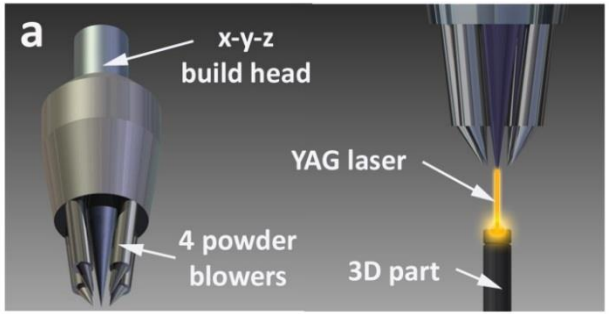
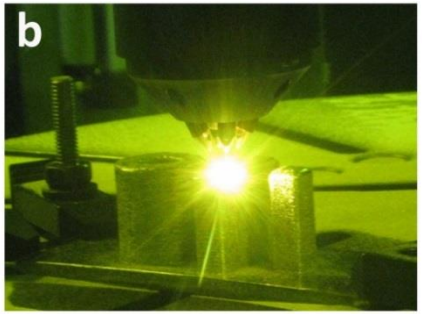
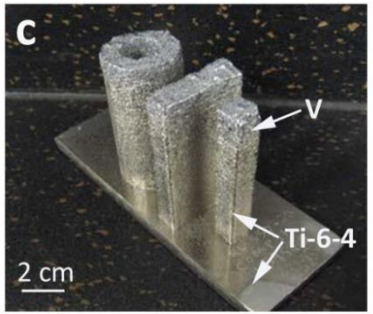
Toward a New Generation of Thermodynamic Models for Alloy Additive Manufacturing

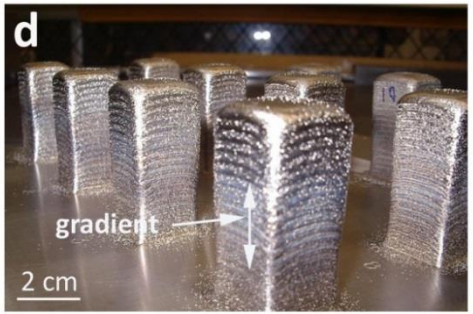
Richard Otis*, Lourdes Bobbio, Allison Beese, Zi-Kui Liu
Pennsylvania State University
*Now at NASA/JPL
March 1, 2017

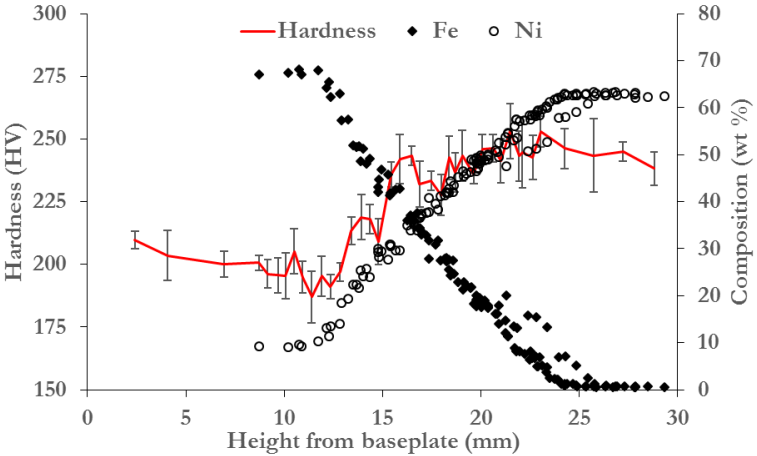
Additive Manufacturing: Graded Alloys

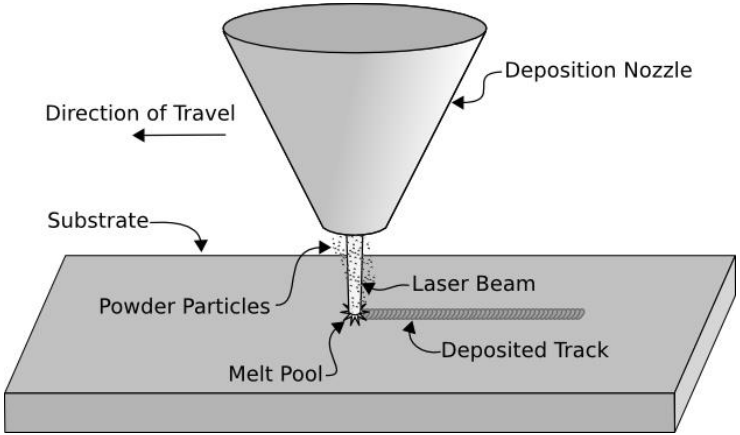






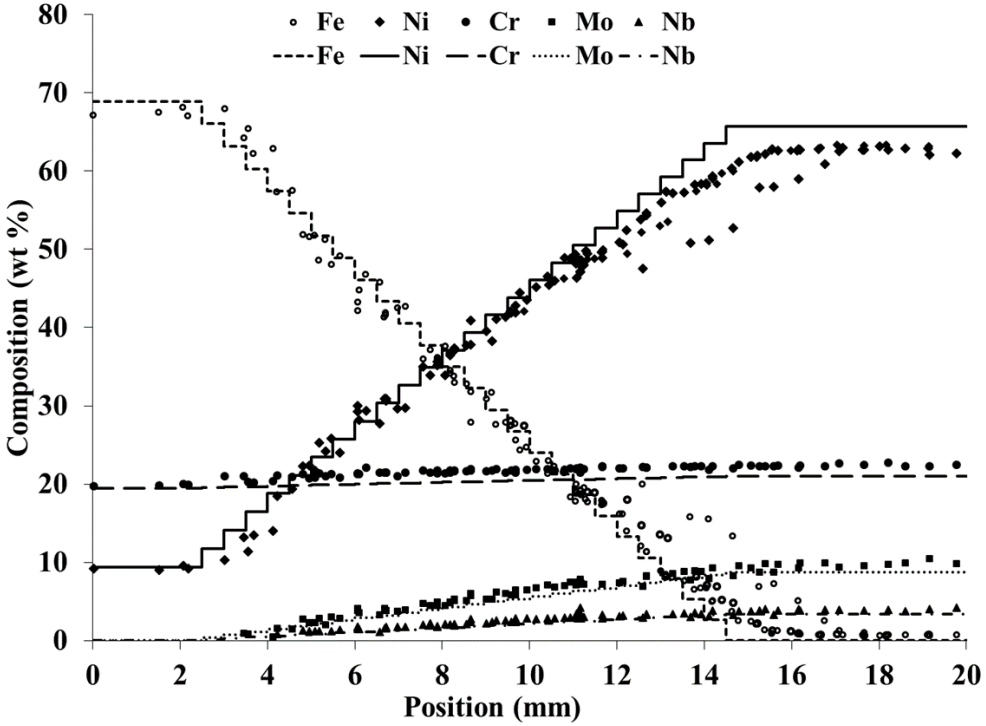


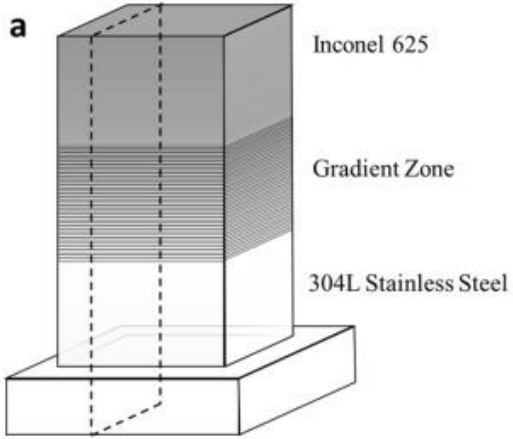




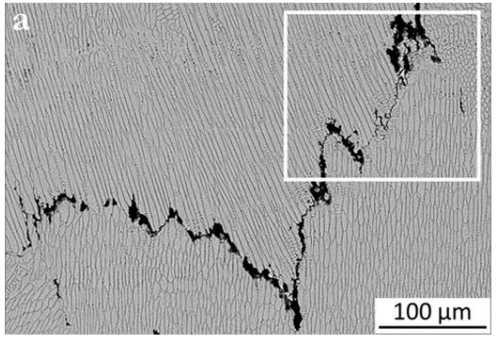
(*bottom right*) Z. Fan and F. Liou, *Numerical Modeling of the Additive Manufacturing (AM) Processes of Titanium Alloy*, 2012.

304L Stainless to Inconel 625

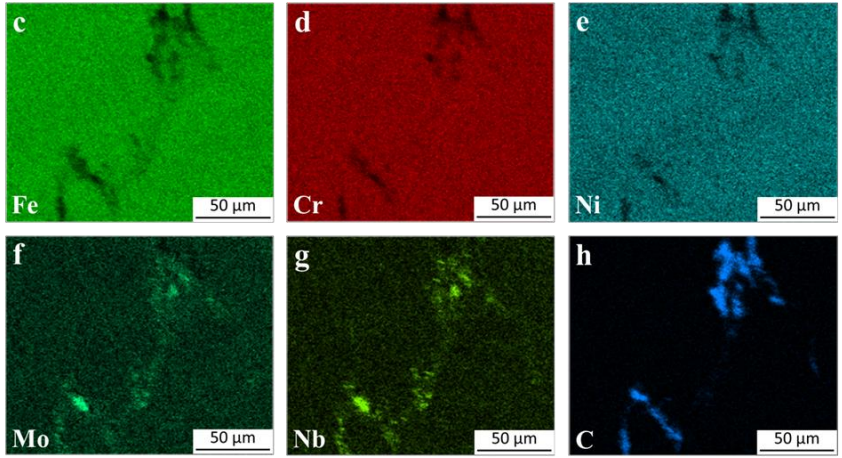




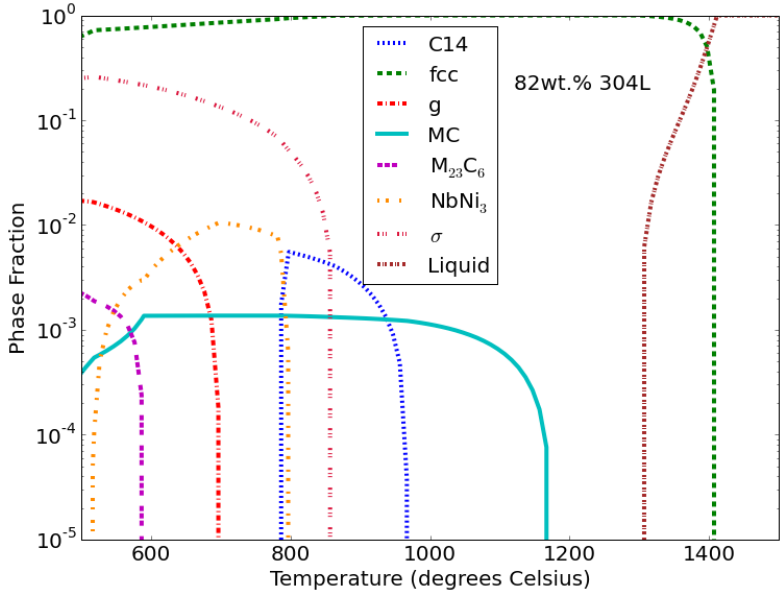
B. E. Carroll, R. A. Otis, J.-P. Borgonia, J. Suh, R. P. Dillon, A. A. Shapiro, D. C. Hofmann, Z.-K. Liu, A. M. Beese, *Acta Mater.*, Vol. 108, 2016, pp. 46-54.

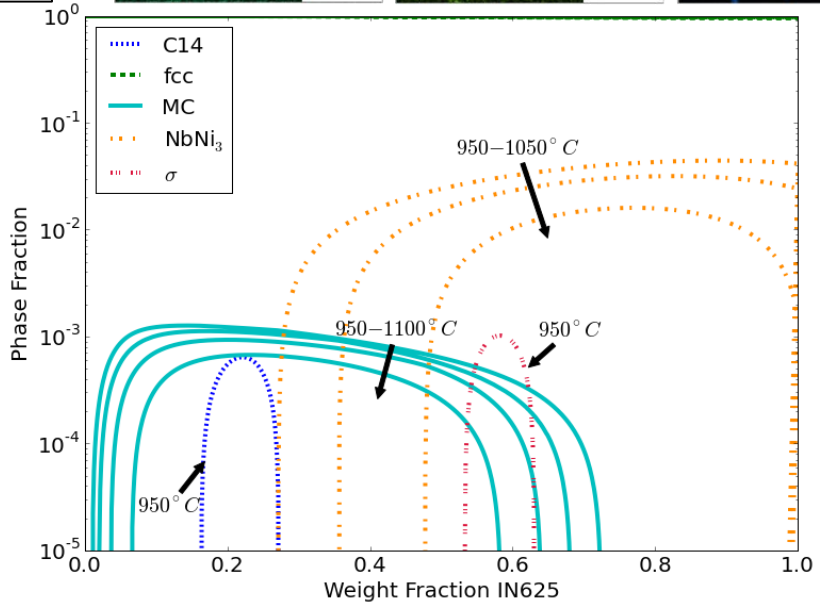


b	Design	Actual
Fe	54.6	56
Cr	19.8	20
Ni	21.1	18
Mo	1.8	2
Mn	1.1	1
Nb	0.7	1

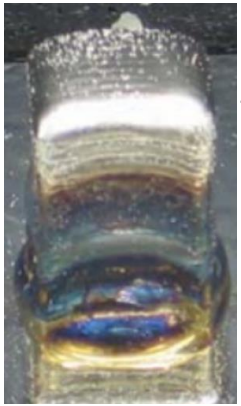


Crack related to **refractory carbide** formation





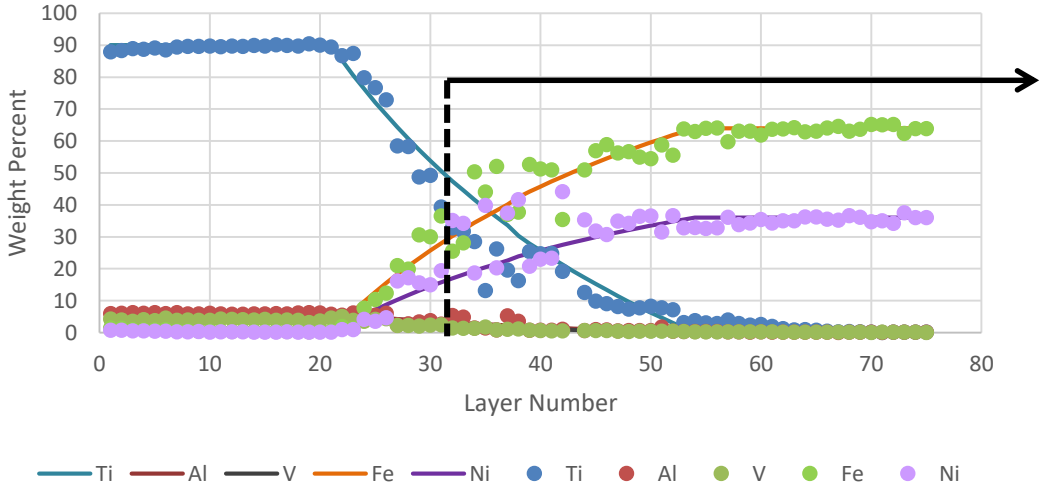
Ti-6Al-4V to Invar 36

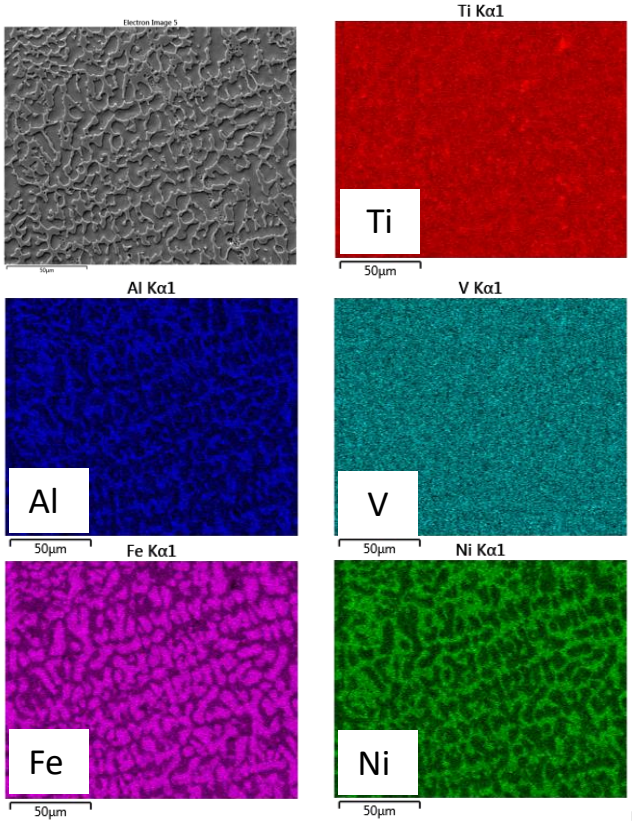


Fe-36Ni

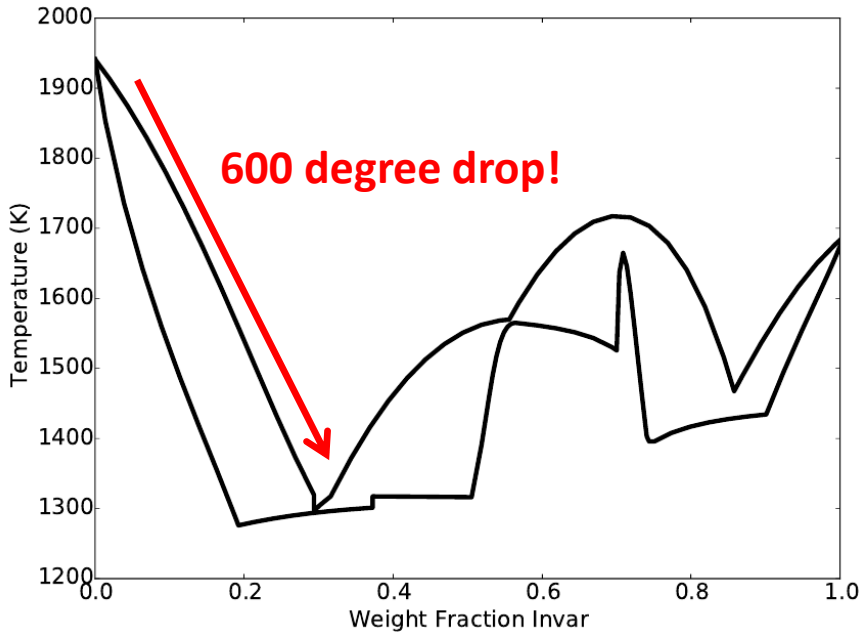
Q: What is the cause of distortion and cracking in the gradient?

Ti-6Al-4V



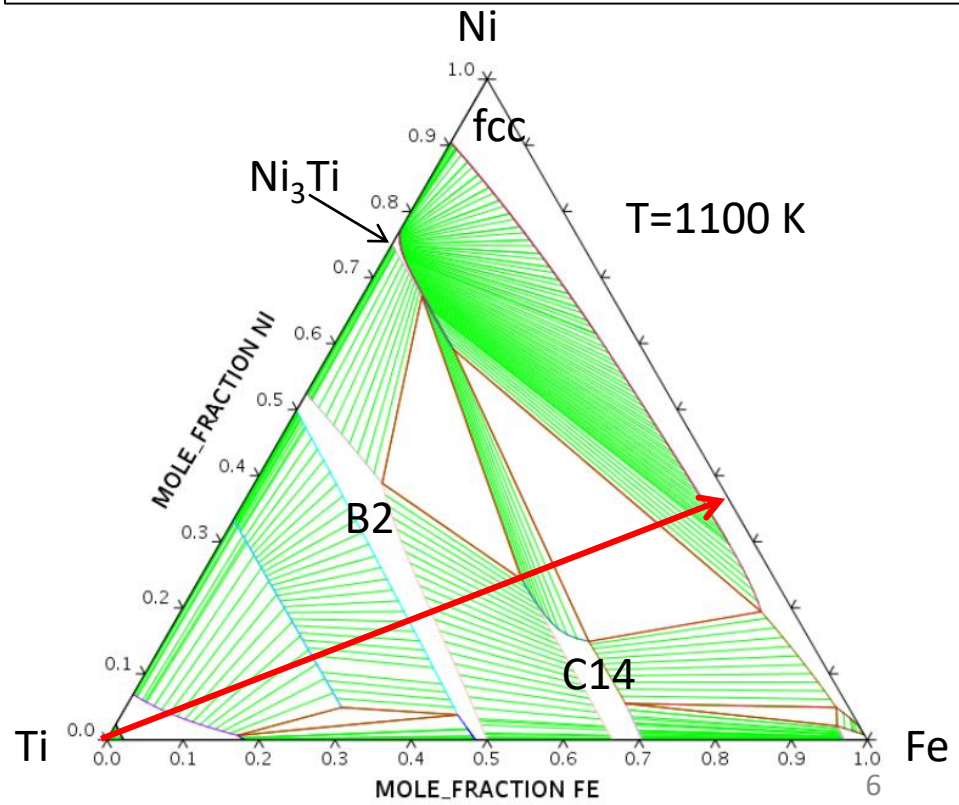


Liquidus and Solidus: Pure Ti to Fe-36Ni

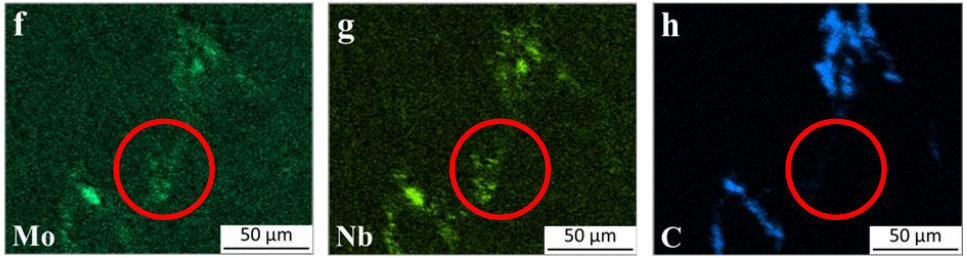


Distortion likely caused by formation of **liquid in the gradient**

Where's Al and V? More on that in a minute…



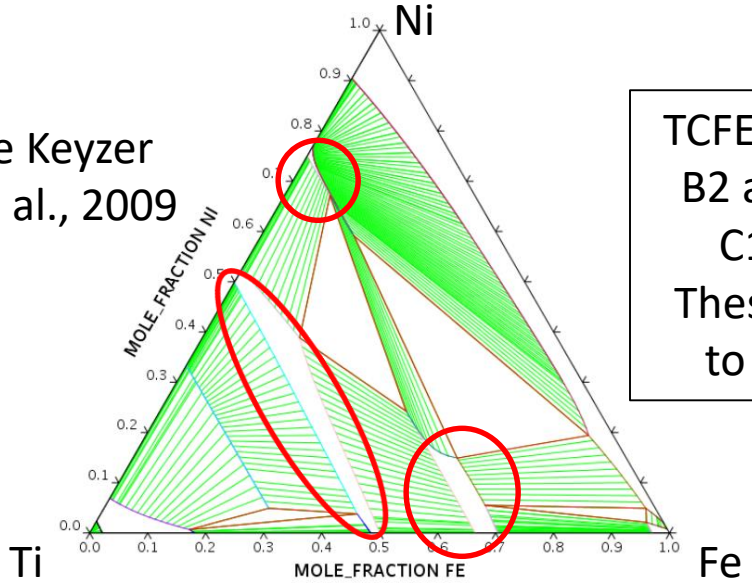
Thermodynamic Analysis: Limitations



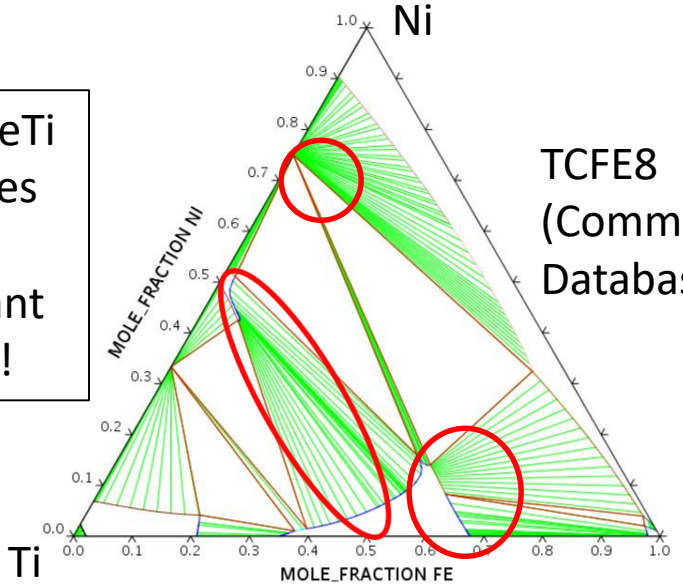
B2 ordering may* exist in Mo-Nb; not on current phase diagram

*C. Jiang et al., *Phys. Rev. B.*, 2004.

De Keyzer et al., 2009



TCFE8 neglects FeTi B2 and overstates C14 solubility. These are relevant to our analysis!



TCFE8 (Commercial Database)

Conclusion So Far

We can **rationalize success/failure of pre-existing samples** like these with existing databases, if we are diligent.

For **real design** purposes, we need to improve the databases.

CALPHAD: CALculation of PHase Diagrams

Phase-based modeling of material properties – equilibrium and non-equilibrium

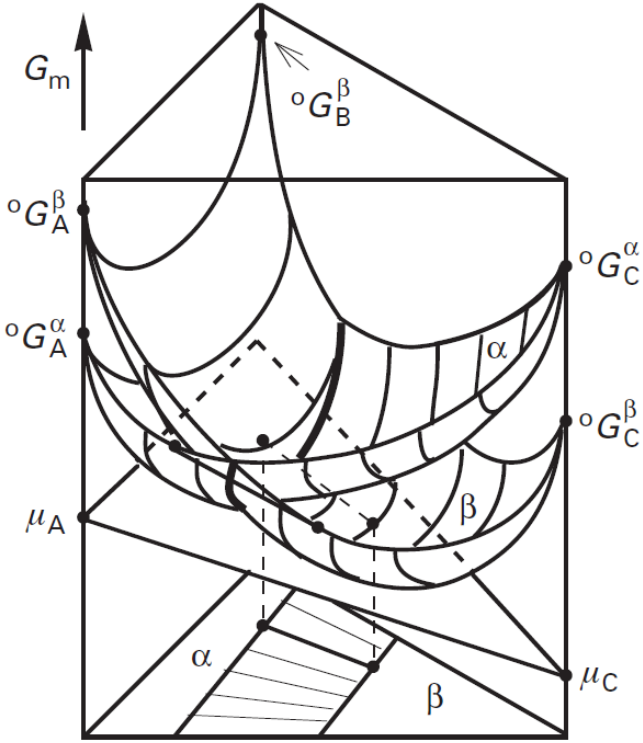
$$\min G(T, P, f_i, y_j^s)$$

$$\sum f_i = 1$$ Phase Fraction Balance

$$\sum y_j^s = 1$$ Sublattice Site Fraction Balance

$$\sum_i f_i \sum_s a^s y_j^s = x_j$$ Mass Balance

L. Kaufman and H Bernstein. *Computer Calculation of Phase Diagrams.* (Academic Press Inc., 1970).



M. Hillert, 2nd ed., 2008, p. 149

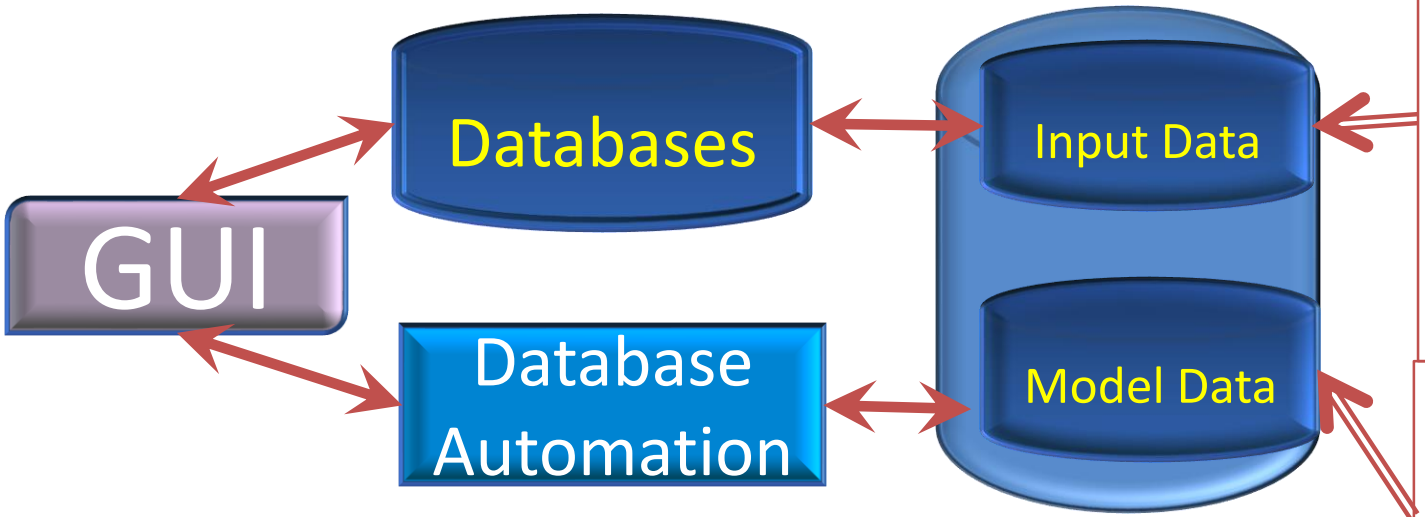
CALPHAD: Example energy function

$$G_m^{\alpha}(T, x_i) = \sum_i x_i(\Delta G_i^{\beta,298.15\rightarrow T} + \Delta G_i^{\beta\rightarrow\alpha} + RT\ln x_i) +$$

$$\sum_{i,j} x_i x_j \sum_m {}^m L_{i,j}(x_i - x_j)^m + \sum_{i,j,k} x_i x_j x_k (x_i I_i^{i,j,k} + x_j I_j^{i,j,k} + x_k I_k^{i,j,k})$$

"L" and "I" parameters are fit to experimental and first-principles data. They are the "secret sauce" of a database and play a significant role in prediction accuracy.

ESPEI: Extensible, Self-optimizing Phase Equilibrium Infrastructure



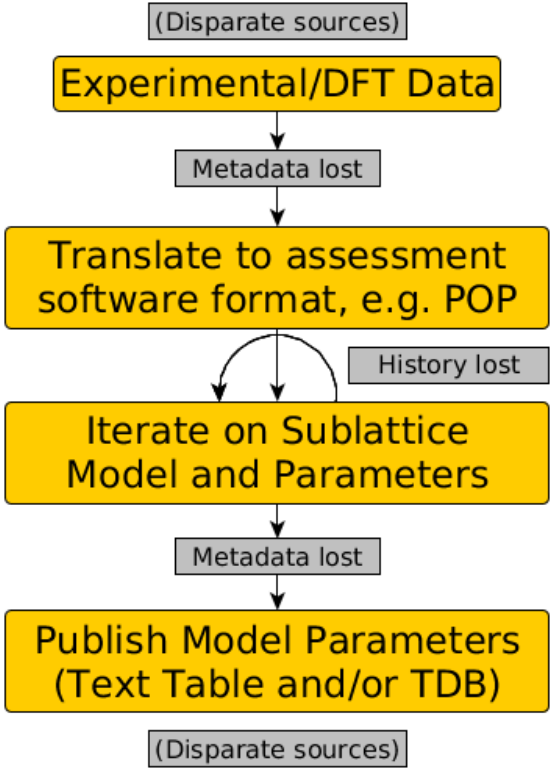
- Crystallographic: atomic positions
- Thermochemical: heat capacity, enthalpy, free energy
- Phase equilibrium: phase boundary

- Gibbs energy models linked to input data used in evaluation of model parameters

Shang, Wang, Liu, Magnesium Technology 2010, p. 617

CALPHAD Data Challenges

- Data fragmentation
- Data standards
- Workflow tools
- Uncertainty quantification
- **Reproducibility**



pycalphad

IP[y]: Notebook BinaryExamples Last Checkpoint: Jan 08 23:25 (autosaved)

File Edit View Insert Cell Kernel Help

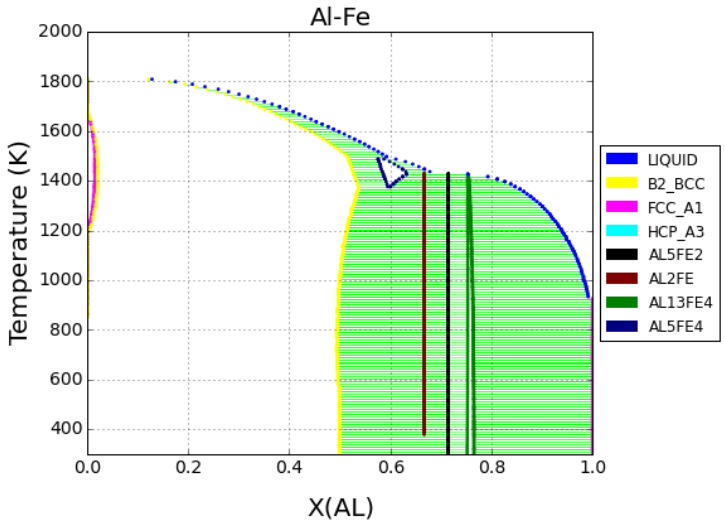
Heading 1 Cell Toolbar: None

Al-Fe (M.Seiersten et al., 1991)

In [3]:
```
db_alfe = Database('alfe_sei.TDB')
my_phases_alfe = ['LIQUID', 'B2_BCC', 'FCC_A1', 'HCP_A3', 'AL5FE2', 'AL2FE', 'AL13FE4', 'AL5FE4']

fig = plt.figure(figsize=(9,6))
pdens = [{'B2_BCC': 20000}, 2000]
%time ax = binplot(db_alfe, ['AL', 'FE', 'VA'] , my_phases_alfe, 'X(AL)', 300, 2000, pdens=pdens, ax=fig.gca())
plt.show()
```

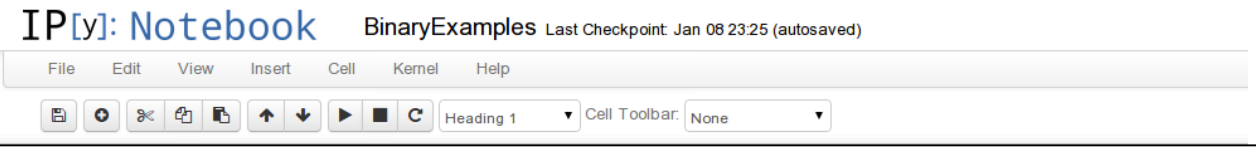
```
CPU times: user 27.5 s, sys: 639 ms, total: 28.1 s
Wall time: 17.1 s
```

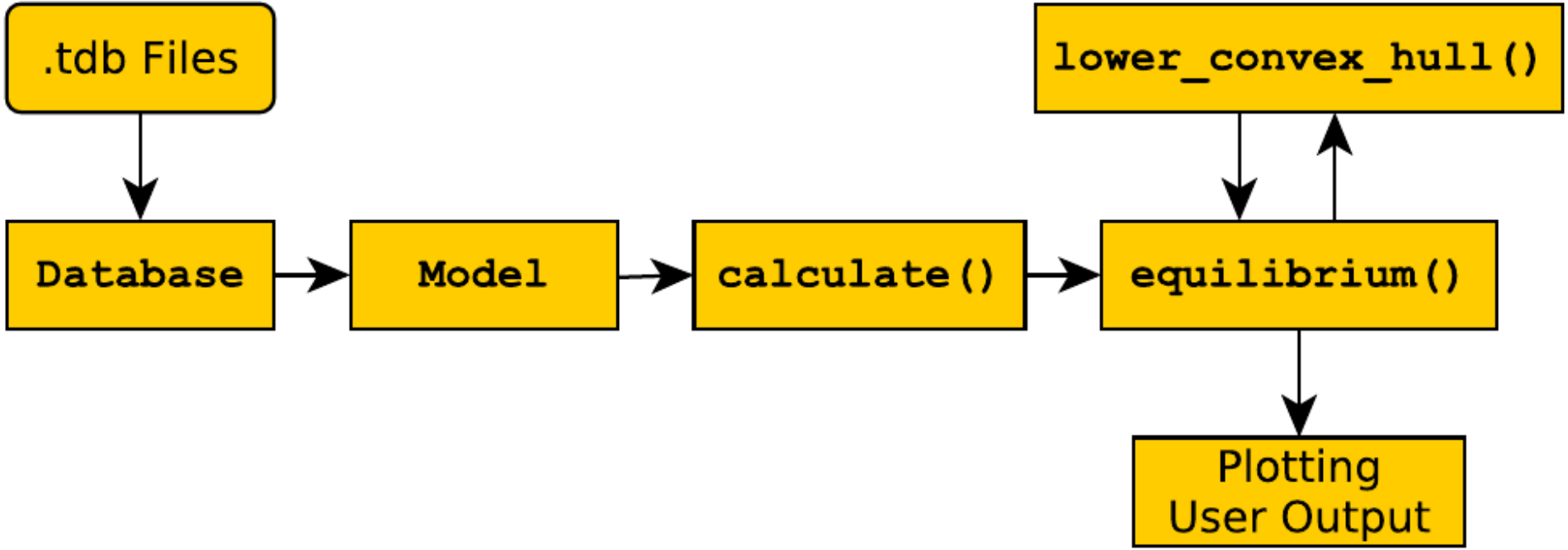


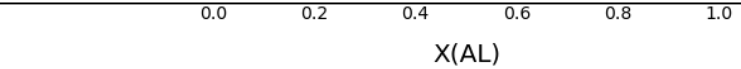
- Written in Python (open source)
- Supports multi-component alloys
- Designed for database development and prototyping

pycalphad.org

pycalphad







pycalphad.org

Uncertainty in CALPHAD Modeling

- Uncertainty Quantification (UQ)
- Uncertainty Propagation
  - Data -> Parameters -> Properties (ΔH, ΔS, T_{melt}, etc.)
- Outlier Detection
- Model Selection
  - How many parameters for a phase? Which type?
  - What does it mean for a model to be “better?”

Scoring Candidate Models

- Which columns do we include?

$$\begin{pmatrix} 1 & T_1 & T_1 \ln T_1 & \cdots & T_1^n \\ 1 & T_2 & T_2 \ln T_2 & \cdots & T_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & T_k & T_k \ln T_k & \cdots & T_k^n \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ \vdots \\ d_n \end{pmatrix} = \begin{pmatrix} \Delta G_1 \\ \Delta G_2 \\ \vdots \\ \Delta G_k \end{pmatrix}$$

Scoring Candidate Models

- Which columns do we include?

Akaike Information Criterion (AIC)

2k + n ln(RSS)

RSS = Residual Sum of Squares

k = # of parameters

n = # of data points

Parameter Optimization: Bayes' Theorem

Posterior distribution

Prior distribution

$$P(\text{Parameters} \mid \text{Data}) \sim P(\text{Data} \mid \text{Parameters}) * P(\text{Parameters})$$

Likelihood

Frequentism vs. Bayesianism

- **Frequentism**: Parameters are “fixed”, data are random variables
- **Bayesianism**: Data are “fixed”, parameters are random variables

Bayesianism and CALPHAD

- Great match because:
  - Expert/prior knowledge is key and can be quantified and added directly
  - We often know a rough acceptable range for parameter values
  - Updating the model does not require all of the old data (choose old posterior as new prior)

Markov Chain Monte Carlo (MCMC)

- Recall Bayes' theorem: How do we address the proportionality constant?
- **Answer**: Metropolis sampling.
  - We walk parameter space by randomly choosing a next step.
  - $P_{i+1}(\text{Data} \mid \text{Parameters}_{i+1}) * P_{i+1}(\text{Model}) / P_i(\text{Data} \mid \text{Parameters}_i) * P_i(\text{Model})$
  - When ratio > 1, always accept. When ratio < 1, sometimes accept.

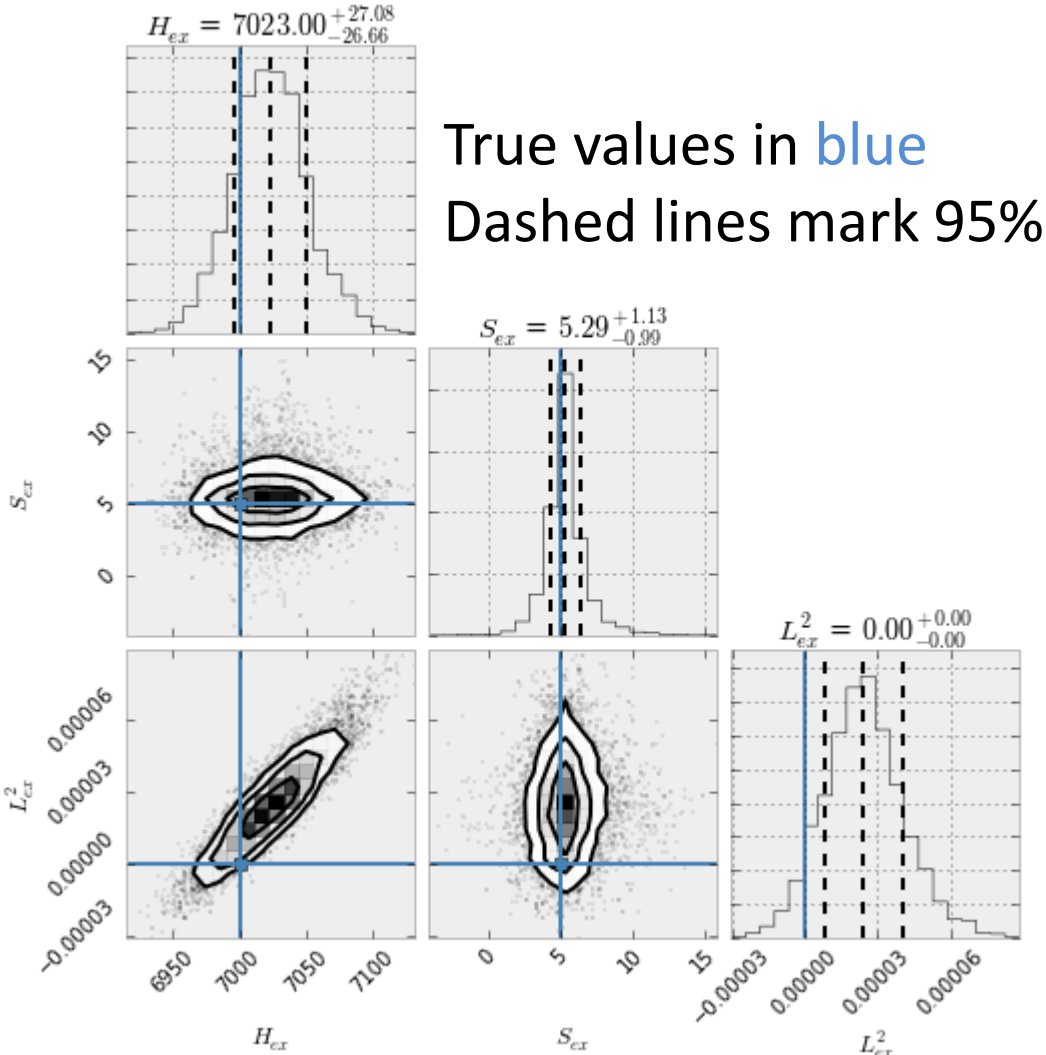
"MCMC sampling for dummies," http://twiecki.github.io/blog/2015/11/10/mcmc-sampling/. Retrieved 10 November 2015.

Constructing a Likelihood Function

- Assume normally distributed errors
- For each dataset
  - $\log L = \langle (y_{pred} - y_{data})^2 \rangle$ / variance
  - Let variance be a random variable
- Total log-likelihood is sum of all log L's
- Dataset variance also acts as a weight

Ideal Case: Binary Solution

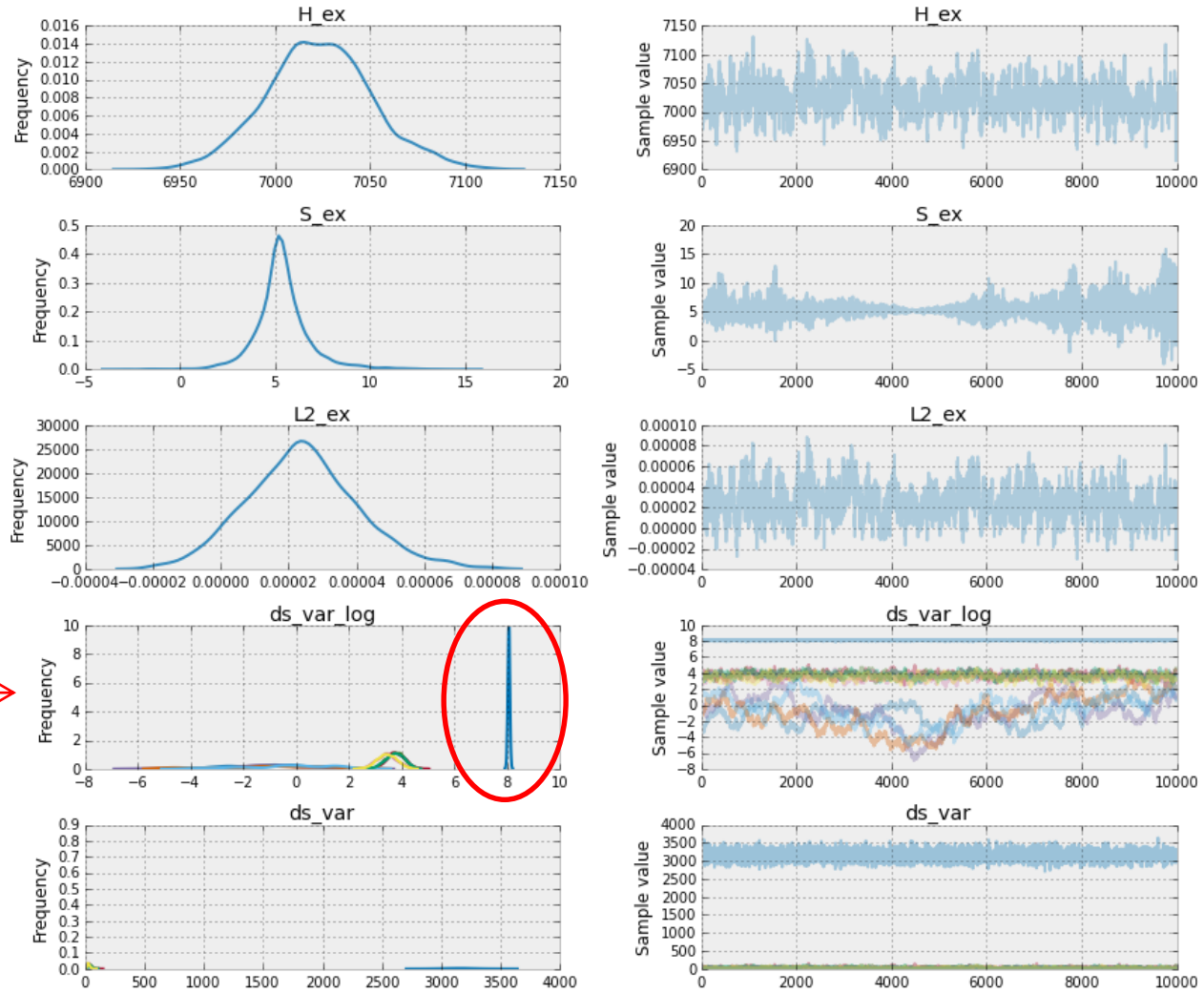
- 10 synthetic datasets with Gaussian noise
  - 5 enthalpy
  - 5 entropy
  - 1 “bad” dataset (systematic bias)
  - Equal values for prior dataset variance
- Three parameters to fit
  - $G^{ex} = (H - T*S + LT^2)x_a x_b$
- True values: H = 7000, S = 5, L = 0



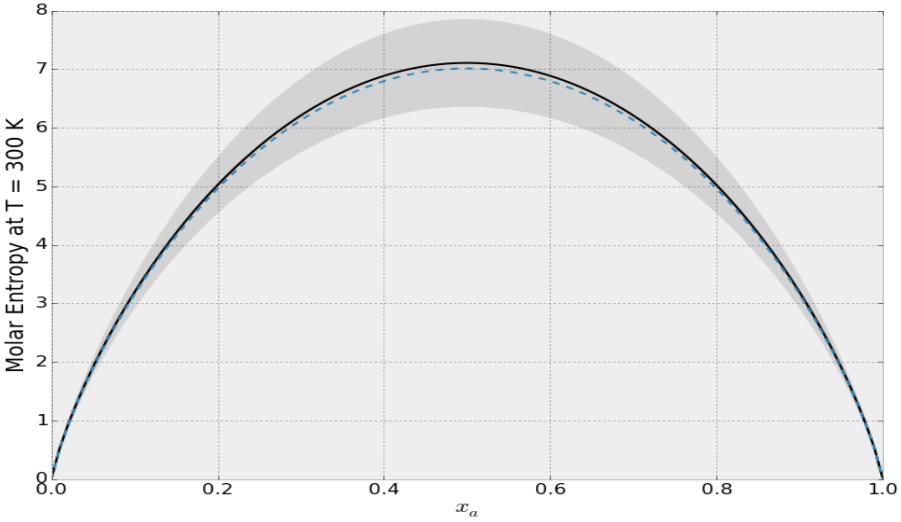
True values in blue

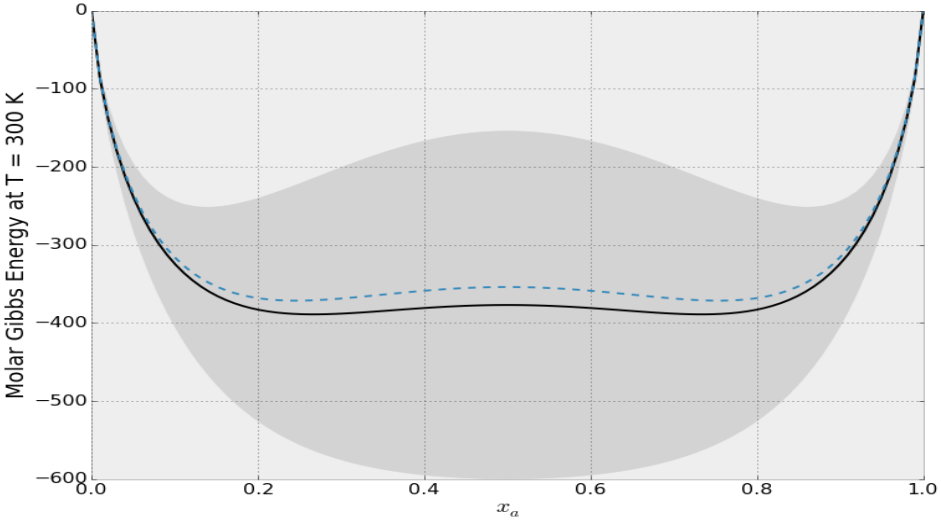
Dashed lines mark 95% credible intervals

We found the outlier dataset!



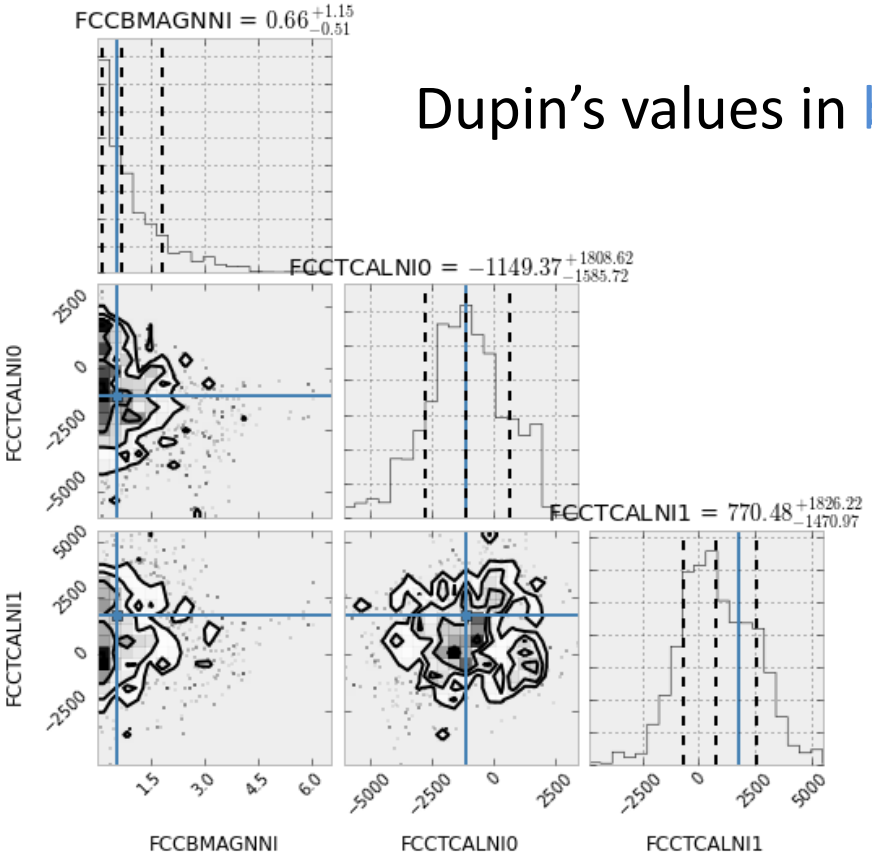
Ideal Case: Uncertainty Propagation



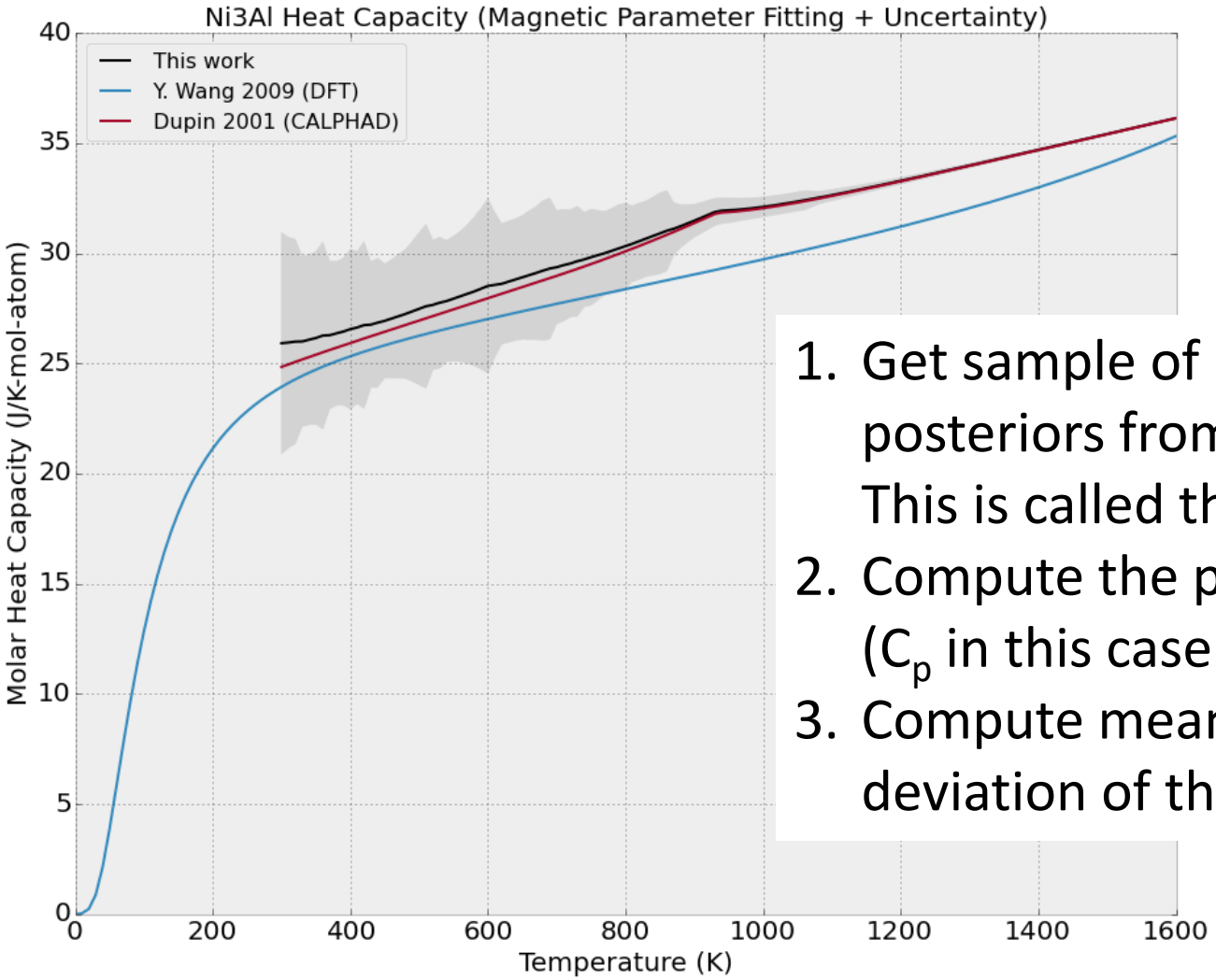


Real Case: Ni_3Al Heat Capacity

- BMAGN(FCC, Ni)
- TC(FCC, Al, Ni; 0)
- TC(FCC, Al, Ni; 1)
- Priors at zero
- Fit to DFT data
- Compare to Dupin



Dupin's values in blue



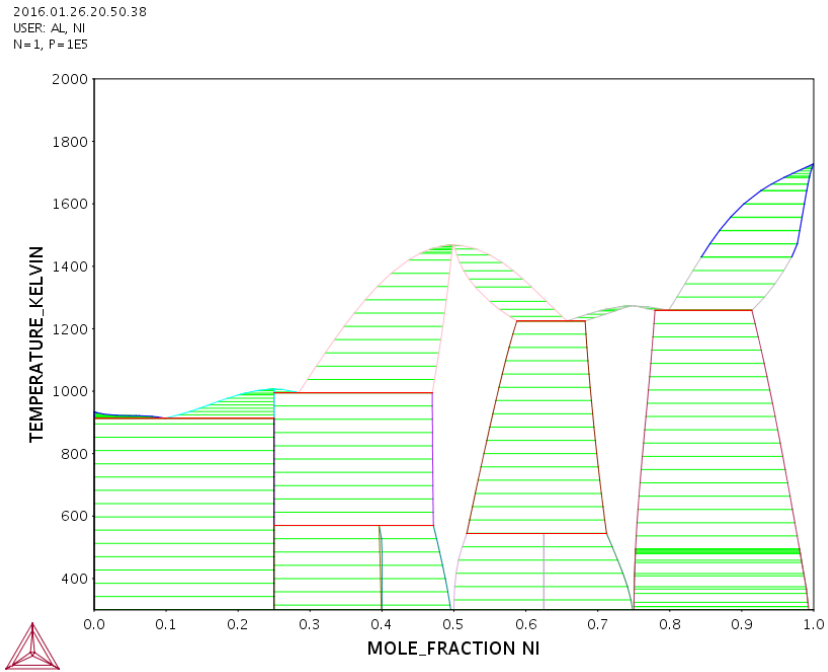
1. Get sample of parameter posteriors from MC simulation. This is called the **trace**.
2. Compute the property of interest (C_p in this case) with the trace.
3. Compute mean and standard deviation of the property.

Model Selection

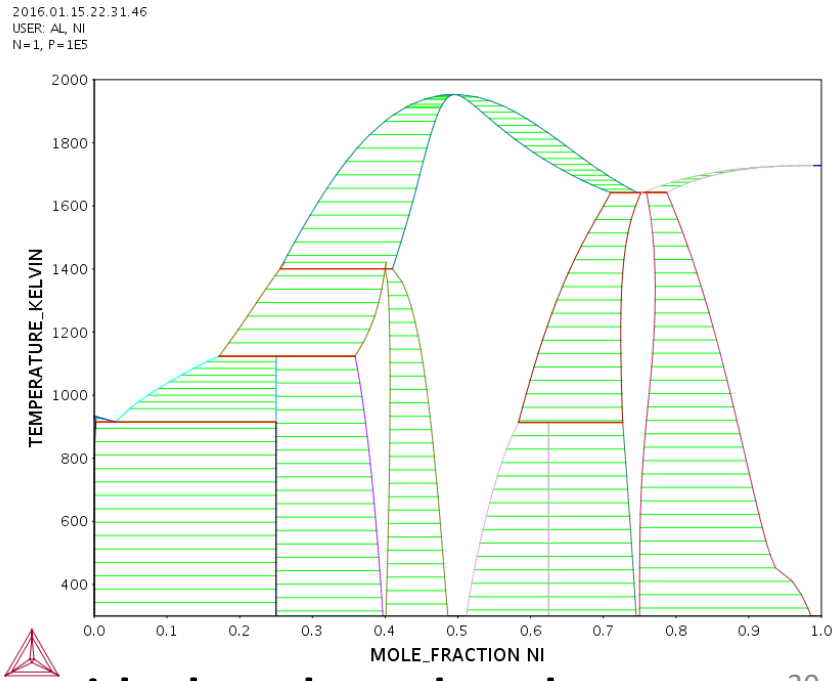
- Run MCMC separately on each candidate model, then calculate a model "score"
  - Lots of methods: BPIC, AIC, BIC, DIC, …
- Use categorical variables and compare multiple models in one MCMC simulation
  - P(Model 1), P(Model 2), etc.
  - Computationally intensive

Automating Al-Ni

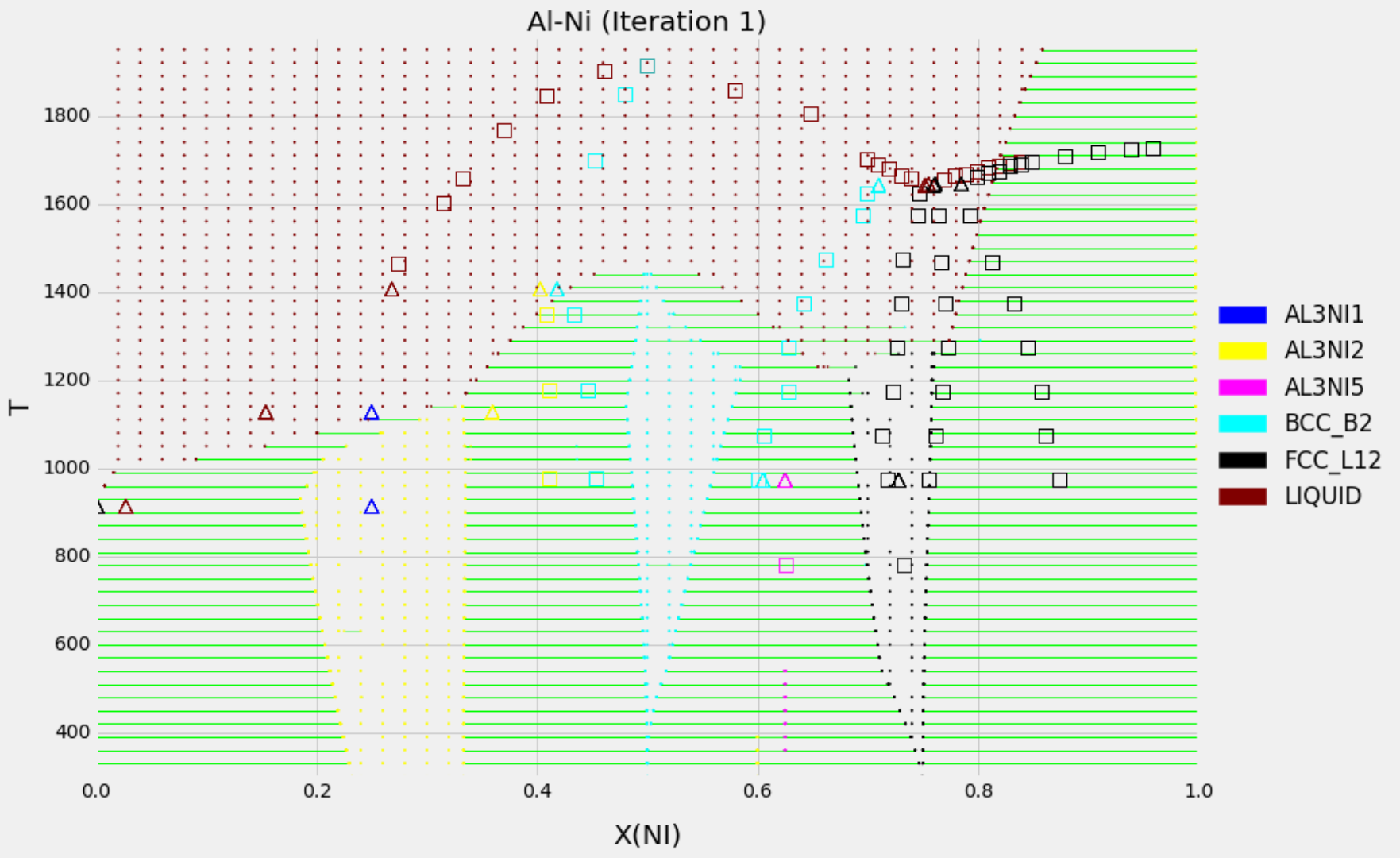
Automated DFT Phase Diagram

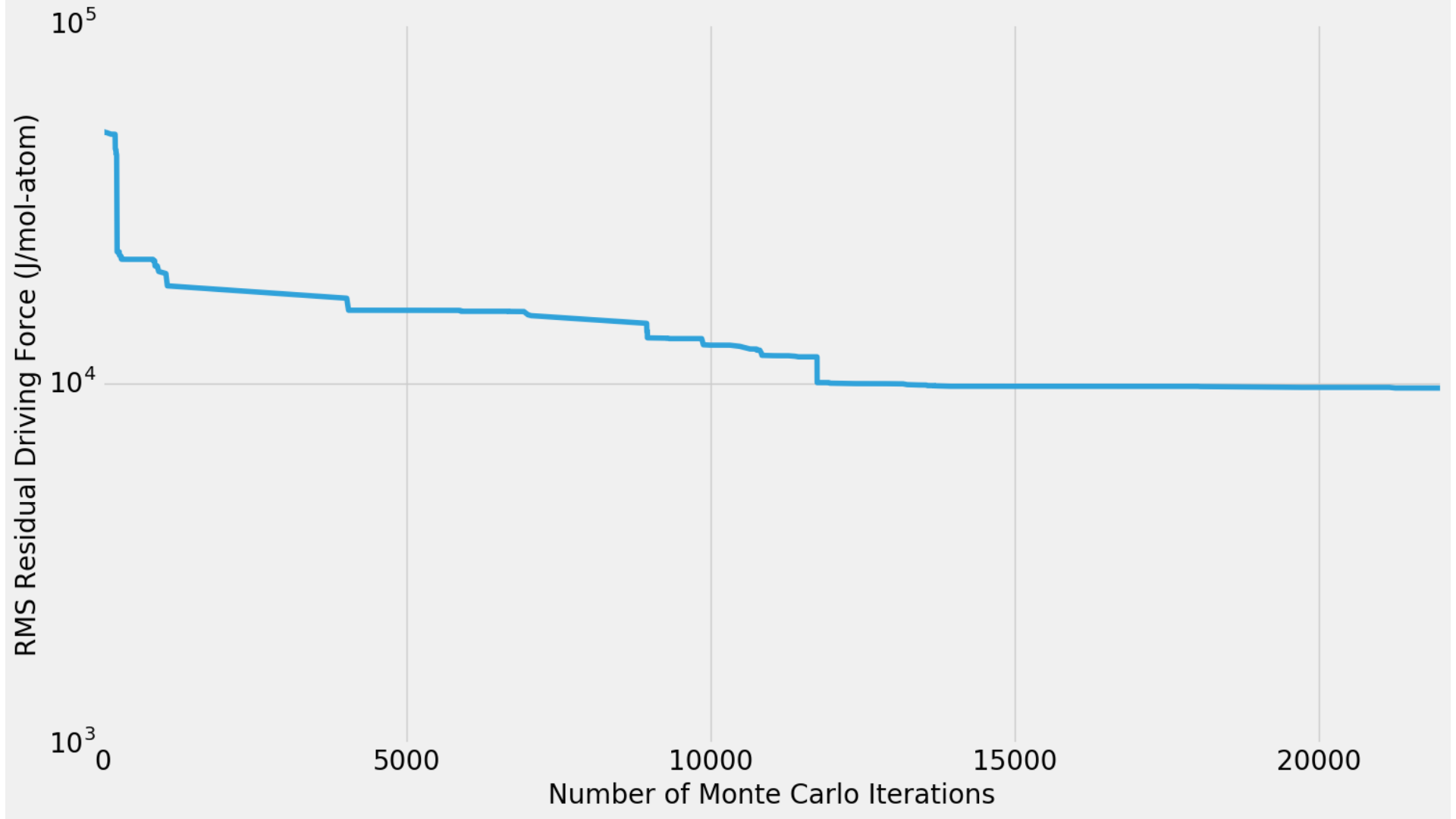


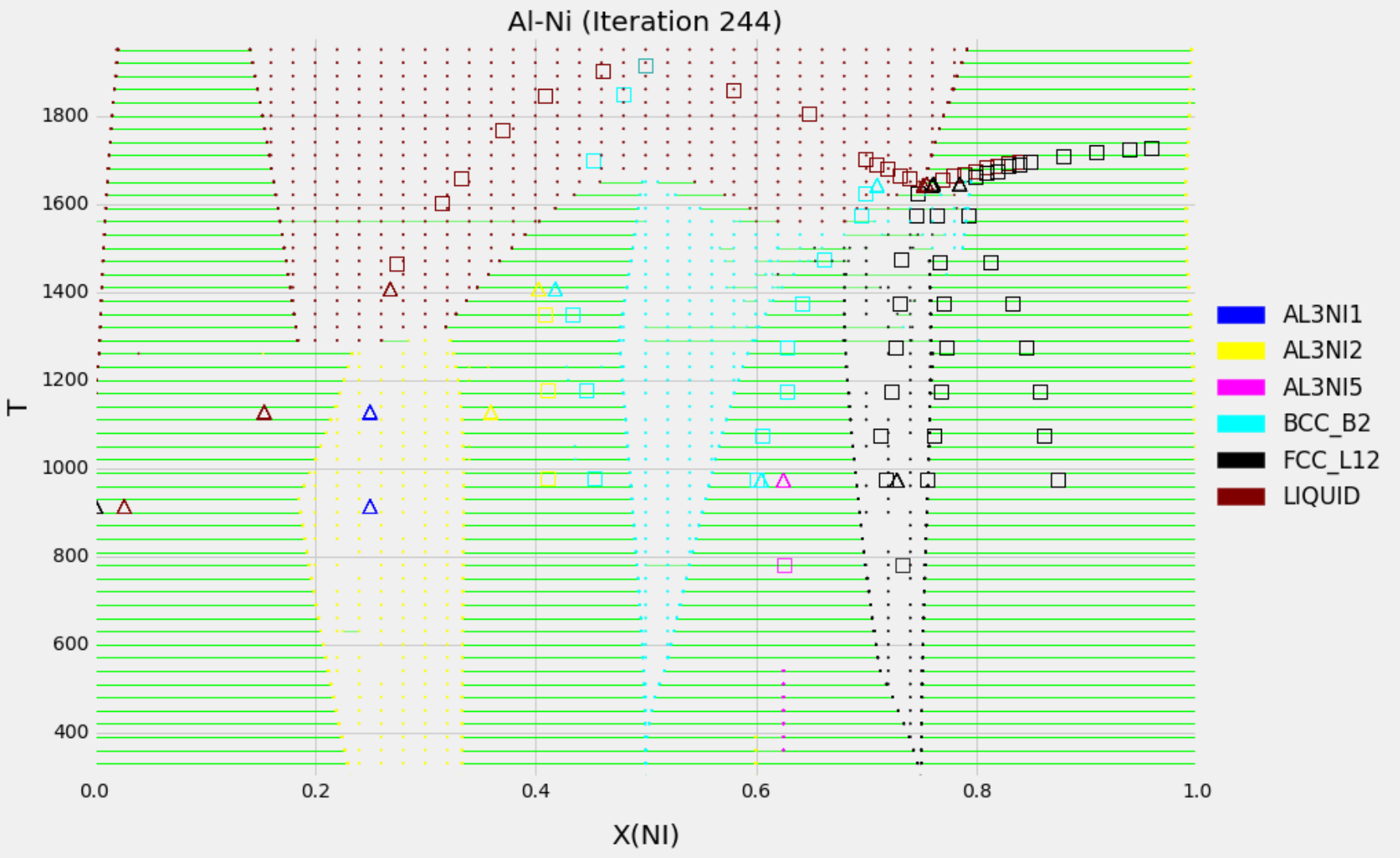
Dupin's Assessment (2001)

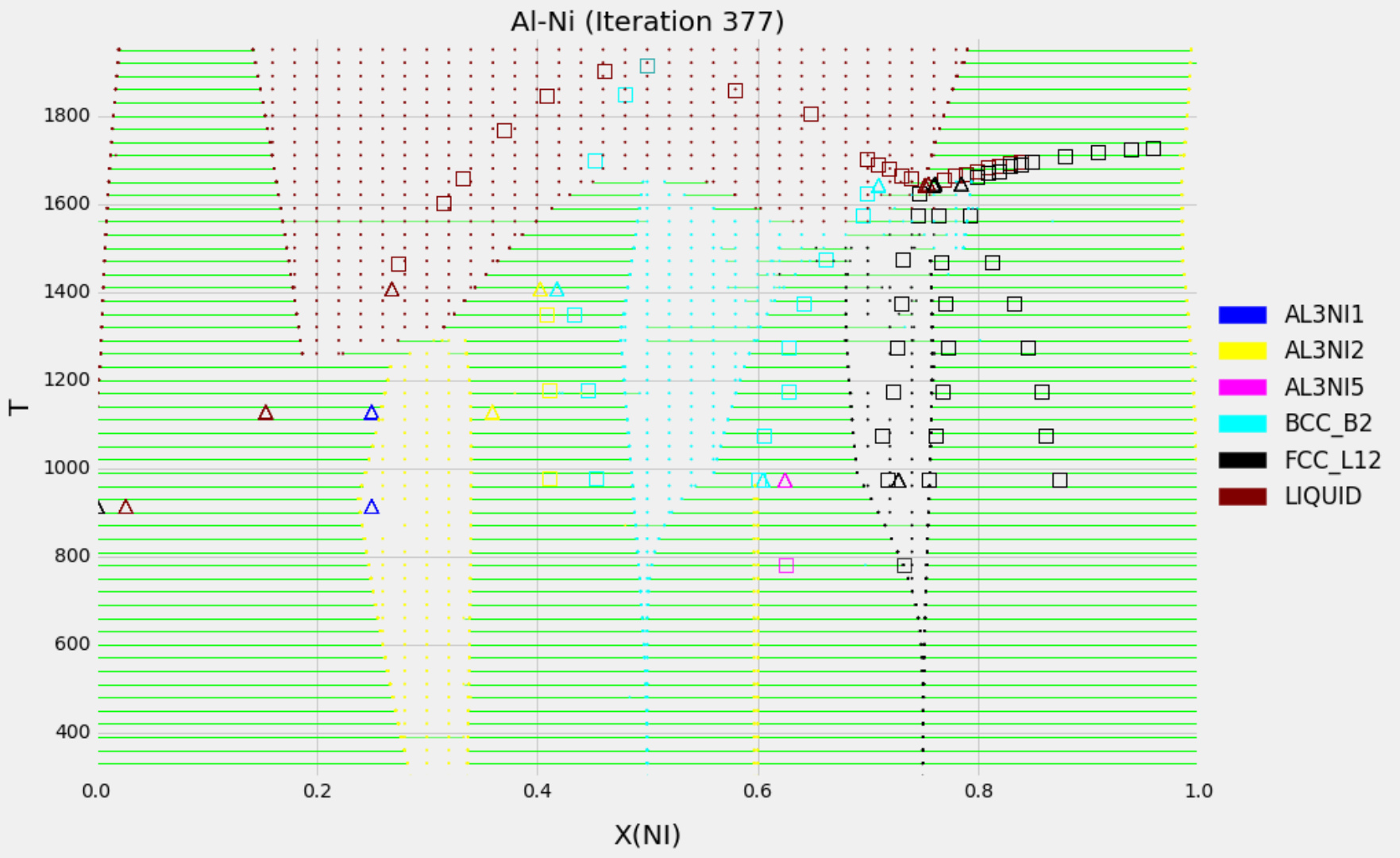


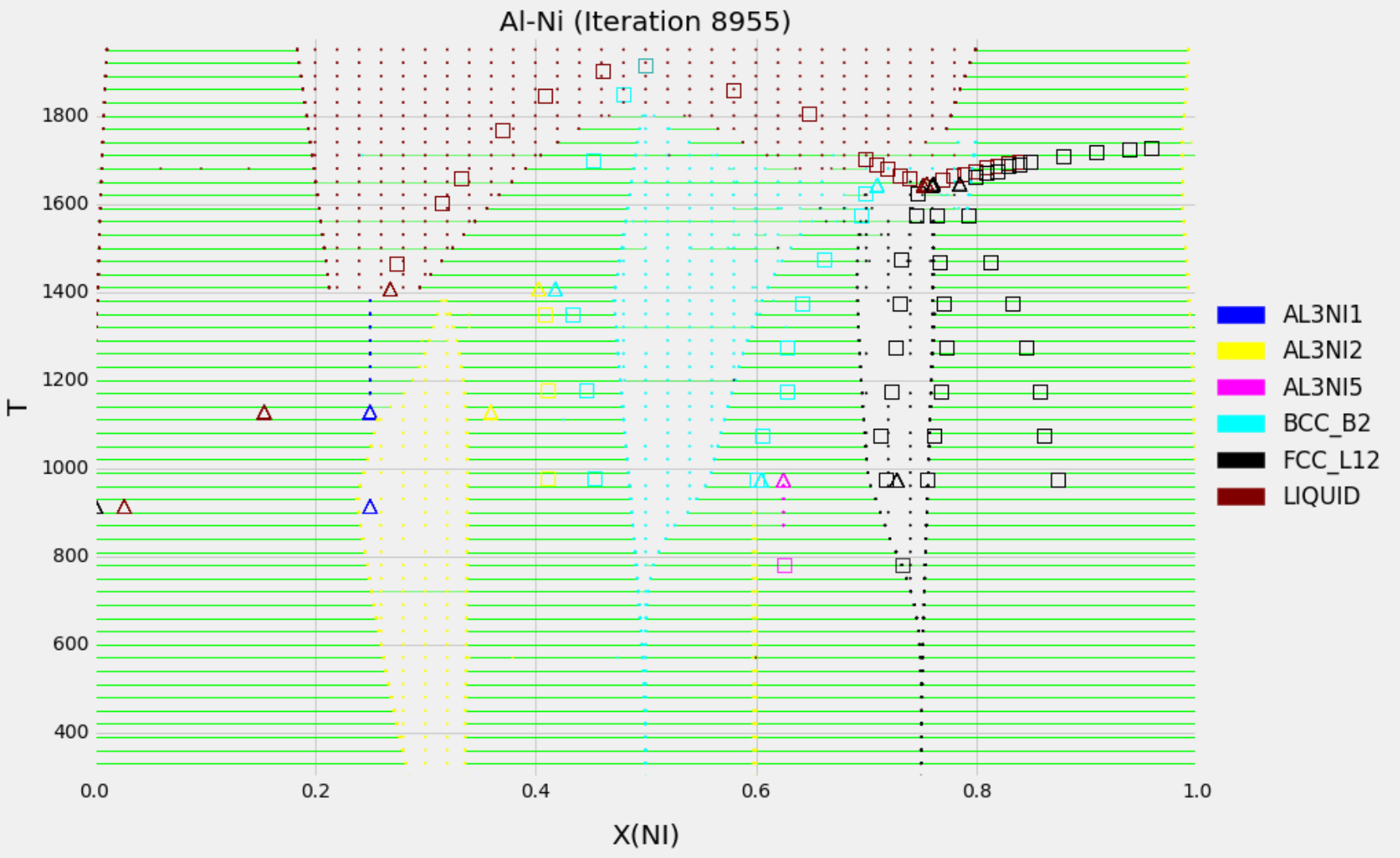
Next Step: Refine parameters with phase boundary data

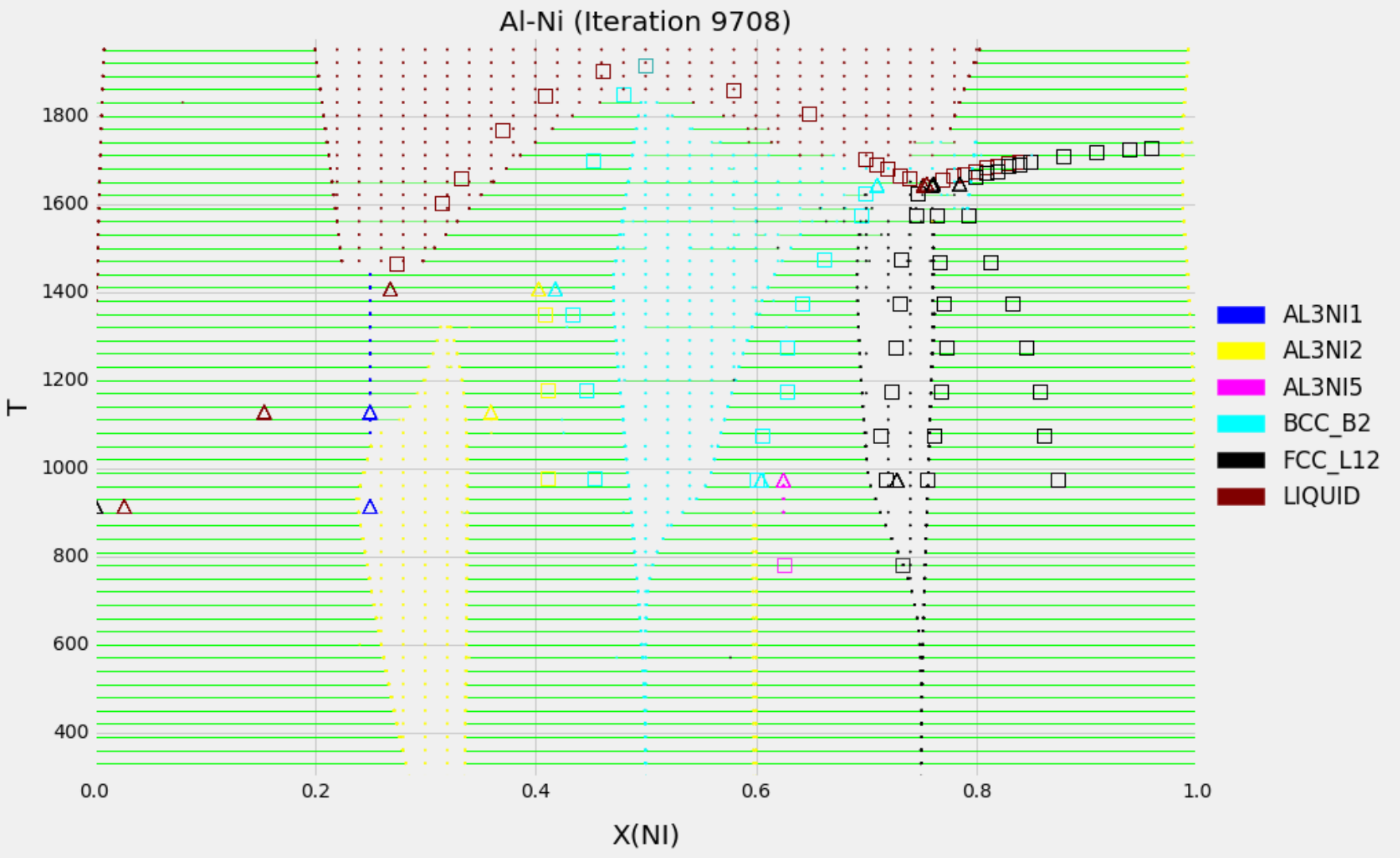

Al-Ni (Iteration 1)
T
X(NI)
1800
1600
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
AL3NI1
AL3NI2
AL3NI5
BCC_B2
FCC_L12
LIQUID

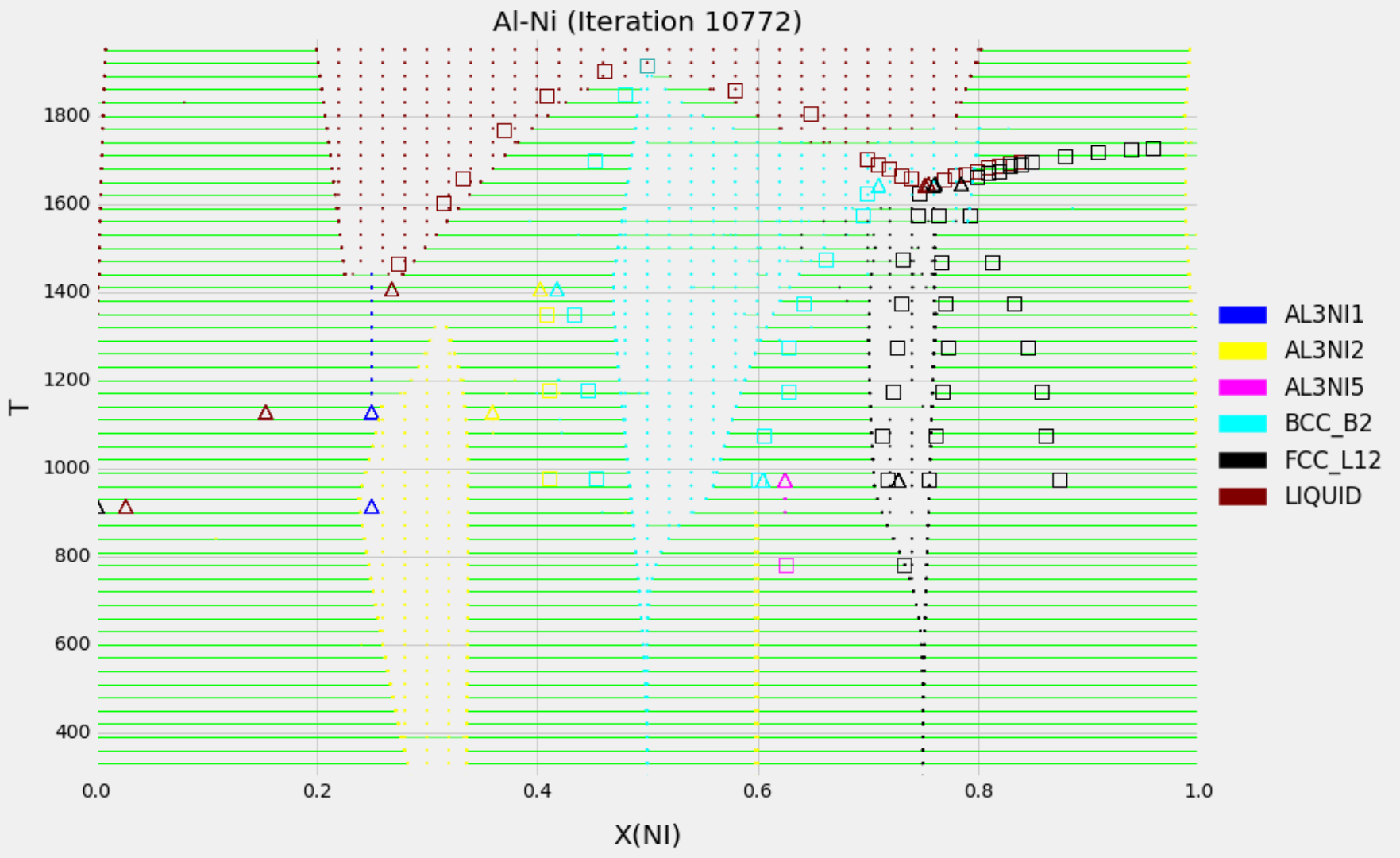

10^5
10^4
10^3
RMS Residual Driving Force (J/mol-atom)
0
5000
10000
15000
20000
Number of Monte Carlo Iterations

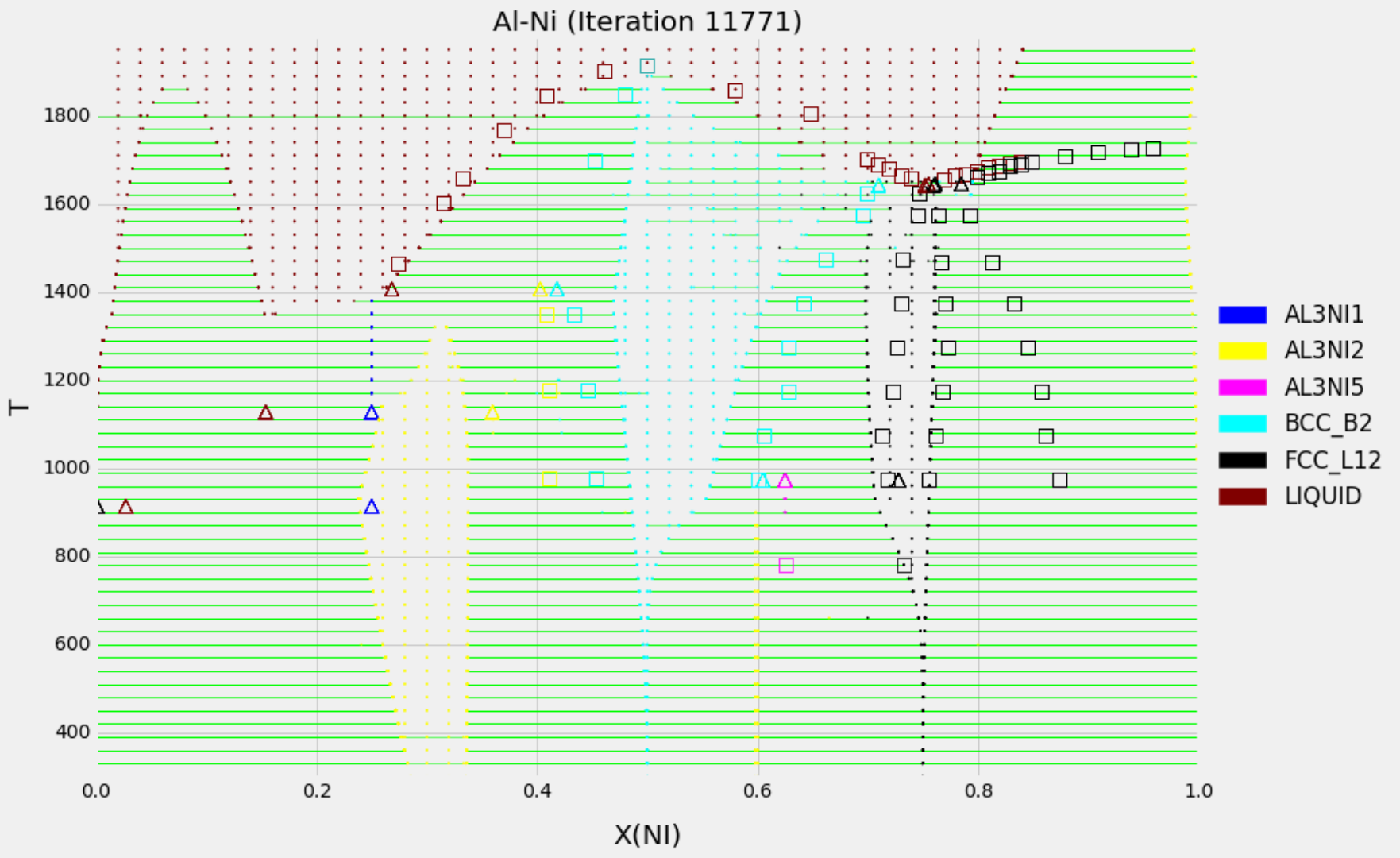

Al-Ni (Iteration 244)
T
X(NI)
1800
1600
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
AL3NI1
AL3NI2
AL3NI5
BCC_B2
FCC_L12
LIQUID

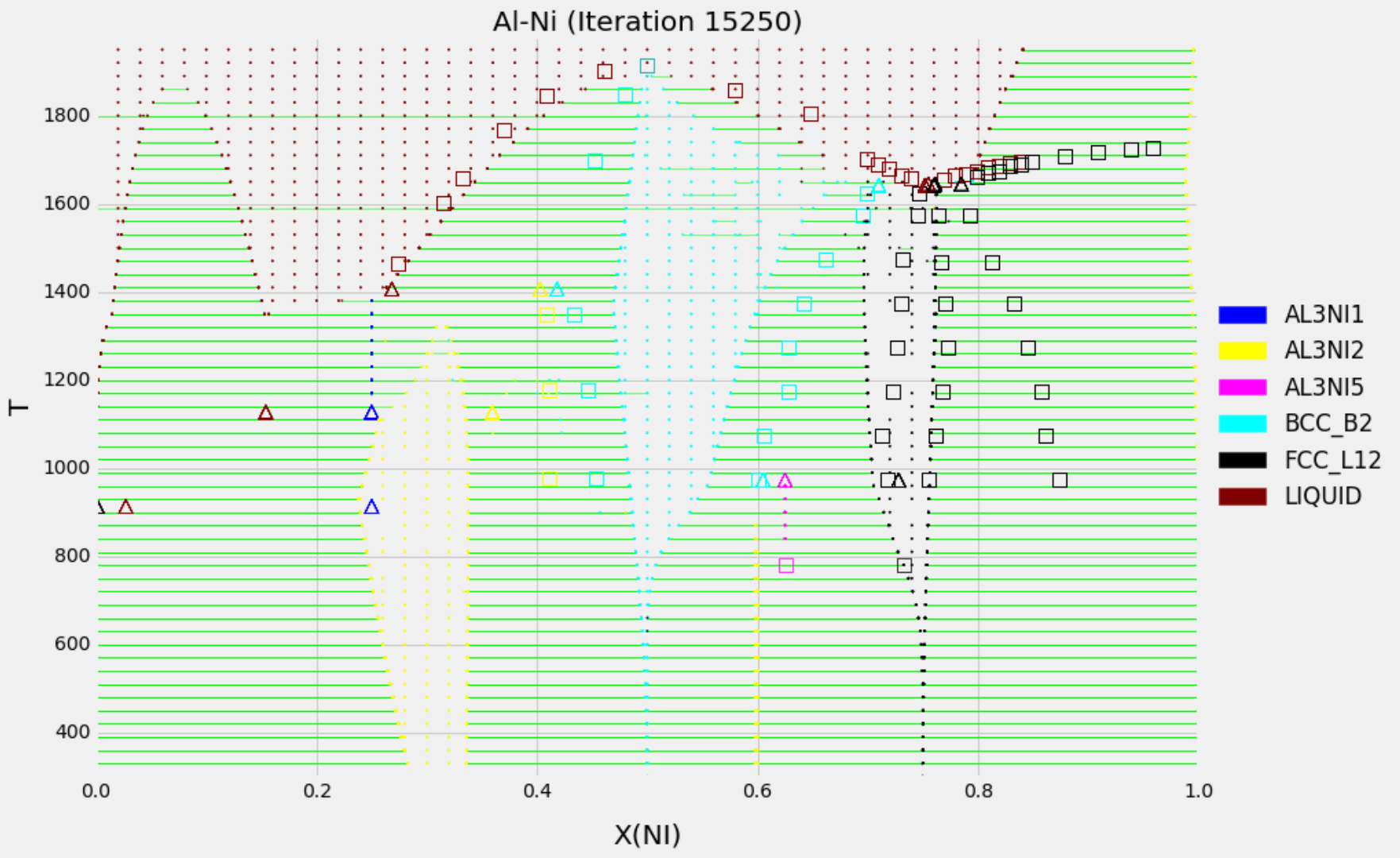

Al-Ni (Iteration 377)
T
X(NI)
1800
1600
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
AL3NI1
AL3NI2
AL3NI5
BCC_B2
FCC_L12
LIQUID

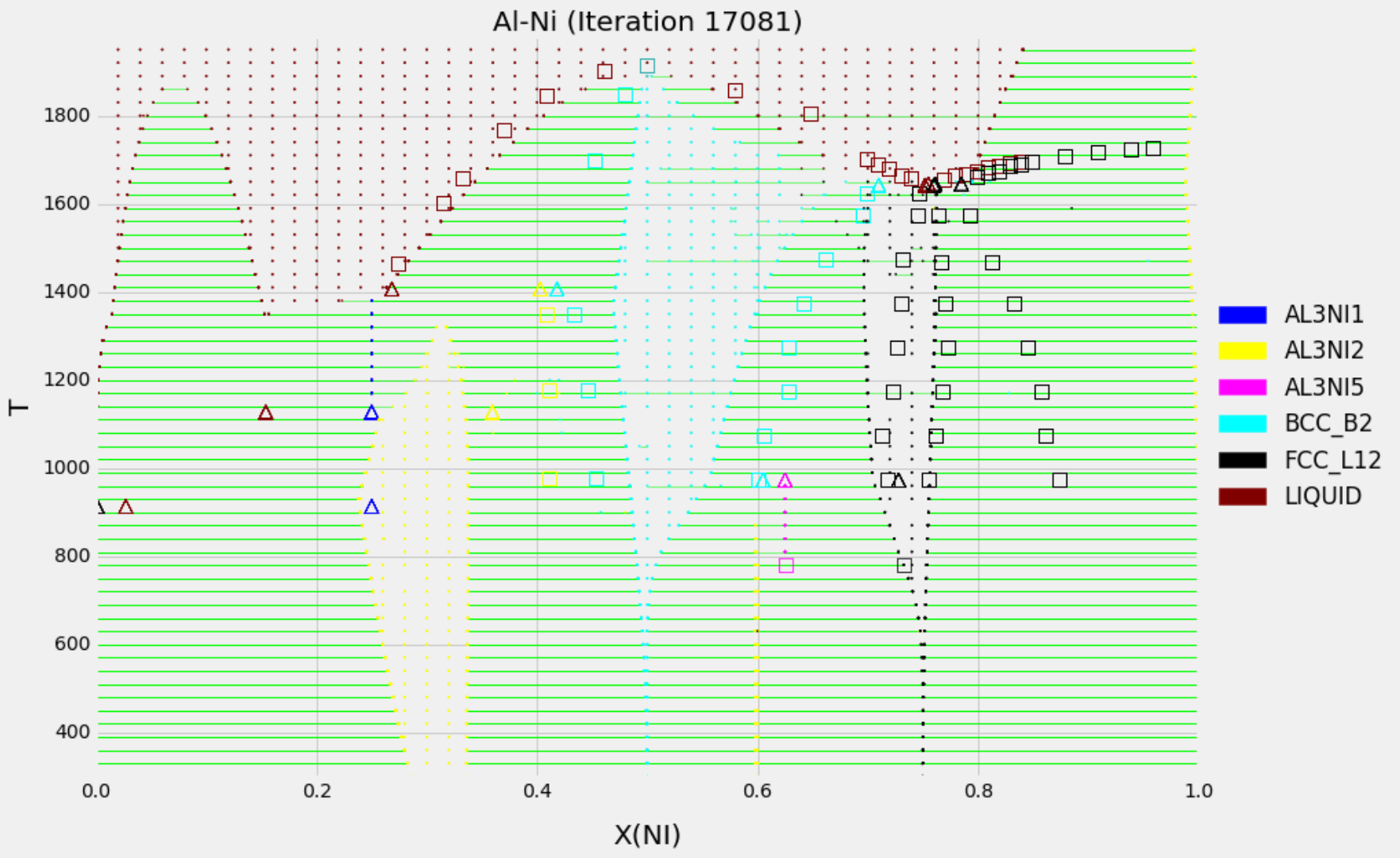

Al-Ni (Iteration 8955)
T
X(NI)
1800
1600
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
AL3NI1
AL3NI2
AL3NI5
BCC_B2
FCC_L12
LIQUID


Al-Ni (Iteration 9708)
T
X(NI)
1800
1600
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
AL3NI1
AL3NI2
AL3NI5
BCC_B2
FCC_L12
LIQUID


Al-Ni (Iteration 10772)
T
X(NI)
1800
1600
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
AL3NI1
AL3NI2
AL3NI5
BCC_B2
FCC_L12
LIQUID


Al-Ni (Iteration 11771)
T
X(NI)
1800
1600
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
AL3NI1
AL3NI2
AL3NI5
BCC_B2
FCC_L12
LIQUID


Al-Ni (Iteration 15250)
T
X(NI)
1800
1600
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
AL3NI1
AL3NI2
AL3NI5
BCC_B2
FCC_L12
LIQUID


Al-Ni (Iteration 17081)
T
X(NI)
1800
1600
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
AL3NI1
AL3NI2
AL3NI5
BCC_B2
FCC_L12
LIQUID

Future of CALPHAD

- We don’t have one “best” model; we have an ensemble of models based on joint probability of parameter distributions
- Parameter distributions are updated over time with new evidence (data and/or beliefs)
- The “best” or most probable models can be automatically determined using MCMC

Thank You

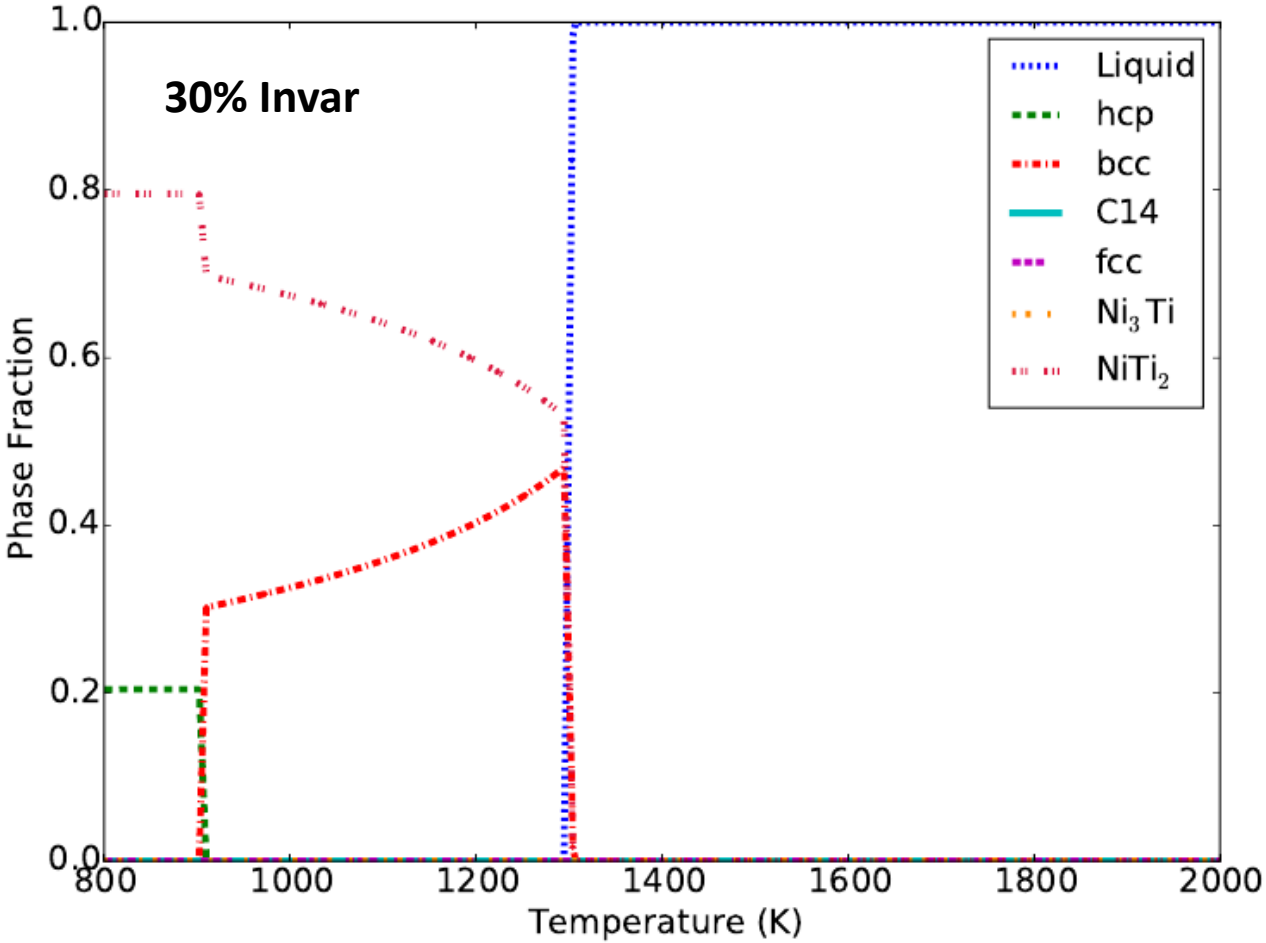
richard.otis@jpl.nasa.gov

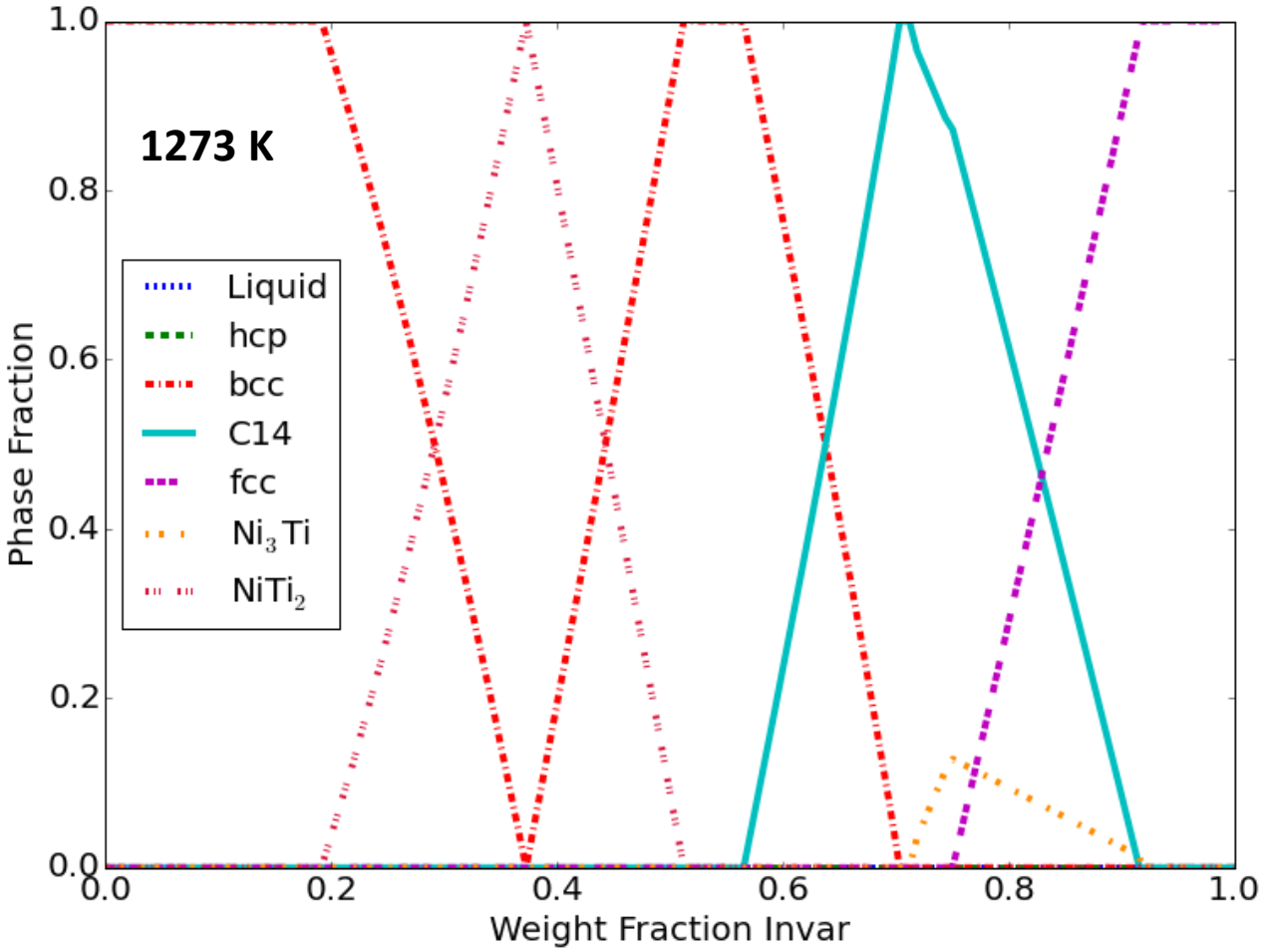
www.github.com/richardotis

This work was supported by a NASA Space Technology Research Fellowship under grant number NNX14AL43H.

Automation Challenges

- Getting Structured Data
  - Tends to be jealously guarded
  - Solution is tedious but tractable
- Sublattice Model Determination
  - For now, humans must perform this step


30% Invar
Liquid
hcp
bcc
C14
fcc
Ni3Ti
NiTi2
Phase Fraction
Temperature (K)
1.0
0.8
0.6
0.4
0.2
0.0
800
1000
1200
1400
1600
1800
2000


1273 K
Liquid
hcp
bcc
C14
fcc
Ni_3Ti
$NiTi_2$
Phase Fraction
Weight Fraction Invar
0.0
0.2
0.4
0.6
0.8
1.0

Previous Work in Bayesian Methods

- ChemSage (1995) was first to incorporate Bayesian methods into CALPHAD modeling
  - Update priors by iterative procedure using covariance matrix
- Stan and Reardon (Calphad, 2003) discussed implementing Bayesian statistics using a genetic algorithm (no public implementation)

Königsberger, E., Eriksson, G., A new optimization routine for chemsage, Calphad, Volume 19, Issue 2, June 1995, Pages 207-214.