Preface

Accelerating materials discovery has been the theme of a number reports from the Department of Energy's Office (DOE) of Basic Energy Science (BES), the National Science Foundation (NSF), the National Academies and other government agencies and professional societies. As a driver for accelerating materials discovery, the Materials Genome Initiative, announced by the President, is part of a bold plan to boost US manufacturing over the next few decades by halving the time it takes to discover and design new materials. In this plan, accelerating discovery relies on using in the material sciences large databases, computation, mathematics, and information science in a manner similar to the way that they were used to make the Human Genome Initiative a success for the biological sciences. Novel approaches are therefore being called for that can explore the enormous phase space presented by complex materials and processes. If we are to achieve the desired performance gains, then we must have a predictive capability that can guide experiments and computations in the most fruitful directions by reducing the possibilities that need to be tried.

Despite advances in computational and experimental techniques to generate large volumes of data to screen the vast search space, it is clear that the outstanding challenge remains to integrate information-theoretic tools and materials knowledge, in the form of constraints imposed by theory, to develop, robust, predictive tools for materials design and discovery. The rapidly emerging field of materials informatics provides the critical methodology that enables the discovery, identification, and harnessing of the materials "genes" for accelerated materials discovery and design.

We provide in this book a collection of articles in this nascent field which integrates contributions from the information sciences and materials communities. The collection is partly derived from a workshop held at Santa Fe, New Mexico, February 4–7, 2014 that was organized by the editors and sponsored with support from the Centers for Nonlinear Studies and Information Science and Technology at Los Alamos National Laboratory and National Science Foundation (Grant #: 13-07811). It outline*s* challenges and opportunities in the use of information-theoretic tools and

evaluates the state of the art on a number of materials-motivated problems. Presented are contrasting but complementary approaches, such as those based on high-throughput calculations or experiments, as well as data-driven discovery, together with the merits and challenges of machine-learning and statistical inference methods to accommodate searches within a high dimensional feature space.

The book is organized into *three* parts. In the *first* part, following a perspective of the state of the art in materials design and discovery, Chaps. 2–6 focus largely on information-theoretic tools and how they apply to specific materials problems. Chaps. 2 and 3 discuss how aspects of decision theory within a Bayesian framework can be used for optimal experimental design. In particular, Chap. 2 discusses how to decide on the best pair of experiments for inferring the parameters of a given model, as well as how to choose an experiment to distinguish between competing models. Chapter 3 discusses strategies based on methods for global optimization for choosing the next experiment to find a material with a desired property. Proceeding from problems involving regression to those requiring classification, Chap. 4 focuses on Bayesian methods for classifying objects, especially in the limit of small samples where classifier design procedures, which work well with large samples, can have problems when data is limited. The first part of this monograph is concluded with Chaps. 5 and 6 which deal with different aspects of clustering. Chapter 5 considers the effectiveness of data visualization algorithms that look for groupings of features and materials. Chapter 6 discusses how community detection, studied in statistical physics, can be used to partition a complex system into decoupled subsets at different spatial and temporal scales.

The focus of the *second* part of the book, Chaps. 7–12, is the application of informatics tools to materials science problems. Chapter 7 discusses how parameters in the additive manufacturing process may be constrained by combining simulations and experiments using feature selection and data-driven models. Learning from high-throughput data generated from electronic structure calculations is the emphasis of Chaps. 8–11. Techniques such as principal component analysis (PCA), support vector regression (SVR), partial least squares, and Kriging using Gaussian process modeling suggest new features and materials with specified properties. Chapter 8 shows how suitable dopants in an oxide may be identified for increasing water-splitting processes. Applications in Chap. 9 include the discovery of cathode materials for lithium-ion batteries and thermoelectrics. Chapter 10 focuses on the layered compounds known as MAX phases, and Chap. 11 discusses *ab initio* methods and applied crystallography tools for descriptor development to establish structure–property relationships. Chapter 12 describes hybrid methods that integrate statistical learning techniques, to extract features from the density of states for predicting elastic properties, such as bulk modulus and yet unexplored chemistries.

The *third* and final part, Chaps. 13 and 14, discusses high-throughput experiments, which generate large amounts of data. With appropriate characterization tools, the idea is to quickly identify the subspace of the large parameter space where a new compound with desired properties may be found. Such experiments, together with informatics tools, provide opportunities for “combinatorial materials science.”

Chap. 13 provides a review in the context of multifunctional materials, and Chap. 14 incorporates aspects of informatics with a focus on solar fuel applications and multicomponent oxide catalysts.

The book is aimed at an interdisciplinary audience as the subject spans aspects of statistics, computer science, and materials science and will be of timely appeal to those interested in learning about this emerging field. We are grateful to all the authors for their articles as well as their support of the editorial process.

Los Alamos, USA — Turab Lookman
Los Alamos, USA — Francis J. Alexander
Buffalo, NY — Krishna Rajan

http://www.springer.com/978-3-319-23870-8

Information Science for Materials Discovery and Design
Lookman, T.; Alexander, F.J.; Rajan, K. (Eds.)
2016, XVII, 307 p. 134 illus., 88 illus. in color.,
Hardcover
ISBN: 978-3-319-23870-8