Electronic Supplementary Information for

Machine learning-enabled band gap prediction of monolayer transition metal chalcogenide alloys

Chan Gao[1,2], Xiaoyong Yang[3], Ming Jiang[1], Lixin Chen[1], Zhiwen Chen[1], Chandra Veer Singh[1,4,*]

[1]*Department of Materials Science and Engineering, University of Toronto, Toronto, ON M5S 3E4, Canada*

[2]*Institute of Nuclear Physics and Chemistry, China Academy of Engineering Physics, Mianyang 621900, China*

[3]*Condensed Matter Theory Group, Materials Theory Division, Department of Physics and Astronomy, Uppsala University, Uppsala 75120, Sweden*

[4]*Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, ON M5S 3G8, Canada*

* Correspondence and request for materials should be addressed to C.V. Singh (chandraveer.singh@utoronto.ca).

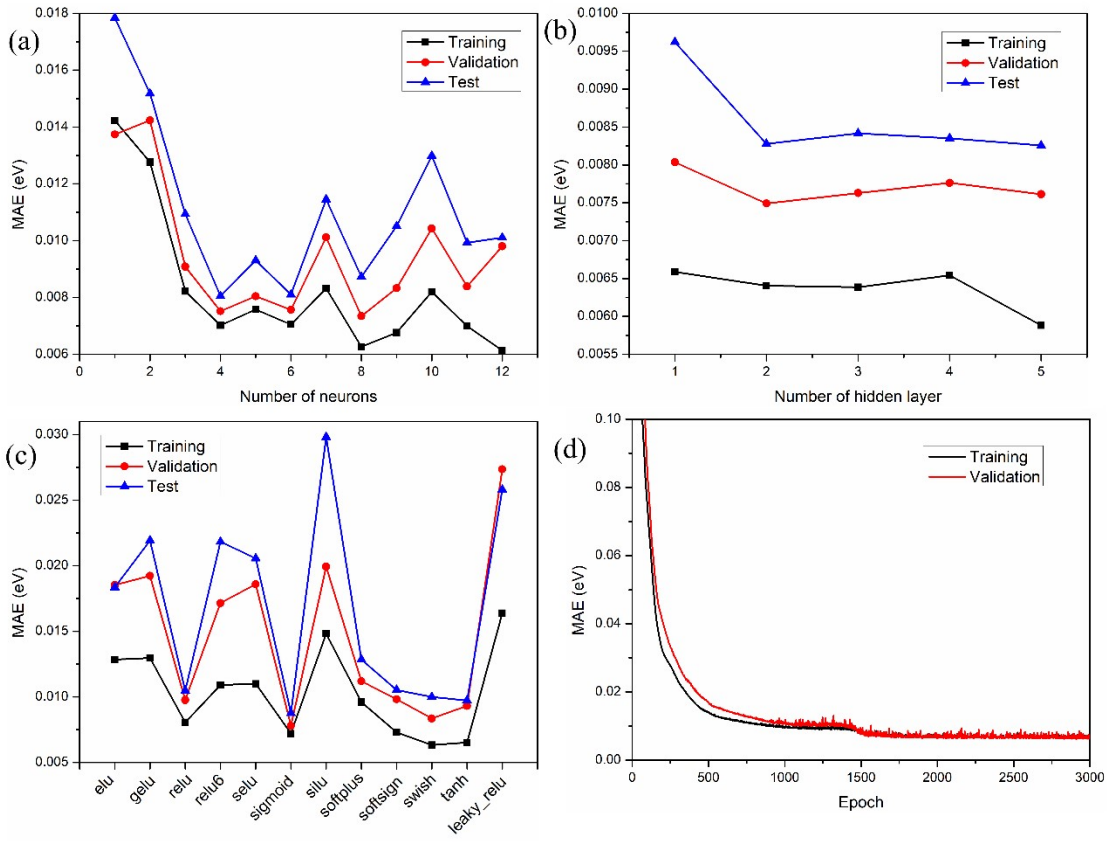

Figure S1. Variation of the mean absolute error (MAE) with (a) the number of neutrons in hidden layer, (b) the number of hidden layers, (c) the activation in hidden layer, and (d) the epoch.

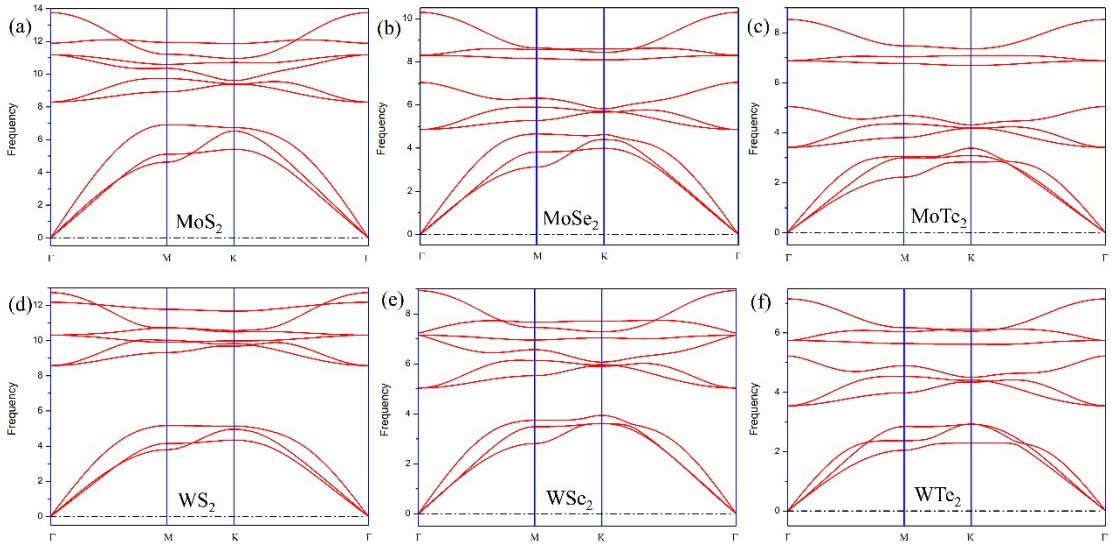

Figure S2. The phonon spectrum in pure monolayer TMD: (a) MoS_2, (b) $MoSe_2$, (c) $MoTe_2$, (d) WS_2, (e) WSe_2, (f) WTe_2.

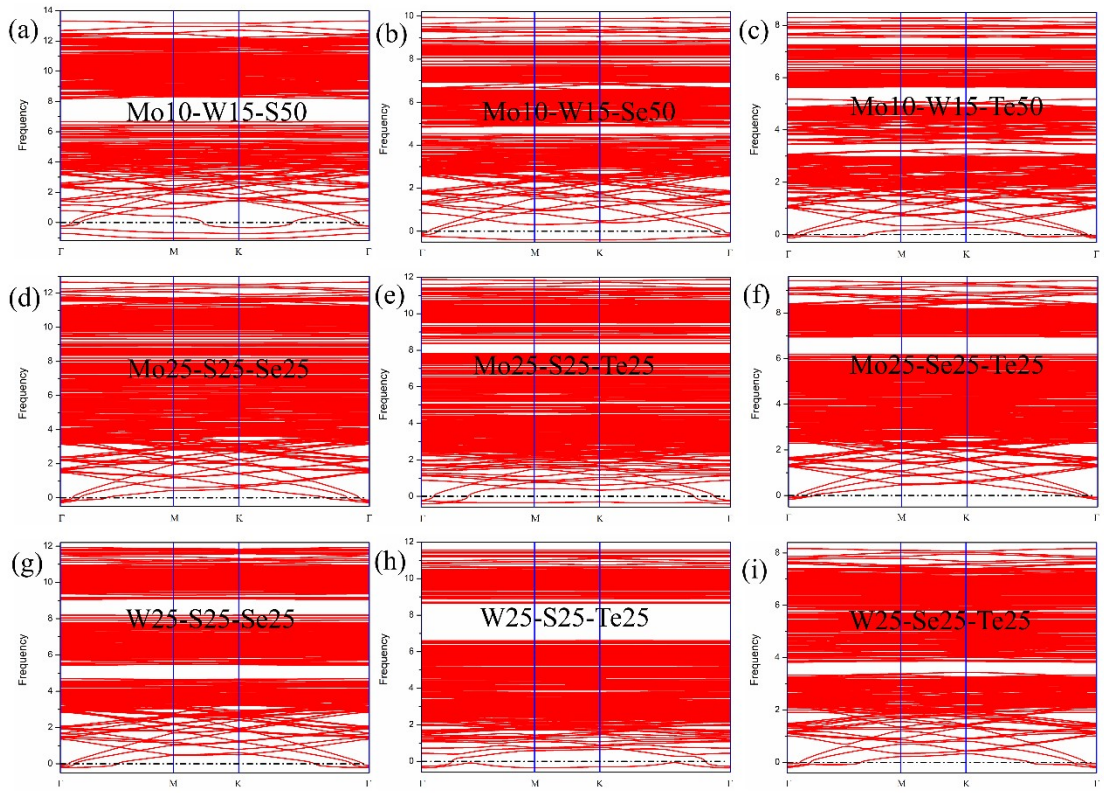

Figure S3. The phonon spectrum in the representative of ternary monolayer TMD alloy: (a) $Mo_xW_{1-x}S_2$: Mo10-W15-S50, (b) $Mo_xW_{1-x}Se_2$: Mo10-W15-Se50, (c) $Mo_xW_{1-x}Te_2$: Mo10-W15-Te50, (d) $MoS_{2x}Se_{2(1-x)}$: Mo25-S25-Se25, (e) $MoS_{2x}Te_{2(1-x)}$: Mo25-S25-Te25, (f) $MoSe_{2x}Te_{2(1-x)}$: Mo25-Se25-Te25, (g) $WS_{2x}Se_{2(1-x)}$: W25-S25-Se25, (h) $WS_{2x}Te_{2(1-x)}$: W25-S25-Te25, (i) $WSe_{2x}Te_{2(1-x)}$: W25-Se25-Te25.

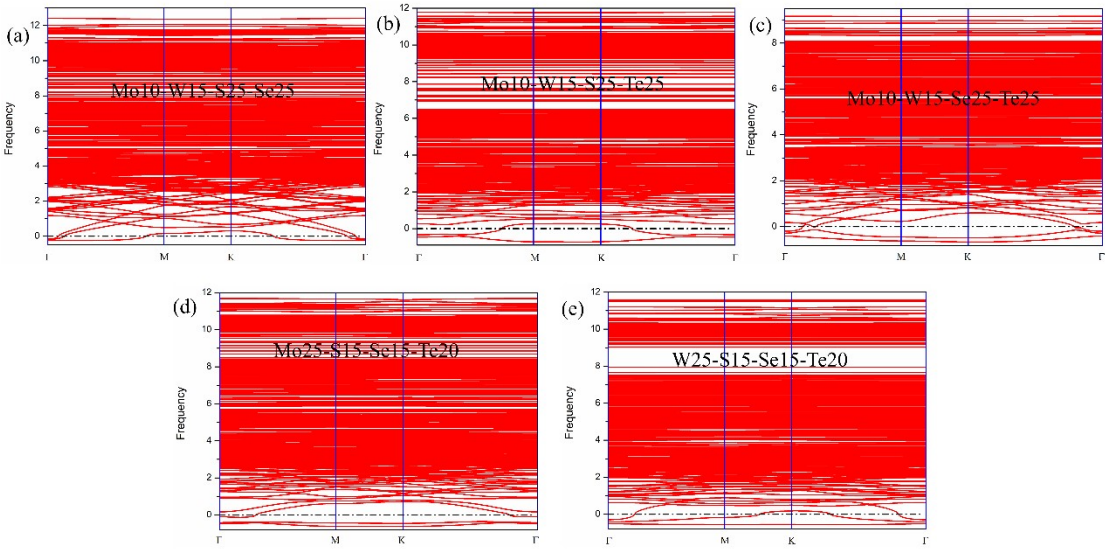

Figure S4. The phonon spectrum in the representative of quaternary monolayer TMD alloy: (a) $Mo_xW_{1-x}S_{2y}Se_{2(1-y)}$: Mo10-W15-S25-Se25, (b) $Mo_xW_{1-x}S_{2y}Te_{2(1-y)}$: Mo10-

W15-S25-Te25, (c) $Mo_xW_{1-x}Se_{2y}Te_{2(1-y)}$: Mo10-W15-Se25-Te25, (d) $MoS_{2x}Se_{2y}Te_{2(1-x-y)}$: Mo25- S15-Se15-Te20, (e) $WS_{2x}Se_{2y}Te_{2(1-x-y)}$: W25- S15-Se15-Te20.

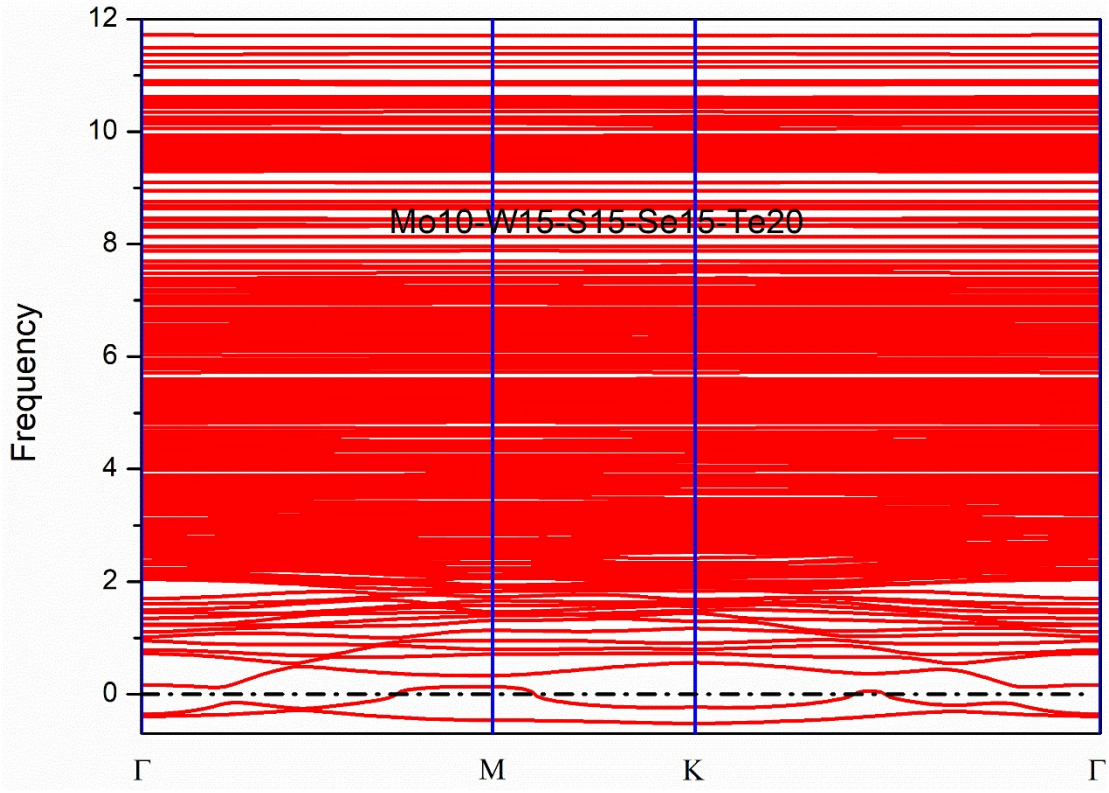

Figure S5. The phonon spectrum in the representative of quinary monolayer TMD alloy, $Mo_xW_{1-x}S_{2y}Se_{2z}Te_{2(1-y-z)}$: Mo10-W15-S15-Se15-Te20.

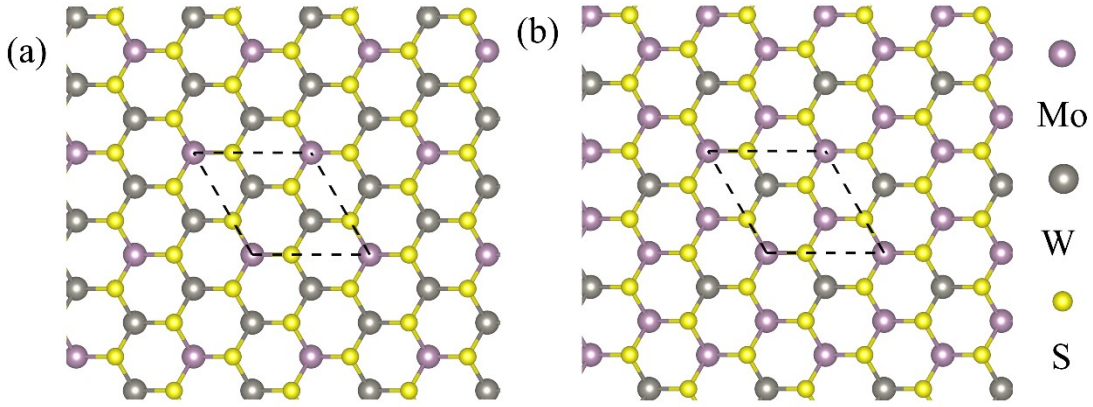

Figure S6. The atomic structure of ternary ordered monolayer $Mo_{1-x}W_xS_2$ alloy, where the unit cell is marked by dashed lines. (a) x = 2/3; (b) x = 1/3.

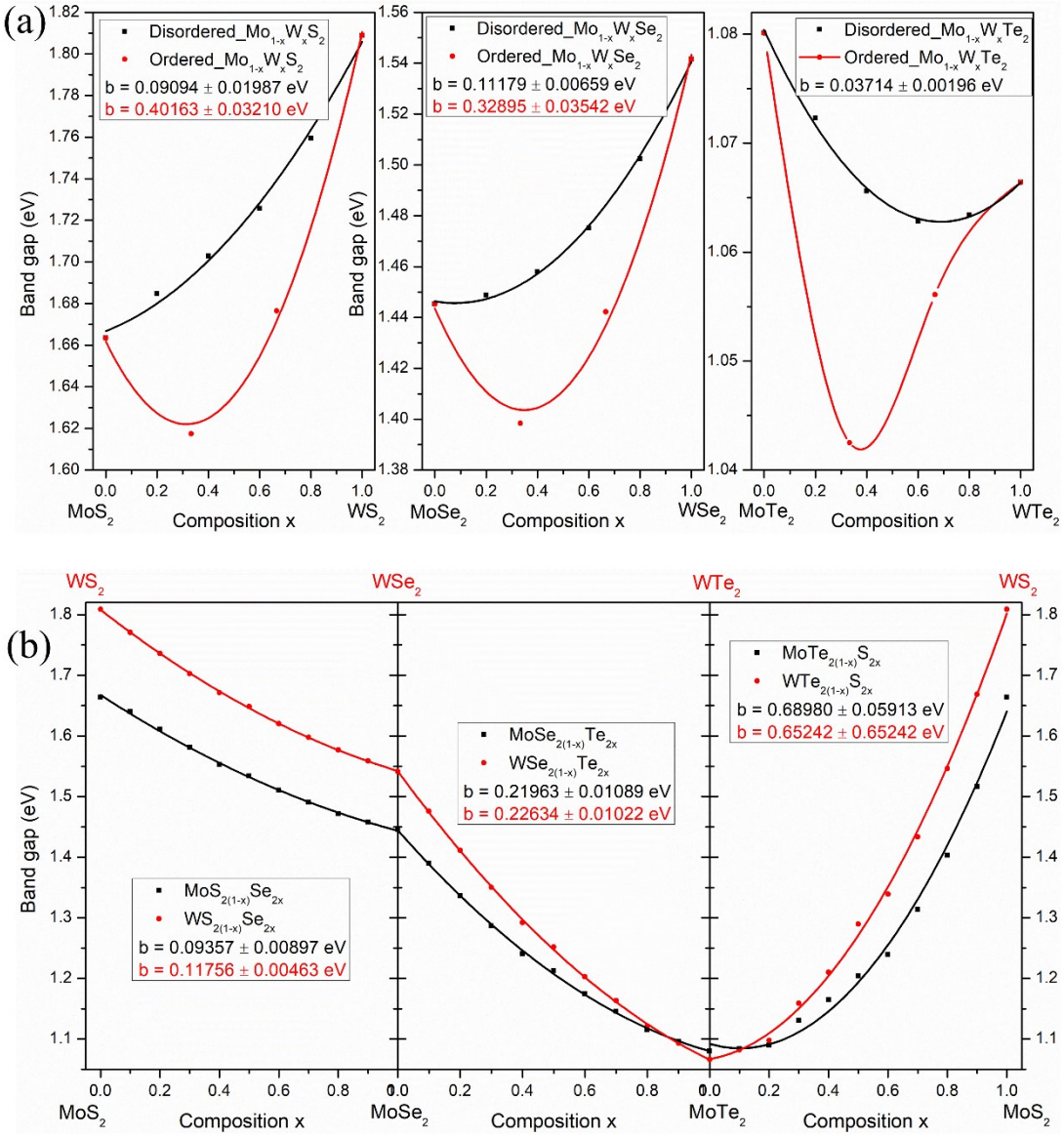

Figure S7. Composition-dependent band gaps of ternary monolayer TMD alloys based on the DFT calculations. The number after the ± sign in the bowing parameter b indicates the standard deviation. (a) Disordered and ordered monolayer TMD alloys with mixing different transition metals ($Mo_{1-x}W_xS_2$, $Mo_{1-x}W_xSe_2$, and $Mo_{1-x}W_xTe_2$); (b) Monolayer TMD alloys with mixing different chalcogen atoms ($MoS_{2(1-x)}Se_{2x}$, $MoSe_{2(1-x)}Te_{2x}$, $MoTe_{2(1-x)}S_{2x}$, $WS_{2(1-x)}Se_{2x}$, $WSe_{2(1-x)}Te_{2x}$, and $WTe_{2(1-x)}S_{2x}$).

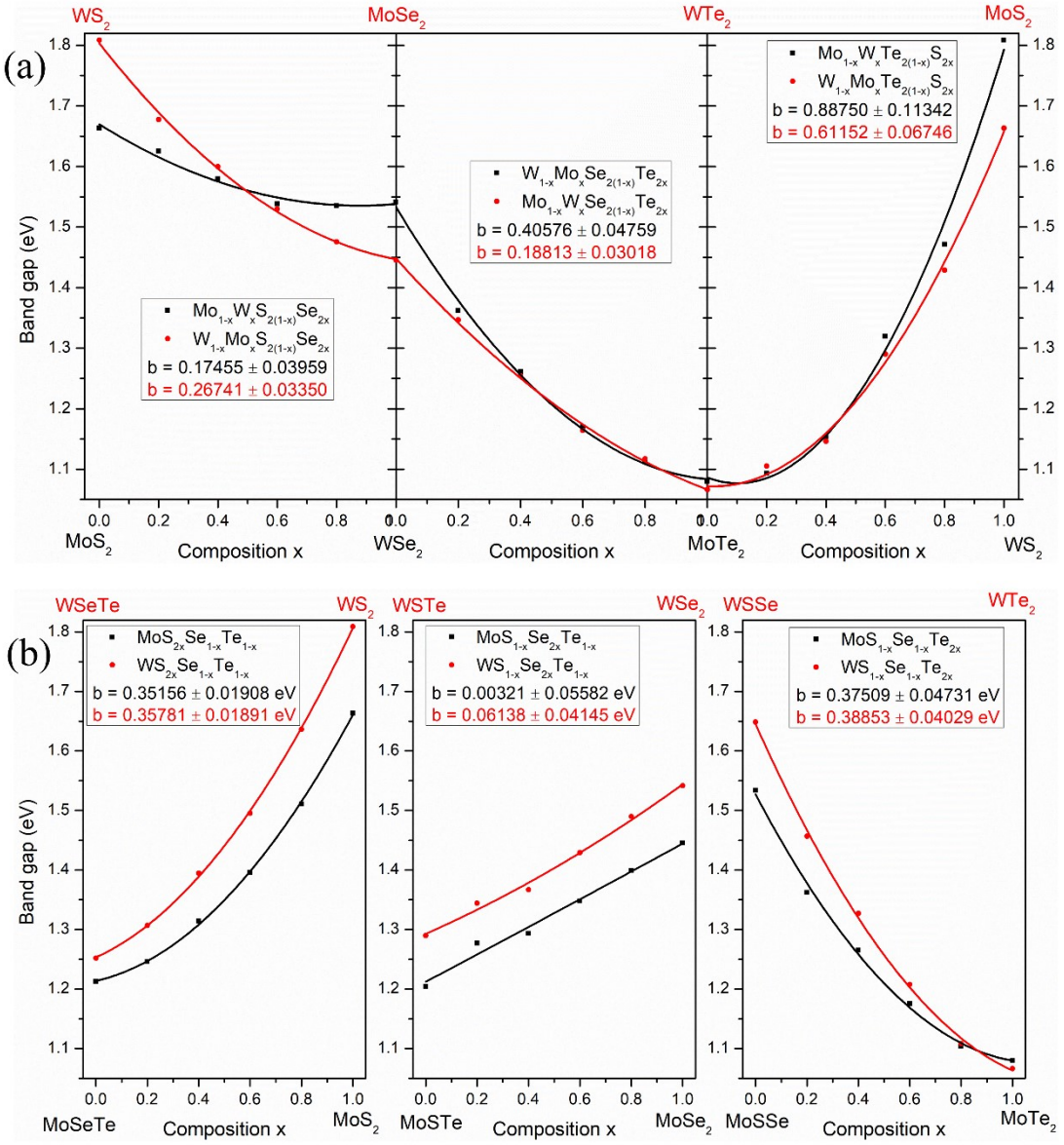

Figure S8. Composition-dependent band gaps of quaternary monolayer TMD alloys based on the DFT calculations. The number after the ± sign in the bowing parameter b indicates the standard deviation. (a) Monolayer TMD alloys with mixing different transition metals and chalcogen atoms ($Mo_{1-x}W_xS_{2(1-x)}Se_{2x}$, $W_{1-x}Mo_xS_{2(1-x)}Se_{2x}$, $W_{1-x}Mo_xSe_{2(1-x)}Te_{2x}$, $Mo_{1-x}W_xSe_{2(1-x)}Te_{2x}$, $Mo_{1-x}W_xTe_{2(1-x)}S_{2x}$, and $W_{1-x}Mo_xTe_{2(1-x)}S_{2x}$); (b) Monolayer TMD alloys with mixing different chalcogen atoms ($MoS_{2x}Se_{1-x}Te_{1-x}$, $WS_{2x}Se_{1-x}Te_{1-x}$, $MoS_{1-x}Se_{2x}Te_{1-x}$, $WS_{1-x}Se_{2x}Te_{1-x}$, $MoS_{1-x}Se_{1-x}Te_{2x}$, and $WS_{1-x}Se_{1-x}Te_{2x}$).

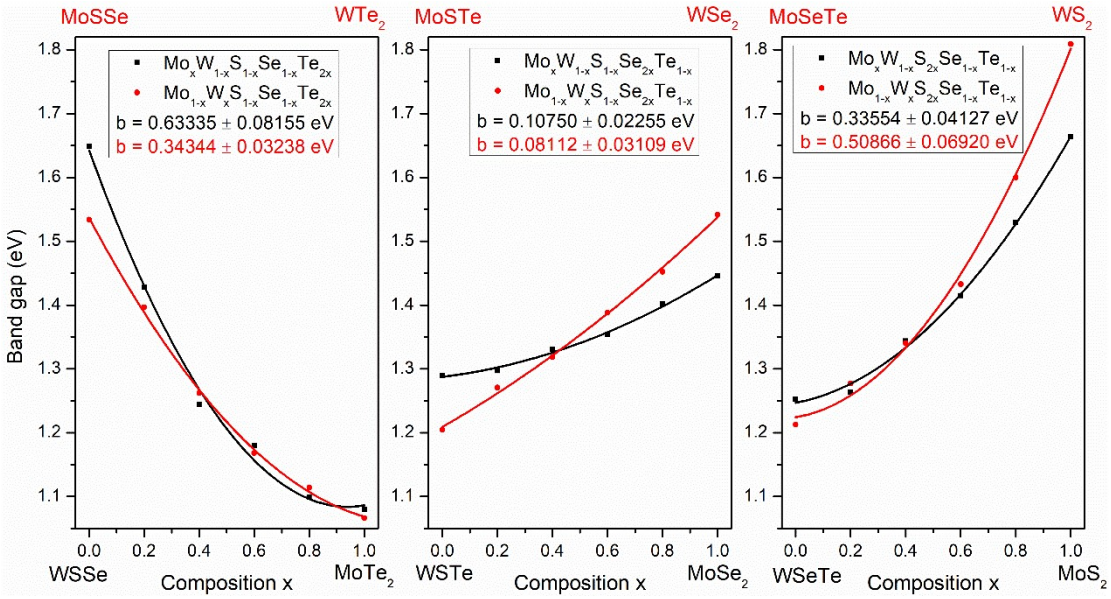

Figure S9. Composition-dependent band gaps of quinary monolayer TMD alloys with mixing different transition metals and chalcogen atoms ($Mo_xW_{1-x}S_{1-x}Se_{1-x}Te_{2x}$, $Mo_{1-x}W_xS_{1-x}Se_{1-x}Te_{2x}$, $Mo_xW_{1-x}S_{1-x}Se_{2x}Te_{1-x}$, $Mo_{1-x}W_xS_{1-x}Se_{2x}Te_{1-x}$, $Mo_xW_{1-x}S_{2x}Se_{1-x}Te_{1-x}$, and $Mo_{1-x}W_xS_{2x}Se_{1-x}Te_{1-x}$) based on the DFT calculations. The number after the ± sign in the bowing parameter b indicates the standard deviation.

Table S1. The band gaps in pure monolayer TMD calculated with GGA-PBE functional and hybrid functional HSE06, and experimental values.

TMDs	GGA-PBE (eV)	HSE06 (eV)	Expt. [71]
MoS_2	1.66	2.02	1.95
$MoSe_2$	1.45	1.72	1.58
$MoTe_2$	1.08	1.27	1.1
WS_2	1.81	1.98	1.99
WSe_2	1.54	1.63	1,65
WTe_2	1.07	1.48	-

Table S2. Bowing parameters of ternary, quaternary and quinary monolayer TMD alloys. The values of standard deviation follow ‘‘±’’ signs. b-ML and b-DFT indicate the bowing parameters of band gaps obtained from machine learning (ML) predictions

and density functional theory (DFT) calculations, respectively.

Alloy	b-ML (eV)	b-DFT (eV)
Ternary alloys		
$Mo_{1-x}W_xS_2$	0.04339 ± 0.00371	0.09094 ± 0.01987
$Mo_{1-x}W_xSe_2$	0.07837 ± 0.00494	0.11179 ± 0.00659
$Mo1\text{-}_xW_xTe_2$	0.01423 ± 0.00035	0.03714 ± 0.00196
$MoS_{2(1-x)}Se_{2x}$	0.10403 ± 0.00184	0.09357 ± 0.00897
$MoSe_{2(1-x)}Te_{2x}$	0.20841 ± 0.00336	0.21963 ± 0.01089
$MoTe_{2(1-x)}S_{2x}$	0.63326 ± 0.01146	0.68980 ± 0.05913
$WS_{2(1-x)}Se_{2x}$	0.08703 ± 0.00127	0.11756 ± 0.00463
$WSe_{2(1-x)}Te_{2x}$	0.25286 ± 0.00470	0.22634 ± 0.01022
$WTe_{2(1-x)}S_{2x}$	0.68036 ± 0.00443	0.65242 ± 0.05913
Quaternary alloys		
$Mo_{1-x}W_xS_{2(1-x)}Se_{2x}$	0.13580 ± 0.00616	0.17455 ± 0.03959
$W_{1-x}Mo_xS2_{(1-x)}Se_{2x}$	0.21403 ± 0.00264	0.26741 ± 0.03350
$W_{1-x}Mo_xSe_{2(1-x)}Te_{2x}$	0.36006 ± 0.00588	0.40576 ± 0.04759
$Mo_{1-x}W_xSe_{2(1-x)}Te_{2x}$	0.16866 ± 0.00624	0.18813 ± 0.03018
$Mo_{1-x}W_xTe_{2(1-x)}S_{2x}$	0.83943 ± 0.01783	0.88750 ± 0.11342
$W_{1-x}Mo_xTe_{2(1-x)}S_{2x}$	0.53977 ± 0.01179	0.61152 ± 0.06746
$MoS_{2x}Se_{1-x}Te_{1-x}$	0.30626 ± 0.01269	0.35156 ± 0.01908
$MoS_{1-x}Se_{2x}Te_{1-x}$	-0.02496 ± 0.00397	0.00321 ± 0.05582
$MoS_{1-x}Se_{1-x}Te_{2x}$	0.37111 ± 0.00474	0.37509 ± 0.04731
$WS_{2x}Se_{1-x}Te_{1-x}$	0.31995 ± 0.00704	0.35781 ± 0.01891
$WS_{1-x}Se_{2x}Te_{1-x}$	-0.01078 ± 0.00128	0.06138 ± 0.04145
$WS_{1-x}Se_{1-x}Te_{2x}$	0.41906 ± 0.00545	0.38853 ± 0.04029
Quinary alloys		
$Mo_xW_{1-x}S_{1-x}Se_{1-x}Te_{2x}$	0.54685 ± 0.00955	0.63335 ± 0.08155
$Mo_{1-x}W_xS_{1-x}Se_{1-x}Te_{2x}$	0.30544 ± 0.00609	0.34344 ± 0.03238
$Mo_{1-x}W_xS_{1-x}Se_{2x}Te_{1-x}$	0.06687 ± 0.00915	0.10750 ± 0.02255

$Mo_xW_{1-x}S_{1-x}Se_{2x}Te_{1-x}$	0.00911 ± 0.00181	0.08112 ± 0.03109
$Mo_xW_{1-x}S_{2x}Se_{1-x}Te_{1-x}$	0.26885 ± 0.01069	0.33554 ± 0.04127
$Mo_{1-x}W_xS_{2x}Se_{1-x}Te_{1-x}$	0.46053 ± 0.01686	0.50866 ± 0.06920