Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplemental information for "Efficient Pourbaix diagrams of many-element compounds"

Anjli M. Patel,[a] Jens K. Nørskov, [b] Kristin A. Persson,[c] Joseph H. Montoya[d]

[a] *Department of Chemical Engineering, Stanford University, Stanford, CA, USA*
[b] *Department of Physics, Technical University of Denmark, Lyngby, Denmark*
[c] *Lawrence Berkeley National Laboratory and Department of Materials Science, University of California, Berkeley, CA*
[d] *Toyota Research Institute, Los Altos, CA.; E-mail: joseph.montoya@tri.global*

1 Pourbaix diagram thermodynamics

In this section, we provide additional information on the thermodynamics of the Pourbaix diagram methodology described in the main text. To generate a Pourbaix diagram in our formalism, we begin with "entries" as termed in pymatgen, which consist of a composition and an attached energy. In the case of the Pourbaix diagram, the energy we attach to the composition is the Gibbs free energy of formation for the given compound. From this energy, one can form the free energy at a given potential, $\Psi(\mathrm{E},\mathrm{pH})$, via the *Pourbaix potential* formalism, as derived in prior reports[1,2]. This may be expressed by eq. (1),

$$\Psi(E,pH) = \frac{1}{\sum N_M}\left(G_f^0 + 0.0591\log(c) - N_O\mu_{\mathrm{H_2O}} - 0.0591\,(2N_O - N_H)\,pH - (2N_O - N_H + Q)\,E\right) \quad (1)$$

where E is the electrochemical potential, pH is the pH, N_O is the number of oxygen atoms in the chemical formula, N_H is the number of hydrogen atoms in the chemical formula, Q is the charge associated with the composition, and c is the concentration of the species for aqueous ions. The free energy is normalized by N_M, which represents the number of atoms in the compound excluding oxygen and hydrogen. This is essentially the formation energy of the species at a variable chemical potential of protons and electrons, as expressed in the scaling of E and pH. The Pourbaix potential is a plane in E-pH space for each species under consideration.

For each species to be considered in the Pourbaix diagram, one must find regions of minimal Pourbaix potential for all species subject to the E-pH scaling. Since they are linear, these regions may be efficiently found by determining the halfspace intersection of the planes represented by eq. (1).

For multi-element entries, one must determine the *minimal combination* of entries, rather than simply the minimal entries. These combinations are subject to a composition constraint[3], that is the Pourbaix

diagram is computed at a fixed ratio of non-oxygen or hydrogen elements. The Pourbaix potential for these combinations of species, termed **MultiEntries**, is expressed as an ideal mixture of the species and their associated Pourbaix potentials, as shown in eq. (2),

$$\Psi_{multi}(E, pH, w_i) = \sum_i w_i \Psi_i(E, pH) \tag{2}$$

where w_i are the normalized weights of each of the single-entry Pourbaix potentials (Ψ_i) that fulfill the composition constraint. The regions of the Pourbaix diagram for multiple elements are similarly found by minimizing this potential over all possible valid stoichiometric combinations of species.

To achieve minimization of the Pourbaix potential for a multi-element diagram, appropriate stoichiometric mixtures of single species (which may have one or more elements) must be enumerated. Stoichiometric coefficients which minimize the formation energy of mixtures are generally determined in the canonical phase diagram, i.e. by finding the convex hull in $E_0^f - x_1 - x_2 - ... - x_n$ space, where E_0^f is the formation energy per atom and x_n are the fractional compositions of constitutive elements. For example, the convex hull and its simplices in $E_0^f - x_O$ space form the species and the tie lines between them in the phase diagram of the Fe-O chemical system sketched in fig. S1. In the case of higher-dimensional diagrams, the simplices will be triangles (3-D), tetrahedra (4-D), etc.

In addition, to find all of the possible energy-minimizing combinations of single entries subject to variable chemical potential of the constitutive reactants (i.e. Fe and O in fig. S1), one may determine the hull without those reference points. Phases may appear or be absent from the hull when one fixes the chemical potentials, but the analysis without the endpoints will reveal all possibilities for the hull. In addition, one may include arbitrary reactants for the formation of these phases, for example, H^+ and e^-, in order to allow for the variable chemical potential of arbitrary species. H^+ and e^- are constitutive reactants for all species in the Pourbaix diagram, and the diagrams shown in the main text's 3-D hull figure correspond to these.

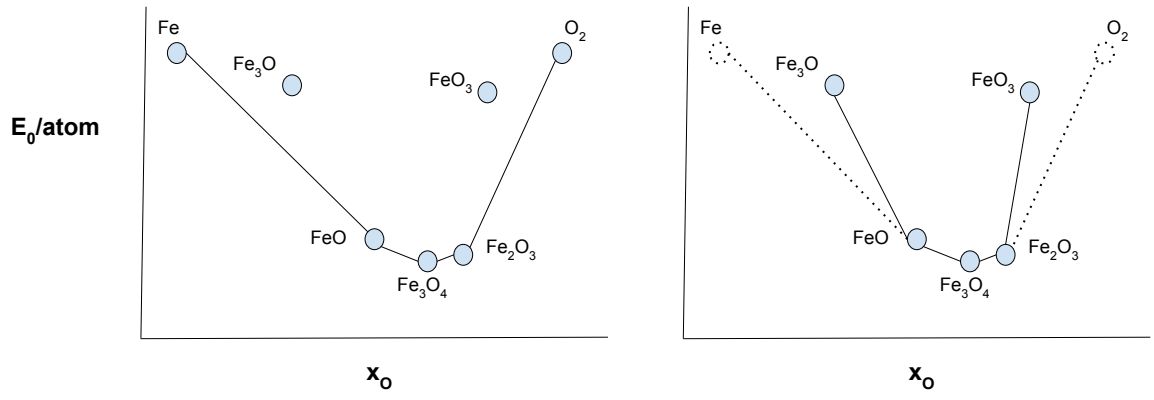

Figure S1 Schematic demonstrating how the convex hull reveals stable combinations in composition space, and how removal of convex hull endpoints allows for determination of "inner" convex hull with all possible phases and phase mixtures enumerated for variable chemical potentials of endpoint species.

When valid entry combinations are determined using the simplices of the convex hull, the halfspace

intersection of the resultant MultiEntries may be constructed in the same manner as that of a single element PourbaixDiagram, since the scaling of the overall formation energy of the mixture is still linear.

2 Algorithmic benchmarking

For additional clarity, we provide a flow diagram of the old and new Pourbaix diagram generation algorithms in fig. S2. Benchmarks for the new algorithm computed on a 2016 MacBook Pro with an 2 GHz Intel i5 processor are shown in table S1. The contents of a script used for these benchmarking are provided below and serve as an example of how to use the Pourbaix diagram generation method in the code:

```
from pymatgen import MPRester
import time
from pymatgen.analysis.pourbaix_diagram import PourbaixDiagram, \
    ELEMENTS_HO, PourbaixPlotter

# Set up rester
mpr = MPRester()  # Enter materials project API key if not configured
entry = mpr.get_entries("mp-1215061")[0]
composition = entry.composition
comp_dict = {str(key): value for key, value in composition.items()
             if key not in ELEMENTS_HO}

# Get data
fetch_start = time.time()
data = mpr.get_pourbaix_entries(list(comp_dict.keys()))
fetch_elapsed = time.time() - fetch_start
print("Fetch takes {} seconds".format(fetch_elapsed))
entry = [entry for entry in data if entry.entry_id == "mp-1215061"][0]

# Construct pourbaix diagram
construct_start = time.time()
pbx = PourbaixDiagram(data, comp_dict=comp_dict, filter_solids=False)
construct_elapsed = time.time() - construct_start
plt = PourbaixPlotter(pbx).plot_entry_stability(
    entry, label_domains=False)
plt.savefig("pourbaix.png")
print("Construct takes {} seconds".format(construct_elapsed))
# Plot
```

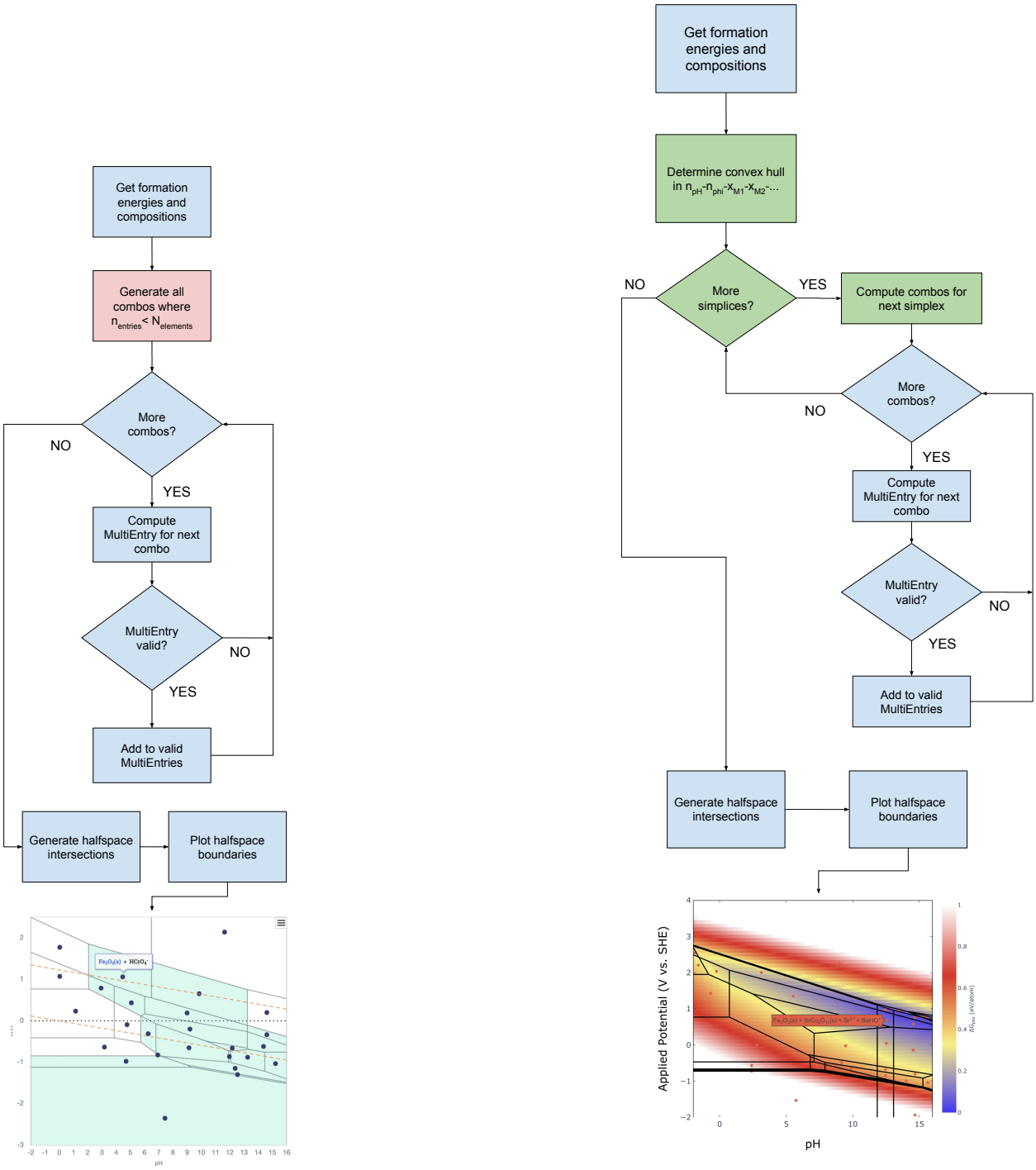

Figure S2 Flowchart representing old (left) and new (right) Pourbaix diagram construction algorithms - In the left flow diagram, the red box representing multi-entry combination candidate generation is the primary bottleneck, and the green sections replacing this in the right diagram are significantly faster.

		New algorithm		Old algorithm	
Formula or chemsys	**mp-id**	**Iterations**	**Elapsed time (s)**	**Iterations**	**Elapsed time (s)**
$BiVO_4$	mp-25122	2.2×10^2	1.0	5.0×10^4	3.4×10^1
$SrIrO_3$	mp-17097	1.0×10^2	0.8	5.0×10^3	4.0
Cr_2FeO_4	mp-24900	2.8×10^2	1.0	2.9×10^4	2.1×10^1
MoS_2	mp-1434	1.9×10^2	0.9	1.9×10^4	1.4×10^1
$FeCN_2$	mp-567933	6.2×10^2	1.5	4.6×10^7	2.9×10^4
$ZnNiAs_2O_7$	mp-1042795	1.1×10^3	1.4	1.1×10^6	8.4×10^2
$TiAl_2Zn$	mp-1217109	8.0×10^2	1.4	6.0×10^6	5.5×10^3
$Ni(C_2N_3)_2$	mp-22405	5.9×10^2	1.4	1.6×10^7	1.4×10^4
$BaSr_7Fe_7CoO_{24}$	mp-1075935	3.5×10^3	4.3	4.1×10^9	3.1×10^6 *
Al-Mn-Fe-Cu-Mg	–	2.4×10^4	2.4×10^1	7.0×10^{12}	5.3×10^9 *
$Ba_2NaTi_2MnRe_2Si_8HO_{26}F$	mp-1215061	1.4×10^5	4.2×10^2	1.2×10^{19}	1.4×10^{14} *

Table S1 Performance Metrics for Pourbaix Algorithms - Starred (*) values indicate estimated evaluation time based on number of iterations and average iteration rate from 2 and 3-element runs.

References

[1] K. A. Persson, B. Waldwick, P. Lazic and G. Ceder, *Physical Review B*, 2012, **85**, 235438.

[2] W. Sun, D. A. Kitchaev, D. Kramer and G. Ceder, *Nature Communications*, 2019, **10**, 573.

[3] W. T. Thompson, M. H. Kaye, C. W. Bale and A. D. Pelton, *Uhlig's Corrosion Handbook: Third Edition*, 2011, pp. 103–109.