ACS APPLIED MATERIALS & INTERFACES

www.acsami.org

Research Article

Current Rerouting Improves Heat Removal in Few-Layer WSe_2 Devices

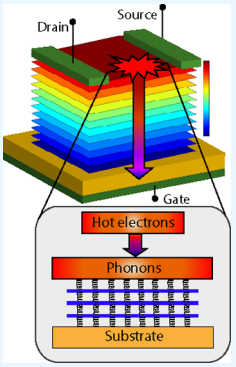
Arnab K. Majee, Zahra Hemmat, Cameron J. Foss, Amin Salehi-Khojin, and Zlatan Aksamija*

Cite This: https://dx.doi.org/10.1021/acsami.9b22039

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Few-layer (FL) transition-metal dichalcogenides have drawn attention for nanoelectronics applications due to their improved mobility, owing to the partial screening of charged impurities at the oxide interface. However, under realistic operating conditions, dissipation leads to self-heating, which is detrimental to electronic and thermal properties. We fabricated a series of FL-WSe_2 devices and measured their $I-V$ characteristics, while their temperatures were quantified by Raman thermometry and simulated from first principles. Our tightly integrated electrothermal study shows that Joule heating leads to a significant layer-dependent temperature rise, which affects mobility and alters the flow of current through the stack. This causes the temperatures in the top layers to increase dramatically, degrading their mobility and causing the current to reroute to the bottom of the FL stack where thermal conductance is higher. We discover that this current rerouting phenomenon improves heat removal because the current flows through layers closer to the substrate, limiting the severity of self-heating and its impact on carrier mobility. We also observe significant lateral heat removal via the contacts because of longer thermal healing length in the top layers and explore the optimum number of layers to maximize mobility in FL devices. Our study will impact future device designs and lead to further improvements in thermal management in van der Waals (vdW)-based devices.



KEYWORDS: *electrothermal, self-heating, few-layered TMDs, thermal boundary conductance, field-effect transistors, mobility, Raman thermometry, phonons*

1. INTRODUCTION

The on-going quest for faster switching, higher density, and better performance, predicted by Moore's law,[1] has led to persistent downscaling of nanoelectronic devices. Shrinking the device dimensions increases their surface-to-volume ratio[2] and introduces atomic-scale disorder at boundaries and interfaces,[3,4] reducing performance and threatening to limit scaling. To avoid these issues, the nanoelectronics community has turned to intrinsically two-dimensional (2D) material platforms. The ultrathin nature of 2D materials facilitates device downscaling[5] and vertical stacking,[6,7] which could extend Moore's law and enable high-density device integration for modern integrated circuits, while their atomic flatness and the absence of dangling bonds[8] prevent scattering of carriers by surface roughness (SR), which is an advantage over ultrathin three-dimensional (3D) devices.[9,10] Despite their immunity to SR scattering, single-layer (SL) transition-metal dichalcogenides (TMDs) typically exhibit degraded carrier mobility[11] relative to bulk due to strong Coulomb scattering from charged interfacial impurities.[12] A viable alternative is to use few-layer (FL) TMDs, in which the bottom layers act akin to encapsulation, screening the layers above from impurities. Due to this self-encapsulating nature, carrier mobility improves significantly in FL-TMDs,[13,14] with the highest room-temperature (RT) field-effect mobility on SiO_2 attained in FL tungsten diselenide (WSe_2).[15]

Despite their superior electrical mobilities, thermal management in devices based on FL-TMDs remains a critical issue in their development and implementation.[16,17] Heat dissipation in FL-TMDs is more complex because each layer carries a different current,[18] resulting in highly nonuniform Joule heating and causing a hotspot. Heat removal in 2D field-effect transistors (FETs) is mainly cross-plane through the substrate, owing to the small thermal healing length (a measure of heat spreading, around 100 nm)[19] and large lateral/vertical aspect ratio.[20,21] Heat dissipated by electrons is carried to the substrate via quantized lattice vibrations, or phonons, whose transmission is further limited by (i) large difference between the phase spaces of the 2D material and the 3D substrate,[22] (ii) small overlap between their vibrational densities of states,[23] and (iii) mechanical mismatch between the stiff substrate and soft out-of-plane flexural phonon modes,[24] which transfer most of the heat across the interface. Consequently, thermal boundary conductance (TBC) between the 2D material and the substrate, typically between 12 and 27 MW m^{-2} K^{-1}, is the main bottleneck for heat removal.[21,24−26] In FL-TMDs, layers

Received: December 9, 2019
Accepted: March 3, 2020
Published: March 3, 2020

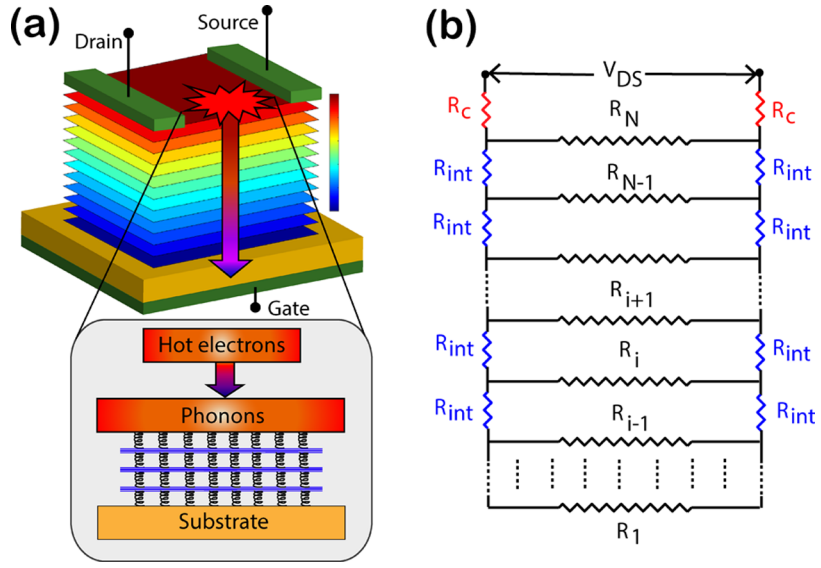


Figure 1. (a) Schematic depiction of an FL 2D FET device configuration showing nonuniform current distribution across various layers. (b) Resistance-network model used to represent the FL 2D FET.

farther from the substrate are partially insulated by layers below them and thus contribute less to TBC,[27] but measurements to date only provide us with the average temperature rise of the entire stack through Raman thermometry.[28] To fully understand heat removal in FL 2D FETs, it is imperative that we quantify the temperature rise and the impact of self-heating on the electrical performance of each layer in FL 2D devices.

Here, we show that Joule heating in FL-WSe_2 leads to significant layer-dependent temperature rise, which, in turn, alters electrical conductivity and causes the current to reroute toward the bottom of the FL stack. We study dissipation and thermal management in FL FETs built from WSe_2 and extract their layer-dependent temperature rise to elucidate its impact on device performance. We employ a resistance-network model to understand the current distribution in an FL-TMD FET and extract parameters for our model by fitting experimentally measured drain current (I_D) versus drain–source voltage (V_{DS}) characteristics of the device operating under low-field conditions. Then, we perform high-field measurements to study the effect of self-heating. We calculate the TBC of each layer from our first-principles phonon transport model[28,29] and use them to obtain the temperature rise in each layer of the device due to self-heating. We validate our results by performing Raman thermometry to measure the average temperature rise and obtain an effective TBC between FL-WSe_2 and substrate. At large bias voltages, the temperature rise in layers near the top of the stack, which carry the majority of the current, increases significantly, which deteriorates the mobility of these layers. We uncover an interesting dual behavior of the hotspot caused by Joule heating: at small drain biases, the hotspot shifts toward the top layers with increasing gate voltage, whereas at high drain biases, the temperature rise and mobility degradation cause the hotspot to reroute toward the bottom layers to minimize self-heating. We also found that, unlike SL TMDs, there is a considerable amount of heat removal through the contacts in FL-TMD FETs, even in those with long channels, due to longer thermal healing length of top layers.

2. RESULTS

2.1. Layer-Dependent Electrical Properties of FL-WSe_2. In this work, we employ a typical back-gated device configuration, shown schematically in Figure 1a. The metal source and drain contacts are deposited on the topmost layer, and the current is injected from the top through contact resistances R_c between the metal and the TMD, assuming an ohmic contact. Each layer in the stack has a different resistance R_i, where "i" refers to the ith layer in an N-layered stack, depending on its local carrier concentration, mobility, and temperature. Current encounters an additional resistance R_{int} to access the next layer. A resistor-network model represents the current flow through the FL stack, as shown in Figure 1b. The extent of current penetration through the stack would depend on the strength of R_{int} with respect to the layer resistances R_i. R_i is the product of carrier concentration n_i and mobility μ_i of that layer, $R_i = [qn_i\mu_i(T)]^{-1}(L/W)$. The total charge on the gate is $Q_{tot} = C_{ox}(V_g - V_T)$, where C_{ox} is the capacitance per unit area of the oxide. V_g and V_T are back-gate voltage and threshold voltage, respectively. However, charge screening causes the distribution of carriers across layers to be highly nonuniform. Carrier concentration is highest in the bottommost layer and decreases exponentially with each additional layer, given by the ratio[18] $n_{i+1}/n_i = \exp(-d_{FL}/\lambda_{TF})$, where λ_{TF} is the Thomas–Fermi (TF) screening length. The total charge induced in the channel is equal to the charge on the gate, $Q_{tot} = q\sum_i n_i$. Analogously, mobility in the bottommost layer (μ_1) is primarily limited by charged-impurity (CI) scattering;[30] mobility increases in the layers away from the substrate as the charged impurities are screened by layer charges, approaching the phonon-limited bulk mobility $\mu_\infty(T_i)$, which depends on the temperature of each layer. Its dependence was experimentally determined to follow a power-law $\mu_\infty(T_i) = \mu_\infty(300) \times (T_i/300)^{-\gamma}$ with $\gamma = 1.9$.[31]

To extract the three parameters R_{int}, λ_{TF}, and μ_∞ from the experiment, we mechanically exfoliated FL-WSe_2 flakes and transferred them onto a Si–SiO_2 substrate. Here, WSe_2 was selected as a representative TMD because of its high carrier mobility[32] and good stability to oxidation in ambient conditions.[33] By using an optical microscope and atomic

https://dx.doi.org/10.1021/acsami.9b22039
ACS Appl. Mater. Interfaces XXXX, XXX, XXX–XXX

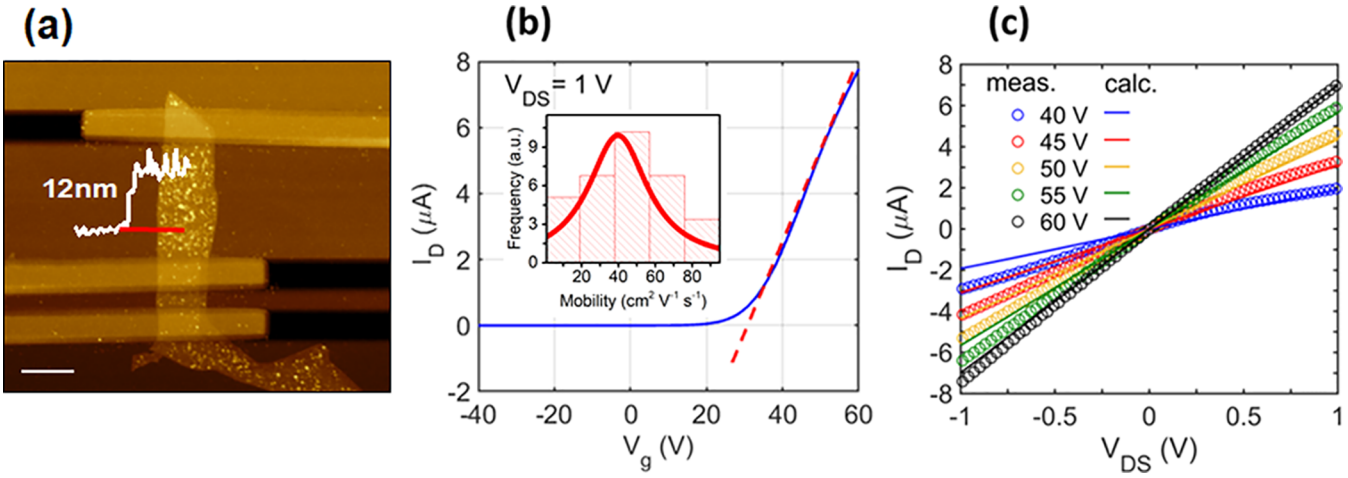


Figure 2. (a) AFM micrograph of the device. Scale bar is 2 μm. The flake thickness is 12 nm. Transfer (I_D−V_g) characteristics at V_{DS} = 1 V are shown in (b), from which V_T is extracted. The inset shows frequency distribution of mobilities for ∼20 WSe_2 devices tested at room temperature. (c) Current−voltage (I_D−V_{DS}) output characteristics obtained between two inner electrodes with different back-gate voltages (V_g) at room temperature. As indicated, I_D−V_{DS} plots are linear around zero showing reasonably good ohmic contacts for the WSe_2 channel. The channel length is 4.9 μm.

force microscopy (AFM), uniform and thin flakes were selected for subsequent experiments. Electron beam lithography and standard nanofabrication processes were employed to pattern the Au/Ag contact electrodes on selected WSe_2 flakes (see Section 4). The use of silver as the electrode metal is due to the ohmic contact between the Ag electrodes and WSe_2 flakes, as reported earlier.[28] An AFM image of a representative device with channel lengths ranging from 0.85 to 7.05 μm is shown in Figure 2a. The width of the contact electrodes is 2.35 μm. Figure 2b plots current−voltage transfer characteristics (I_D−V_g) by applying V_{DS} of 1 V between the two inner electrodes and sweeping back-gate voltage (V_g) from −60 to 60 V. The threshold voltage is extracted from the extrapolation method in the linear regime[34] of the I_D−V_g characteristics of this device and is found to be 30 V. The current−voltage (I_D−V_{DS}) output characteristics obtained between these two leads (inner electrodes) at different V_g are linear around zero (Figure 2c), implying ohmic contacts for the channel. Extrinsic electron mobility for the three different FETs was about 40−77 cm^2 V^{-1} s^{-1}, while additional fabrication and measurements on 20 devices produced the mobility distribution in the inset of Figure 2b. We vary the unknown parameters to fit our calculated I_D−V_{DS} curves to experimentally measured data, shown by the solid lines and symbols in Figure 2b, and obtain R_{int} = 5958 Ω μm, μ_∞ = 114 cm^2 V^{-1} s^{-1}, and a TF screening length λ_{TF} = 13 nm (see Section S1 in the Supporting Information for more details on λ_{TF} of WSe_2). The same parameters also fit the electrical characteristics measured between the outer electrodes (channel length of 7.05 μm) of the WSe_2 FET, which are provided in Figure S2b (see Section S2 in the Supporting Information).

2.2. Raman Thermometry. To measure the average temperature rise and extract the total TBC(G) at the FL-WSe_2/SiO_2 interface, Raman thermometry was employed during the transistor performance (electrical self-heating). Figure 3a depicts the Raman spectra of the exfoliated WSe_2 at the center of an unbiased transistor channel. Three main Raman-active peaks of high-frequency A_{1g}, E^1_{2g}, and low-frequency E^2_{2g} were observed, corresponding to the out-of-

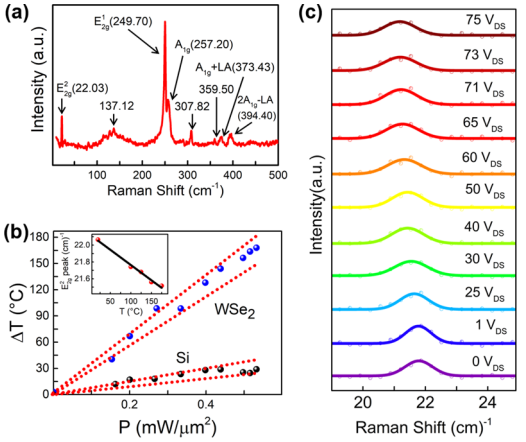


Figure 3. (a) Raman spectra at the center of an unbiased WSe_2 FET channel labeling typically observed high-frequency Raman modes of A_{1g}, E^1_{2g}, and low-frequency modes of E^2_{2g}. (b) Measured temperature rise (ΔT) of few-layer WSe_2 channel and its corresponding Si substrate versus applied electrical power. The inset shows the frequency shift of E^2_{2g} Raman mode at different stage temperature. (c) Raman peak shift of E^2_{2g} at the center of the WSe_2 channel at varying electrical bias (V_{DS}).

plane, in-plane, and interlayer vibration of atoms, respectively. Thermal characteristics of WSe_2 can be obtained from both the low-frequency and high-frequency Raman-active peaks. However, as reported earlier, two high-frequency Raman peaks of E^1_{2g} and A_{1g} can impose substantial errors in the thermometry measurements due to the peak broadening and intensity weakening. This broadening increases significantly at elevated temperatures and is primarily driven by anharmonic decay.[28] Therefore, the low-frequency E^2_{2g} shear mode was selected for accurate thermal measurements. For this purpose, Horiba's ultralow-frequency (ULF) Bragg filter was employed to probe the low-frequency Raman peak of the WSe_2 (see Section 4 for more details). For thermometry measurements, the Raman peak shift of E^2_{2g} and silicon substrate versus temperature were first calibrated. For this purpose, the laser intensity was kept low (<7.5 μW μm^{-2}, a laser spot diameter of about 1.3 μm) to

https://dx.doi.org/10.1021/acsami.9b22039
ACS Appl. Mater. Interfaces XXXX, XXX, XXX−XXX

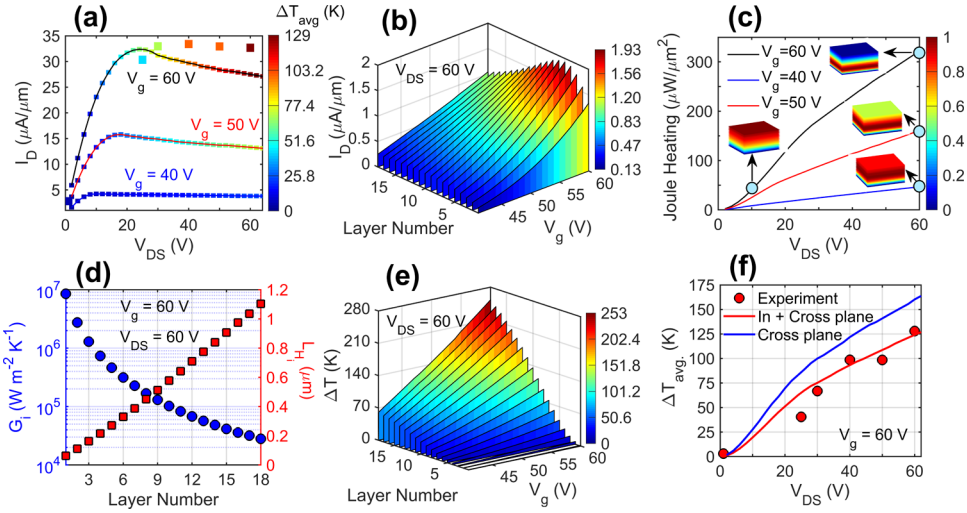


Figure 4. (a) I_D−V_{DS} output characteristics in 18-layer WSe_2 with different V_g calculated at room temperature (300 K). The color of square markers represents the rise in average device temperature ΔT_{avg} at each V_{DS} and V_g. Large square colored markers denote experimentally measured I_D, and ΔT_{avg} with V_{DS} at V_g equals 60 V. (b) Drain current per layer for different gate voltages V_g at V_{DS} = 60 V. (c) Joule heating versus V_{DS} at different V_g. The insets show Joule heating among layers at different V_{DS} and V_g. (d) Thermal boundary conductance G_i (left y-axis) and thermal healing length L_H (right y-axis) of each layer in an 18-layer WSe_2 stack. (e) Temperature rise in each layer for different V_g at V_{DS} equal to 60 V. (f) Calculated (red curve) and measured (red circles) average temperature rises of the stack versus V_{DS} at V_g equal 60 V. The blue curve represents the average temperature rise if the heat removal was entirely cross-plane.

measure the Raman peaks while the sample is placed on a hot stage and allowed to reach a thermal equilibrium condition. The inset of Figure 3b represents the frequency shift of E^2_{2g} Raman mode versus stage temperature. The temperature rise of the channel was then obtained by measuring the Raman peak shifts at different applied power, using the calibration curves. The temperature calibration and voltage-dependent Raman spectra for Si substrate are shown in Section S3 of the Supporting Information.

It is worth noting that, by employing Raman thermometry, the temperature rise of the Si substrate and WSe_2 stack can be measured simultaneously. Figure 3b shows the average temperature rise (ΔT) in the WSe_2 channel and the underlying Si substrate versus electrical power density. Also, the peak shift of E^2_{2g} Raman mode at various applied voltages (V_{DS}) obtained at the center of the operating FET channel is shown in Figure 3c, exhibiting self-heating of the channel. Knowing the thermal resistances of the SiO_2 thin film and Si substrate, G at the interface of the WSe_2 channel and the substrate (directly beneath the channel) can be calculated from an equivalent thermal circuit shown in Figure S4. The total thermal resistance (R_{TH}) and the Si thermal resistance (R_{Si}) are calculated from the slope of ΔT versus power density (P) curves for the WSe_2 channel and the Si substrate, respectively. Furthermore, since device dimensions and the thermal properties of SiO_2 are known, we can calculate the R_{SiO_2}. Details of the thermal analyses are provided in Section S4 of the Supporting Information. We found G to be 13.52 MW m^{-2} K^{-1} at room temperature for the tested WSe_2 device, in good agreement with previous measurements.[28]

2.3. Joule Heating and Layer-Resolved Temperatures. Once we extract the values of both intra- and interlayer resistances, we calculate the current distributions across all of the layers and obtain their layer-specific Joule heating $Q_i = I_i^2 R_i$. Self-heating is significant in two operating conditions: when the overdrive voltage ($V_g - V_T$) is smaller than V_{DS} and the carriers are pinched-off near the drain contact, and when $V_g - V_T \geq V_{DS}$ but the lateral field (V_{DS}/L) is large enough for the carrier drift velocity to reach the saturation region, $v_d \rightarrow v_{sat}$. The average lateral field beyond which the device is driven into velocity saturation is given by[35] $(V_g - V_T)/L \approx 2v_{sat}/\mu_{eff}$. Thus, the critical device length (L_{crit}) for velocity saturation in FL-WSe_2 is about 3.5 μm, where v_{sat} in FL-TMDs is about 3 × 10^6 cm s^{-1}.[36,37] In this work, we chose a long-channel device (L = 4.9 μm > L_{crit}) such that the carriers are not driven into velocity saturation. Moreover, beyond a critical field, the current is dominated by impact ionization which leads to a thickness-dependent avalanche breakdown in FL-TMD FETs. The maximum average electric field in our device channel is 0.122 MV cm^{-1}, which is 1 order magnitude smaller than the reported critical breakdown field for FL-MoS_2 with similar thickness.[38] Figure 4a shows the drain−source characteristics for different gate voltages at RT (300 K). With the increase in V_{DS}, there is increased Joule heating, and the average device temperature ΔT_{avg}, denoted by the color of square markers on top of the $I_D - V_{DS}$ lines in Figure 4a, rises. The large square markers represent the experimentally measured drain−source characteristics at V_g = 60 V, and the marker color denotes the measured average temperature rise of WSe_2 with respect to the substrate, which was obtained from Raman thermometry. We found that, at large V_g (=60 V), there is an increase in Joule heating due to higher carrier concentration, and the ΔT_{avg} reaches as high as 130 K. In Figure 4b, we show that the current distribution among layers is nonmonotonic for large V_{DS} = 60 V, especially at large gate bias. Due to strong TF screening and R_{int}, the top layers carry higher currents at large V_g, resulting in an increased Joule heating and temperature in these layers. This deteriorates their mobilities, and the current penetrates lower into the stack as V_g is increased. In Figure 4c, we plot Joule heating against V_{DS} for the same gate voltages as in Figure 4a. We find that the distribution of Joule heating among layers shows a distinct variation with V_g and V_{DS}, as shown by the insets in Figure 4c, where Joule heating, normalized by its maximum value in the

https://dx.doi.org/10.1021/acsami.9b22039
ACS Appl. Mater. Interfaces XXXX, XXX, XXX−XXX

stack, is schematically depicted for different gate and source−drain voltages. While for small V_g, heat is mostly concentrated in the top layers, increasing V_g causes dissipation to become concentrated in the bottom layers.

To understand these trends, we next delve into the temperatures ΔT_i in each layer. We used layer-dependent TBC values (G_i), calculated from our first-principles FL thermal transport model,[28,29] discussed further in Section 4. In Figure 4d, we plot G_i at RT and observe that most conductance is contributed by the bottom layers, with G_i decaying approximately quadratically with distance from the substrate. The total $G_{tot} = \sum_i G_i$ = 15 MW m^{-2} K^{-1}, in good agreement with Raman measurements, and the TBCs exhibit only a weak increase with T_i above RT.[24,39] The TBC of the top layer is about 2 orders of magnitude lower than the layers close to the substrate because heat has to traverse all of the weak van der Waals (vdW) bonds of the layers in series below it. The average temperature rise of each layer, calculated from Joule heating via eq 2 in Section 4 (see Section S5 of the Supporting Information for more details), is plotted for varying gate-bias conditions at V_{DS} = 60 V in Figure 4e. ΔT_i shows a strong layer dependence for large gate voltages (V_g = 60 V), indicating significant self-heating near the top of the FL stack, far exceeding the average. Although there is significant Joule heating in the bottom layers (Figure 4c), the temperatures of these layers do not increase significantly because of their higher TBC.

The thermal healing length L_H, which represents the lateral length along the channel over which the temperature increases away from the contact, also shows a prominent layer dependence, plotted on the right axis of Figure 4d. L_H of the bottom layer is about 80 nm, comparable to SL TMDs,[21] indicating that there is negligible lateral heat conduction in the bottom layers, while L_H for the top layer reaches about 1.1 μm, indicating that there is considerable heat removal through the contacts from the top layers. The ratio of power dissipation through cross-plane and in-plane is plotted against the channel length in Figure S6b; we find Q_{cross}/Q_{in} = 5 for L = 4.9 μm. Although heat removal through contacts is non-negligible in the upper layers due to their large L_H, cross-plane thermal conduction still plays a dominant role and the temperatures of those layers increase tremendously because of their poor TBC. In Figure 4f, the average device temperature rise ΔT_{avg}, shown by the solid red line, is plotted as a function of V_{DS} for V_g = 60 V. ΔT_{avg} shows a sharp rise for small V_{DS}, but for large drain−source voltage, Joule heating shifts toward the bottom layers, which efficiently conduct heat into the substrate, resulting in a smaller rise in ΔT_{avg} with V_{DS}. Our calculated average device temperatures show a good agreement with those obtained from Raman thermometry, represented by red circular markers. The solid blue line represents the average device temperature with V_{DS} if the heat removal was entirely cross-plane into the substrate. The difference stems from the lateral heat removal through the contacts, mainly from the top layers due to their large L_H.

2.4. Mobility Degradation and Current Rerouting. Earlier, we showed that I_D does not increase quadratically with V_g for a large V_{DS} (Figure 4a), indicating that carrier mobility in the device is not independent of gate bias, especially for large V_{DS}. To elaborate, we plot the intrinsic mobility of each layer for different gate biases at V_{DS} equal to 60 V in Figure 5a. For small V_g, the mobility is low for the bottom layers because of dominant CI scattering from the substrate. It increases

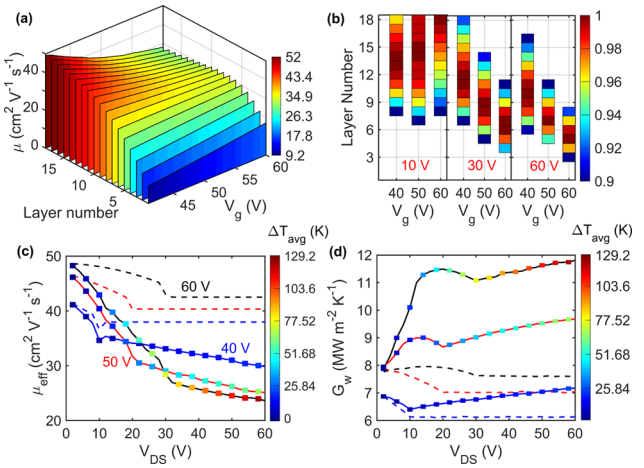


Figure 5. (a) Effect of self-heating on the mobility of each layer for different V_g. (b) Shift in the hotspot in 18-layer WSe_2 with V_g for different V_{DS} (10, 30, and 60 V). Color corresponding to each layer denotes the amount of Joule heating in that layer with respect to the maximum Joule heating. The color bar is scaled from 0.9 to 1 to show the migration of the hotspot in the stack. The effective mobility and weighted TBC of the stack are plotted against V_{DS} for different gate voltages under normal operating conditions with self-heating (solid lines) and isothermal case (dashed lines) in (c) and (d), respectively. The blue, red, and black lines represent V_g equal to 40, 50, and 60 V, respectively. The average device temperatures due to self-heating are represented by the marker colors in (c) and (d).

monotonically toward the bulk value with layer number due to charge screening from the bottom layers. For large gate-bias conditions, the temperature rises significantly for top layers due to self-heating (shown earlier in Figure 4c). Since layer mobility in eq 1 is temperature-dependent, it decreases with V_g, showing a peak in the sixth layer when V_g = 60 V. To demonstrate the effect of self-heating, we compare mobility during device operation with the isothermal case (no self-heating); see Section S6 for more details. We find that self-heating has a negligible impact on mobilities of the bottom layers because of their excellent TBC. However, mobilities of top layers exhibit severe degradation due to self-heating with a 57% reduction in the topmost layer at V_g = 60 V, shown in Figure S7.

Interestingly, we found that the migration of the hotspot with V_g exhibits different trends for low and high V_{DS}. Figure 5b shows that, at V_{DS} equal to 10 V, the hotspot shifts toward the top layers with increasing gate bias because of the increase in charge concentration in these layers, which was also found earlier in FL-MoS_2.[18] However, at large V_{DS}, the location of the hotspot shows an opposite trend—it migrates toward the bottom layers with increasing V_g, even though charge concentration increases in the top layers. This is a consequence of the mobility reduction in top layers due to self-heating, which forces the current, and hence the hotspot, to move toward the bottom layers. Figure 5b also shows that, at small V_{DS}, the layers contributing at least 90% of the peak Joule heating are spread over a wide number of layers with a weak dependence on V_g, while at large V_{DS} the spread decreases from 11 to 6 layers wide with increasing gate bias. Under these conditions, both current and hotspot become localized in a few bottom layers, typically between layers 3 and 9.

To observe the effect of hotspot migration on device performance, we calculate an effective device mobility[35] μ_{eff} =

https://dx.doi.org/10.1021/acsami.9b22039
ACS Appl. Mater. Interfaces XXXX, XXX, XXX−XXX

$L/(WC_{ox}V_{DS/Dsat})$ $(\partial I_D/\partial V_g|_{V_g})$ and plot it in Figure 5c. Relative to the isothermal cases (dashed curves), μ_{eff} of the device decreases significantly due to the average temperature rise with V_{DS} (shown by the colored markers). With increasing V_{DS}, the hotspot migrates toward the bottom layers, whose mobilities are primarily limited by CI scattering, resulting in a decrease of the effective mobility of the device. While μ_{eff} reduces from the hotspot shifting to the bottom layers with V_{DS}, the effective capacity of the FL device to remove heat, which we capture by a dissipation-weighted TBC $G_w = \sum_{i=1}^{N} Q_i G_i/\max_i(Q_i)$, increases, as shown in Figure 5d. We conclude that Joule heating increases temperature and degrades mobility, but the trend is weaker than quadratic because current reroutes into layers closer to the substrate, improving heat removal. There is a trade-off between effective mobility and TBC; we explore computationally how this trade-off varies with the thickness of the FL stack in Section S7 of the Supporting Information and find that drive current improves with increasing thickness but effective mobility saturates around 6−9 layers. For comparison, we repeat our calculations for a hypothetical material with a shorter TF screening length of 1 nm and observe improvements in both current and mobility, with a clear peak around 6 layers.

3. CONCLUSIONS

In summary, we have developed a tightly integrated experimental and numerical study of Joule heating in FL-WSe_2 FETs, allowing us to extract the heat dissipation, thermal conduction, and temperature rise in each layer of the FL stack. The combination of nonuniform heating, caused by top layers having less CI scattering and higher mobility, and layer-dependent TBC results in significant temperature rise toward the top of the FL stack, far exceeding the average. This dramatic layer dependence is driven by the approximately quadratic drop in thermal conductance with distance from the substrate, caused by heat traversing the series of relatively weak interlayer vdW bonds, and offset only partially by the increased role of lateral heat spreading from the top layers into the contacts. Ultimately, this temperature rise reduces mobilities in layers near the top of the stack, where mobility is phonon-limited and exhibits a strong temperature dependence, and causes the current to reroute into layers nearer the substrate as their TBC is higher. In contrast to isothermal and low-field operation, when layer-wise Joule heating and TBC are fully accounted for, the dissipation hotspot moves down as gate voltage, and with it carrier concentration, increases. This means mobility gains from additional layers in the stack are partially offset by self-heating. However, current rerouting improves effective heat removal as the layers closer to the substrate have higher TBCs. The optimal balance between mobility and heat removal occurs in stacks 6−9 layers thick. Short-channel FL-TMD devices could benefit somewhat from heat removal through the contacts, but a more complete solution may require a carefully designed heterostructure stack possessing both stronger screening and substrate/interlayer vdW coupling.

4. METHODS

4.1. Device Fabrication Process. Few-layer tape-exfoliated WSe_2 flakes were transferred onto a Si/SiO_2 substrate (300 nm oxide thickness). A layer of poly(methyl methacrylate) (PMMA) was spin-coated on the sample, serving as both a protective layer and a resist for electron beam lithography (EBL) process. After patterning the electrical contacts, metal electrodes were fabricated by the deposition of a 5/40 nm Ag/Au via electron beam evaporation process followed by a contact lift-off in acetone. Next, to circumvent ambient degradation, WSe_2 devices were passivated by atomic layer deposition (ALD) of about 20 nm AlO_x at 150 °C corresponding to 200 cycles of trimethylaluminum (TMA) and H_2O pulses (Ultratech/Cambridge Savannah ALD System). Before the ALD process, ∼2−4 nm of SiO_x layer was deposited as the seeding layer (Angstrom EvoVac E-Beam Evaporator system), which guarantee the material's passivation against any ambient degradation. After the Al_2O_3 deposition, another lithography process was performed to establish access to the electrical pads. Finally, AlO_x was etched from the contact-pads by Plasma-Therm ICP (inductively coupled plasma) chlorine etchant process.

4.2. Sample Characterization. Electrical characterizations were performed under a high-vacuum environment in a cryogenic probe station (chamber pressure = 10^{-8}−10^{-7} Torr). After AlO_x encapsulation, all thermometry measurements were performed in ambient air at room temperature. Electrical power was applied to the FETs during thermal measurements in a homebuilt probe setup with a Keithley 2612A System SourceMeter. The surface morphology and the thickness of the flakes were imaged using the tapping mode of the Bruker Dimension Icon atomic force microscope. SEM images were taken by using the Raith e-LiNE electron beam lithography system.

4.3. Raman Spectroscopy. Raman measurements were performed by a Horiba LabRam HR evolution confocal Raman microscope equipped with a 532 nm laser source for excitation. The diameter of the laser spot was approximately 1.3 μm with the objective of 50× at laser intensities of less than 7.5 $\mu W\ \mu m^{-2}$. With this level of laser illumination intensities, heating due to laser and photocurrent will be negligible compared to the Joule heating in the WSe_2 channel. Also, ultralow-frequency (ULF) Bragg filter of the Horiba Scientific equipment was employed to detect the low-frequency Raman peak of WSe_2.

4.4. Coupled Electrothermal Model. The mobility of individual layers is expressed as[18]

$$\mu_i(n_i, T_i) = \mu_1(n_1, n_{imp}) + [\mu_\infty(T_i) - \mu_1(n_1, n_{imp})]\left[1 - \exp\left(-\frac{(i-1)d_{FL}}{\lambda_{TF}}\right)\right] \quad (1)$$

where μ_∞ and μ_1 are the bulk and SL mobilities of WSe_2, while T_i and n_i are the temperature and carrier concentration in the ith layer, respectively. The SL mobility is calculated based on our previous work,[40] which includes acoustic, optical, and remote surface optical phonons, as well as CI scattering. Using this resistive network model, we fit our calculated I_D−V_{DS} curves at low lateral fields for different gate voltages to experimentally measured values to obtain R_{int}, λ_{TF}, and μ_∞, which are inputs to the model. To capture both lateral and cross-plane heat conduction, and their impact on the mobility of each layer, we calculate the temperature rise in each layer of the FL stack from

$$\Delta T_i = \left(\frac{Q_i}{WL}\right) \times \frac{1}{G_i}\left[1 - \frac{2L_{H_i}}{L}\tan h\left(\frac{L}{2L_{H_i}}\right)\right] \quad (2)$$

where WL is the surface area of the WSe_2 layers, and L_H is the thermal healing length. We further elaborate the derivation of this equation in the Supporting Information (Section S5). Healing length is calculated as $L_{H_i} = \kappa_{in} \times d_{SL}/G_i$, where κ_{in} is the in-plane thermal conductivity of WSe_2, which is 53 $W\ m^{-1}\ K^{-1}$.[41] Depending on the average temperature rise of each layer ΔT_i, we update the layer mobilities μ_i in eq 1 and then recalculate the layer resistances R_i. Based on the updated R_i, we recompute Q_i and ΔT_i. Using this electrothermal coupling, we set up an iterative loop, where we update ΔT_i based on the $\mu_i(n_i, T_i)$ of the previous iteration. We continue until the total Joule heating $\sum_i Q_i$ reaches convergence, typically taking between 5 and 15 iterations.

4.5. Thermal Boundary Conductance Model. We employ a combination of first-principles density functional theory and phonon

https://dx.doi.org/10.1021/acsami.9b22039
ACS Appl. Mater. Interfaces XXXX, XXX, XXX−XXX

transport simulations to calculate the thermal boundary conductance of the FL-WSe_2/SiO_2 interface. First, the phonon dispersion of monolayer WSe_2 is calculated from density functional perturbation theory (DFPT). It is then modified to account for gapping imposed by vdW coupling to the substrate and split into N sub-branches by interlayer interactions of the few-layer WSe_2 stack as $\omega_{i,\mathrm{ZA}}(q) = \sqrt{\tilde{\omega}_{\mathrm{ZA}}(q) + \omega_0^2 + \frac{2K_z}{M_{\mathrm{WSe2}}}\sin^2\left[\frac{(i-1)\pi}{2N}\right]}$, where $\tilde{\omega}_{\mathrm{ZA}}(\vec{q})$ is the phonon dispersion of the ZA branch of WSe_2 calculated from DFPT, ω_0 is the gapping of the first monolayer in the stack $\omega_0 = \sqrt{K_{\mathrm{sub}}/M_{\mathrm{Se}}}$, K_{sub} is the vdW spring coupling constant between the bottommost WSe_2 layer and the substrate, K_z (=8.24 N m^{-1}) is the interlayer spring coupling constant between adjacent layers of WSe_2, and M_{Se} is the atomic mass of Se. The modified dispersion is then input into our FL thermal boundary conductance model, where we resolve the contribution of each layer (G_i) in the FL stack to the overall TBC (G). Interfacial phonon transport is governed by two primary mechanisms: (1) flexural out-of-plane acoustic (ZA) phonons transferring energy across the interface, which depopulates the ZA phonons in the FL-WSe_2, and (2) internal scattering of phonons in the FL-WSe_2 that repopulates the ZA phonons,[29] which includes phonon−phonon scattering, phonon−boundary scattering, and phonon scattering with the AlO_x encapsulating layer. The former mechanism is an external resistance governed by a substrate scattering rate, derived from Fermi's golden rule and perturbation theory,[28] $\Gamma_{i,\mathrm{sub}}(\omega) = \frac{\pi}{2}\left[\frac{M_{\mathrm{Se}}}{K_{\mathrm{sub}}} + \frac{(i-1)M_{\mathrm{WSe_2}}}{K_z}\right]^{-2}\frac{M_{\mathrm{Se}}}{M_{\mathrm{sub}}}\frac{D_{\mathrm{sub}}(\omega)}{\omega^2}$. The rates contribute to a total $\Gamma_i(\omega)$, which is input into a Landauer model for thermal boundary conductance of each layer[23,24] $G_i(T) = \int \hbar\omega \frac{\mathrm{d}N^0(\omega,T)}{\mathrm{d}T} D_{\mathrm{2D}}(\omega)\Gamma_i(\omega)\mathrm{d}\omega$, with $N^0(\omega,T)$ being the equilibrium Bose−Einstein distribution function for phonons $N^0(\omega,T) = 1/[e^{\hbar\omega/k_BT} - 1]$. Further details of the first-principles phonon dispersion calculation and the effect of encapsulation on the TBC are described in Section S8 of the Supporting Information.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsami.9b22039.

Thomas−Fermi (TF) screening length of WSe_2; transfer and output characteristics of other devices; temperature calibration and voltage-dependent Raman spectroscopy of silicon substrate; thermal resistance circuit and TBC extraction using Raman thermometry; layer-wise temperature rise and heat dissipation through contacts and substrate; effect of self-heating on layer-wise mobilities; thickness and screening length dependence of drive current and mobility; and modeling thermal boundary conductance in few-layered 2D devices (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Zlatan Aksamija − *Department of Electrical and Computer Engineering, University of Massachusetts Amherst, Amherst, Massachusetts 01003-9292, United States;* orcid.org/0000-0001-9085-9641; Email: zlatana@engin.umass.edu

Authors

Arnab K. Majee − *Department of Electrical and Computer Engineering, University of Massachusetts Amherst, Amherst, Massachusetts 01003-9292, United States*

Zahra Hemmat − *Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, Chicago, Illinois 60607, United States;* orcid.org/0000-0002-6463-8404

Cameron J. Foss − *Department of Electrical and Computer Engineering, University of Massachusetts Amherst, Amherst, Massachusetts 01003-9292, United States*

Amin Salehi-Khojin − *Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, Chicago, Illinois 60607, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsami.9b22039

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Science Foundation Emerging Frontiers & Multidisciplinary Activities grant 1542864 and Division of Materials Research award 1902352.

■ REFERENCES

(1) Moore, G. E. Cramming more components onto integrated circuits. *Electronics* **1965**, *38*, 114−117.

(2) Novoselov, K. S.; Fal'ko, V. I.; Colombo, L.; Gellert, P. R.; Schwab, M. G.; Kim, K. A roadmap for graphene. *Nature* **2012**, *490*, 192−200.

(3) Luo, L.-B.; Yang, X.-B.; Liang, F.-X.; Xu, H.; Zhao, Y.; Xie, X.; Zhang, W.-F.; Lee, S.-T. Surface Defects-Induced p-type Conduction of Silicon Nanowires. *J. Phys. Chem. C* **2011**, *115*, 18453−18458.

(4) Kislitsyn, D. A.; Kocevski, V.; Mills, J. M.; Chiu, S.-K.; Gervasi, C. F.; Taber, B. N.; Rosenfield, A. E.; Eriksson, O.; Rusz, J.; Goforth, A. M.; Nazin, G. V. Mapping of Defects in Individual Silicon Nanocrystals Using Real-Space Spectroscopy. *J. Phys. Chem. Lett.* **2016**, *7*, 1047−1054.

(5) Peng, B.; Zheng, W.; Qin, J.; Zhang, W. Two-Dimensional MX_2 Semiconductors for Sub-5 nm Junctionless Field Effect Transistors. *Materials* **2018**, *11*, No. 430.

(6) Dean, C. R.; Young, A. F.; Meric, I.; Lee, C.; Wang, L.; Sorgenfrei, S.; Watanabe, K.; Taniguchi, T.; Kim, P.; Shepard, K. L.; Hone, J. Boron nitride substrates for high-quality graphene electronics. *Nat. Nanotechnol.* **2010**, *5*, 722−726.

(7) Yankowitz, M.; Xue, J.; Cormode, D.; Sanchez-Yamagishi, J. D.; Watanabe, K.; Taniguchi, T.; Jarillo-Herrero, P.; Jacquod, P.; LeRoy, B. J. Emergence of superlattice Dirac points in graphene on hexagonal boron nitride. *Nat. Phys.* **2012**, *8*, 382−386.

(8) Chhowalla, M.; Shin, H. S.; Eda, G.; Li, L.-J.; Loh, K. P.; Zhang, H. The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets. *Nat. Chem.* **2013**, *5*, 263−275.

(9) Prunnila, M.; Ahopelto, J.; Gamiz, F. Electron mobility in ultrathin silicon-on-insulator layers at 4.2 K. *Appl. Phys. Lett.* **2004**, *84*, 2298−2300.

(10) Cao, W.; Jiang, J.; Xie, X.; Pal, A.; Chu, J. H.; Kang, J.; Banerjee, K. 2-D Layered Materials for Next-Generation Electronics: Opportunities and Challenges. *IEEE Trans. Electron Devices* **2018**, *65*, 4109−4121.

(11) Novoselov, K. S.; Jiang, D.; Schedin, F.; Booth, T. J.; Khotkevich, V. V.; Morozov, S. V.; Geim, A. K. Two-dimensional atomic crystals. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10451−10453.

(12) Ong, Z.-Y.; Fischetti, M. V. Mobility enhancement and temperature dependence in top-gated single-layer MoS_2. *Phys. Rev. B* **2013**, *88*, No. 165316.

(13) Kim, S.; Konar, A.; Hwang, W.-S.; Lee, J. H.; Lee, J.; Yang, J.; Jung, C.; Kim, H.; Yoo, J.-B.; Choi, J.-Y.; Jin, Y. W.; Lee, S. Y.; Jena, D.; Choi, W.; Kim, K. High-mobility and low-power thin-film transistors based on multilayer MoS_2 crystals. *Nat. Commun.* **2012**, *3*, No. 1011.

(14) Das, S.; Chen, H.-Y.; Penumatcha, A. V.; Appenzeller, J. High Performance Multilayer MoS_2 Transistors with Scandium Contacts. *Nano Lett.* **2013**, *13*, 100−105.

(15) Pradhan, N. R.; Rhodes, D.; Memaran, S.; Poumirol, J. M.; Smirnov, D.; Talapatra, S.; Feng, S.; Perea-Lopez, N.; Elias, A. L.; Terrones, M.; Ajayan, P. M.; Balicas, L. Hall and field-effect mobilities in few layered p-WSe_2 field-effect transistors. *Sci. Rep.* **2015**, *5*, No. 8979.
(16) Pop, E. Energy Dissipation and Transport in Nanoscale Devices. *Nano Res.* **2010**, *3*, 147−169.
(17) Grosse, K. L.; Bae, M.-H.; Lian, F.; Pop, E.; King, W. P. Nanoscale Joule heating, Peltier cooling and current crowding at graphene-metal contacts. *Nat. Nanotechnol.* **2011**, *6*, 287−290.
(18) Das, S.; Appenzeller, J. Where Does the Current Flow in Two-Dimensional Layered Systems. *Nano Lett.* **2013**, *13*, 3396−3402.
(19) Suryavanshi, S. V.; Pop, E. S2DS: Physics-based compact model for circuit simulation of two-dimensional semiconductor devices including non-idealities. *J. Appl. Phys.* **2016**, *120*, No. 224503.
(20) Yasaei, P.; Foss, C. J.; Karis, K.; Behranginia, A.; El-Ghandour, A.; Fathizadeh, A.; Majee, A. K.; Olivares, J.; Foster, C.; Khalili-Araghi, F.; Aksamija, Z.; Salehi-Khojin, A. Interfacial Thermal Transport in Monolayer Graphene- and MoS_2-Based Devices. *Adv. Mater. Interfaces* **2017**, *4*, No. 1700334.
(21) Yalon, E.; McClellan, C. J.; Smithe, K. K. H.; Rojo, M. M.; Xu, R. L.; Suryavanshi, S. V.; Gabourie, A. J.; Neumann, C. M.; Xiong, F.; Farimani, A. B.; Pop, E. Energy Dissipation in Monolayer MoS_2 Electronics. *Nano Lett.* **2017**, *17*, 3429−3433.
(22) Ong, Z.-Y.; Cai, Y.; Zhang, G. Theory of substrate-directed heat dissipation for single-layer graphene and other two-dimensional crystals. *Phys. Rev. B* **2016**, *94*, No. 165427.
(23) Correa, G. C.; Foss, C. J.; Aksamija, Z. Interface thermal conductance of van der Waals monolayers on amorphous substrates. *Nanotechnology* **2017**, *28*, No. 135402.
(24) Foss, C. J.; Aksamija, Z. Quantifying thermal boundary conductance of 2D-3D interfaces. *2D Mater.* **2019**, *6*, No. 025019.
(25) Yasaei, P.; Behranginia, A.; Hemmat, Z.; El-Ghandour, A. I.; Foster, C. D.; Salehi-Khojin, A. Quantifying the limits of through-plane thermal dissipation in 2D-material-based systems. *2D Mater.* **2017**, *4*, No. 035027.
(26) Vaziri, S.; Yalon, E.; Rojo, M. M.; Suryavanshi, S. V.; Zhang, H.; McClellan, C. J.; Bailey, C. S.; Smithe, K. K. H.; Gabourie, A. J.; Chen, V.; Deshmukh, S.; Bendersky, L.; Davydov, A. V.; Pop, E. Ultrahigh thermal isolation across heterogeneously layered two-dimensional materials. *Sci. Adv.* **2019**, *5*, No. eaax1325.
(27) Ong, Z.-Y. Thickness-dependent Kapitza resistance in multi-layered graphene and other two-dimensional crystals. *Phys. Rev. B* **2017**, *95*, No. 155309.
(28) Behranginia, A.; Hemmat, Z.; Majee, A. K.; Foss, C. J.; Yasaei, P.; Aksamija, Z.; Salehi-Khojin, A. Power Dissipation of WSe_2 Field-Effect Transistors Probed by Low-Frequency Raman Thermometry. *ACS Appl. Mater. Interfaces* **2018**, *10*, 24892−24898.
(29) Yasaei, P.; Hemmat, Z.; Foss, C. J.; Li, S. J.; Hong, L.; Behranginia, A.; Majidi, L.; Klie, R. F.; Barsoum, M. W.; Aksamija, Z.; Salehi-Khojin, A. Enhanced Thermal Boundary Conductance in Few-Layer Ti_3C_2 MXene with Encapsulation. *Adv. Mater.* **2018**, *30*, No. 1801629.
(30) Ma, N.; Jena, D. Charge Scattering and Mobility in Atomically Thin Semiconductors. *Phys. Rev. X* **2014**, *4*, No. 011043.
(31) Xu, S.; Wu, Z.; Lu, H.; Han, Y.; Long, G.; Chen, X.; Han, T.; Ye, W.; Wu, Y.; Lin, J.; Shen, J.; Cai, Y.; He, Y.; Zhang, F.; Lortz, R.; Cheng, C.; Wang, N. Universal low-temperature Ohmic contacts for quantum transport in transition metal dichalcogenides. *2D Mater.* **2016**, *3*, No. 021007.
(32) Fang, H.; Chuang, S.; Chang, T. C.; Takei, K.; Takahashi, T.; Javey, A. High-Performance Single Layered WSe_2 p-FETs with Chemically Doped Contacts. *Nano Lett.* **2012**, *12*, 3788−3792.
(33) Liu, W.; Kang, J.; Sarkar, D.; Khatami, Y.; Jena, D.; Banerjee, K. Role of Metal Contacts in Designing High-Performance Monolayer n-Type WSe_2 Field Effect Transistors. *Nano Lett.* **2013**, *13*, 1983−1990.
(34) Ortiz-Conde, A.; Sànchez, F. J. G.; Liou, J. J.; Cerdeira, A.; Estrada, M.; Yue, Y. A review of recent MOSFET threshold voltage extraction methods. *Microelectron. Reliab.* **2002**, *42*, 583−596.
(35) Smithe, K.; English, C. D.; Suryavanshi, S. V.; Pop, E. High-Field Transport and Velocity Saturation in Synthetic Monolayer MoS_2. *Nano Lett.* **2018**, *18*, 4516−4522.
(36) Fiori, G.; Szafranek, B. N.; Iannaccone, G.; Neumaier, D. Velocity saturation in few-layer MoS_2 transistor. *Appl. Phys. Lett.* **2013**, *103*, No. 233509.
(37) Jin, Z.; Li, X.; Mullen, J. T.; Kim, K. W. Intrinsic transport properties of electrons and holes in monolayer transition-metal dichalcogenides. *Phys. Rev. B* **2014**, *90*, No. 045422.
(38) Pak, J.; Jang, Y.; Byun, J.; Cho, K.; Kim, T.-Y.; Kim, J.-K.; Choi, B. Y.; Shin, J.; Hong, Y.; Chung, S.; Lee, T. Two-Dimensional Thickness-Dependent Avalanche Breakdown Phenomena in MoS_2 Field-Effect Transistors under High Electric Fields. *ACS Nano* **2018**, *12*, 7109−7116.
(39) Yalon, E.; Aslan, O. B.; Smithe, K. K. H.; McClellan, C. J.; Suryavanshi, S. V.; Xiong, F.; Sood, A.; Neumann, C. M.; Xu, X.; Goodson, K. E.; Heinz, T. F.; Pop, E. Temperature-Dependent Thermal Boundary Conductance of Monolayer MoS_2 by Raman Thermometry. *ACS Appl. Mater. Interfaces* **2017**, *9*, 43013−43020.
(40) Behranginia, A.; Yasaei, P.; Majee, A. K.; Sangwan, V. K.; Long, F.; Foss, C. J.; Foroozan, T.; Fuladi, S.; Hantehzadeh, M. R.; Shahbazian-Yassar, R.; Hersam, M. C.; Aksamija, Z.; Salehi-Khojin, A. Direct Growth of High Mobility and Low-Noise Lateral MoS_2 Graphene Heterostructure Electronics. *Small* **2017**, *13*, No. 1604301.
(41) Gu, X.; Yang, R. Phonon transport in single-layer transition metal dichalcogenides: A first-principles study. *Appl. Phys. Lett.* **2014**, *105*, No. 131903.

https://dx.doi.org/10.1021/acsami.9b22039
ACS Appl. Mater. Interfaces XXXX, XXX, XXX−XXX