Transient hot-carrier dynamics and intrinsic velocity saturation in monolayer MoS$_2$

J. Nathawat,[1,*] K. K. H. Smithe,[2] C. D. English,[2] S. Yin,[1] R. Dixit,[1] M. Randle,[1] N. Arabchigavkani,[3] B. Barut,[3]
K. He,[1] E. Pop,[2,†] and J. P. Bird[1,3,‡]

[1]*Department of Electrical Engineering, University at Buffalo, the State University of New York, Buffalo, New York 14260-1900, USA*
[2]*Department of Electrical Engineering, Stanford University, Stanford, California 94305, USA*
[3]*Department of Physics, University at Buffalo, the State University of New York, Buffalo,*
New York 14260-1500, USA

Drift velocity saturation (at some characteristic value, v_d^{sat}) is a critical process that limits the ultimate current-carrying capacity of semiconductors at high electric fields ($\sim 10^4$ V/cm) . With the recent emergence of two-dimensional (2D) semiconductors, there is a need to understand the manner in which velocity saturation is impacted when materials are thinned to the monolayer scale. Efforts to determine v_d^{sat} are typically hampered, however, by self-heating effects that arise from undesirable energy loss from the active 2D layer to the dielectric substrate that supports it. In this work, we explore this problem for an important 2D semiconductor, namely monolayer molybdenum disulfide (MoS$_2$). By applying a strategy of rapid (nanosecond duration), single-shot, pulsing, we are able to probe the true hot-carrier dynamics in this material, free of the influence of self-heating of its SiO$_2$ substrate. Our approach allows us to realize high current densities ($\sim$mA/μm) in the MoS$_2$ layers, representing a significant enhancement over prior studies. We similarly infer values for the saturated drift velocity ($v_d^{sat} \sim 5 - 7 \times 10^6$ cms^{-1}) that are higher than those reported in earlier works, in which the influence of self-heating (and carrier injection into oxide traps) could not be excluded. In fact, our estimates for v_d^{sat} are somewhat close to the ideal velocity expected for normal (parabolic) semiconductors. Since a proper knowledge of this parameter is essential to the design of active electronic and optoelectronic devices, the insight into velocity saturation provided here should provide useful guidance for such efforts.

I. INTRODUCTION

The study of hot-carrier dynamics in semiconductors has provided the focal point for experimental and theoretical investigations for more than half a century, yielding profound impact on our fundamental understanding of transport phenomena in these materials, and wide-ranging technologies that have transformed our daily lives. With the recent emergence of two-dimensional (2D) semiconductors, prominent among which are graphene and the transition-metal dichalcogenides [1] (TMDs), there is a need to understand the manner in which hot-carrier action is modified when materials are thinned to the monolayer scale. One of the critical problems here concerns the details of drift-velocity saturation, a phenomenon that arises from the spontaneous emission of optical phonons by hot carriers at high electric fields ($\sim 10^4$ V/cm) [2,3]. This dissipative process results in a saturation of the drift velocity under high bias, at a value (v_d^{sat}) that reflects the details of the carrier and phonon dispersions in the conducting material. A proper knowledge of v_d^{sat} is essential to the design of active electronic and optoelectronic devices, making this parameter a crucial one for 2D materials.

Attempts to study the intrinsic aspects of hot-carrier action in 2D semiconductors are typically complicated by a number of extrinsic factors, which may mask the true electrical properties of these materials. Many of these problems arise from the need to support the 2D layer on an appropriate (dielectric) substrate, the introduction of which may give rise to undesirable remote- and interfacial-phonon scattering [4–6], and to deep traps that can capture energetic carriers from the ultrathin channel [7,8]. Another criticial problem is that of self-heating, in which the application of large (static) electric fields along the 2D channel results in significant current flow and Joule heating, and to a subsequent transfer of energy to the dielectric layers in contact with the channel. (This phonon-mediated mechanism should be contrasted with the dynamics of carriers in the channel itself; these very quickly equilibrate at an elevated, or "hot", temperature [2], reaching this steady state in as little as a few picoseconds.)

The various problems described above are well understood for graphene [9–15], for which it has been shown [14] that substrate-related complications may be avoided at high fields by using a strategy of rapid electrical pulsing, rather than static (DC) biasing, to excite the graphene. The key idea of this approach is to exploit the large thermal mismatch between graphene and SiO$_2$, with the much lower thermal conductivity of the latter giving rise to a long (~ 100 ns) time constant for the transfer of heat [13] from the 2D layer to its substrate. By driving the graphene with pulsed voltages that are much shorter (~ 1 ns) than this time scale, it is

_____
*jubinnat@buffalo.edu
†epop@stanford.edu
‡jbird@buffalo.edu

014002-1

therefore possible to reveal the true electrical characteristics of this material, independent of self-heating (and other extrinsic factors) [14].

While the details of velocity saturation in graphene are now fairly well understood [11–15], the situation is less clear for the TMDs. Among these, MoS_2 has been the most extensively studied, yet little consensus exists with regards to the details of its high-field behavior. In early studies of this material, various (indirect) approaches [16–18] were applied to determine v_d^{sat}, yielding estimates in the range of $\sim 1-4 \times 10^6 \, cms^{-1}$. In order to address these differing results, Smithe *et al.* [19]. have analyzed the details of current saturation in monolayer MoS_2, using temperature-dependent studies to account for the influence of self-heating. By combining their measurements with a theoretical model, for heat flow from the MoS_2 to the SiO_2/Si substrate, these authors showed how backing out the influence of self-heating results in an increase of the inferred saturation velocity, from $\sim 1.3-2.5 \times 10^6 \, cms^{-1}$ to $\sim 3.4 \times 10^6 \, cms^{-1}$.

As noted above, in the work of Ref. [19], the influence of self-heating was accounted for by combining the results of temperature-dependent experiments, with a thermal model for the Joule heating under high bias. A more ideal approach, however, would be to implement an experiment in which the influence of self-heating is strongly suppressed through deliberate design. It is this objective that we realize here, where we apply a strategy of rapid, single-shot pulsing to suppress self-heating and to study the true current-saturation characteristics of monolayer MoS_2. By varying the pulse duration in these experiments, we also identify the existence of a characteristic time scale in transient transport, beyond which the injection of hot carriers into deep oxide traps compromises the measured characteristics. As the electric field applied along the channel is varied from a few, to a few tens of, kV/cm, this characteristic time scale decreases from a few hundred, to a few tens of, nanoseconds. By ensuring that the pulse duration is kept much shorter than the lower end of this range (i.e., no more than a few nanoseconds), we are therefore able to investigate the true aspects of velocity saturation in MoS_2, free of extrinsic influences. At a carrier concentration of $7.5 \times 10^{12} \, cm^{-2}$, we determine a corresponding saturation velocity $v_d^{sat} \sim 5-7 \times 10^6 \, cms^{-1}$, values that are considerably higher than those reported previously [16–19]. In fact, our estimates are somewhat close to the ideal value ($v_d^{sat} = 1.2 \times 10^7 \, cms^{-1}$) expected [20,21], from a high-field energy-balance analysis for normal (parabolic) semiconductors. We therefore suggest that our findings may be attributed to the suppressed role of self-heating (and carrier capture) in our experiments, which reveal the true hot-carrier transport in monolayer MoS_2. It is also worth pointing out that, in Ref. [19], it was suggested that current densities in MoS_2 should approach $1 \, mA/\mu m$ in the absence of self-heating. We achieve such high current densities here, representing a significant improvement over prior work [17,19] and furthermore pointing to a suppression of self-heating in our experiments. Since a proper knowledge of v_d^{sat} is essential to the design of active electronic and optoelectronic devices based on MoS_2, the results of our study should provide useful guidance for such efforts.

II. EXPERIMENTAL METHODS

Single-crystal, monolayer MoS_2 was synthesized on SiO_2 substrates by chemical vapor deposition, using an approach that is described in more detail in Refs. [22,23]. The growth yielded triangular crystals with grain sizes exceeding $100 \, \mu m$, which were patterned with a gentle O_2 plasma to form wide ($\sim 150-200 \, \mu m$) channels for the transient measurements. These structures were then contacted in a two-step process, the first step of which involved the deposition (under high vacuum) of the low-resistance [24] Ag/Au contacts shown in Fig. 1(a). In previous work, performed on devices of channel length 5 μm, this was found to yield contacts that contribute less than 5% to the overall device resistance [19,24]. Following this, a much larger (Cr/Au) contact pads, designed to provide 50-Ω matching to the substrate, were defined by electron-beam lithography and metal liftoff. As described in our earlier work [14], realizing the impedance-matched environment that is required for these measurements necessitates the use of a high-resistivity Si/SiO_2 (0.5 mm/300 nm) substrate. This allows the metallized bottom facet of the Si wafer to be utilized as a ground plane, thereby making the design of the on chip striplines manageable. Completed samples were mounted on FR-4 boards that enabled SMA connections to be made to the external electronics; see Ref. [14] for further details. A schematic of the pulsed-measurement setup is indicated in Fig. 1(a), which shows how a trigger unit (Model 577; Berkley Nucleonics Corp.) is used to generate an isolated pulse that triggers a much faster single-shot pulse (of peak amplitude V_{in}) from a second generator (AVMP Series; Avtech Electrosystems Ltd.) After passing through the sample, the pulse is fed into the 50-Ω input of a mixed-signal oscilloscope (DSO-X 6000A Series; 6-GHz bandwidth; Keysight Technologies, Inc.), whose output voltage (V_{out}) therefore represents the time-dependent variation of the current through the device. The measurements reported here were made at room temperature, with the samples mounted inside a homemade light-tight enclosure that was maintained at a background pressure of a few mbar. A total of five different devices were measured for this study (D1–D5), with source-drain separation (i.e., channel length) of 1 μm and various widths (D1–154 μm; D2–180 μm; D3–210 μm; D4–160 μm; D5–164 μm).

III. RESULTS AND ANALYSIS

We begin our investigations by establishing the conditions under which the pulsed measurements may be performed, without suffering from the degradative influence of self-heating and hot-carrier trapping. To illustrate the issues involved here, in the left panel Fig. 1(b) we plot the output pulse detected in a series of experiments, in which we apply 4.0-V pulses of varying duration (#1–#7, in the sequence indicated) to one of our devices. As the pulse duration is initially increased from 40 ns to 1 μs (pulses #1–#3; rise time of 5 ns), the output voltage decreases by some 20%, implying a similar decrease in the corresponding current level. While the aforementioned self-heating and hot-carrier trapping are two possible sources of such a reduction [25,26], the outcome
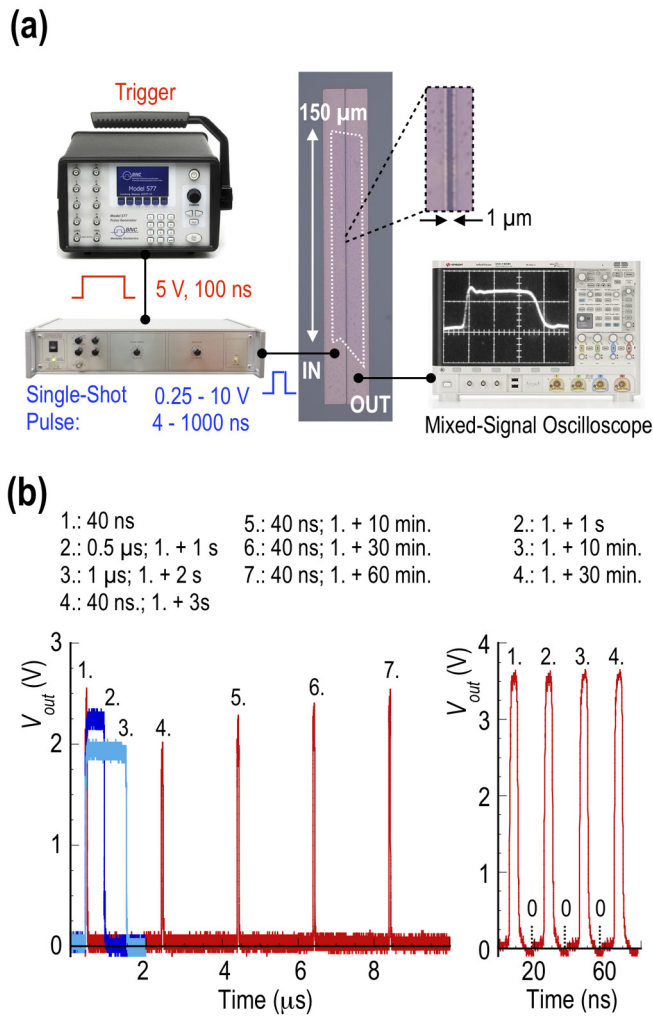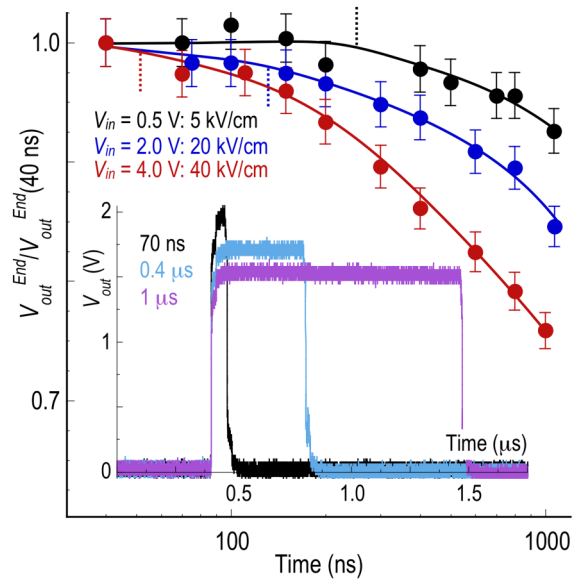
FIG. 2. The main panel summarizes the results of measurements of the output pulse amplitude as a function of pulse duration. The different color symbols represent the results of measurements obtained for input pulses of different amplitude (V_{in}, indicated in legend). The ordinate axis plots the output voltage (V_{out}^{End}) detected at the oscilloscope at the end of the pulse, normalized to this value for the shortest (40-ns) pulse measured. The dotted lines in the figure very roughly identify the crossover from a duration-independent, to a duration-dependent, pulse amplitude with increasing pulse length. The inset indicates the influence of increasing pulse duration on the output-pulse line shape. Measurements are for device D4, for an input-pulse amplitude of $V_{in} = 4.0$ V.

FIG. 1. (a) Schematic diagram, illustrating the main features of the transient measurement setup used to investigate velocity saturation in MoS_2. The trigger unit is a pulse generator that supplies a single-shot pulse to a second supply, triggering it to generate a single pulse that may be as short as 4 ns. The current associated with this pulse is then detected by a mixed-signal oscilloscope with 6 GHz bandwidth. The central portion of the figure shows an MoS_2 crystal that is contacted with Ag/Au ohmic contacts, prior to the deposition of its microstrip line. (b) The left panel plots the output voltage transient detected at the oscilloscope, for single-shot pulses of various duration ($V_{in} = 5.0$ V). Pulses are applied in the sequence 1.0–7.0. The right panel shows the result of applying four single-shot pulses, each of amplitude 5 V and length 4 ns, subject to the delays indicated in the plot. See main text for further details. Measurements are for device D5. In both panels of (b), it should be noted that the positions of the different pulses are indicated arbitrarily; the exact temporal relationship among the different pulses is indicated in the legends to the two panels.

of the experiment involving the application of pulses #4–#7 indicates that the latter mechanism is primarily at play. In this latter sequence, we apply a series of 40-ns pulses, at various times subsequent to the conclusion of pulse #3. These delays range from roughly 1 s for pulse #4, to 1 h for pulse #7. From these data it is clear that it is only for pulse #7 that the original current level, observed for pulse #1, is recovered. Clearly,

this long recovery time is inconsistent with any effect due to heating, and points instead to a mechanism in which charge trapped in deep oxide levels during the application of the pulse, is slowly released over a time scale of tens of minutes. Such behavior is, in fact, consistent with the results of our prior investigations of pulsed transport in graphene-on-SiO_2 [14].

Having identified the importance of hot-carrier capture by deep oxide traps, in Fig. 2 we explore how this process depends on the characteristics of the applied voltage pulse. The main panel summarizes the results of three different experiments, in which the pulse duration was varied for three fixed pulse amplitudes (0.5, 2.0, and 4.0 V, corresponding to black, blue, and red data, respectively). Plotted on the ordinate axis is the output voltage measured at the end of the pulse (V_{out}^{End}), normalized relative to that of the shortest (the 40-ns) pulse. (The manner in which the output wave form varies with pulse duration is indicated in the inset to the figure.) Starting with the behavior exhibited for a pulse amplitude of 0.5 V, the data in the main panel indicated that the output pulse amplitude remains essentially unchanged up to a pulse duration that is estimated to lie (as identified by the black dotted line) somewhere in the range of 200–300 ns. Turning to the (blue and red) data exhibited for the larger input pulses, it is clear that a similar effect is observed; as the pulse duration is initially increased, the output amplitude remains almost unchanged, before a crossover to a new regime occurs

where the output voltage becomes strongly pulse duration dependent. As the pulse amplitude is increased, it is apparent from the blue and red dotted lines that the crossover occurs at ever shorter durations, with the threshold corresponding to just a few tens of ns for the 4.0-V input.

The data of Fig. 2 provide important insight as to how to mitigate the influence of hot-carrier capture by deep traps on our pulsed measurements. Essentially, they indicate that this nonvolatile function does not arise instantaneously, but rather requires a minimum characteristic time that is dependent upon the pulse amplitude. In a sense, this is much like the influence of self-heating noted already, although we note that the carrier capture can occur on a significantly quicker time (just a few tens of ns) when the applied pulse is sufficiently large. In order to avoid the influence of such effects, it is crucial to make use of extremely short pulses, and in the discussion that follows in this paper we therefore present the results of experiments obtained for a pulse duration of just 4 ns. In the right panel of Fig. 1(b), we demonstrate the absence of charge-memory effects in a series of such measurements performed over a time interval of half an hour. It is absolutely clear that these data are not compromised in any way by the influence of the oxide traps.

Having established the appropriate protocol for the pulsed measurements, we now move to summarize the results of our investigations of the intrinsic aspects of high-field transport in monolayer MoS_2. Pulsed current-voltage characteristics of the five different devices are presented in Fig. 3(a), in which the data are plotted to show the variation of current density vs electric field. The various devices show a consistent trend, with the onset of current saturation apparent for electric fields approaching 100 kV/cm. At these fields the current reaches as much as 0.7 mA/μm, corresponding to a total current of 126 mA. These levels are much higher than those reported in earlier studies [17,27–31], an observation that we attribute to two key factors. The first of these is the low contact resistance achieved through the use of the Ag/Au contacts to our devices [24], which minimizes their total, two-terminal resistance. Indeed, closer inspection of the pulse waveforms in Figs. 1 and 2 indicates that this resistance was typically on the order of several tens of ohms, a feature that makes the devices also well matched to the 50-Ω pulsed environment. The second important factor is the use of our pulsed measurement scheme, which, as we have noted already, significantly suppresses the influence of the extrinsic processes of self-heating and hot-carrier trapping.

For an estimate of the drift velocity from the data of Fig. 3(a), knowledge of the carrier concentration in the MoS_2 layer is required. To determine this quantity, we have used similar MoS_2 crystals, synthesized on the same wafer as those utilized in our pulsed measurements, and have patterned these into Hall bars [see the upper inset of Fig. 3(a)] by a combination of electron-beam lithography and oxygen-plasma etching. A representative Hall measurement is indicated in the lower inset to Fig. 3(a), from which we infer a carrier concentration of $7.5 \pm 0.5 \times 10^{12}$ cm^{-2}. A similar value was obtained in measurements of a second Hall device, and is consistent with the findings of a separate study performed using the same material [19]. Using this value for the carrier concentration, we determine the velocity-field characteristics
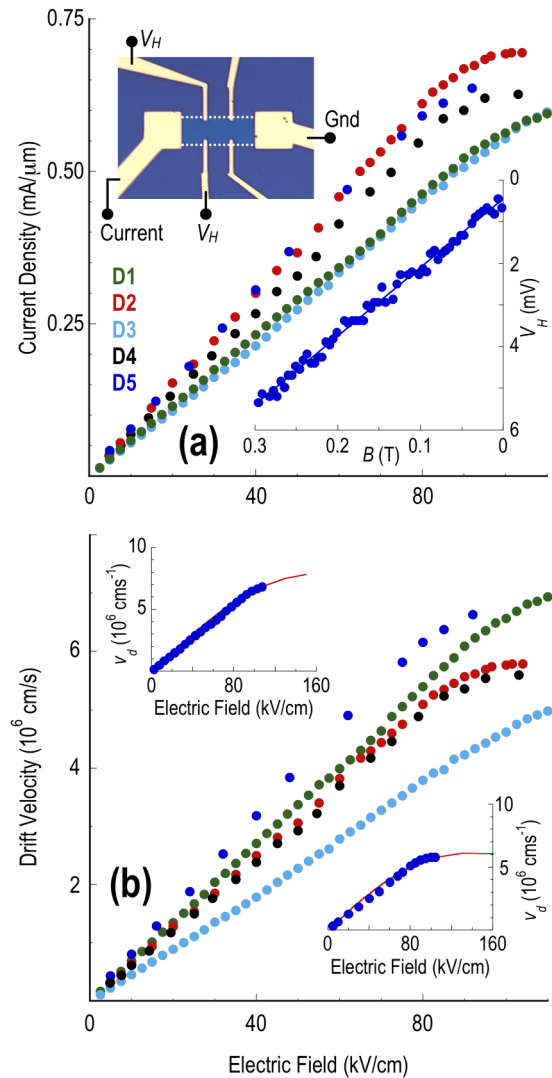

FIG. 3. (a) The main panel shows the current-field characteristics of the different MoS_2 crystals, measured using single-shot pulses 4-ns long. The upper inset is an optical micrograph of one of the fabricated Hall bars, in which the separation of the voltage probes along the same edge is 6 μm. The lower inset shows the measured Hall voltage (V_H) for one of the Hall bars as a function of magnetic field, allowing the electron concentration to be determined. (b) Variation of drift velocity with electric field, determined from the data of panel (a) (symbols correspond to those in that panel), and from measurements of the carrier concentration via the Hall effect. The upper and lower insets show the variation of drift velocity for D1 and D2, respectively, along with fits (solid lines) to the form of Eq. (1).

of our devices as demonstrated in Fig. 3(b). From this figure, it is clear that the drift velocity saturates at a value of around $5-7 \times 10^6$ cms^{-1}.

IV. DISCUSSION

Recently, Ferry has calculated the characteristics of hot carriers in monolayer MoS_2 using an ensemble Monte Carlo approach that includes all relevant phonon processes, the transfer of electrons to the conduction-band side valleys,

and the influence of surface polar optical phonons of a SiO_2 substrate [2]. In this way he predicts a saturated velocity of $5 - 6 \times 10^6 \, \mathrm{cms^{-1}}$ at a carrier concentration of $5 \times 10^{11} \, \mathrm{cm^{-2}}$, a value that is consistent with our measurements. We also note that our estimate for v_d^{sat} is within the range predicted in the earlier theoretical study of Ref. [32], which utilized a density functional theory calculation combined with a full-band Monte Carlo analysis.

As for related experiments on MoS_2, while the number is relatively small they have systematically yielded lower values for the saturation velocity than we obtain here. Fiori *et al.* undertook an extensive investigation of velocity saturation in multilayer devices, using a DC-biasing scheme [16], and found a saturation velocity of $\sim 3 \times 10^6 \, \mathrm{cm/s}$ at room temperature. This decreased by around an order of magnitude on raising the ambient temperature to 500 K, demonstrating the impact that self-heating may separately exert when induced by DC biasing. In work by He *et al.* [17], we extended the studies of Fiori *et al.* to monolayer MoS_2, measuring its transistor curves at various temperatures and subject to similar DC biasing. Through a combination of experiment and modeling (using the SILVACO Atlas package), we inferred a saturation velocity of $3 \times 10^6 \, \mathrm{cm/s}$, consistent with the results of Fiori *et al.* Without doubt, self-heating arising from the DC biasing would have exerted an influence on these measurements. Finally, in a most-recent study by Smithe *et al.*, we demonstrated very clearly how reduction of the substrate temperature may be used to suppress (but not eliminate) the role of heating [19]. For ambient temperature (300 K), however, saturation velocities were again in the range of $3 - 4 \times 10^6 \, \mathrm{cm/s}$, with thermal modeling suggesting that self-heating raises the substrate temperature by as much as 200 K during high-field DC measurements. It therefore seems clear that the much higher saturation velocities that we determine here should arise from the single-shot pulsing strategy that we apply; this allows us to generate fields well in excess of 100 kV/cm, while dissipating less than a nJ of energy per applied pulse (an average power some 8–9 orders of magnitude smaller than that dissipated in DC studies).

In Ref. [19] we used the Caughey-Thomas model [33] to describe the variation of drift velocity (v_d) as a function of electric field (F):

$$v_d(F) = \frac{\mu_{LF} F}{\left[1 + \left(\frac{\mu_{LF} F}{v_d^{\mathrm{sat}}} \right)^\gamma \right]^{1/\gamma}}, \tag{1}$$

where μ_{LF} is the low-field mobility, and γ is an empirical fitting parameter. In the upper and lower insets to Fig. 3(b), we show two examples where have fitted the observed variation of drift velocity to this form. The fits use μ_{LF}, v_d^{sat} and γ as free parameters, with μ_{LF} being determined by a fit to the linear variation of the drift velocity near zero field. The plots shown in the insets indicate that this model is able to capture the variations observed in experiment, and a summary of the different fitting parameters is provided in Table I. The

TABLE I. Parameters for fitting to the Caughey-Thomas model.

	D1	D2	D3	D4	D5
$\boldsymbol{\mu_{LF}}$ (cm²/Vs)	62	64	46	59	78
$\boldsymbol{v_d^{sat}}$ (10⁶ cm/s)	6.1	4.9	5.6	7.2	6.1
$\boldsymbol{\gamma}$	2.6	2.5	2.1	2.1	2.0

mobility values (μ_{LF}) indicated here are around a factor of two larger than those reported in Ref. [19]; this presumably also reflects the reduced influence of charge trapping and self-heating in the pulsed measurements. As for the value of the free parameter (γ), which ranges from 2.0 to 2.6, this is lower than that utilized in Ref. [19], where the need for values of this parameter approaching as much as five was attributed to the influence of self-heating. As we have noted several times already, this should not be an issue in our measurements.

Finally, it is worth mentioning that Taniguchi *et al.* [34]. recently used an approach based on gate pulsing to investigate the interfacial trap density for monolayer MoS_2 supported on quartz substrates. They revealed that the trapping of carriers by interfacial states (when the pulse is applied) can be much faster than the reverse one of emission (when the pulse is removed). Their observations were made on a time scale of a tens to hundreds of microseconds, however, and so did not address the issue of much longer, nonvolatile, oxide trapping discussed here.

V. CONCLUSIONS

In conclusion, we have investigated the details of high-field drift velocity saturation in monolayer MoS_2, using a strategy of rapid (nanosecond-duration) electrical pulsing to probe hot-carrier transport in this material, free of the influence of self-heating. Our approach yields values for the saturated drift velocity ($v_d^{\mathrm{sat}} \sim 5 - 7 \times 10^6 \, \mathrm{cms^{-1}}$) that are considerably higher than those reported previously [16–19], and close to the ideal value expected [20,21] for normal (parabolic) semiconductors. We furthermore observed current densities as high as 0.7 mA/μm, again representing an improvement over prior work [17,19]. Our studies therefore highlight how strategies of rapid pulsing can be used to manage heating effects in 2D materials, when they are subject to high electric fields. Since a proper knowledge of v_d^{sat} is essential to the design of active electronic and optoelectronic devices based on MoS_2, the results of our study will also be important for such efforts.

[1] Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, Nat. Nanotechnol. **7**, 699 (2012).

[2] D. K. Ferry, Semicond. Sci. Technol. **32**, 085003 (2017).

[3] R. S. Shishir and D. K. Ferry, J. Phys.: Condens. Matter **21**, 344201 (2009).

[4] J.-H. Chen, C. Jang, S. Xiao, M. Ishigami, and M. S. Fuhrer, Nat. Nanotechnol. **3**, 206 (2008).

[5] S. Fratini and F. Guinea, Phys. Rev. B **77**, 195415 (2008).

[6] A. M. DaSilva, K. Zou, J. K. Jain, and J. Zhu, Phys. Rev. Lett. **104**, 236601 (2010).

[7] Y. G. Lee, C. G. Kang, U. J. Jung, J. J. Kim, H. J. Hwang, H. J. Chung, S. Seo, R. Choi, and B. H. Lee, Appl. Phys. Lett. **98**, 183508 (2011).

[8] K. Majumdar, S. Kallatt, and N. Bhat, Appl. Phys. Lett. **101**, 123505 (2012).

[9] I. Meric, M. Y. Han, A. F. Young, B. Ozyilmaz, P. Kim, and K. L. Shepard, Nat. Nanotechnol. **3**, 654 (2008).

[10] A. Barreiro, M. Lazzeri, J. Moser, F. Mauri, and A. Bachtold, Phys. Rev. Lett. **103**, 076601 (2009).

[11] V. E. Dorgan, M.-H. Bae, and E. Pop, Appl. Phys. Lett. **97**, 082112 (2010).

[12] I. Meric, C. R. Dean, A. F. Young, N. Baklitskaya, N. J. Tremblay, C. Nuckolls, P. Kim, and K. L. Shepard, Nano Lett. **11**, 1093 (2011).

[13] S. Islam, Z. Li, V. E. Dorgan, M.-H. Bae, and E. Pop, IEEE Electron Dev. Lett. **34**, 166 (2013).

[14] H. Ramamoorthy, R. Somphonsane, J. Radice, G. He, C.-P. Kwan, and J. P. Bird, Nano Lett. **16**, 399 (2015).

[15] M. A. Yamoah, W. Yang, E. Pop, and D. Goldhaber-Gordon, ACS Nano **11**, 9914 (2017).

[16] G. Fiori, B. N. Szafranek, G. Iannaccone, and D. Neumaier, Appl. Phys. Lett. **103**, 233509 (2013).

[17] G. He, K. Ghosh, U. Singisetti, H. Ramamoorthy, R. Somphonsane, G. Bohra, M. Matsunaga, A. Higuchi, N. Aoki, S. Najmaei, Y. Gong, X. Zhang, R. Vajtai, P. M. Ajayan, and J. P. Bird, Nano Lett. **15**, 5052 (2015).

[18] A. Sanne, R. Ghosh, A. Rai, M. N. Yogeesh, S. H. Shin, A. Sharma, K. Jarvis, L. Mathew, R. Rao, D. Akinwande, and S. Banerjee, Nano Lett. **15**, 5039 (2015).

[19] K. K. H. Smithe, C. D. English, S. V. Suryavanshi, and E. Pop, Nano Lett. **18**, 4516 (2018).

[20] In such semiconductors, an energy-balance analysis yields a saturation velocity $v_d^{\mathrm{sat}} = \sqrt{\hbar\omega_{OP}/m^*}$, where $\hbar\omega_{OP}$ is the optical-phonon energy and m^* is the effective mass of the relevant carriers. For details, see, P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors: Physics and Materials Properties*, 4th ed. (Springer-Verlag, Berlin, 2010), pp. 227–228.

[21] For monolayer MoS$_2$, $\hbar\omega_{OP} = \sim 37$ meV and $m^* = 0.48m_0$. See, for example, K. Kaasbjerg, K. S. Thygesen, and K. W. Jacobsen, Phys. Rev. B **85**, 115317 (2013).

[22] K. K. H. Smithe, C. D. English, S. V. Suryavanshi, and E. Pop, 2D Mater. **4**, 011009 (2017).

[23] K. K. H. Smithe, S. V. Suryavanshi, M. Muñoz Rojo, A. D. Tedjarati, and E. Pop, ACS Nano **11**, 8456 (2017).

[24] C. D. English, G. Shine, V. E. Dorgan, K. C. Saraswat, and E. Pop, Nano Lett. **16**, 3824 (2016).

[25] J. Nathawat, M. Zhao, C.-P. Kwan, S. Yin, N. Arabchigavkani, M. Randle, H. Ramamoorthy, G. He, R. Somphonsane, N. Matsumoto, K. Sakanashi, M. Kida, N. Aoki, Z. Jin, Y. Kim, G.-H. Kim, K. Watanabe, T. Taniguchi, and J. P. Bird, ACS Omega **4**, 4082 (2019).

[26] H. Ramamoorthy, R. Somphonsane, J. Radice, G. He, J. Nathawat, C.-P. Kwan, M. Zhao, and J. P. Bird, Semicond. Sci. Technol. **32**, 084005 (2017).

[27] S. Kim, A. Konar, W.-S. Hwang, J. H. Lee, J. Lee, J. Yang, C. Jung, H. Kim, J.-B. Yoo, J.-Y. Choi, Y. W. Jin, S. Y. Lee, D. Jena, W. Choi, and K. Kim, Nat. Commun. **3**, 1011 (2012).

[28] M. Amani, M. L. Chin, A. G. Birdwell, T. P. O'Regan, S. Najmaei, Z. Liu, and P. M. Ajayan, Appl. Phys. Lett. **102**, 193107 (2013).

[29] H.-Y. Chang, W. Zhu, and D. Akinwande, Appl. Phys. Lett. **104**, 113504 (2014).

[30] T. Roy, M. Tosun, J. S. Kang, A. B. Sachid, S. B. Desai, M. Hettick, C. C. Hu, and A. Javey, ACS Nano **8**, 6259 (2014).

[31] S. Das and J. Appenzeller, Phys. Status Solidi RRL **7**, 268 (2013).

[32] X. Li, J. T. Mullen, Z. Jin, K. M. Borysenko, M. Buongiorno Nardelli, and K. W. Kim, Phys. Rev. B **87**, 115418 (2013).

[33] D. M. Caughey and R. E. Thomas, Proc. IEEE **55**, 2192 (1967).

[34] K. Taniguchi, N. Fang, and K. Nagashio, Appl. Phys. Lett. **113**, 133505 (2018).