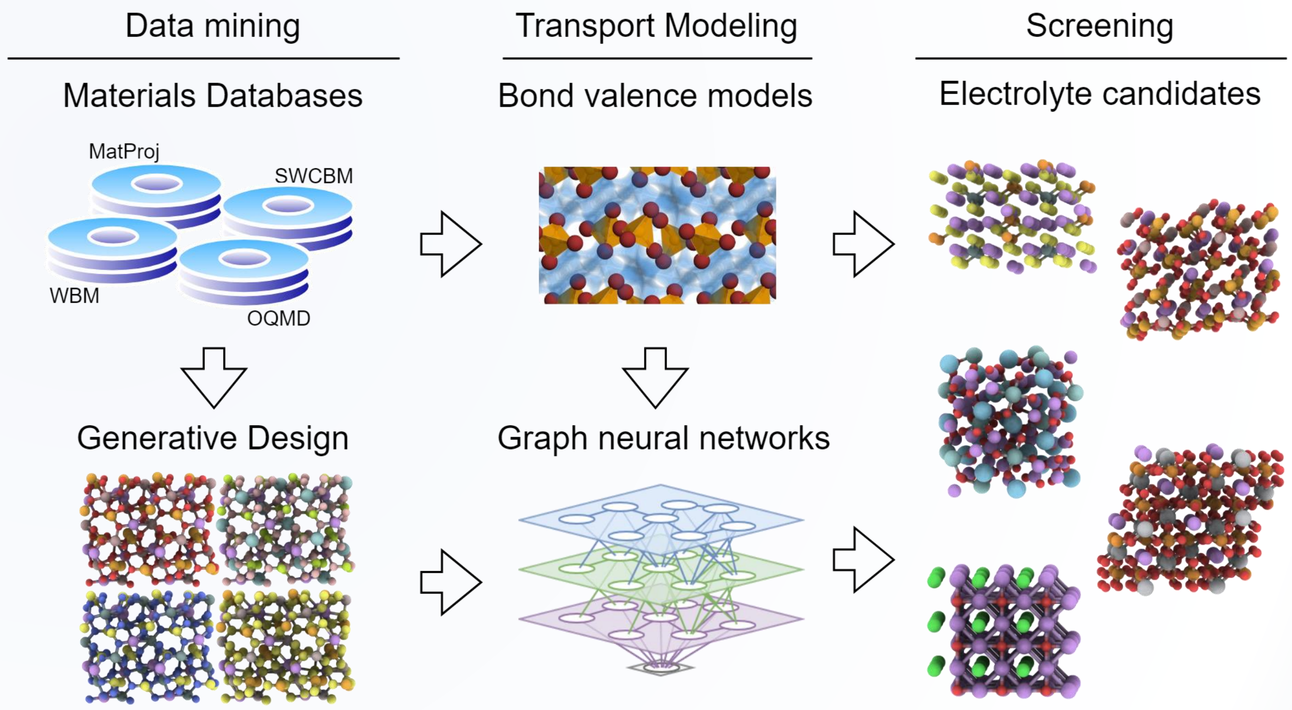
Screening for Li-based solid electrolytes using bond-valence methods and graph neural networks

Stephen R. Xie[1], Shreyas J. Honrao[1], John W. Lawson[2]

[1] KBR Inc at NASA Ames Research Center

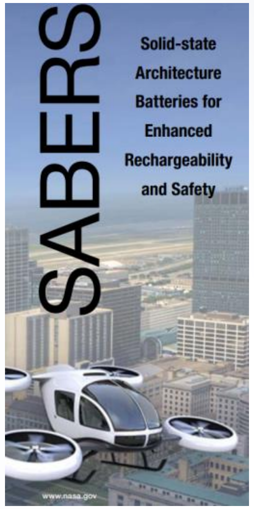
[2] NASA Ames Research Center, Moffett Field, CA, USA

- Pipeline for high-throughput ionic conductivity predictions
- Machine learning with conduction graphs from bond valence methods
- Medium-throughput screening with machine-learned interatomic potential

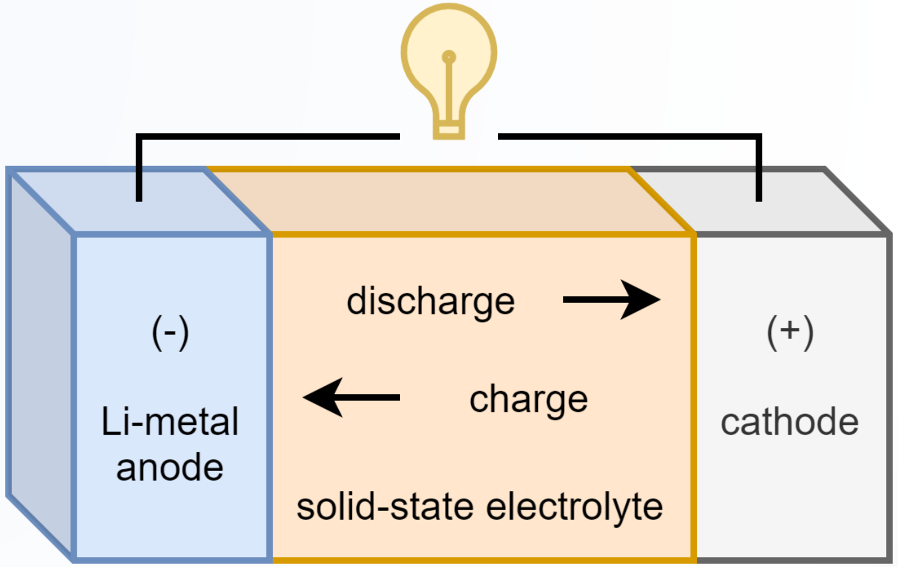


Background

- Safe, all-electric aircraft
- All-solid-state batteries
- Solid electrolyte
  - Low electrical conductivity
  - Good electrochemical stability
  - High ionic conductivity
  - Low-cost, manufacturable, …

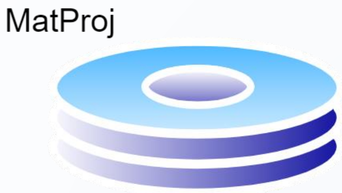


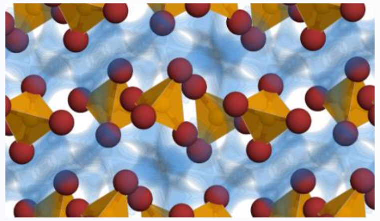
nasa.gov

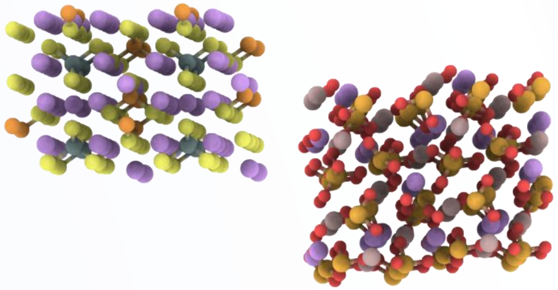


Honrao, S.J., Yang, X., Radhakrishnan, B. *et al.* Discovery of novel Li SSE and anode coatings using interpretable machine learning and high-throughput multi-property screening. *Sci Rep* **11**, 16484 (2021). https://doi.org/10.1038/s41598-021-94275-5

Data mining	Transport Modeling	Screening
Materials Databases	Bond valence models	Electrolyte candidates







- Database of 15,000 Li-containing materials
- Migration barriers
- Thermodynamic stability, electrochemical stability, and band gap
- Interpretable Gradient Boosting Models (GBMs)

Honrao, S.J., Yang, X., Radhakrishnan, B. *et al.* Discovery of novel Li SSE and anode coatings using interpretable machine learning and high-throughput multi-property screening. *Sci Rep* **11**, 16484 (2021). https://doi.org/10.1038/s41598-021-94275-5

Can we directly, efficiently compute ionic conductivity for screening?

Predicting transport through atomistic simulations

- Mean squared displacement of Li^+

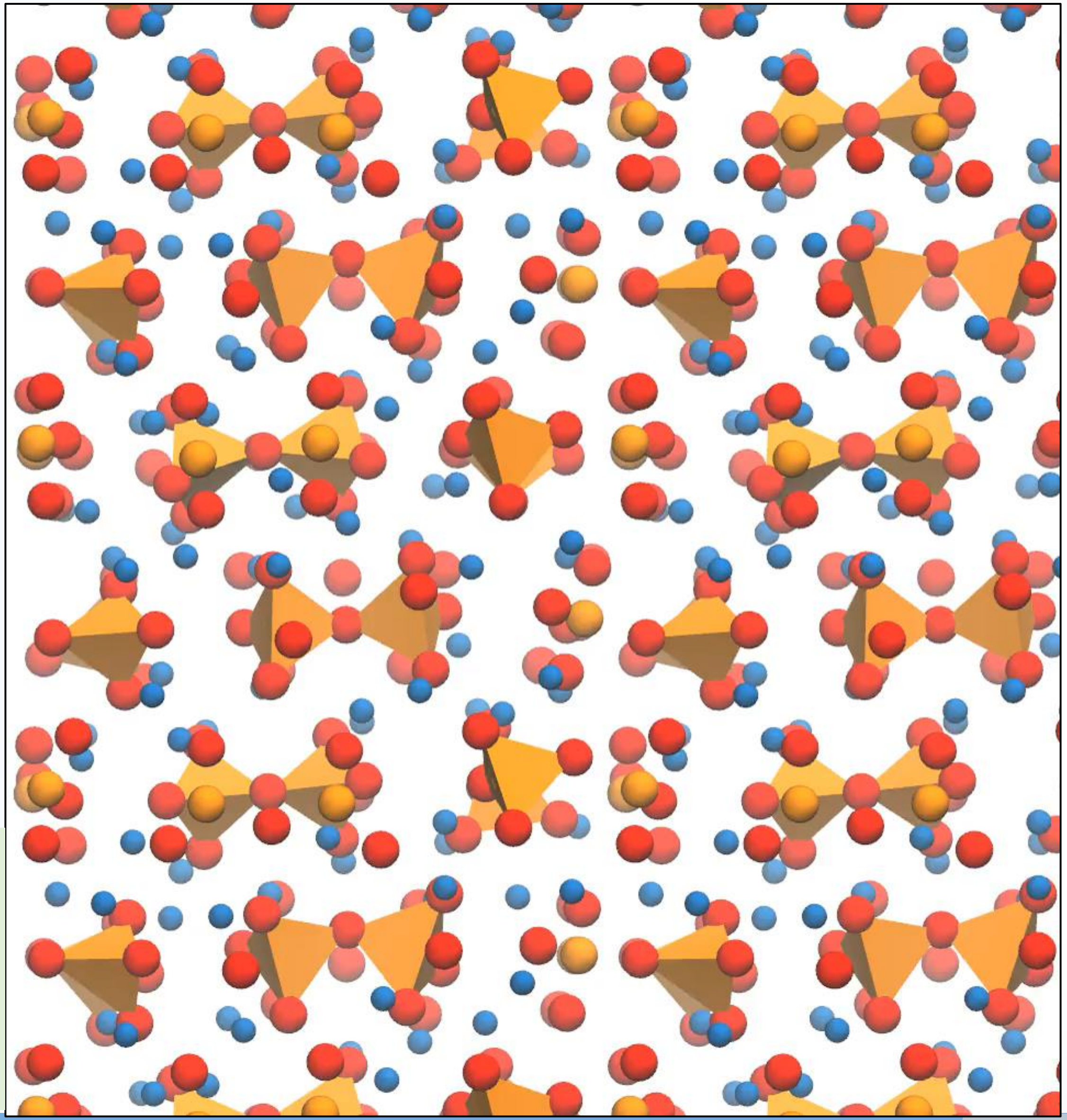
$$\mathrm{MSD} = \frac{1}{N}\sum_{i=1}^{N}\langle||\Delta r_i||^2\rangle$$

- Diffusion coefficient

$$D_{\mathrm{tracer}} = \lim_{t\to\infty}\left(\frac{\mathrm{MSD}}{6t}\right), \quad D(T) = D(\infty)\exp\left(\frac{-E_a}{k_B T}\right)$$

- Ionic Conductivity

$$\sigma(T) = \frac{\rho^2 z^2}{k_B T} D(T)$$



- 100-1000 ps of simulation time for convergence
- *ab initio* molecular dynamics (AIMD)
  - 3,000-40,000 CPU-hours (\$25 - \$414)

Simulating diffusion from first principles is expensive

Bond valence methods with softBV

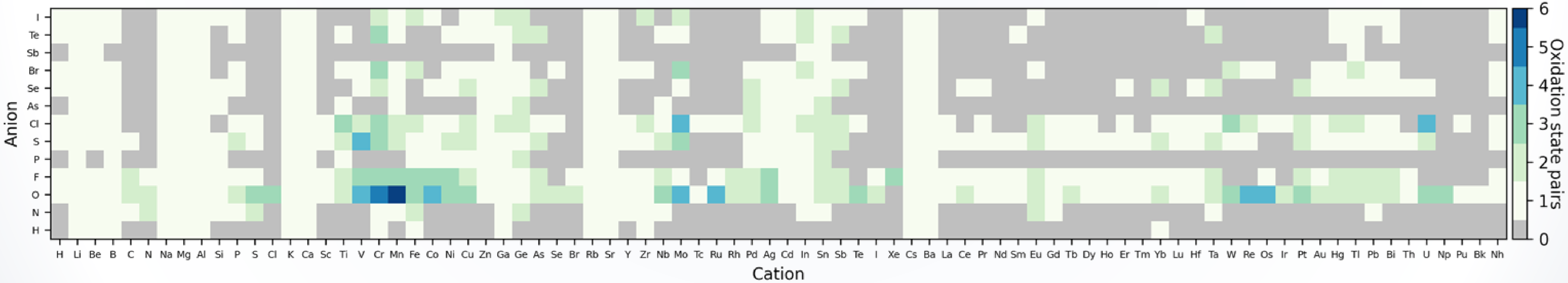
- Fast empirical force field for battery materials
- Morse-like term for cation-anion interactions
- Coulomb term for same-charge interactions
- Parametrized for hundreds of cation-anion combinations

$$E_{\text{BV,Li-X}}(r_{\text{Li-X}}) = \epsilon\left(\left(\exp\left[a\left(r_{\text{min,Li-X}} - r\text{Li-X}\right)\right] - 1\right)^2 - 1\right)$$

$$= \epsilon\left[\left(\frac{s - s_{\text{min}}}{s_{\text{min}}}\right)^2 - 1\right]$$

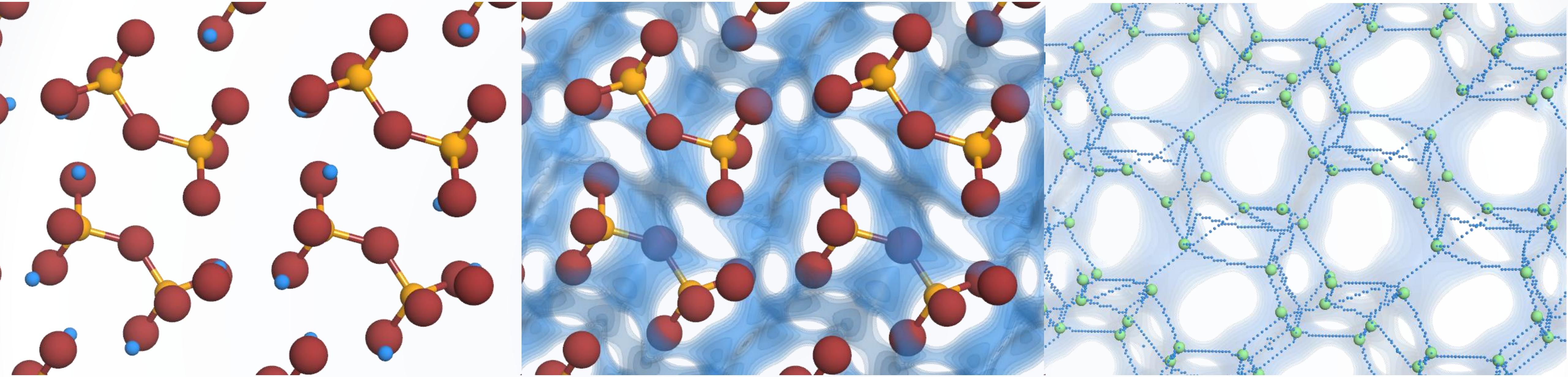
$$E_{\text{Coulomb,Li-B}}(r_{\text{Li-B}}) = \frac{q_{\text{Li}} q_{\text{B}}}{R}\text{erfc}\left(\frac{r_{\text{Li-B}}}{f \cdot (r_{\text{Li}} + r_{\text{B}})}\right)$$

$$s_{\text{Li-X}} = \exp\left(\frac{r_{0,\text{Li-X}} - r_{\text{Li-X}}}{b_{\text{Li-X}}}\right), \quad \sum_X s_{\text{Li-X}} \approx V_{0,\text{Li}}$$

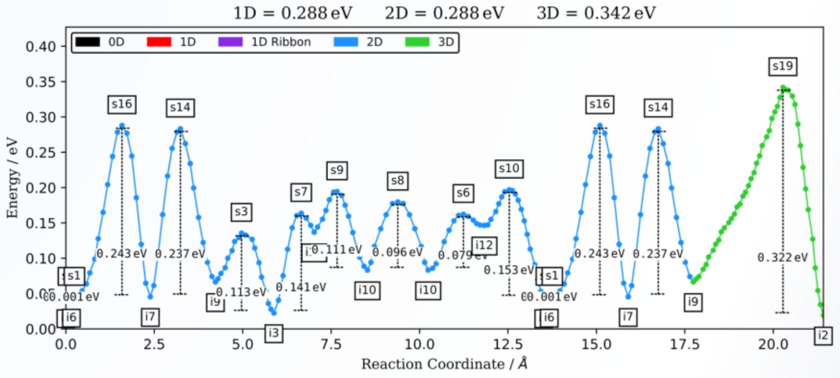


H. Chen, L. L. Wong and S. Adams. SoftBV – a software tool for screening the materials genome of inorganic fast ion conductors. Acta Cryst. (2019). B75, 18-33.

Bond valence methods with softBV

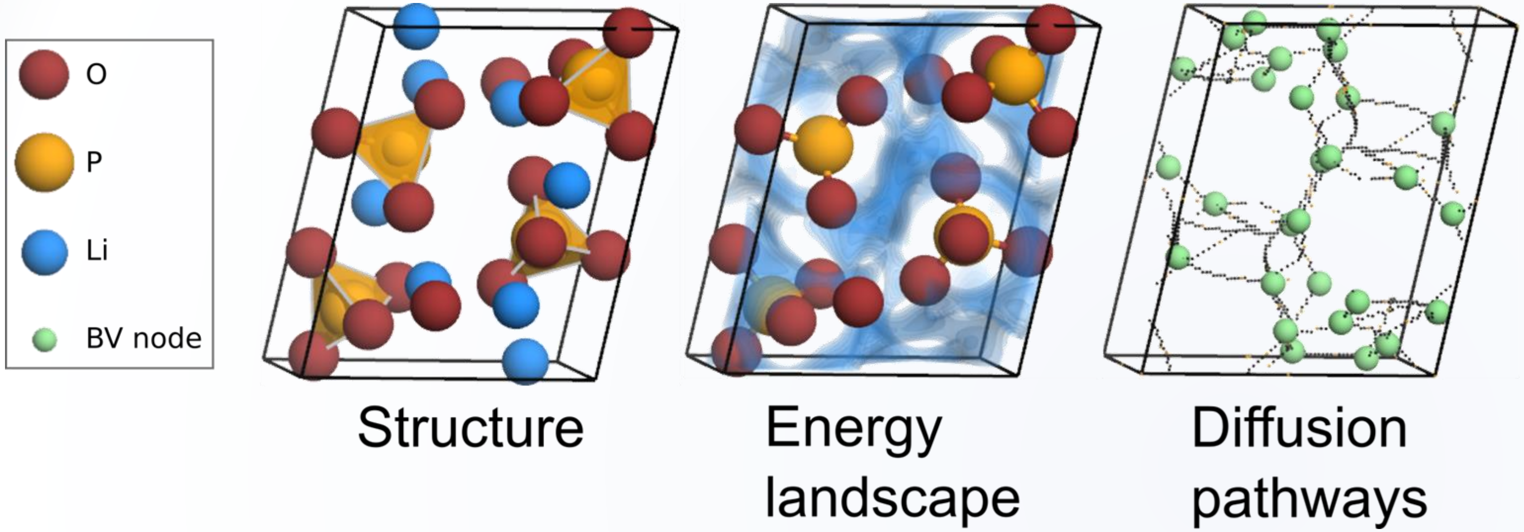


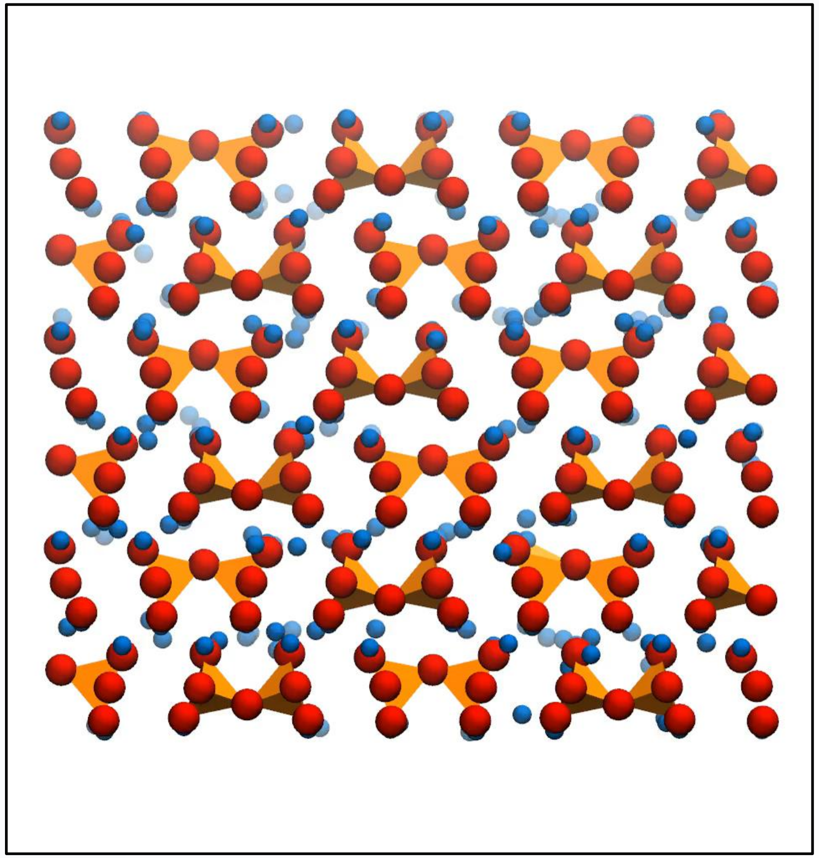
- Calculate energy landscape of Li^+ ion using the bond valence potential energy function
- Reveals low-energy pathways for Li-ion diffusion
- **No Li^+-Li^+ interaction, no movement for non-Li^+ ions**



H. Chen, L. L. Wong and S. Adams. SoftBV – a software tool for screening the materials genome of inorganic fast ion conductors. Acta Cryst. (2019). B75, 18-33.

Kinetic Monte Carlo with softBV





- Mean squared displacement of Li^+

$$\mathrm{MSD} = \frac{1}{N}\sum_{i=1}^{N}\langle||\Delta r_i||^2\rangle$$

- Ionic Conductivity

$$\sigma(T) = \frac{\rho^2 z^2}{k_B T} D(T)$$

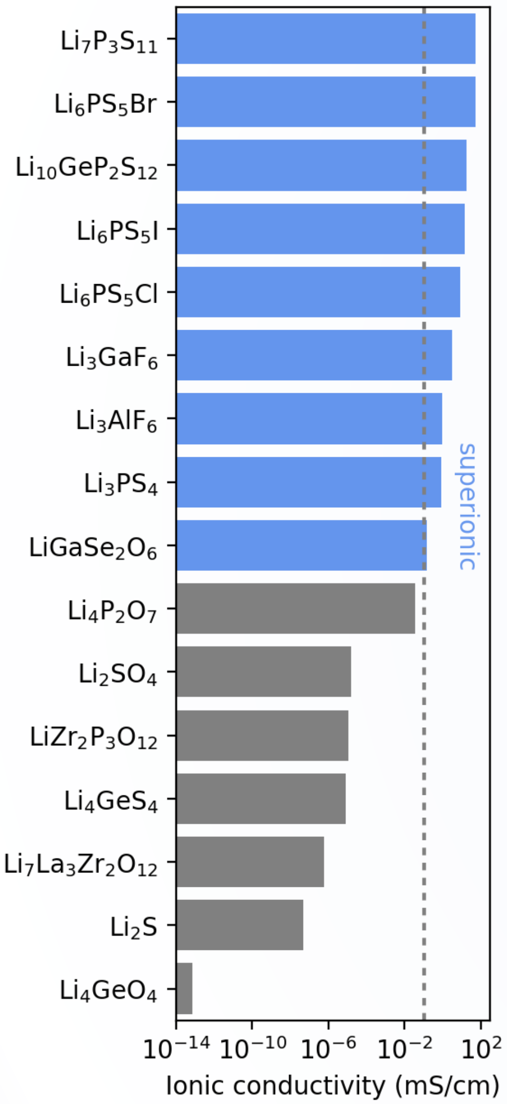
- Site-to-site transition rate

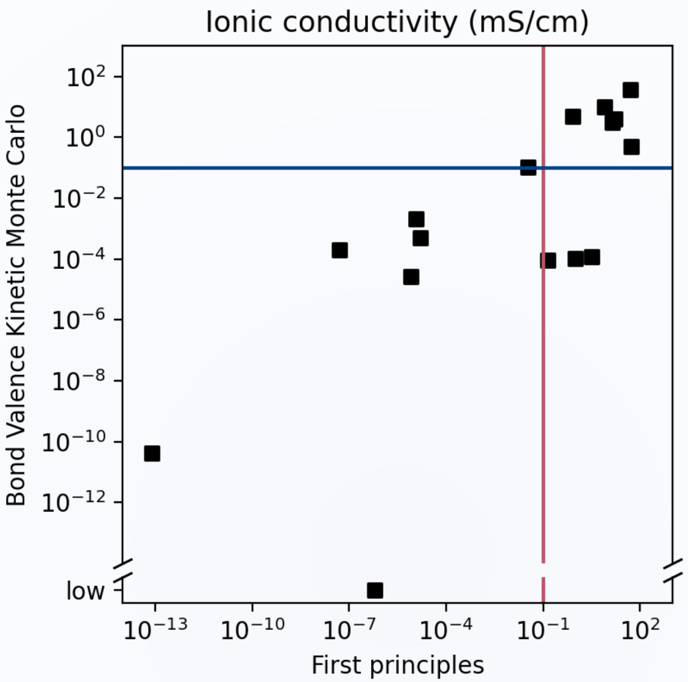
$$p = \omega \exp\left(\frac{E_m + \Delta E}{k_B T}\right)$$

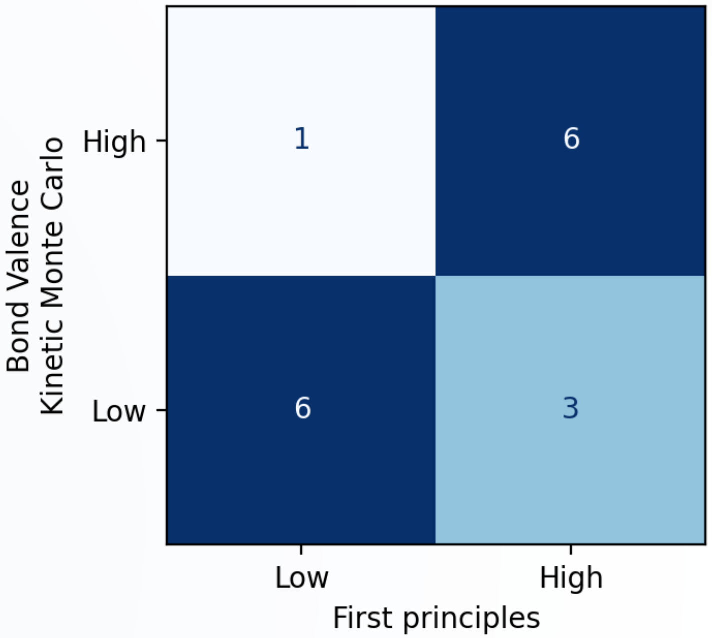
- Li^+ – Li^+ interactions included in ΔE
- **No movement for non-Li^+ ions**

H. Chen, L. L. Wong and S. Adams. SoftBV – a software tool for screening the materials genome of inorganic fast ion conductors. Acta Cryst. (2019). B75, 18-33.

Benchmarking Kinetic Monte Carlo



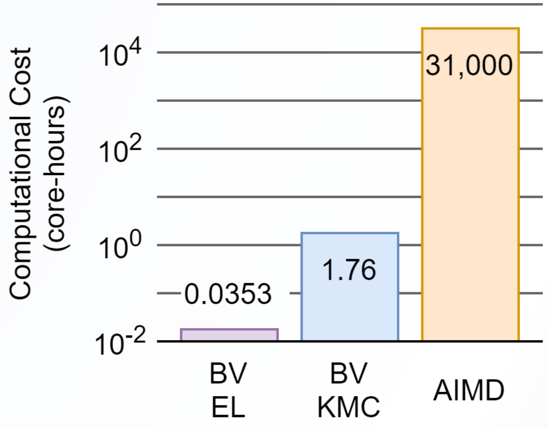




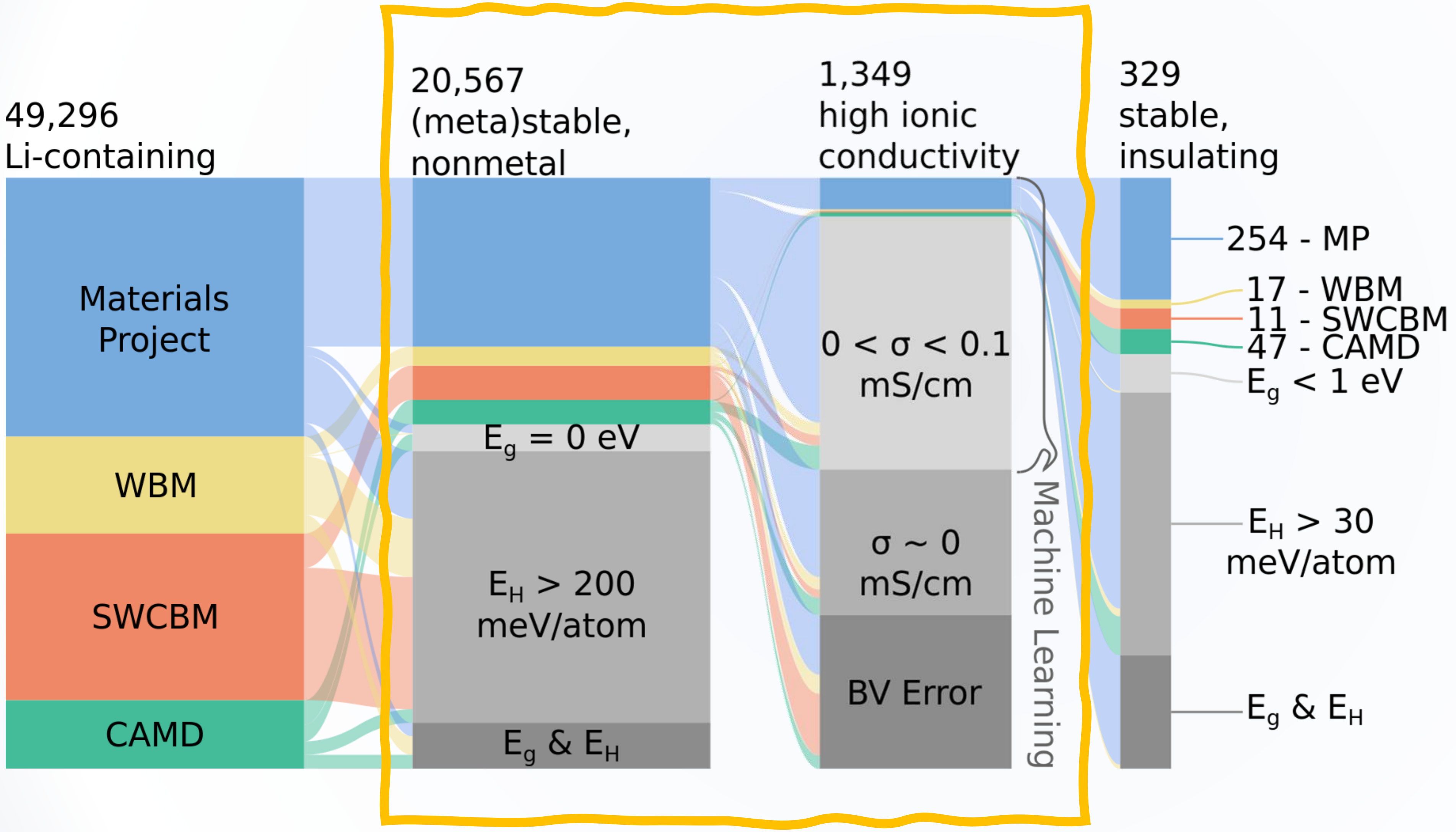
- 1 false positive
- 3 false negatives

- Precision = 0.86
- Recall = 0.67
- *F1* = 0.75

KMC distinguishes between high and low ionic conductivity with suitable precision for screening for 1/10,000th cost of AIMD



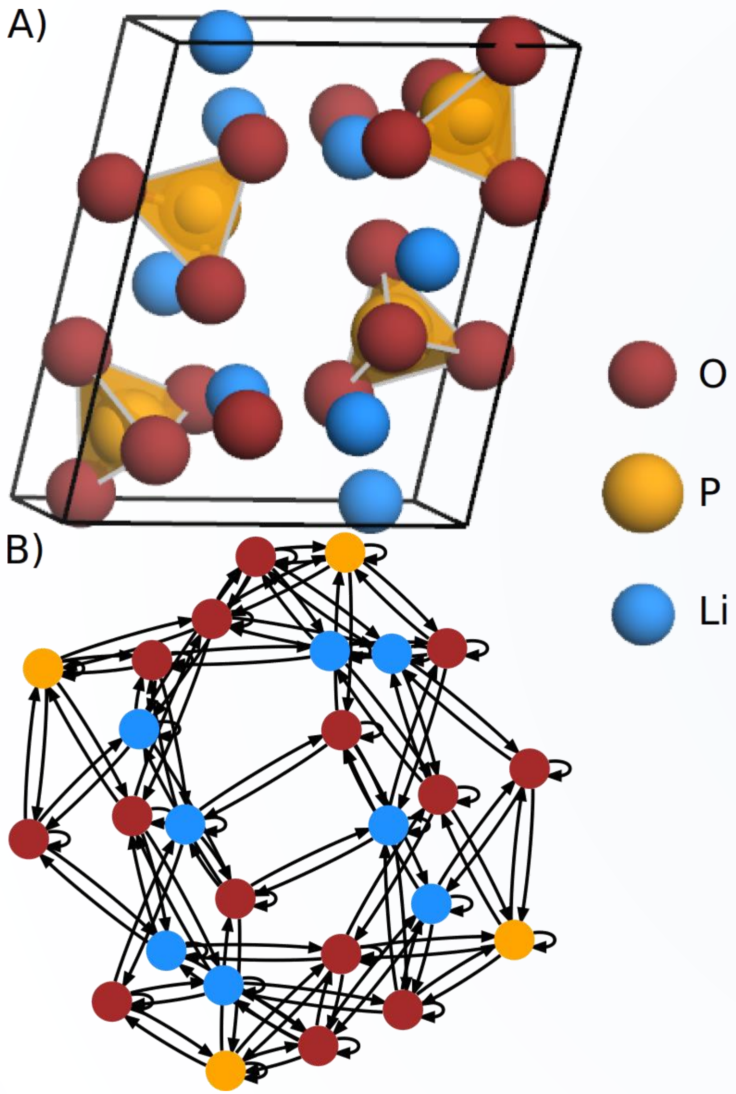
High-throughput screening with softBV



- **49,296** materials from existing databases
- **329** promising candidates with high KMC ionic conductivity
- **10,148** KMC simulations
↓
training data for machine learning

J. Schmidt, H.-C. Wang, T. F. T. Cerqueira, S. Botti, and M. A. L. Marques, "A dataset of 175k stable and metastable materials calculated with the pbesol and scan functionals," Scientific Data 9 (2022).
W. Ye, X. Lei, M. Aykol, and J. H. Montoya, "Novel inorganic crystal structures predicted using autonomous simulation agents," Scientific Data 9 (2022).
J. H. Montoya, K. T. Winther, R. A. Flores, T. Bligaard, J. S. Hummelshj, and M. Aykol, "Autonomous intelligent agents for accelerated materials discovery," Chemical Science 11 (2020).

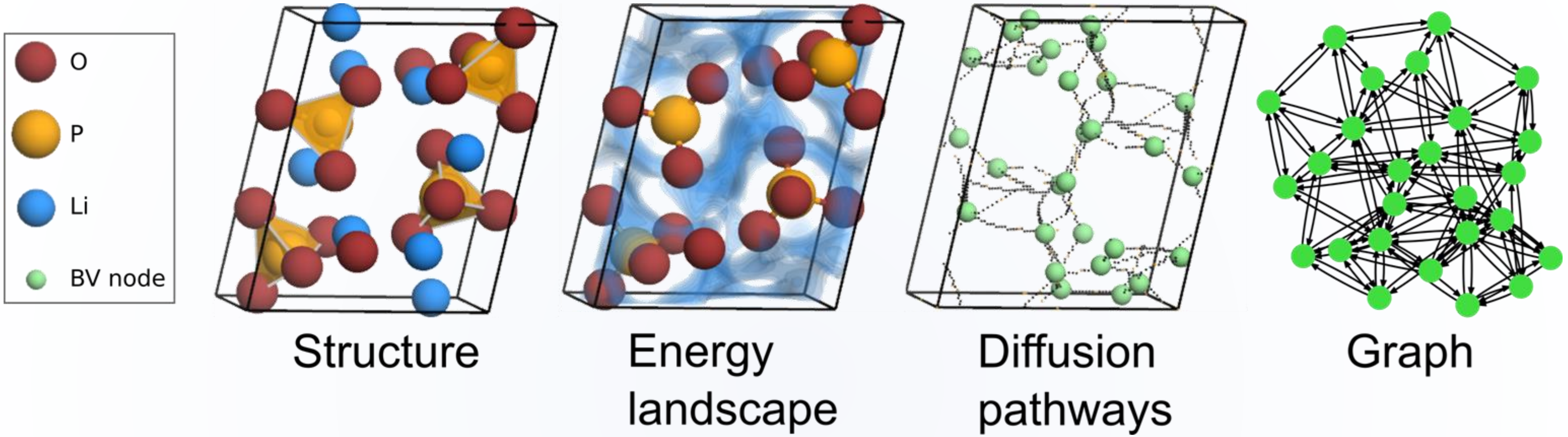
Machine learning ionic conductivity from crystal structure



- Gradient boosting models (GBM)
  - physical quantities e.g. sublattice packing factor
  - compositional quantities e.g. Li atomic fraction
- M3GNet graph neural network
  - node vectors encoding atomic information
  - edge vectors encoding interatomic distances

Chen, C., Ong, S.P. A universal graph deep learning interatomic potential for the periodic table. Nat Comput Sci 2, 718–728 (2022).
Honrao, S.J., Yang, X., Radhakrishnan, B. *et al.* Discovery of novel Li SSE and anode coatings using interpretable machine learning and high-throughput multi-property screening. *Sci Rep* **11**, 16484 (2021).

Machine learning ionic conductivity from BV pathways



- Bond valence conduction graph
  - Nodes: minima in energy landscape
  - Edges: pathway segments
- Same graph topology used for KMC
- Can a graph convolutional network (GCN) replace KMC?

Graph Convolutional Network

- Convolution function propagates information from nodes and edges to neighboring nodes
  - Node vectors v
  - Edge vectors u
  - Learnable weights W and biases B

$$\boldsymbol{v}_i^{(t+1)} = \boldsymbol{v}_i^{(t)} + \sum_{j,k} \sigma\left(\boldsymbol{z}_{(i,j)_k}^{(t)} \boldsymbol{W}_f^{(t)} + \boldsymbol{b}_f^{(t)}\right) \odot g\left(\boldsymbol{z}_{(i,j)_k}^{(t)} \boldsymbol{W}_s^{(t)} + \boldsymbol{b}_s^{(t)}\right)$$

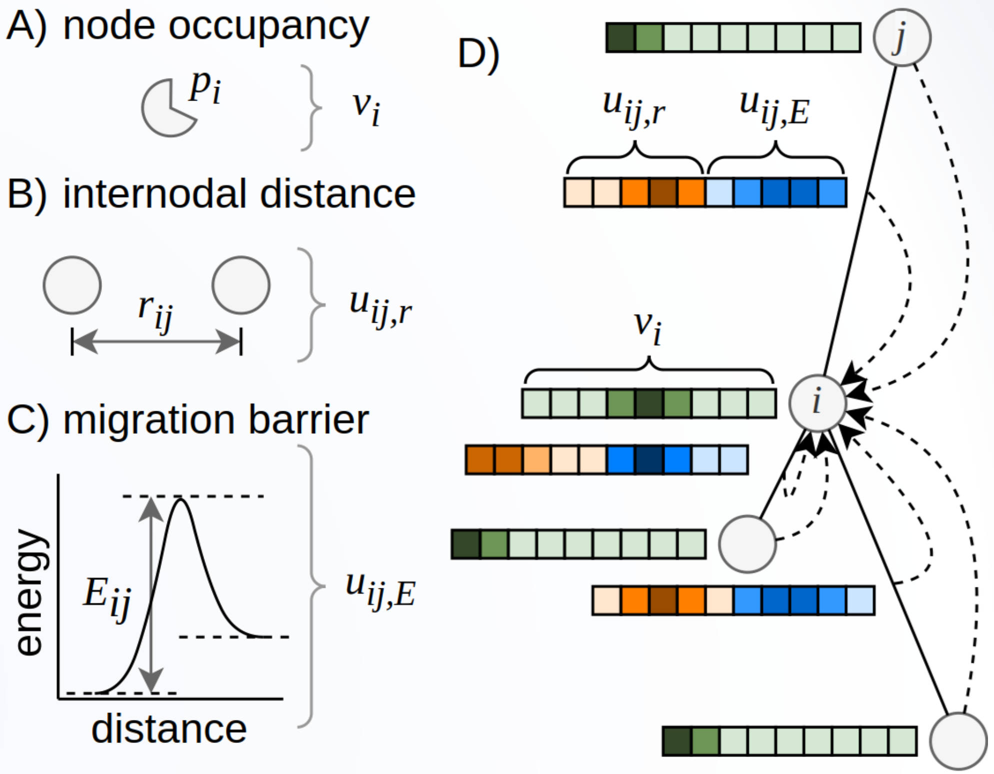
- Vectors are concatenated

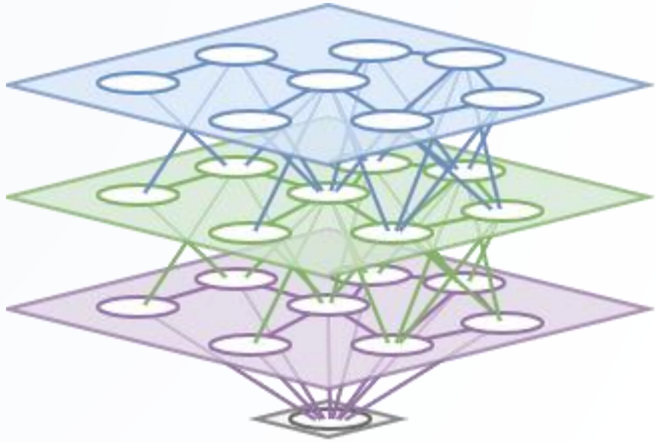
$$\boldsymbol{z}_{(i,j)_k}^{(t)} = \boldsymbol{v}_i^{(t)} \oplus \boldsymbol{v}_j^{(t)} \oplus \boldsymbol{u}_{(i,j)_k,E} \oplus \boldsymbol{u}_{(i,j)_k,r}$$

- Node vectors after R convolutions are averaged, yielding vector representing entire graph

$$\boldsymbol{v}_c = \frac{1}{N} \sum_i^N \boldsymbol{v}_i^{(R)}$$

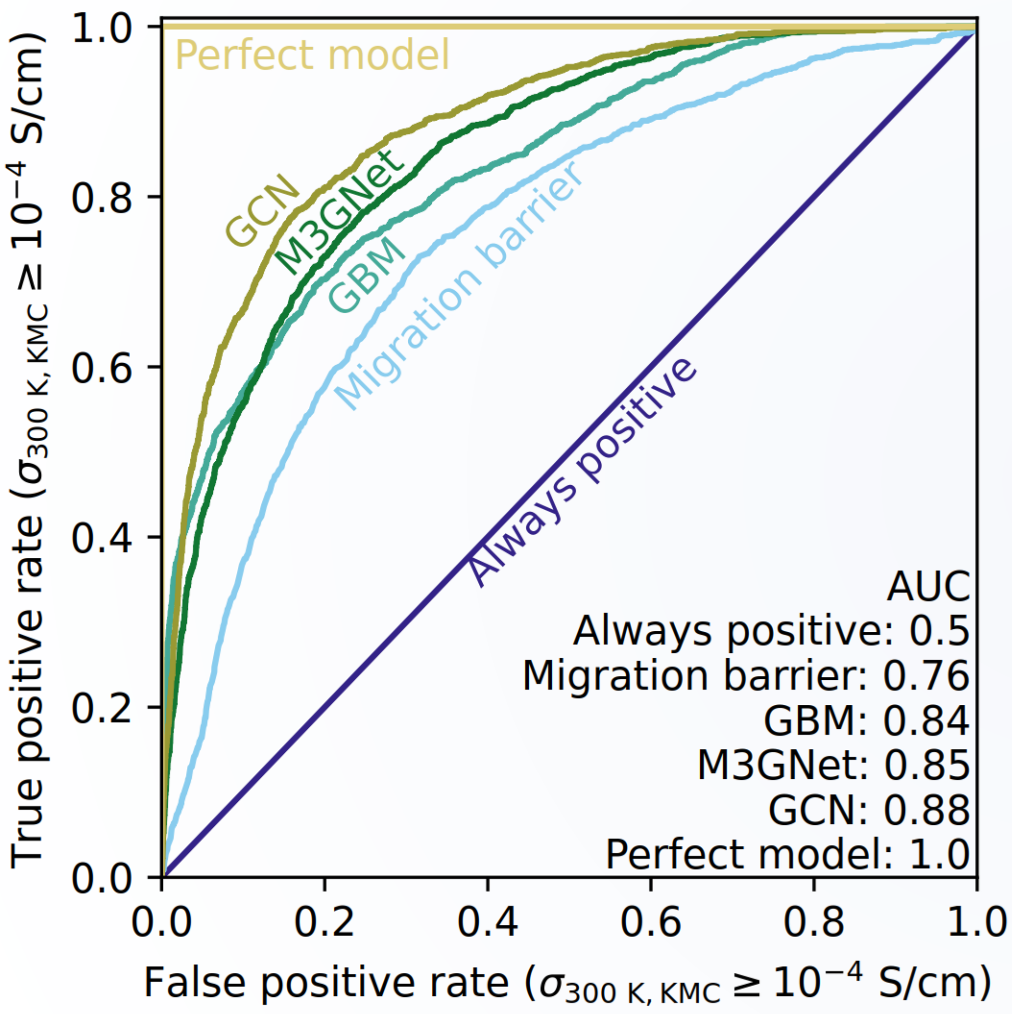
- Vector used as input to feedforward neural network

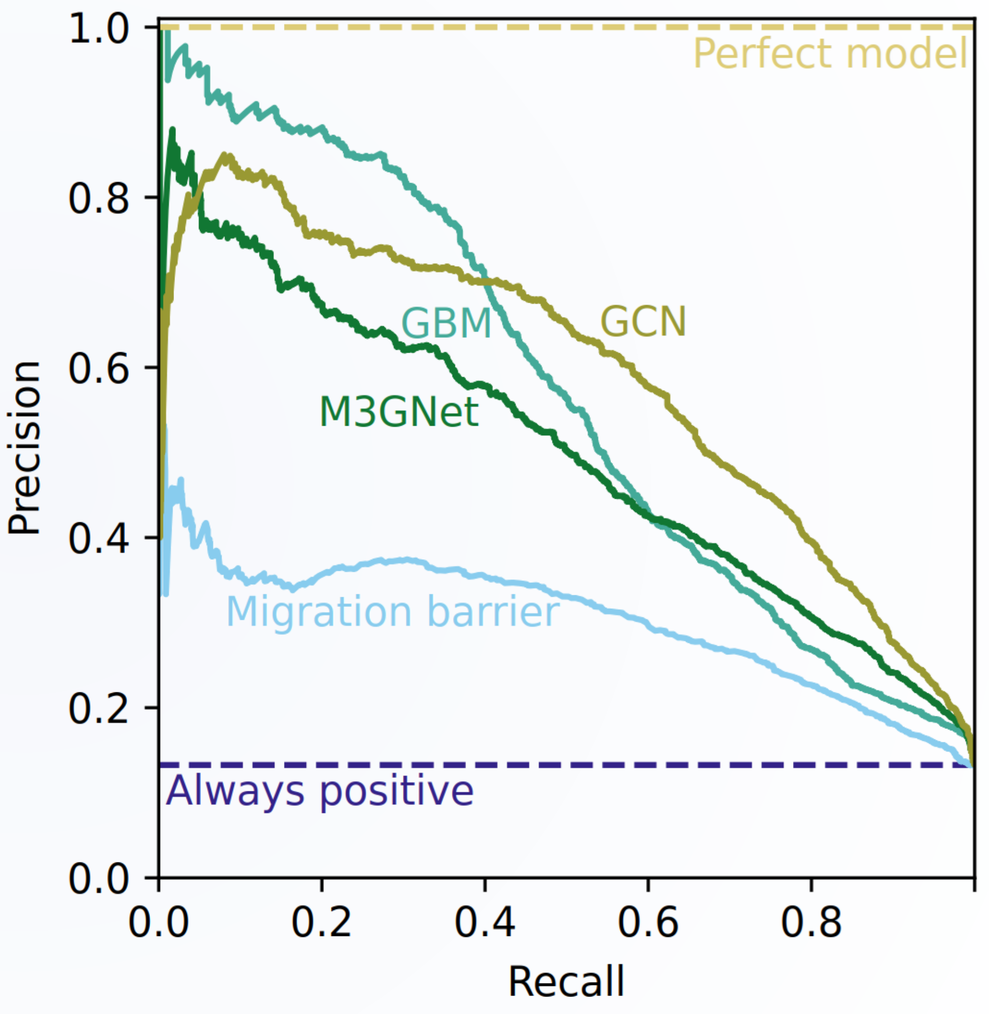




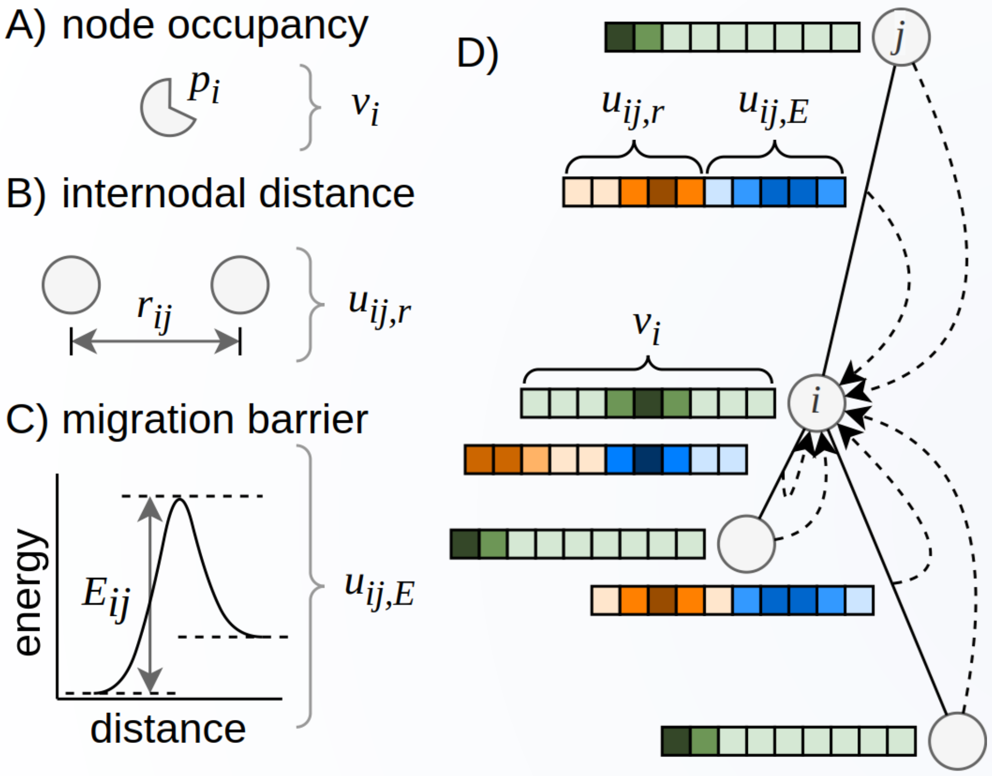
T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," arXiv (2016).
T. Xie and J. C. Grossman, "Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties," Physical Review Letters 120 (2018)

Machine learning ionic conductivity





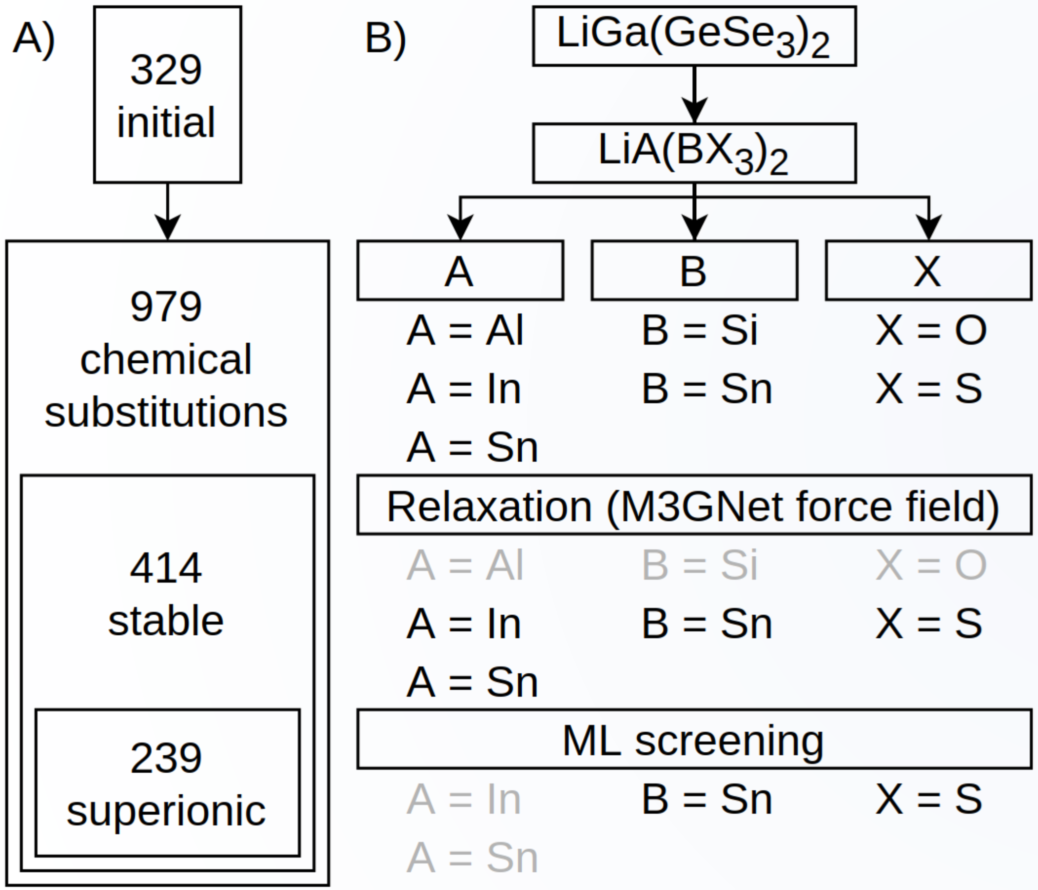
Graph Convolutional Network



Model performance ($\sigma_{300\text{ K, KMC}}$)				
	R^2	MAE	Precision Recall=0.8	Average Precision
Baseline	0.0	3.638	0.133	0.133
E_m linear fit	0.039	3.506	0.227	0.304
M3GNet	0.560	1.937	0.306	0.499
GBM	0.608	2.411	0.268	0.569
GCN	0.691	1.558	0.395	0.587
GCN, no E_{ij}	0.453	2.074	0.290	0.455
GCN, no p_i	0.602	1.794	0.290	0.498
GCN, no r_{ij}	0.587	1.777	0.362	0.557

- Ablation study
  - Site-to-site barriers: affect transition rates
  - Occupancies: fully occupied or completely unoccupied paths do not contribute to ionic conductivity
  - Distances: length of hops contributing to MSD

Generative design and Machine learning



- Isovalent substitution
  - Parent structures from high-throughput screening candidates
  - Substitutions from chemical similarity rules by Wang *et al.*

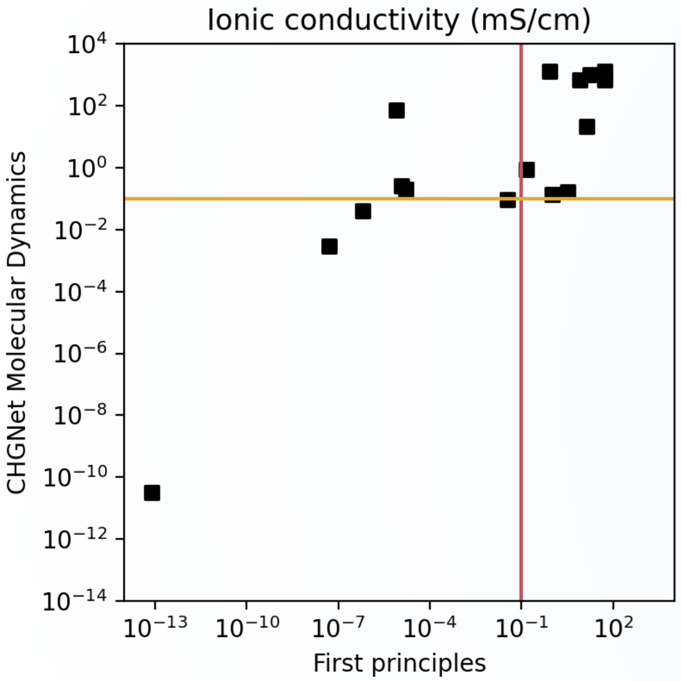
- **979** materials generated from 329 candidates from high-throughput

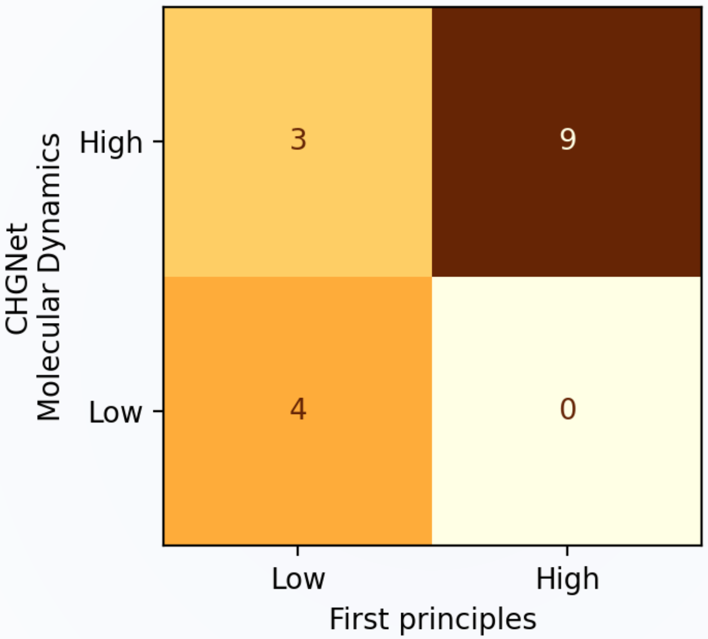
- **239** predicted to have high ionic conductivity by GCN

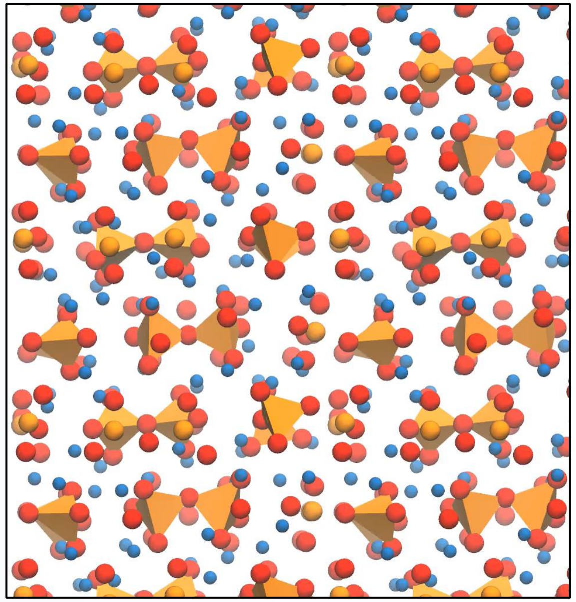
Wang, HC., Botti, S. & Marques, M.A.L. Predicting stable crystalline compounds using chemical similarity. npj Comput Mater 7, 12 (2021).

Benchmarking CHGNet potential for molecular dynamics

CHGNet is a pretrained universal neural network potential.
Molecular dynamics with CHGNet accounts for more physics than bond-valence kinetic Monte Carlo.







- 3 false positives
- 0 false negatives

- Precision = 0.75
- Recall = 1.00
- *F1* = 0.86

- CHGNet outperforms KMC
- 100x faster than *ab initio* molecular dynamics
- 100x slower than KMC

Deng, B., Zhong, P., Jun, K. *et al.* CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nat Mach Intell* **5**, 1031–1041 (2023). https://doi.org/10.1038/s42256-023-00716-3

High-throughput screening

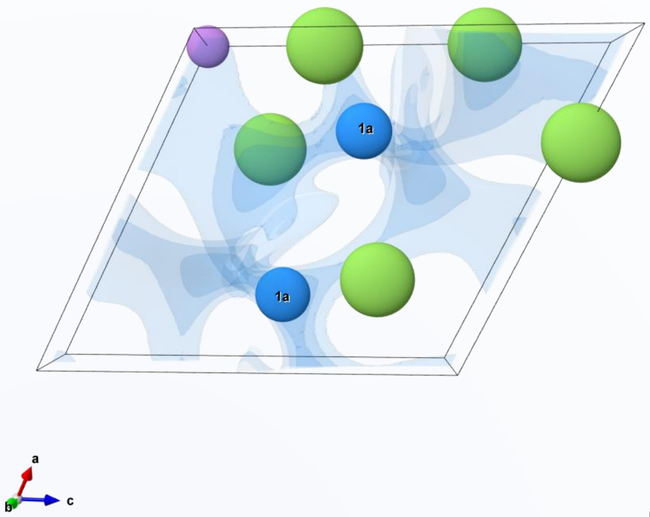
	Source	Ionic Conductivity $\log_{10}(S/cm)$ KMC	Ionic Conductivity $\log_{10}(S/cm)$ CHGNet	Haven Ratio KMC	Haven Ratio CHGNet	Correlated Hopping	Anisotropic Diffusion	Hull Energy (eV/atom)	Reduction Potential (V)	ECS Window (V)
Li_2SbF_5	WBM	0.80	-3.42	0.004	0.639		✓	0.000	2.738	0.672
Li_2AsF_5	WBM	-0.12	-3.39	0.004			✓	0.024	3.030	0.639
$LiZnPSe_4$	WBM	-1.11	-0.35	0.544	0.376	✓		0.001	2.119	0.062
$LiHgPSe_4$	WBM	-1.39	-0.25	0.626		✓		0.003		
$Li_7Cl_3O_2$	WBM	-3.41		0.428		✓		0.028	0.000	2.767
$LiBSe_2$	SWCBM	-3.58	-1.05	0.358		✓		0.000	1.802	0.018
$LiCuBr_2$	SWCBM	-3.98	-3.04	0.570				0.004	2.637	0.054
Li_6PBrO_5	mp-1103940	-3.16	-3.62	0.235	0.564	✓		0.017	0.686	2.087
$Li_2Ta_2O_3F_6$	mp-561011	-3.69	-17.91	0.013	0.789	✓		0.000	1.701	2.920
$Li_{10}Mg_7Cl_{24}$	mp-530738	-3.89	-1.62	0.009	0.205	✓		0.009	0.886	2.902

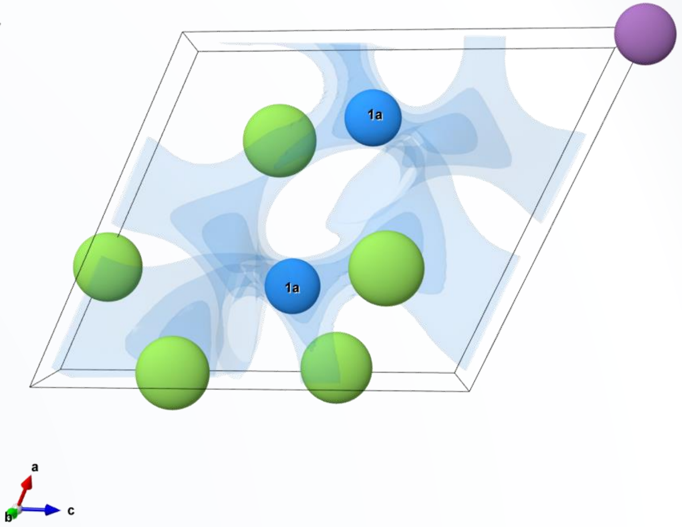
Machine learning

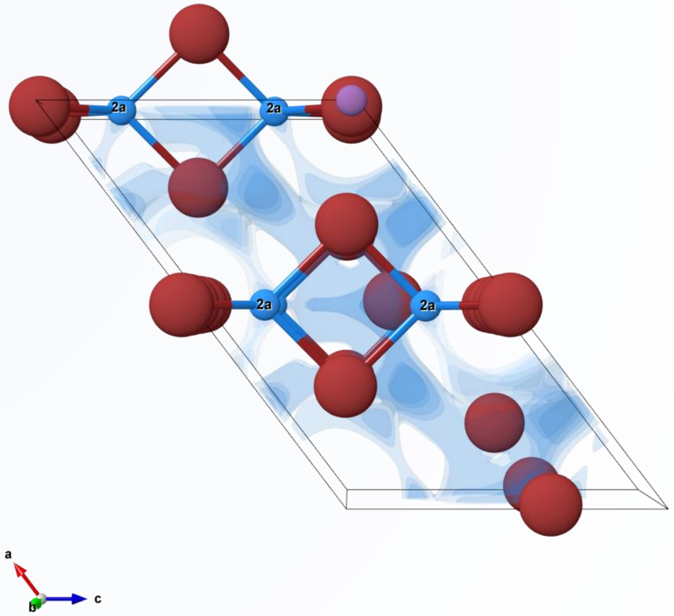
	Source	Ionic Conductivity $\log_{10}(S/cm)$ KMC	Ionic Conductivity $\log_{10}(S/cm)$ CHGNet	Haven Ratio KMC	Haven Ratio CHGNet	Correlated Hopping	Anisotropic Diffusion	Hull Energy (eV/atom)	Reduction Potential (V)	ECS Window (V)
Li_2SbBr_5	WBM	1.03	-0.01	0.009	0.556		✓	0.028	2.327	0.961
Li_5I_2N	WBM	-0.99	-2.14	0.318	0.803	✓	✓	0.000	0.000	0.477
$LiGaF_4$	WBM	-2.86	-5.80	0.291		✓	✓	0.009	2.578	3.614
$LiCuI_2$	SWCBM	-2.96	-0.61	0.550				0.007	2.159	0.414
$LiGaSi_4O_{10}$	CAMD	-3.23		0.065	0.639	✓	✓	0.023	1.814	1.950
$Li_3AlB_2O_6$	CAMD	-3.72		0.097	0.210			0.027	1.072	2.226
$LiBI_4$	MatProj	-0.09		0.014	0.873		✓	0.000	2.366	0.260
$LiMn_2I_9$	MatProj	-0.62		0.046		✓	✓	0.000	2.410	2.427
$LiVBr_5$	MatProj	-0.96	0.02	0.286	0.385	✓	✓	0.000	2.646	2.219
Li_2TiF_6	MatProj	-3.32	-1.27	0.073	0.246	✓	✓	0.026	2.130	4.170

Li_2BF_5 analogs

High-throughput	Machine Learning
$\boxed{Li_2SbF_5}$	$\boxed{Li_2SbBr_5}$
$\boxed{Li_2AsF_5}$	Li_5I_2N
$LiZnPSe_4$	$LiGaF_4$
$LiHgPSe_4$	$LiCuI_2$
$Li_7Cl_3O_2$	$LiGaSi_4O_{10}$
$LiBSe_2$	$Li_3AlB_2O_6$
$LiCuBr_2$	$LiBI_4$
Li_6PBrO_5	$LiMn_2I_9$
$Li_2Ta_2O_3F_6$	$LiVBr_5$
$Li_{10}Mg_7Cl_{24}$	Li_2TiF_6

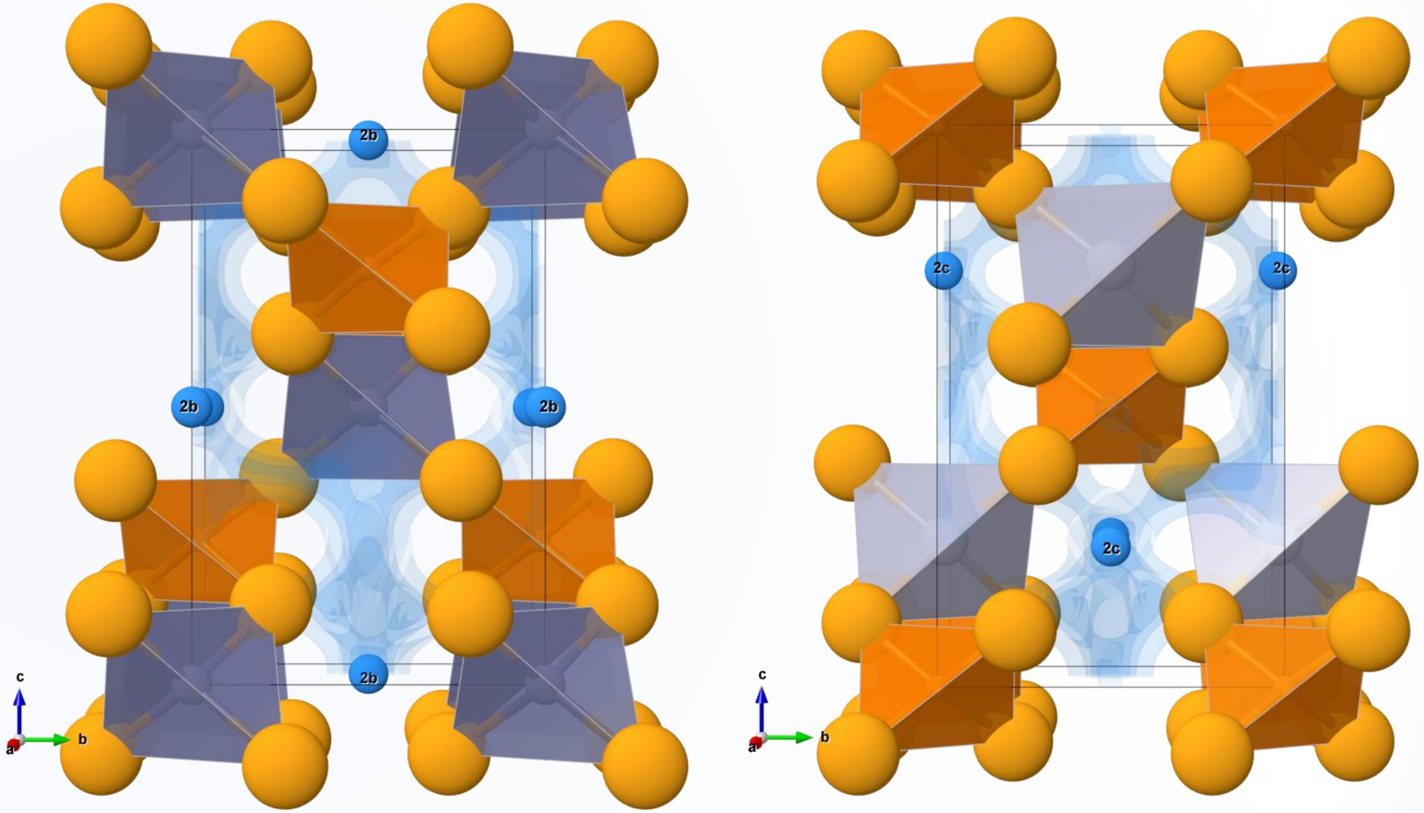






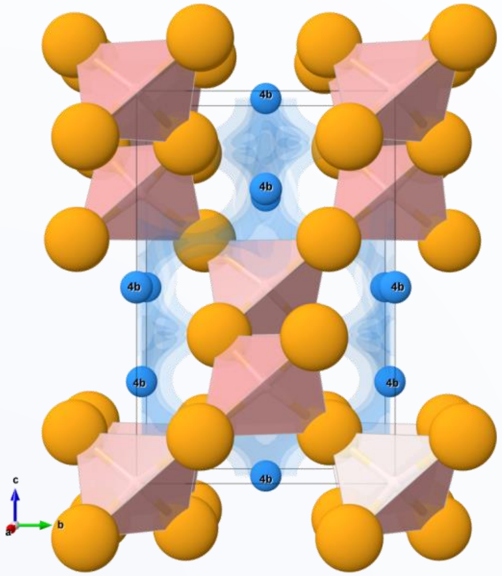
$LiZnPS_4$ analogs

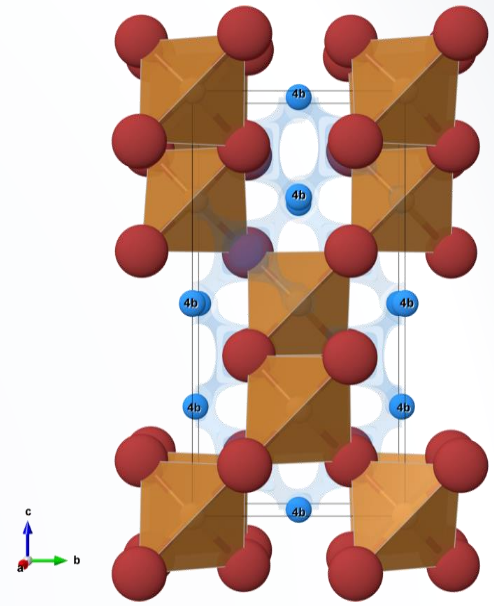
High-throughput	Machine Learning
Li_2SbF_5	Li_2SbBr_5
Li_2AsF_5	Li_5I_2N
$\boxed{LiZnPSe_4}$	$LiGaF_4$
$\boxed{LiHgPSe_4}$	$LiCuI_2$
$Li_7Cl_3O_2$	$LiGaSi_4O_{10}$
$LiBSe_2$	$Li_3AlB_2O_6$
$LiCuBr_2$	$LiBI_4$
Li_6PBrO_5	$LiMn_2I_9$
$Li_2Ta_2O_3F_6$	$LiVBr_5$
$Li_{10}Mg_7Cl_{24}$	Li_2TiF_6

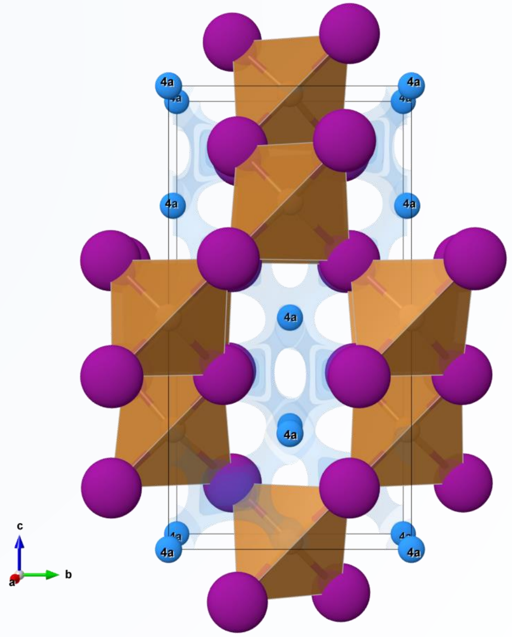


Wurtzites

High-throughput	Machine Learning
Li_2SbF_5	Li_2SbBr_5
Li_2AsF_5	Li_5I_2N
$LiZnPSe_4$	$LiGaF_4$
$LiHgPSe_4$	$\boxed{LiCuI_2}$
$Li_7Cl_3O_2$	$LiGaSi_4O_{10}$
$\boxed{LiBSe_2}$	$Li_3AlB_2O_6$
$\boxed{LiCuBr_2}$	$LiBI_4$
Li_6PBrO_5	$LiMn_2I_9$
$Li_2Ta_2O_3F_6$	$LiVBr_5$
$Li_{10}Mg_7Cl_{24}$	Li_2TiF_6

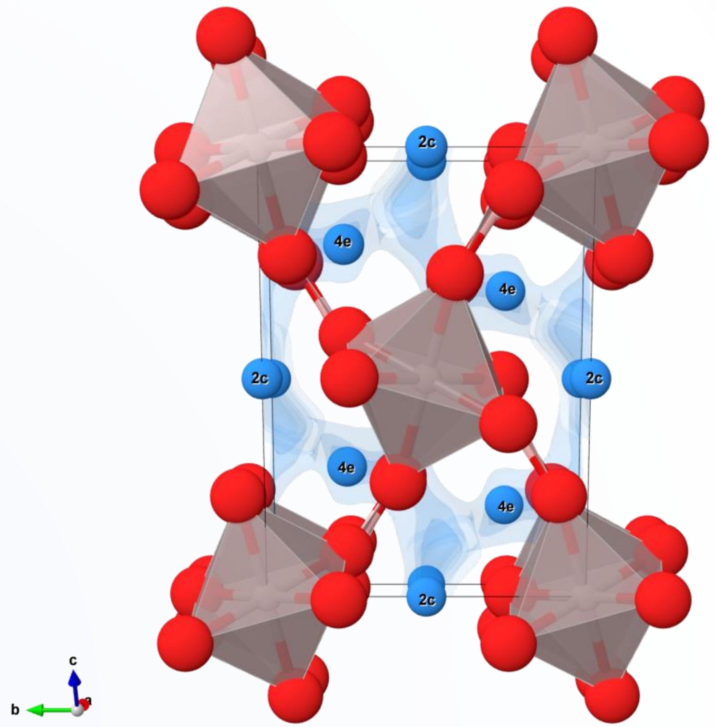


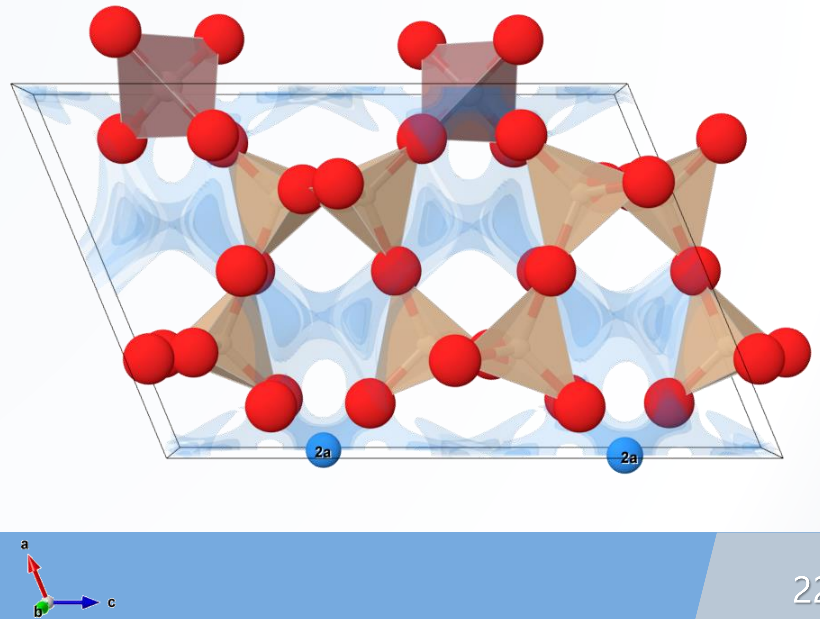




Oxides

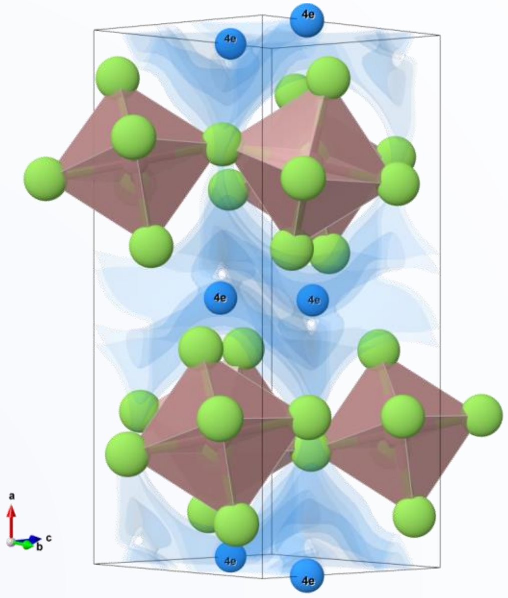
High-throughput	Machine Learning
Li_2SbF_5	Li_2SbBr_5
Li_2AsF_5	Li_5I_2N
$LiZnPSe_4$	$LiGaF_4$
$LiHgPSe_4$	$LiCuI_2$
$Li_7Cl_3O_2$	$\boxed{LiGaSi_4O_{10}}$
$LiBSe_2$	$\boxed{Li_3AlB_2O_6}$
$LiCuBr_2$	$LiBI_4$
Li_6PBrO_5	$LiMn_2I_9$
$Li_2Ta_2O_3F_6$	$LiVBr_5$
$Li_{10}Mg_7Cl_{24}$	Li_2TiF_6

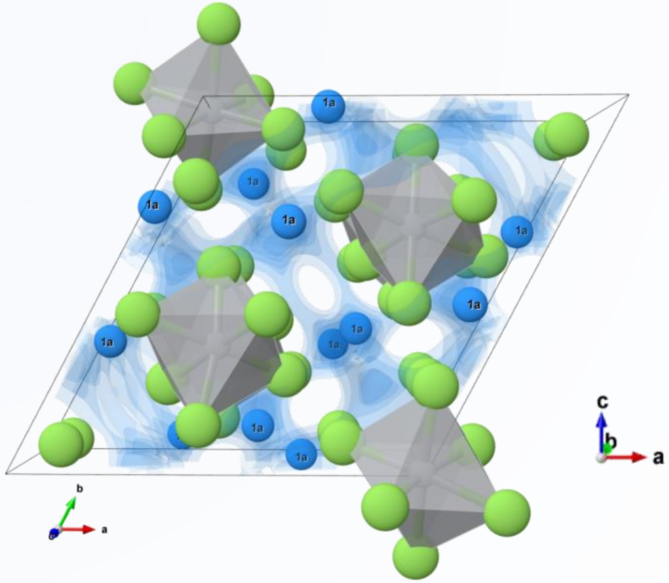


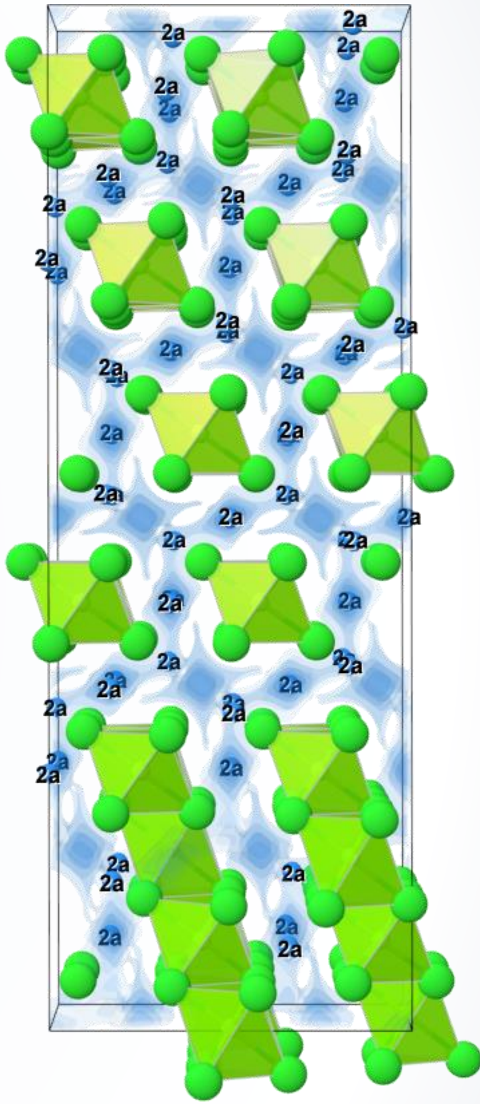


Fluorides, Chlorides

High-throughput	Machine Learning
Li_2SbF_5	Li_2SbBr_5
Li_2AsF_5	Li_5I_2N
$LiZnPSe_4$	$\boxed{LiGaF_4}$
$LiHgPSe_4$	$LiCuI_2$
$Li_7Cl_3O_2$	$LiGaSi_4O_{10}$
$LiBSe_2$	$Li_3AlB_2O_6$
$LiCuBr_2$	$LiBI_4$
Li_6PBrO_5	$LiMn_2I_9$
$Li_2Ta_2O_3F_6$	$LiVBr_5$
$\boxed{Li_{10}Mg_7Cl_{24}}$	$\boxed{Li_2TiF_6}$

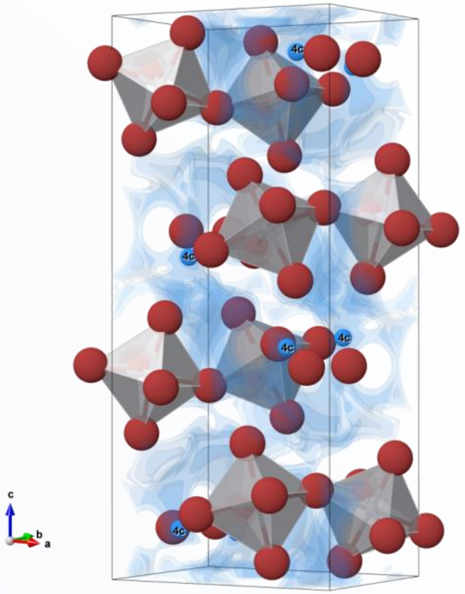


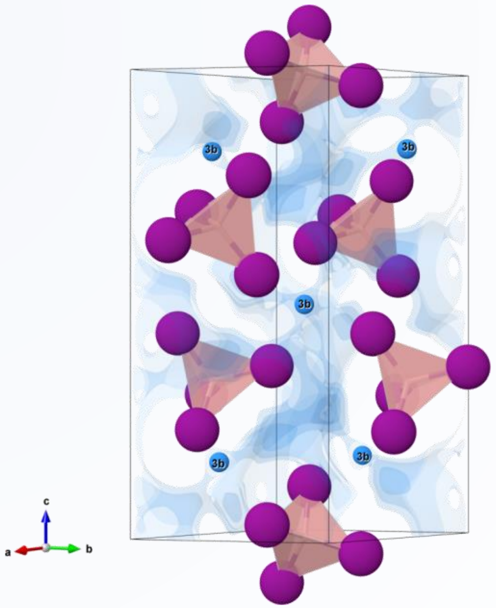


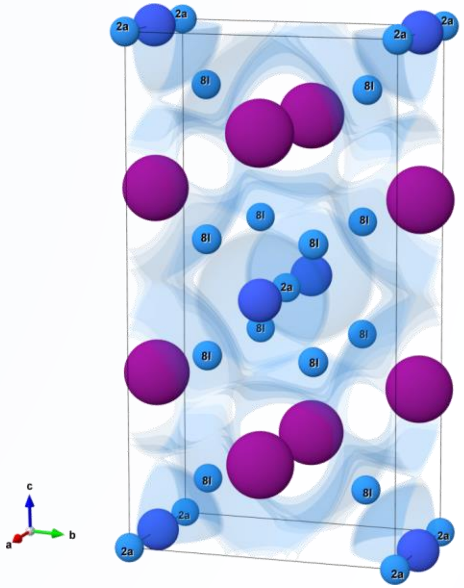


Bromides, Iodides

High-throughput	Machine Learning
Li_2SbF_5	Li_2SbBr_5
Li_2AsF_5	$\boxed{Li_5I_2N}$
$LiZnPSe_4$	$LiGaF_4$
$LiHgPSe_4$	$LiCuI_2$
$Li_7Cl_3O_2$	$LiGaSi_4O_{10}$
$LiBSe_2$	$Li_3AlB_2O_6$
$LiCuBr_2$	$\boxed{LiBI_4}$
Li_6PBrO_5	$LiMn_2I_9$
$Li_2Ta_2O_3F_6$	$\boxed{LiVBr_5}$
$Li_{10}Mg_7Cl_{24}$	Li_2TiF_6

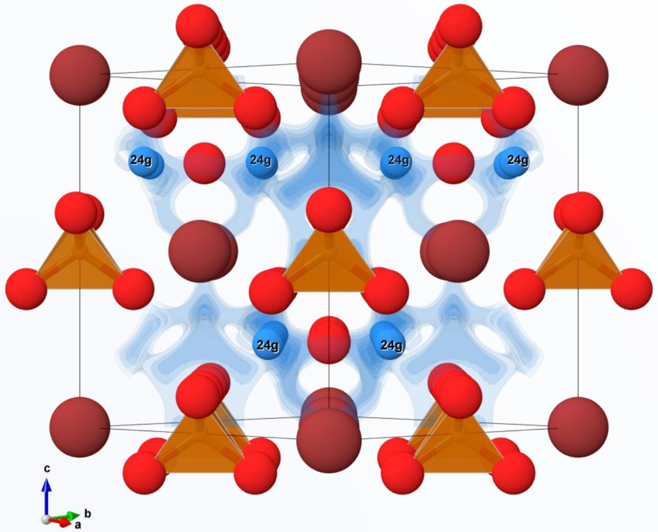


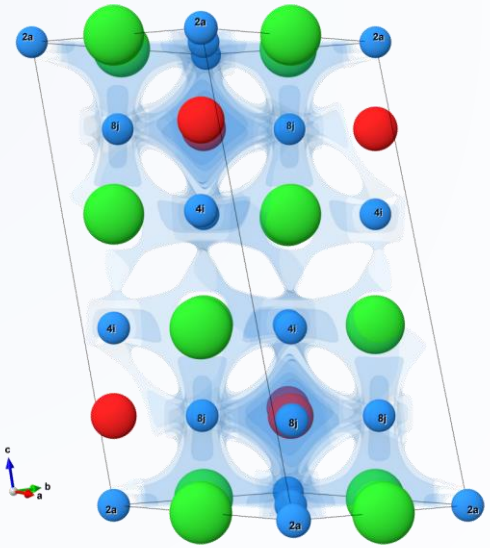


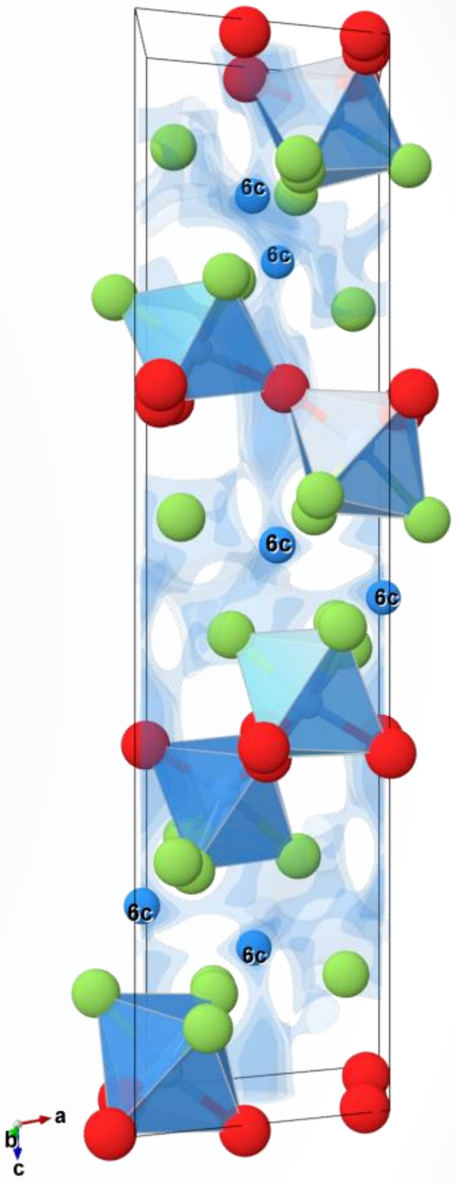


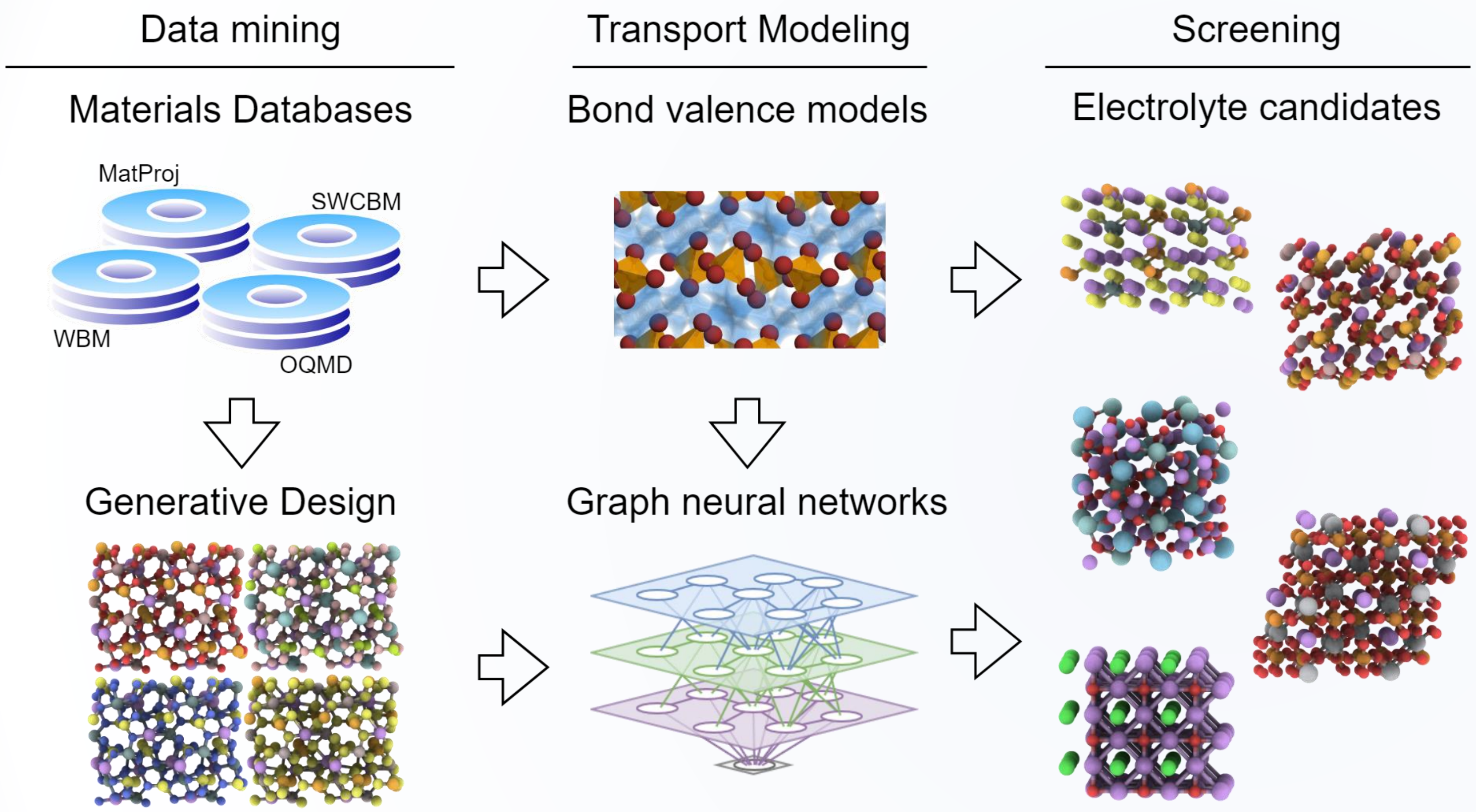
Oxohalides

High-throughput	Machine Learning
Li_2SbF_5	Li_2SbBr_5
Li_2AsF_5	Li_5I_2N
$LiZnPSe_4$	$LiGaF_4$
$LiHgPSe_4$	$LiCuI_2$
$\boxed{Li_7Cl_3O_2}$	$LiGaSi_4O_{10}$
$LiBSe_2$	$Li_3AlB_2O_6$
$LiCuBr_2$	$LiBI_4$
$\boxed{Li_6PBrO_5}$	$LiMn_2I_9$
$\boxed{Li_2Ta_2O_3F_6}$	$LiVBr_5$
$Li_{10}Mg_7Cl_{24}$	Li_2TiF_6





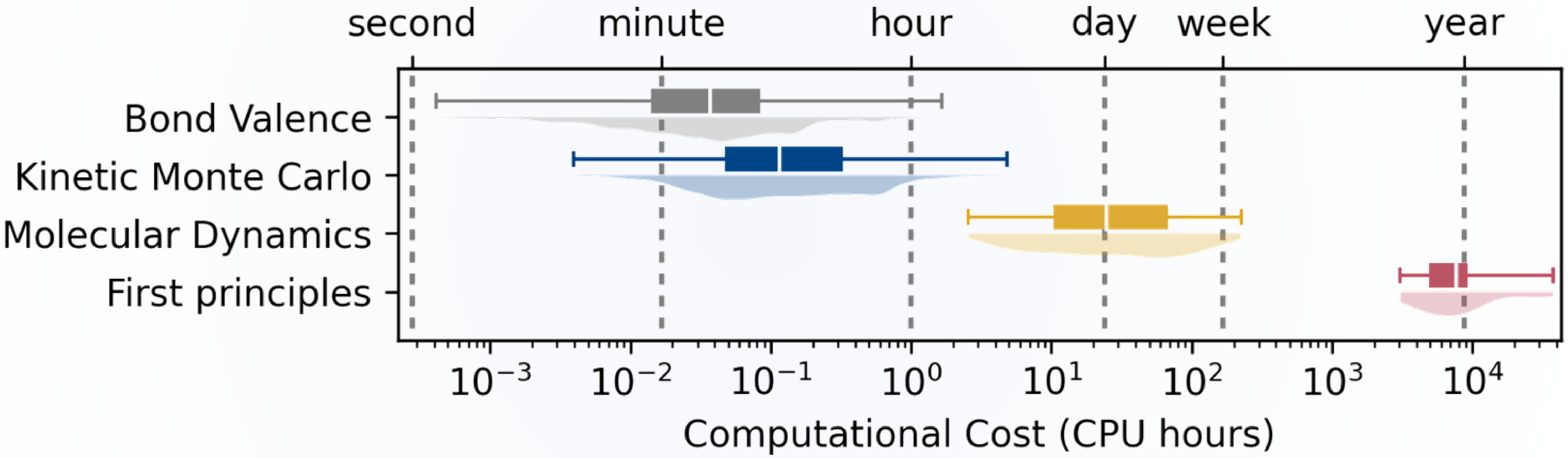




AMES RESEARCH CENTER

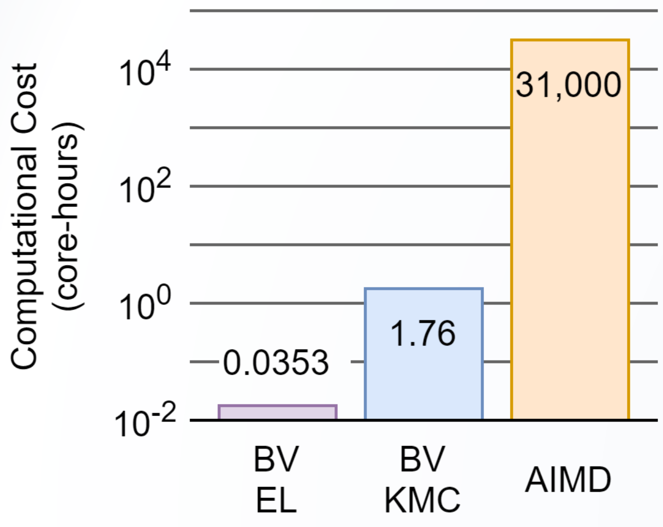
- Transformational Tools & Technology (TTT)
- Solid-state Architecture Batteries for Enhanced Rechargeability and Safety (SABERS)
- Ames Research Center
  - Junsoo Park
  - Joakim Stenlid
  - Zhigang Wu
  - Mohit Mehta
- West Virginia University
  - David Mebane

Computational cost

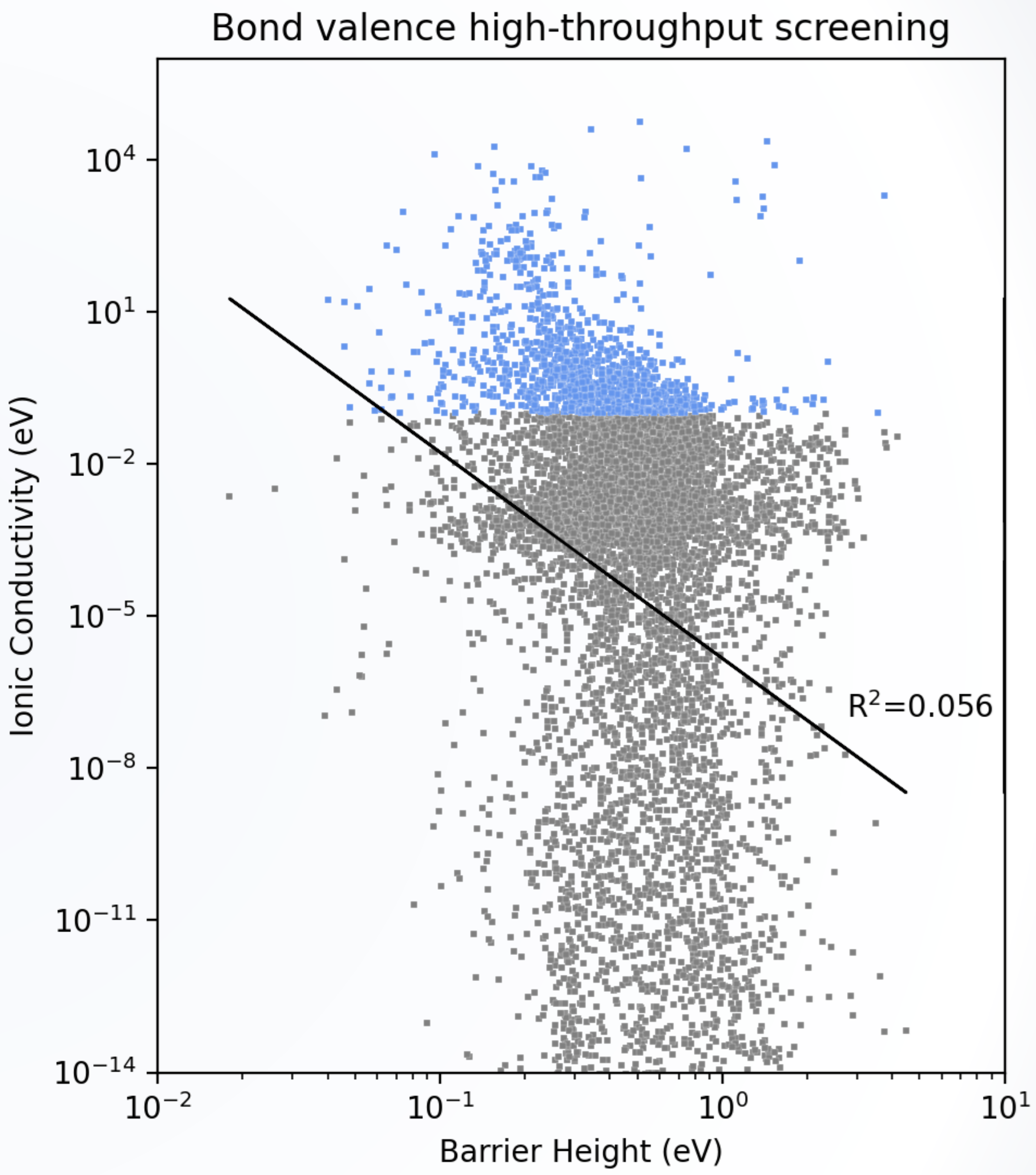


	Energy landscape	Kinetic Monte Carlo	Molecular dynamics	First principles
Single-lithium picture	X			
Li^+ - Li^+ interactions		X	X	X
Non-lithium motion			X	X
Ionic and electronic DoF				X

Kinetic Monte Carlo with softBV



$$p = \omega \exp\left(\frac{E_m + \Delta E}{k_B T}\right)$$



- Energy landscape is a single-Li^+ picture
  - ignores Li^+-Li^+ interactions
  - ionic conductivity depends on mixture of occupied and unoccupied sites in pathways
- Transition rate also depends on vibration frequencies (pre-exponential factor)

Validation test of GCN trained on Kinetic Monte Carlo