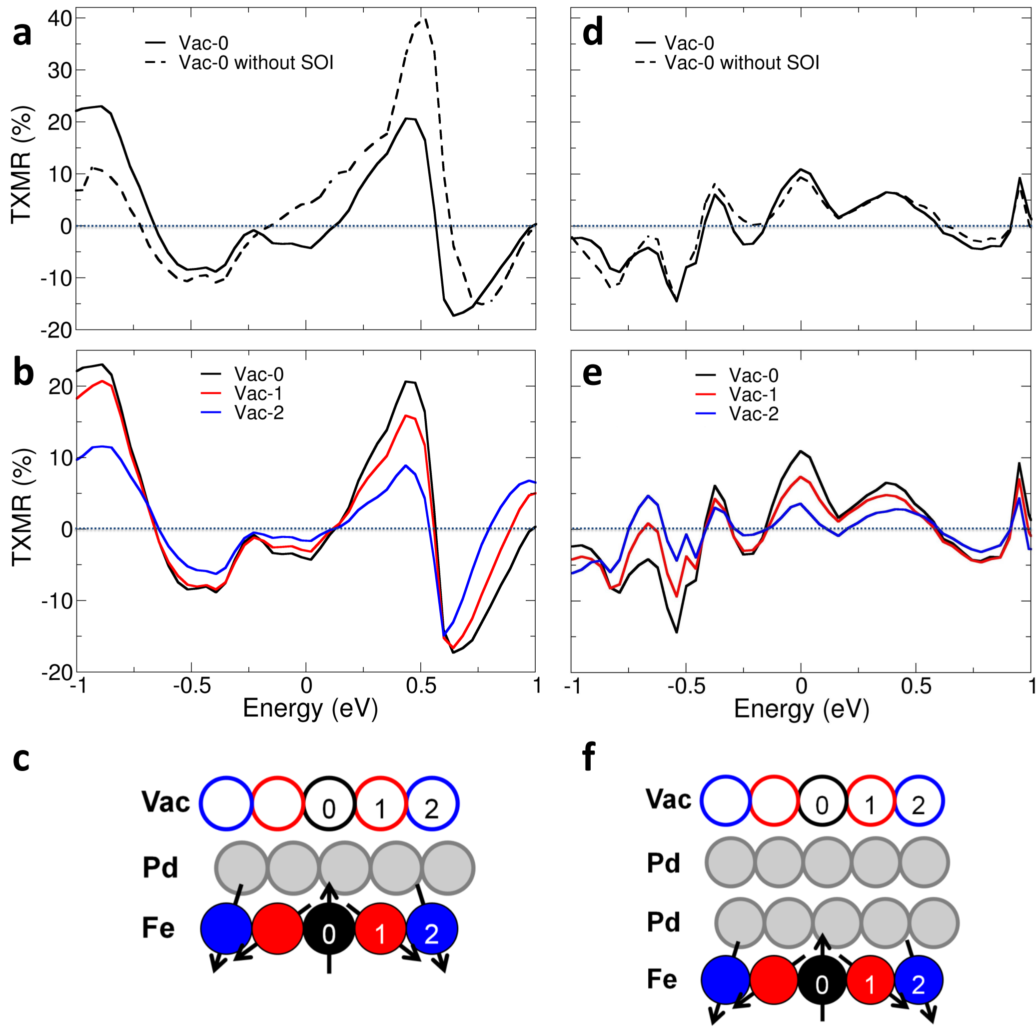


Figure 4.9 – Tunneling spin-mixing magnetoresistance. (a) Energy-resolved TXMR signals measured at the skyrmion's core comparing the effects of SOI in a $D_{\mathrm{Sk}} \approx 1.7$ nm skyrmion in Pd/Fe/Ir. **(b)** TXMR signals again, but for a line of atoms extending radially from the center of the skyrmion, illustrating the spatial dependence of the effect. **(c)** Illustrative legend for (**a**, **b**) where the spheres are numbered and color-coded to identify the appropriate curves to the corresponding vacuum domains above the skyrmion. **(d-f)** Plots corresponding to (**a-c**) but in the Pd/Pd/Fe/Ir system.

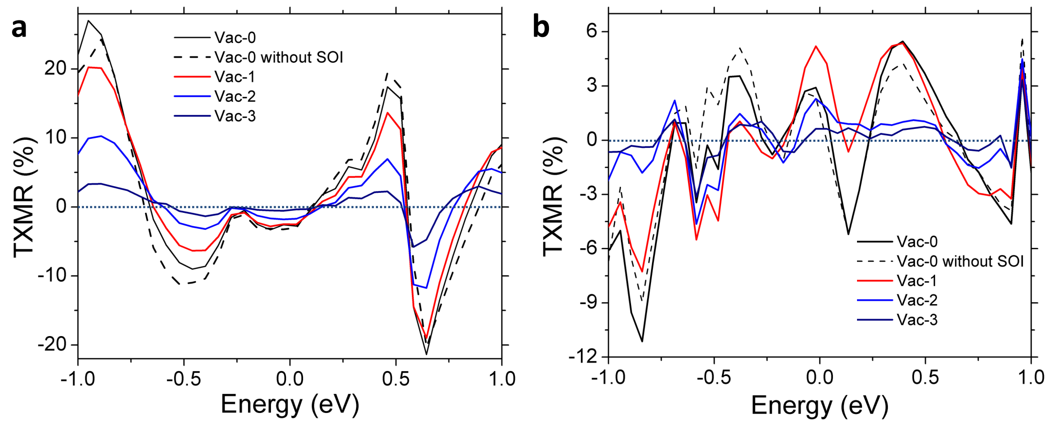


Figure 4.10 – Spacially-varying TXMR in 2.2 nm skyrmions. (a) We plot the TXMR in vacuum for Pd/Fe/Ir beginning in the core of the skyrmion (Vac-0, black) and then moving radially outwards (Vac-1, red; Vac-2, blue; Vac-3, navy) towards the confining FM environment. The vacuum resonance near 0.5 eV (see Fig. 4.6b) contributes a strong peak in the TXMR near the same energy. **(b)** Again but in Pd/Pd/Fe/Ir, where the peak-signal strength appears near -0.8 eV, as predicted by Fig. 4.8b.

4.4.2 All-electrical skyrmion detection

The vacuum-resonance we found in Fig. 4.5b appears now as a large TXMR signal at the same energy in the Pd/Fe/Ir(111) system (Fig. 4.9a,b). Thus, an experimentalist probing the surface in a STS experiment, having set the ripple-bias voltage near $V_{\mathrm{bias}} \approx 0.5\mathrm{V}$, would see an electrical contrast as visualized in Fig. 4.11 when approaching a skyrmion of similar size in the single-Pd heterostructure.

Within a reasonable bias voltage range, the TXMR effect is smaller in the double-Pd case ($\sim$10%) when compared to the single-Pd case ($\sim$20%). The additional Pd-overlayer changes the resonance nature around 0.5 eV compared to the single-Pd case, and states with a high tunneling cross-section into the

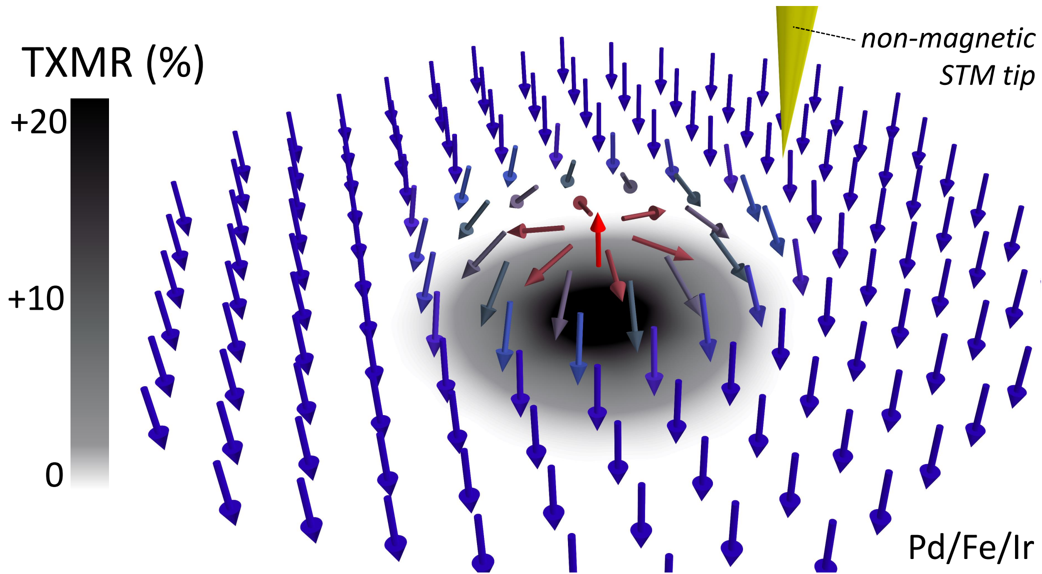


Figure 4.11 – Perpendicular reading of single confined magnetic skyrmions. Expected STS-signal when approaching a skyrmionic defect in the single-Pd case. The electrical contrast has been projected onto the plane below the skyrmion. Near the injection energy $eV_{\text{bias}} \approx 0.5$ eV, there are about $\sim 20\%$ fewer tunneling states in the skyrmion's core compared to the unperturbed FM-environment. This increases the local magnetoresistance, allowing for the reliable spin-averaged electrical detection of skyrmions in a CPP-geometry.

vacuum are distributed over a wider energy (see previous Section 4.3.5). As a consequence, the TXMR is reduced by nearly half when compared to the single-Pd case, and a bias energy near -0.8 eV is experimentally more favorable.

4.5 Discussion

We have studied realistic- and experimentally-observable confined nano-skyrmions within metallic thin-films of Pd/Fe/Ir(111) and Pd/Pd/Fe/Ir(111) completely from first-principles. We established how the

combined effects of local inhomogeneous magnetic non-collinearity and SOI in nano-skyrmions can alter the atomistic electronic structure in a magnetically active Fe-film, and, via hybridization with additional surface layers, the electrons which tunnel into the vacuum.

The change in the LDOS can be understood in terms of the rotation parameters of the magnetic moment of the considered atom. The largest spin-mixing contribution comes from non-collinearity and depends on the relative canting between magnetic moments on neighboring sites, $\mathrm{d}\theta$. The dependence on the absolute polar angle of the magnetic moment with respect to the substrate, θ, comes in as a second-order term to the change in the LDOS, but can become important if the impact of the SOI is large.

Finally, we have shown in detail how such a physical interplay could induce a sizeable electrical conduction anisotropy as a function of position and energy in realistic nano-skyrmions, up to $\sim$20% in the single-Pd case. The manifestation of this TXMR effect could possibly be exploited in an all-electrical tunneling spectroscopy experiment.

In addition, the changes in the magnetoresistance on the nanometer scale of skyrmions can possibly be engineered to design advanced magnetic memory devices. Typical memory circuits require at least one control device (either transistor or diode) in each memory cell. Instead, technologies based on spin-mixing in single skyrmions could have potentially hundreds of bits stored in nanometer-sized magnetic racetracks needing only a single read-out element to detect the contents of each array (see the upcoming Section 4.5.1

and Fig. 4.12).

Such a mature magnetic device technology would have to be relatively impervious to interface quality, film surface roughness, and various point-defect impurities, however. Now while the TXMR effect reported here in this work lead to an impressive magnitude of conduction anisotropy in epitaxial magnetic thin-films, future challenges will be related ultimately to the robustness of the TXMR in realistic devices.

4.5.1 Skyrmion racetracks for dense magnetic memories

Spin-transfer torque magnetic random access memory (STT-MRAM) circuits reliably read-out bit-states depending on a tunneling magnetoresistance anisotropy of ∼30-50% in some structures,[136] with a hope to achieve a magnetoresistance ratio $R_{\mathrm{ON}}/R_{\mathrm{OFF}} \approx 200\%$ by 2022.[65] A TXMR effect as large as ∼20% as we have shown in this work should be enough to provide adequate read-margin for scaled technologies, and is larger than the $< 2\%$ change in resistance found in widespread commercially-used hard disk read heads based on anisotropic magnetoresistance alone.[137] Smaller changes in magnetoresistance just means there should be a more sensitive read-out circuit. Typically this means a few extra control- and boost-transistors and does not substantially increase the footprint of the memory, i.e. incorporating more sensitive read hardware does not degrade packing density considerably.

In potential skyrmion-based devices using CPP-TXMR, a $R_{\mathrm{ON}}/R_{\mathrm{OFF}} \approx 120\%$ could feasably be well-worth the tradeoff when considering the possible

performance gains with regards to: (1) potentially very low power dissipation due to small currents needed to manipulate the magnetic textures; (2) fast speed operation due to reduced read/write latencies associated with nanoscopic size; and (3) large increases in packing density. Let's consider points (2) and (3) in greater detail.

One issue with racetrack memories is that they are not random access. In a random access memory (RAM), any read/write operation can access any bit with roughly the same access time since the word and bit access lines (WL and BL), which are connected to the set/reset elements and read elements, are also connected in parallel with the individual memory cells (see Fig. 4.12a). In a racetrack memory, the situation is different. In practice, if a read/write were requested for an address whose representative bit were at the end of the racetrack, there would be additional latency associated with moving subsequent magnetic domains out the way until the requested bit were under the read or write device (see Fig. 4.12b). But does this make racetrack memory intrinsically slower than RAM?

This is not clear. It depends on the velocities of the skyrmions in the racetrack. Consider 200 nm length racetracks populated with 2 nm in diameter quasiparticles in an array of 100 domains (or 100 bits). With velocity[94] $v_{\mathrm{Sk}} = 100$ m/s , the furthest bit would acquire an additional delay in time t_{delay}

$$t_{\mathrm{delay}} = \frac{L_{\mathrm{RT}}}{v_{\mathrm{Sk}}} \approx \frac{\left(2\,\frac{\mathrm{nm}}{\mathrm{Sk}}\right)\left(1\,\frac{\mathrm{Sk}}{\mathrm{bit}}\right)(100 \text{ bits})}{100\left(\frac{\mathrm{m}}{\mathrm{s}}\right)} = 2 \text{ ns} \quad , \tag{4.19}$$

a small number compared to the total access time needed to complete the

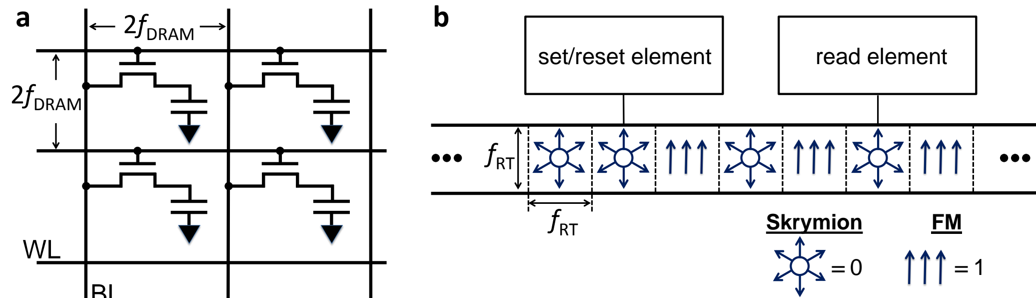


Figure 4.12 – Concept spin-mixing magneto-memory versus DRAM. (a) 2×2 1-T 1-C DRAM unit. In a random access memory, any memory cell can be accessed in roughly the same time since the WLs and BLs are connected to each bit in parallel. The ideal minimum packing footprint goes as $4f^2_{\mathrm{DRAM}}/\mathrm{bit}$ in units of area, where f_{DRAM} is the node generation. **(b)** Spin-mixing racetracks are not random access but acquire a latency associated with moving the bits sequentially out of the way to reach the requested address. However, with realistic skyrmion velocities, the access time could be roughly the same as in DRAM, while dramatically increasing the packing density.

read or write operation, which in modern dynamic RAM (DRAM) is in the range 20–50 ns.[138] We do note, however, that larger in-plane currents will be required to accelerate the quasiparticles up to a velocity such as $v_{\mathrm{Sk}} = 100$ m/s, meaning there will be a tradeoff between t_{delay} and power consumption.

Regardless, by incorporating skyrmion racetracks based on spin-mixing, there seems to be at first glance negligible additional acquired access latency – in fact, we may learn to find in the end that racetracks can be potentially faster than traditional RAM in certain geometries and biasing conditions, due to the nano-scopic size of the skyrmion quasiparticles.

With regards to circuit layouts, it is clear that moving to racetracks will provide large gains in packing density. As an example, let's compare the workhouse 1-transistor 1-capacitor (1-T 1-C) DRAM unit to our racetrack-

based spin-mixing magneto-memory. Considering the lithographical node generation, or minimum feature size, f, DRAM memory minimum packing requirements for a single bit's footprint in the ideal case goes as surface area $\mathrm{SA}_{\mathrm{DRAM}} \approx 4f^2$ per bit (Fig. 4.12a). The current technology node for DRAM in 2015 is given as $f_{\mathrm{DRAM}} \approx 22$ nm by the International Technology Roadmap for Semiconductors (ITRS).[65] In a skyrmion racetrack, the potential effective per-bit surface area could be reduced to possibly $\mathrm{SA}_{\mathrm{RT}} \approx f_{\mathrm{RT}}^2$ per bit, where f_{RT} is the diameter of the magnetic skyrmions in the racetrack (Fig. 4.12b). In our study thus far we considered realistic skyrmions of order $D_{\mathrm{Sk}} \approx 2$ nm. Comparing against current DRAM arrays, we define the potential gain in packing density Γ as

$$\Gamma \approx \frac{SA_{\mathrm{DRAM}}}{SA_{\mathrm{RT}}} \approx \frac{4f_{\mathrm{DRAM}}^2}{f_{\mathrm{RT}}^2} = \frac{(4)(22\ \mathrm{nm})^2}{(2\ \mathrm{nm})^2} = 500 \ . \tag{4.20}$$

Even including more intricate sense amplifiers for read/write operations in skyrmion racetracks, it is clear that incorporation of sub-5 nm skyrmions as data-carriers would substantially increase packing density compared to current state-of-the-art technologies.

Possibly the gains in packing density could even be larger introducing vertical racetracks,[109] which are difficult to fabricate thus far.

4.6 Methods

4.6.1 Computational details

The electronic structure was determined employing density functional theory (DFT) in the local spin density approximation.[139] Calculations were executed by means of the screened Korringa-Kohn-Rostoker full-potential relativistic Green function method.[130] A full-potential method is important to accurately describe the nature of the complex spin-texture and rapidly-decaying vacuum states of the tunneling electrons.

For the calculations we chose an angular momentum cutoff of $l_{max} = 3$ for the orbital expansions of the Green functions. The energy contour for numeric integration of the spin and charge density contained 40 grid points in the upper complex plane (including 7 Matsubara poles) with a Brillouin zone mesh of 30×30 k-points. The FM-slab LDOS and skyrmion impurity cluster LDOS were obtained by one-shot calculations using the FM-state or skyrmion-state as starting points, respectively. We found that increasing the k-mesh to 200×200 was sufficiently adequate to numerically stabilize the relevant observables.

4.6.2 Thin-film slab configurations

The magnetic thin-film slab configurations follow, where positive percentages refer to inward relaxation with respect to the Ir(111) interlayer separation. We consider only fcc-stacking in all cases, which is in fact energetically favorable compared to hcp crystal growth.[124]

Pd/Fe/Ir: 44 total layers (3 vacuum + 1 vacuum (-1%) + 1 Pd (8%) + 1 Fe (7%) + 1 Ir (1%) + 33 Ir + 4 vacuum).

Pd/Pd/Fe/Ir: 44 total layers (3 vacuum + 1 Pd (-1%) + 1 Pd (8%) + 1 Fe (7%) + 1 Ir (1%) + 33 Ir + 4 vacuum).

We choose 34 Ir layers since it was the minimum thickness by which we completely decoupled any wave function penetration from top-to-bottom surface. We obtained the relaxation parameters as optimized and reported by Dupé *et al.*[124]

4.6.3 Calculating whole skyrmions within density-functional theory

To stabilize skyrmions after determining the 2D-FM slabs, the slab Green functions were harvested and a single spin-flipped Fe-atom was embedded in the FM-background. We then allowed three-layer cylindrical ring-like stacks of atoms within the skyrmion impurity cluster to update their potentials and magnetic moments (Fe-layer + 1 Pd-layer above and 1 Ir-layer below). The effect of the FM-background was included self-consistently by the slab Green function (G_0), which connects the skyrmion impurity cluster (G_{imp}) to the host via the Dyson-like equation: $G_{\text{imp}} = G_0 + G_0\,\Delta V\,G_{\text{imp}}$, where ΔV represents the modified atomic potential as compared to the unperturbed slab Green function potential, V. In such a manner a real-space defect can be perfectly embedded in an otherwise periodic crystal. After converging the different sized skyrmionic profiles, observables were calculated as mentioned previously.

Chapter 5

Conclusion

5.1 Dissertation recap

After a short introductory Chapter 1 which motivated the need for high-performance simulation of end-of-the-roadmap and potential beyond CMOS materials and devices, a particle-based ensemble Monte Carlo simulator employing advanced quantum corrections was presented in Chapter 2. We illustrated the significance of quantum confinement and carrier degeneracy effects through simulation of n-channel Si and III-V FinFETs. Original contributions included the treatment of far-from-equilibrium degenerate statistics and QC-based modeling of surface-roughness scattering, as well as considering quantum-confined phonon and impurity scattering in 3D. Typical MC simulations approximate degenerate carrier populations as Fermi distributions to model the Pauli-blocking (PB) of scattering to occupied final states. To allow for increasingly far-from-equilibrium non-Fermi carrier distributions in ultra-scaled and III-V devices, we instead generate the final-state occupa-

tion probabilities used for PB by sampling the local carrier populations as a function of energy valley, energy, and propagation direction. This process is aided by the use of fractional carriers or sub-carriers, which minimizes classical carrier-carrier scattering. Quantum confinement effects are addressed through quantum-correction potentials (QCPs) generated from coupled Schrödinger-Poisson solvers, as commonly done. However, we use our valley- and orientation-dependent QCPs not just to redistribute carriers in real space, or even among energy valleys, but also to calculate confinement-dependent phonon, impurity, and surface-roughness scattering rates. Collectively, these quantum effects can substantially reduce and even eliminate otherwise expected benefits of considered $In_{0.53}Ga_{0.47}As$ FinFETs over otherwise identical Si FinFETs, despite higher thermal velocities in $In_{0.53}Ga_{0.47}As$.

In Chapter 3, an analysis of possible future logic device scaling trends was detailed where high performance n-channel Si and InGaAs FinFETs were simulated down to channel lengths $L_{\mathrm{ch}} = 9$ nm using ensemble Monte Carlo with state-of-the-art quantum corrections. Si-channel devices exhibit somewhat superior short channel behavior as evidenced by lower subthreshold swing and drain-induced barrier lowering. In addition, InGaAs devices exhibit just comparable drive currents to Si. The less-than-hoped-for InGaAs ON-state performance is a result of more severe adverse quantum degeneracy and confinement effects, whose origins were illuminated in Chapter 2. In addition, it was shown that the ON-state performance of n-channel FinFETs may be sensitive to the gate oxide's complex band structure, especially with light-mass III-

V channel materials, such as $In_{0.53}Ga_{0.47}As$. We studied this effect again using the ensemble semi-classical Monte Carlo device simulator of Chapter 2. Our simulations suggest that using a surface oxide with a heavy effective mass may lower the channel carrier confinement energies, somewhat mitigating unwanted quantum side-effects that hinder device performance. Ultimately, future high-k stacks for III-Vs may benefit from oxide gate stack heterostructures balancing effective mass and dielectric permittivity considerations. Si-based devices, which often have SiO_2 interfacial layers sandwiched between the Si channel and high-k gate oxide, already may be benefiting from such heterostructures.

Finally, in Chapter 4, single magnetic skyrmions were investigated via the first ever study of entire skyrmions executed completely within density-functional theory. A technique was shown to detect isolated nano-skyrmions with a current-perpendicular-to-plane geometry, which has immediate implications for device concepts. The physics behind such a mechanism was explored by studying the atomistic electronic structure of the magnetic quasiparticles. Investigation from first-principles showed how the isolated skyrmion local-density-of-states, from which electrons tunnel into the vacuum, when compared to that of the ferromagnetic background, is modified by the site-dependent spin-mixing of electronic states with different relative canting angles. Local transport properties are sensitive to this effect, as an atomistic conductance anisotropy of up to $\sim$20% was discovered for magnetic skyrmions in Pd/Fe/Ir(111) thin-films. In single skyrmions, engineering this spin-mixing magnetoresistance possibly could be incorporated in future magnetic storage

technologies.

5.2 Future work

There are viable pathways forward to continue the work established by this thesis. In terms of ensemble Monte Carlo studies, our simulator is viable to study other material systems of interest. Two possibilities are:

- *Germanium.* Ge may be the only semiconductor that could dethrone Si in future CMOS nodes. Not only does Ge have superior pMOS performance, but its respectable nMOS performance might be better than has been shown in past experiments. Certainly the possibility for all-Ge n/p CMOS is attractive and simple to implement versus III-V materials, which are more challenging to integrate. Improvements in Ge interface processing and demonstrations of Ge-FinFETs bolster the material's viability for future device nodes. Furthermore, Ge has an interesting balance of desirable qualities. Its large L-valley DOS means it may be robust to degeneracy effects, while its highly anisotropic transverse versus longitudinal effective masses could lead to optimized FinFET sidewall orientations, providing reduced quantum-confinement effects and large along-channel thermal velocities. Specifically, in such orientations, the Ge transport mass could approach $m^*_{\mathrm{L},t} = 0.08\ m_e$, while the confinement mass could approaches $m^*_{\mathrm{L},l} = 1.64\ m_e$. In this way, Ge may be able to mix large Si-like DOS with light III-V-like conductivity effective masses, while having smaller quantum-confinement effects than even

Si. However, given its non-orthogonal band-edge L-valley orientations, small intervalley separations such that L-, Γ-, and Δ-valleys all must be included in transport calculations, and at least 3 possible sidewall orientations to consider, use of Ge leads to a complex device system requiring Monte Carlo simulation to investigate its possible benefits.

- *2D material systems.* 2D materials such as transition metal dichalcogenides (TMDs) and topological insulators (TIs) may provide enhanced transistor behavior considering their ultimate electrostatic gate control over a truly 2D channel. Such materials may also provide novel device applications. Within the employed non-parabolicity approximation, we could consider parabolic through Dirac-like band structures in such 2D systems with our Monte Carlo simulator. For such simulations, however, motion in the confinement direction and associated quantum-confinement effects would need to be turned off, and consideration of additional bands and associated valleys may be needed. The biggest challenges would be obtaining the relevant band structures, phonon dispersion relations, and carrier-phonon interaction potentials. Density-functional calculations may prove useful for this purpose. For interaction potentials, careful comparison to experimental mobilities may be the best way to infer their values.

There also is room for inclusion of still more physics within our Monte Carlo simulator. Two possibilities are:

- *Carrier-carrier scattering.* We showed in this dissertation that classical molecular dynamic modeling of carrier-carrier scattering effects—intentionally or merely as a by-product of solving Poisson's equations in 3D in a particle based system—are not acceptable in degenerate systems because of the inherent lack of Pauli-blocking and associated thermal relaxation of carriers toward a *Boltzmann* distribution. We largely removed this issue by the introduction of sub-carriers. However, the role of carrier-carrier scattering to thermalize carrier populations toward *Fermi* distributions, particularly in the source and drain, may remain important. Although not clear at present, there may be a practical way of calculating carrier-carrier scattering rates within the Pauli-blocking framework, introducing associated thermal relaxation of carriers toward a Fermi distribution while also conserving overall energy and momentum. However, all permutations of interactions among particles would quickly become computationally unfeasible. A generalized carrier-carrier scattering rate based on the local electron density is needed, which could then could be applied to the individual particles while still conserving energy and momentum, at least on average.

- *Exchange effects.* In the degenerate limit, exchange effects could be approximated simply according to a local density approximation for the exchange potential V_{ex}. Then exchange effects could be included in a similar manner as our quantum-confinement effects—via effective potentials. In such a scheme, electrons would move within a total potential

$V_{\text{tot}} = V_{\text{el}} + V_{\text{QC}} + V_{\text{ex}}$, where V_{el} is the regular electrostatic potential given the solution of Poisson's equation, and V_{QC} is the quantum-correction potential as discussed in detail in Chapter 2.

The physics of magnetic skyrmions, which is really in its infancy but moving rapidly, could be explored in the following ways:

- *Skyrmion-defect interactions.* The simulations of Chapter 4 could be repeated but with an added near-surface defect close to the skyrmion. Possible defects could be metallic atoms, such as Cu, or even just simply an atomic vacancy. The purpose of this study would be to understand if skyrmions preferentially nucleate or pin at defects. If this were true, then defects could be intentionally engineered as skyrmion collection or nucleation centers for future skyrmion racetrack devices.

- *Skyrmion-skyrmion interactions.* Are skyrmions attracted to or repelled by other skyrmions? This question represents a potentially important issue about which relatively little is known. To execute this study, Heisenberg interactions of magnetic exchange could be extracted from the first-principles calculations of Chapter 4 and then used in a spin-lattice Hamiltonian to calculate the total energy as a function of skyrmion-skyrmion distance. The exchange tensor must be extracted for at least 3rd nearest neighbor interactions, however, making this a challenging, albeit possible, calculation.

Appendix

Monte Carlo simulation parameters

Listed are the simulated band structure and scattering parameters for Si and $In_{0.53}Ga_{0.47}As$ including the lattice constant (a_0), mass density (ρ), speed of sound (v_s), relative dielectric permittivity (ε_r), electron affinity ($q\chi$) , non-parabolicity constant (α), valley effective mass (m), acoustic deformation potential (Δ_{ac}), deformation field (DK), phonon energy ($\hbar\omega$), valley-wise bowing parameter (C_i), and intervalley separation (E_{ij}).

	Si	Units
a_0	5.43 ^{a}	Å
ρ	2.33 ^{a}	g/cm
v_s^l	9.18 ^{a}	$\times 10^5$ cm/s
v_s^t	4.70 ^{a}	$\times 10^5$ cm/s
ε_r^0	11.7 ^{a}	–
$q\chi$	4.05 ^{b}	eV
α_Δ	0.5 ^{c}	eV^{-1}
m_t^Δ	0.191 ^{c}	m_e
m_l^Δ	0.983 ^{c}	m_e
Δ_{ac}	5.0 (1.7 ^{a}) (9.0 ^{c})	eV
$(D_tK)_{g1}^\Delta$ (TA)	0.40 (0.50 ^{c})	$\times 10^8$ eV/cm
$\hbar\omega_{g1}^\Delta$	140 ^{c}	K
$(D_tK)_{g2}^\Delta$ (LA)	0.64 (0.80 ^{c})	$\times 10^8$ eV/cm
$\hbar\omega_{g2}^\Delta$	215 ^{c}	K
$(D_tK)_{g3}^\Delta$ (LO)	8.73 (11.00 ^{c})	$\times 10^8$ eV/cm
$\hbar\omega_{g3}^\Delta$	720 ^{c}	K
$(D_tK)_{f1}^\Delta$ (TA)	0.24 (0.30 ^{c})	$\times 10^8$ eV/cm
$\hbar\omega_{f1}^\Delta$	220 ^{c}	K
$(D_tK)_{f2}^\Delta$ (LA)	1.59 (2.00 ^{c})	$\times 10^8$ eV/cm
$\hbar\omega_{f2}^\Delta$	550 ^{c}	K
$(D_tK)_{f3}^\Delta$ (TO)	1.59(2.00 ^{c})	$\times 10^8$ eV/cm
$\hbar\omega_{f3}^\Delta$	685 ^{c}	K

[a]Fischetti, Ref. 52
[b]Goldberg, Ref. 140
[c]Jacoboni, Ref. 38

	InAs	GaAs	$In_{0.53}Ga_{0.47}As$	Units
a_0	6.04 [a]	5.64 [a]	5.85	Å
ρ	5.67 [a]	5.36 [a]	5.52	g/cm
v_s^l	4.28 [a]	5.24 [a]	4.73	$\times 10^5$ cm/s
v_s^t	2.65 [a]	2.47 [a]	2.57	$\times 10^5$ cm/s
ε_r^0	15.15 [a]	12.90 [a]	14.09	–
ε_r^∞	12.75 [a]	10.92 [a]	11.88	–
$q\chi$	4.90 [b]	4.07 [b]	4.51	eV
C_Γ	–	–	0.48 [g]	eV
C_L	–	–	0.58 (0.33 – 0.72) [g]	eV
C_X	–	–	1.09 (0.08 – 1.40) [g]	eV
$E_{\Gamma L}$	0.711 [c]	0.290 [c]	0.487	eV
$E_{\Gamma X}$	1.011 [c]	0.481 [c]	0.610	eV
$(DK)_{po}$	1.06 [a]	13.6 (2.1 [a])	6.95	$\times 10^8$ eV/cm
$\hbar\omega_{po}$	348 [a]	417 [a]	380	K
α_Γ	1.39 [e]	0.69 [f]	1.06	eV^{-1}
m_Γ	0.023 [b]	0.067 [d]	0.044	–
Δ_{ac}^Γ	10.0 (5.8) [a]	10.0 (5.0) [a]	10.0	eV
$DK_{(\Gamma\leftrightarrow L)}$	8.16 (5.59) [a]	5.25 [a]	6.81	$\times 10^8$ eV/cm
$\hbar\omega_{(\Gamma\leftrightarrow L)}$	347 [a]	322 [a]	335	K
$DK_{(\Gamma\leftrightarrow X)}$	8.16 (6.35) [a]	5.28 (5.48) [a]	6.81	$\times 10^8$ eV/cm
$\hbar\omega_{(\Gamma\leftrightarrow X)}$	347 [a]	347 [a]	347	K
α_L	0.54 [e]	0.65 [f]	0.59	eV^{-1}
m_t^L	0.286 [e]	0.075 [d]	0.187	–
m_l^L	0.286 [e]	1.900 [d]	1.04	–
Δ_{ac}^L	10.0 (5.8) [a]	9.2 (5.0) [a]	9.62	eV
$DK_{(L\leftrightarrow X)}$	4.76 (5.59) [a]	13.6 (5.01) [a]	8.91	$\times 10^8$ eV/cm
$\hbar\omega_{(L\leftrightarrow X)}$	341 [a]	341 [a]	341	K
$DK_{(L\leftrightarrow L)}$	5.28 (6.35) [a]	13.6 (5.94) [a]	9.19	$\times 10^8$ eV/cm
$\hbar\omega_{(L\leftrightarrow L)}$	338 [a]	338 [a]	338	K
α_X	0.90 [e]	0.36 [f]	0.65	eV^{-1}
m_t^X	0.640 [e]	0.19 [d]	0.429	–
m_l^X	0.640 [e]	1.9 [d]	1.23	–
Δ_{ac}^X	10.0 (5.8) [a]	9.7 (5.0) [a]	9.86	eV
$DK_{(X\leftrightarrow X)}$	4.76 (3.36) [a]	13.6 (2.99) [a]	8.91	$\times 10^8$ eV/cm
$\hbar\omega_{X\leftrightarrow X}$	347 [a]	347 [a]	347	K

[a] Fischetti, Ref. 52
[b] Goldberg, Ref. 140
[c] Adachi, Ref. 141
[d] Blakemore, Ref. 142
[e] Brennan, Ref. 143
[f] Brennan, Ref. 144
[g] Vurgaftman, Ref. 54

Bibliography

[1] J. A. del Alamo *Nature*, vol. 479, p. 317, 2011.

[2] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices.* Cambridge, UK: Cambridge University Press, 1998.

[3] S. K. Banerjee, L. F. Register, E. Tutuc, D. Basu, S. Kim, D. Reddy, and A. H. MacDonald *Proc. IEEE*, vol. 98, p. 2032, 2010.

[4] K. Bernstein, R. K. Cavin, W. Porod, A. Seabaugh, and J. Welser *Proc. IEEE*, vol. 98, p. 2169, 2010.

[5] R. R. Sinden, *DNA structure and function.* San Diego, CA: Academic Press, Inc., 1st ed., 1994.

[6] D. M. Crum, M. Bouhassoune, J. Bouaziz, B. Schweflinghaus, S. Blügel, and S. Lounis *Nature Commun.*, vol. 6, p. 8541, 2015.

[7] D. M. Crum, A. Valsaraj, J. K. David, L. F. Register, and S. K. Banerjee, "Ensemble Monte Carlo for III-V and Si n-channel FinFETs considering non-equilibrium degenerate statistics and quantum-confined scattering," *arXiv:1604.00085*, 2016.

[8] D. M. Crum, A. Valsaraj, L. F. Register, S. K. Banerjee, B. Sahu, Z. Krivakopic, S. Banna, and D. Nayak, "Impact of Gate Oxide Complex Band Structure on n-Channel III-V FinFETs," in *IEEE Int. Conf. Simul. Semicond. Process. Devices*, p. 250, 2015.

[9] D. Hisamoto, W. C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T. J. King, J. Bokor, and C. Hu *IEEE Trans. Electron Devices*, vol. 47, p. 2320, 2000.

[10] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices.* Hoboken, New Jersey: John Wiley & Sons Ltd, 3rd ed., 2007.

[11] M. Radosavljevic, G. Dewey, D. Basu, J. Boardman, B. Chu-Kung, J. M. Fastenau, S. Kabehie, J. Kavalieros, V. Le, W. K. Liu, D. Lubyshev, M. Metz, K. Millard, N. Mukherjee, L. Pan, R. Pillarisetty, W. Rachmady, U. Shah, H. W. Then, and R. Chau, "Electrostatics Improvement in 3-D Tri-gate Over Ultra-Thin Body Planar InGaAs Quantum Well Field Effect Transistors with High-K Gate Dielectric and Scaled Gate-to-Drain / Gate-to-Source Separation," in *IEEE Int. Electron Device Meet.*, p. 765, 2011.

[12] S. Ramey, A. Ashutosh, C. Auth, J. Clifford, M. Hattendorf, J. Hicks, R. James, A. R. V. Sharma, A. S. Amour, and C. Wiegand, "Intrinsic transistor reliability improvements from 22nm tri-gate technology," in *IEEE Int. Reliab. Phys. Symp. Proc.*, p. 4C.5.1, 2013.

[13] K. W. Ang, J. Barnett, W. Y. Loh, J. Huang, B. G. Min, P. Y. Hung, I. Ok, J. H. Yum, G. Bersuker, M. Rodgers, V. Kaushik, S. Gausepohl, C. Hobbs, P. D. Kirsch, and R. Jammy, "300mm finfet results utilizing conformal, damage free, ultra shallow

junctions (xj 5nm) formed with molecular monolayer doping technique," in *IEEE Int. Electron Device Meet.*, p. 35.5.1, 2011.

[14] G. Dewey, M. Radosavljevic, and N. Mukherjee, "Iii-v quantum well field effect transistors on silicon for future high performance and low power logic applications," in *IEEE Comp. Semicond. Int. Circ. Symp.*, p. 1, 2011.

[15] N. Mukherjee, J. Boardman, B. Chu-Kung, G. Dewey, A. Eisenbach, J. Fastenau, J. Kavalieros, W. K. Liu, D. Lubyshev, M. Metz, K. Millard, M. Radosavljevic, T. Stewart, H. W. Then, P. Tolchinsky, and R. Chau, "Movpe iiiâĂŞv material growth on silicon substrates and its comparison to mbe for future high performance and low power logic applications," in *IEEE Int. Electron Device Meet.*, p. 35.1.1, 2011.

[16] S. Lee, C. Y. Huang, D. Cohen-Elias, J. J. M. Law, V. Chobpattanna, S. Krämer, B. J. Thibeault, W. Mitchell, S. Stemmer, A. C. Gossard, and M. J. W. Rodwell *Appl. Phys. Letters*, vol. 103, p. 233505, 2013.

[17] M. Radosavljevic, B. Chu-Kung, S. Corcoran, G. Dewey, M. K. Hudait, J. M. Fastenau, J. Kavalieros, W. K. Liu, D. Lubyshev, M. Metz, K. Millard, N. Mukherjee, W. Rachmady, U. Shah, and R. Chau, "Advanced High-K Gate Dielectric for High-Performance Short-Channel In0.7Ga0.3As Quantum Well Field Effect Transistors on Silicon Substrate for Low Power," in *IEEE Int. Electron Device Meet.*, p. 319, 2009.

[18] N. Seoane, M. Aldegunde, K. Kalna, and A. J. Garcia-Loureiro, "MC/DD study of metal grain induced current variability in a nanoscale InGaAs FinFET," in *IEEE Int. Conf. Simul. Semicond. Process. Devices*, p. 253, 2014.

[19] M. Choi, V. Moroz, L. Smith, and J. Huang, "Extending Drift-Diffusion Paradigm into the Era of FinFETs and Nanowires," in *IEEE Int. Conf. Simul. Semicond. Process. Devices*, p. 242, 2015.

[20] M. Lundstrom, "Drift-diffusion and computational electronics - still going strong after 40 years!," in *IEEE Int. Conf. Simul. Semicond. Process. Devices*, p. 12, 2015.

[21] M. Lundstrom *IEEE Electron Device Lett.*, vol. 18, p. 361, 1997.

[22] M. Lundstrom and Z. Ren *IEEE Trans. Electron Devices*, vol. 49, p. 133, 2002.

[23] A. Svizhenko and M. P. Anantram *IEEE Trans. Electron Devices*, vol. 50, p. 1459, 2003.

[24] S. Datta, *Quantum Transport: Atom to Transistor.* Cambridge, UK: Cambridge University Press, 2005.

[25] A. Martinez, M. Bescond, J. R. Barker, A. Svizhenko, M. P. Anantram, C. Millar, and A. Asenov *IEEE Trans. Electron Devices*, vol. 54, p. 2213, 2007.

[26] S. M. Amoroso, V. P. Georgiev, L. Gerrer, E. Towie, W. Xingsheng, C. Riddet, A. R. Brown, and A. Asenov *IEEE Trans. Electron Devices*, vol. 61, p. 4014, 2014.

[27] K. M. Liu, W. Chen, L. F. Register, and S. K. Benerjee *J. Appl. Phys.*, vol. 104, p. 114515, 2008.

[28] W. Chen, L. F. Register, and S. K. Banerjee *J. App. Phys.*, vol. 103, p. 024508, 2008.

[29] D. M. Crum, A. Valsaraj, L. F. Register, and S. K. Banerjee, "Semi-Classical Ensemble Monte Carlo Simulator Using Innovative Quantum Corrections for Nano-Scale

n-Channel FinFETs," in *IEEE Int. Conf. Simul. Semicond. Process. Devices*, p. 109, 2014.

[30] M. V. Fischetti and S. E. Laux *Phys. Rev. B*, vol. 38, p. 9721, 1988.

[31] J. M. M. Pantoja and J. L. S. Franco *IEEE Electron Device Lett.*, vol. 18, p. 258, 1997.

[32] J. Mateos, T. González, D. P. V. Hoel, H. Happy, and A. Cappy *IEEE Trans. Electron Devices*, vol. 47, p. 250, 2000.

[33] A. Islam and K. Kalna *Semicond. Sci. and Technol.*, vol. 10, p. 055007, 2011.

[34] K. Kalna, M. Boriçi, L. Yang, and A. Asenov *Semicond. Sci. Technol.*, vol. 19, p. S202, 2004.

[35] A. Islam, B. Benbakhti, and K. Kalna *IEEE Trans. Nanotech.*, vol. 10, p. 1424, 2011.

[36] J. K. David, L. F. Register, and S. K. Banerjee *IEEE Trans. Electron Devices*, vol. 59, p. 976, 2012.

[37] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*. Wien, Austria: Springer-Verlag, 1989.

[38] C. Jacoboni and L. Reggiani *Rev. Mod. Phys.*, vol. 55, p. 645, 1983.

[39] B. Winstead and U. Ravaioli *IEEE Trans. Electron Devices*, vol. 50, p. 440, 2003.

[40] G. Kathawala, B. Winstead, and U. Ravaioli *IEEE Trans. Electron Devices*, vol. 50, p. 2467, 2003.

[41] X. F. Fan, X. Wang, B. Winstead, L. F. Register, U. Ravaioli, and S. K. Banerjee *IEEE Trans. Electron Devices*, vol. 51, p. 962, 2004.

[42] L. F. Register and N. Shi *J. Comput. Electron.*, vol. 9, p. 201, 2010.

[43] N. Shi, L. F. Register, and S. K. Banerjee, "On strain and scattering in deeply scaled n-channel MOSFETs: a quantum-corrected semiclassical monte carlo analysis," in *IEEE Int. Electron Device Meet.*, p. 1, 2008.

[44] N. Shi, L. F. Register, and S. K. Banerjee, "Semiclassical Monte Carlo with Quantum-Confinement Enhanced Scattering: Quantum Correction and Application to Short-Channel Device Performance Vs. Mobility for Biaxial-Tensile-Strained Silicon nMOSFETs," in *IEEE Int. Conf. Simul. Semicond. Process. Devices*, p. 1, 2009.

[45] J. K. David and L. F. Register and S. K Banerjee *Solid State Electron.*, vol. 61, p. 7, 2011.

[46] J. Lindberg, M. Aldegunde, D. Nagy, W. G. Dettmer, K. Kalna, A. J. García-Loureiro, and D. Perić *IEEE Trans. Electron Devices*, vol. 61, p. 423, 2014.

[47] D. Nagy, M. A. Elmessary, M. Aldegunde, R. Valin, A. Martinez, J. Lindberg, W. G. Dettmer, D. Perić, A. J. García-Loureiro, and K. Kalna *IEEE Trans. Nanotech.*, vol. 14, p. 93, 2015.

[48] G. W. Charache, D. M. DePoy, J. L. Egley, R. J. Dziendziel, M. J. Freeman, P. F. Baldasaro, B. C. Campbell, P. R. Sharps, M. L. Timmons, R. E. Fahey, K. Zhang, and J. M. Borrego, "Electrical and optical properties of degenerately-doped n-type ingaas," in *Proc. 3rd NREL Conf. Thermovoltaic Gener. Electr.*, p. 215, 1997.

[49] U. Singisetti, M. A. Wistey, J. D. Zimmerman, B. J. Thibeault, M. J. W. Rodwell, A. C. Gossard, and S. R. Bank *Appl. Phys. Lett.*, vol. 93, p. 183502, 2008.

[50] S. E. Laux *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.*, vol. 15, p. 1266, 1996.

[51] M. Aldegunde and K. Kalna, "Self-forces in 3D finite element Monte Carlo simulations of a 10.7 nm gate length SOI FinFET," in *IEEE Int. Conf. Simul. Semicond. Process. Devices*, p. P9, 2014.

[52] M. V. Fischetti *IEEE Trans. Electron Devices*, vol. 38, p. 634, 1991.

[53] J. A. Rowlette and K. E. Goodson *IEEE Trans. Electron Devices*, vol. 55, p. 220, 2008.

[54] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan *J. Appl. Phys.*, vol. 89, p. 5815, 2001.

[55] W. Porod and D. K. Ferry *Phys. Rev. B*, vol. 27, p. 2587(R), 1983.

[56] K. Y. Cheng, A. Y. Cho, S. B. Christman, T. P. Pearsall, and J. E. Rowe *Appl. Phys. Lett.*, vol. 40, p. 423, 1982.

[57] T. P. O'Regan and P. K. Hurley *Appl. Phys. Lett.*, vol. 99, p. 163502, 2011.

[58] G. Greene-Diniz, M. V. Fischetti, and J. C. Greer *J. Appl. Phys.*, vol. 119, p. 055707, 2016.

[59] D. K. Ferry *Phys. Rev. B*, vol. 14, p. 1605, 1976.

[60] H. Brooks and C. Herring *Phys. Rev. B*, vol. 83, p. 879, 1951.

[61] K. Hess, *Monte Carlo Device Simulation: Full Band and Beyond*. New York, New York: Kluwer Academic Publishers, 1991.

[62] M. A. Littlejohn, J. R. Hauser, and T. H. Glisson *Solid State Electronics*, vol. 21, p. 107, 1978.

[63] J. L. Thobel, L. Baudry, A. Cappy, P. Bourel, and R. Fauquembergue *Appl. Phys. Letters*, vol. 56, p. 346, 1990.

[64] V. Balynas, A. Krotkus, A. Stalnionis, A. T. Gorelionok, N. M. Shmidt, and J. A. Tellefsen *Appl. Phys. A*, vol. 51, p. 357, 1990.

[65] https://www.itrs.net, "International Technology Roadmap for Semiconductors," 2013.

[66] D. K. Ferry, *Semiconductor Transport*. New York, NY, USA: Taylor Francis, 2000.

[67] D. K. Ferry *Superlattices Microstruct.*, vol. 27, p. 61, 2000.

[68] D. K. Ferry and H. Grubin *Solid State Phys.*, vol. 49, p. 283, 1995.

[69] C. Heitzinger and C. Ringhofer *J. Comput. Electron.*, vol. 6, p. 401, 2007.

[70] C. Heitzinger, C. Ringhofer, S. Ahmes, and D. Vasileska *J. Comput. Electron.*, vol. 6, p. 15, 2007.

[71] M. G. Ancona and G. Iafrate *Phys. Rev. B*, vol. 39, p. 9536, 1989.

[72] A. J. García-Loureiro, N. Seoane, M. Aldegunde, R. Valin, A. Asenov, A. Martinez, and K. Kalna *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.*, vol. 30, p. 841, 2011.

[73] S. Yamakawa, H. Ueno, K. Taniguchi, and C. Hamaguchi *J. Appl. Phys.*, vol. 79, p. 911, 1996.

[74] T. Ando *J. Phys. Soc. Jpn.*, vol. 43, p. 1616, 1977.

[75] S. Mori and T. Ando *Phys. Rev. B*, vol. 19, p. 6433, 1979.

[76] T. Ando, A. B. Fowler, and F. Stern *Rev. Mod. Phys.*, vol. 54, p. 437, 1982.

[77] S. Jin, M. V. Fischetti, and T. W. Tang *IEEE Trans. Electron Devices*, vol. 54, p. 2191, 2007.

[78] K. Uchida and S. Takagi *Appl. Phys. Lett.*, vol. 82, p. 2916, 2003.

[79] Q. Liu, M. Vinet, J. Gimbert, N. Loubet, R. Wacquez, L. Grenouillet, Y. L. Tiec, A. Khakifirooz, T. Nagumo, K. Cheng, H. Kothari, D. Chanemougame, F. Chafik, S. Guillaumet, J. Kuss, F. Allibert, G. Tsutsui, J. Li, P. Morin, S. Mehta, R. Johnson, L. F. Edge, S. Ponoth, T. Levin, S. Kanakasabapathy, B. Haran, H. Bu, J. L. Bataillon, O. Weber, O. Faynot, E. Josse, M. Haond, W. Kleemeier, M. Khare, T. Skotnicki, S. Luning, B. Doris, M. Celik, and R. Sampson, "High performance UTBB FDSOI devices featuring 20nm gate length for 14nm node and beyond," in *IEEE Int. Electron Device Meet.*, p. 9.2.1, 2013.

[80] V. S. Basker, T. Standaert, H. Kawasaki, C. C. Yeh, K. Maitra, T. Yamashita, J. Faltermeier, H. Adhikari, H. Jagannathan, J. Wang, H. Sunamura, S. Kanakasabapathy, S. Schmitz, J. Cummings, A. Inada, C. H. Lin, P. Kulkarni, Y. Zhu, J. Kuss, T. Yamamoto, A. Kumar, J. Wahl, A. Yagishita, L. F. Edge, R. H. Kim, E. Mclellan, S. J. Holmes, R. C. Johnson, T. Levin, J. Demarest, M. Hane, M. Takayanagi, M. Colburn, V. K. Paruchuri, R. J. Miller, H. Bu, B. Doris, D. McHerron, E. Leobandung, and J. O'Neill, "A 0.063 Âţm2 FinFET SRAM cell demonstration with conventional lithography using a novel integration scheme with aggressively scaled fin and gate pitch," in *IEEE Symp. VLSI Tech.*, p. 19, 2010.

[81] N. Waldron, C. Merckling, L. Teugels, P. Ong, S. A. U. Ibrahim, F. Sebaai, A. Pourghaderi, K. Barla, N. Collaert, and A. V. Y. Thean *IEEE Electron Device Lett.*, vol. 35, p. 1097, 2014.

[82] S. H. Park, Y. Liu, N. Kharche, M. S. Jelodar, G. Klimeck, M. S. Lundstrom, and M. Luisier *IEEE Trans. Electron Devices*, vol. 59, p. 2107, 2012.

[83] K. D. Cantley, Y. Liu, H. S. Pal, T. Low, S. S. Ahmed, and M. S. Lundstrom, "Performance Analysis of III-V Materials in a Double-Gate nano-MOSFET," in *IEEE Int. Electron Device Meet.*, p. 113, 2007.

[84] D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur, and H. S. P. Wong *Proc. IEEE*, vol. 89, p. 259, 2001.

[85] T. V. Perevalov, A. V. Shaposhnikov, V. A. Gritsenko, H. Hong, J. H. Han, and C. W. Kim *JETP Letts.*, vol. 85, p. 165, 2007.

[86] S. Monaghan, P. K. Hurley, K. Cherkaoui, M. A. Negara, and A. Schenk *Solid-State Electronics*, vol. 53, p. 438, 2009.

[87] G. Allan, G. Bastard, N. Boccara, M. Lannoo, and M. Voos, *Heterojunctions and Semiconductor Superlattices*. Proceedings of the Winter School Les Houches, France: Springer, 1986.

[88] C. Chappert, A. Fert, and F. N. V. Dau, "The emergence of spin electronics in data storage," *Nature Mater.*, vol. 6, pp. 813–823, 2007.

[89] D. A. Thompson and J. S. Best, "The future of magnetic data storage techology," *IBM Journal of Research and Development*, vol. 44, pp. 311–322, May 2000.

[90] T. H. R. Skyrme, "A unified field theory of mesons and baryons," *Nucl. Phys.*, vol. 31, pp. 556–569, 1962.

[91] U. K. Rößler, A. N. Bogdanov, and C. Pfleiderer, "Spontaneous skyrmion ground states in magnetic metals," *Nature*, vol. 442, pp. 791–801, 2006.

[92] X. Z. Yu, Y. Onose, N. Kanazawa, J. H. Park, J. H. Han, Y. Matsui, N. Nagaosa, and Y. Tokura, "Real-space observation of a two-dimensional skyrmion crystal," *Nature*, vol. 465, pp. 901–904, 2010.

[93] X. Z. Yu, N. Kanazawa, Y. Onose, K. Kimoto, W. Z. Zhang, S. Ishiwata, Y. Matsui, and Y. Tokura, "Near room-temperature formation of a skyrmion crystal in thin-films of the helimagnet fege," *Nature Mater.*, vol. 10, pp. 106–109, 2011.

[94] A. Fert, V. Cros, and J. Sampaio, "Skyrmions on the track," *Nature Nanotech.*, vol. 8, pp. 152–156, 2013.

[95] M. Kläui, H. Ehrke, U. Rüdiger, T. Kasama, R. E. Dunin-Borkowski, D. Backes, L. J. Heyderman, C. A. F. Vaz, J. A. C. Bland, G. Faini, E. Cambril, and W. Wernsdorfer, "Direct observation of domain-wall pinning at nanoscale constrictions," *Appl. Phys. Lett.*, vol. 87, p. 102509, 2005.

[96] E. Martinez, L. Lopez-Diaz, O. Alejos, L. Torres, and C. Tristan, "Thermal effects on domain wall depinning from a single notch," *Phys. Rev. Lett.*, vol. 98, p. 267202, 2007.

[97] Y. Jang, S. Yoon, K. Lee, S. Lee, C. Nam, and B. K. Cho, "Current-induced domain wall nucleation and its pinning characteristics at a notch in a spin-valve nanowire," *Nanotechnology*, vol. 20, p. 125401, 2009.

[98] J. Iwasaki, M. Mochizuki, and N. Nagaosa, "Universal current-velocity relation of skyrmion motion in chiral magnets," *Nature Commun.*, vol. 4, p. 1463, 2013.

[99] J. Iwasaki, M. Mochizuki, and N. Nagaosa, "Current-induced skyrmion dynamics in constricted geometries," *Nature Nanotech.*, vol. 8, pp. 742–747, 2013.

[100] J. Sampaio, V. Cros, S. Rohart, A. Thiaville, and A. Fert, "Nucleation, stability, and current-induced motion of isolated magnetic skyrmions in nanostructures," *Nature Nanotech.*, vol. 8, pp. 839–844, 2013.

[101] F. Jonietz, S. Mühlbauer, C. Pfleiderer, A. Neubauer, W. Münzer, A. Bauer, T. Adams, R. Georgii, P. Böni, R. A. Duine, K. Everschor, M. Garst, and A. Rosch, "Spin transfer torques in MnSi at ultralow current densities," *Science*, vol. 330, pp. 1648–1651, 2010.

[102] X. Z. Yu, N. Kanazawa, W. Z. Zhang, T. Nagai, T. Hara, K. Kimoto, Y. Matsui, Y. Onose, and Y. Tokura, "Skyrmion flow near room temperature in an ultralow current density," *Nature Commun.*, vol. 3, p. 988, 2012.

[103] A. Thiaville, Y. Nakatani, J. Miltat, and Y. Suzuki, "Micromagnetic understanding of current-driven domain wall motion in patterned nanowires," *Europhys. Lett.*, vol. 69, pp. 990–996, 2005.

[104] A. V. Khalkovskiy, V. Cros, D. Apalkov, V. Nikitin, M. Krounbi, K. A. Zvezdin, A. Anane, J. Grollier, and A. Fert, "Matching domain-wall configuration and spin-orbit torques for efficient domain-wall motion," *Phys. Rev. B*, vol. 87, p. 020402, 2013.

[105] A. V. Khalkovskiy, K. A. Zvezdin, Y. V. Gorbunov, V. Cros, J. Grollier, A. Fert, and A. K. Zvezdin, "High domain wall velocities due to spin currents perpendicular to the plane," *Phys. Rev. Lett.*, vol. 102, p. 067206, 2009.

[106] J. Müller and A. Rosch, "Capturing of a magnetic skyrmion with a hole," *Phys. Rev. B*, vol. 91, p. 054410, 2015.

[107] S. Woo, K. Litzius, M. Y. I. B. Krüger, L. Caretta, K. Richter, M. Mann, A. Krone, R. Reeve, M. Weigand, P. Agrawal, P. Fischer, M. Kläui, and G. S. D. Beach, "Observation of room temperature magnetic skyrmions and their current-driven dynamics in ultrathin matallic ferromagnets," *Nature Mat.*, p. doi:10.1038/nmat4593, 2016.

[108] N. Romming, A. Kubetzka, C. Hanneken, K. von Bergmann, and R. Wiesendanger, "Field-dependent size and shape of single magnetic skyrmions," *Phys. Rev. Lett.*, vol. 114, p. 177203.

[109] S. S. P. Parkin, M. Hayashi, and L. Thomas, "Magnetic domain-wall racetrack memory," *Science*, vol. 320, pp. 190–194, 2008.

[110] N. S. Kiselev, A. N. Bogdanov, R. Schäffer, and U. K. Rößler, "Chiral skyrmions in thin magnetic films: new objects for magnetic storage technologies?," *J. Phys. D*, vol. 44, p. 392001, 2011.

[111] N. Romming, C.Hanneken, M. Menzel, J. E. Bickel, B. Wolter, K. von Bergmann, A. Kubetzka, and R. Wiesendanger, "Writing and deleting single magnetic skyrmions," *Science*, vol. 341, pp. 636–639, 2013.

[112] B. Dupé, G. Bihlmayer, S. Blügel, and S. Heinze, "Engineering skyrmions in transition-metal multilayers for spintronics," *arXiv:1503.08098*, 2015.

[113] T. Schulz, R. Ritz, A. Bauer, M. Halder, M. Wagner, C. Franz, C. Pfleiderer, K. Everschor, M. Garst, and A. Rosch, "Emergent electrodynamics of skyrmions in a chiral magnet," *Nature Phys.*, vol. 8, pp. 301–304, 2012.

[114] A. Neubauer, C. Pfleiderer, B. Binz, A. Rosch, R. Ritz, P. G. Niklowitz, and P. Böni, "Topological hall effect in the A Phase of MnSi," *Phys. Rev. Lett.*, vol. 102, p. 186602, 2009.

[115] N. Kanazawa, M. Kubota, A. Tsukazaki, Y. Kozuka, K. S. Takahashi, M. Kawasaki, M. Ichikawa, F. Kagawa, and Y. Tokura, "Discretized topological hall effect emerging from skyrmions in constricted geometry," *Phys. Rev. B*, vol. 91, p. 041122, 2015.

[116] F. Freimuth, R. Bamler, Y. Mokrousov, and A. Rosch, "Phase-space berry phases in chiral magnets: Dzyaloshinskii-moriya interaction and the charge of skyrmions," *Phys. Rev. B*, vol. 88, p. 214409, 2013.

[117] C. Franz, F. Freimuth, A. Bauer, R. Ritz, C. Schnarr, C. Duvinage, T. Adams, S. Blügel, A. Rosch, Y. Mokrousov, and C. Pfleiderer, "Real-space and reciprocal-space Berry Phases in the Hall effect of MnFeSi," *Phys. Rev. Lett.*, vol. 112, p. 186601, 2014.

[118] W. P. Pratt, S. F. Lee, J. M. Slaughter, R. Loloee, P. A. Schroeder, and J. Bass, "Perpendicular giant magnetoresistances of Ag/Co multilayers," *Phys. Rev. Lett.*, vol. 66, p. 3060, 1991.

[119] S. Zhang and P. M. Levy, "Conductivity perpendicular to the plane of multilayered structures," *J. Appl. Phys.*, vol. 69, p. 4786, 1991.

[120] M. Bode, S. Heinze, A. Kubetzka, O. Pietzsch, X. Nie, G. Bihlmayer, S. Blügel, and R. Wiesendanger, "Magnetization-direction-dependent local electronic structure probed by scanning tunneling spectroscopy," *Phys. Rev. Lett.*, vol. 89, p. 237205, 2002.

[121] C. Gould, C. Ruster, T. Jungwirth, E. Girgis, G. M. Schott, R. Giraud, K. Brunner, G. Schmidt, and L. W. Molenkamp, "Tunneling anisotropic magnetoresistance: a spin-valve-like tunnel magnetoresistance using a single magnetic layer," *Phys. Rev. Lett.*, vol. 93, p. 117203, 2004.

[122] K. von Bergmann, M. Menzel, D. Serrate, Y. Yoshida, S. Schröder, P. Ferriani, A. Kubetzka, R. Wiesendanger, and S. Heinze, "Tunneling anisotropic magnetoresistance on the atomic scale," *Phys. Rev. B*, vol. 86, p. 134422, 2012.

[123] A. O. Leonov, T. L. Monchesky, N. Romming, A. Kubetzka, A. N. Bogdanov, and R. Wiesendanger, "The properties of isolated chiral skyrmions in thin magnetic films," *arXiv:1508.02155*, 2015.

[124] B. Dupé, M. Hoffmann, C. Paillard, and S. Heinze, "Tailoring magnetic skyrmions in ultra-thin transition metal films," *Nature Commun.*, vol. 5, p. 4030, 2014.

[125] E. Simon, K. Palotás, L. Rózsa, L. Udvardi, and L. Szunyogh, "Formation of magnetic skyrmions with tunable properties in PdFe bilayer deposited on Ir(111)," *Phys. Rev. B*, vol. 90, p. 094410, 2014.

[126] B. E. Feldman, B. Krauss, J. H. Smet, and A. Yacoby, "Unconventional sequence of fractional quantum hall states in suspended graphene," *Science*, vol. 337, pp. 1196–1199, 2012.

[127] I. Dzyaloshinskii, “A thermodynamic theory of ’weak’ ferromagnetism of antiferromagnets,” *J. Phys. Chem. Solids*, vol. 4, pp. 241–255, 1958.

[128] T. Moriya, “Anisotropic superexchange interaction and weak ferromagnetism,” *Phys. Rev.*, vol. 120, pp. 91–98, 1960.

[129] A. Crépieux and C. Lacroix, “Dzyaloshinskii-Moriya interactions induced by symmetry breaking at a surface,” *J. Magn. Magn. Mater.*, vol. 182, pp. 341–349, 1988.

[130] D. Bauer, P. Mavropoulos, R. Zeller, and S. Blügel, *Development of a relativistic full-potential first-principles multiple scattering Green function method applied to complex magnetic textures of nano structures at surfaces.* PhD thesis, RWTH Aachen University, 2014.

[131] S. Heinze, K. von Bergmann, M. Menzel, J. B. A. Kubetzka, R. Wiesendanger, G. Bihlmayer, and S. Blügel, “Spontaneous atomic-scale magnetic skyrmion lattice in two dimensions,” *Nature Phys.*, vol. 7, pp. 713–718, 2011.

[132] J. Tersoff and D. R. Hamann, “Theory and application for the scanning tunneling microscope,” *Phys. Rev. Lett.*, vol. 50, p. 1998, 1983.

[133] N. M. Caffrey, S. Schröder, P. Ferriani, and S.Heinze, “Tunneling anisotropic magnetoresistance effect of single adatoms on a noncollinear magnetic surface,” *J. Phys.: Condens. Matter*, vol. 26, p. 394010, 2014.

[134] S. Alexander and P. W. Anderson, “Interaction between localized states in metals,” *Phys. Rev. A*, vol. 133, pp. A1594–A1603, 1964.

[135] A. Oswald, R. Zeller, P. J. Braspenning, and P. H. Dederichs, “Interaction of magnetic impurities in Cu and Ag,” *J. Phys. F. Met. Phys.*, vol. 15, pp. 193–212, 1985.

[136] S. Tehrani, J. M. Slaughter, M. Deherrera, B. N. Engel, N. D. Rizzo, J. Salter, M. Durlam, R. W. Dave, J. Janesky, R. Butcher, K. Smith, and G. Grynkewich, "Magnetoresistive random access memory using magnetic tunnel junctions," *Proc. IEEE*, vol. 91, pp. 703–714, 2003.

[137] A. V. Pohm, J. S. T. Huang, J. M. Daughton, D. R. Krahn, and V. Mehra, "The design of a one megabit nonvolatile M-R memory chip using 1.5x5 micron cells," *IEEE Trans. Magn.*, vol. 24, pp. 3117–3119, 1988.

[138] Y. H. Son, O. Seongil, Y. Ro, J. W. Lee, and J. H. Ahn, "Reducing memory access latency with asymmetric DRAM bank organizations," in *IEEE Int. Symp. Comp. Arch.*, pp. 380–391, 2013.

[139] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis," *Can. J. Phys.*, vol. 58, p. 1200, 1980.

[140] Y. A. Goldberg and N. M. Schmidt, *Handbook Series on Semiconductor Parameters*, vol. 2nd. London, England: World Scientific, 1999.

[141] S. Adachi, *Properties of Semiconductor Alloys.* West Sussex, UK: John Wiley & Sons Ltd, 2009.

[142] J. S. Blakemore *J. Appl. Phys.*, vol. 53, p. R123, 1982.

[143] K. Brennan and K. Hess *Solid-State Electron.*, vol. 27, p. 347, 1984.

[144] K. Brennan, D. H. Park, K. Hess, and M. A. Littlejohn *J. Appl. Phys.*, vol. 63, p. 5004, 1988.

Vita

Dax M. Crum is an Electrical Engineer specializing in semiconductor physics and a National Science Foundation (NSF) Graduate Research Fellow, a University of Texas at Austin (UT) Microelectronics and Computer Development (MCD) Fellow, and a UT Cockrell School of Engineering (CSE) Fellow. His primary work is the simulation, modeling, and characterization of end-of-the-roadmap CMOS and beyond CMOS nanoelectronic materials and devices.

Mr. Crum was born and raised in Phoenix, Arizona, and spent most of his youth submersed in athletics. His athletic career culminated in a full Division I scholarship to play quarterback for the University of Wyoming, where he earned his undergraduate degree (B.S.E.E.) in Electrical and Computer Engineering in 2011.

Mr. Crum graduates with a doctorate of Electrical Engineering in May 2016 from UT before beginning his career at Intel Corporation within their device modeling and research group in Hillsboro, Or., where he hopes to contribute to the scientific and secular communities by pushing the state-of-the-art for CMOS technology and beyond.

Permanent address: dcrum@utexas.edu

This dissertation was typeset with LaTeX[†] by Dax Michael Crum.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.