Journal of Materials Chemistry A

A Bayesian Method for Selecting Data Points for Thermodynamic Modeling of Off-Stoichiometric Metal Oxides

Journal:	*Journal of Materials Chemistry A*
Manuscript ID	TA-ART-10-2023-006627.R2
Article Type:	Paper
Date Submitted by the Author:	29-Mar-2024
Complete List of Authors:	Wilson, Steven; Arizona State University, SEMTE Muhich, Christopher; Arizona State University Ira A Fulton Schools of Engineering, Chemical Engineering

A Bayesian Method for Selecting Data Points for Thermodynamic Modeling of Off-Stoichiometric Metal Oxides

Steven A. Wilson[1], Christopher L. Muhich[1,2]*

[1] Chemical Engineering, School for the Engineering of Matter, Transport, & Energy, Arizona State University, 551 E. Tyler Mall, Tempe Arizona, 85287, USA

[2] Materials Science & Engineering, School for the Engineering of Matter, Transport, & Energy, Arizona State University, 551 E. Tyler Mall, Tempe Arizona, 85287, USA

* Corresponding Author

Abstract

Thermodynamic characterization of metal oxide reduction/re-oxidation plays a vital role in material identification and optimization of many chemical processes. However, this characterization generally requires significant data collection (spanning several hundred T, pO_2, and composition (X) combinations) to appropriately sample phase space and identify key inflection zones that are not known *a priori* and are missed without the sampling of a fine mesh grid of T, pO_2, and X combinations. Here we have coupled our previously reported CrossFit Compound Energy Formalism algorithm for reduction/re-oxidation thermodynamic model fitting with Bayesian Inference techniques to build an *optimized* data selection scheme. Using the $Ba_xSr_{1-x}FeO_{3-\delta}$ system as a proof of concept, we show that our Bayesian data selection technique required less than half (44) data points to achieve the same accuracy as a mesh grid of 100 T, pO_2, and X point combinations. Our method has errors of < 2 kJ/mol in reduction enthalpy $\left(\frac{\partial H}{\partial \delta}\right)$ and < 3 J/(mol K) difference in reduction entropy $\left(\frac{\partial S}{\partial \delta}\right)$ compared to the full data set. Further, randomly selected

44 T, pO_2 and X data points only reproduced the ground truth model 5% of the time, demonstrating the power of our approach. Our method offers a human free, physically informed, data collection approach and paves the way for a high-throughput active data selection process for metal oxide reduction/re-oxidation thermodynamics.

1 Introduction

The cyclic reduction/re-oxidation of metal oxide (M_xO_y) materials forms the backbone of a vast number of chemical processes, such as gas reforming[1, 2], gas separation and pumping[3-9], energy production and storage[10-17], and many catalytic processes, particularly those which operate on Mars-van Krevelen-like mechanisms[18-24]. The performance of these materials is directly related to their partial molar reduction thermodynamics, i.e. enthalpies and entropies of reduction as a function of reduction ($M_xO_{y-\delta}$). Importantly, the subtle differences between reduction enthalpies and entropies across varying compositions of the same M_xO_y can have a measurable effect on material performance[12, 25, 26]. Therefore, accurate thermodynamic modeling of these M_xO_y materials underpins the design and optimization of M_xO_y for their respective chemical processes. This paper describes a new method for selecting data points which should be collected to extract these crucial thermodynamic parameters.

Detailed thermodynamic characterization of M_xO_y reduction/re-oxidation generally consists of careful thermogravimetric analysis (TGA) of mass loss as a function of temperatures and pressures which are fit to a thermodynamic model. The most commonly used model is based on van't Hoff analysis[27-31] , which is relatively simple to implement but does not provide temperature dependent information. Conversely, the compound energy formalism (CEF) is more robust albeit more complicated to fit. Using either model, fitting of thermodynamic parameters to TGA experimental data requires the measurement of non-stoichiometry across a large array of temperatures (T) and oxygen partial pressures (pO_2). It is unclear *a priori* what points should be examined, aside from the need to span a large range of operating points. The difficulty in selecting points arises because the non-stoichiometry, and the underlying thermodynamics are often highly non-linear[11, 14] and obviously unknown *a priori*. Thus, the use of evenly spaced points in

measurement points, T, pO_2 and compositional changes, generically referred to here as X, which we call (TpOX), either require fine search meshes, and thus significant time in synthesis and characterization, or luck to adequately sample inflection zones. Therefore, the development of a model informed data collection scheme for M_xO_y reduction thermodynamics would substantially decrease the time needed for material synthesis and TGA characterization and enable a high throughput material screening process. A clear benefit to such a scheme is that it will highlight important compositional changes to investigate as these bear the highest cost in experimental data collection. However, while TGA programming may be easily adjusted from one run to the next, specific X, T, and pO_2 points around inflection zones are of great interest, but not known *a priori*.

Active data selection techniques have been studied for several decades in a variety of scientific disciplines[32-35] such as medical disease analysis[36, 37], weather prediction[38], and material design[39-41]. To date, these have not been applied or refined for thermodynamic characterization of M_xO_y reduction/re-oxidation cycle materials. In this work we aim to fill this methodological gap through a novel algorithm which systematically selects the most important experimental data points to collect without prior knowledge of the system behavior. This method uses a combination of our recently developed cross-fit experimental and computational CEF model[42] and Bayesian inference[43-46] to select the data points to be experimentally examined and guide the fitting of robust thermodynamic models. In our algorithm, we leverage *ab initio* methods to estimate material enthalpies while the complex temperature and entropy interactions arise from TGA data. Overall, we provide a hands-off approach that: 1) requires the collection of less computational and experimental data than would be necessary if collected systematically on a grid, and 2) informs the experimentalist when additional data is required to build a more robust thermodynamic model. Furthermore, we note not all M_xO_y behave alike, with some varying greatly in reduction

thermodynamics as a function of composition or even undergoing phase transitions. The Bayesian data selection approach is inherently built to "seek out" this variation in trends as data is added during each iteration guiding the researcher towards these behavior altering regions.

As a demonstration of this method, we examine $Ba_xSr_{1-x}FeO_{3-\delta}$ (BSF) because we have a large existing experimental data set[47] and have previously established its thermodynamics. BSF reduces and oxidizes according to eqn 1& eqn 2

$$Ba_xSr_{1-x}FeO_{3-\delta 1} \rightarrow Ba_xSr_{1-x}FeO_{3-\delta 2} + \frac{\Delta\delta}{2}O_2 \quad \text{eqn 1}$$

$$Ba_xSr_{1-x}FeO_{3-\delta 2} + \frac{\Delta\delta}{2}O_2 \rightarrow Ba_xSr_{1-x}FeO_{3-\delta 1} \quad \text{eqn 2}$$

Where $\Delta\delta$ is the change in the non-stoichiometry of BSF at new TpOX conditions. We will detmerine how our Bayesian informed data selection scheme requires less TpOX points to derive the same thermodynamic results as a ground truth model utilizing all available TpOX data. Furthermore, we will show that the Bayesian informed approach performs better than pure random sampling as we repeatedly modeled a set of randomly selected experimental data points and compared the derived thermodynamics to the Bayesian selected set and the ground truth models.

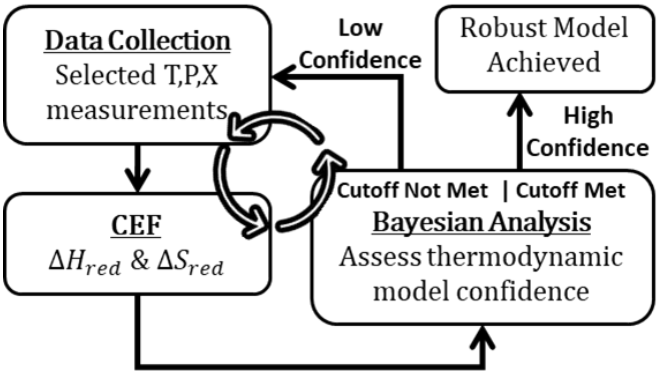

Figure 1: Schematic of the information flow in the proposed data selection scheme.

2 Methods

We propose using two Bayesian approaches to identify the appropriate thermodynamic model and optimize data collection: Bayesian Information Criterion (BIC)[48, 49] and a new Bayesian implementation for data selection. The BIC determines 1) which CEF model is the most

appropriate from an array of optimized models with varying numbers of excess terms and 2) when sufficient data has been collected, i.e. when the CEF model is trustworthy. The new proposed Bayesian determines the next TpOX sample point based on the collected data points and the associated uncertainty in the predicted experimental data across the current array CEF models. Figure 1 shows the proposed iterative cycle.

In this section we first briefly explain the CEF construction and the associated generation of multiple thermodynamic models from our previously described cross-fitting algorithm.[42] We then describe the use of BIC for selecting the "best" model. Subsequently, we outline our new Bayesian approach for selecting the next TpOX data point to be sampled, and the associated stopping criteria for the iterative loop. Our intention is not to provide detailed derivation of Bayesian Inference techniques but rather to outline the integration of the method into our workflow. For an in depth understanding of Bayesian Inference as it applies to model comparison, we direct the reader to the excellent O'Hagan and Foster text on the subject[50]. Finally, we give a brief explanation of the data sources for our models.

2.1 Overview of CEF

We use the CEF[51, 52] approach to represent the Gibbs free energy of a solid as a solution on a set of sub-lattices via the summation of three terms(eqn 3): 1) a linear combination of the Gibbs free energies of the so-called endmember compounds representing the composition of the solid solution, 2) a configurational entropy term, and 3) an excess term that accounts for interactions on and between the sub-lattices and accounts for any deviations from ideality as prescribed by the endmember and configurational entropy terms. We note that the configurational entropy term describes full disorder. Real materials may have a significant extent of short-range ordering (SRO), and thus lower entropy. While methods exits that incorporate SRO within the CEF[53], accounting

for such effects is not trivial and beyond the scope of our studies. In both $G^{endmembers}$ and G^{excess} the free energy is expressed as an expansion with constant heat capacity (Cp = C) (eqn 4).

$$G^{soln} = G^{endmembers} - T * S_{config} + G^{excess} \quad \text{eqn 3}$$

$$G = A + B * T + C * T * \ln(T) \quad \text{eqn 4}$$

Where the parameters A, B, and C, in eqn 4, are derived from the integration of a constant heat capacity (Cp = C). The enthalpy integrated from heat capacity as, H = $\int CdT = c1 + CT$ and the entropy integrated from heat capacity as: S = $\int \frac{C}{T} dT = c2 + Cln(T)$. Utilizing the state function $G = H - TS$ and combining like terms to simplify the parameter space we arrive at eqn 4.

The terms from eqn 3 and eqn 4 are defined by eqn 5 - eqn 8:

$$G^{endmembers} = \sum_{i}^{N} \prod \gamma_M^z G_i^{endmember} \quad \text{eqn 5}$$

$$S_{config} = -R \sum_{z} n_z \sum_{X} \gamma_M^z * ln(\gamma_M^z) \quad \text{eqn 6}$$

$$G^{excess} = \sum_{h} \gamma_1^h \gamma_2^h \sum_{k \neq h} \sum_{M=1}^{2} \gamma_M^k \sum_{l \neq h \neq k} \sum_{M=1}^{2} \gamma_M^l \mathcal{L}_{h:k:l} \quad \text{eqn 7}$$

$$\mathcal{L}_{h:k:l} = \sum_{\nu=0}^{1} (\gamma_1^h - \gamma_2^h)^{\nu} L_{h:k:l}^{\nu} \quad \text{eqn 8}$$

where γ is the site fraction of a species on a sublattice site, N is the total number of endmember terms, n is the total number of sites, z is a particular sub-lattice, and M counts over the components that can occupy a site on sublattice z. eqn 8 is configured for the BSF system which is a three sublattice model with two components per sublattice, and h, k, and l are the sublattices. A Redlich-Kister expansion expresses the $\mathcal{L}$ terms in G^{excess} [54] of the γ site fraction terms up to order m, taken

to be ν=1. For further development and fundamental analysis of the CEF model, we direct the reader to Ref. [42, 51, 52, 55-58]. In the SI, Table SI-1, we list the Redlich-Kister expansion $\mathcal{L}$ terms.

Constructing the CEF model as shown for any given M_xO_y systems results in several excess $L^{\nu}_{h:k:l}$ terms each carrying their own A, B, and C parameters to be optimized. Thus, the CEF fitting model provides a range of thermodynamic prediction models based on how many excess terms are included. We have previously demonstrated how to down select which excess terms should be included up to q excess terms using cross-fitting and a down selection logic model[42]; however, the fitting procedure does not identify which q best describes the data, i.e., how many excess terms should be included.

We show the $\frac{\partial H}{\partial \delta}$ and $\frac{\partial S}{\partial \delta}$ model trends for all optimized models on the full experimental (ground truth) BSF dataset in Figure 2a-b. Commonly, the selection of the most appropriate model is based on visual comparison of thermodynamic trends and error comparisons. This human intervention can lead to biasing of the model based on the choices of the user. Therefore, we propose using a statistical BIC method to select the most appropriate CEF model.

2.2 Determining the most appropriate model using BIC

The proper dimensionality of a model when fit to a non-predetermined functional form is often difficult to discern[59]. The BIC circumvents the unknown dimensionality problem by balancing the model accuracy (based of the observed data) against the number of parameters utilized by the model, seen as eqn 9:

$$BIC = \kappa * ln(N) - 2 * \ln(\ell) \quad \text{eqn 9}$$

where κ is the number of parameters in the model, N is the number of data points, and ℓ is the log-likelihood function for the given model. The optimal balance of complexity and accuracy results in the lowest BIC value. Figure 1c depicts an example BIC calculation for the ground truth model. The likelihood function comes from Bayes' Theorem as the sampling probability ($P(\boldsymbol{D} \mid M)$ also written as $\mathcal{L}(M)$), where **D** is the dataset for the given model M, or probability that the data resulting from physics is accurately described by a given model. In the case that the $\mathcal{L}(M)$ is

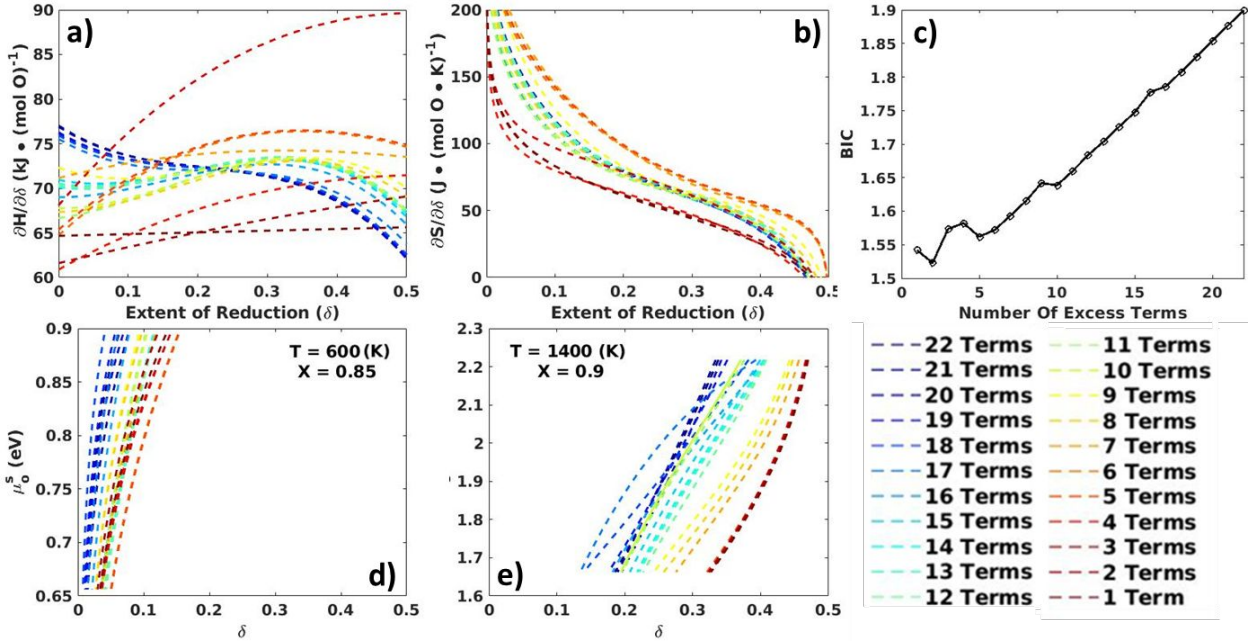

Figure 2: a) The enthalpy of reduction $\left(\frac{\partial H}{\partial \delta}\right)$ of BSF for all 22 models generated during the CrossFit CEF algorithm, b) The entropy of reduction $\left(\frac{\partial S}{\partial \delta}\right)$ of all 22 models, c) BIC results for the ground truth BSF models, d-e) Data points with low (d) and high (e) δ range predictions versus chemical potential where T and X are held constant for each point and pO_2 varies. The legend indicates the number of excess ($\mathcal{L}$) terms for each model.

utilized in BIC, all model likelihoods are calculated on the same dataset with which all models were optimized.

We construct our comparisons assuming that error arises from both data noise and model error, and therefore, the forms of the likelihood functions are based on two normal distributions, as outlined by Gregory et al.[60] The convolution of the two distributions gives the probability that data point Y_i resulted from the given model $m(x_i \mid \boldsymbol{\theta})$ where $\boldsymbol{\theta}$ are the optimized model parameters. When all Y_i are independent, the likelihood $P(\boldsymbol{D} \mid M) = P(Y_1,Y_2,Y_3,...,Y_n \mid M,\theta) = \prod_{i=1}^{N} P(Y_i|M,\theta)$. Therefore, the construction of $\mathcal{L}(M)$ for a model is:

$$\mathcal{L}(M) = (2\pi)^{-N/2}\left(\prod_{i=1}^{N}\left(\sigma_i^2 + \sigma_m^2\right)^{-1/2}\right)exp\left(\sum_{i=1}^{N} -\frac{\left(y_i - m(x_i \mid \theta)\right)^2}{2\left(\sigma_i^2 + \sigma_m^2\right)}\right) \quad \text{eqn 10}$$

Where σ_m is the standard distribution of model error, and σ_i is the standard distribution of the i^{th} data point noise.

We use the natural logarithm of the $\mathcal{L}(M)$ in constructing the log-likelihood function $\ell(M)$ (eqn 11) which makes the calculations more numerically stable by transforming products into sums, which have better behaved derivatives. We note that maximizing the likelihood $\ell(M)$ is equal to maximizing the log-likelihood $\mathcal{L}(M)$. The final form of the log likelihood used is:

$$\ell(M) = log(\mathcal{L}(M)) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{N} log\left(\sigma_i^2 + \sigma_m^2\right) - \frac{1}{2}\left(\sum_{i=1}^{N}\frac{\left(y_i - m(x_i \mid \theta)\right)^2}{\left(\sigma_i^2 + \sigma_m^2\right)}\right) \quad \text{eqn 11}$$

2.3 Using Bayesian Approaches for Model and Data Selection – Comparing Log-Likelihoods

2.3.1 Selection of new TpOX points for data collection

The multiple models with which we conduct our statistical analysis arise from the G^{excess} term in eqn 3. G^{excess} describes the interaction of the various species on one sublattice given the composition of the other sublattices. Most of the excess terms have negligible contributions, or there is insufficient data to fully define them. Therefore, we use our recently developed cross-fitting algorithm which adopts a top-down approach to select and optimize the parameters associated with G^{excess}. This process generates multiple optimized models, which vary in the number of excess terms included, q, for a given TpOX dataset. The differing number of excess terms results in different predictions of thermodynamic states and expected δ (off stoichiometric) values. We exploit the variance in the predicted thermodynamics and δ to select the appropriate model by statistical methods. Likewise, the uncertainty from the multiple models enables a Bayesian approach to select the TpOX point which, if sampled next, would add the most information and thus generate the most robust/accurate next model.

We base our data point selection methods on Gaussian Process Regression (GPR)[61, 62]. Through GPR, the variation between models collapses as more data is observed. In GPR one determines points for data collection by considering which points have the greatest variance across all models. In our method, each of our CEF models (22 of them) predicts a δ value across a range of unsampled TpOX points. TpOX points with a narrow range in predicted δ indicates CEF model agreement and insensitivity to the parameters. Thus, sampling at these points is not expected to improve the model fit. Conversely, TpOX points with a wide range in predicted δ denotes disagreement between models, and thus an ability to delineate effects between excess terms. We

elect to use the range (or spread) of predicted δ values to determine model consistency because the range is more sensitive to differences in outlying models than other statistical quantities. An example of two data points with low and high range predictions where T and X are held constant for each point and pO_2 varies is seen in Figure 2d-e.

In order to create a more robust modeling algorithm we select a point that has the highest predicted range in δ across the models to sample next. When the new data point is selected and the model parameters re-optimized with the additional data point, the various models lower their prediction range at that point as all models have now been optimized to include the new observed data. As a result, the models may differ in points which were previously well described by older models. This is a desired trait, allowing the algorithm to "hunt" for inflections and curvature in thermodynamic trends due to new TpOX sample points. The hunting nature prevents the algorithm from agreeing with itself too early in the data collection process and preventing inadequate sampling of the data.

2.3.2 Stopping Criteria Utilizing the Likelihood Ratio Test

The stopping criteria of our iterative data selection process is based on a comparison of the models from one iteration to the next via a likelihood function through the so-called Bayes Factor [63] and a cutoff threshold for the number of sequential models which are statistically the same. Thus, building high confidence in our model.

Instead of using an obscure kernel as a prior, which is non-physical, we use the uncertainty in the existing optimized CEF models which vary in the number of excess terms. The optimized CEF models are all equally likely to represent the data they were optimized for, resulting in all prior probabilities being equal. Because the model parameters are optimized in each model and

there are no unknown parameters, the Bayes Factor simplifies to the ratio of likelihoods in what is known as a likelihood-ratio test[50, 63-66] (eqn 12).

$$\mathcal{B}_{LR} = -2\left(\ell(M_{i-1}) - \ell(M_i)\right) \quad \text{eqn 12}$$

The likelihood ratio test is used to compare one model to another in the Bayesian iterative process. In our iterative approach, we add data then optimize thermodynamic parameters to that data through the cross-fitting CEF approach, creating several models of varying parameter amounts. We then utilize BIC to determine the appropriate model for the current dataset in Bayesian iterative process (e.g during the i^{th} iteration $\ell(M_i)$ utilizes dataset $\boldsymbol{D^{1:i}}$ to determine the best model for this iteration). Lastly, we assess whether we have reached our optimal model with the least amount of data required via the likelihood-ratio test ($\mathcal{B}_{LR}$).[67, 68] When two models would fit the data with the same trends their likelihoods would be almost identical, i.e $\mathcal{B}_{LR} \approx 0$.

Because the likelihood ratio test requires consistent data points between the candidate models, we evaluate both the current model (model *i*) and the previous model (model *i*-1) with all available data (i.e. $D^{1:i}$). Thus, the $i^{th} - 1$ model is assessed on a data point it has not been optimized on. When the *i-1* model has the same likelihood as the i^{th} model even at a TpOX it was not optimized for, the confidence in the models grows. We follow Kass and Raftery's[63] work and take a value of $|\mathcal{B}_{LR}| \leq 2$ to indicate that there is very little evidence of having dissimilar models. We note that $\mathcal{B}_{LR}$ can be negative or positive simply based on which model better describes the dataset. To ensure confidence in our model, we require several consecutive iterations which achieve $|\mathcal{B}_{LR}| \leq 2$., i.e. we impose a cutoff value. Too low of a cutoff value may stop the Bayesian iterative process prematurely, particularly early in the cycle, while too large of a cutoff would cause unnecessary data collection.

2.4 Data Sources

This work uses both computational and experimental BSF off-stoichiometry data from our previous work[42] and Bush et al,[47] respectively. Because these data sets have been previously described in detail, we provide only a brief summary of their collection method and characteristics to provide context.

DFT calculations were conducted to obtain the energy of $Ba_{1-x}Sr_xFeO_{3-\delta}$ compositions using the Vienna Ab Initio Simulation Package (VASP)[69, 70] with the meta-GGA strongly constrained and appropriately normed (SCAN)[71] exchange/correlation functional and a Hubbard correction[72, 73] (SCAN+U) U value of 3.1 eV for the Fe d-orbitals, [74]. Various δ values were achieved via supercell size variations and meticulous comparison of all vacancy interaction pairs to find the lowest ground state configuration for each δ value considered. In this study, we note we use all available density functional theory (DFT) data previously gathered and leave the application of the Bayesian data selection method to DFT data as future work. However, our method could easily be extended to drive computational data collection as well.

The experimental data consists of 2200 thermogravimetric (TGA) data points spanning four temperatures (T=400°C, 538°C, 747°C, 1100°C), five oxygen partial pressures (pO_2=0.9, 0.29, 0.095, 0.03, and 0.01 bar) and five Ba mole fractions (x=0, 0.05, 0.10, 0.15, 0.20). Measurements were taken at each TpOX, providing 100 TpOX combinations and

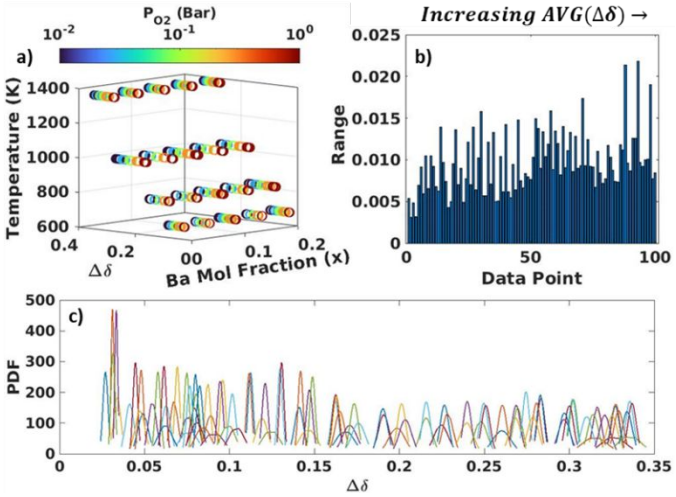

Figure 3: a) BSF TGA data provided by Bush et al, b) Range of Δδ values for each TpOX combination of TGA data points sorted by increasing average Δδ, c) Normal PDF plots for each TpOX data point illustrating increasing uncertainty with increasing Δδ values.

were corrected by 10 background (i.e., "blank run") measurements. Additionally, the average of the 10 subtractions for each point was included yielding 11 total subtractions and 1100 datapoints. The same 100 material measurements and 10 background measurements were repeated, resulting in a total dataset of 2200 TpOX points. TGA measured the change in δ (Δδ) from a reference measurement (δ_0) at 300°C and $pO_2 = 0.9$ bar. The δ_0 value has not been determined experimentally for this dataset. Figure 3a shows a plot of Δδ values at each TpOX. We note that the variance and range of the measured Δδ is correlated with its mean value, with higher values of Δδ having more uncertainty. Figure 3b shows the range of each 100 measured data points sorted by increasing average Δδ value. Figure 3c shows the probability density of each 100 measured data points based on their respective mean and standard deviation created by the 20 total background subtraction from each measured data point.

3 Results and Discussion

In this section we first report on the performance of the Bayesian approach to selecting a robust model with the least amount of data required. We then compare our approach to randomly selected data from the pool of available experimental data points to determine the robustness of our method. We define the ground truth model as the model that was optimized on the full set of available experimental TpOX data.

3.1 Iterative Bayesian Algorithm Applied to Example BSF System.

Through the Bayesian data selection process, we assess the predictive confidence of the current model based on the range of the predicted δ value at an unsampled set of TpOX across the optimized CEF using different numbers of excess terms. We apply this method to the BSF system with an available pool of 100 unique TpOX data points. The Bayesian method iterates through the available data points assessing at each iteration which point to add next that would build

confidence in the model. We show an example of the range in predicted δ during data point selection in Figure 4a-b with a model fit on 44 TpOX points. Red "hot zones" in Figure 4a indicate TpOX points with the highest range in predicted δ values while blue "cold regions" indicate where the models mostly agree. For ease of visualization, in Figure 4b, we only display the data points with the largest (top 20%) range in δ, thus highlighting the sampling regions of interest. Once the algorithm has produced a statistically equivalent model multiple times in a row, based on the cutoff value; the algorithm stops, and a high confidence thermodynamic model is reached.

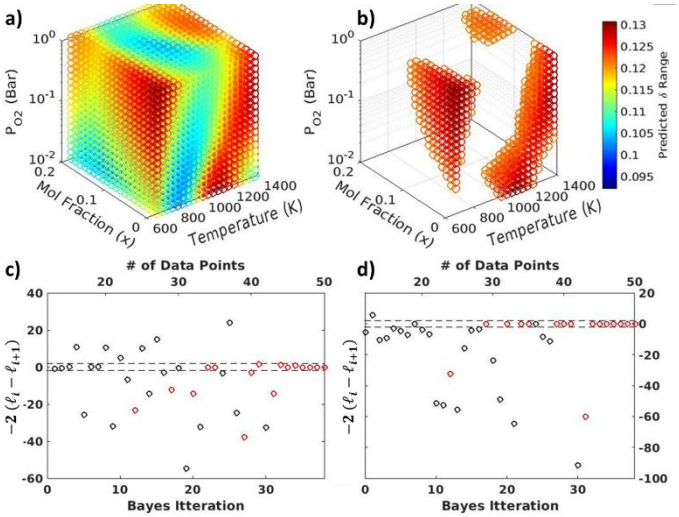

Figure 4:a) Sampling TpOX mesh grid illustrating calculated range in δ values across each model. Higher range means higher uncertainty across all models. b) Top 20% of predicted δ range, c) Likelihood ratio test for Bayesian iterative data selection process (dashed lines at $|\mathcal{B}_{LR}| \leq 2$indicate statistically similar models), d) Likelihood ratio test for model comparison to ground truth model. Red circles indicate iterations that resulted in models that match the ground truth model in number of excess terms and selected excess terms.

For the remainder of this work, we only select TpOX point for which we have existing experimental data. However, in practice the full range of points would be considered for selection. Additionally, multiple points from the different hot regions could be selected for simultaneous experimental evaluation to facilitate data collection. In the selection of multiple points Bayesian Inference can be used to determine the best choice in multiple points; however, such an implementation is left to future work.

We initialized the Bayesian iterative process with 12 experimental TpOX points and all DFT data points. We used every combination of three temperature points (T = 673, 811 & 1373 K) (the required minimum number of temperature points to fit the free energy model given the

constant heat capacity assumption), two oxygen partial pressures (pO_2 = 0.9 & 0.01 bar) and two mol fractions (x=0 & 0.2). Based on initial testing, we recommend that the initial T and pO_2 point selection include the minimum and maximum conditions experimentally obtainable with the third T point halfway between the extremes in order to initially sample a wide range of the chemical potential phase space. The thermodynamic model based on these 12 initial data points is named iteration 0.

The likelihood-ratio test between subsequent iterations determines when the model is likely to be accurate and thus sufficient data has been collected. This test is used after each iteration as the stopping criteria because under experimental scenarios the true thermodynamics are unknown *a priori.* Figure 4c shows the likelihood ratios from one iteration to the next out to 50 Bayesian iterations. Using a convergence criterion of five sequential iterations with $|\mathcal{B}_{LR}| \leq 2$, terminates the Bayesian process after iteration 36 (48 total data points). As demonstrated by the red circles in Figure 4c, the models converge to the same excess terms as the ground truth model. The points selected are displayed in Figure 5. We note that these points do not form an obvious pattern, nor are they seemingly systematic.

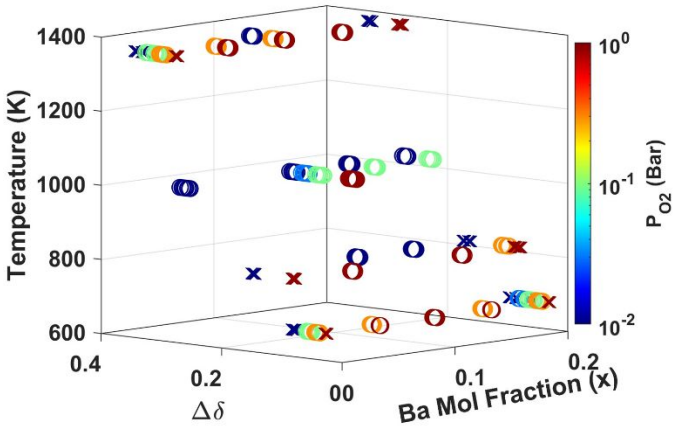

Figure 5: Data selected up to Iteration 32 (44 data points). Circles indicates point selected in Bayesian iterations. X points indicates initial starting points selected for iteration 0.

Figure 4c also illustrates that overly loose convergence criteria can produce poor results. For example, a convergence criterion of three sequential iterations with $|\mathcal{B}_{LR}| \leq 2$, would stop our algorithm after iteration 3, 15 data points total, but iteration 3's thermodynamics do not match the ground truth. We hypothesize that early convergence of the Bayesian iterations is most likely early

in sampling iterations where the data is most limited. Furthermore, we conducted a likelihood ratio test of the first 50 models to the ground truth model illustrated in Figure 4d. We find that the same models that met the stopping criteria are not statistically different than the ground truth model.

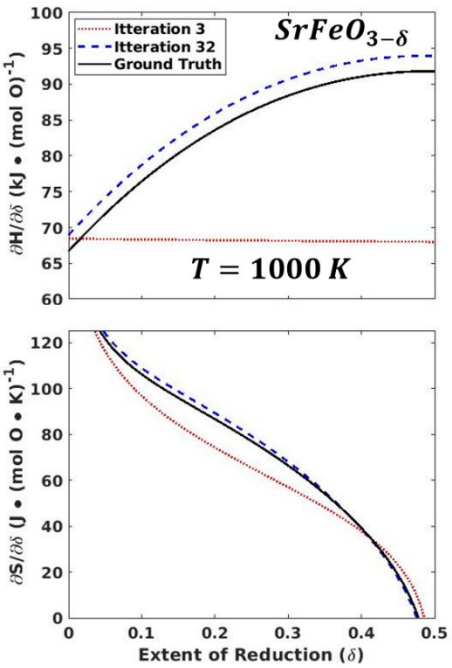

Figure 6: Comparison of enthalpy of reduction $\left(\frac{\partial H}{\partial \delta}\right)$ (top) and entropy of reduction $\left(\frac{\partial S}{\partial \delta}\right)$ (bottom) for Bayesian iteration 3, 32 and ground truth at T = 1000K.

We compare the accuracy of the predicted thermodynamics for $SrFeO_{3-\delta}$ at iteration 3, 32 and our ground truth in Figure 6. We select iteration 32 as the point of comparison because it is the first of five in the cutoff iterations and, via our likelihood ratio test, the next four models after iteration 32 are statistically equivalent. One can see in Figure 6 that at iteration 3 the thermodynamic results are vastly different from the ground truth results. Conversely, the thermodynamic model constructed by iteration 32 closely matches the ground truth model with only a 2 kJ/mol error in reduction enthalpy $\left(\frac{\partial H}{\partial \delta}\right)$ and < 3 J/mol K difference in reduction entropy $\left(\frac{\partial S}{\partial \delta}\right)$ normalized to mol O. Furthermore, the model from iteration 32 matches the ground truth, both in number of excess terms selected by BIC, q= 2, and the selected excess terms, L17 and L18. The excess term identity and number selected in iteration 32 – 36 match, and that the optimized parameters are the same order of magnitude and sign. A table of all optimized parameters from iteration 32 – 36 and the ground truth can be found in the SI, Table SI-2. We note, the complex linear combination of terms allows for some tradeoff between parameters B and C in eqn 4 so that they may differ between models and still result in the same thermodynamic trends. We leave the assessment of state function parameter setup to prevent this

strict parameter trade off to future work. In this work, careful selection of parameter initialization values is conducted to prevent inaccurate trade off as described in our previous work.[42]

The cost effectiveness of running this algorithm directly arises from the algorithm speed, experimental data collection time, frequency and ability to update the data collection program. In an unoptimized state, the MATLAB implementation of this method is < 2 hours; however, we note that the code can be further optimized, parallelized, and migrated to a faster language, significantly accelerating its execution in future versions. The experimental approach requires the synthesis of materials to investigate compositional changes and the collection non-stoichiometry at T and pO_2 points. Amongst these, the synthesis of new materials represents a significant bottleneck, as such, the ability of our method to pre-select important compositions significantly decreases the number of expensive synthesis steps. The benefits of selecting T and pO_2 points will depend on the kinetics of the materials, where slower reducing/oxidizing materials will benefit more than those with near instantaneous equilibrium. In general, lower temperature and pO_2 points are likely to benefit the most as kinetics are slower. Because the users of this method choose the candidate points the model select as the most important, they can weigh their relative cost of T, pO_2, or X data collection and adjust the candidate points accordingly. Considering these facts, the <2 hour run time for the Bayesian scheme heavily outweighs the time needed to synthesize and establish fine mesh grids of T and pO_2 for the traditional systematic approach to data collection.

3.2 Bayesian Data Selection Outperforms Random Chance

We test the Bayesian selected data approach to one in which data is randomly selected as an example where the experimentalist is unsure of the critical points that control the thermodynamics of the material. We created 20 datasets each starting with the same 12 initial data points as our Bayesian iterative process and then selected 32 additional data points at random from

the remaining 88 available. We note that there are almost 1 septillion (10^{24}) possible combinations of data points to choose from. Therefore, the probability that an experimentalist would choose the Bayseian selected data points is effectively 0.

Utilizing the likelihood-ratio test we compare the 20 models optimized to their respective randomly selected data points to the Bayesian selected dataset, illustrated in Figure 7. In order to make an appropriate comparison the models must be compared on the same dataset therefore we calculate the error of all 20 models on the dataset the Bayesian approach selected, $\boldsymbol{D^{1:44}}$. We note the dataset used has no bearing on the comparison, only that the comparison is done on the same dataset. We find that only one of the twenty models is statistically the same as the Bayesian selected data model, i.e $|\mathcal{B}_{LR}| < 2$. Thus, the Bayesian approach outperformed 95% of the randomly selected data.

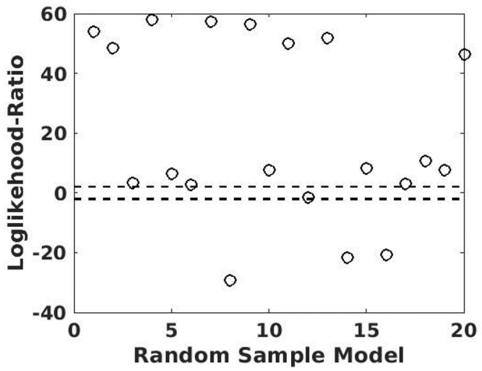

Figure 7: Likelihood-ratio test comparing 20 models generated from randomly selected data compared to the Bayesian method selected data at iteration 32 (44 data points).

4 Conclusions

In this work we successfully implemented a Bayesian inference approach for model and data selection for reduction thermodynamic analysis. We utilize the Bayesian Information Criterion to identify the proper dimensionality (i.e. how many excess terms to include, q) of the CEF model. Through an iterative process, we add data until a cutoff metric is met via the likelihood-ratio test. The data to be added at each iteration is selected through a Bayesian approach where the TpOX data point with the highest predicted range in δ among CEF models with various excess terms is selected. The aim of the Bayesian approach to data selection is to lower the spread of predicted values over the entire sample space. We show that through the employment of our Bayesian informed data selection scheme less than half the experimental data set (44 out of 100 points) is required to reach

the same thermodynamic results. Importantly, the data selected by the Bayesian method is not overly represented in the low pO_2 and high temperature regions. Rather, in its BSF application, only ~22% of the selected points were in the lowest pO_2 of the available dataset. Therefore, this methodology effectively decreases the number of total data points required as well as the number of difficult data points to sample. We note, however, it this method does not guarantee that no low pO_2 points are needed if that is where critical information is located, but it does restrict the number of datapoints at low pO_2 to those which are necessary in building a robust model. Furthermore, we show that the Bayesian informed approach does better than pure random sampling of data as only 5% of the models optimized on 44 randomly selected data points identified the same thermodynamic trends as the model optimized on the full 100 data points. Overall, implementation of our method will save significant time in data collection, allowing for more materials to be investigated and lower research costs. Further, it opens the possibility of a hands-off high-throughput process for M_xO_y material selection and design and while we have not directly interfaced this code with TGA or other data acquisition software, there is no reason that it could not be easily implemented.

5 Associated Content

Supporting Information.

Contains list of Redlich-Kister expansion of $\mathcal{L}$ terms. A table of all optimized parameters from iterations 32 – 36 and the ground truth. MATLAB code for all algorithms and data used can be found at https://github.com/MuhichLab/Bays_Data_Selection.git.

6 Funding Sources

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship under Award Number DE-SC0022158. This work is also supported by the U.S. Department of Energy's Energy Efficiency & Renewable Energy office under Award Number DE-EEDE-EE0010732.

7 Acknowledgement

We gratefully acknowledge the group of Prof. Peter Loutzenhiser for sharing their TGA data on BSF, Nhu Nguyen and Tyler Farr for conducting the experiments and processing the data and sharing it in spreadsheet form. All fitting calculations were conducted in MATLAB© 2021a. The authors acknowledge Research Computing at Arizona State University for providing HPC resources that have contributed to the research results reported within this paper.

8 Corresponding Author Information

Corresponding Authors: Christopher Muhich

Email: cmuhich@asu.edu

Telephone: 480-965-2673

Address: 551 E. Tyler Mall, ERC 257, Arizona State University, Tempe AZ

REFERNCE LIST

1. Guo, Q.; Geng, J.; Pan, J.; Zou, L.; Tian, Y.; Chi, B.; Pu, J., Brief review of hydrocarbon-reforming catalysts map for hydrogen production. *Energy Reviews* **2023,** *2* (3), 100037.
2. Ahmad, Y. H.; Mohamed, A. T.; Kumar, A.; Al-Qaradawi, S. Y., Solution combustion synthesis of Ni/La2O3 for dry reforming of methane: tuning the basicity via alkali and alkaline earth metal oxide promoters. *RSC Advances* **2021,** *11* (53), 33734-33743.
3. Bulfin, B.; Lapp, J.; Richter, S.; Gubàn, D.; Vieten, J.; Brendelberger, S.; Roeb, M.; Sattler, C., Air separation and selective oxygen pumping via temperature and pressure swing oxygen adsorption using a redox cycle of SrFeO3 perovskite. *Chemical Engineering Science* **2019,** *203*, 68-75.
4. Krzystowczyk, E.; Haribal, V.; Dou, J.; Li, F., Chemical Looping Air Separation Using a Perovskite-Based Oxygen Sorbent: System Design and Process Analysis. *ACS Sustainable Chemistry & Engineering* **2021,** *9* (36), 12185-12195.
5. Cai, R.; Dou, J.; Krzystowczyk, E.; Richard, A.; Li, F., Chemical looping air separation with Sr0.8Ca0.2Fe0.9Co0.1O3-δ perovskite sorbent: Packed bed modeling, verification, and optimization. *Chemical Engineering Journal* **2022,** *429*, 132370.
6. Gu, X.-K.; Samira, S.; Nikolla, E., Oxygen Sponges for Electrocatalysis: Oxygen Reduction/Evolution on Nonstoichiometric, Mixed Metal Oxides. *Chemistry of Materials* **2018,** *30* (9), 2860-2872.
7. Xu, M.; Ermanoski, I.; Stechel, E. B.; Deng, S., Oxygen pumping characteristics of YBaCo4O7+δ for solar thermochemical cycles. *Chemical Engineering Journal* **2020,** *389*, 124026.
8. Bulfin, B.; Vieten, J.; Starr, D.; Azarpira, A.; Zachäus, C.; Hävecker, M.; Skorupska, K.; Schmücker, M.; Roeb, M.; Sattler, C., Redox chemistry of CaMnO 3 and Ca 0.8 Sr 0.2 MnO 3 oxygen storage perovskites. *Journal of Materials Chemistry A* **2017,** *5* (17), 7912-7919.
9. Vieten, J.; Bulfin, B.; Call, F.; Lange, M.; Schmücker, M.; Francke, A.; Roeb, M.; Sattler, C., Perovskite oxides for application in thermochemical air separation and oxygen storage. *Journal of Materials Chemistry A* **2016,** *4* (35), 13652-13659.
10. Mane, R.; Kim, H.; Han, K.; Kim, H.; Lee, S. S.; Roh, H.-S.; Lee, C.; Jeon, Y., Important factors of the A-site deficient Mn perovskites design affecting the CO oxidation activity. *Catalysis Today* **2023**, 114347.
11. Wexler, R. B.; Sai Gautam, G.; Bell, R. T.; Shulda, S.; Strange, N. A.; Trindell, J. A.; Sugar, J. D.; Nygren, E.; Sainio, S.; McDaniel, A. H.; Ginley, D.; Carter, E. A.; Stechel, E. B., Multiple and nonlocal cation redox in Ca–Ce–Ti–Mn oxide perovskites for solar thermochemical applications. *Energy & Environmental Science* **2023,** *16* (6), 2550-2560.
12. Hashimoto, J.; Bayon, A.; Tamburro, O.; Muhich, C. L., Thermodynamic and Structural Effects of Fe Doping in Magnesium Manganese Oxides for Thermochemical Energy Storage. *Energy & Fuels* **2023,** *37* (6), 4692-4700.
13. Bayon, A.; Hashimoto, J.; Muhich, C., Chapter Two - Fundamentals of solar thermochemical gas splitting materials. In *Advances in Chemical Engineering*, Lipiński, W., Ed. Academic Press: 2021; Vol. 58, pp 55-90.

14. Sai Gautam, G.; Stechel, E. B.; Carter, E. A., Exploring Ca–Ce–M–O (M = 3d Transition Metal) Oxide Perovskites for Solar Thermochemical Applications. *Chemistry of Materials* **2020,** *32* (23), 9964-9982.
15. Park, J.; Xu, B.; Pan, J.; Zhang, D.; Lany, S.; Liu, X.; Luo, J.; Qi, Y., Accurate prediction of oxygen vacancy concentration with disordered A-site cations in high-entropy perovskite oxides. *npj Computational Materials* **2023,** *9* (1), 29.
16. van de Krol, R.; Liang, Y.; Schoonman, J., Solar hydrogen production with nanostructured metal oxides. *Journal of Materials Chemistry* **2008,** *18* (20), 2311-2320.
17. Mastronardo, E.; Qian, X.; Coronado, J. M.; Haile, S. M., The favourable thermodynamic properties of Fe-doped CaMnO3 for thermochemical heat storage. *Journal of Materials Chemistry A* **2020,** *8* (17), 8503-8517.
18. Liu, D.; Dou, Y.; Xia, T.; Li, Q.; Sun, L.; Huo, L.; Zhao, H., B-site La, Ce, and Pr-doped Ba0.5Sr0.5Co0.7Fe0.3O3-δ perovskite cathodes for intermediate-temperature solid oxide fuel cells: Effectively promoted oxygen reduction activity and operating stability. *Journal of Power Sources* **2021,** *494*, 229778.
19. Fuks, D.; Mastrikov, Y.; Kotomin, E.; Maier, J., Ab initio thermodynamic study of (Ba, Sr)(Co, Fe) O 3 perovskite solid solutions for fuel cell applications. *Journal of Materials Chemistry A* **2013,** *1* (45), 14320-14328.
20. Barry, T. I.; Dinsdale, A. T.; Gisby, J. A.; Hallstedt, B.; Hillert, M.; Jansson, B.; Jonsson, S.; Sundman, B.; Taylor, J. R., The compound energy model for ionic solutions with applications to solid oxides. *Journal of Phase Equilibria* **1992,** *13* (5), 459-475.
21. Kolodiazhnyi, T.; Sakurai, H.; Belik, A.; Gornostaeva, O., Unusual lattice evolution and magnetochemistry of Nb doped CeO2. *Acta Materialia* **2016,** *113*, 116-123.
22. Rousseau, R.; Glezakou, V.-A.; Selloni, A., Theoretical insights into the surface physics and chemistry of redox-active oxides. *Nature Reviews Materials* **2020,** *5* (6), 460-475.
23. Teh, L. P.; Setiabudi, H. D.; Timmiati, S. N.; Aziz, M. A. A.; Annuar, N. H. R.; Ruslan, N. N., Recent progress in ceria-based catalysts for the dry reforming of methane: A review. *Chemical Engineering Science* **2021,** *242*, 116606.
24. Young, S. D.; Chen, J.; Sun, W.; Goldsmith, B. R.; Pilania, G., Thermodynamic Stability and Anion Ordering of Perovskite Oxynitrides. *Chemistry of Materials* **2023,** *35* (15), 5975-5987.
25. Arifin, D.; Ambrosini, A.; Wilson, S. A.; Mandal, B.; Muhich, C. L.; Weimer, A. W., Investigation of Zr, Gd/Zr, and Pr/Zr – doped ceria for the redox splitting of water. *International Journal of Hydrogen Energy* **2020,** *45* (1), 160-174.
26. Bulfin, B.; Vieten, J.; Agrafiotis, C.; Roeb, M.; Sattler, C., Applications and limitations of two step metal oxide thermochemical redox cycles; a review. *Journal of Materials Chemistry A* **2017,** *5* (36), 18951-18966.
27. Hoes, M.; Muhich, C. L.; Jacot, R.; Patzke, G. R.; Steinfeld, A., Thermodynamics of paired charge-compensating doped ceria with superior redox performance for solar thermochemical splitting of H2O and CO2. *Journal of Materials Chemistry A* **2017,** *5* (36), 19476-19484.
28. Takacs, M.; Hoes, M.; Caduff, M.; Cooper, T.; Scheffe, J. R.; Steinfeld, A., Oxygen nonstoichiometry, defect equilibria, and thermodynamic characterization of LaMnO3 perovskites with Ca/Sr A-site and Al B-site doping. *Acta Materialia* **2016,** *103*, 700-710.
29. Panlener, R. J.; Blumenthal, R. N.; Garnier, J. E., A THERMODYNAMIC STUDY OF NONSTOICHIOMETRIC CERIUM DIOXIDE. *Journal of Physics and Chemistry of Solids* **1975,** *36*, 1213-1222.

30. Bayon, A.; Hashimoto, J.; Muhich, C., Fundamentals of solar thermochemical gas splitting materials. In *Advances in Chemical Engineering*, Elsevier: 2021; Vol. 58, pp 55-90.
31. Krug, R. R.; Hunter, W. G.; Grieger, R. A., Enthalpy-entropy compensation. 1. Some fundamental statistical problems associated with the analysis of van't Hoff and Arrhenius data. *The Journal of Physical Chemistry* **1976,** *80* (21), 2335-2341.
32. MacKay, D. J. C., Information-Based Objective Functions for Active Data Selection. *Neural Computation* **1992,** *4* (4), 590-604.
33. Seo, S.; Wallat, M.; Graepel, T.; Obermayer, K. In *Gaussian Process Regression: Active Data Selection and Test Point Rejection*, Mustererkennung 2000, Berlin, Heidelberg, 2000//; Sommer, G.; Krüger, N.; Perwass, C., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2000; pp 27-34.
34. Zhang, B.-T.; Cho, D.-Y. In *Genetic Programming with Active Data Selection*, Simulated Evolution and Learning, Berlin, Heidelberg, 1999//; McKay, B.; Yao, X.; Newton, C. S.; Kim, J.-H.; Furuhashi, T., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 1999; pp 146-153.
35. Tsamardinos, I.; Charonyktakis, P.; Papoutsoglou, G.; Borboudakis, G.; Lakiotaki, K.; Zenklusen, J. C.; Juhl, H.; Chatzaki, E.; Lagani, V., Just Add Data: automated predictive modeling for knowledge discovery and feature selection. *npj Precision Oncology* **2022,** *6* (1), 38.
36. Khan, N. M.; Abraham, N.; Hon, M., Transfer Learning With Intelligent Training Data Selection for Prediction of Alzheimer's Disease. *IEEE Access* **2019,** *7*, 72726-72735.
37. Tomida, N.; Tanaka, T.; Ono, S.; Yamagishi, M.; Higashi, H., Active Data Selection for Motor Imagery EEG Classification. *IEEE Transactions on Biomedical Engineering* **2015,** *62* (2), 458-467.
38. Dunbar, O. R. A.; Howland, M. F.; Schneider, T.; Stuart, A. M., Ensemble-Based Experimental Design for Targeting Data Acquisition to Inform Climate Models. *Journal of Advances in Modeling Earth Systems* **2022,** *14* (9), e2022MS002997.
39. Eng, A. Y. S.; Soni, C. B.; Lum, Y.; Khoo, E.; Yao, Z.; Vineeth, S. K.; Kumar, V.; Lu, J.; Johnson, C. S.; Wolverton, C.; Seh, Z. W., Theory-guided experimental design in battery materials research. *Science Advances* **2022,** *8* (19), eabm2422.
40. Lin, Y.; Li, M.; Watanabe, Y.; Kimura, T.; Matsunawa, T.; Nojima, S.; Pan, D. Z., Data Efficient Lithography Modeling With Transfer Learning and Active Data Selection. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* **2019,** *38* (10), 1900-1913.
41. Lookman, T.; Balachandran, P. V.; Xue, D.; Yuan, R., Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Computational Materials* **2019,** *5* (1), 21.
42. Wilson, S. A.; Stechel, E. B.; Muhich, C. L., Overcoming significant challenges in extracting off-stoichiometric thermodynamics using the compound energy formalism through complementary use of experimental and first principles data: A case study of Ba1-xSrxFeO3-δ. *Solid State Ionics* **2023,** *390*, 116115.
43. Todorović, M.; Gutmann, M. U.; Corander, J.; Rinke, P., Bayesian inference of atomistic structure in functional materials. *Npj computational materials* **2019,** *5* (1), 35.
44. Lei, B.; Kirk, T. Q.; Bhattacharya, A.; Pati, D.; Qian, X.; Arroyave, R.; Mallick, B. K., Bayesian optimization with adaptive surrogate models for automated experimental design. *npj Computational Materials* **2021,** *7* (1), 194.
45. Vigliotti, A.; Csányi, G.; Deshpande, V. S., Bayesian inference of the spatial distributions of material properties. *Journal of the Mechanics and Physics of Solids* **2018,** *118*, 74-97.

46. Fang, L.; Guo, X.; Todorović, M.; Rinke, P.; Chen, X., Exploring the Conformers of an Organic Molecule on a Metal Cluster with Bayesian Optimization. *Journal of Chemical Information and Modeling* **2023,** *63* (3), 745-752.
47. Bush, H. E.; Nguyen, N. P.; Farr, T.; Loutzenhiser, P. G.; Ambrosini, A., Air separation via a two-step solar thermochemical cycle based on (Ba,La)xSr1-xFeO3-δ: Thermodynamic analysis. *Solid State Ionics* **2021,** *368*, 115692.
48. Neath, A. A.; Cavanaugh, J. E., The Bayesian information criterion: background, derivation, and applications. *WIREs Computational Statistics* **2012,** *4* (2), 199-203.
49. Schwarz, G., Estimating the Dimension of a Model. *The Annals of Statistics* **1978,** *6* (2), 461-464, 4.
50. O'Hagan, A., *Kendall's Advanced Theory of Statistic 2B Chapter 7*. John Wiley & Sons: 2010; Vol. 2B.
51. Hillert, M.; Staffansson, L., Regular-solution model for stoichiometric phases and ionic melts. *Acta chem. scand.* **1970,** *24* (10), 3618-3626.
52. Hillert, M., The compound energy formalism. *Journal of Alloys and Compounds* **2001,** *320* (2), 161-176.
53. Sundman, B.; Chen, Q.; Du, Y., A Review of Calphad Modeling of Ordered Phases. *Journal of Phase Equilibria and Diffusion* **2018,** *39* (5), 678-693.
54. Redlich, O.; Kister, A. T., Algebraic Representation of Thermodynamic Properties and the Classification of Solutions. *Industrial & Engineering Chemistry* **1948,** *40* (2), 345-348.
55. Hillert, M., Some properties of the compound energy model. *Calphad* **1996,** *20* (3), 333-341.
56. Spencer, P., A brief history of CALPHAD. *Calphad* **2008,** *32* (1), 1-8.
57. Ji, Y.; Abernathy, H. W.; Chen, L.-Q., Thermodynamic models of multicomponent nonstoichiometric solution phases using internal process order parameters. *Acta Materialia* **2022,** *223*, 117462.
58. Cacciamani, G., An introduction to the calphad method and the compound energy formalism (CEF). *Tecnologia em Metalurgia, Materiais e Mineração* **2016,** *13* (1), 16-24.
59. Madin, O. C.; Boothroyd, S.; Messerly, R. A.; Fass, J.; Chodera, J. D.; Shirts, M. R., Bayesian-Inference-Driven Model Parametrization and Model Selection for 2CLJQ Fluid Models. *Journal of Chemical Information and Modeling* **2022,** *62* (4), 874-889.
60. Gregory, P., *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support*. Cambridge University Press: Cambridge, 2005.
61. Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G., Gaussian Process Regression for Materials and Molecules. *Chemical Reviews* **2021,** *121* (16), 10073-10141.
62. Schulz, E.; Speekenbrink, M.; Krause, A., A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology* **2018,** *85*, 1-16.
63. Kass, R. E.; Raftery, A. E., Bayes factors. *Journal of the american statistical association* **1995,** *90* (430), 773-795.
64. Montoya, J. A.; Díaz-Francés, E.; Sprott, D. A., On a criticism of the profile likelihood function. *Statistical Papers* **2009,** *50* (1), 195-202.
65. Fan, J.; Zhang, C.; Zhang, J., Generalized likelihood ratio statistics and Wilks phenomenon. *The Annals of statistics* **2001,** *29* (1), 153-193.

66. Wilks, S. S., The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The annals of mathematical statistics* **1938,** *9* (1), 60-62.
67. Woolf, B., The log likelihood ratio test (the G ‐ test). *Annals of human genetics* **1957,** *21* (4), 397-409.
68. Kent, J. T., Robust properties of likelihood ratio tests. *Biometrika* **1982,** *69* (1), 19-27.
69. Kresse, G.; Furthmüller, J., Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science* **1996,** *6* (1), 15-50.
70. Kresse, G.; Furthmüller, J., Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B* **1996,** *54* (16), 11169.
71. Sun, J.; Ruzsinszky, A.; Perdew, J. P., Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Physical Review Letters* **2015,** *115* (3), 036402.
72. Hubbard, J., Electron Correlations in Narrow Energy Bands. *Proceedings of the Royal Society of London. Series A, Mathematical and physical sciences* **1963,** *276* (1365), 238-257.
73. Dudarev, S. L.; Botton, G. A.; Savrasov, S. Y.; Humphreys, C. J.; Sutton, A. P., Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study. *Physical Review B* **1998,** *57* (3), 1505-1509.
74. Sai Gautam, G.; Carter, E. A., Evaluating transition metal oxides within DFT-SCAN and $\text{SCAN}+U$ frameworks for solar thermochemical applications. *Physical Review Materials* **2018,** *2* (9), 095401.