Design and Process Co-Optimization of 2-D Monolayer Transistors via Machine Learning

Chin-Cheng Chiang, Hao-Yu Lan, *Graduate Student Member, IEEE*, Lina Liu, Yong P. Chen, Dmitry Zemlyanov, Joerg Appenzeller, *Fellow, IEEE*, and Zhihong Chen, *Fellow, IEEE*

***Abstract*— Machine learning (ML) algorithms have been widely adopted in the industry for optimizing semiconductor process modules. However, the final device performance may depend on a myriad of variables, including process conditions and device design parameters in an entangled fashion. This makes gaining detailed physical insights necessary to disentangle all these factors effectively. Here, we propose a design and process co-optimization framework using ML to improve the performance of 2-D transistors in analogy to conventional process optimization-design of experiments (DOEs). In particular, we utilize the "feature importance score" to quantitatively evaluate the impact of each process step or design feature on the final device performance. Example given, through the utilization of a random forest ML algorithm, we can design distinct threshold voltages by combining suitable process and design parameters. This framework aims at unrevealing the ultimate performance of 2-D field-effect transistors (FETs) through an expedited process that allows for quick experimental turnaround.**

***Index Terms*— 2-D semiconductors, charge trapping, dielectric interface, doping, monolayer transition metal dichalcogenides (1L-TMDs), transistors.**

I. Introduction

AS SILICON (Si) metal–oxide–semiconductor field-effect transistors (MOSFETs) approach the ∼5-nm gate length range, undesirable short channel effects, including, but not limited to, source-to-drain tunneling, drain-induced barrier lowering (DIBL), and the loss of electrostatic gate control severely degrade the device performance, thus limiting the scaling of Si MOSFETs [1]. On the other hand, 2-D semiconducting transition metal dichalcogenides (2D TMDs) are highly promising as a next-generation channel material for ultrascaled field-effect transistors (FETs), as their atomically thin nature ensures excellent electrostatic control of the channel potential [2], [3], [4], [5], [6], [7], [8]. Recent progress has been made on both n-type [9], [10], [11], [12], [13], [14] and p-type [15], [16], [17], [18], [19], [20] TMDs for realizing high-performance FETs exhibiting hundreds of μA/μm current densities at a *bias* of 1 V. Despite steady progress over the last decade, a substantial amount of open questions regarding the ultimate performance of TMD transistors remain, many of them related to the fact that these devices are "interface-only" transistors and are much more sensitive to the surrounding environment, fabrication processes, and treatments. Moreover, as apparent from advanced semiconductor manufacturing process flows that involve thousands of processes, it has become a more challenging and time-consuming task to optimize device performance metrics, merely relying on human judgment.

To address these problems, Chen et al. [21] adopted machine learning (ML) to evaluate and co-optimize key process parameters that impact the final TMD device performance. However, the parameters that were selected to evaluate were rather limited in scope, while it is more and more clear that a much larger set of "entangled" parameters needs to be considered, which should include detailed process conditions and device design parameters. In particular, optimizing the channel geometry plays a significant role in the overall device performance of nanosheet transistors (NSFETs) [22]. For example, the nanosheet width can have a significant impact on the carrier mobility and transconductance, due to edge conduction differing from the one in the middle of the channel, and the nanosheet thickness is important for electrostatic control and quantum confinement in the 2-D channels, both leading to complex dependencies on the device performance. Here, we have implemented a design and process co-optimization framework using ML that involves 16 independently controlled parameters and more than 1000 device characteristics to identify the desired parameter space for optimum monolayer (1L) TMD transistor performance. The total processing time from the initial electrical data extraction, the ML model training, to the final prediction only takes a minute, which greatly reduces the turnaround time for predicting the desired set of parameters to develop optimized TMD transistors.

Manuscript received 15 May 2023; revised 3 August 2023 and 26 August 2023; accepted 29 August 2023. Date of publication 12 September 2023; date of current version 24 October 2023. This work was supported in part by the Semiconductor Research Corporation (SRC) and National Institute of Standards and Technology (NIST) through the NEW LIMITS Center under Award 70NANB17H041. The review of this article was arranged by Editor J. Xu. *(Corresponding author: Zhihong Chen.)*

Chin-Cheng Chiang, Hao-Yu Lan, Joerg Appenzeller, and Zhihong Chen are with the Elmore Family School of Electrical and Computer Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907 USA (e-mail: lan36@purdue.edu; zhchen@purdue.edu).

Lina Liu and Yong P. Chen are with the Department of Physics and Astronomy and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907 USA.

Dmitry Zemlyanov is with Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TED.2023.3310942.

Digital Object Identifier 10.1109/TED.2023.3310942

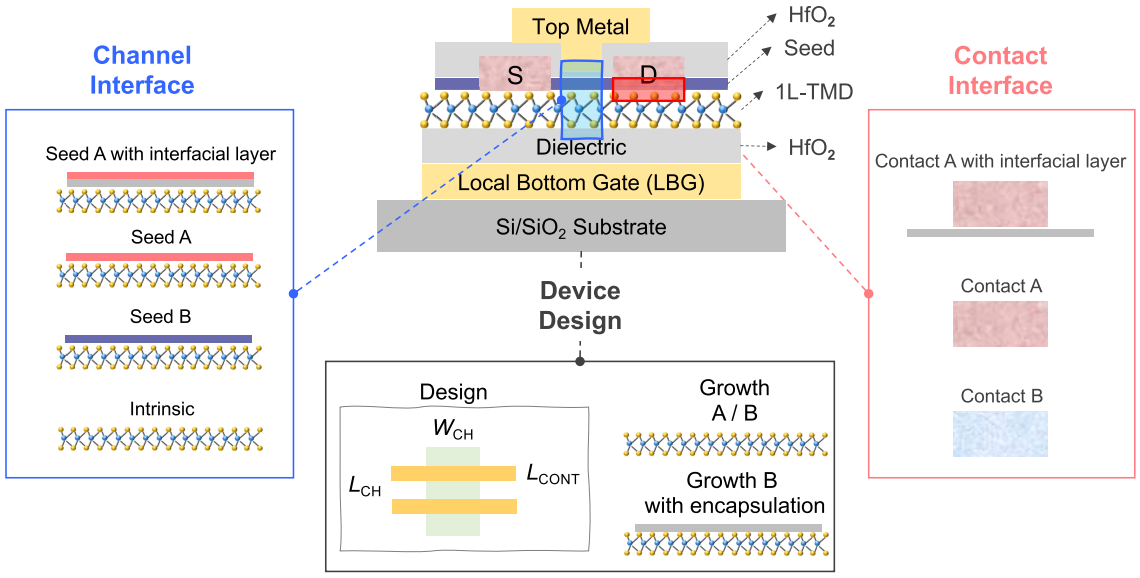


Fig. 1. Features determining the electrical characteristics of 2-D 1L TMD transistors schematic cross section of a 1L-TMD FET with a LBG and a top gate. Various components that govern the device performance are categorized, including the channel region where the top gate-stack is formed, FET device design parameters (L_{CH}: channel length, W_{CH}: channel width, and L_{CONT}: contact length), and the interface between TMD and contact metals.

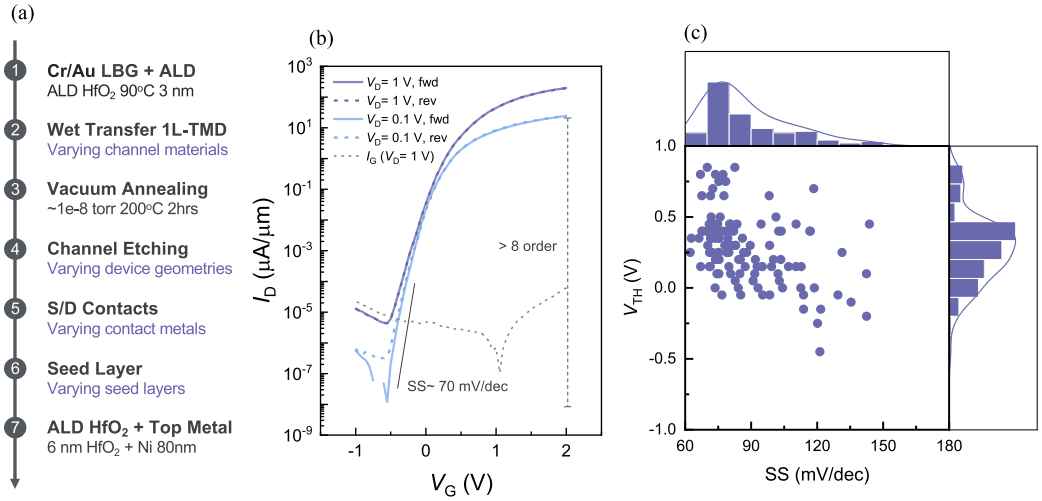


Fig. 2. Device fabrication and 1L-MoS2 LBG FETs. (a) Process flow for fabricating 1L-TMD FETs. (b) Transfer characteristics (I_D–V_G) of a representative 1L-MoS_2 transistor made of a 55-nm L_{CH} with Ni contacts, reaching ~184 μA/μm at $V_G = 2$ V and $V_D = 1$ V, a high ON/OFF ratio of ~eight orders of magnitude, a good SS of 70 mV/dec, and negligible hysteresis. (c) Distribution of SS and $V_{TH, CC}$. Each data point represents a different device fabricated by the same recipe: MBE MoS_2 as the channel material with short channel lengths (<100 nm), channel width of 860 nm, and S/D contacts of Ni.

II. Features Determining the Electrical Characteristics of 2-D 1L TMD Transistors

Fig. 1 illustrates the key features impacting the electrical performance of 1L transition metal dichalcogenide (1L-TMD) FETs, including the top-gated dielectric stack, device design parameters (L_{CH}: channel length, W_{CH}: channel width, and L_{CONT}: contact length), and the interface between the TMD and the contact metals. Here, we choose 1L-MoS_2 as a representative example of 1L-TMD. All 1L-MoS_2 FETs were fabricated using a process flow illustrated in Fig. 2(a). A Cr/Au bottom gate electrode was defined and deposited onto a 90-nm SiO_2/silicon substrate employing e-beam lithography and lift-off, followed by the deposition of a 3 nm atomic layer deposition (ALD) HfO_2 gate dielectric at 90 °C. molecular-beam epitaxy (MBE)-grown 1L-MoS_2 crystals on a sapphire substrate (Growth A) or chemical vapor deposition (CVD)-grown ones on a silicon substrate (Growth B) were transferred onto the local bottom gates (LBGs) and then vacuum annealed at a pressure of $\sim 1 \times 10^{-8}$ torr at 200 °C for 2 h to remove polymer residues from the transfer process and eliminate any air gap between the 1L-MoS_2 and the HfO_2 dielectric. Note that for the encapsulation step, the same transfer and annealing processes were used for CVD-grown 1L-hBN films that get inserted at both the channel and contact interfaces. At the contact interface, the 1L-hBN interfacial layer serves as a tunneling barrier to reduce the effect of

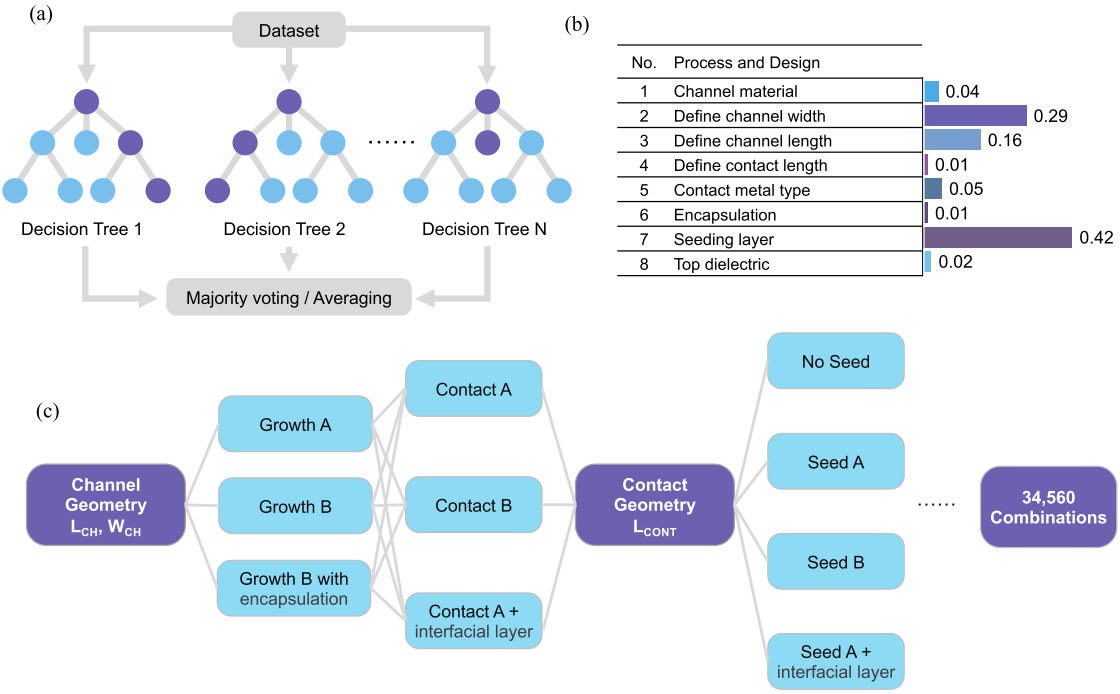


Fig. 3. Process and design optimization using ML. (a) Schematic of the random forest algorithm. A random forest regressor consists of a group of decision tree learners. Each tree randomly chooses features and iterates all possible conditions to optimize the information gain. (b) Feature importance score for a combination of SS and I_D at the same V_{OV}. (c) Grid search optimization. Device performance scores can be predicted by ML for all possible combinations using a grid-search method. Random forest is an ensemble learning method for classification and regression tasks. It combines predictions from multiple decision trees to enhance accuracy and robustness. During training, a large number of decision trees are built, and their outputs are aggregated to make the final prediction, reducing overfitting and improving generalization. Random feature selection randomly considers a subset of features for each split node to avoid overfitting. Decision tree construction utilizes the MAPE for regression tasks, minimizing it at each node. Voting for predictions combines all decision trees' predictions, averaging them for the final prediction.

Fermi-level pinning, while it acts as a passivation layer to reduce the impact from defects introduced by the top-gate dielectric integration in the channel region.

After the TMD transfer, reactive ion etching using Cl_2/O_2 at a power of 40 W was employed to define the device width W_{CH}. Next, e-beam lithography was employed again to define source/drain (S/D) contacts (Contact A = Ni and Contact B = Bi), in this way creating devices with a range of geometries (L_{CH} and L_{CONT}). Finally, a seeding layer (Seed A = Al and Seed B = Ta) was deposited, followed by the deposition of 6-nm ALD HfO_2 gate dielectric at 90 °C and patterning of a Ni top metal gate using e-beam lithography and lift-off. All electrical measurements were performed in a probe station under vacuum ($\sim 1 \times 10^{-5}$ torr) at room temperature.

III. Two-Dimensional TMD FETs Fabrication and Characteristics

Fig. 2(b) shows the transfer characteristics of a representative 1L-MoS_2 device with $L_{CH} = 55$ nm. The device metrics extraction was performed through an automation script. The minimum subthreshold swing (SS) was calculated, and the threshold voltage ($V_{TH,CC}$) was extracted at a constant current of 100 nA/um. One of the important device performance metrics is the ON-state current at a given overdrive voltage, $V_{OV} = V_G - V_{TH,LIN}$, where $V_{TH,LIN}$ stands for the linear threshold voltage. The figure-of-merit (FOM) is constituted by the scores of SS and ON-state current at the same V_{OV} with a certain weight. The weight of these two metrics can be adjusted according to the desired specs of the 2-D FETs, e.g., for high-performance or low-power applications. The device in Fig. 2(b) reaches ~184 μA/μm at $V_G = 2$ V and $V_D = 1$ V and a high ON/OFF ratio of ~eight order. The ultrathin dielectric (3 nm) and atomically thin 1L-TMD ensure good electrostatic control of the channel, leading to a good SS of 70 mV/dec, and negligible hysteresis. Detailed device characterization can be seen in [13]. Fig. 2(c) shows the distribution of SS and $V_{TH,CC}$ values for 100 electrical characteristics from devices that were fabricated with the same design and under the same processing conditions. Each data point represents a single device.

Among a plethora of ML algorithms, such as the support vector machine (SVM) [23], linear regression model, and convolutional neural network (CNN) [24], we choose a decision tree-based model [25], since it can efficiently handle discrete data. The model randomly selects features and builds hundreds of decision trees. In particular, we use Gini impurity to measure the likelihood of mislabeling a randomly chosen element from a set, based on the distribution of labels within that set. For each decision tree, all possible conditions are iterated, the Gini impurity at each node is calculated, and the information gain is optimized. Finally, majority voting is conducted to average out the result as illustrated in Fig. 3(a) and (b). This method not only demonstrates the highest accuracy of 93% via the mean absolute percentage error (MAPE) but also shows several other advantages if compared with other conventional ML algorithms, including having a lower chance of overfitting, fewer hyperparameters to be determined, and being

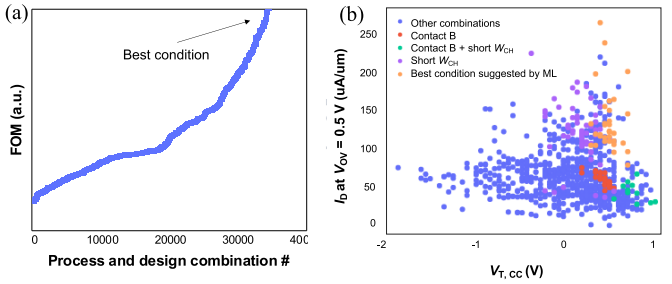


Fig. 4. Grid search optimization and prediction from ML. (a) Process and design combinations ranked by ascending FOM to simultaneously achieve high ON-currents and positive $V_{TH, CC}$. (b) Scatter plot of I_D at $V_{OV} = 0.5$ V versus $V_{TH,CC}$ from ~1000 1L-TMD FETs. The color corresponds to different design or process conditions. The green, red, and purple points are three batches of devices designed by human experiences based on step-by-step optimization as discussed in the main text. The best condition shown in orange color is the result of the recipe in Fig. 4(a) pointed by the black arrow. Noted, all data points shown were based on the experimental fabrication of devices.

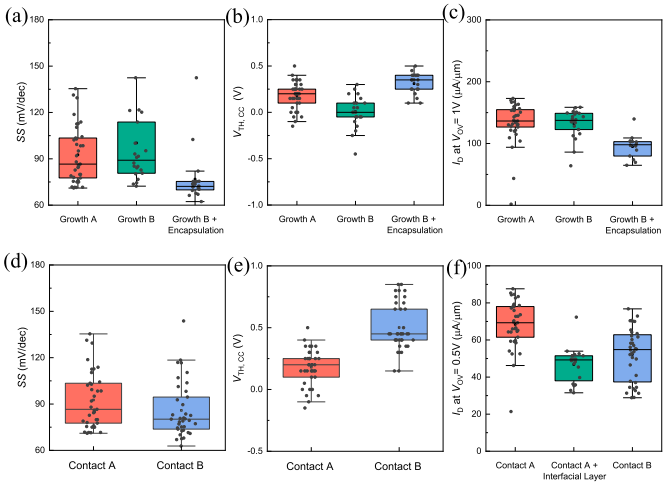


Fig. 5. Impact of channel material and metal contact. (a)–(c) Channel material impact. Box plots of (a) SS, (b) $V_{TH,CC}$, and (c) I_D at $V_{OV} = 1$ V, comparing 1L-MoS_2 grown by MBE, CVD, and CVD with 1L-hBN encapsulation. (d)–(f) Metal contact impact. Box plots of (d) SS, (e) $V_{TH,CC}$, and (f) I_D at $V_{OV} = 0.5$ V, comparing Ni, Bi, and Ni/1L-hBN contacted 1L-MoS_2 devices.

immune to missing values and unnormalized data. Although ML represents a computer-aided learning process from a black box containing large amounts of device data, the optimization process through ML reveals some physical mechanisms behind the experimental data, which will be discussed in the following sections.

IV. Framework and Algorithm

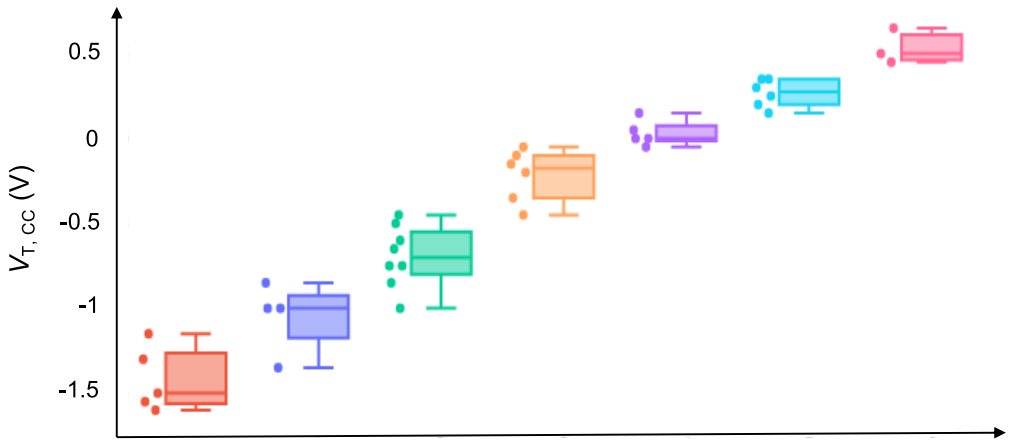
Based on the training results, one can simply extract the normalized feature importance score [26] with regard to the predicted label, as shown in Fig. 3(b). With these importance scores in place, we can quantitatively evaluate the impact of each process step or design feature on the final device performance. Next, the grid search method is adopted to iterate through all possible combinations (34 560 in this case), as illustrated in Fig. 3(c). In this way, a prediction is made about which process combination gives rise to the highest FOM for the targeted application. This allows device fabrication to adopt the optimal process parameters based on the recipe suggested by ML, as shown in Fig. 4(a). To avoid outlier data to prevent a meaningful analysis, every design of experiments (DOEs) is conducted judiciously with at least several device characteristics to ensure data uniformity. The outlier data are thus not included for further analysis.

A scatter plot of around 1000 1L-TMD FETs is shown in Fig. 4(b). For instance, one recipe using a particular contact metal (red dots, Contact B) shows positive $V_{TH,CC}$, while another recipe pushing for small channel widths (purple dots) gives higher ON-state currents per width at the same V_{OV}. However, combining these two recipes (green dots) did not yield both, high current density and positive $V_{TH,CC}$, as apparent from the plot. This is counterintuitive and cannot be explained by a simple logic analysis. The underlying reason is that individual processes and design parameters are highly coupled and those connections are not yet fully understood, making process optimization of 2-D transistors complicated. With the growing complexity of future transistor technology nodes, the difficulty in device process optimization will increase exponentially. Therefore, utilizing ML in this context is highly beneficial. This is evident from the fact that devices that yield both high ON-states and positive $V_{TH,CC}$ are obtained (orange dots) when exactly following the process and design combination predicted by the grid search optimization with the highest FOM [see the "best condition" black arrow in Fig. 4(a)]. This specific recipe suggests using 1L-MoS_2 grown by MBE as the channel material (instead of CVD) with small channel widths and S/D contact of Ni. Note that the devices denoted by red dots actually used Bi contacts.

V. Statistical Analysis of Chanel Materials and Metal Contacts

With more than 1000 1L-MoS_2 devices, we are able to use our data to manually draw many conclusions about the impact of materials and processes on certain device metrics, when the correlation is straightforward. Below, two examples are given. Fig. 5(a)–(c) compares the ON- and OFF-states performance of 1L-MoS_2 devices using channel materials grown by different synthesis methods, i.e., MBE (Growth A) and CVD (Growth B). It is understood that SS, $V_{TH,CC}$, and ON-currents heavily rely on the material qualities, such as defect density, doping level, and so on. Our statistics reveal that MBE MoS_2 offers a slightly better ON-state current at the same overdrive voltage and a ~10% better SS compared to CVD MoS_2. By incorporating an encapsulation layer after CVD growth, a substantial improvement in SS is observed. Fig. 5(d)–(f) compares the impact of using different metal contacts, Ni (Contact A) and Bi (Contact B), on the ON- and OFF-states performance. Particularly, Ni contacts outperform Bi contacts in ON-state performance. Adding an 1L-hBN interfacial layer did not improve the Fermi level pinning, but rather decreased current injection from Ni contacts. The lower current is a result of electrons injected from the contacts being "blocked" by the 1L-hBN interfacial layer in between the channel and the contacts.

The individual steps of gate-stack formation are very important to both the ON- and OFF-states performance. For example,



Channel	Growth B	Growth B	Growth A	Growth A	Growth B	Growth A	Growth A
L_{CH} (nm)	75	55	75	55	75	75	75
W_{CH} (nm)	860	860	860	860	115	115	115
Contact metal	Contact A	Contact A	Contact A	Contact B	Contact A	Contact A	Contact B
Top seeding	Seed A	Seed A	Seed B	Seed B	No Seed	No Seed	Seed B
Top dielectric	HfO_2	No	HfO_2	HfO_2	No	No	No

Fig. 6. V_{TH} tuning through different process combinations. Seven different threshold voltages can be achieved in short-channel 1L-MoS_2 devices, following the recipes predicted by ML optimization of the process conditions and design parameters.

depositing an Al seeding layer onto the intrinsic channel followed by HfO_2 to complete the top gate dielectric stack can substantially change the TMD channel properties, such as the degradation of SS and significant shift of V_{TH} [12], [13]. However, by adding an encapsulation layer onto the channel region, the detrimental impact of the gate dielectric stack formation can be greatly mitigated, as shown in our previous work [12]. It is clear from these examples that the combining different process steps can result in different effects on the final device performance. This is why ML can be a powerful approach to optimize the parameter set needed to achieve the desired performance for a particular target application. More importantly, some device metrics are correlated by some hidden materials or process parameters that are not readily accessible to an analysis by a human. Note that the demands from the application side are increasingly high on the semiconductor industry, e.g., semiconductor foundries nowadays have to provide several threshold voltage options to meet customers' circuit design requirements, we show in the next section how ML can help identifying parameter sets to satisfy said customer needs.

VI. Threshold Voltage Tunning via ML Optimization

Fig. 6 illustrates how seven different $V_{TH,CC}$, ranging from 0.5 to −1.5 V, can be achieved using ML by deliberately choosing process conditions and design parameters collected from statistical data presented here. This level of V_{TH} tuning has not been reported previously and is here presented for the first time in TMD-based devices through ML-based design and process co-optimization. Note that in Si transistors, usually different work functions of gate-metal stack are employed to tune V_{TH}. Differently, since 1L-TMD transistors are "interface" only transistors, the variation of process conditions and design parameters can significantly impact the threshold voltage.

VII. Conclusion

In this article, we have presented a novel ML-based framework for improving and predicting the performance of 2-D semiconductor transistors by co-optimizing process conditions and design features. This framework allows to expedite the process of identifying the parameters that unlock the ultimate performance of TMD transistors by utilizing a feature importance score to evaluate the impact of each process step or design feature. This framework is intended to reveal the potential performance of 2-D FETs in an efficient manner that allows for rapid experimental iterations. Future work will add more data and conditions to enrich the present database and expand the search space. The framework presented here underscores the potential of ML to advance 2-D electronics.

References

[1] T. N. Theis and H.-S. P. Wong, "The end of Moore's law: A new beginning for information technology," *Comput. Sci. Eng.*, vol. 19, no. 2, pp. 41–50, Mar. 2017, doi: 10.1109/MCSE.2017.29.

[2] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, "Single-layer MoS_2 transistors," *Nature Nanotechnol.*, vol. 6, no. 3, pp. 147–150, Mar. 2011, doi: 10.1038/nnano.2010.279.

[3] S. B. Desai et al., "MoS_2 transistors with 1-nanometer gate lengths," *Science*, vol. 354, no. 6308, pp. 99–102, Oct. 2016, doi: 10.1126/science.aah4698.

[4] Y. Liu, X. Duan, H.-J. Shin, S. Park, Y. Huang, and X. Duan, "Promises and prospects of two-dimensional transistors," *Nature*, vol. 591, no. 7848, pp. 43–53, Mar. 2021, doi: 10.1038/s41586-021-03339-z.

[5] S. Manzeli, D. Ovchinnikov, D. Pasquier, O. V. Yazyev, and A. Kis, "2D transition metal dichalcogenides," *Nature Rev. Mater.*, vol. 2, no. 8, Jun. 2017, doi: 10.1038/natrevmats.2017.33.

[6] M. Chhowalla, D. Jena, and H. Zhang, "Two-dimensional semiconductors for transistors," *Nature Rev. Mater.*, vol. 1, no. 11, p. 16052, Aug. 2016, doi: 10.1038/natrevmats.2016.52.

[7] A. Allain, J. Kang, K. Banerjee, and A. Kis, "Electrical contacts to two-dimensional semiconductors," *Nature Mater.*, vol. 14, no. 12, pp. 1195–1205, Dec. 2015, doi: 10.1038/nmat4452.

[8] C. Liu et al., "Two-dimensional materials for next-generation computing technologies," *Nature Nanotechnol.*, vol. 15, no. 7, pp. 545–557, Jul. 2020, doi: 10.1038/s41565-020-0724-3.

[9] P.-C. Shen et al., "Ultralow contact resistance between semimetal and monolayer semiconductors," *Nature*, vol. 593, no. 7858, pp. 211–217, May 2021, doi: 10.1038/s41586-021-03472-9.

[10] C. J. McClellan, E. Yalon, K. K. H. Smithe, S. V. Suryavanshi, and E. Pop, "High current density in monolayer MoS_2 doped by AlOx," *ACS Nano*, vol. 15, no. 1, pp. 1587–1596, Jan. 2021, doi: 10.1021/acsnano.0c09078.

[11] W. Li et al., "Approaching the quantum limit in two-dimensional semiconductor contacts," *Nature*, vol. 613, no. 7943, pp. 274–279, Jan. 2023, doi: 10.1038/s41586-022-05431-4.

[12] H.-Y. Lan, J. Appenzeller, and Z. Chen, "Dielectric interface engineering for high-performance monolayer MoS_2 transistors via hBN interfacial layer and Ta seeding," in *IEDM Tech. Dig.*, Dec. 2022, pp. 7.7.1–7.7.4, doi: 10.1109/IEDM45625.2022.10019439.

[13] H.-Y. Lan, V. P. Oleshko, A. V. Davydov, J. Appenzeller, and Z. Chen, "Dielectric interface engineering for high-performance monolayer MoS_2 transistors via TaOxInterfacial layer," *IEEE Trans. Electron Devices*, vol. 70, no. 4, pp. 2067–2074, Apr. 2023, doi: 10.1109/TED.2023.3251965.

[14] J. Jiang, L. Xu, C. Qiu, and L.-M. Peng, "Ballistic two-dimensional InSe transistors," *Nature*, vol. 616, no. 7957, pp. 470–475, Mar. 2023, doi: 10.1038/s41586-023-05819-w.

[15] C.-C. Chiang, H.-Y. Lan, C.-S. Pang, J. Appenzeller, and Z. Chen, "Air-stable p-doping in record high-performance monolayer WSe_2 devices," *IEEE Electron Device Lett.*, vol. 43, no. 2, pp. 319–322, Feb. 2022, doi: 10.1109/LED.2021.3135312.

[16] R. Wu et al., "Bilayer tungsten diselenide transistors with on-state currents exceeding 1.5 milliamperes per micrometre," *Nature Electron.*, vol. 5, no. 8, pp. 497–504, Aug. 2022, doi: 10.1038/s41928-022-00800-3.

[17] Y. Wang et al., "P-type electrical contacts for 2D transition-metal dichalcogenides," *Nature*, vol. 610, no. 7930, pp. 61–66, Oct. 2022, doi: 10.1038/s41586-022-05134-w.

[18] C. Pang et al., "Atomically controlled tunable doping in high-performance WSe_2 devices," *Adv. Electron. Mater.*, vol. 6, no. 8, Aug. 2020, Art. no. 1901304, doi: 10.1002/aelm.201901304.

[19] Y. Jung et al., "Transferred via contacts as a platform for ideal two-dimensional transistors," *Nature Electron.*, vol. 2, no. 5, pp. 187–194, May 2019, doi: 10.1038/s41928-019-0245-y.

[20] H. J. Chuang et al., "Low-resistance 2D/2D ohmic contacts: A universal approach to high-performance WSe_2, MoS_2, and $MoSe_2$ transistors," *Nano Lett.*, vol. 16, no. 3, pp. 1896–1902, Mar. 2016, doi: 10.1021/acs.nanolett.5b05066.

[21] X. Chen et al., "Wafer-scale functional circuits based on two dimensional semiconductors with fabrication optimized by machine learning," *Nature Commun.*, vol. 12, no. 1, p. 5953, Oct. 2021, doi: 10.1038/s41467-021-26230-x.

[22] H. Xu et al., "A machine learning approach for optimization of channel geometry and source/drain doping profile of stacked nanosheet transistors," *IEEE Trans. Electron Devices*, vol. 69, no. 7, pp. 3568–3574, Jul. 2022, doi: 10.1109/TED.2022.3175708.

[23] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2560–2574, Aug. 2007, doi: 10.1016/j.ymssp.2006.12.007.

[24] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6, doi: 10.1109/ICEngTechnol.2017.8308186.

[25] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, and B. P. Feuston, "Random forest: A classification and regression tool for compound classification and QSAR modeling," *J. Chem. Inf. Comput. Sci.*, vol. 43, no. 6, pp. 1947–1958, Nov. 2003, doi: 10.1021/ci034160g.

[26] K. J. Archer and R. V. Kimes, "Empirical characterization of random forest variable importance measures," *Comput. Statist. Data Anal.*, vol. 52, no. 4, pp. 2249–2260, Jan. 2008, doi: 10.1016/j.csda.2007.08.015.