Spinel nitride solid solutions: charting properties in the configurational space via machine learning

Pablo Sánchez-Palencia,[1,2] Said Hamad,[3] Pablo Palacios,[1,4] Ricardo Grau-Crespo,[5] Keith T. Butler[6*]

[1] *Instituto de Energía Solar, ETSI Telecomunicación, Universidad Politécnica de Madrid, Ciudad Universitaria, s/n, 28040, Madrid, Spain.*

[2] *Departamento de Tecnología Fotónica y Bioingeniería, ETSI Telecomunicación, Universidad Politécnica de Madrid, Ciudad Universitaria, s/n, 28040 Madrid, Spain.*

[3] *Department of Physical, Chemical and Natural Systems, Universidad Pablo de Olavide, 41013 Sevilla, Spain.*

[4] *Departamento de Física aplicada a las Ingenierías Aeronáutica y Naval. ETSI Aeronáutica y del Espacio, Universidad Politécnica de Madrid, Pz. Cardenal Cisneros, 3, 28040 Madrid, Spain.*

[5] *Department of Chemistry, University of Reading, Reading RG6 6DX, United Kingdom.*

[6] *SciML, Scientific Computing Department, Rutherford Appleton Laboratory, Harwell OX11 0QX, United Kingdom.*

**E-mail: keith.butler@stfc.ac.uk*

Supplementary Material

Table S1 Calculated values of lattice constant and bandgap for the reference compounds with the different U-values tested in comparison with experimental values. Bold values are those in better agreement with experiments.

U value	Lattice constant (Å)	PBE+U bandgap (eV)	HSE bandgap (eV)
		Sn_3N_4	
0	9.140	0.240	1.345
1.5	9.075	0.518	**1.570**
2	9.051	0.624	1.657
2.5	**9.023**	0.741	1.752
Expt	**9.03**	**-**	**1.6**
		Ge_3N_4	
0	8.311	1.929	3.311
1.5	8.257	2.243	3.463
2	8.238	2.295	**3.496**
2.5	**8.217**	2.337	3.531
Expt.	**8.21**	**-**	**3.5**

Table S2 Number of total and calculated configurations within the whole $GeSn_2N_4$ configurational space for the different inversion degrees.

Inversion degree	GGA calculations	HSE calculations	Whole space
0	1	1	1
0.125	2	2	2
0.25	10	3	31
0.375	50	4	186

0.5	100	5	762
0.625	400	20	1337
0.75	300	15	1291
0.875	100	5	515
1	50	4	97
Total	**1013**	**59**	**4222**

Table S3 List of the main parameters used for the three statistical models (linear regressor, gradient boosted decision tree and multilayer perceptron) employed in this work.

Linear regressor (LR)	
Type of regularization	L1 (LASSO)
Alpha (reg. coef.)	10^{-12}
Python package	Scikit-learn
Gradient boosted decision tree (GBDT)	
Loss function	Squared error
Learning rate	0.01
# of estimators	1000
Min. samples to split	5
Min. samples per leaf	1 (default)
Max. depth	4
Python package	Scikit-learn
Multilayer perceptron (MLP)	
Loss function	Squared error
Learning rate	10^{-4}
# of layers	5
Activation function	ReLU*
Interm. Normalization	Yes
Interm. Dropout	Yes
Dropout rate	0.1
Optimizer	Adam (stochastic gradient descent)
Max. epochs	10000
# samples per batch	32
Python package	Tensorflow/Keras

*Linear activation function was used for final layer

Table S4 Details of clusters with identifier number, order (k) and maximum distance between atoms in the cluster (l).

Cluster index	k	l	Cluster index	k	l
1	1	0	**18**	3	5.750
2	1	0	**19**	3	5.750
3	2	3.130	**20**	3	5.750
4	2	3.670	**21**	3	5.750
5	2	3.833	**22**	3	5.750
6	2	5.421	**23**	3	6.260
7	2	5.750	**24**	3	6.260
8	2	6.260	**25**	3	6.260
9	2	6.260	**26**	3	6.260
10	3	3.130	**27**	3	6.260
11	3	3.670	**28**	3	6.260
12	3	3.670	**29**	3	6.260
13	3	3.833	**30**	3	6.260
14	3	5.421	**31**	3	6.260
15	3	5.421	**32**	3	6.260
16	3	5.421	**33**	3	7.257
17	3	5.421	**34**	3	7.340