

Supplemental Information

S1. Feature calculations.

This section describes the methods used for calculating the 20 atomistic features given in Table 2. To account for periodic boundary conditions, feature extraction is performed for all atoms within a $1 \times 1 \times 1$ computational cell at the center of a $3 \times 3 \times 3$ supercell. Features **1.6**, **1.7**, **1.9**, **1.12**, and **1.13** described here comprise the ionic conductivity classification model with regression coefficients given in Table 2. For reference, the values of all features for lithium iodide and lithium phosphide are given in Table S3.

S1.1 Average atomic volume, AAV (\AA^3)

Sum the number of atoms in the unit cell N_{cell} and divide by the unit cell volume V_{cell} .

$$AAV = \frac{N_{\text{cell}}}{V_{\text{cell}}} \quad (\text{S1})$$

S1.2 Standard deviation in Li neighbor count, $SDLC$ (dimensionless)

For each of the N_{Li} lithium atoms in the unit cell, count the number of atoms within a distance $r_{ij} < 4 \text{\AA}$ in any direction. This distance may include nearest and greater distance neighbor atoms. Take the standard deviation of this value over all Li atoms in the unit cell.

$$SDLC = \sqrt{\frac{\sum_{i \in \{\text{Li}\}}^{N_{\text{Li}}} \left(\sum_{j=1}^{N_{\text{atoms}}} \mathbf{1}\{r_{ij} < 4\text{\AA}\} - LNC \right)^2}{N_{\text{Li}} - 1}} \quad (\text{S2})$$

where LNC is the average lithium neighbor count, given in S1.5. Note $\mathbf{1}\{X\}$ is the indicator function, which evaluates to 1 if the argument X is true and 0 if X is false.

S1.3 Standard deviation in Li bond ionicity, $SDLI$ (dimensionless)

Find all lithium bonds in the crystal, defined as any connection between Li and another atom separated by less than 4\AA . For each Li-X bond i in the unit cell,

evaluate the Pauling electronegativity^{1,2} difference (“ionicity”) between Li and its neighbor, $I_i = |EN_{Li} - EN_X| \mathbf{1}\{r_{Li-X} \leq 4\text{\AA}\}$. Take the standard deviation of these N_{bonds} bonds.

$$SDLI = \sqrt{\frac{\sum_{i \in \{Li-X\}}^{N_{bonds}} (I_i - \bar{I})^2}{N_{bonds} - 1}} \quad (S3)$$

S1.4 Average Li bond ionicity, LBI (dimensionless)

Take the average electronegativity difference of all N_{bonds} Li-X bonds, as described in S1.3.

$$LBI = \frac{\sum_{i \in \{Li-X\}}^{N_{bonds}} I_i}{N_{bonds} - 1} \quad (S4)$$

S1.5 Average Li neighbor count, LNC (dimensionless)

Take the average of the distribution described in S1.2, i.e. take the average of each Li atom’s 4Å neighbor count.

$$LNC = \frac{1}{N_{Li}} \sum_{Li \in \{Li\}}^{N_{Li}} \sum_{j \neq i}^{N_{atoms}} \mathbf{1}\{r_{ij} \leq 4\text{\AA}\} \quad (S5)$$

Note this is not necessarily the lithium coordination, as the cutoff of 4Å may reach into second nearest neighbors. This relatively large value of cutoff is chosen to minimize the number of cases for which this feature evaluates to zero.

S1.6 Average Li-Li bonds per Li, LLB (dimensionless)

For each Li atom in the unit cell, count the number of Li atoms within 4Å. Average this value over all Li in the unit cell.

$$LLB = \frac{1}{N_{Li}} \sum_{Li_i \in \{Li\}}^{N_{Li}} \sum_{Li_j \in \{Li\}, j \neq i}^{N_{Li}} \mathbf{1}\{r_{ij} \leq 4\text{\AA}\} \quad (S6)$$

S1.7 Average bond ionicity of sublattice, SBI (dimensionless)

Similar to S1.4 but for all non-Li atoms in the lattice; e.g. loop through all non-Li atoms (X) in the unit cell, and find all X-Y bonds, where X and Y are separated by at most 4Å and X is not Li but Y may be Li. Calculate the bond ionicity of all N_{bonds} of these bonds and take the average:

$$SBI = \frac{\sum_{i \in \{X-Y\}}^{N_{bonds}} I_i}{N_{bonds}} \quad (S7)$$

S1.8 Average sublattice neighbor count, SNC (dimensionless)

Similar to the Li neighbor count (S1.5) but for all non-Li atoms in the lattice. For all atoms in the unit cell except any Li atoms, count the number of neighboring atoms (including Li) within 4Å. Average this value over all non-Li atoms.

$$SNC = \frac{1}{N_{atoms} - N_{Li}} \sum_{i \notin \{Li\}}^{N_{atoms} - N_{Li}} \sum_{j \neq i}^{N_{atoms}} 1\{r_{ij} \leq 4\text{Å}\} \quad (S8)$$

S1.9 Anion framework coordination, AFC (dimensionless)

First, find the atom in the lattice with the highest Pauling electronegativity value and designate this as the lattice anion, A . Ignoring all other atoms in the lattice, calculate the coordination of the anions with each other. This is done with the following procedure: for each anion in the lattice i , find the nearest anion j at a distance r_{ij_0} . Then count the number of anions in the supercell within a distance r_{ij} from atom i such that $r_{ij_0} \leq r_{ij} \leq r_{ij_0} + 1\text{Å}$. Average this over all anions in the unit cell.

$$AFC = \frac{\sum_{i \in \{A\}}^{N_A} \sum_{j \in \{A\}_{PBC}}^{N_{A,PBC}} 1\{r_{ij_0} \leq r_{ij} \leq r_{ij_0} + 1\text{Å}\}}{N_A} \quad (S9)$$

S1.10 Average shortest anion-anion separation distance, AASD (Å)

After finding the anion-anion coordination for each anion i , take the average distance to its nearest neighbors, NN_i . Average this distance over all N_A anions in the unit cell.

$$AASD = \frac{\sum_{i \in \{A\}}^{N_A} NN_i}{N_A} \quad (S10)$$

S1.11 Volume per anion, VPA (Å³)

Divide the unit cell volume V_{cell} by the total number of anions in the cell N_A .

$$VPA = V_{\text{cell}} / N_A \quad (S11)$$

S1.12 Average shortest Li-anion separation distance, LASD (Å)

For each of the Li atoms in the cell, find the distance r_{ij} to the nearest anion A in the supercell. Take the average of these values over all N_{Li} Li atoms.

$$LASD = \frac{\sum_{i \in \{\text{Li}\}}^{N_{\text{Li}}} \min_{j \in \{A\}} \{r_{ij}\}}{N_{\text{Li}}} \quad (S12)$$

S1.13 Average shortest Li-Li separation distance, LLSD (Å)

For each Li atom in the unit cell, find the Euclidean distance to the nearest Li atom in the supercell. Take the average value of this distribution over all Li atoms.

$$LLSD = \frac{\sum_{i \in \{\text{Li}\}}^{N_{\text{Li}}} \min_{j \in \{\text{Li}\}, j \neq i} \{r_{ij}\}}{N_{\text{Li}}} \quad (S13)$$

S1.14 Average electronegativity of sublattice, ENS (dimensionless)

Take the average of the electronegativity of all atoms $\{EN\}$ in the unit cell excluding Li.

$$ENS = \frac{\sum_{i \notin \{Li\}}^{N_{atoms} - N_{Li}} EN_i}{N_{atoms} - N_{Li}} \quad (S14)$$

S1.15 Packing fraction of full crystal, PF (dimensionless)

For each atom, first choose an atomic radius based on the bonding environment. If the bond ionicity of the atom i , $(END)_i$ if greater than 2, choose the ionic radius in Table S1; if the bond ionicity is less than 2, choose the covalent radius. Once all effective radii r_i^{eff} have been assigned, the packing fraction is calculated using a quadrature method. We populate the unit cell with points chosen uniformly at random, and then calculate the ratio of points within the effective atomic radii of any atom to the total number of points. This ratio is the packing fraction. The set of points $\{r_{ES}\}$ within the effective radius of any atom is the set of

points for which $\prod_{j=1}^{N_{atoms,PBC}} 1\{r_{ij} > r_j^{eff}\} = 0$. We use a quadrature approach to account for possible overlap between atoms.

$$PF = \frac{\sum_{i=1}^{N_{points}} [1 - \prod_{j=1}^{N_{atoms,PBC}} 1\{r_{ij} > r_j^{eff}\}]}{N_{points}} \quad (S15)$$

Randomly located points are added to the unit cell until convergence is reached. The packing fraction is evaluated after every addition of 1,000 new points; convergence is reached when the packing fraction changes by less than 1% between successive evaluations.

S1.16 Packing fraction of sublattice, SPF (dimensionless)

Similar to S1.16, but first remove all Li atoms from the unit cell.

$$PF = \frac{\sum_{i=1}^{N_{points}} [1 - \prod_{j \notin \{Li\}}^{N_{atoms,PBC} - N_{Li,PBC}} 1_{\{r_{ij} > r_j^{eff}\}}]}{N_{points}} \quad (S16)$$

S1.17 Average straight-line path width, *SLPW* (Å)

This feature finds the maximum radius of the cylinder that connects each lithium with its nearest lithium neighbor. For each Li atom i , first find the nearest Li atom within the supercell, as performed in S1.14. Draw a straight line connecting the two Li atoms. Loop through all other atoms in the supercell and project them onto the connecting line. If atom j does not project onto the line segment between the two Li atoms, discard atom j . Of the remaining atoms, evaluate the distance between the connecting line and the edge of the hard sphere defined by that atom (using the effective radius of each atom, determined as described in S1.16). If an atom is blocking the pathway completely, assign the distance to zero. Find the atom that gives the smallest distance. This distance, which represents the radius of the largest cylinder that connects Li atom i and its nearest neighbor j , makes up the contribution to the *SLPW* feature for Li atom i . To calculate the *SLPW* feature value, evaluate this distance for each Li atom and take the average. A schematic of this calculation is shown in Figure S2.

Note that this metric does not necessarily reflect the “bottleneck” width for conduction, as the minimum-energy diffusion pathway for lithium need not follow a straight line. This feature may be interpreted, however, as a crude indicator for the energy change and/or scattering required for Li to conduct.

S1.18 Average straight-line path electronegativity, *SLPE* (dimensionless)

After finding the atom responsible for the minimum straight-line path width for each Li in the unit cell using the procedure described in S1.18, take the electronegativity of this atom. Repeat this process for all Li atoms in the cell and take the average.

S1.19 Ratio of average Li bond ionicity to average sublattice bond ionicity, RBI
(dimensionless)

Take the ratio of the average Li bond ionicity calculated in S1.4 and the average sublattice bond ionicity calculated in S1.7: $RBI = LBI/SBI$.

S1.20 Ratio of average Li neighbor count to average sublattice neighbor count, RNC
(dimensionless)

Take the ratio of the average Li neighbor count calculated in S1.5 and the average sublattice neighbor count calculated in S1.8: $RNC = LNC/SNC$.

S2. Training structures with fractional coordinates

Of the 40 training structures, 23 are given in the ICSD database with fractional atomic occupancies. In these cases, the exact equilibrium positions of some atoms in the unit cell have not been resolved due to high disorder or, in the case of highly mobile atoms, are ill-defined. Our feature extraction algorithm maps a structure to a vector based on the atomistic positions, and thus structures with fractional occupancies require some additional definition. In these cases, we employ a probabilistic sampling approach in which we evaluate the expectation value of the feature vector over the distribution of likely structures: $\langle X \rangle = \int XP(X)dX$. The probability function $P(X)$ is the probability of seeing a feature vector X . Since the feature extraction process f maps a structure x to a unique feature vector X , $f:x \rightarrow X$, we instead evaluate the expectation value over structures:

$$\langle X \rangle = \int f(x)P(x)dx \quad (\text{S18})$$

In this case, the function $P(x)$ is the probability of observing a particular structure with some sites occupied and others not. Provided in the structure files from the ICSD are the individual site occupation probabilities $P_{occ}(x_i)$ for all sites i ; we take the probability of seeing a particular occupation to be a Bernoulli random variable: $P(x_i) = P_{occ}(x_i)^{n_i} [1 - P_{occ}(x_i)]^{(1-n_i)}$, where $n_i=1$ if site i is occupied and $n_i=0$ otherwise. The probability of seeing a particular combination of occupied and unoccupied sites in a crystal is not generally equal to the product of the individual site occupation probabilities of the occupied sites, as these occupations are likely to

be correlated; i.e. the joint occupation probability $P(x) = P(\{x_i\}) \neq \prod_i P(x_i)$. To account for this correlation, we add a hard sphere term to ensure that sites separated by less than a cutoff radius r_c are not simultaneously occupied. Beyond this radius, we assume electronic screening negates correlation. We define the structural realization probability thus:

$$P(\{x_{ij}\}) = \prod_i \prod_{j \neq i} P_{occ}(x_i)^{n_i} [1 - P_{occ}(x_i)]^{(1-n_i)} 1_{\{r_{ij} \geq r_c\}} \quad (\text{S19})$$

This ensures that all mobile atoms are always separated by some minimum distance. To evaluate the integral S18, we use a Metropolis Monte Carlo-inspired scheme to evolve the structure and sample the features in accordance with the distribution $P(\{x_{ij}\})$. An initial structure is drawn randomly from the site occupation probabilities. If any two fractionally occupied atoms are separated by less than r_c , the structure is discarded. We initially set $r_c = 2 \text{ \AA}$ for all structures. However, some structures do not have any possible realizations in which all fractionally occupied atoms are separated by more than 2 \AA . In this case, if an acceptable structure is not drawn after 1,000 attempts, r_c is decremented by 0.1 \AA for the next 1,000 attempts, and so on, until a feasible structure is realized.

Once a starting structure is chosen, in each Monte Carlo step we generate trial structures by choosing two atoms and/or vacancies uniformly at random and swapping them. If this swap creates a permissible new structure, we accept the trial structure with probability $P_{\text{accept}} = P(\{x_{ij}\}_{\text{new}})/P(\{x_{ij}\}_{\text{old}})$. This is accomplished using a Metropolis type algorithm, i.e. accept with probability one if P_{accept} evaluates to greater than one. See Figure S3 for a flowchart detailing this structure evolution process. At each step, the features are extracted and added to the running average. This continues until convergence is reached. Convergence is reached when the running mean of all features change by less than 1% for 100 consecutive steps. All training set structures with fractional occupancies converge within 1,000 steps total. A plot of the convergence thresholds for each feature relative to the standard deviation of feature values across the training set is given in Fig. S4.

There are several classes of swaps that create invalid or redundant structures. For example, swapping two interchangeable atoms in an alloyed material, e.g. Ge and V in $\text{Li}_{3.5}\text{Ge}_{0.5}\text{V}_{0.5}\text{O}_4$, creates a valid new structure. Swapping the locations of Li and Ge, however, creates an invalid structure, as there are no sites that either Li or Ge can occupy. For this invalid structure $P(\{x\})=0$, so it makes no contribution to the integral S18 and we therefore do not bother to extract features.

Similarly, some swaps give rise to completely unchanged structures, e.g. swapping two vacancies, or two occupied Li sites in $\text{Li}_{3.5}\text{Ge}_{0.5}\text{V}_{0.5}\text{O}_4$. We take the

atoms and vacancies to be indistinguishable particles and thus reject trial structures that result in unchanged structures and do not extract their features.

S3. Motivation for logistic regression

The choice to classify based on ionic conductivity rather than predict a value of ionic conductivity is due to the relatively small size of the training set and as a precaution against overfitting to experimental noise, as ionic conductivity measurements on the same material can often vary by an order of magnitude or more due to sample and microstructure variability. Logistic regression is a robust low-variance (low flexibility) classification method, well suited to small training sets, which produces an easy-to-interpret classification function.

In our analysis, we studied the performance of models built with several other regression and classification methods, including ordinary least squares regression, robust regression (with several different cost functions), locally-weighted least squares, support vector machine classification (with several different kernels), multi-class (>2) classification, and neural networks. None of these models displayed strong predictive power as logistic regression in leave-one-out trials. Compared to logistic regression, these methods are relatively high variance, which can result in high cross-validation errors when the data set is small. Logistic regression performs well on this data, but its utility is limited by the simple good-or-bad-conductor nature of classification. As the data set grows, more powerful higher variance methods that can predict a numerical value of ionic conductivity (i.e. regression) may become appropriate.

Additionally, we avoid nonparametric methods like support vector machines because these do not provide an easily interpretable model. Parametric models, although generally less robust, can provide scientific intuition on which crystal features are most important to superionic conductor design. This is beneficial in any case where the underlying dynamics are not well understood. One particular nonparametric method that performs well on this dataset is locally-weighted regression. However, in addition to not providing an interpretable model, locally-weighted regression cannot extrapolate effectively beyond the region of training data, thus making it ineffective for a large proportion of the screened Materials Project data.

We seek to reduce the data dimensionality by throwing out individual features, as opposed to performing dimension reduction through regularization or a basis rotation algorithm, e.g. partial least squares³ because of the small dataset and the physical nature of the features. By throwing out individual features we can gain improved scientific understanding on which features are important and which are unimportant to improve ionic conductivity. Additionally, regularized models with many parameters generally have lower predictive power than unregularized models with a smaller number of features. The small dataset enables the usage of a relatively expensive combinatorial approach to model selection (all feature combinations considered) rather than regularization or forward/backward propagation. With a larger dataset, dimension reduction would likely need to be done with propagation or regularization.

S4. Classification threshold effects

The CVMR for both the true models and the randomized models (red curve in Fig. 5) approach a minimum at either end of the x -axis, when the number of training examples in one class goes to zero, while the difference between the X -randomized CVMR and the true CVMR peaks near the 50/50 class split. A more even split in classes in the training data is desirable, as this gives the model the optimum number of examples to learn what each class looks like. This is likely to give us the best performance on unseen data. Thus, in choosing our ionic conductivity threshold value, we must weigh three competing factors: (a) we want the lowest CVMR value, which tends to occur at the extremes in class division, (b) we want the best performance on unseen data, which occurs when the difference between the true CVMR and random CVMR is maximized (which occurs near the 50/50 split in Fig. 5), and (c) we need a physically relevant cutoff to make our model scientifically useful. We see in Fig. S5 that the 10^{-4} S/cm cutoff, originally chosen to satisfy criterion (c), performs well on criteria (a) and (b); the CVMR is low and the misclassification rate is approximately four times smaller than the random data case.

The percentage of the predicted superionics that are actually false positives mirrors the CVMR trend for the true models; this is represented by the blue curves in Fig. S5. Here the false positive rate (FPR) is defined as the probability that a material in the training set is non-superionic, given that it is predicted by the model as superionic:

$$FPR = P(\text{observed non-superionic} \mid \text{predicted superionic}) \quad (\text{S20})$$

Under randomized conditions, the probability of a material in the training set holding the non-superionic label and the probability of that same material being classified (uniformly at random) as superionic are independent; thus, FPR should simply equal the proportion of the training set that is non-superionic. The X -randomization process confirms this in Fig. S5 by tracing out the curve $FPR = 1 - (\% \text{ training set in positive class}) = (\% \text{ training set in negative class})$.

Predicting the exact misclassification rate of the screened materials is difficult, as it depends on the unknown underlying distribution of ionic conductivities across this set. Given the historical difficulty with which new

superionics are discovered, it is unlikely that 27.5% of the universe of Li-containing materials is superionic, as is the case in the training set. However if we assume this optimistic distribution does hold and the training data is a representative sample of the MP structures, and all MP structures are within the domain of applicability of the model, we would expect our model's superionic predictions would be correct 82% of the time. If the unknown true distribution of superionic materials is 1% for example, an investigator would need to test 100 materials before identifying a superionic, while our model may cut this down fourfold to $100/4 = 25$. In the absence of knowledge on the underlying ionic conductivity distribution, one cannot make this claim with certainty.

55. Computational investigation of promising electrolyte $\text{Li}_5\text{B}_7\text{S}_{13}$

We perform theoretical studies of the ionic conductivity of $\text{Li}_5\text{B}_7\text{S}_{13}$ ⁴, one of the promising new electrolyte candidates from Table 3. This provides additional validation of the model beyond the correct identification of superionic materials that are not included in the training set, discussed in Section 4.5. Our DFT studies on the structure $\text{Li}_5\text{B}_7\text{S}_{13}$ suggest that RT superionic lithium conduction is likely. We perform DFT molecular dynamics (MD) simulations at high temperature to speed up diffusion; MD at 900 K shows the lithium diffusion in $\text{Li}_5\text{B}_7\text{S}_{13}$ to be liquid-like while the sublattice remains stationary. At this temperature, we find ionic conductivity $\sigma_{900\text{ K}} = 3\text{ S/cm}$. Extrapolating down to room temperature, we estimate an ionic conductivity in the range of $\sigma_{\text{RT}} = 2.6 \times 10^{-2}$ to 1.6 S/cm depending on the diffusion mechanism. The details of this calculation are discussed below.

For the DFT calculations we use the Vienna Ab Initio Simulation Package⁵ (VASP) with the generalized gradient approximation (GGA) of Perdew-Burke-Ernzerhof⁶ (PBE) and the projector augmented wave⁷ (PAW) method. We use the following pseudopotentials: PAW_PBE Li_sv, PAW_PBE B, and PAW_PBE S. We employ a plane wave cutoff energy of 499 eV, a 2fs time step, and a gamma-point only k -mesh.

For the MD simulations, we begin with a single unit cell (100 atoms) and remove 1 of the 20 Li atoms for a Li vacancy concentration of 5%. We measure the diffusion coefficient from the mean-squared displacement of lithium to be $D = 1.31\text{ \AA}^2/\text{ps}$, while no other atom species are observed diffusing. We can calculate the energy barrier for lithium motion E_a using the following general diffusion formula:

$$E_a = -k_B T \ln\left(\frac{D}{f a^2 \nu_0 p_{\text{occ}} z^j}\right) \quad (\text{S21})$$

where k_B is Boltzmann's constant, T is absolute temperature, f is the diffusion correlation coefficient, $a = 3.673\text{ \AA}$ is the average hopping distance, $\nu_0 = 10\text{ ps}^{-1}$ is the Li attempt frequency, and $j = 1/6$ for 3-dimensional diffusion. We take $f = 1$. The product $p_{\text{occ}} z$ represents the fraction of Li atoms that are available to diffuse, and the

terms depend on the diffusion mechanism. If diffusion is vacancy driven, then p_{occ} is simply the vacancy concentration $x_v^0 = 0.05$, and $z = 2.7$ is the average number of Li atoms surrounding a Li vacancy (feature *LLB*); thus $p_{occ}^z = 0.135$. If the mechanism is interstitial, all atoms can diffuse and $p_{occ}^z = 1$. We calculate the attempt frequency ν_0 from the Li velocity autocorrelation function.

Given the speed of conduction, we cannot infer the mechanism from the MD results and we therefore calculate a range of possible diffusion barriers. For the purely vacancy driven case, we find $E_a = 0.065$ eV; for the interstitial case we find $E_a = 0.22$ eV. Utilizing the Einstein relationship, we relate the diffusion barrier to the Li conductivity as a function of temperature:

$$\sigma(T) = \frac{Dnq^2}{k_B T} = \frac{(fa^2\nu_0 p_{occ}^z j)nq^2}{k_B T} e^{-E_a/k_B T} \quad (S22)$$

where n is the lithium concentration and q is the fundamental charge. At 900 K, regardless of the mechanism, we calculate a liquid-like conductivity of $\sigma_{900 K} = 3$ S/cm. At RT, we estimate the ionic conductivity to fall between the range of $\sigma_{293 K} = 0.026$ S/cm for the interstitial mechanism case, and $\sigma_{293 K} = 1.6$ S/cm for the vacancy case. We believe that the interstitial mechanism is more likely, so we expect the RT ionic conductivity to be closer to the value of 0.026 S/cm. Regardless, we note that the RT ionic conductivity appears to be greater than the superionic threshold of 0.1 mS/cm; this result suggests that our model has correctly identified $\text{Li}_5\text{B}_7\text{S}_{13}$ as a promising new electrolyte material.

Table S1: References for atomic mass⁸, Pauling electronegativity^{1,2}, covalent radius⁹, ionic radius of most common oxidation state¹⁰, materials cost per kg¹¹, and ranking of abundance in Earth's crust¹². For some elements, this data is unavailable. The sublattice bond ionicity (SBI) feature in the superionic predictor uses the ionicity difference between atoms, so structures including elements without electronegativity values (He, Ne, Ar, Pm, Eu, Tb, Yb, Rn) are eliminated from the screening. Costs are not calculated for materials including elements without reliable cost values (Ar, Kr, Tc, Pm, Po, At, Rn, Fr, Ra, Ac, Th) or earth abundance indices (He, Ne, Ar, Kr, Tc, Xe, Pm, Po, At, Rn, Fr, Ra, Ac, Pa), but these materials are not eliminated from the screening.

Symbol	Z	Atomic mass	Pauling electronegativity	Covalent radius	Ionic radius	Cost per kg	Earth abundance rank
H	1	1.008	2.20	0.31	0.31	3	11
He	2	4.003	–	0.28	0.28	52	–
Li	3	6.94	0.98	1.28	0.9	270	34
Be	4	9.012	1.57	0.96	0.59	7480	50
B	5	10.81	2.04	0.84	0.41	11140	38
C	6	12.011	2.55	0.76	0.3	24	10
N	7	14.007	3.04	0.71	1.32	4	32
O	8	15.999	3.44	0.66	1.26	3	1
F	9	18.998	3.98	0.57	1.19	1900	14
Ne	10	20.180	–	0.58	0.58	330	–
Na	11	22.990	0.93	1.66	1.16	250	7
Mg	12	24.305	1.31	1.41	0.86	5.8	6
Al	13	26.982	1.61	1.21	0.68	1.31	3
Si	14	28.085	1.90	1.11	0.54	2.37	2
P	15	30.974	2.19	1.07	0.52	300	13
S	16	32.059	2.58	1.05	1.7	500	15
Cl	17	35.45	3.16	1.02	1.67	1.5	19
Ar	18	39.948	–	1.06	1.06	–	–
K	19	39.098	0.82	2.03	1.52	1000	8
Ca	20	40.078	1.00	1.76	1.14	200	5
Sc	21	44.956	1.36	1.70	0.89	14000	31
Ti	22	47.867	1.54	1.60	0.85	66.22	9
V	23	50.942	1.63	1.53	0.78	14.33	18
Cr	24	51.996	1.66	1.39	0.74	9.59	20
Mn	25	54.938	1.55	1.39	0.75	1.5	12
Fe	26	55.845	1.83	1.32	0.77	72	4
Co	27	58.933	1.88	1.26	0.76	46	29
Ni	28	58.693	1.91	1.24	0.72	40	22
Cu	29	63.546	1.90	1.32	0.82	7	24

Zn	30	65.38	1.65	1.22	0.88	3.5	23
Ga	31	69.723	1.81	1.22	0.76	525	33
Ge	32	72.63	2.01	1.20	0.77	940	54
As	33	74.9216	2.18	1.19	0.66	1.433	48
Se	34	78.96	2.55	1.20	1.84	85	67
Br	35	79.904	2.96	1.20	1.82	50	45
Kr	36	83.798	3.00	1.16	1.16	–	–
Rb	37	85.468	0.82	2.20	1.66	12000	25
Sr	38	87.62	0.95	1.95	1.32	1000	16
Y	39	88.906	1.22	1.90	1.04	45	30
Zr	40	91.224	1.33	1.75	0.86	1570	21
Nb	41	92.906	1.60	1.64	0.82	180	35
Mo	42	95.96	2.16	1.54	0.78	30	57
Tc	43	98	1.90	1.47	0.74	–	–
Ru	44	101.07	2.20	1.46	0.76	14000	75
Rh	45	102.906	2.28	1.42	0.75	80000	76
Pd	46	106.42	2.20	1.39	0.88	58330	69
Ag	47	107.868	1.93	1.45	1.09	1022	65
Cd	48	112.411	1.69	1.44	1.09	4.19	64
In	49	114.818	1.78	1.42	0.94	590	63
Sn	50	118.71	1.96	1.39	0.83	12	47
Sb	51	121.76	2.05	1.39	0.82	6	62
Te	52	127.6	2.10	1.38	2.07	240	74
I	53	126.904	2.66	1.39	2.06	83	60
Xe	54	131.293	2.60	1.40	1.40	1200	–
Cs	55	132.905	0.79	2.44	1.81	11000	49
Ba	56	137.327	0.89	2.15	1.49	100	17
La	57	138.905	1.10	2.07	1.17	2	27
Ce	58	140.116	1.12	2.04	1.08	2	26
Pr	59	140.908	1.13	2.03	1.06	25	37
Nd	60	144.242	1.14	2.01	1.12	25	28
Pm	61	145	–	1.99	1.11	–	–
Sm	62	150.36	1.17	1.98	1.10	3	40
Eu	63	151.964	–	1.98	1.20	300	51
Gd	64	157.25	1.20	1.96	1.08	19.5	42
Tb	65	158.925	–	1.94	0.98	600	58
Dy	66	162.5	1.22	1.92	1.13	70	39
Ho	67	164.930	1.23	1.92	1.04	8600	55
Er	68	167.259	1.24	1.89	1.03	5400	44
Tm	69	168.934	1.25	1.90	1.10	70000	61
Yb	70	173.054	–	1.87	1.08	14000	46
Lu	71	174.967	1.27	1.87	1	2400	78

Hf	72	178.49	1.30	1.75	0.85	1200	43
Ta	73	180.948	1.50	1.70	0.82	150	53
W	74	183.84	2.36	1.62	0.77	24	56
Re	75	186.207	1.90	1.51	0.71	5000	72
Os	76	190.23	2.20	1.44	0.71	77000	73
Ir	77	192.217	2.20	1.41	0.77	42000	77
Pt	78	195.084	2.28	1.36	0.81	54000	70
Au	79	196.967	2.54	1.36	1.07	53000	71
Hg	80	200.592	2.00	1.32	1.24	17.4	66
Tl	81	204.38	1.62	1.45	1.33	480	59
Pb	82	207.2	2.33	1.46	1.12	2.75	36
Bi	83	208.980	2.02	1.48	1.04	18.7	68
Po	84	209	2.00	1.40	0.94	-	-
At	85	210	2.20	1.50	0.76	-	-
Rn	86	222	-	1.50	1.5	-	-
Fr	87	223	0.70	2.60	1.94	-	-
Ra	88	226	0.90	2.21	1.62	-	-
Ac	89	227	1.10	2.15	1.26	-	-
Th	90	232.038	1.30	2.06	1.08	-	41
Pa	91	231.036	1.50	2.00	1.04	280000	-
U	92	238.029	1.38	1.96	0.99	165	52

Table S2: *Feature values for select training structures.*

Feature index	Feature name	LiI	Li ₃ P
1	AAV	27.35 Å ³	14.63 Å ³
2	SDLC	0.0	1.55
3	SDLI	0.0	0.54
4	LBI	1.68	0.48
5	LNC	6.0	15
6	LLB	0.0	11.33
7	SBI	1.68	1.21
8	SNC	6.0	11
9	AFC	12.0	12
10	AASD	4.26 Å	4.36 Å
11	VPA	54.70 Å ³	58.51 Å ³
12	LASD	3.02 Å	2.50 Å
13	LLSD	4.26 Å	2.59 Å
14	ENS	2.66	2.19
15	PF	0.37	0.54
16	SPF	0.21	0.09
17	SLPW	0.74 Å	1.09 Å
18	SLPE	2.66	0.98
19	RBI	1.0	0.39
20	RNC	1.0	1.36

Table S3: Screening results with looser constraints. A list of the results of the Materials Project screening with looser screening constraints is given with all material specifications, ranked in the table by overall performance across all specifications. P_{LR} is the superionic probability; d , ϵ , and A are confidence metrics for this prediction; E_{gap} is the DFT computed band gap; V_{ox} is the upper bound oxidation potential; the “cost per area” column gives approximate materials cost per m^2 per $10 \mu m$ thickness; E_{hull} is the energy above the convex hull; I_A is the earth abundance index; to understand model performance, the final column provides references to studies on the ionic conductivity of these or similar materials. The screening criteria employed to generate this list are: $P_{LR} \geq 0.5$, $E_{gap} \geq 0.5$ eV; $V_{ox} \geq 3$ V; $E_{hull} \leq 0.1$ eV/atom; $T = 0$ (no transition metals present). Materials already listed in Table 3 have been removed. *The training set includes these materials or structurally similar derivatives as superionic examples and thus these materials should not be considered successful model generalizations (see Table 1 in main text for data references). †These materials are reported poor conductors, i.e. they are confirmed false positive predictions. ‡These materials are reported fast conductors, i.e. they are correct model predictions. #These materials are low confidence predictions given $d > 2$. Note that several of these materials have more than one anion type, making the model prediction susceptible to the mixed anion effect; see Section 4.3 in main text.

MPID	Chemical formula	P_{LR}	d	ϵ	A	E_{gap} (eV)	V_{ox} (V)	Cost per area, 10 μm thick (USD/ m^2)	E_{hull} (eV/atom)	I_A	Related study
mp-768229	Li_2BiS_3	0.869	0.986	0.031	1	0.888	3.195	7.30	0.068	68	–
mp-753429	$Li_4Bi_2S_7$	0.978	1.579	0.009	1	1.093	3.312	7.40	0.058	68	–
mp-696123	$Li_{10}Sn(PS_6)_2$	0.793	0.79	0.037	1	2.164	3.151	7.60	0.023	47	*
mp-4495	$KLiTe$	0.982	1.742	0.008	1	2.413	3.24	13	0	74	–
mp-985583	Li_3PS_4	0.622	0.775	0.033	1	2.951	3.376	7.90	0	34	*
mp-9703	$K_2LiGaAs_2$	0.97	1.545	0.013	1	0.755	3.52	13	0	48	–
mp-505431	$K_2LiInAs_2$	0.978	1.725	0.010	1	0.619	3.428	15	0	63	–
mp-695365 [#]	$LiAl_3Si_9(N_7O)_2$	0.84	3.157	0.136	0	3.874	44.999	0.19	0.046	34	–
mp-35205	$LiErSe_2$	0.645	1.46	0.051	1	1.605	6.71	170	0.017	67	–
mp-30249 [†]	Li_4GeS_4	0.781	0.802	0.040	1	2.483	3.058	14	0	54	[¹³]
mp-30959	$LiHS$	0.813	0.898	0.035	1	3.547	3.125	6.20	0	34	–
mp-37747	$LiHS$	0.801	0.89	0.036	1	3.496	3.122	6.20	0.001	34	–
mp-867656 [#]	LiH_2N_3O	1.000	3.603	0.000	0	3.651	5.8	0.49	0.062	34	–
mp-8756	$KLiSe$	0.886	1.44	0.034	1	2.525	3.658	11	0	67	–
mp-864755 [#]	$LiAcTe_2$	0.997	3.014	0.002	0	0.557	4.855	–	0	–	–

mp-29008‡	Li ₆ MgBr ₈	0.88	1.35	0.027	1	4.764	3.856	2.10	0.02	45	[14]
mp-14591	LiSb ₂	0.759	1.15	0.041	1	1.184	3.972	6.40	0.004	62	[15]
mp-721236	Li ₁₀ Sn(PS ₆) ₂	0.678	0.798	0.039	1	2.443	3.181	7.50	0.011	47	*
mp-690708 [#]	KLiPH ₂ O ₄ F	1.000	4.521	0.000	0	4.904	23.93	12	0.039	34	-
mp-555874	LiAs ₂	0.843	1.216	0.033	1	1.066	3.763	7	0	48	[16]
mp-769129 [#]	Li ₃ SbN ₃ (O ₃ F) ₃	0.966	2.258	0.015	0	3.338	10.583	7.40	0.096	62	-
mp-841 [†]	Li ₂ O ₂	0.948	2.918	0.029	0	1.974	3.3	1.90	0	34	[17]
mp-28592	Li ₇ Br ₃ O ₂	0.572	1.719	0.077	1	4.448	3.076	2.40	0.028	45	-
mp-9250	RbLiSe	0.77	1.625	0.059	1	2.309	3.562	210	0	67	-
mp-33526	LiBiS ₂	0.929	1.398	0.021	1	0.811	4.231	7.20	0	68	-
mp-720554 [#]	LiPH ₄ NO ₃ F	1.000	3.971	0.000	0	5.104	20.554	7.80	0.014	34	-
mp-36791	LiNdS ₂	0.887	1.38	0.030	1	1.857	8.422	6.90	0.019	34	-
mp-1002069 [#]	NaLiCl	0.85	3.392	0.073	0	4.187	6.435	2.90	0.083	60	-
mp-995393 [#]	LiS ₄	0.997	2.504	0.002	0	2.159	3.741	1.00	0	34	-
mp-28603	NaLiSe	0.505	1.154	0.057	1	2.438	3.517	4.00	0	67	-
mp-1001069	Li ₄₈ P ₁₆ S ₆₁	0.705	0.765	0.035	1	2.297	3.254	7.60	0.017	34	-
mp-6450	K ₂ LiAlP ₂	0.962	1.42	0.015	1	1.885	3.882	12	0	34	-
mp-696138	Li ₁₀ Ge(PS ₆) ₂	0.68	0.776	0.039	1	2.476	3.201	10.00	0.019	54	*
mp-33667	LiDyS ₂	0.871	1.269	0.030	1	1.974	8.713	9.10	0.006	39	-
mp-29463	LiBeN	0.598	0.908	0.054	1	2.686	3.014	45	0	50	-
mp-985585‡	Li ₃ ClO	0.565	1.777	0.085	1	4.677	3.439	1.60	0.022	34	[18]
mp-644419	LiHS	0.804	0.893	0.036	1	3.525	3.123	6.20	0.001	34	-
mp-696129	Li ₁₀ Si(PS ₆) ₂	0.788	0.749	0.036	1	2.383	3.283	7.80	0.011	34	*
mp-29009 [†]	Li ₂ MgBr ₄	0.786	1.258	0.038	1	4.527	5.37	1.90	0.015	45	[19]
mp-644271	LiHS	0.823	0.917	0.034	1	3.498	3.124	6.20	0	34	-
mp-558731	Li ₂ TeS ₃	0.77	1.138	0.038	1	1.457	3.037	11	0	74	[16]
mp-559971 [#]	LiS ₂ N(O ₂ F) ₂	1.000	4.113	0.000	0	5.9	17.609	12	0.093	34	-
mp-556562 [#]	LiAsH ₆ (OF ₂) ₃	1.000	3.594	0.000	0	3.784	33.978	21	0.015	48	-
mp-641703	Li ₇ P ₃ S ₁₁	0.779	0.886	0.033	1	2.487	3.529	8.20	0.02	34	*
mp-9722	Na ₂ LiGaAs ₂	0.937	1.361	0.022	1	0.684	3.37	6.50	0	48	-
mp-777963	Li ₃ SbH ₁₈ S ₄ O ₉	0.661	1.137	0.047	1	2.454	14.658	2.70	0.087	62	-

mp-10618	LiInSe ₂	0.952	1.464	0.016	1	0.838	3.199	16	0.082	67	-
mp-35591	LiErS ₂	0.827	1.238	0.035	1	1.989	8.73	190	0.008	44	-
mp-23259 [†]	LiBr	0.949	1.651	0.015	1	5.091	3.096	2.30	0.025	45	[²⁰]
mp-684082	LiSb(PO ₃) ₄	0.507	1.103	0.065	1	4.004	44.831	2.30	0.047	62	-
mp-695635 [#]	Li ₃ Al ₃ P ₃ H ₂ O ₁₄ F	1.000	3.353	0.000	0	5.771	24.935	4.70	0.01	34	-
mp-767137	Li ₃ Sb ₁₇ S ₂₇	0.601	1.272	0.093	1	0.508	11.426	6.50	0.011	62	-
mp-5614	Li ₃ BS ₃	0.901	1.206	0.026	1	3.050	3.092	22	0	38	[²¹]
mp-9719	Na ₂ LiAlP ₂	0.956	1.382	0.017	1	1.885	3.852	4.60	0	34	-
mp-720509	Li ₁₀ Si(PS ₆) ₂	0.687	0.762	0.039	1	2.564	3.311	7.80	0	34	*
mp-696128	Li ₁₀ Ge(PS ₆) ₂	0.79	0.753	0.036	1	2.214	3.173	10	0.03	54	*

Table S4: Screening results for stable superionics. A list of the results of the Materials Project screening based on superionic probability and stability only is given with all material specifications. All metrics are the same as in Table S3 with the addition of T , a Boolean that indicates if a transition metal is present in the structure ($T = 1$) or not ($T = 0$). The screening criteria employed to generate this list are: $P_{LR} \geq 0.95$, $d \leq 1.5$; energy above hull = 0 eV/atom. Note most materials have small or zero band gaps, low oxidative decomposition voltages, and transition metals present, rendering them poorly suited for electrolyte applications. None of these materials are included in our training set, nor do any have reported ionic conductivity values to our knowledge.

MPID	Chemical formula	P_{LR}	d	ϵ	A	E_{gap} (eV)	V_{ox} (V)	Cost per area, 10 μm thick (USD/m ²)	T	I_A
mp-29720	Li ₂₁ Si ₅	0.984	1.283	0.006	1	0	0.275	1.70	0	34
mp-569115	Li ₂ ZnGe	0.978	1.426	0.009	1	0	0.667	18	1	54
mp-21471	LiFeAs	0.981	1.367	0.007	1	0	1.329	2.3	1	48
mp-14209	LiYGe	0.961	1.47	0.016	1	0	2.2	20	1	54
mp-8132	Li(TiSe ₂) ₃	0.979	1.357	0.007	1	0	11.606	3.40	1	67
mp-6450	K ₂ LiAlP ₂	0.962	1.42	0.015	1	1.885	3.882	12	0	34
mp-8864	Li(Co ₃ P ₂) ₂	0.975	1.499	0.01	0	0	7.111	7.40	1	34
mp-1777	Li ₁₅ Ge ₄	0.978	1.299	0.008	1	0	0.413	16	0	54
mp-8331	Li ₁₂ Mg ₃ Si ₄	0.994	1.322	0.003	1	0	0.365	1.30	0	34
mp-27454	BaLi ₂ Si	0.989	1.406	0.005	1	0	0.545	2.90	0	34
mp-8406	Li ₃ NdAs ₂	0.961	1.253	0.014	1	0.036	1.99	1.40	0	48
mp-530480	Li ₁₉ (FeAs) ₂₀	0.978	1.346	0.008	1	0.001	1.351	2.20	1	48
mp-11489	Li ₃ Pd	0.966	1.446	0.014	1	0	0.415	1700	1	69
mp-9562	LiBeAs	0.977	1.307	0.008	1	1.008	1.461	26.00	0	50
mp-733455	Li ₃ H ₆ Rh	0.995	1.468	0.002	1	2.425	1.647	1600	1	76
mp-865874	Li ₂ AgHg	0.982	1.364	0.007	1	0	0.595	28	1	66
mp-20698	LiCoAs	0.98	1.376	0.007	1	0	1.491	1.80	1	48
mp-20049	LiFeP	0.973	1.269	0.009	1	0	1.973	6.30	1	34
mp-19755	Li(TiS ₂) ₃	0.96	1.201	0.012	1	0	17.422	8.20	1	34
mp-15820	LiHoGe	0.958	1.461	0.017	1	0	2.143	400	0	55
mp-865965	Li ₂ CaSi	0.981	1.376	0.008	1	0.425	0.643	2.80	0	34
mp-865666	Li ₂ TlAg	0.987	1.447	0.005	1	0	0.453	48	1	65

mp-14208	LiYSi	0.965	1.49	0.014	1	0	1.839	1.60	1	34
mp-30100	Li ₂ B ₂ Se ₅	0.952	1.289	0.015	1	1.776	2.308	27	0	67
mp-862658	LiCu ₃	0.976	1.455	0.008	0	0.117	0.123	1.10	1	34
mp-13474	LiErGe	0.956	1.454	0.017	1	0	2.132	270	0	54
mp-861651	Li ₂ PdPb	0.955	1.246	0.015	1	0	0.767	1500	1	69
mp-571566	Li ₂ ZnSi	0.981	1.43	0.008	1	0	0.56	1.00	1	34
mp-569373	Li ₈ Zn ₂ Ge ₃	0.988	1.314	0.005	1	0	0.567	18	1	54
mp-976114	Li ₂ CdHg	0.987	1.453	0.006	1	0	0.632	1.70	1	66
mp-865935	Li ₂ MgSn	0.985	1.444	0.006	1	0	0.663	1.20	0	47
mp-14210	LiGdGe	0.966	1.497	0.014	1	0	2.105	19	0	54
mp-9915	LiBeP	0.965	1.206	0.011	1	1.159	1.763	33	0	50
mp-9719	Na ₂ LiAlP ₂	0.956	1.382	0.017	1	1.885	3.852	4.60	0	34

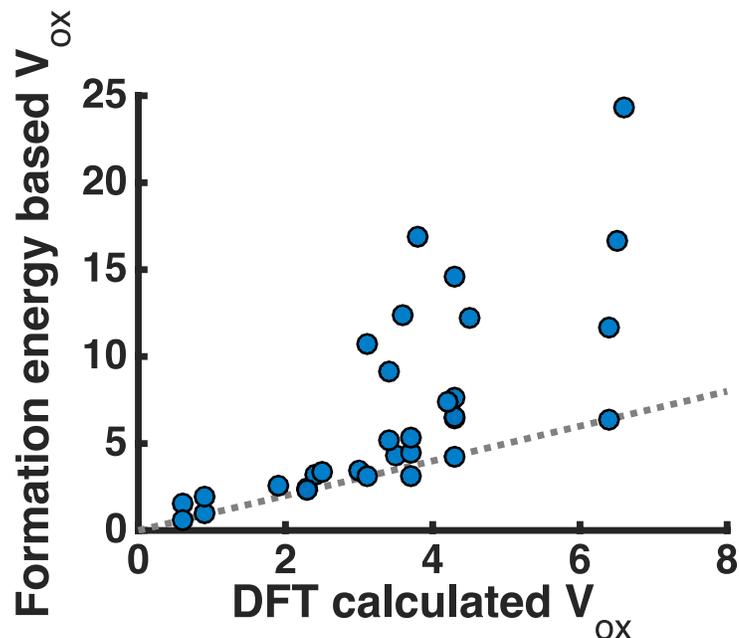


Figure S1: *Confirmation of upper bound on oxidation potential.* Above we see the formation energy-based oxidative decomposition potential estimates we employ in this work versus rigorously calculated oxidative decomposition potentials reported in Ref. [22]. With two slight exceptions, the formation energy based potentials are always greater than the DFT-calculated potentials (the potentials are equal on the dotted $y=x$ reference line). This suggests that the simple formation energy based estimates are generally upper bounds on the true decomposition potentials, as we expect from our analysis in Section 2.3. The estimates are especially accurate in the interval $[0, 3]$, the range most relevant to LIB operation.

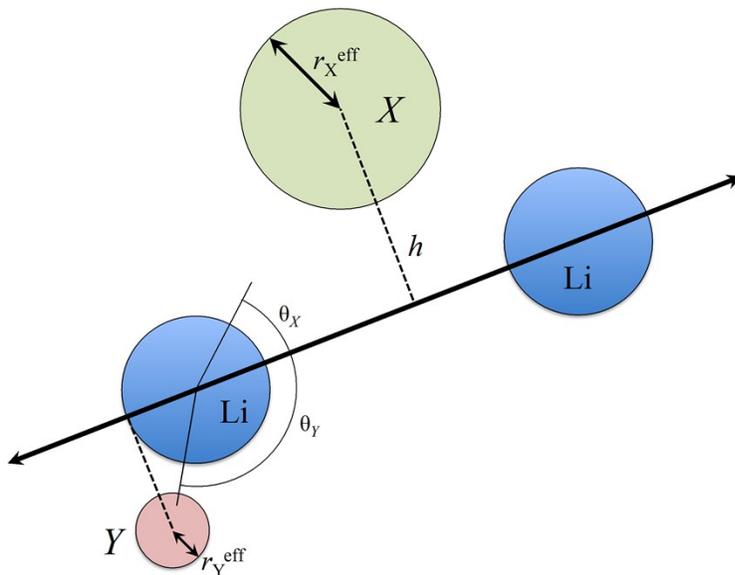


Figure S2: *Straight-line path width and electronegativity calculation.* To calculate features *SLPW* and *SLPE* (1.17 and 1.18), we draw a straight line connecting each lithium atom i with its nearest lithium atom j in the supercell. We project the coordinates of all other atoms in the lattice onto this line and find the atom that gives the minimum distance connecting the straight line to the edge of the hard sphere defined by the atom. Atoms that do not project down onto a point between Li atoms i and j are ignored, e.g. atom Y is ignored (since $\theta_Y > 90^\circ$) but atom X is not ($\theta_X < 90^\circ$). The minimum distance h is considered the straight-line path width for Li atom i . To calculate feature *SLPW*, we perform this width calculation for all Li atoms and take the average value. To calculate feature *SLPE*, we evaluate the Pauling electronegativity of the atom responsible for the minimum path width (atom X).

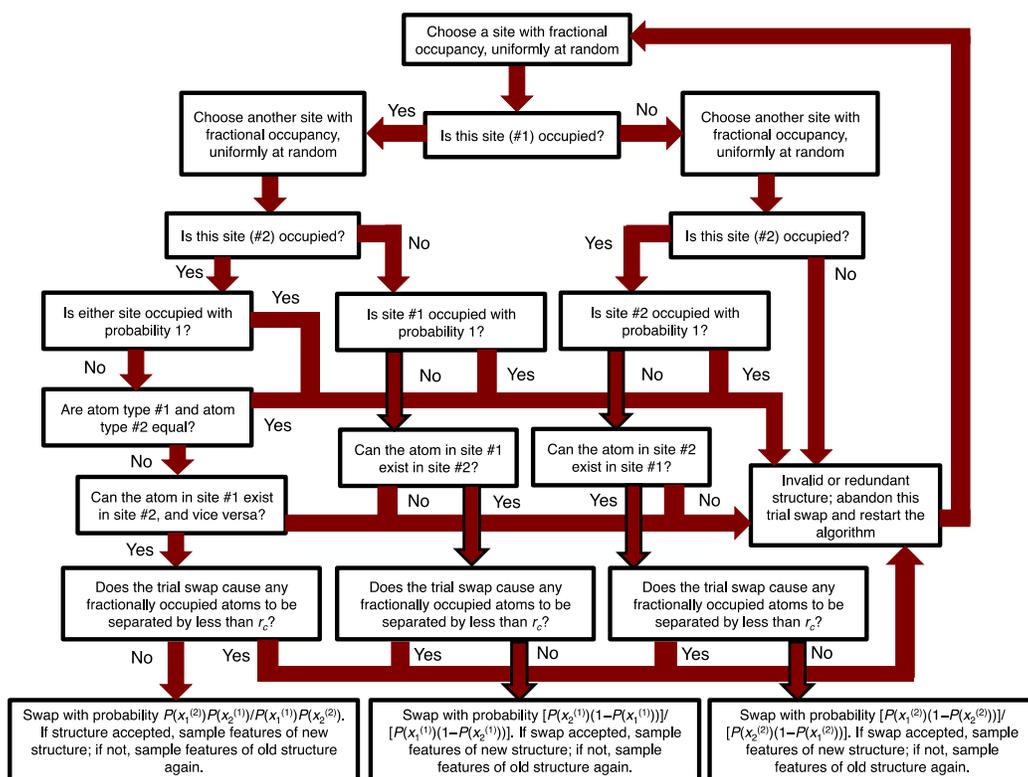


Figure S3: Flowchart for Metropolis Monte Carlo-style structural evolution. The above flowchart describes one step in the Monte Carlo scheme for swapping two atoms and/or vacancies in a structure with fractionally occupied sites. After each step, whether the trial structure is accepted or not, the feature vector of the resulting structure is calculated and added to the running average until convergence is reached. $P(x_i^{(j)})$ represents the occupation probability of atom type j in site i . Note that choosing the two sites 1 and 2 uniformly at random preserves detailed balance. Trial structures that are identical to the previous structure (i.e. two identical atoms or two vacancies are swapped) or are otherwise unfeasible structures (i.e. two non-interchangeable atoms are swapped), are ignored and their features are never sampled.

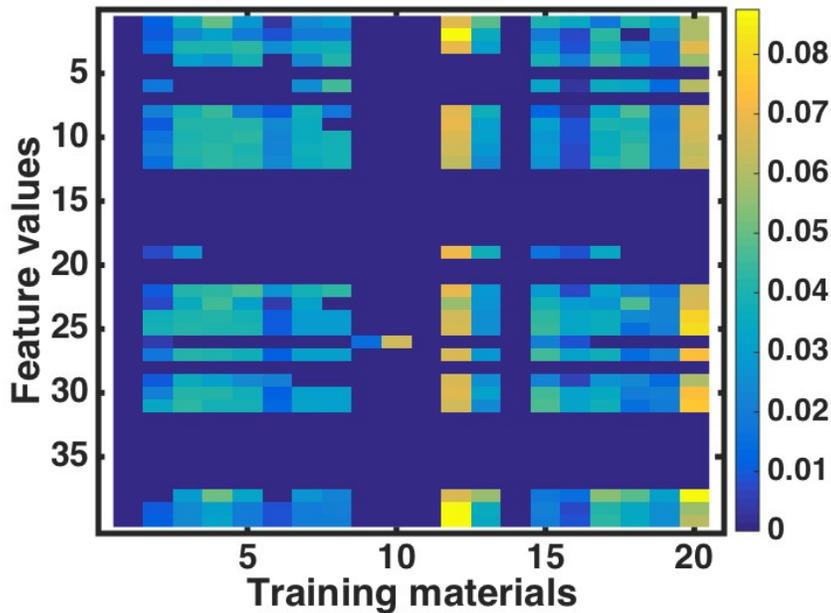


Figure S4: *Metropolis convergence thresholds.* The Metropolis structure evolution algorithm ends when convergence is reached. Convergence is defined as the point in which the running feature means of all features deviate by less than 1% for 100 consecutive steps. To visualize this degree of convergence relative to the spread of feature values across the training set, the 1%-of-mean values for all converged features for all training materials with fractional occupancies are normalized by the standard deviation of that feature value across the entire training set. Note that all features are converged within 10% of the training set feature spread.

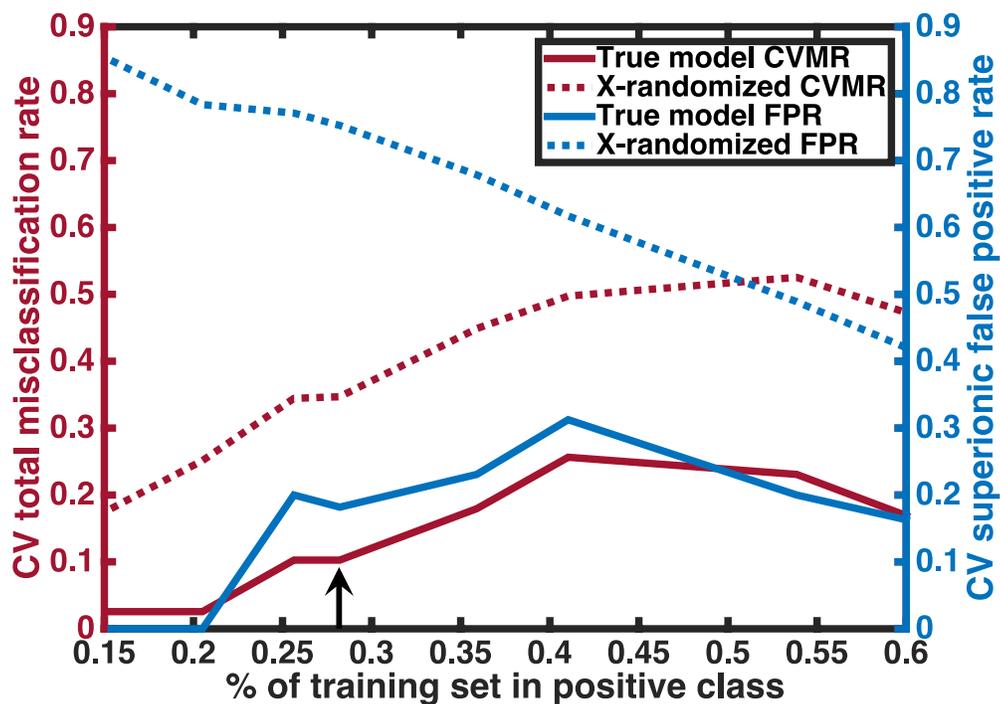


Figure S5: Conductivity threshold effects. The conductivity threshold for positive and negative classification affects the CVMR rate of the optimal model. The CVMR peaks when approximately 40% of the training set examples belong to the positive class, corresponding to a threshold of 10^{-5} S/cm. The threshold value of 10^{-4} S/cm we use in our screening is chosen due to technological requirements. Note the CVMR for the X-randomized case is always significantly higher than the CVMR of the true models. The false positive rate for the X-randomized case follows the trend of $y=1-x$, which we would expect from random drawing of materials. The false positive rate for the 10^{-4} S/cm cutoff model (indicated on the graph with the black arrow) is four times lower than that of the X-randomized models, suggesting a fourfold improvement over trial-and-error materials testing.

References

- 1 A. L. Allred, *J. Inorg. Nucl. Chem.*, 1961, **17**, 215–221.
- 2 *Wikipedia Free Encycl.*, 2016.
- 3 H. Abdi, *Wiley Interdiscip. Rev. Comput. Stat.*, 2010, **2**, 97–106.
- 4 F. Hiltmann, P. Zum Hebel, A. Hammerschmidt and B. Krebs, *Z. Für Anorg. Allg. Chem.*, 1993, **619**, 293–302.
- 5 G. Kresse and J. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169–11186.
- 6 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
- 7 P. E. Blöchl, *Phys. Rev. B*, 1994, **50**, 17953–17979.
- 8 J. Meija, T. B. Coplen, M. Berglund, W. A. Brand, B. P. De, M. Gröning, N. E. Holden, J. Irrgeher, R. D. Loss, T. Walczyk and T. Prohaska, *Pure Appl. Chem.*, 2016, **88**, 265–291.
- 9 B. Cordero, V. Gómez, A. E. Platero-Prats, M. Revés, J. Echeverría, E. Cremades, F. Barragán and S. Alvarez, *Dalton Trans.*, 2008, 2832–2838.
- 10 R. D. Shannon, *Acta Crystallogr. Sect. A*, 1976, **32**, 751–767.
- 11 *Wikipedia Free Encycl.*, 2016.
- 12 *Wikipedia Free Encycl.*, 2016.
- 13 R. Kanno, T. Hata, Y. Kawamoto and M. Irie, *Solid State Ion.*, 2000, **130**, 97–104.
- 14 M. Schneider, P. Kuske and H. D. Lutz, *Z. Für Naturforschung B*, 1993, **48**, 1–6.
- 15 US20110274984 A1, 2011.
- 16 Y. Wang, W. D. Richards, S. P. Ong, L. J. Miara, J. C. Kim, Y. Mo and G. Ceder, *Nat. Mater.*, 2015, **advance online publication**.
- 17 O. Gerbig, R. Merkle and J. Maier, *Adv. Mater.*, 2013, **25**, 3129–3133.
- 18 Y. Zhao and L. L. Daemen, *J. Am. Chem. Soc.*, 2012, **134**, 15042–15047.
- 19 R. Kanno, Y. Takeda, O. Yamamoto, C. Cros, W. Gang and P. Hagemuller, *J. Electrochem. Soc.*, 1986, **133**, 1052–1056.
- 20 Y. Haven, *Recl. Trav. Chim. Pays-Bas*, 1950, **69**, 1471–1489.
- 21 H. Wada, M. Menetrier, A. Lévassieur and P. Hagemuller, *Mater. Res. Bull.*, 1983, **18**, 189–193.
- 22 W. D. Richards, L. J. Miara, Y. Wang, J. C. Kim and G. Ceder, *Chem. Mater.*, 2016, **28**, 266–273.