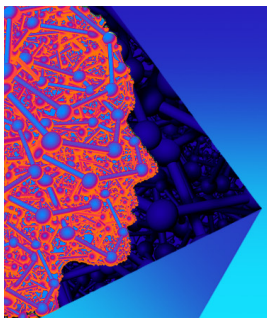
RESEARCH ARTICLE | OCTOBER 06 2023

Testing the limits of the global instability index

Special Collection: Challenges and Perspectives in Materials Chemistry—A Celebration of Prof. Sir Anthony K. Cheetham's 75th Birthday

Kyle D. Miller ; James M. Rondinelli

APL Mater. 11, 101108 (2023)
https://doi.org/10.1063/5.0140480



Testing the limits of the global instability index

Cite as: APL Mater. 11, 101108 (2023); doi: 10.1063/5.0140480
Submitted: 28 December 2022 • Accepted: 20 September 2023 •
Published Online: 6 October 2023

Kyle D. Miller[a] and James M. Rondinelli[b]

AFFILIATIONS

Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208, USA

Note: This paper is part of the Special Topic on Challenges and Perspectives in Materials Chemistry—A Celebration of Prof. Sir Anthony K. Cheetham's 75th Birthday.
[a]**Electronic mail:** kmiller@u.northwestern.edu
[b]**Author to whom correspondence should be addressed:** jrondinelli@northwestern.edu

ABSTRACT

The global instability index (GII) is a computationally inexpensive bond valence-based metric originally designed to evaluate the total bond strain in a crystal. Recently, the GII has gained popularity as a feature of data-driven models in materials research. Although prior studies have proven that GII is an effective predictor of structural distortions and decomposition energy when applied to small datasets, the wider use of GII as a global indicator of structural stability has yet to be evaluated. To that end, we compute GII for thousands of compounds in inorganic structure databases and partition compounds by chemical interactions underlying their stability to understand the GII values and their variations. Our results show that the GII captures *relative* chemical trends, such as electronegativity, even beyond the intended domain of strongly ionic compounds. However, we also find that GII magnitudes vary significantly with factors such as chemistry (cation–anion identities and bond character), geometry (connectivity), data source, and model bias, making GII suitable for comparisons within controlled datasets but unsuitable as an *absolute*, *universal* metric for structural feasibility.

I. INTRODUCTION

The global instability index (GII or *R*1) measures the aggregated deviation of a structure's bond lengths from their respective "ideal" lengths via comparison of bond valence sums with the formal valence,[1,2]

$$\mathrm{GII} = \sqrt{\frac{\sum_{i=1}^{N} \left(V_{\mathrm{VSR},i} - V_i\right)^2}{N}}, \tag{1}$$

where $V_{\mathrm{VSR},i}$ is the bond valence sum of atom i according to the electrostatic valence sum rule (VSR), V_i is the formal valence state of atom i, and N is the number of cations in a compound's unit cell. The GII metric was developed as a measure of bond strain in a structure.[3] Brown showed how this metric could be used to predict structural distortions in perovskite oxides,[1] and Salinas–Sanchez used the metric to rationalize formability in a superconducting cuprate family.[2] Since then, GII has gained popularity as a tool for understanding the distortion and formation of crystals, for example, as a feasibility check during structural refinement,[4,5] and as part of a perovskite structure prediction software.[6]

GII has also found use as a predictor of properties beyond steric effects. Rao *et al.* used GII to create electronic and magnetic phase diagrams for perovskite manganites,[7] Feng *et al.* used GII as a predictor of thermal stability of perovskite oxides,[8] and Georgescu *et al.* found that GII is an important feature in predicting a compound's metallicity.[9] Kapera and Koleżyński used GII to inform structural stability across doping in a thermoelectric materials candidate.[10]

Clearly, GII can be a cost-effective, powerful tool to help predict bonding-related properties of materials. However, most of these studies focus on a limited family of compounds with relatively well-known electronic characteristics, as is consistent with Brown's original work. Brown excluded structures with electronic distortions (compounds containing certain transition metals and Pearson's soft acids and bases[11]), noting that electronic effects often cooperate with steric effects to stabilize such distortions. The presence of electronic effects beyond simple local charge transfer yields a more complicated physical picture less easily captured by the bond valence model.

As the popularity and size of open access materials databases increase and large-scale materials studies become more common, GII becomes even more attractive to use because of its relatively low computational cost, interpretability, and simplicity as a single scalar value. Many papers have touted Brown's observed maximum GII of 0.2 valence units (v.u.), as a universal threshold, above which the

feasibility of a structure should be doubted.[6,7,12–14] Several papers, however, have noted that while GII is correlated with lattice energy, the 0.2 v.u. threshold is far from comprehensive.[8,15] In fact, Brown specifically cautions readers against assuming the 0.2 v.u. threshold for arbitrary structures.[1] Given the diversity of factors affecting the bonding behavior, such as covalency, metal–metal bonding, electron filling, coordination, and polyhedral connectivity, it follows that one simple metric is unlikely to capture a property as complex as structural stability for all inorganic crystals. Thus, we seek to explore the limits of the global instability index. We address this goal by focusing on four aspects of a compound: chemistry (cation–anion identities and bond character), geometry (connectivity), data source, and model bias. We find that GII depends on all of these aspects to varying degrees in different compounds, making it less useful for global predictions across diverse compositions and structure types and better suited for materials-chemistry-informed analyses within controlled datasets.

II. METHODS

All structures are stored as Crystallographic Information Files (CIFs)[16] and manipulated using Pymatgen[17] and visualized using VESTA.[18]

A. Glossary of terms

We use the following terms throughout to analyze GII dependencies on various underlying contributions to atomic structure, composition, and stability:

- Energy above hull (E_{hull}): The energy per atom separating a given compound from the DFT-computed convex hull of ground state structures, sourced from the Materials Project (v2021.05.13).[19]
- Kernel density estimate (KDE): Kernel-smoothed, normalized estimate of a probability distribution used in this work to convey the relative distributions of materials properties, such as GII and energy above hull. We use SciPy's `gaussian_kde` function.[20]

B. Data sets

We used two datasets: one derived from the Materials Project (MP)[19] and the other derived from the Inorganic Crystal Structure Database (ICSD).[21] The MP dataset includes binary through quinternary compounds reported as synthesizable for which all components can be assigned a meaningful integer oxidation state. Figure 1 shows an overview of the chemical makeup of the MP dataset, revealing a strong bias toward ternary and quaternary oxides. The ICSD dataset includes binary and ternary compounds containing at least one transition, alkali, or alkaline earth metal as well as at least one of the following anions: N, O, F, P, S, Cl, As, Se, Br, Te, and I. These structures are all reported in the ICSD under the following conditions: 283 K $\leq T \leq$ 303 K and 0.9 atm $\leq P \leq$ 1.1 atm. A corresponding chemical composition plot for the ICSD dataset can be found in the supplementary material.

We also generated two *intersection datasets*, consisting of subsets of the MP and ICSD datasets for which there are matching entries in the other dataset. Cross-dataset matches were obtained using the `icsd_ids` feature from MP or by matching formula and AFLOW prototype.[22–24] In cases where there are multiple matching ICSD entries, we choose the most recent entry. The GII distributions

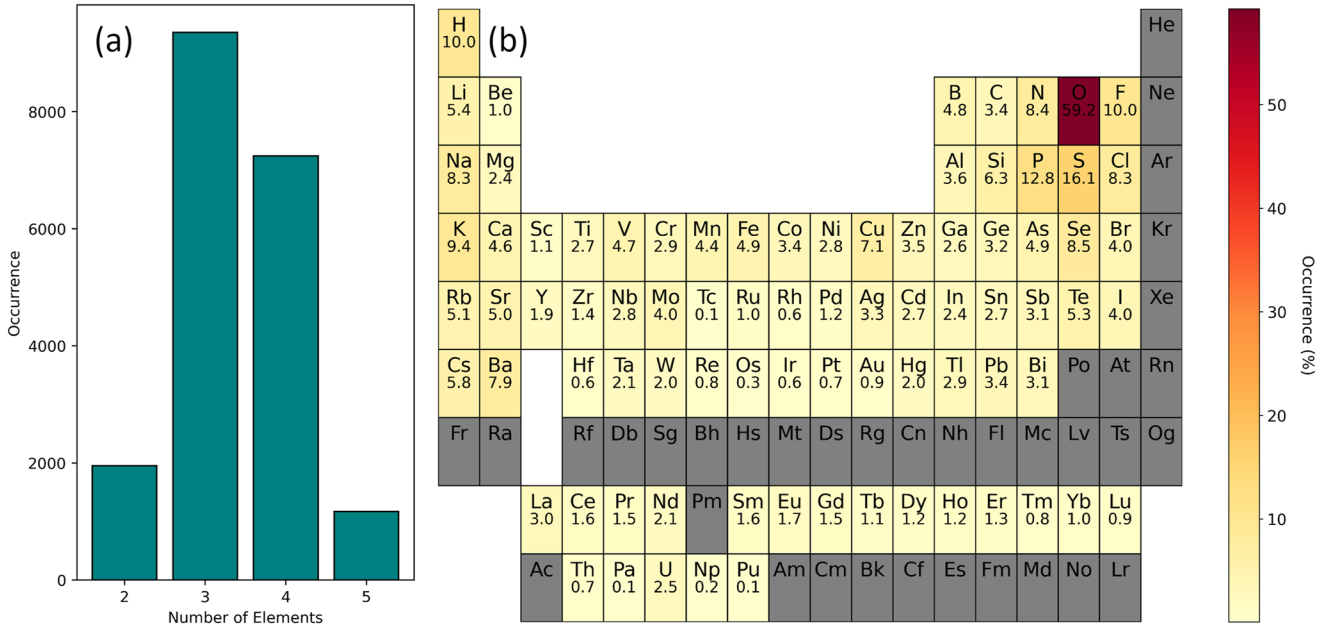


FIG. 1. (a) Histogram showing the population by number of elements and (b) a heatmap showing the frequency of elements in the MP dataset used for our GII assessments. The number below each element corresponds to the percentage of compounds that contain that element.

are nearly unaffected by choice of ICSD entry in these cases. Unless otherwise specified, the figures shown were generated from the MP dataset and the corresponding figures for the ICSD data can be found in the supplementary material.

C. Data processing

We employed the following workflow for data augmentation and cleaning:

1. Retrieve E_{hull} values from the Materials Project (for MP-derived entries only).
2. Standardize file formats and remove problematic files.
3. Predict oxidation states.
4. Calculate GII (both conventional and softBV variants).
5. Assign structural prototype tags to all structures to remove duplicates and to indicate isopointal groupings, i.e., groupings that exhibit the same space-group type (or belong to enantiomorphic space-group pairs) and the same (fully or partially) occupied Wyckoff sites.

Standardization was achieved by first determining the prototype and parameters of each structure using AFLOW-XtalFinder's `--prototype` function and then rewriting each structure using AFLOW-XtalFinder's `--proto` function. Files that triggered errors for either step were discarded. The oxidation states were predicted using Pymatgen's `ValenceIonicRadiusEvaluator` class. We calculated the conventional GII using the bond valence parameters from Brown's 2020 bond valence parameter set.[25] Structures for which any bond valence parameter was absent from the bond valence parameter set were left out. The first coordination shell was defined using CrystalNN[26] as implemented in Pymatgen.[17] For the softBV model, we calculated GII using the `cal-gii` function of the `softBV` executable (version 1.2.7).[27] We assigned structural prototype (isopointal grouping) tags by categorizing structures using the `--compare_structures` function from AFLOW-XtalFinder.[24,28]

III. RESULTS AND DISCUSSION

A. GII as a proxy for stability

In addition to the applications of GII for structural validation and phase stability, GII has been formulated as a pseudo-energy and applied successfully to simple compounds. Models using GII as a proxy for lattice energy or component of a semiempirical interatomic potential have found success in predicting the ground-state structures of MgO surfaces[29] and perovskite oxides.[15,30] However, GII may have limitations in smaller energy regimes. One study found GII to be insufficient to predict anion-order in perovskite oxides.[31] Furthermore, all of these studies are limited to relatively small datasets. None has performed a large-scale examination of GII and its dependence on stability-related features.

In Fig. 2(a), we find that the spread of the E_{hull} distribution increases with GII. In Fig. 2(b), we find that the percentage of materials predicted to be stable at room temperature, indicated by $E_{hull} = kT$ and $T = 300$ K, decreases with larger GII. Both of these trends indicate that GII is at least weakly correlated with stability. However, our data also show that many stable compounds exhibit high GII values and that not all compounds with very low GII values are stable in the specified structure near room temperature. We further find that compounds with GII ≤0.2 v.u. are not significantly closer to the convex hull on average [Fig. 2(a)]. The majority of experimentally observed structures exhibit GII values above the 0.2 v.u. threshold [Fig. 2(b)]. The existence of significant tails in the distributions with low GII values and the significant number of stable compounds with high GII values demonstrates that, at least on a broad scale, a universal GII threshold is *not* a reliable *global* predictor of phase stability. As Brown said,[1] "there is no particular reason to assume that (the 0.20 v.u.) limit will apply to all crystals."

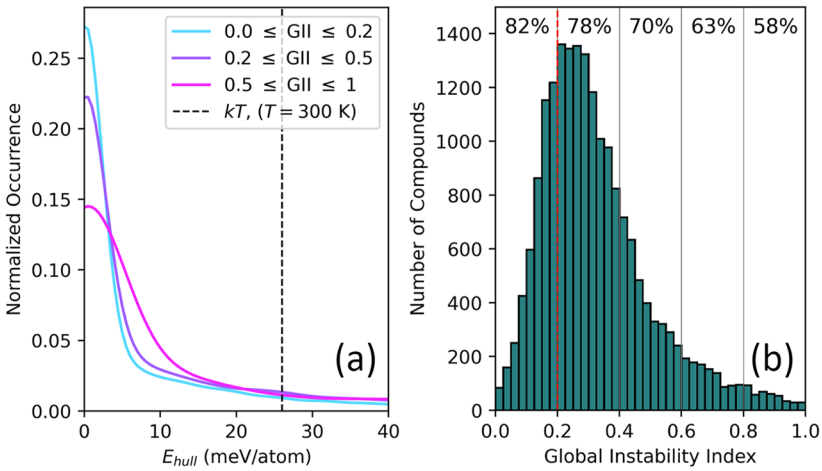


FIG. 2. Two plots showing that the 0.2 GII threshold is *not* strongly related to crystal stability. (a) KDE of computed energies above the convex hull for various segments of the GII distribution from the MP dataset. (b) Histogram of GII values segmented into regions for which the percentage of materials lying within 26 meV of the convex hull is shown at the top. The 0.2 v.u. GII threshold is represented by the red dotted line.

While GII remains a cost-effective, powerful tool for bonding analysis in limited chemical families, researchers should be aware of its limitations when applied to larger datasets.

B. GII dependence on chemistry

GII distributions also depend on the ligand species coordinating the metal atoms in a structure. Since the bond valence model underlying the GII measures conformity to Pauling's second rule, which assume ionic "bonding," i.e., electrostatic-based charge transfer, we would expect GII to be negatively correlated with the electronegativity of the anion. This hypothesis is supported in part by the GII distributions for our datasets, consistent with a chemist's expectation when moving across a row of the Periodic Table but not when moving down a column of the Periodic Table. In Figs. 3(a)–3(c), we show that the GII distributions generally skew to smaller values moving from group 7 to group 9 anions, i.e., upon increasing electronegativity. In Fig. 3(d), we collapse the GII distributions to median GII values to confirm that there is clear negative correlation between GII and electronegativity moving to the right in the Periodic Table. However, the correlation is absent when moving down a column. Both the KDEs in Figs. 3(a)–3(c) and the median GII values in Fig. 3(d) show inconsistency with the hypothesized GII values moving down the Periodic Table, i.e., from N → P → As, O → S → Se → Te, and F → Cl → Br → I. Rather than a negative correlation with electronegativity, our data suggest that, overall, GII is negatively correlated with the anion oxidation state. Further nuance likely exists within these chemical groups related to *sp* hybridization, lone pairs, and σ- vs π-bonding.

The presence of this correlation in the dataset as a whole does not mean it holds for individual structural families. By limiting our analysis to the rock salt structural family, for example, we can control for structural influences on GII to isolate the chemical contribution. In Fig. 4, we show GII KDE comparisons across rows of main group anions for rock salt compounds. Compounds with anions in row 2 support the overall trend [Fig. 4(a)], with GII tending to decrease as we move to the right in the Periodic Table. However, in rows 3 and 4, the trend is less well-supported [Figs. 4(b) and 4(c)], with the distributions being more similar and the GII peaks of the chalcogenides being lower and further toward large GII than those of the pnictides [Fig. 4(b)].

We showed on a large scale that the GII reflects broad bonding behavior trends across the anion oxidation state even beyond the domain of strongly ionic compounds but does not capture vertical periodic trends. These trends also show that average GII values increase markedly as we move inward in the Periodic Table—affirming Brown's warning about a universal GII threshold.[1] We conclude that GII can be useful in relative comparisons of any one factor in bonding behavior, e.g., covalency, bond strain, and bond multiplicity, even beyond strongly ionic systems, as long as the other factors are controlled.

C. GII dependence on geometry

We observe that GII distributions vary significantly between different structural prototypes, indicating that connectivity and bonding environment are important factors. Figure 5 shows that while some prototypes may exhibit similar GII distributions to each other, e.g., caswellsilverite, orthorhombic perovskite, and spinel, other prototypes, e.g., cubic perovskite, rock salt, and orthorhombic double perovskite, are both unique and different from the overall GII distribution.

To provide a more controlled analysis of the role of structure, we perform a comparison among the following four AB prototypes: NiAs, rock salt, wurtzite, and zinc blende. The main crystal chemistry features for these close-packed structure types are given in Table I. Examining the GII distributions in these prototypes (Fig. 6), we see that the NiAs family exhibits the largest GII values followed by rock salt, with wurtzite and zinc blende both featuring relatively low GII values. Referencing Table I and Fig. 6, we see that compounds with a higher coordination number and greater connectivity, i.e., face > edge > corner, feature a higher GII.

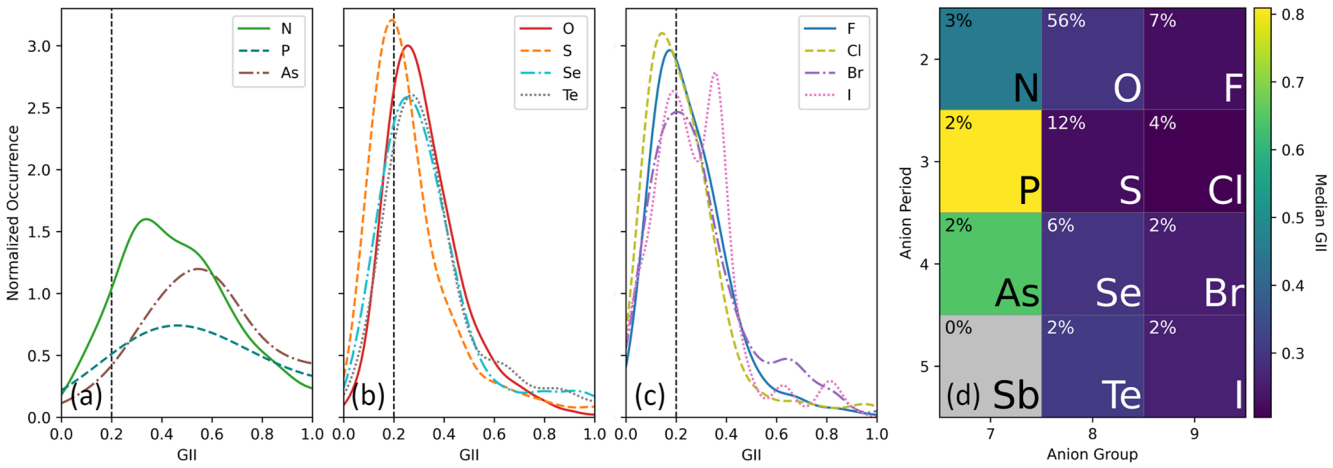


FIG. 3. KDEs of GII values for all (a) pnictides, (b) oxides and chalcogenides, and (c) halides in the MP dataset. The commonly adopted GII threshold is indicated by the black dashed line. (d) Median GII by anion. Percent occurrences of each anion are shown in the top left corners. The MP dataset contained too few antimonides ($N < 10$) to provide a meaningful distribution.

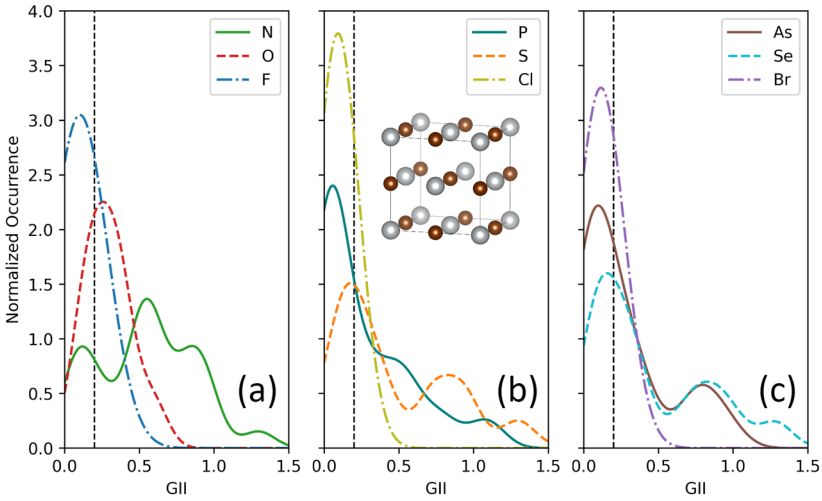


FIG. 4. KDEs of GII for rock salt-structure compounds with anions from (a) row 2, (b) row 3, and (c) row 4.

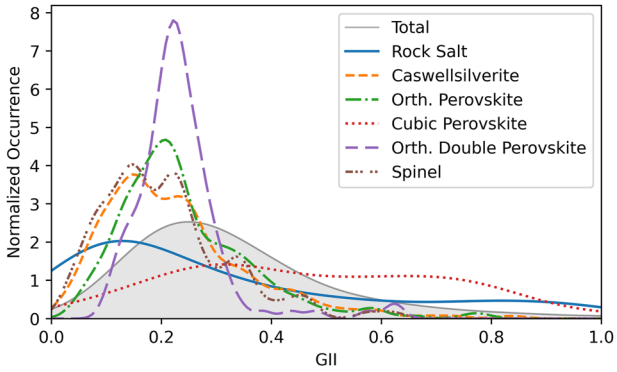


FIG. 5. KDEs of GII values for the top 6 most common structural prototypes in the MP dataset.

Possible sources of these correlations include metal–metal bonding, steric effects, and sampling bias in bond valence parameters. Face connectivity and edge connectivity provide more opportunity for bonding interactions between cations than corner connectivity. Since metal–metal bonding is not accounted for by the bond valence model, any such interactions would tend to increase the GII. For example, the first coordination shell of NiAs consists of 2.5 Å Ni–As bonds and the Ni–Ni distances are only 1 Å larger, close enough that some coordination environment calculators include those Ni–Ni interactions in the first shell.[26] The bond valence model, which does not account for cation–cation interactions, cannot capture the stabilizing influence of the metal–metal bonding that occurs between the face-sharing octahedra in the NiAs structure.[32] We find similar effects that occur in rutile-structured compounds. For example, NbO_2,[33] VO_2,[34] MoO_2,[35] RuO_2,[36,37] and IrO_2[36] all feature strong metal–metal interactions along the *c*-directed axis and, as a result, have inflated GII values (all above 0.14 v.u.).

TABLE I. Select characteristics of four common close-packed AB structures.

Structure	Coordination	Anion packing	Connectivity
NiAs	Octahedral	Hexagonal	Face
Rock salt	Octahedral	Cubic	Edge
Wurtzite	Tetrahedral	Hexagonal	Corner
Zinc blende	Tetrahedral	Cubic	Corner

Figure 7 displays a comparison of the electron localization functions (ELF) and GII values for two representative crystals. The first is MgO, an ionic crystal with well-separated electron clouds evidenced by the white regions of very low ELF values (<0.10), indicating the absence of electron sharing between neighboring ions in Fig. 7. This ionic material has strong, spherical localization around the anions and is well within the assumptions of bond valence theory. Accordingly, MgO exhibits a very low GII value of 0.002. MoO_2 provides an example of behavior beyond the scope of the bond valence model with more covalent cation–anion bonds and metal–metal bonding shown by the gray regions of larger ELF values ranging from 0.3 to 0.8 between Mo–O and Mo–Mo neighbors. Values in this range are consistent with electron localization expected for an electron gas and metallic bonding. This departure from the ionic model caused by these extended Mo-4*d* orbitals contributes to bond length deviation from the empirical bond valence parameters and, therefore, to a high GII value of 0.243. The lack of clear correlation and the significant population of structures on the hull for all GII values demonstrate that on large scales, GII is not an effective predictor of stability. This supports the idea that other factors contribute to the GII and its scope is more limited.

In addition, ions found in octahedral coordination tend to have radius ratios closer to unity, making alternate phases with larger coordination, such as CsCl, more competitive.[38] Since the bond valence parameters are calculated empirically, any bias present in the bond valence parameter set will appear in the GII values. Bond lengths show a significant variance, skew, and even multimodal-

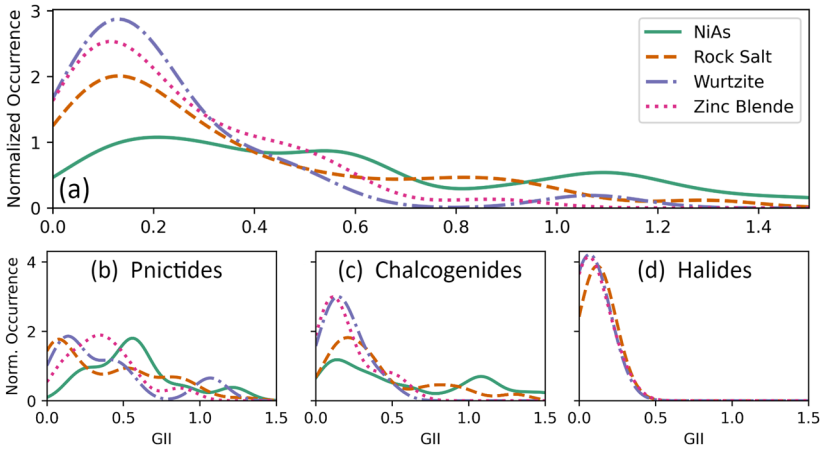


FIG. 6. KDEs of GII values for four AB structural prototypes shown for (a) all anions, (b) pnictides, (c) chalcogenides, and (d) halides.

ity (i.e., multiple maxima in the distribution of bond lengths) even for a single anion with a specific oxidation state and coordination number.[39–42] Since a single scalar value cannot accommodate such a variation, bond valence models necessarily sacrifice accuracy for simplicity and generality, especially for ions that are found in a variety of bonding environments. Gagné and Hawthorne observed this phenomenon directly in their analysis of bond length variation in oxides.[41] This simplification of the ideal bond length underlying the bond valence model is another explanation for the multimodality and inflated values of the GII distributions shown in Figs. 5 and 7.

Other sources of bond length variation identified by Gagné and Hawthorne include bond-topological effects, the Jahn–Teller effect (JTE), the pseudo-Jahn–Teller effect (PJTE), and the π-bond formation. These mechanisms for variations in GII appear in both

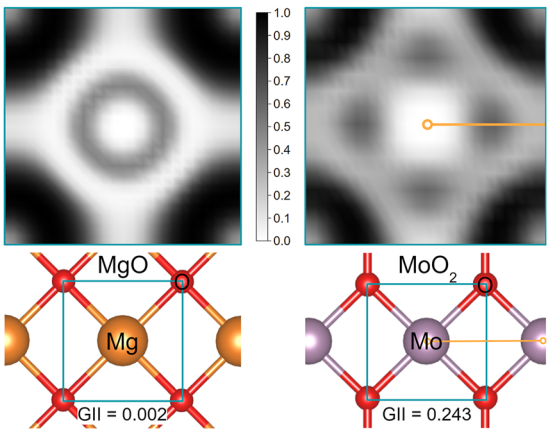


FIG. 7. Electron localization function (ELF) and crystal structure for (left) rock salt MgO and (right) rutile MoO_2. The teal outlines indicate the corresponding frames for the ELF and the crystal structure. The yellow lines indicate a metal–metal bond in MoO_2.

TABLE II. Examples of structures with large GII values due to the bond length variation.

Mechanisms	Example (MP ID)	GII (v.u.)
Bond topology	$SrV_2P_2O_9$ (mp-21614)	0.28
Jahn–Teller effect (JTE)	Cr_5S_8 (mp-1181961)	0.21
Pseudo-JTE	$DyTa_7O_{19}$ (mp-15492)	0.19
π-bond formation	$BaMo_2(PO_4)_4$ (mp-555876)	0.16

of our datasets. As shown in Table II, each of these phenomena can result in large GII values, above the commonly used 0.1 v.u. threshold for stable structures and even above the 0.2 v.u. threshold for incorrect structures.

As a response to the widespread use of GII as a measure of structural feasibility, we present a use case evaluating the efficacy of GII as a relative structural stability metric in a tightly controlled dataset. Figure 8 showcases the ability of GII to accurately predict the lower energy structural phase for a given compound. If GII were a good predictor of phase stability, all green and orange circles would lie in the green and orange regions, respectively. Instead, we observe an inconsistency of agreement between GII and E_{hull} for different pairwise comparisons. In Fig. 8(d), there is some agreement between GII and E_{hull}, with most of the green and orange points lying in the correspondingly colored region. However, the low number of common compounds with zinc blende ground states means that the agreement in this plot offers very weak support for the predictive power of GII on phase stability. The other plots in Fig. 8 show insufficient data and/or a clear lack of reliable correlation between GII and phase stability—manifesting as orange points in green regions and vice versa. Furthermore, the substantial spread along the diagonal of each plot reinforces that the variation in GII due to bonding type is much larger than the 0.2 v.u. threshold.

Separating GII distributions into more narrow materials classes, e.g., controlling for data source, BV model, structure type,

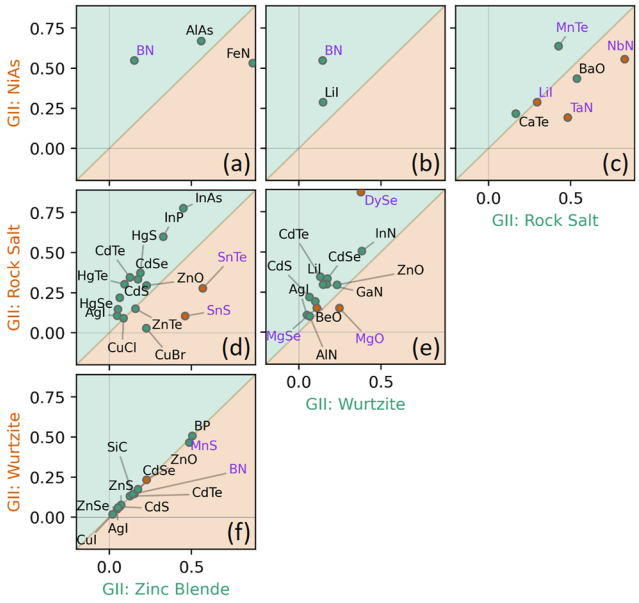


FIG. 8. Pairwise comparison of GII values for compounds that have been computed in at least two of the four AB prototypes. For any point on a plot, the *x* and *y* coordinates correspond to the GII value of the compound in that structure listed on the corresponding axis. Coloring of the point protectors indicates the lower energy structure according to DFT calculation. Coloring of the plot area indicates the region in which the relative GII values imply stability for the correspondingly colored phase, i.e., orange (green) regions indicate lower GII for the structural phase on the *y*axis (*x*axis). Thus, an orange point protector in the orange region indicates a compound for which relative GII *correctly* predicts the more stable structural phase, while an orange point protector in the green region indicates a compound for which relative GII *incorrectly* predicts the more stable phase. Compounds for which the more stable phase lies on the convex hull are labeled with black texts; otherwise, they are labeled with purple texts.

and ligand, generally tightens the spread of GII but does not necessarily remove multimodality, suggesting that additional factors remain. In Fig. 7, we find that the multimodality of a structure type's total GII distribution can be reduced to some extent by breaking down further by the anion oxidation state. We observe the expected dependence of the GII distribution, with halides featuring the lowest GII, clustered below 0.5, and pnictides featuring the lowest relative population of low GII structures. However, all of the subfamilies shown in Figs. 7(b) and 7(c) still exhibit a significant spread in their GII distributions.

This tendency of the distributions to tighten and move toward a monomodal distribution in the case of halides, as we control for more factors, suggests that GII can capture physical trends. The remaining variance indicates that the various factors affecting GII should first be disentangled before useful information can be extracted. So far, this has indicated that GII is only reliable for use in small datasets in which only one or two of the aforementioned factors are allowed to vary. To move beyond such narrow applications, one might derive a set of GII baselines to account for different chemical and structural regimes. If successful, these baselines could allow GII comparison on a larger scale more suitable for high-throughput materials informatics work.

D. GII dependence on model bias

As shown in Eq. (1), the GII depends on bond valence sums V_{VSR}, which take the form

$$V_{\mathrm{VSR},i} = \sum_{j=1}^{M_i} \exp\left(\frac{R_{ij} - R_0}{B}\right),$$

where M_i is the number of anions in the first coordination shell of cation i, R_{ij} is the distance between atoms i and j, and R_0 and B are the so-called bond valence parameters (BVPs). These bond valence parameters must be known *a priori* and are generally derived empirically from experimental observations. Several different ways of obtaining them have been proposed, which differ in the dataset used for derivation and the approximations made. Brown and Altermatt proposed a universal B parameter and optimized R_0 values for the standard deviation in R_0 to obtain the most consistent fit to the relevant cation environments in their ICSD dataset.[44] Gagné and Hawthorne optimized both B and R_0 for the root-mean-squared deviation between bond valence sum and the formal valence of the central cation—essentially the GII for a single polyhedron.[39] Chen and Adams moved beyond the consideration of only the first coordination shell and incorporated bond softness into their version (softBV), building off Pearson's concept of hard and soft acids and bases.[11,12,43,45] Although the best way to parameterize the bond valence model has been debated, the consequences of these nuances have not been extended to the GII.

Here, we provide a comparison between two of the most different models: the original GII based on Brown's conventional bond valence model and GII based on softBV, a revised model that captures bond softness[11] and considers interactions beyond the first coordination shell.[12,27,45] In Fig. 9(a), we show that GII values produced by the softBV model have a similar distribution shape but are generally lower, with the Brown and softBV models yielding median GIIs of 0.28 and 0.19, respectively. On a per-structure basis, we show in Fig. 9(b) differences between the GII values from each BVP model. The two GII types produce significantly different values for many structures. Most notably, the softBV model yields far fewer GII values above the 0.2 v.u. threshold. The tails beyond ±0.2 in Fig. 9(b) indicate that a single stability criterion, such as the 0.2 v.u. cutoff, would find contradictions between the two models. Out of the 16 070 structures for which both the Brown and softBV BVP sets can be applied to our MP dataset, 5644 (≈35%) cross over the 0.2 v.u. threshold when changing between bond valence models. In short, GII is sensitive to the approximations and simplifications built into the underlying model.

E. GII dependence on dataset characteristics

Both experiment- and simulation-sourced materials datasets have utility in ongoing efforts to characterize broad inorganic phase spaces. Simulated-sourced data can offer more standardized data and access to regions of phase space unavailable to experiment but contain bias from any approximations in the theory underlying the simulation. Experiment-sourced data can offer a higher fidelity but can contain variations due to differences in synthesis and/or measurement conditions. These trade-offs are reflected in our GII observations for both the MP and ICSD datasets.

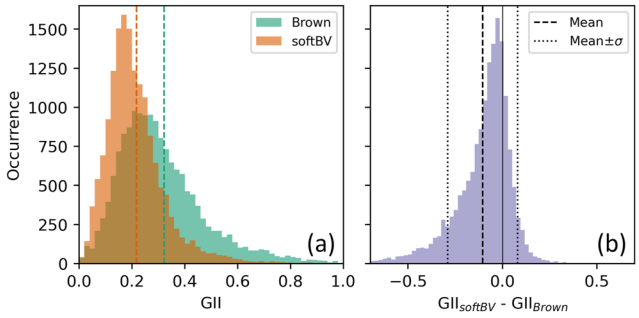


FIG. 9. (a) Histogram of GII values calculated using two different bond valence models for the MP dataset: the traditional model with Brown's bond valence parameters[25] and the softBV model.[12,43] Dotted lines represent the median GII. (b) Histogram of the disparities between GII values from the two bond valence models. Entries plotted are limited to the subset of compounds for which BVPs are available from both sources.

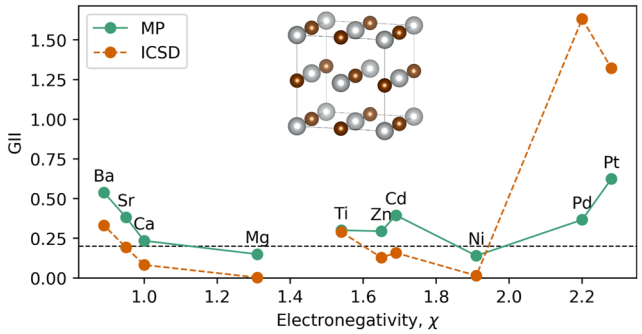


FIG. 11. GII dependencies on electronegativity for select rock salt structured oxides with the MP (DFT-calculated) and ICSD (experimental) datasets. The horizontal black dotted line represents the commonly adopted 0.2 v.u. GII threshold.

In Fig. 10, the GII distributions of the two datasets show us that structures in the ICSD dataset tend to have lower GII values on the whole and on a per-structure basis. Considering that the bond valence parameters used to evaluate GII were empirically determined using the ICSD data, we expect that the ICSD versions of structures tend to have a lower GII. In addition, any GII values calculated using *ab initio* structural data include the errors of the model used to perform geometric optimization, e.g., the complexity and completeness of exchange and correlation energies of the DFT functional. In some cases, this DFT error manifests systematically, allowing for trends to be preserved.[46] For example, the series of rock-salt-structured alkaline earth metal oxides shown in the left side of Fig. 11 display identical trends in both the ICSD and MP data. The MP data are approximately translated by a constant value. Given that these materials have large electronegativity differences and relatively simple electronic structures, it makes sense that the bond valence model would accurately capture the electronegativity trend and that DFT could replicate those relationships. While the trend is reproduced, these data further undermine the concept of the 0.2 v.u. threshold. In those nine compounds alone, four (SrO, CaO, ZnO, and CdO) yield GII values, which sit on either side of the threshold. This is an interesting observation because two researchers using the same threshold but different data sources for these compounds would produce predictions that disagree almost half the time.

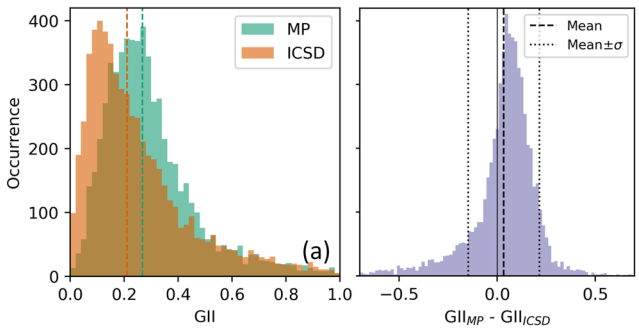


FIG. 10. (a) Histograms of GII values for the intersection datasets. The dotted lines represent the median GII for each dataset. (b) Histogram of the per-structure GII disparities between the MP and ICSD versions of the structures in the intersection datasets.

In compounds with more complex electronic structures, there is even less consistency in agreement between the MP (DFT-calculated) and ICSD (experimental) data. In the right half of Fig. 11, we show the GII values for several transition metal oxides, of which all feature electrons in *d* orbitals near their Fermi levels. For the transition metal oxides, we find a range of agreements between MP and experiment: from excellent agreement in TiO to systematic disagreement in ZnO, CdO, and NiO, to extreme disagreement in PdO and PtO. Notably, PdO and PtO also receive larger GII values for their experimental structures than their DFT structures in our MP dataset. Since PdO and PtO are unstable phases, both decomposing their dioxide counterparts, the high GII values can be ascribed to phase instability and difficulty in modeling the electron–electron interactions endemic to Pd and Pt metal oxides.[47–49]

Our comparisons of the MP and ICSD datasets have shown that both DFT and experimental GII values can capture chemical trends for limited datasets for strongly ionic crystals, consistent with the original assumptions of the bond valence model. Consistency between datasets is not guaranteed for compounds containing cations near the center of the Periodic Table where covalent bonding is expected to be stronger. It may be possible to develop a renormalization scheme to enable mixing of data from different sources, but the large variance in the per-structure residuals in Fig. 10(b) foreshadows a difficult task. We have identified the source of the structural data as yet another factor affecting GII.

IV. CONCLUSION

GII has successfully been used in a diverse set of contexts to explore bonding-related behaviors and capture the key factors of crystal structure stability. We highlight some additional factors that affect GII and suggest several precautions for its use. We have demonstrated that the various ways a compound can defy the bond valence model translate to variations in global instability index values across chemistry, structure, data source, and model bias. We also showed that GII depends on the partitioning of chemical bonds,

specifically the choice of coordination shells to be included, and the source of the structural data. Accordingly, GII can be used as a proxy for many materials properties but is only reliable when the other factors are controlled. We offer two caveats: (1) GII should not be used as an absolute metric but rather as a relative metric for comparison and (2) GII should only be used for limited datasets in which a baseline value can be established by controlling for the chemical, electronic, and structural factors. Moving forward, we see an opportunity in the development of a normalization scheme for GII to extend its generality. This could establish a middle ground between the simplicity of GII and the complexity of interatomic potentials—a regime well-suited to high-throughput screening.

In addition to its utility as a cost-effective, simple metric for small-scale, comparative investigations of bonding behavior, we posit that GII may also be useful as an indicator of electronic instabilities. In both our AB compound data and rutile analysis, we observed that GII is strongly correlated with nonionic interactions, e.g., metal–metal covalency, a hallmark of Peierls-type metal–insulator transitions. The search for such metal–insulator transitions constitutes a needle-in-the-haystack problem, where we must search a vast materials space for materials exhibiting a very rare property. In cases like these, any metric that can indicate a subset of materials as more likely to contain the property of interest can be useful as a way to focus the search, even if there are false positives and false negatives.

SUPPLEMENTARY MATERIAL

Additional computational methods, link to data and featurization code, and figures of various GII dependencies from the ICSD.

ACKNOWLEDGMENTS

We would like to thank Dr. David Hicks for answering our questions about AFLOW-XtalFinder. J.M.R. thanks K. Poeppelmeier for bringing his attention to the utility of GII for understanding crystal chemistry of inorganic compounds and providing comments on this manuscript. This work was supported, in part, by the National Science Foundation (NSF) under Award Nos. DMR-2011208 and DGE-1842165.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Kyle D. Miller: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Funding acquisition (equal); Investigation (equal); Methodology (equal); Visualization (lead); Writing – original draft (lead). **James M. Rondinelli**: Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material.

REFERENCES

[1] I. D. Brown, "Chemical and steric constraints in inorganic solids," Acta Crystallogr., Sect. B: Struct. Sci. **48**, 553–572 (1992).
[2] A. Salinas-Sanchez, J. L. Garcia-Muñoz, J. Rodriguez-Carvajal, R. Saez-Puche, and J. L. Martinez, "Structural characterization of R_2BaCuO_5 (R = Y, Lu, Yb, Tm, Er, Ho, Dy, Gd, Eu and Sm) oxides by x-ray and neutron diffraction," J. Solid State Chem. **100**, 201–211 (1992).
[3] *Bond Valences*, edited by I. D. Brown and K. R. Poeppelmeier, Structure and Bonding Vol. 158 (Springer, Berlin, Heidelberg, 2014).
[4] R. L. Withers, S. Schmid, and J. G. Thompson, "Compositionally and/or displacively flexible systems and their underlying crystal chemistry," Prog. Solid State Chem. **26**, 1–96 (1998).
[5] H. A. Evans, L. Mao, R. Seshadri, and A. K. Cheetham, "Layered double perovskites," Annu. Rev. Mater. Res. **51**, 351–380 (2021).
[6] M. W. Lufaso and P. M. Woodward, "Prediction of the crystal structures of perovskites using the software program *SPuDS*," Acta Crystallogr., Sect. B: Struct. Sci. **57**, 725–738 (2001).
[7] G. H. Rao, K. Bärner, and I. D. Brown, "Bond-valence analysis on the structural effects in magnetoresistive manganese perovskites," J. Phys.: Condens. Matter **10**, L757–L763 (1998).
[8] W. Feng, R. Zhao, X. Wang, B. Xing, Y. Zhang, X. He, and L. Zhang, "Global instability index as a crystallographic stability descriptor of halide and chalcogenide perovskites," J. Energy Chem. **70**, 1–8 (2022).
[9] A. B. Georgescu, P. Ren, A. R. Toland, S. Zhang, K. D. Miller, D. W. Apley, E. A. Olivetti, N. Wagner, and J. M. Rondinelli, "Database, features, and machine learning model to identify thermally driven metal–insulator transition compounds," Chem. Mater. **33**, 5591–5605 (2021).
[10] K. Kapera and A. Koleżyński, "First-principles study of structural disorder, site preference, chemical bonding and transport properties of Mg-doped tetrahedrite," Comput. Mater. Sci. **213**, 111681 (2022).
[11] R. G. Pearson, "Hard and soft acids and bases," J. Am. Chem. Soc. **85**, 3533–3539 (1963).
[12] S. Adams, O. Moretzki, and E. Canadell, "Global instability index optimizations for the localization of mobile protons," Solid State Ionics **168**, 281–290 (2004).
[13] I. Etxebarria, J. M. Perez-Mato, A. García, P. Blaha, K. Schwarz, and J. Rodriguez-Carvajal, "Comparison of empirical bond-valence and first-principles energy calculations for a complex structural instability," Phys. Rev. B **72**, 174108 (2005).
[14] L. L. Wong, K. C. Phuah, R. Dai, H. Chen, W. S. Chew, and S. Adams, "Bond valence pathway analyzer—An automatic rapid screening tool for fast ion conductors within softBV," Chem. Mater. **33**, 625–641 (2021).
[15] I. Yamada, A. Takamatsu, and H. Ikeno, "Complementary evaluation of structure stability of perovskite oxides using bond-valence and density-functional-theory calculations," Sci. Technol. Adv. Mater. **19**, 101–107 (2018).
[16] S. R. Hall, F. H. Allen, and I. D. Brown, "The crystallographic information file (CIF): A new standard archive file for crystallography," Acta Crystallogr., Sect. A: Found. Crystallogr. **47**, 655–685 (1991).
[17] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, "Python materials genomics (pymatgen): A robust, open-source python library for materials analysis," Comput. Mater. Sci. **68**, 314–319 (2013).
[18] K. Momma and F. Izumi, "VESTA: A three-dimensional visualization system for electronic and structural analysis," J. Appl. Crystallogr. **41**, 653–658 (2008).
[19] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, "Commentary: The materi-

als Project: A materials genome approach to accelerating materials innovation," APL Mater. **1**, 011002 (2013).
[20]P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, and P. van Mulbregt, SciPy 1.0 Contributors, "SciPy 1.0: Fundamental algorithms for scientific computing in Python," Nat. Methods **17**, 261–272 (2020).
[21]G. Bergerhoff and I. D. Brown, "Inorganic crystal structure database," in *Crystallographic Databases*, edited by F. H. Allen, G. Bergerhoff, and R. Sievers (International Union of Crystallography, Chester, 1987), pp. 77–95.
[22]M. J. Mehl, D. Hicks, C. Toher, O. Levy, R. M. Hanson, G. Hart, and S. Curtarolo, "The AFLOW library of crystallographic prototypes: Part 1," Comput. Mater. Sci. **136**, S1–S828 (2017).
[23]D. Hicks, M. J. Mehl, E. Gossett, C. Toher, O. Levy, R. M. Hanson, G. Hart, and S. Curtarolo, "The AFLOW library of crystallographic prototypes: Part 2," Comput. Mater. Sci. **161**, S1–S1011 (2019).
[24]D. Hicks, C. Toher, D. C. Ford, F. Rose, C. D. Santo, O. Levy, M. J. Mehl, and S. Curtarolo, "AFLOW-XtalFinder: A reliable choice to identify crystalline prototypes," Npj Comput. Mater. **7**, 30 (2021).
[25]I. D. Brown, Bond Valence Parameters, 2020, https://www.iucr.org/resources/data/datasets/bond-valence-parameters.
[26]H. Pan, A. M. Ganose, M. Horton, M. Aykol, K. A. Persson, N. E. R. Zimmermann, and A. Jain, "Benchmarking coordination number prediction algorithms on inorganic crystal structures," Inorg. Chem. **60**, 1590–1603 (2021).
[27]H. Chen, L. L. Wong, and S. Adams, "SoftBV—A software tool for screening the materials genome of inorganic fast ion conductors," Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater. **75**, 18–33 (2019).
[28]R. Allmann and R. Hinek, "The introduction of structure types into the inorganic crystal structure database ICSD," Acta Crystallogr., Sect. A: Found. Crystallogr. **63**, 412–417 (2007).
[29]J. A. Enterkin, A. E. Becerra-Toledo, K. R. Poeppelmeier, and L. D. Marks, "A chemical approach to understanding oxide surfaces," Surf. Sci. **606**, 344–355 (2012).
[30]S. Liu, I. Grinberg, and A. M. Rappe, "Development of a bond-valence based interatomic potential for $BiFeO_3$ for accurate molecular dynamics simulations," J. Phys.: Condens. Matter **25**, 102202 (2013).
[31]J. K. Harada, K. R. Poeppelmeier, and J. M. Rondinelli, "Predicting the structure stability of layered heteroanionic materials exhibiting anion order," Inorg. Chem. **58**, 13229–13240 (2019).
[32]S. Somiya, F. Aldinger, R. M. Spriggs, K. Uchino, K. Koumoto, and M. Kaneno, *Handbook of Advanced Ceramics: Materials, Applications, Processing and Properties* (Academic Press, 2003).
[33]A. O'Hara and A. A. Demkov, "Nature of the metal-insulator transition in NbO_2," Phys. Rev. B **91**, 094305 (2015).
[34]Z. Hiroi, "Structural instability of the rutile compounds and its relevance to the metal–insulator transition of VO_2," Prog. Solid State Chem. **43**, 47–69 (2015).
[35]D. O. Scanlon, G. W. Watson, D. J. Payne, G. R. Atkinson, R. G. Egdell, and D. S. L. Law, "Theoretical and experimental study of the electronic structures of $MoO_3$and MoO_2," J. Phys. Chem. C **114**, 4636–4645 (2010).
[36]J. Riga, C. Tenret-Noël, J. J. Pireaux, R. Caudano, J. J. Verbist, and Y. Gobillon, "Electronic structure of rutile oxides TiO_2, RuO_2, and IrO_2 studied by x-ray photoelectron spectroscopy," Phys. Scr. **16**, 351 (1977).
[37]N. Beatham and A. F. Orchard, "X-ray and UV photoelectron spectra of the oxides NbO_2, MoO_2 and RuO_2," J. Electron Spectrosc. Relat. Phenom. **16**, 77–86 (1979).
[38]A. Michmerhuizen, K. Rose, W. Annankra, and D. A. Vander Griend, "Radius ratio rule rescue," J. Chem. Educ. **94**, 1480–1485 (2017).
[39]O. C. Gagné and F. C. Hawthorne, "Comprehensive derivation of bond-valence parameters for ion pairs involving oxygen," Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater. **71**, 562–578 (2015).
[40]O. C. Gagné and F. C. Hawthorne, "Bond-length distributions for ions bonded to oxygen: Alkali and alkaline-earth metals," Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater. **72**, 602–625 (2016).
[41]O. C. Gagné and F. C. Hawthorne, "Mean bond-length variations in crystals for ions bonded to oxygen," Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater. **73**, 1019–1031 (2017).
[42]O. C. Gagné and F. C. Hawthorne, "Bond-length distributions for ions bonded to oxygen: Results for the transition metals and quantification of the factors underlying bond-length variation in inorganic solids," IUCrJ **7**, 581–629 (2020).
[43]H. Chen and S. Adams, "Bond softness sensitive bond-valence parameters for crystal structure plausibility tests," IUCrJ **4**, 614–625 (2017).
[44]I. D. Brown and D. Altermatt, "Bond-valence parameters obtained from a systematic analysis of the inorganic crystal structure database," Acta Crystallogr., Sect. B: Struct. Sci. **41**, 244–247 (1985).
[45]S. Adams, "Relationship between bond valence and bond softness of alkali halides and chalcogenides," Acta Crystallogr., Sect. B: Struct. Sci. **57**, 278–287 (2001).
[46]N. Wagner, D. Puggioni, and J. M. Rondinelli, "Learning from correlations based on local structure: Rare-earth nickelates revisited," J. Chem. Inf. Model. **58**, 2491–2501 (2018).
[47]K.-T. Park, D. L. Novikov, V. A. Gubanov, and A. J. Freeman, "Electronic structure of noble-metal monoxides: PdO, PtO, and AgO," Phys. Rev. B **49**, 4425–4431 (1994).
[48]R. Ahuja, S. Auluck, B. Johansson, and M. A. Khan, "Optical properties of PdO and PtO," Phys. Rev. B **50**, 2128–2132 (1994).
[49]L. K. Lamontagne, G. Laurita, M. W. Gaultois, M. Knight, L. Ghadbeigi, T. D. Sparks, M. E. Gruner, R. Pentcheva, C. M. Brown, and R. Seshadri, "High thermopower with metallic conductivity in *p*-type Li-substituted $PbPdO_2$," Chem. Mater. **28**, 3367–3373 (2016).