Ab-initio quantum transport with a basis of unit-cell restricted Bloch functions and the NEGF formalism

Marco G. Pala, David Esseni

Ab-initio quantum transport with a basis of unit-cell restricted Bloch functions and the NEGF formalism

Marco Pala
Université Paris-Saclay, CNRS
Centre de Nanosciences et de Nanotechnologies
91120 Palaiseau, France
marco.pala@c2n.upsaclay.fr

David Esseni
DPIA
University of Udine
33100 Udine, Italy
david.esseni@uniud.it

***Abstract*—This invited contribution illustrates the theory and application of a first-principle transport methodology employing a basis set obtained directly from the Bloch functions computed with a plane wave (PW) *ab-initio* solver. We start from a PW density functional theory (DFT) Hamiltonian, use a unitary transformation to real space in the transport direction, and then discuss a basis of Bloch functions enabling a huge reduction of the size of the Hamiltonian blocks and an effective suppression of possible unphysical states. Our methodology enables *ab-initio* transport simulations with a good computational efficiency, and we here present results for self-consistent simulations of a single-gate monolayer $PtSe_2$ field effect transistor.**

***Index Terms*—Density functional theory, quantum transport, NEGF, 2D materials**

I. Introduction

In recent years many innovations in nanoelectronics devices have been driven by the exploitation of new materials, such as atomically thin 2D materials and their heterostructures [1]–[5]. *Ab-initio* methods based on the DFT have been shown to be quite effective to investigate the electronic and optical properties of these new materials, but their direct use for quantum transport calculations has been hindered by the size and complexity of their Hamiltonians. Remarkable exceptions are the contributions employing Hamiltonians based on the linear combination of local orbitals (LCAO) [6], or on the maximally localized Wannnier functions (MLWF) [7], [8].

Here we discuss an alternative approach to *ab-initio* quantum transport [9], that leverages directly on the Hamiltonian matrix and Bloch functions obtained with plane-wave DFT solvers [10], [11]. We believe that such a methodology, differently from the widely used MLWF approach, can be naturally integrated into the workflow of an *ab-initio* solver, and thus provide a fairly automatic pathway to *ab-initio* quantum transport simulations.

II. Methodology

In order to obtain the DFT Hamiltonian, we start by self-consistently solving the Kohn-Sham (KS) equations

$$H_{KS}\,\Psi_n = E_n\,\Psi_n, \qquad H_{KS} = T + V_{scf} \tag{1}$$

where the KS Hamiltonian H_{KS} is the sum of the operators corresponding to the kinetic energy T and to the self-consistent potential $V_{scf} = V_{eI} + V_H + V_{xc}$, where V_{eI} is the electron-ion interaction potential, V_H the Hartree electrostatic potential, V_{xc} the exchange-correlation potential. To this goal we use a PW basis set and pseudopotentials to represent the valence electron-nuclei interactions. The KS Hamiltonian is then written as

$$\begin{aligned} H(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}') &= \frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2\delta_{\mathbf{G},\mathbf{G}'} + V_L(\mathbf{G}-\mathbf{G}') \\ &+ V_{NL}(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}') \end{aligned} \tag{2}$$

where $V_L(\mathbf{G}-\mathbf{G}')$ is the Fourier transform of the local part of the total potential and $V_{NL}(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}')$ is the spatially non-local contribution of the pseudopotential, that for most atoms is short ranged and vanishes for $|\mathbf{r}-\mathbf{r}'|>r_c$, with r_c being the radius beyond which pseudo and true KS orbitals coincide [12].

For transport calculations it is convenient to use an orthorombic unit cell [9], where the real space unit vectors can be written as $\mathbf{a_1}=(a_x,0,0)$, $\mathbf{a_2}=(0,a_y,0)$, $\mathbf{a_3}=(0,0,a_z)$, with x being the transport direction. Hence the unit vectors of the reciprocal lattice are $\mathbf{b_1}=(2\pi/a_x,0,0)$, $\mathbf{b_2}=(0,2\pi/a_y,0)$, $\mathbf{b_3}=(0,0,2\pi/a_z)$. The Brillouin zone (BZ) is therefore given by the conditions $-\pi/a_s < k_s \leq \pi/a_s$ (with $s=x,y,z$).

A. Discussion of the reduced basis set

The DFT Hamiltonian in Eq. (2) is a dense matrix with a rank equal to the number, N_G, of reciprocal lattice vectors, and thus it cannot be directly employed in transport calculations. Thus our first step is a unitary transformation from PW to real space along the transport direction x, that provides the Hamiltonian in the hybrid $x\mathbf{K}_{yz}$ basis given by real-space along x and plane waves in the (y,z) plane [13]. In this representation the Hamiltonian matrix is a block tri-diagonal matrix, which is crucially important in order to exploit recursive algorithms for non-equilibrium Green's function (NEGF) calculations [14]. In this hybrid $x\mathbf{K}_{yz}$ basis each block describes an a_x long spatial region consisting of N_{dx} discretization points, so that the matrix blocks have a rank $N_G{=}N_{dx}N_{Gy}N_{Gz}$ (with $N_{dx}{=}N_{Gx}$). Even if the $x\mathbf{K}_{yz}$ basis has been used to perform quantum transport calculations as in [13], [15], the size N_G of the blocks is large and it increases by enlarging the cutoff energy used in DFT calculations. Huge computational

advantages can thus be obtained by moving to a basis of unit-cell restricted Bloch functions, that preserves the block tri-diagonal structure of the Hamiltonian matrix [9], [16].

To this purpose we first argue that, due to the block tridiagonal form of the Hamiltonian, the unit cell restricted Bloch functions $\{\Psi^0_{\mathbf{k}}\}$ corresponding to the wave vector $\mathbf{k}$=(k_x,$\mathbf{k}_{yz}$) satisfy the secular equation

$$\left[\mathbf{H}^{\dagger}_{0,1}\, e^{-ik_x a_x} + \mathbf{H}_{0,0} + \mathbf{H}_{0,1}\, e^{ik_x a_x} \right] \{\Psi^0_{\mathbf{k}}\} = E(\mathbf{k})\, \{\Psi^0_{\mathbf{k}}\} \tag{3}$$

with $\mathbf{H}_{0,0}$, $\mathbf{H}_{0,1}$ being the Hamiltonian blocks describing respectively one unit cell and its interaction with the nearest neighbor cell. Our reduced basis set consists of a subset of the $\{\Psi^0_{n\mathbf{k}}\}$ corresponding to a few k_x values in the reduced zone $-\pi/a_x < k_x \leq \pi/a_x$ along the transport direction and, for each k_x, to some tens of bands $E_n(\mathbf{k})$ up to the energies of interest for the problem at study.

It is noteworthy that the $\{\Psi^0_{n\mathbf{k}}\}$ are obtained directly from the PW representation provided by the DFT solver, namely *we do not need to solve Eq. (3)* to find the basis functions.

The size of the basis can be written as

$$N_B = \sum_{n=n_i}^{n_f} N_{kB,n} \tag{4}$$

where $N_{kB,n}$ denotes the number of k_x values included in the basis for the n-th band, with n ranging between n_i and n_f. The $\{\Psi^0_{n\mathbf{k}}\}$ functions are not orthogonal for different k_x, but they can be orthogonalized and we denote by Φ the basis set orthonormalized over a unit cell. The Hamiltonian blocks in the Φ basis have a size N_B that is typically much smaller than the size N_G in the hybrid $x\mathbf{K}_{yz}$ basis.

A first validation of the reduced basis set requires that it allows us to calculate the band structure for any k_x in the reduced zone, which is discussed below in Sec.II-B.

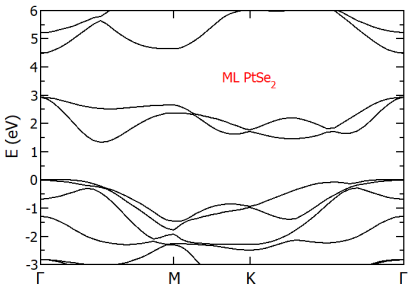

Fig. 1. Bandstructure of the monolayer $PtSe_2$.

B. Material system and validation of the reduced basis set

In this paper we exemplify our methodology by focusing on the monolayer $PtSe_2$, which is a noble metal dichalcogenure identified as a promising channel material for MOSFETs due to the small effective mass and large density of states (DOS) [17]. Figure 1 reports the corresponding band structure along the high-symmetry points of the primitive BZ obtained from the DFT Hamiltonian in the PW basis. The DFT calculation was performed with the Quantum ESPRESSO code [11], by using the Perdew-Burke-Ernzerhof [18] approximation to the exchange-correlation functional and a norm-conserving pseudopotential [19]. We employed a $12 \times 12 \times 1$ Monkhorst-Pack k-points grid and a cutoff energy of $E_w = 60$ Ry.

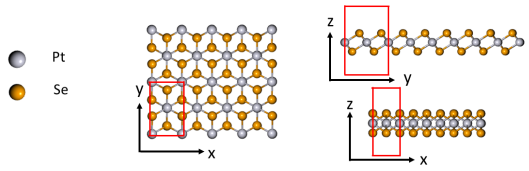

Fig. 2. Orthorombic unit cell of the 1T phase monolayer $PtSe_2$.

For transport calculations we first define the orthorombic cell composed of 6 atoms and sketched in Fig. 2. The unit vectors are $\mathbf{a_1}$=$(a_x, 0, 0)$, $\mathbf{a_2}$=$(0, a_y, 0)$, $\mathbf{a_3}$=$(0, 0, a_z)$ with a_x=3.712 Å, a_y=6.429372 Å and a_z=32 Å, while the vertical distance between the Pt and Se atoms is 1.312 Å. In order to validate the reduced Bloch state basis, we compared the bands computed with the Hamiltonian blocks in the reduced basis with the results of the plane-wave DFT solver. The accuracy of the reduced basis depends on the included number of k_x Bloch functions [9], hence we optimized the basis by using a larger $N_{kB,n}$ for bands with energies in the range of interest, and smaller $N_{kB,n}$ for bands at much lower or higher energies.

In Fig. 3 we show an excellent band-structure reconstruction obtained by using $N_{kB,n}$=4 for the bands close to the energygap and $N_{kB,n}$=2 for smaller energy bands, and for an overall number N_B=74 of Bloch functions. Such a definition of the reduced basis results in no observable unphysical states [20], [21].

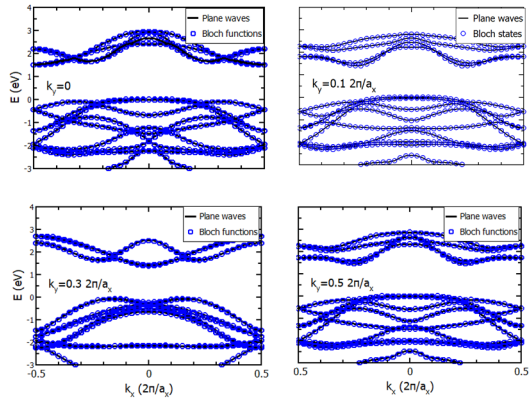

Fig. 3. Bandstructure of the monolayer $PtSe_2$ for various k_y values computed with (circles) the Bloch function basis and (lines) the PW basis.

C. Carrier densities and electron current

All relevant physical quantities are defined in terms of the retarded(advanced), $[\mathbf{G}^{r(a)}]_{\Phi}$, and the lesser(greater)-than Green's functions, $[\mathbf{G}^{<(>)}]_{\Phi}$, computed in the reduced Bloch

functions basis. In order to compute the carrier concentrations in the real space, we first compute $[\mathbf{G}^{<(>)}]_\Phi$, then we reconstruct the Green's function in the hybrid basis according to the expansion

$$\mathbf{G}^{<(>)}(x_j\mathbf{K}_{yz}, x_j\mathbf{K}'_{yz}; E) = \sum_{n,m=1}^{N_B} \mathbf{G}^{<(>)}(n,m;E) \times \Phi_n(x_j, \mathbf{G}_{yz})\, \Phi^*_m(x_j, \mathbf{G}'_{yz}) \quad (5)$$

and finally evaluate electron and hole densities as described in Ref. [9]. Thanks to this procedure we are able to evaluate the 3D electron and hole concentrations, $n(\mathbf{r})$ and $p(\mathbf{r})$, with atomistic accuracy as shown in the example of Fig. 4.

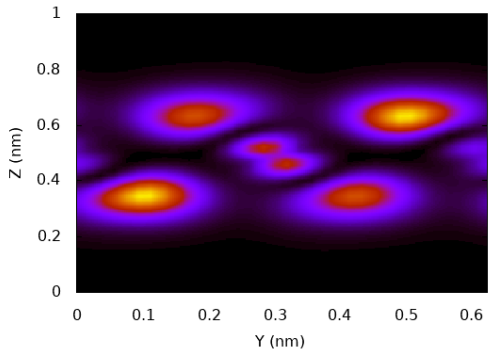

Fig. 4. Contourplot of the cross section of the electron density at an abscissa x corresponding to the center of the orthorombic unit cell. Colors are in arbitrary units.

However, by leveraging the fact that the electrostatic potential varies over larger spatial scales compared to the discretization used in Eq. (5), in the self-consistent solution of the NEGF transport equations and the Poisson equation we evaluate $n(\mathbf{r})$ and $p(\mathbf{r})$ on a coarser mesh. For the 3D Poisson problem we assume Diriclect conditions at the metal boundaries and Neumann conditions elsewhere.

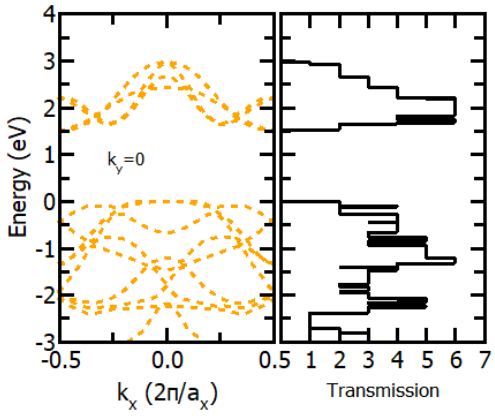

Fig. 5. Bandstructure of an infinitely-long monolayer $PtSe_2$ and the corresponding transmission probability obtained with the Hamiltonian in the reduced basis consisting of unit-cell-restricted Bloch states. $k_y = 0$.

The electron current is obtained within the elastic approximation as

$$I_D = \frac{ge}{h} \int T(E)\,[f_D(E) - f_S(E)]\; dE \quad (6)$$

with g being the degeneracy and $f_{S(D)}$ the Fermi-Dirac function at the source (drain) contact. Figure 5 illustrates the bandstructure and transmission of an infinitely long monolayer $PtSe_2$ at $k_y = 0$ calculated by using the reduced basis and showing that, as expected, the transmission at a given energy E equals the number of available bands.

The device under investigation is sketched in Fig. 6. It is a single-gate FET with doped source and drain access regions having a length $L_{S,D}$≈11.1 nm and an undoped channel region with L_G≈18.5 nm. The donor concentration in the source and drain region is N_D=10^{20} cm^{-3}. The equivalent oxide thickness (EOT) of the high-k dielectric layer is 0.65 nm. In the device width direction, y, we assume periodic boundary conditions, described by sampling the BZ along the k_y direction with a step Δk_y=$0.1 \times 2\pi/a_y$.

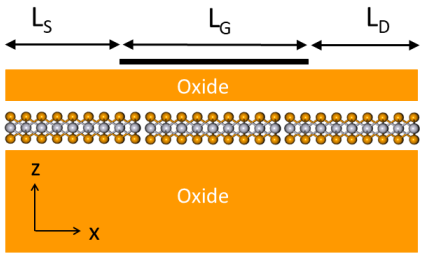

Fig. 6. Sketch of the single-gate monolayer $PtSe_2$ FET.

Figure 7 shows the transfer characteristics for different V_{DS} values. In the sub-threshold voltage regime we observe an ideal sub-threshold swing of about 60 mV/dec, which stems from the excellent electrostatic integrity of the device due, in turn, to the ultra-thin channel and the aggressive EOT. The device exhibits also a high transconductance resulting in a high on-state current, that confirms the promising properties of this material for high-performance MOSFETs. Of course the transconductance and on-state current in Figure 7 should be considered as upper limit figures, that can be significantly degraded by series resistance and scattering [22].

Quite interestingly, in Figure 7 we can also observe a non-monotonic I_D dependence on V_{DS}, in fact the I_D at large V_{TG} increases with V_{DS} for V_{DS} up to 0.3 V, but then it decreases for larger V_{DS}. This behavior has been already observed in simulated FETs with a monolayer MoS_2 channel [23], and it has been ascribed to peculiar features of the DOS. We verified that a similar explanation applies to our results, in fact Fig. 8 reports the local density of states (LDOS) at V_{DS}=0.3 V and 0.6 V in the on-state regime (i.e. at V_{TG}=0.8 V), showing that in the drain region the LDOS for energies close to the top of the barrier is larger at V_{DS}=0.3 V than it is at V_{DS}=0.6 V.

Such a non-monotonic I_D dependence on V_{DS}, however, is expected to disappear in the presence of significant inelastic scattering mechanisms, such as electron-phonon scattering.

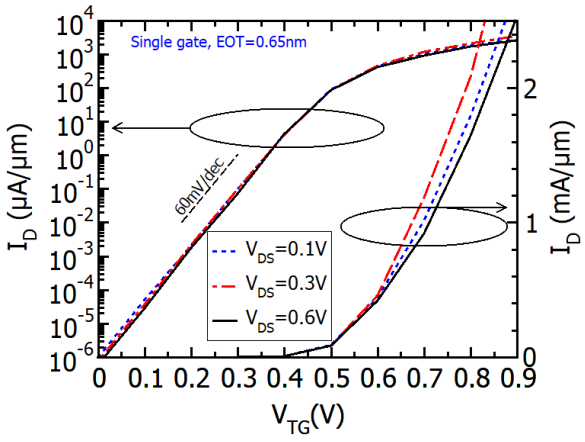

Fig. 7. Transfer characteristics of the monolayer $PtSe_2$ FET for V_{DS}=0.1, 0.3 and 0.6 V.

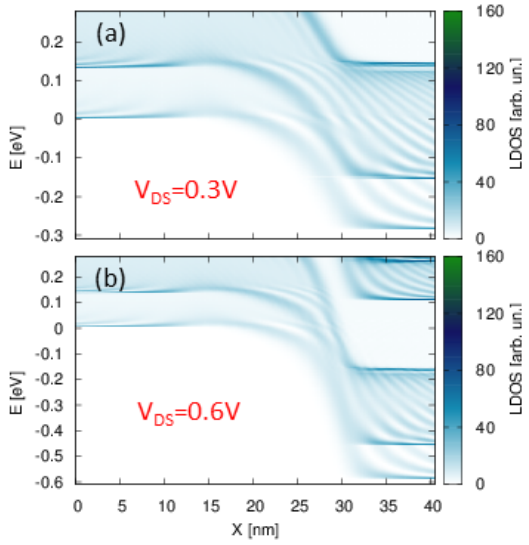

Fig. 8. LDOS computed for (a) V_{DS}=0.3 V and (b) V_{DS}=0.6 V and for $V_{TG} = 0.8$ V. The transverse wave-vector is $k_y = 0.1 \times 2\pi/a_y$.

III. Conclusions and Outlook

We have shown that a unit-cell restricted Bloch functions basis enables band structure and transport calculations with a significant reduction of the computational burden compared to the original plane-wave DFT formulation. This methodology is a viable approach for *ab initio* and semi-empirical quantum transport simulations, and it can be seen as an alternative to the MLWF approach. In fact the Bloch functions can be obtained directly from the DFT solvers, thus bypassing the extraction of a MLWF basis and of the corresponding tight-binding Hamiltonian. We foresee that further developments of the methodology will enable us to deal with non homogenous systems (i.e. including hetero-jnctions), and to include inelastic scattering mechanisms.

References

[1] L. Mingda, *et al.*, "Single Particle Transport in Two-dimensional Heterojunction Interlayer Tunneling Field Effect Transistor," *Journal of Applied Physics*, vol. 115, p. 074508, 2014.

[2] D. Sarkar, *et al.*, "A subthermionic tunnel field-effect transistor with an atomically thin channel," *Nature*, vol. 526, no. 7571, pp. 91–95, 2015.

[3] J. Cao, *et al.*, "Operation and Design of van der Waals Tunnel Transistors: A 3-D Quantum Transport Study," *IEEE Transactions on Electron Devices*, vol. 63, no. 11, pp. 4388–4394, 2016.

[4] D. Logoteta, *et al.*, "A steep-slope MOS_2-nanoribbon MOSFET based on an intrinsic cold-contact effect," *IEEE Electron Device Letters*, vol. 40, no. 9, pp. 1550–1553, 2019.

[5] E. G. Marin, *et al.*, "Lateral heterostructure field-effect transistors based on two-dimensional material stacks with varying thickness and energy filtering source," *ACS Nano*, vol. 14, no. 2, pp. 1982–1989, 2020.

[6] J. M. Soler, *et al.*, "The SIESTA method for ab initio order-n materials simulation," *Journal of Physics: Condensed Matter*, vol. 14, no. 11, pp. 2745–2779, mar 2002.

[7] N. Marzari and D. Vanderbilt, "Maximally localized generalized Wannier functions for composite energy bands," *Phys. Rev. B*, vol. 56, p. 12847, Nov 1997.

[8] A. Szabó, R. Rhyner, and M. Luisier, "Ab initio simulation of single- and few-layer MoS_2 transistors: Effect of electron-phonon scattering," *Phys. Rev. B*, vol. 92, p. 035435, Jul 2015.

[9] M. G. Pala, P. Giannozzi, and D. Esseni, "Unit cell restricted Bloch functions basis for first-principle transport models: Theory and application," *Phys. Rev. B*, vol. 102, p. 045410, Jul 2020.

[10] G. Kresse and J. Furthmüller, "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set," *Computational Materials Science*, vol. 6, no. 1, pp. 15 – 50, 1996.

[11] P. Giannozzi *et al.*, "Quantum Espresso: a modular and open-source software project for quantum simulations of materials," *Journal of Physics: Condensed Matter*, vol. 21, no. 39, p. 395502 (19pp), 2009.

[12] P. Giannozzi *et al.*, "Advanced capabilities for materials modelling with Quantum Espresso," *Journal of Physics: Condensed Matter*, vol. 29, no. 46, p. 465901, 2017.

[13] M. G. Pala and D. Esseni, "Full-band quantum simulation of electron devices with the pseudopotential method: Theory, implementation, and applications," *Phys. Rev. B*, vol. 97, p. 125310, 2018.

[14] M. P. Anantram, M. S. Lundstrom, and D. E. Nikonov, "Modeling of Nanoscale Devices," *Proceedings of the IEEE*, vol. 96, no. 9, pp. 1511–1550, Sept 2008.

[15] M. Pala, O. Badami, and D. Esseni, "NEGF based transport modeling with a full-band, pseudopotential Hamiltonian: Theory, Implementation and Full Device Simulations," *IEEE International Electron Devices Meeting*, pp. 35.1.1–35.1.4, 2017.

[16] M. L. Van de Put, M. V. Fischetti, and W. G. Vandenberghe, "Scalable atomistic simulations of quantum electron transport using empirical pseudopotentials," *Computer Physics Communications*, vol. 244, pp. 156–169, 2019.

[17] A. AlMutairi, D. Yin, and Y. Yoon, "$PtSe_2$ field-effect transistors: New opportunities for electronic devices," *IEEE Electron Device Letters*, vol. 39, no. 1, pp. 151–154, 2018.

[18] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.*, vol. 77, pp. 3865–3868, Oct 1996.

[19] D. R. Hamann, "Optimized norm-conserving Vanderbilt pseudopotentials," *Phys. Rev. B*, vol. 88, p. 085117, Aug 2013.

[20] G. Mil'nikov, N. Mori, and Y. Kamakura, "Equivalent transport models in atomistic quantum wires," *Phys. Rev. B*, vol. 85, p. 035317, Jan 2012.

[21] M. Shin, W. J. Jeong, and J. Lee, "Density functional theory based simulations of silicon nanowire field effect transistors," *Journal of Applied Physics*, vol. 119, no. 15, p. 154505, 2016.

[22] D. Esseni, P. Palestri, and L. Selmi, *"Nanoscale MOS Transistors"*, 1st ed. Cambridge University Press., 2011.

[23] J. Chang, L. F. Register, and S. K. Banerjee, "Atomistic full-band simulations of monolayer MoS_2 transistors," *Applied Physics Letters*, vol. 103, no. 22, p. 223509, 2013.