Generation and transport of valley-polarized current in transition-metal dichalcogenides

Lei Zhang,[1] Kui Gong,[1,2] Jingzhe Chen,[1] Lei Liu,[3] Yu Zhu,[3] Di Xiao,[4] and Hong Guo[1]
[1]*Department of physics, McGill University, Montreal, H3A 2T8, Canada*
[2]*School of Materials Science and Engineering, University of Science and Technology Beijing, Beijing, 100083 China*
[3]*Nanoacademic Technologies Inc., Brossard, QC, J4Z 1A7, Canada*
[4]*Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*
(Received 21 July 2014; revised manuscript received 3 November 2014; published 19 November 2014)

In two-dimensional crystals of transition-metal dichalcogenides (TMDC) having strong spin-orbit interaction such as monolayer WSe_2, quantum states can be labeled by a valley index τ defined in the reciprocal space and the spin index s. We developed a first-principles theoretical formalism to both qualitatively and quantitatively predict nonequilibrium quantum transport of valley-polarized currents. We propose a WSe_2 TMDC transistor to selectively deliver net valley- and spin-polarized current $I_{\tau,s}$ to the source or drain by circularly polarized light under external bias. Due to the lack of translational symmetry of the real-space device, we predict a depolarization effect that increases with the decrease of the channel length of the transistor.

DOI: 10.1103/PhysRevB.90.195428 PACS number(s): 73.63.−b, 85.60.−q, 73.23.Ad

I. INTRODUCTION

The notion of valleytronics has attracted great attention due to its fundamental interest to the physics community as well as potential for future electronics [1–14]. The valley degree of freedom is of course well known for semiconductors where a valley means a local energy extrema in the Brillouin zone (BZ). The excitement about valleytronics is the idea to store and manipulate bits of information encoded in the valley quantum index, namely using *discrete* values of the crystal momentum for quantum information. Many systems have been investigated both theoretically and experimentally to realize the generation and detection of valley information, including Si [1,6,11], graphene [2–4,14], bi-layer graphene [13], carbon nanotube [8,9], bismuth [5], diamond [10] and others.

Most recently, fabrication of two dimensional (2D) transition-metal dichalcogenides (TMDC) has provided a new area of valleytronics [15–17], where *valley* and *spin* degrees of freedom are intimately coupled due to strong spin-orbit interactions (SOI). Monolayer TMDC such as MoS_2 and WSe_2 have a honeycomb lattice, hence there are two well-separated valleys labeled as K and K' in the BZ [17,18]. It was shown in both theory and experiment that a finite valley polarization can be created and detected by circularly polarized light in periodic monolayer TMDC lattices [17–20]. Such a dynamical control of valley index clearly demonstrated that the new type of electronics, valleytronics, may well be emerging.

The next immediate question is how to transfer the valley polarization once it is created. In particular, a valleytronics transistor would be interesting where valley-polarized current is turned on or off selectively. However, since the valley index is a concept living in the reciprocal space, how to generate and deliver a valley-polarized electric current to the outside world that can be received by a circuit is a challenging problem. A transistor has a source, a drain, a channel, and the existence of interfaces as well as bias voltages break translational symmetry (TS), hence it will be very interesting to investigate the generation and delivery of valley-polarized current in such a system.

Given the experimental success in TMDC phototransistors, photodetectors [21–24], and optical selectivity of the valleys [17–20], it should be possible, as we show in this work, to realize TMDC transistors to deliver a *net* valley-polarized electric current to the outside world. Interestingly and importantly, however, we find that even in the absence of any disorder, the lack of TS causes a valley depolarization effect. This effect would be absent in systems having TS. The depolarization depends on the bias voltage and the length of the transistor channel and must be taken into account in the design of valleytronic devices.

Figure 1 shows the TMDC transistor structure we consider, which is a two-probe open system. Reference [13] proposed a valley-filtering device for valley-polarized current in bilayer graphene under intense terahertz radiation in the presence of a transverse electric field. In contrast to graphene, the strong SOI in TMDC permits both spin current and valley current to be generated at the same time. Without losing generality, the transistor is made of monolayer 2D TMDC material WSe_2. The channel is controlled by a back-gate voltage V_g. In light of the optical selectivity rule [17,18], valid for periodic TMDC lattice, we shall consider a monochromatic incident light impinging on the transistor channel. The source and drain extend to $y = \mp\infty$, where a bias voltage V_{ds} is applied and current $I_{\tau,s}$ collected. The transistor is driven to nonequilibrium when $V_{ds} \neq 0$. To calculate $I_{\tau,s}$, a first-principles technique that can quantitatively determine the effects of SOI and the light—in addition to calculating quantum transport at *nonequilibrium*—is required. To our knowledge, such a technique does not yet exist and is developed in this work. We predict that even though V_{ds} breaks the TS, a net valley-spin-polarized current $I_{K,\uparrow/K',\downarrow}$ can still be excited by circularly polarized light and, importantly, can be delivered to the drain of the transistor under nonequilibrium transport conditions. This way, the quantum information associated with valley polarization is transferred from the reciprocal space to the real-space circuitry.

The rest of the paper is organized as follows. In Sec. II, the theoretical formalism is presented. In Sec. III, numerical results of valley-polarized (and spin-polarized) current are discussed. Finally, the paper is summarized in Sec. IV. In the Appendix, we present further technical details and present effective transmission coefficient $T_D(\mathbf{k} = K', E)$ at three different bias voltages.

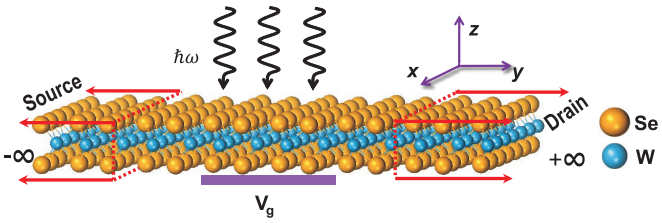

FIG. 1. (Color online) Schematic plot of the two-probe monolayer WSe_2 transistor having a source, a channel where light impinges, and a drain. The source and drain regions extend to $y = \mp\infty$. The channel can be controlled by a back-gate voltage V_g. The light has photon energy $\hbar\omega$.

II. THEORETICAL FORMALISM

Our system is described by a Hamiltonian $H = H_e + H_{e\text{-ph}}$ where H_e is calculated by a state-of-the-art first-principles method, i.e., density functional theory (DFT) carried out within the Keldysh nonequilibrium Green's function (NEGF) formalism [25–27]. The NEGF-DFT includes SOI and noncollinear spin under external voltages. Further technical details are summarized in Appendix A; we also refer interested readers to Refs. [25,26] and computation details to Ref. [28]. Having obtained H_e, we treat $H_{e\text{-ph}}$ by the first Born approximation. Here $H_{e\text{-ph}} = \frac{e}{m}\mathbf{A}\cdot\mathbf{P}$, where $\mathbf{A}$ is the electromagnetic vector potential and $\mathbf{P}$ the momentum of the electron [29]. For circularly polarized light, $\mathbf{A} = (\frac{\hbar\sqrt{\tilde{\mu}_r\tilde{\epsilon}_r}}{2N\omega\tilde{\epsilon}\epsilon}I_\omega)^{1/2}(\mathbf{e}_p b e^{-i\omega t} + \mathbf{e}_p^* b^\dagger e^{i\omega t})$. Here, I_ω is the photon flux defined as the number of photons per unit time per unit area; ω is the frequency of the light; c is the speed of light; $\tilde{\mu}_r$ is the relative magnetic susceptibility and $\tilde{\epsilon}_r$ is the relative dielectric constant; ϵ is the dielectric constant; N is the number of photons; and b and $b^\dagger$ are bosonic annihilation and creation operators. In general, $\mathbf{e}_p$ is a complex unit vector characterizing the polarization of the light. For circularly polarized light $\sigma_\pm$ along the z direction (see Fig. 1), $\mathbf{e}_p = \frac{1}{\sqrt{2}}(1,\pm i,0)$. In the Hilbert space of atomic orbital [28], the self-energy due to electron-photon interaction is [29–31]:

$$\begin{aligned}\Sigma^>_{\text{ph}} &= [NM^\dagger G_0^>(E^+)M + (N+1)MG_0^>(E^-)M^\dagger],\\ \Sigma^<_{\text{ph}} &= [NMG_0^<(E^-)M^\dagger + (N+1)M^\dagger G_0^<(E^+)M],\end{aligned} \tag{1}$$

where $E^\pm = E \pm \hbar\omega$. Matrix elements of M are defined as

$$M_{ln} \equiv \frac{e}{m}\left(\frac{\hbar\sqrt{\tilde{\mu}_r\tilde{\epsilon}_r}}{2N\omega\tilde{\epsilon}c}I_\omega\right)^{1/2}\langle l|\mathbf{P}\cdot\mathbf{e}_p|n\rangle, \tag{2}$$

where l,n label atomic orbital. In the rest of the analysis we neglect the emission process, which is the second term of Eq. (1), because the photocurrent is excited by the absorption process. Using the Keldysh equation [36], to zeroth order the lesser and greater Green's functions can be written as

$$G_0^{>/<} = G_0^r\Sigma^{>/<}G_0^a, \tag{3}$$

where we have used the lesser and greater self-energies

$$\Sigma^< = \sum_\alpha i\Gamma_\alpha f_\alpha;\quad \Sigma^> = \sum_\alpha i\Gamma_\alpha(f_\alpha - 1), \tag{4}$$

here α labels the source and drain, Γ_α is the line-width function, and f_α is the Fermi function. In Eq. (3), the retarded and advanced Green's functions are the ones without $H_{e\text{-ph}}$, namely $G_0^r = [G_0^a]^\dagger = (E + i\eta - H_e - \sum_\alpha \Sigma_\alpha^r)^{-1}$, where Σ_α^r is the retarded self-energy due to the presence of source and drain electrodes.

Finally, the source and drain current is obtained as [36]

$$I_{\alpha,\tau,s} = \frac{e}{\hbar}\int\frac{dE}{2\pi}\sum_{\mathbf{k}\in\tau}T_\alpha(E,\mathbf{k},s), \tag{5}$$

which is in terms of the effective transmission coefficient

$$T_\alpha(E,\mathbf{k},s) = \text{Tr}\{i\Gamma_\alpha(E,\mathbf{k})[(1-f_\alpha)G^<_{\text{ph}} + f_\alpha G^>_{\text{ph}}]\}_{\text{ss}}. \tag{6}$$

Here, the Green's functions including the contribution from voltage (the first term) and photons (the second term) are obtained as

$$G^{>/<}_{\text{ph}} = G_0^r(\Sigma^{>/<} + \Sigma^{>/<}_{\text{ph}})G_0^a. \tag{7}$$

Our first-principles calculation proceeds in two steps. First, for the open transistor structure under V_g, V_{ds}, we calculate H_e by NEGF-DFT[27] without photons. Second, using H_e and perturbatively treating electron-photon interaction $H_{e\text{-ph}}$ as presented, we calculate the lesser and greater Green's functions by Eq. (7) and determine the current by Eqs. (5) and (6). A further discussion of the calculation procedure for the photocurrent can be found in Appendix B.

III. VALLEY-POLARIZED (AND SPIN-POLARIZED) CURRENT

We investigate transport current $I_{\tau,s}$ along the armchair direction of WSe_2 and consider $\sigma_\pm$ circularly polarized light. The photon energy is taken to equal to the energy gap, $\hbar\omega = E_{\text{gap}} = 1.37\,\text{eV}$, of monolayer WSe_2 obtained by DFT. As shown in Fig. 2(a), the unit cell in the transport calculation is rectangular (not hexagon anymore). Therefore, the corresponding BZ is folded as shown in the right of Fig. 2(a), and in particular the K and K' points of the original infinite lattice are folded into two points along the k_x direction. A bias voltage across the source and drain, V_{ds}, is needed to deliver $I_{\tau,s}$. Figure 2(b) describes the qualitative process, namely a channel electron absorbs a photon to cross the gap and flows into the drain, leaving a hole behind; at the same time, an electron from the source fills the hole, resulting in an overall flow of current between the source and drain.

Next, we fix $V_{\text{ds}} = 0.3\text{V}$ and $V_g = 0$ to calculate the valley current. Note we choose $V_{\text{ds}} < E_{\text{gap}}$ and zero gate voltage; there is essentially no DC current, therefore any detected current in the source or drain must be generated by the light. The calculated transmission at K,K' valleys are shown in Figs. 2(c)–2(f). Let's first consider electrons excited by σ_+-polarized light, shown as dashed lines in Fig. 2 versus electron energy. The transmission coefficients $T_{S/D}(\mathbf{k} = K',E,s = \uparrow,\downarrow)$ and $T_{S/D}(\mathbf{k} = K,E,\downarrow)$ are almost zero; only $T_{S/D}(\mathbf{k} = K,E,\uparrow)$ are nonzero. Thus, σ_+ light essentially only excites electrons and holes in the K valley of the transistor channel, still obeying the optical selection rule of the periodic lattice [18]. Due to V_{ds}, these excited electrons and holes flow out of the channel to drain and source. Therefore, summing over $\mathbf{k}$ as in Eq. (5), the current in the drain carries the K-valley index. The collected K-valley current is also

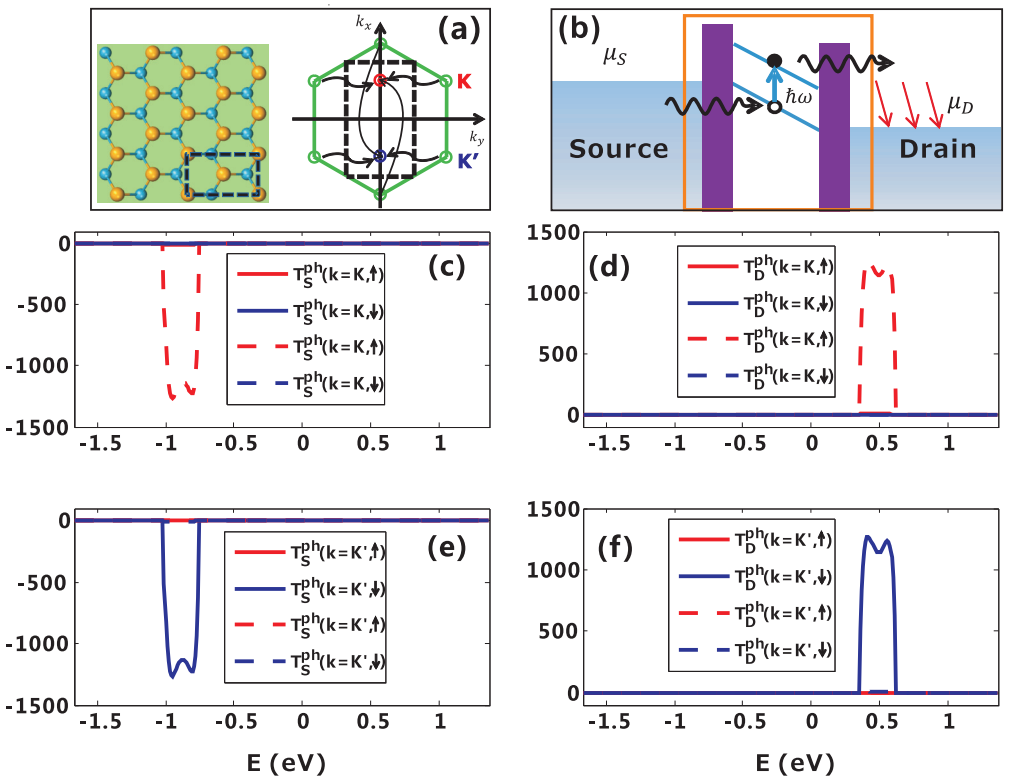

FIG. 2. (Color online) (a) Top view of monolayer WSe_2. Black dash line in left figure indicates a supercell along the armchair direction. The corresponding folded BZ is indicated by the dark dashed line in the right figure. Black curved arrows illustrate folding of the K points. (b) Schematic plot showing the qualitative physics of photocurrent. (c–f) Effective transmission $T_{S/D}(\mathbf{k},E,s)$ at K and K' points of the transistor at $V_{\rm ds}=0.3\,{\rm V}$ and $V_g=0\,{\rm V}$. Dashed and solid lines distinguish the transmission produced by σ_+ and σ_- circularly polarized incident light. The reference energy in Figs. 2(e)–2(f) is the Fermi energy of the system at equilibrium.

spin polarized due to the locked valley and spin indices of monolayer TMDC. Namely, the K-valley current is not only valley polarized but also spin polarized; one gets $I_{S/D,K,\uparrow}$ from the transistor. A similar calculation for σ_--polarized light gives currents shown in Figs. 2(e)–2(f), where the current carries the other valley index K' and spin-down electrons, $I_{S/D,K',\downarrow}$. Importantly, we found that the transmission coefficients are nonzero only when k is near the K and K' points for $\hbar\omega=E_{\rm gap}$, which means the valley indices are good quantum numbers as far as the transistor is concerned. Finally, in Appendix C, Fig. 6 compares transmission coefficient $T_D(\mathbf{k}=K',E)$ for three different values of $V_{\rm ds}$.

Having understood the transmission coefficients $T_\alpha(E,\mathbf{k})$ in $\mathbf{k}$ space, we now investigate the current $I_{\alpha,\tau,s}$ [37]. We still restrict $V_{\rm ds}<E_{\rm gap}$ so that there is no current without photons. Without losing generality, we focus on the valley current generated by σ_--polarized light; results are presented in Fig. 3. The valley current is obtained by summing the spin index s, $I_{\alpha,\tau}=I_{\alpha,\tau,\uparrow}+I_{\alpha,\tau,\downarrow}$; similarly, the spin current is obtained by summing the valley index τ. As shown in Fig. 3(a), the K valley currents $I_{S/D,K}$ are extremely small in comparison to the K' valley current $I_{S/D,K'}$, since we use σ_- light. The total current is obtained by a further summation over K,K', $I_{S/D}=I_{S/D,K}+I_{S/D,K'}$, which is plotted as black lines in Fig. 3(a). Note, we have $I_S=-I_D$ because total current must be conserved. The central result is that a *net* valley-polarized current $I^\tau_{S/D}$ (and spin current) is delivered to the source and drian:

$$I^\tau_{S/D}=(I_{S/D,K}-I_{S/D,K'})/2,$$
$$I^s_{S/D}=(I_{S/D,\uparrow}-I_{S/D,\downarrow})/2\,. \tag{8}$$

As shown in Fig. 3(b), the net valley currents are nonzero and satisfy the continuity relation $I^\tau_S=-I^\tau_D$. Since for σ_- light $I_{S/D,K}\approx 0$ as discussed above, the net valley current $I^\tau_{S/D}$ delivered by the transistor is composed of electrons having quantum index K'. Similarly, the valley current generated by σ_+ light is composed of electrons having index K. Hence, the transistor under a moderate bias (0.3 V) is efficient in operating as a valleytronic device. As discussed in the Introduction, for a pure monolayer WSe_2 lattice, the valleys K,K' are locked to spins $s=\uparrow,\downarrow$. For the transistor under bias, we found the valley-spin locking is still largely intact, $I^\tau_\alpha\approx I^s_\alpha$. Hence the valley and spin currents are simultaneously delivered by the transistor.

Valley polarization is a key quantity to characterize the valley as information carrier. At nonequilibrium ($V_{\rm ds}\neq 0$), valley polarization becomes a function of the bias and is defined by

$$\eta(V_{\rm ds})=\left|\frac{I_{\alpha,K}-I_{\alpha,K'}}{I_{\alpha,K}+I_{\alpha,K'}}\right|. \tag{9}$$

As shown in Fig. 3(c), $\eta(V_{\rm ds})$ is found to decrease with bias. The reason is that bias breaks TS and increases the intervalley

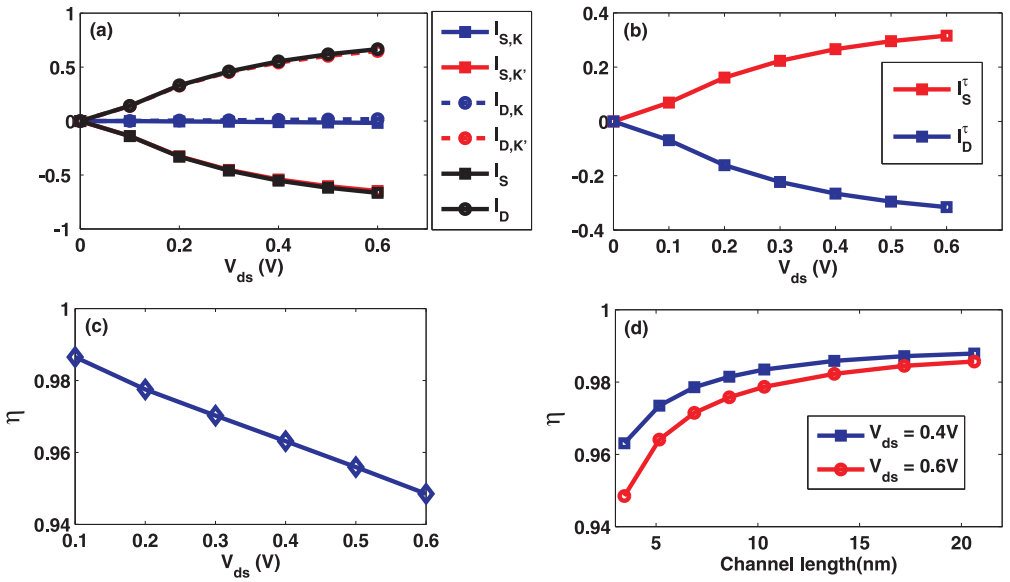

FIG. 3. (Color online) (a) Different valley components and total photocurrent in the source and drain of the transistor. (b) The net valley current in the source and drain of the phototransistor. (c) Valley polarization η versus bias voltage $V_{\rm ds}$. (d) Valley polarization η versus the channel length of the transistor. σ_- light is shed on the entire central region.

scattering (K to K') to reduce the difference between $I_{\alpha,K}$ and $I_{\alpha,K'}$. This depolarization effect was not known before because TS exists in periodic lattices. Clearly, one expects $\eta(V_{\rm ds}) \to 100\%$ when the channel length of the transistor $L \to \infty$. Fixing bias $V_{\rm ds} = 0.4$ and 0.6 V, we have calculated $\eta(V_{\rm ds})$ versus L up to $L = 20$ nm, results shown in Fig. 3(d), and a clear trend is decerned. By extrapolating the curves, we found η reaches 99.9% when L is $50 \sim 100$ nm depending on $V_{\rm ds}$. This length scale is a lower limit for the optical selection rule [17,18] to become almost perfect for the real-space transistor having no TS. Further breaking TS (e.g., by disorder etc.) is expected to enhance this length scale.

IV. SUMMARY

Valleytronics is a very interesting idea but the valley degree of freedom is a reciprocal-space concept. In this work we propose and investigate real-space *transport valleytronics*, the WSe_2 phototransistor, that selectively delivers a net valley-polarized electric current and its concomitant spin current to the outside world in the absence of translational symmetry. To this end we have developed an atomistic first-principles theory and computation technique that makes parameter-free predictions, including SOI, noncollinear spin, and electron-photon interactions. The lack of TS induces a valley depolarization effect that depends on device parameters such as the bias voltage and channel length. We estimate that for the WSe_2 phototransistor, the lower bound of the channel length by which perfect valley polarization can be achieved is about $50 \sim 100$ nm depending on the applied bias. At such a scale, the optical selection rule [18] for the k-space bands become essentially perfect. Finally, due to similarity of properties of many TMDC materials, our predicted valley-polarized current should be observable in other TMDC transistors.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support by NSERC of Canada (H.G.) and the China Scholarship Council (K.G.). We thank CalcuQuebec and Compute Canada for computation facilities.

APPENDIX A

In this Appendix, we discuss further details of the NEGF-DFT technique, the computation procedure for the photocurrent, and provide effective transmission coefficient $T_D(\mathbf{k} = K', E)$ at three different bias voltages.

1. The NEGF-DFT approach

As presented in the main text, the system is described by a Hamiltonian $H = H_e + H_{e\text{-ph}}$, where H_e is the Hamiltonian of the two-probe device without photons and is calculated by a state-of-the-art first-principles method, i.e., DFT carried out within the Keldysh NEGF formalism [25]. After H_e is obtained, the electron-photon interaction $H_{e\text{-ph}}$ is treated by the first Born approximation.

In the familiar semiclassical Boltzmann transport theory, one solves the Newtonian equation in combination with nonequilibrium statistical mechanics. In quantum transport theory, one solves the Schrödinger equation in combination with nonequilibrium statistical mechanics, and this is realized by the NEGF-DFT formalism. The original NEGF-DFT approach was developed for problems without photons, as first reported in Ref. [25]. Consider a two-terminal device shown in Fig. 4 where the two electrodes extend to electron reservoirs at infinity along the transport direction (Z direction). The system is infinitely large and has no translational symmetry due to contacts and/or externally applied bias voltages. This is also a

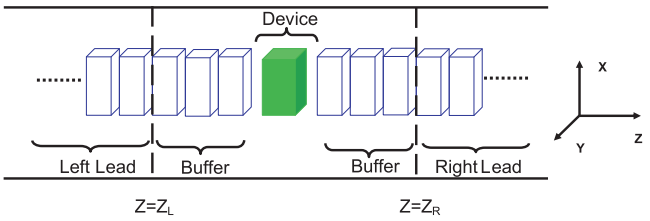

FIG. 4. (Color online) Schematic plot of a two-terminal device consisting of a central scattering region enclosed in the simulation box (between Z_L and Z_R), and left and right electrodes extending to electron reservoirs at $Z = \pm\infty$, where bias voltages are applied and electric current is collected. The scattering region includes many layers of the electrode atoms as indicated by the buffer.

nonequilibrium problem because external bias voltages drive a current flow.

In practical implementions [25], the two-terminal system is partitioned into three parts, the central region and the left and right electrodes. The central region, $Z_L < Z < Z_R$, (see Fig. 4) consists of the device itself plus left and right buffer regions, which are part of the electrodes. Outside the central region are the electrodes extending to reservoirs at $Z = \pm\infty$. At equilibrium, the electrochemical potentials of the two reservoirs (and thus the two electrodes) are equal, $\mu_L = \mu_R$. When a bias voltage V is applied, $|\mu_L - \mu_R| = eV$, where e is the electron charge, a DC current is driven through.

To self-consistently calculate H_e, which is the Hamiltonian of the two-terminal device without photons, we determine the density matrix at nonequilibrium (due to bias) by NEGF, namely the density matrix is calculated from the lesser Green's function $G^<(E)$ as

$$\rho = -i\int_{-\infty}^{\infty}\frac{dE}{2\pi}G^<(E),$$
$$G^<(E) = G^R\Sigma^<G^A, \tag{A1}$$

where the retarded Green's function and lesser self-energy $\Sigma^<(E)$ are defined as

$$G^r = \left(E + i\eta - H_e - \sum_\alpha \Sigma_\alpha^r\right)^{-1}$$
$$\Sigma^<(E) = i\sum_\alpha f_\alpha(E)\Gamma_\alpha(E), \tag{A2}$$

Here, the Fermi distribution of the αth electrode is $f_\alpha(E) \equiv f_\alpha(E - qV_\alpha)$; the line width function $\Gamma_\alpha(E) \equiv \Gamma_\alpha(E - qV_\alpha) = i[\Sigma_\alpha^r(E - qV_\alpha) - \Sigma_\alpha^a(E - qV_\alpha)]$. As shown in Eq. (A2), the electronic Hamiltonian H_e of the scattering region is needed to calculate the Green's functions. In NEGF-DFT, H_e is calculated by a self-consistent DFT.

DFT is described by the Kohn-Sham (KS) Hamiltonian [38,39] where the potential is a functional of the electronic density,

$$H_e = -\frac{\nabla^2}{2} + V_{\text{eff}}, \tag{A3}$$

where

$$V_{\text{eff}} \equiv V_{\text{ext}}(\mathbf{r}) + \int d\mathbf{r}'\frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{xc}(\mathbf{r}). \tag{A4}$$

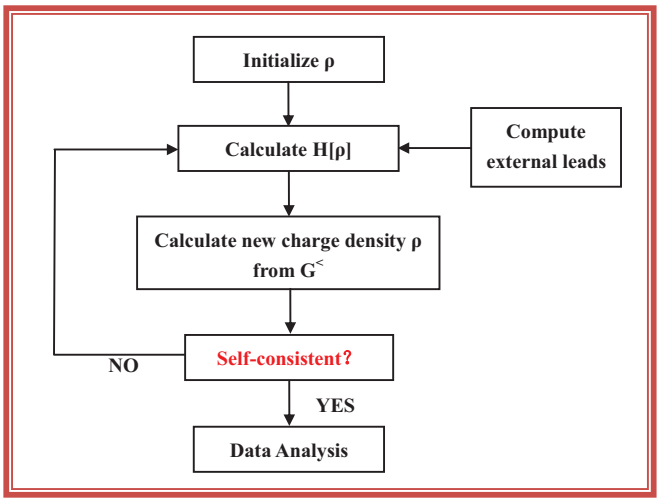

FIG. 5. (Color online) Flowchart of the NEGF-DFT self-consistent procedure.

The first term in Eq. (A4) is the external potential that includes all static background potentials, in particular the nuclear potential from all the atoms. The second term of Eq. (A4) is the Hartree potential, which can be obtained by solving a Poisson equation with proper boundary conditions. For the two-terminal structure, the boundary condition is to match the potentials at Z_L and Z_R (see Fig. 4) since the potentials of the electrodes can be calculated first. The exchange-correlation potential $V_{\text{xc}}(\mathbf{r})$ contains the nonclassical terms. While the exact functional form of V_{xc} is unknown, the local density approximation (LDA) and the generalized gradient approximation (GGA) are the popular forms of V_{xc}. For in-depth discussions of DFT, we refer interested readers to the vast literature [38,39].

In NEGF-DFT, the density matrix that enters H_e is obtained from NEGF via Eq. (A1). A practical implementation is shown in Fig. 5, showing the flowchart of the NEGF-DFT self-consistent procedure. The initial electronic density ρ is taken as that of the collection of isolated atoms forming the device. Starting from ρ, H_e is constructed using Eqs. (A3) and (A4), which is used, together with self-energies calculated from the electrodes [25], to determine the retarded, the advanced, and the lesser Green's functions. $G^<$ is then integrated over energy as in Eq. (A1) to give a new electronic density, which in turn gives a new H_e. This process is repeated until numerical convergence is achieved. This way, the Hamiltonian of the scattering region is calculated self-consistently including the external voltages and the open boundary conditions. This way, self-consistently solving the DFT-like Hamiltonian H_e within the NEGF, nonequilibrium quantum transport can be analyzed including the microscopic details of the device material.

APPENDIX B: PHOTOCURRENT

Having determined H_e, which is the electronic Hamiltonian of the two-terminal device without photons, the full Hamiltonian of our problem, $H = H_e + H_{e\text{-ph}}$, can be investigated by the first Born approximation on the electron-photon (e-ph) interaction $H_{e\text{-ph}}$.

In particular, Eq. (1) of the main text describes the self-energy due to e-ph interaction, which contains the momentum matrix Eq. (2) and zeroth-order lesser Green's function of Eq. (3) (all in the main text). The lesser or greater Green's function $G_0^{>/<}$ contains the information of external bias from lesser or greater self-energy of the electrodes in Eq. (4) of the main text. The source (and drain) current of Eq. (5) of the main text is composed of two parts: one is by external bias (e.g., when no photons) and the other by e-ph interaction, as clearly expressed from the composition of the Green's function in Eq. (7) of the main text. Namely, the first term in Eq. (7) is due to external bias and the second term describes the contribution from the e-ph interaction. If there is no light, only the first term exists and the formulation recovers the known result of DC transport [25,36].

For the semiconducting WSe_2, in the calculations we apply $V_{ds} < E_{gap}$ and $V_g = 0$, hence there is no current without photon which means the first term of Eq. (7) does not contribute due to the gap of the material.

APPENDIX C: NONEQUILIBRIUM EFFECTIVE TRANSMISSION

As shown in Fig. 3(a) of the main text, the absolute value of the photocurrent increases with bias V_{ds}. This is consistent with Eq. (5) (main text), as a larger bias means a larger integration range over the effective transmission coefficient. Quantitatively, in Fig. 6 we compare transmission curves for three V_{ds} values.

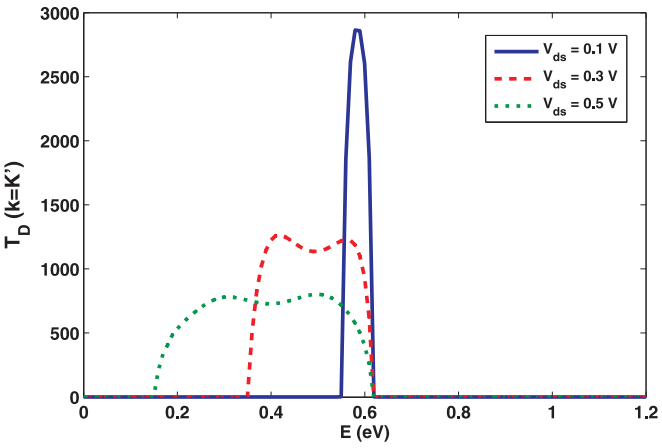

FIG. 6. (Color online) Transmission $T_D(\mathbf{k} = K', E)$ for three values of V_{ds}.

[1] K. Takashina, Y. Ono, A. Fujiwara, Y. Takahashi, and Y. Hirayama, Phys. Rev. Lett. **96**, 236801 (2006).

[2] A. Rycerz, J. Tworzydlo, and C. W. J. Beenakker, Nature Phys. **3**, 172 (2007).

[3] D. Xiao, W. Yao, and Q. Niu, Phys. Rev. Lett. **99**, 236809 (2007).

[4] D. Gunlycke and C. T. White, Phys. Rev. Lett. **106**, 136806 (2011).

[5] Z. Zhu, A. Collaudin, B. Fauqué, W. Kang, and K. Behnia, Nature Phys. **8**, 89 (2011).

[6] D. Culcer, A. L. Saraiva, B. Koiller, X. Hu, and S. DasSarma, Phys. Rev. Lett. **108**, 126804 (2012)

[7] N. Rohling and G. Burkard, New J. Phys. **14**, 083008 (2012).

[8] András Pályi and Guido Burkard, Phys. Rev. Lett. **106**, 086801 (2011).

[9] E. A. Laird, F. Pei, and L. P. Kouwenhoven, Nature Nanotechnol. **8**, 565 (2013).

[10] J. Isberg, M. Gabrysch, J. Hammersberg, S. Majdi, K. K. Kovi, and D. J. Twitchen, Nature Mater. **12**, 760 (2013).

[11] C. H. Yang, A. Rossi, R. Ruskov, N. S. Lai, F. A. Mohiyaddin, S. Lee, C. Tahan, G. Klimeck, A. Morello, and A. S. Dzurak, Nat. Commun. **4**, 2069 (2013).

[12] X. Xu, W. Yao, D. Xiao, and T. F. Heinz, Nature Phys. **10**, 343 (2014).

[13] D. S. L. Abergel and T. Chakraborty, Appl. Phys. Lett. **95**, 062107 (2009).

[14] A. R. Akhmerov and C. W. J. Beenakker, Phys. Rev. Lett. **98**, 157003 (2007).

[15] A. Splendiani, L. Sun, Y. Zhang, T. Li, J. Kim, C.-Y. Chim, G. Galli, and F. Wang, Nano Lett. **10**, 1271 (2010).

[16] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Phys. Rev. Lett. **105**, 136805 (2010).

[17] H. Zeng, J. Dai, W. Yao, D. Xiao, and X. Cui, Nat. Nanotech. **7**, 490 (2012).

[18] D. Xiao, G.-B. Liu, W. Feng, X. Xu, and W. Yao, Phys. Rev. Lett. **108**, 196802 (2012).

[19] K. F. Mak, K. He, J. Shan, and T. F. Heinz, Nat. Nanotech. **7**, 494 (2012).

[20] T. Cao, G. Wang, W. P. Han, H. Q. Ye, C. R. Zhu, J. R. Shi, Q. Niu, P. H. Tan, E. G. Wang, B. L. Liu, and J. Feng, Nat. Commun. **3**, 887 (2012).

[21] Z. Y. Yin, H. Li, H. Li, L. Jiang, Y. M. Shi, Y. H. Sun, G. Lu, Q. Zhang, X. D. Chen, and H. Zhang, ACS Nano **6**, 74 (2012).

[22] Q. H. Wang, K. K-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, Nature Nanotech. **7**, 699 (2012).

[23] H. S. Lee, S. W. Min, Y. G. Chang, M. K. Park, T. Nam, H. J. Kim, J. H. Kim, S. M. Ryu, and S. Im, Nano Lett. **12**, 3695 (2012).

[24] O. L. Sanchez, D. Lembke, M. Kayci, A. Radenovic, and A. Kis, Nature Nanotech. **8**, 497 (2013).

[25] J. Taylor, H. Guo, and J. Wang, Phys. Rev. B **63**, 245407 (2001).

[26] Derek Waldron, Paul Haney, Brian Larade, Allan MacDonald, and Hong Guo, Phys. Rev. Lett. **96**, 166804 (2006).

[27] For details of the NanoDcal quantum tranpsort package, see http://www.nanoacademic.ca.

[28] In our NEGF-DFT self-consistent calculations, a linear combination of atomic orbital basis (LCAO) at the double-ζ polarization (DZP) level is used to expand physical quantities; standard norm-conserving nonlocal pseudopotentials [32] are used to define the atomic core; the SOI is handled at the atomic level [33,34] and generalized gradient approximation (GGA)

is used for the exchange-correlation potential [35]. The WSe_2 phototransistor is periodic in the x direction, transport is along the y direction (see Fig. 1). A k mesh of 12×1 is employed for $\mathbf{k}$ sampling along the x and z directions, and the energy cutoff for the real-space grid is taken at 200 Ry. The NEGF-DFT self-consistency is deemed achieved when monitored quantities such as every element of the Hamiltonian and density matrices differ less than 10^{-5} a.u. between iteration steps. The length of the central scattering region (channel length) is 3.44 nm for Figs. 2(c)–2(f) and 3(a)–3(c); and it is increased up to 20 nm in Fig. 3(d). The light is shed in the whole central region.

[29] L. E. Henrickson, J. Appl. Phys. **91**, 6273 (2002).

[30] R. Lake and S. Datta, Phys. Rev. B **45**, 6670 (1992).

[31] J. Chen, Y. Hu, and H. Guo, Phys. Rev. B **85**, 155441 (2012).

[32] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[33] G. Theurich and N. A. Hill, Phys. Rev. B **64**, 073106 (2001).

[34] L. Fernández-Seivane, M. A. Oliveria, S. Sanvito, and J. Ferrer, J. Phys: Condens. Matter **18**, 7999 (2006).

[35] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[36] H. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors* (Springer-Verlag, New York, 1998).

[37] The calculated photocurrent has been normalized to $eF_{\mathrm{ph}}a^2\sqrt{\mu_r/\epsilon_r}$ and is dimensionless, where e is the charge of an electron, F_{ph} is the photon flux defined as number of photons per unit time per unit area, a^2 is the square of Bohr length, μ_r is relative susceptibility, and ϵ_r is relative dielectric constant.

[38] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[39] See, for example, R. G. Parr and W. T. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).