Research Update: Towards designed functionalities in oxide-based electronic materials

James M. Rondinelli,[1,a] Kenneth R. Poeppelmeier,[2,b] and Alex Zunger[3,c]

[1]*Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208, USA*
[2]*Department of Chemistry, Northwestern University, Evanston, Illinois 60208, USA*
[3]*University of Colorado, Boulder, Colorado 80309, USA*

One of the grand challenges facing materials-by-design approaches for complex oxide deployment in electronic devices is how to balance transformative first-principles based predictions with experimental feasibility. Here, we briefly review the functionality-driven approach (inverse design) for materials discovery, encapsulated in three modalities for materials discovery (m^3D) that integrate experimental feedback. We compare it to both traditional theoretical and high-throughput database-directed approaches aimed at advancing oxide-based materials into technologies. ○

integer stoichiometry in compounds or continuous variable, "subscript engineering," in alloys and ceramics), and **S**tructure (i.e., where the atoms are as pertaining to crystallographic structure, or broadly, from nanoscale to microscale and mesoscopic morphology).

Property prediction has made tremendous strides owing to (i) advanced material-dependent electronic structure methods that (unlike material agnostic model Hamiltonian theories) accept ACS as input: primarily, density functional theory (DFT) and its ACS-recognizing extensions that incorporate many-electron e ects (Configuration Interaction, Dynamical Mean Field Theory, GW, and Quantum Monte Carlo methods);[5–9] and (ii) new protocols to compute complex functionalities, combining a few target properties such as dopeability,[10,11] coexistence of transparency and conductivity,[12–14] and topological insulation.[15,16]

discovery of *unreported* compounds ("missing materials") with desirable properties (so, stability and stable crystal structures need to be computed). The distinction between the various modalities stems in part from the different theoretical approaches required: $m1$ focuses on properties of an artificially grown, kinetically frozen layer sequence which need not be globally stable as long as it is locally stable, while $m2$ (sometimes) and $m3$ (always) require thermodynamic structure stability calculations. Also, $m1$ involves exploration of the almost infinite number of possible assemblies of building blocks (as pulsed-laser deposition (PLD) and molecular beam epitaxy (MBE) layer sequences) and thus cannot be addressed, even in principle by database approaches that attempt to enumerate and compute all structures. Instead, $m1$ would require *search methods* that do not require

50 configurations to construct a robust cluster expansion.[35] The outcome was that new configurations consistent with recent design guidelines for obtaining additional contributions to ferroelectric polarizations in perovskites through improper mechanisms[38] immediately appear. Thus, $m1$ makes it possible to both identify optimal chemical sequences to be recommended for experimentation and to discover novel physical mechanisms in cases where such knowledge may not exist for property enhancement. Feedback from experiment to theory includes observations on the degree on interfacial intermixing—a fact that can be incorporated in the next step theory as a constraint on the search procedure (such as "find configurations with the target properties *and* interfacial sharpness not exceeding ").

m2—Searching Chemical Compound Space for Target Functionality. Electronic based properties of most oxides such as photovoltaic absorption, transparent conductivity, photo-electro-chemical water-splitting ability, or thermoelectricity are intimately tied to metal-oxide coordination preferences and polyhedral connectivity, which change with cation-anion ratios and structure in a material. Modality 2 explores bulk materials where structures and compositions are known, e.g., from previous studies, but the properties for a set of atomic identities are unknown. The concept here is to search through ACS-space (albeit, within a specific compositional group such as A_2BO_4, ABO_3, and ABO_2) to design a material with the set attributes that give the optimal value to the functionality metric. The steps then involve (a) identifying a calculable FM that is simpler to compute than the final functionality sought; (b) calculating the FM in the ACS space while (c) navigating in the space of compounds to specifically identify the "best of class" compounds deemed to have the highest probability for laboratory realization through some evaluation of material stability. The rationale for the above functionality-driven strategy is that it is much faster to screen the functionality assuming a given crystal structure than to investigate the existence of a functionality and the ground state crystal structures simultaneously.

For example, a quantum-mechanical generalization of the Shockley-Queisser criteria, called the "Spectroscopic-Limited Maximum E ciency" (SLME) functionality metric, was created in step (a). It was tested on several hundred ternary materials using $m2$ to identify candidate photovoltaic absorbers,[18] followed by exploration of a select chemical group of semiconductors in search of maximal SLME. Note that $m2$ is distinct from making *all* possible chemical substitutions and optimizing over that set as in high-throughput searches. In $m2$, there is no need to compute many properties of a broad range of compounds, albeit that information could be useful to data-driven discovery approaches. In $m2$, one thus uses the strategy of narrowing down a large number of an *a priori* set of compounds by using the FM as a preliminary initial filter. This filter does not address all aspects of the functionality (such as material performance, durability, and economic factors), but focuses on a number of "first-filter," *critical attributes* needed for a particular functionality (such as FM = dopeability for making junctions or FM = strong SLME absorption for solar cells, depending on the need). Following such initial screening, one is left with a handful of candidate materials. The second step involves application of more accurate high-level methods (such as beyond-DFT methods). Now the calculation of a broader range of material functionality that transcends the simpler principles involved in the FM may occur; e orts to include such features as durability, reliability, and economics are already appearing.[17] It is important to emphasize that this di cult detailed screening occurring in the second step is required only for a much smaller group of "best candidates," rather than for all materials, thus providing a significant discovery acceleration factor. In the last step, the "best of class" materials are handed over to experimental validation, which very often leads to a refined FM, to be used in the next iteration. This protocol defines an iterative loop, whereby theory and experiment work hand in hand.

Examples of $m2$ applications include both successful confirmation of known absorbers and identification of previously unappreciated new candidate compounds that have been proposed and experimentally realized.[39] Another example is the use of FM = dopeability to search the space of

data quickly. Materials selection is made *ex post facto* and based on sorting/ranking a pre-defined figure-of-merit and occur within a limited ACS space.[55] Thus, high-throughput approaches make it possible to identify previously known but unidentified functional materials from a database of known compounds. New understanding may then be gleaned from the down-selected compound and become useful for practitioners of first-principles materials-by-design methods that fall outside the mentioned modalities to predict new functional materials.[56] The functionality-directed approach m^3D is distinct from the high-throughput approach in that if it focuses directly on the subset of materials that has the required target property rather than attempting all combinations. Computationally speaking, the m^3D approach then focuses on optimization theory and search methods, whereas the high-throughput focuses on fast direct computations and databases.

Given availability of such data, more researchers are applying big data science analytics to address the materials chemistry and structure-property challenges.[57] We contend, however, that it remains to be determined whether the real challenges in oxide-based electronics are of the sort suitable for only data-driven scientific discovery approaches. For instance, the subject of m1—di erent configurations of artificially grown structures—cannot be captured even in principle by any database because of the limitless configuration space. Second, the data sets, features, and prior knowledge in condensed matter systems are significantly smaller than those available to biology and health, communication networks, and finance, where big data science dominates. One therefore needs to consider nuances in the informatics methods for materials science data[58–60] and whether

[1] I. Osborne, M. Lavine, and R. Coontz, Science **327**, 1595 (2010).

[2] G. Thomas, Nature **389**, 907 (1997).

[3] This became particularly evident at two workshops focused on oxide electronics in 2014: The 21st International Workshop on Oxide Electronics held in Bolton Landing, New York, and the Advances in oxide materials: Preparation, properties, performance hosted by University of California, Santa Barbara, California, USA. Surely dozens or more related symposia have explored similar topics, but these meetings provided the impetus for the perspective which is by no means intended to be an exhaustive review of computational materials discovery approaches.

[4] J. F. Nye, *Physical Properties of Crystals: Their Representation by Tensors and Matrices* (Oxford University Press, New York, NY, 1985).

[5] M. Ediger, G. Bester, A. Badolato, P. M. Petro, K. Karrai, A. Zunger, and R. J. Warburton, Nat. Phys. **3**, 774 (2007).

[6] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986); M. van Schilfgaarde, T. Kotani, and S. Faleev,