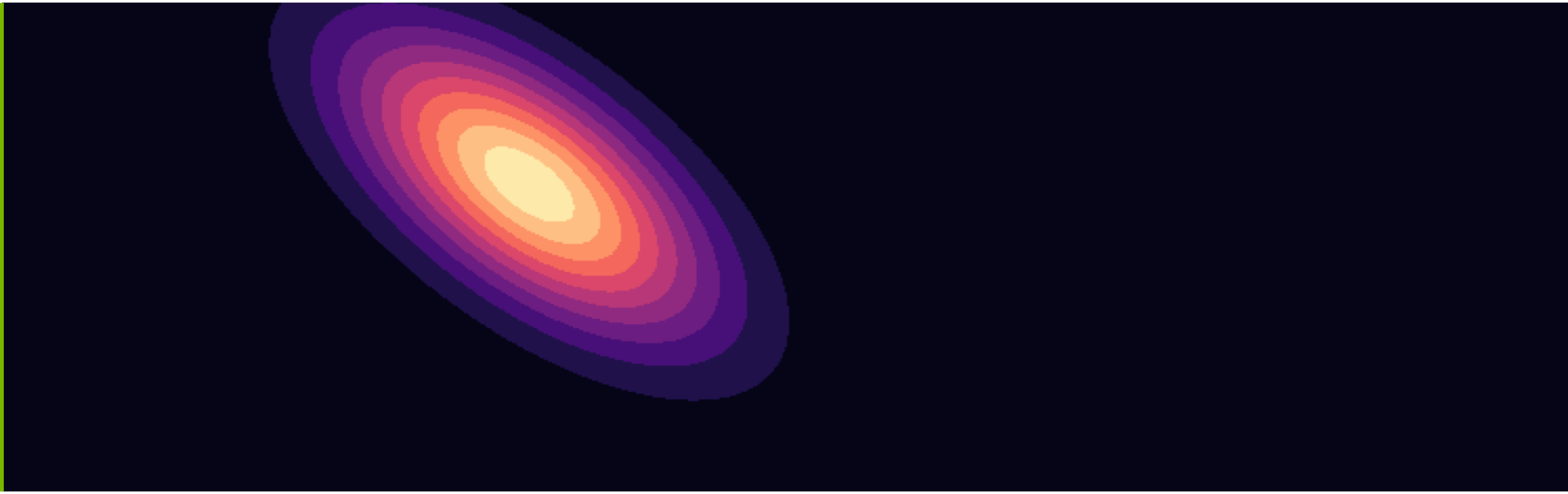
SIAM CSE, 2019
SPOKANE, WA, FEBRUARY 25, 2019



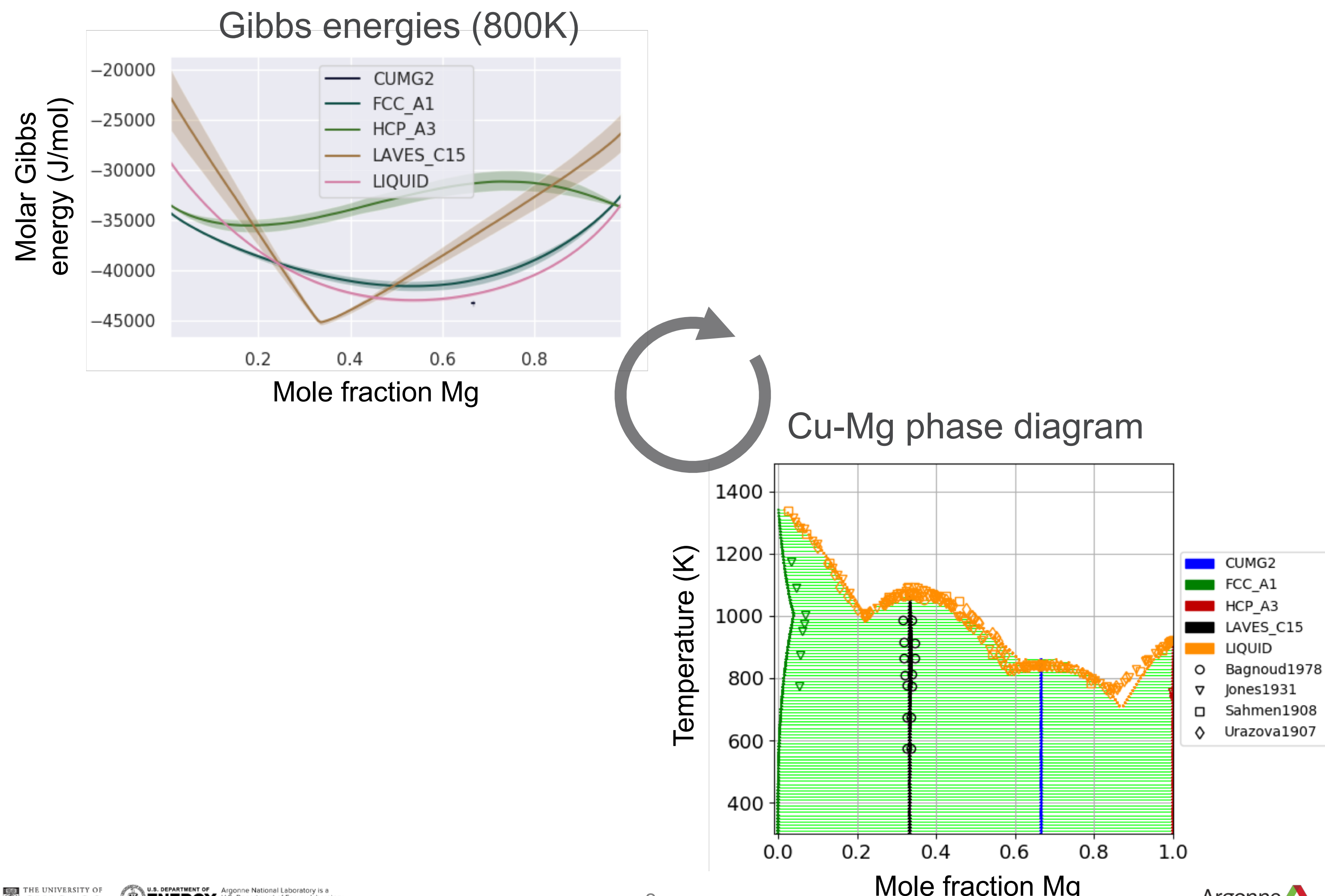
BAYESIAN MODEL SELECTION, CALIBRATION AND UNCERTAINTY QUANTIFICATION OF THERMODYNAMIC PROPERTIES

NOAH H. PAULSON
Postdoctoral Research Associate
Argonne National Laboratory
Applied Materials Division
Lemont, IL

ELISE JENNINGS
Computer Scientist
Argonne National Laboratory
Leadership Computing Facility
Lemont, IL

MARIUS STAN
Senior Scientist
Argonne National Laboratory
Applied Materials Division
Lemont, IL

MOTIVATION


Gibbs energies (800K)
Molar Gibbs energy (J/mol)
−20000
−25000
−30000
−35000
−40000
−45000
CUMG2
FCC_A1
HCP_A3
LAVES_C15
LIQUID
0.2
0.4
0.6
0.8
Mole fraction Mg
Cu-Mg phase diagram
Temperature (K)
1400
1200
1000
800
600
400
0.0
0.2
0.4
0.6
0.8
1.0
Mole fraction Mg
CUMG2
FCC_A1
HCP_A3
LAVES_C15
LIQUID
Bagnoud1978
Jones1931
Sahmen1908
Urazova1907

BACKGROUND
BAYESIAN STATISTICS

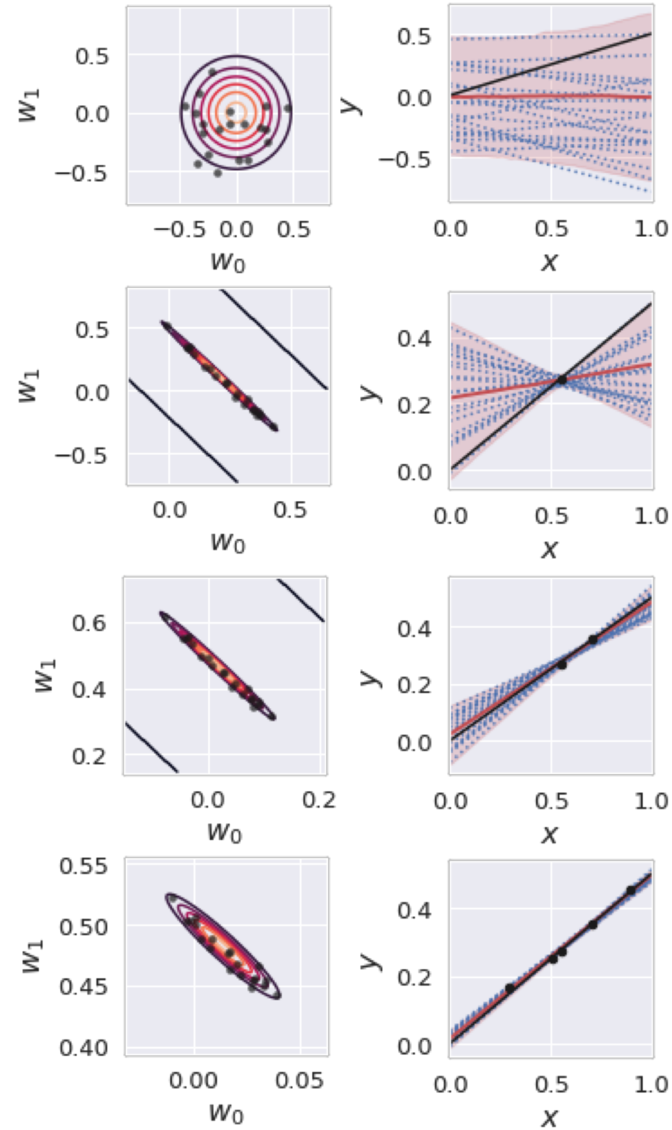
- Bayesian statistics provides a framework for model selection, calibration and UQ
- for some model, M, Bayes' theorem states:

posterior parameter distribution | data Likelihood | prior parameter distribution | marginal Likelihood

$$\Pr(\boldsymbol{\theta}|\mathbf{D},M) = \frac{\Pr(\mathbf{D}|\boldsymbol{\theta},M)\ \Pr(\boldsymbol{\theta}|M)}{\Pr(\mathbf{D}|M)}$$

$$R = \frac{\Pr(M_A|\mathbf{D})}{\Pr(M_B|\mathbf{D})} = \frac{\Pr(\mathbf{D}|M_A)\Pr(M_A)}{\Pr(\mathbf{D}|M_B)\Pr(M_B)}$$

$$y = w_0 + x w_1$$



Bishop 2006

THE UNIVERSITY OF CHICAGO
U.S. DEPARTMENT OF ENERGY
Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.
Argonne NATIONAL LABORATORY

BACKGROUND
BAYESIAN STATISTICS

- Bayesian statistics provides a framework for model selection, calibration and UQ
- for some model, M, Bayes' theorem states:

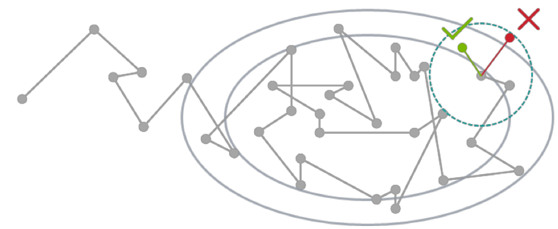
posterior parameter distribution	data Likelihood	prior parameter distribution

$$\Pr(\boldsymbol{\theta}|\mathbf{D},M) = \frac{\Pr(\mathbf{D}|\boldsymbol{\theta},M)\ \Pr(\boldsymbol{\theta}|M)}{\Pr(\mathbf{D}|M)}$$

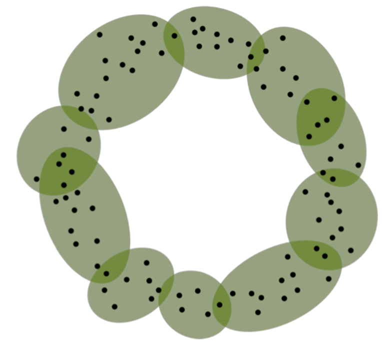
marginal Likelihood

$$R = \frac{\Pr(M_A|\mathbf{D})}{\Pr(M_B|\mathbf{D})} = \frac{\Pr(\mathbf{D}|M_A)\Pr(M_A)}{\Pr(\mathbf{D}|M_B)\Pr(M_B)}$$

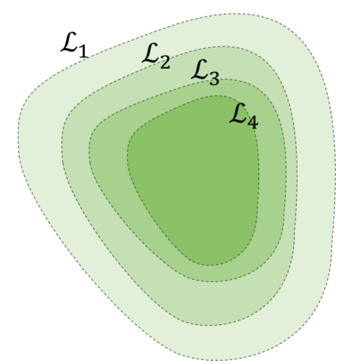
Metropolis-Hastings Markov chain Monte Carlo



affine invariant ensemble sampler / kombine



MultiNest



Bishop 2006

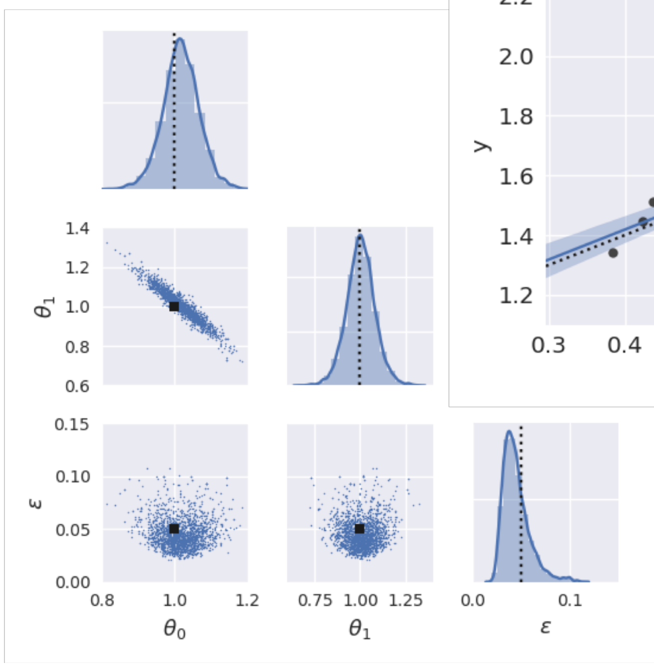
BAYESIAN FRAMEWORK
FITTING A MODEL

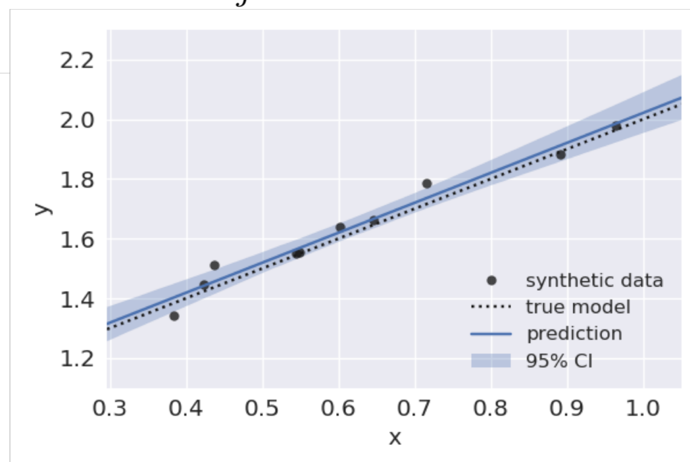
Ex: $C_p^{liq}(T) = c_1 + c_2 T$

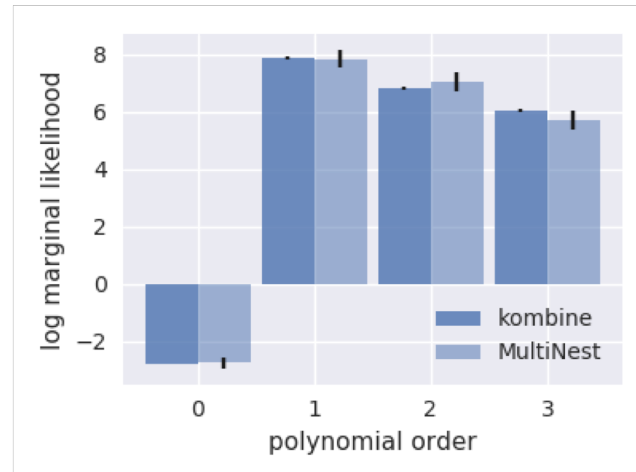
Polynomial Model: $M\left(x, \mathbf{\Theta}\right) = \sum_{i=o}^{N} \Theta_i * x^i$

Priors: $Pr\left(\Theta_i | M\right) = \mathcal{U}\left(\Theta_i | -2, 2\right)$
$Pr\left(\varepsilon | M\right) = \mathcal{U}\left(\varepsilon | 0, 1\right)$

Likelihood: $Pr\left(\mathbf{D} | \mathbf{\Theta}, \varepsilon, M\right) = \prod_j \mathcal{N}\left(y_j | M\left(x_j, \mathbf{\Theta}\right), \varepsilon\right)$







Paulson 2018

BAYESIAN FRAMEWORK
SPECIALIZED METHODS

Thermodynamic Consistency

$$Pr\left(\mathbf{D}|\mathbf{\Theta}, M_H, M_{C_p}\right) =$$

$$\prod_i \mathcal{N}\left(H_i | M_H\left(T_i, \mathbf{\Theta}\right), \varepsilon_i\right) \cdot$$

$$\prod_j \mathcal{N}\left(C_{p\,j} | M_{C_p}\left(T_j, \mathbf{\Theta}\right), \varepsilon_j\right)$$

Rescaling of Errors

$$Pr\left(\mathbf{D}|\mathbf{\Theta}, \alpha, M\right) =$$

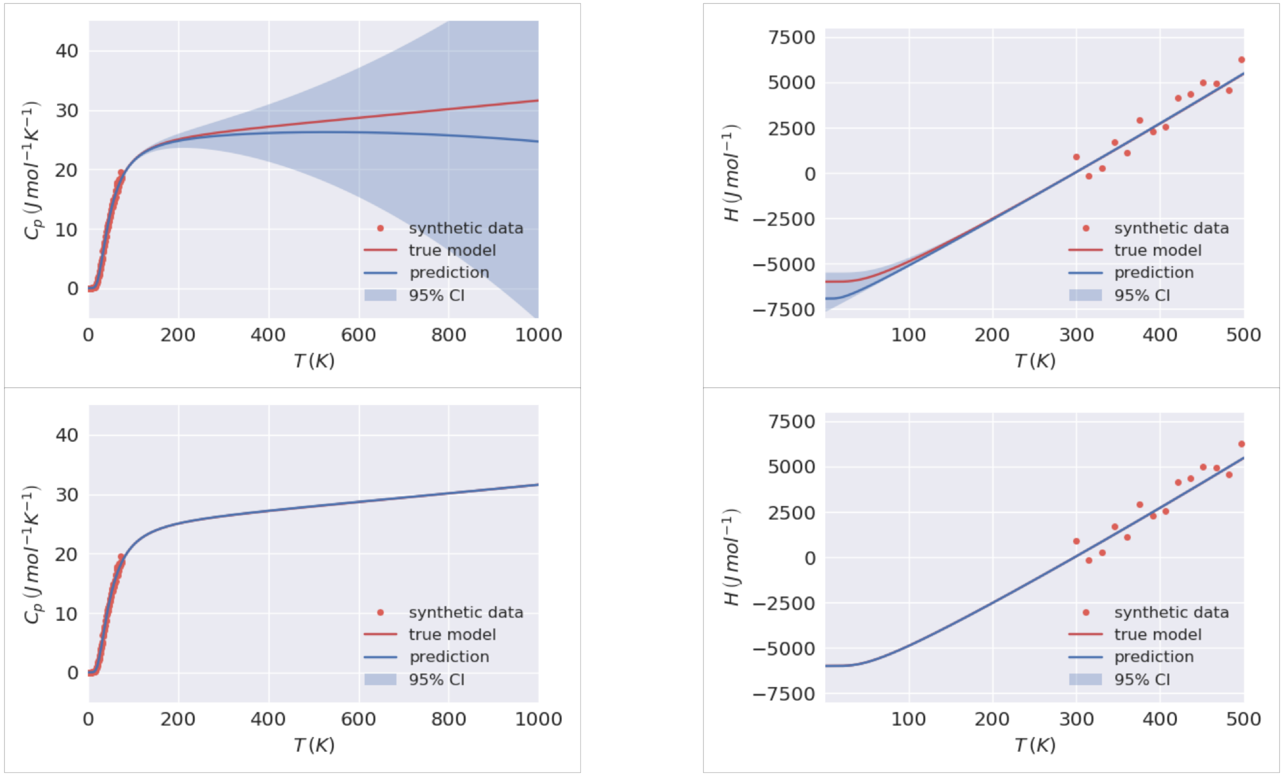
$$\prod_i \prod_j \mathcal{N}\left(y_j^i | M\left(x_j^i, \mathbf{\Theta}\right), \varepsilon_j^i / \alpha^i\right)$$

Gelman 2014, Ma 2014, Paulson 2018

BAYESIAN FRAMEWORK
THERMODYNAMIC CONSISTENCY

Einstein Model

$$C_p(T) = \frac{d}{dT}H(T) = \underbrace{3R\left(\theta/T\right)^2 \frac{e^{\theta/T}}{\left(\theta/T - 1\right)^2}}_{\text{Einstein Model}} + aT + bT^2$$



BAYESIAN FRAMEWORK
DATA WEIGHTING

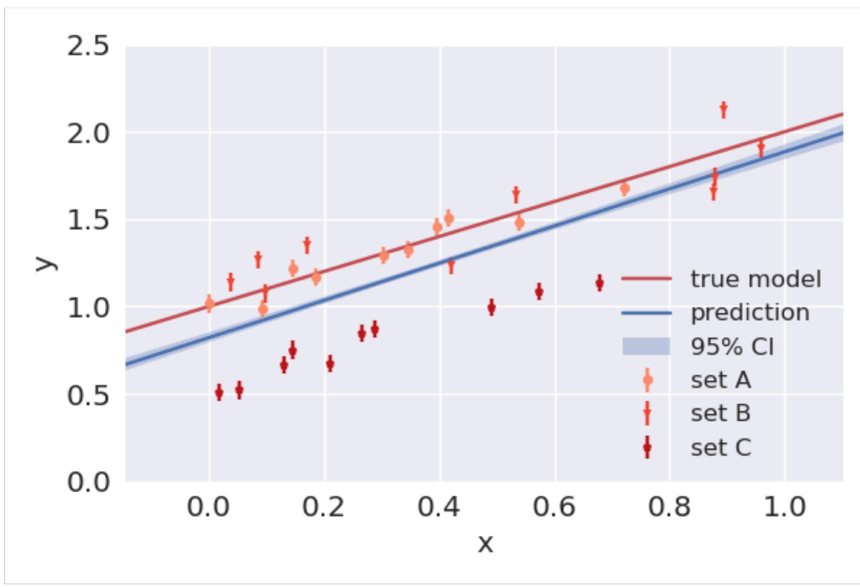
Likelihood: $$\Pr(\mathbf{D}|\boldsymbol{\Theta},\boldsymbol{\alpha},M)=\prod_{\mathrm{i}}\prod_{\mathrm{j}}N\left(\mathrm{y}_{\mathrm{j}}^{\mathrm{i}}\middle|M\left(\mathrm{x}_{\mathrm{j}}^{\mathrm{i}},\boldsymbol{\Theta}\right),\varepsilon_{\mathrm{j}}^{\mathrm{i}}/\alpha^{\mathrm{i}}\right)$$

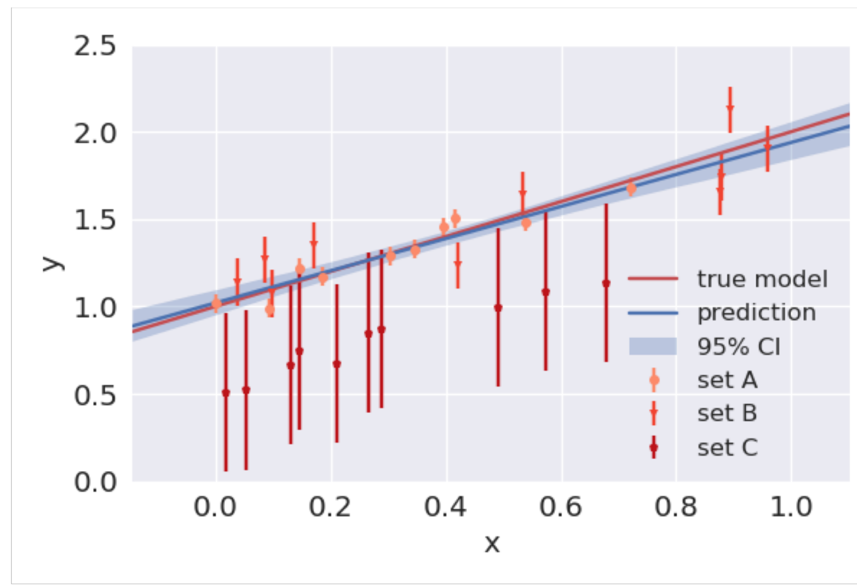
reported error

dataset index

data point index

Bayesian hyperparameter

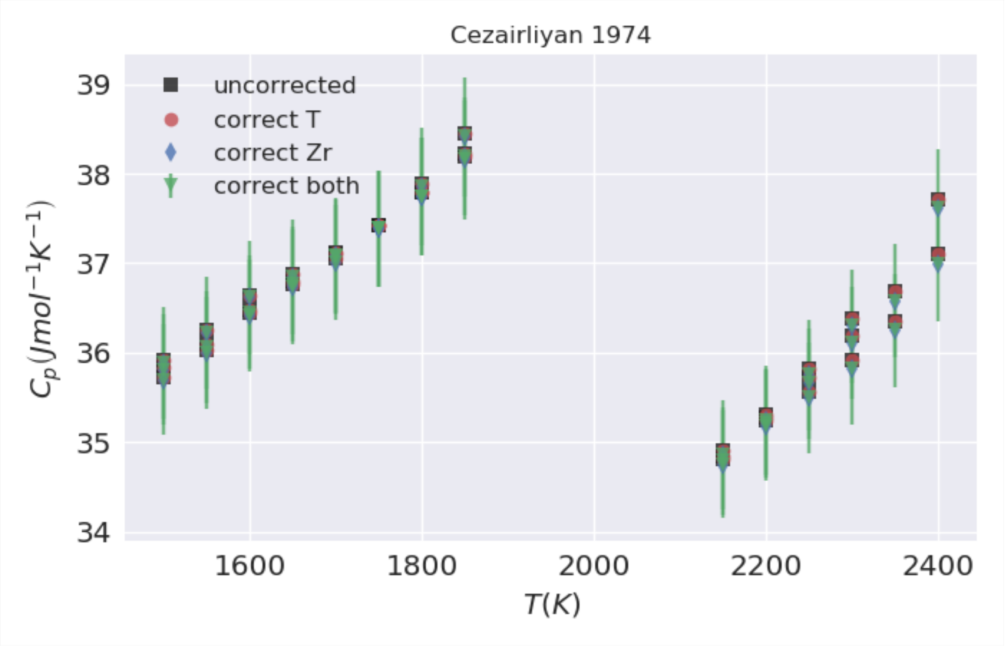




Ma 2014

CASE STUDY
HAFNIUM METAL

- experimental thermodynamic measurements of α, β and liquid phase Hf
- 20 total datasets obtained
  - 17 for α, 9 for β, 10 for liquid
  - 14 for C_p, 6 for H, 0 for S, 0 for G
- data corrected for temperature scale, Zr content
- reported error bars converted to standard errors (GUM)



Goldberg 1992, GUM 1995, Arblaster 2013, Arblaster 2014, Paulson 2018

CASE STUDY
MODEL SELECTION

Model	Log Marginal Likelihood	Bayes' Factor
α phase		
Einstein	-1744.2	~0
Debye	-1262.9	~0
Debye + Linear	-1072.6	~0
Debye + Quadratic	-813.2	~0
Debye + Cubic	-640.2	~0
Debye + Quartic	-623.1	1
Debye + Quintic	-627.4	1.4×10^{-2}
Debye + SR	-629.7	1.4×10^{-3}
β phase		
Constant	-534.2	~0
Linear	-511.1	3.0×10^{-3}
Quadratic	-505.3	1
Cubic	-518.5	1.9×10^{-6}
Liquid Phase		
Constant	-491.4	~0
Linear	-471.0	1
Quadratic	-476.0	6.7×10^{-3}

$$C_p^{Deb}(T,\theta_D) = 9R\left(\frac{T}{\theta_D}\right)^3 \int_0^{\theta_D/T} \frac{x^4 e^x}{(e^x-1)^2}dx$$

$$C_p^{\alpha}(T) = C_p^{Deb}(T,\theta_D)$$
$$+a_2T + a_3T^2 + a_4T^3 + a_5T^4$$

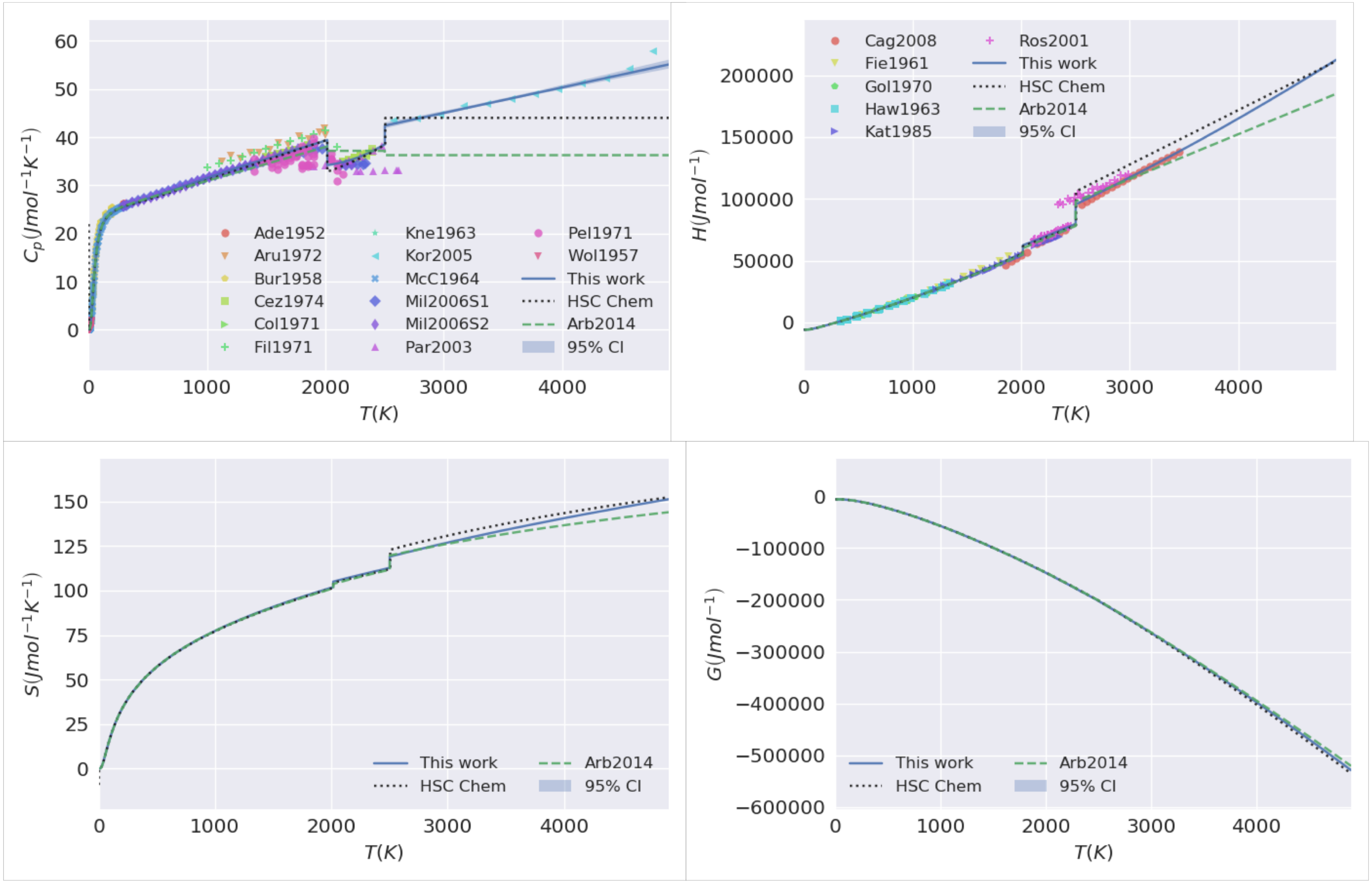
$$H^{\alpha}(T) - H^{\alpha}(298.15\mathrm{K}) = \int_0^T C_p^{Deb}(\tilde{T},\theta_D)d\tilde{T}$$
$$+a_2\frac{T^2}{2} + a_3\frac{T^3}{3} + a_4\frac{T^4}{4} + a_5\frac{T^5}{5}$$

$$C_p^{\beta}(T) = b_1 + b_2T + b_3T^2$$
$$H^{\beta}(T) = b_0 + b_1T + b_2\frac{T^2}{2} + b_3\frac{T^3}{3}$$

$$C_p^{liq}(T) = c_1 + c_2T$$
$$H^{liq}(T) = c_0 + c_1T + c_2\frac{T^2}{2}$$

Roslyakova 2016, Paulson 2018

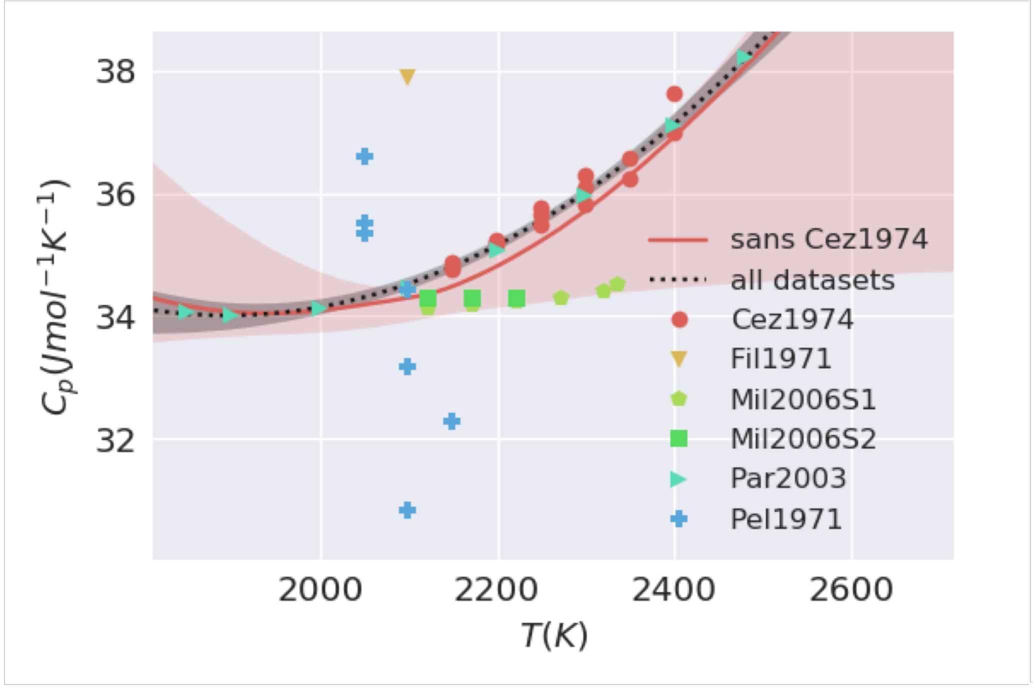
CASE STUDY **FINAL MODEL**

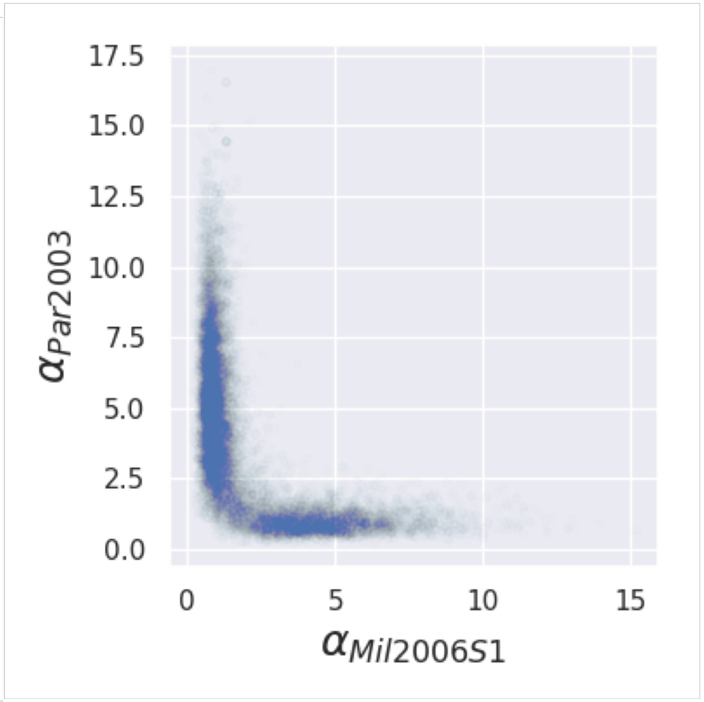


Arblaster 2014, Paulson 2018

CASE STUDY
FINAL MODEL

Investigate effect of removing individual datasets in Bayesian analysis:





Paulson 2018

CONCLUSIONS

- Comprehensive framework for model **selection**, **calibration** and **UQ** for thermodynamic property models

- Intuitive modifications address common issues
  - ensuring **thermodynamic consistency**
  - **automated weighting** of data sets

- Framework demonstrated in construction of models for C_p, H, S and G of Hf metal for α, β and liquid phases

ACKNOWLEDGEMENTS

This research was supported by:

CHIMaD

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

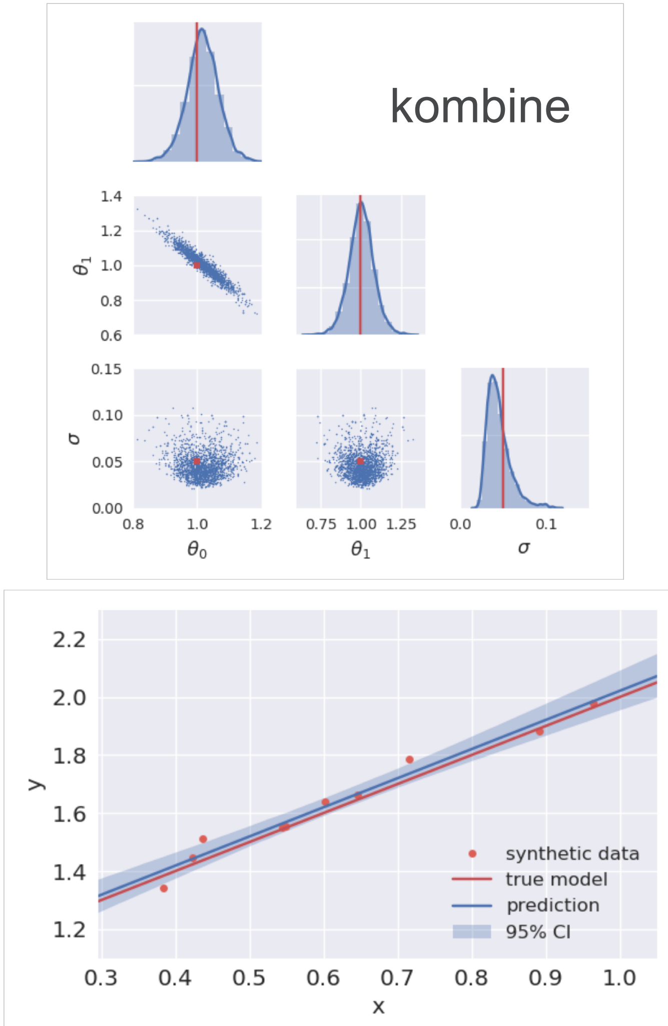
arXiv preprint:

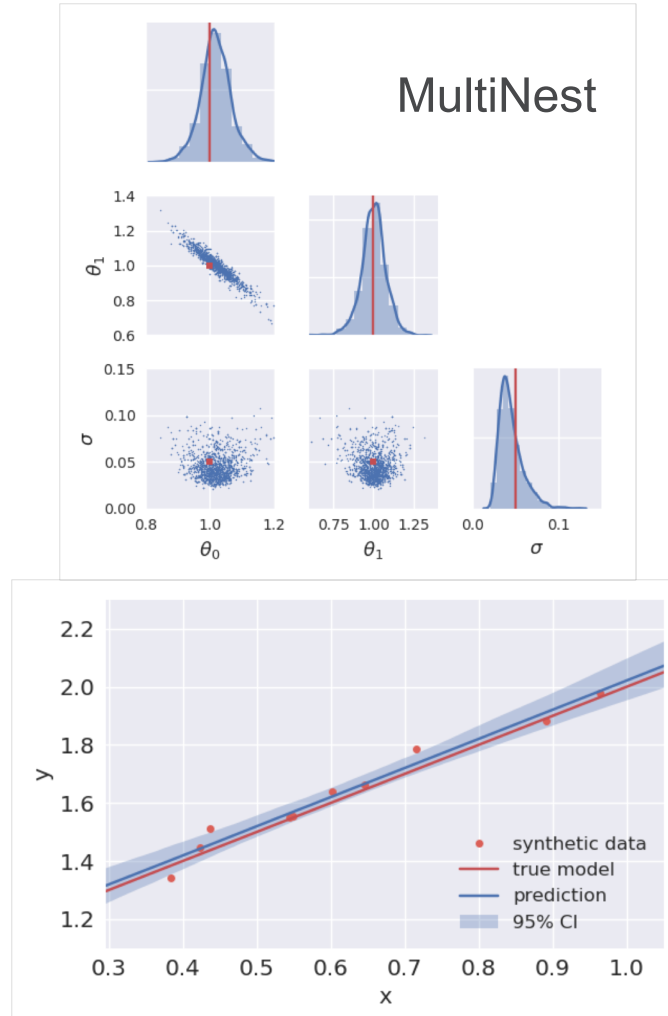
arXiv:1809.07365v

ABSTRACT

Models of the thermodynamic properties of materials form the basis for technological applications including the calculation of phase diagrams and simulation of microstructure evolution during processing – both of which play an important role in the design of materials for improved performance. Currently, the weighting of datasets, removal of outliers and the selection of model forms rely on expert judgements and do not provide uncertainty intervals. In this work we present a Bayesian framework for the selection, calibration and uncertainty quantification of thermodynamic property models. The framework is enabled by recent advances in numerical sampling methods. In addition, we present intuitive modifications that automatically weight datasets, improve robustness of outlier treatments, and ensure consistency of thermodynamically related models. We demonstrate the power of the approach through the construction of models for the specific heat, enthalpy, entropy and Gibbs free energy of Hafnium metal for the alpha, beta and liquid phases at temperatures ranging between 0 and 4900K.

BAYESIAN FRAMEWORK
FITTING A MODEL





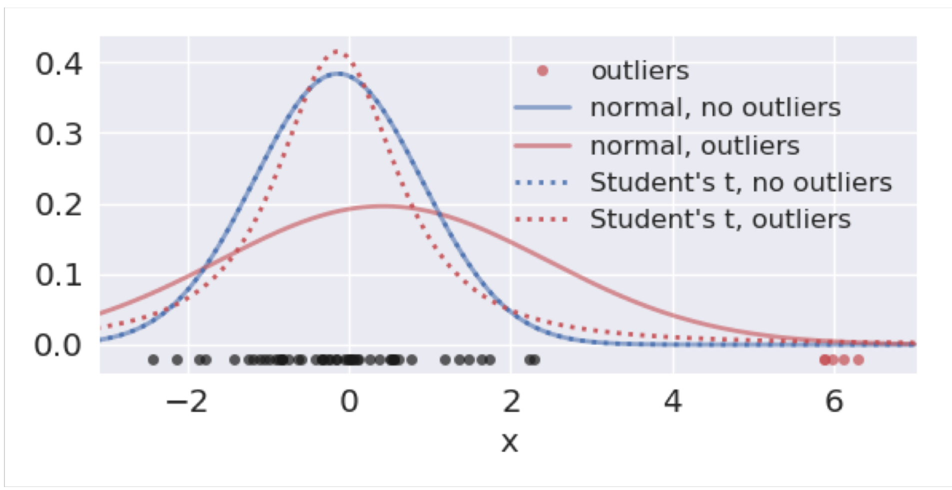
BAYESIAN FRAMEWORK
ROBUSTNESS TO OUTLIERS

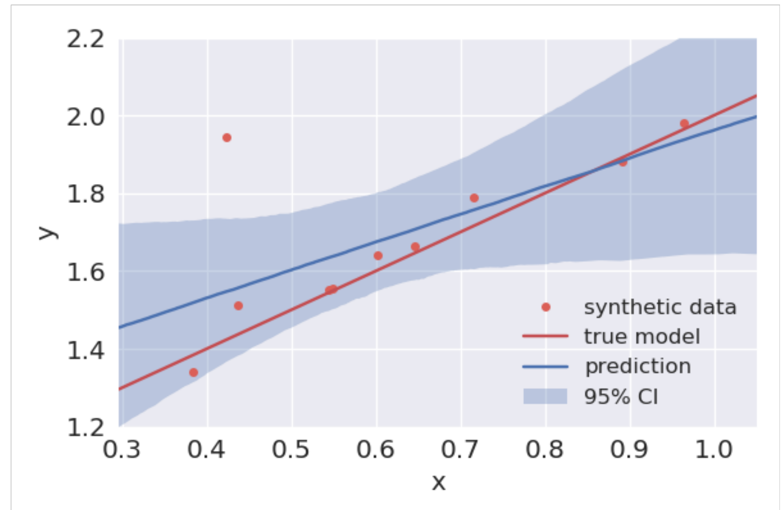
Likelihood:

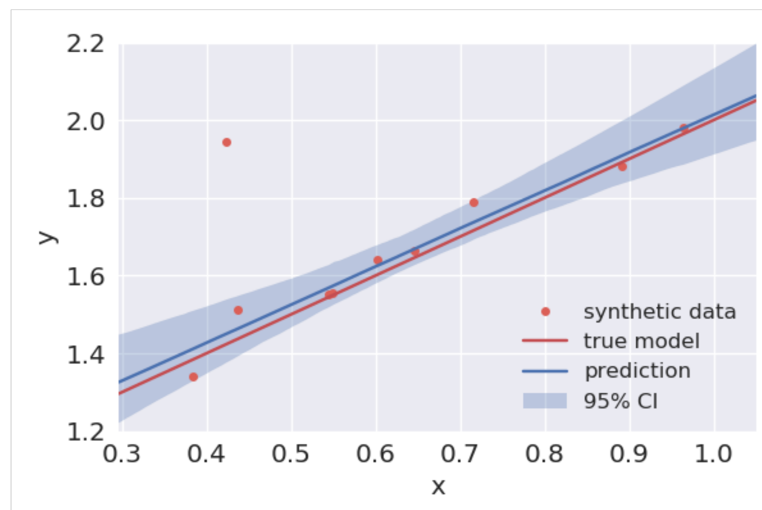
$$Pr\left(\mathbf{D}|\mathbf{\Theta},n,M\right)=\prod_{j}t\left(y_j|n,M\left(x_j,\mathbf{\Theta}\right),\varepsilon_j\right)$$

d.o.f.

Student's t

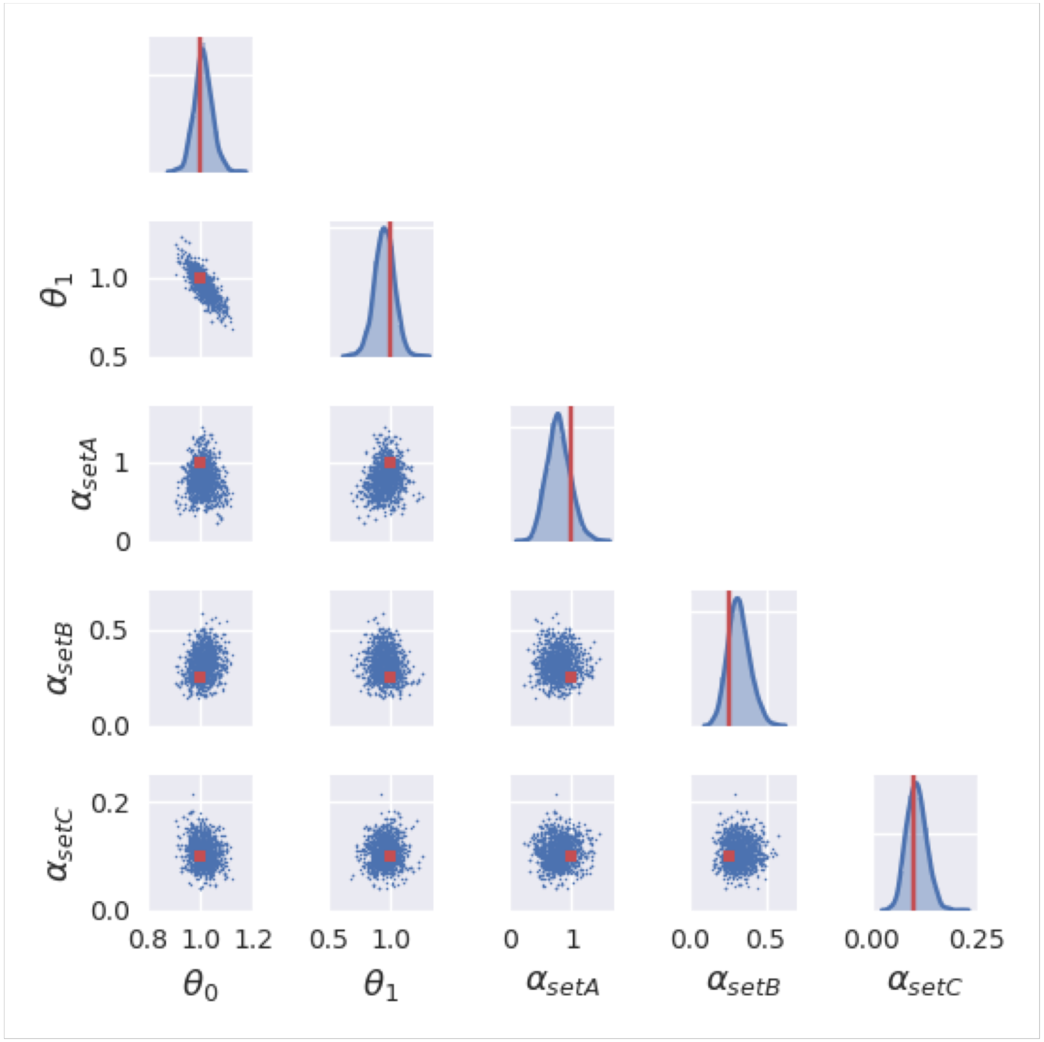






Gelman 2014

BAYESIAN FRAMEWORK
DATA WEIGHTING



BAYESIAN FRAMEWORK
THERMODYNAMIC CONSISTENCY

Models:

Einstein Model

$$C_p(T) = \frac{d}{dT} H(T) = \overbrace{3R\left(\theta/_T\right)^2 \frac{e^{\theta/_T}}{\left(\theta/_T - 1\right)^2}}^{\text{Einstein Model}} + aT + bT^2$$

$$H(T) - H(298.15K) = \frac{3R\theta}{e^{\theta/_T} - 1} + a\frac{T^2}{2} + b\frac{T^3}{3}$$

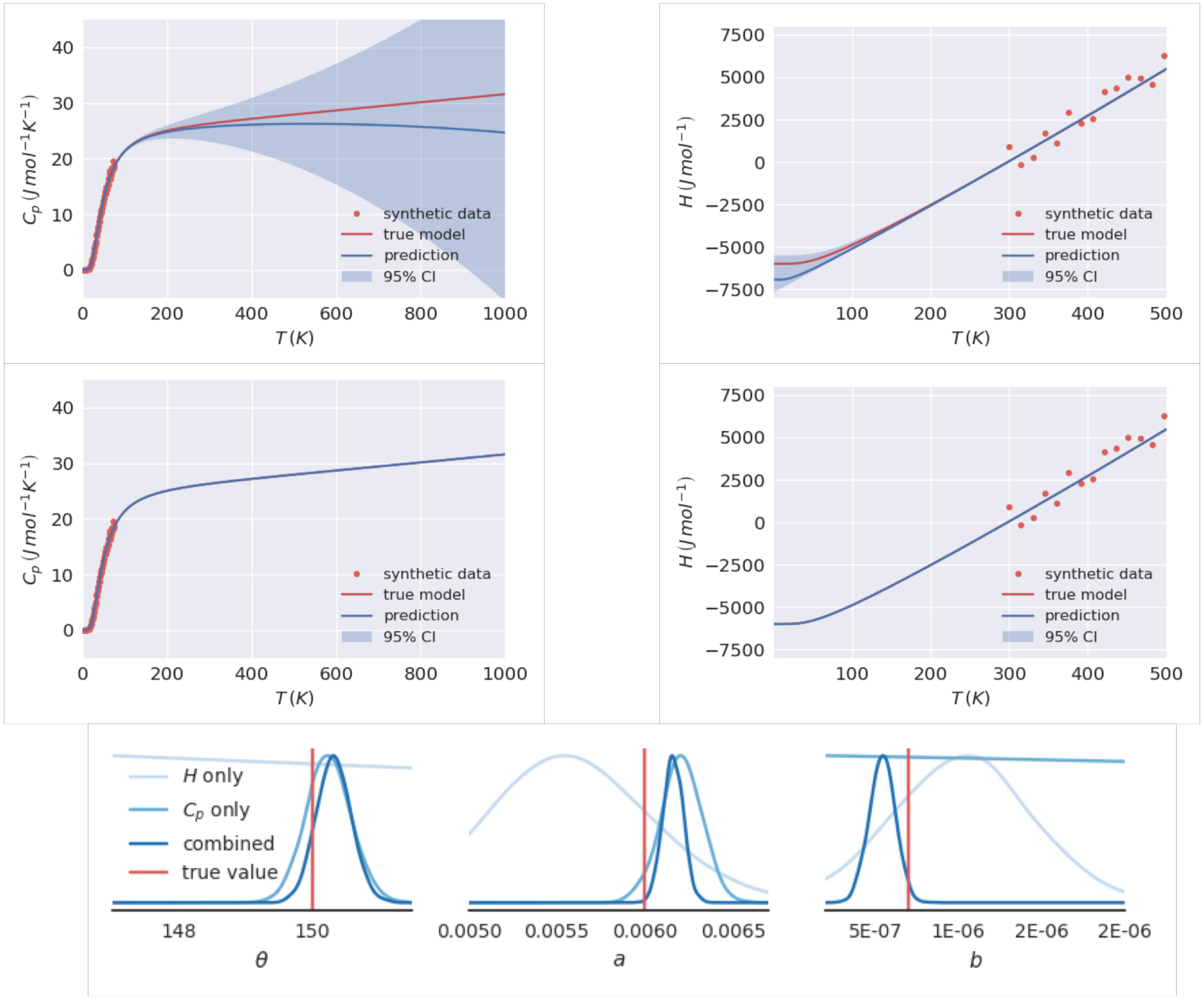
Priors:

$$Pr\left(\theta|M\right) = \mathcal{U}\left(\theta|145, 155\right) \quad Pr\left(a|M\right) = \mathcal{U}\left(a|0.003, 0.009\right) \quad Pr\left(b|M\right) = \mathcal{U}\left(b|-5\cdot 10^{-6}, 5\cdot 10^{-6}\right)$$

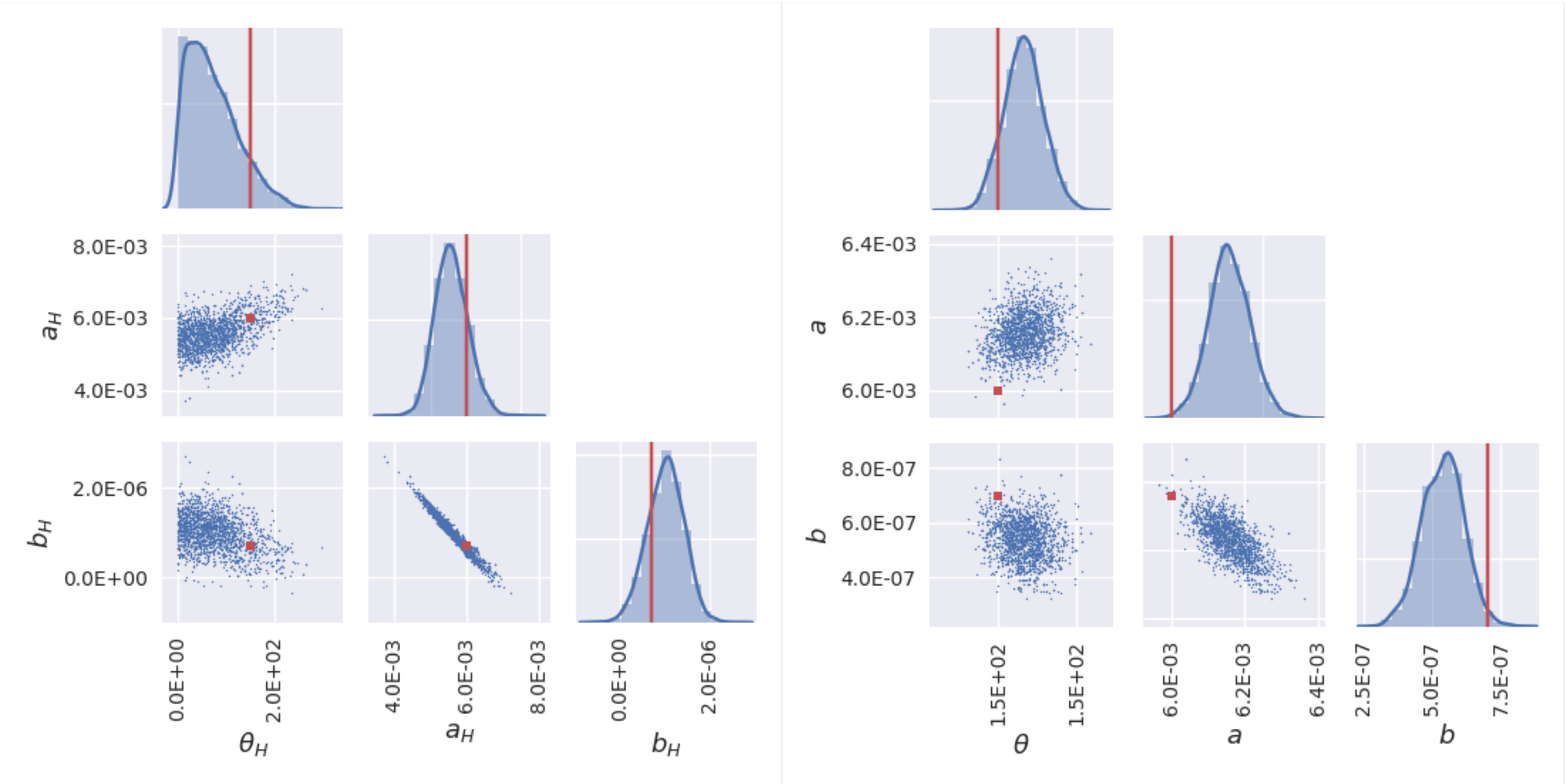
Likelihood: $$Pr\left(\mathbf{D}|\mathbf{\Theta}, M_H, M_{C_p}\right) = \prod_i \mathcal{N}\left(H_i|M_H\left(T_i, \mathbf{\Theta}\right), \varepsilon_i\right) \cdot \prod_j \mathcal{N}\left({C_p}_j|M_{C_p}\left(T_j, \mathbf{\Theta}\right), \varepsilon_j\right)$$

Chen 2001

BAYESIAN FRAMEWORK
THERMODYNAMIC CONSISTENCY



BAYESIAN FRAMEWORK
THERMODYNAMIC CONSISTENCY



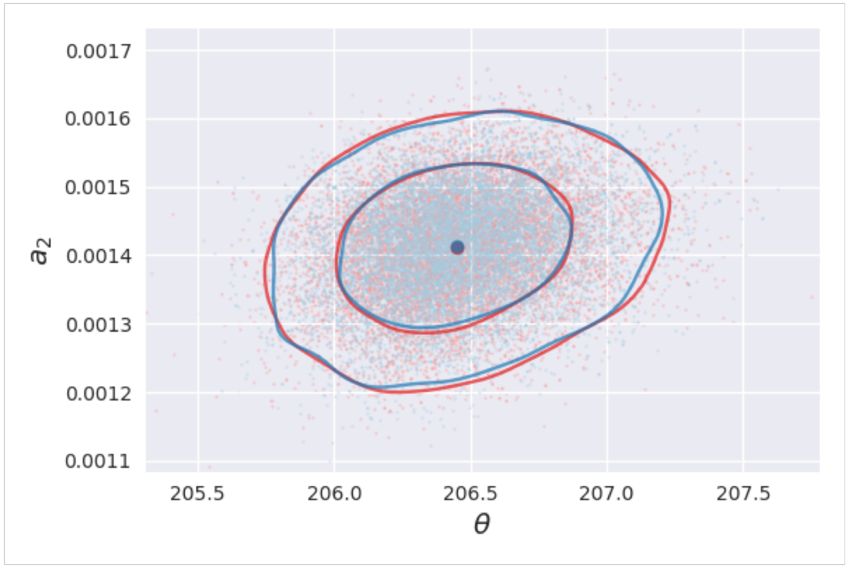
CASE STUDY
COMPUTATIONAL METHODOLOGY

Sampling:
pymultinest with 800 live points

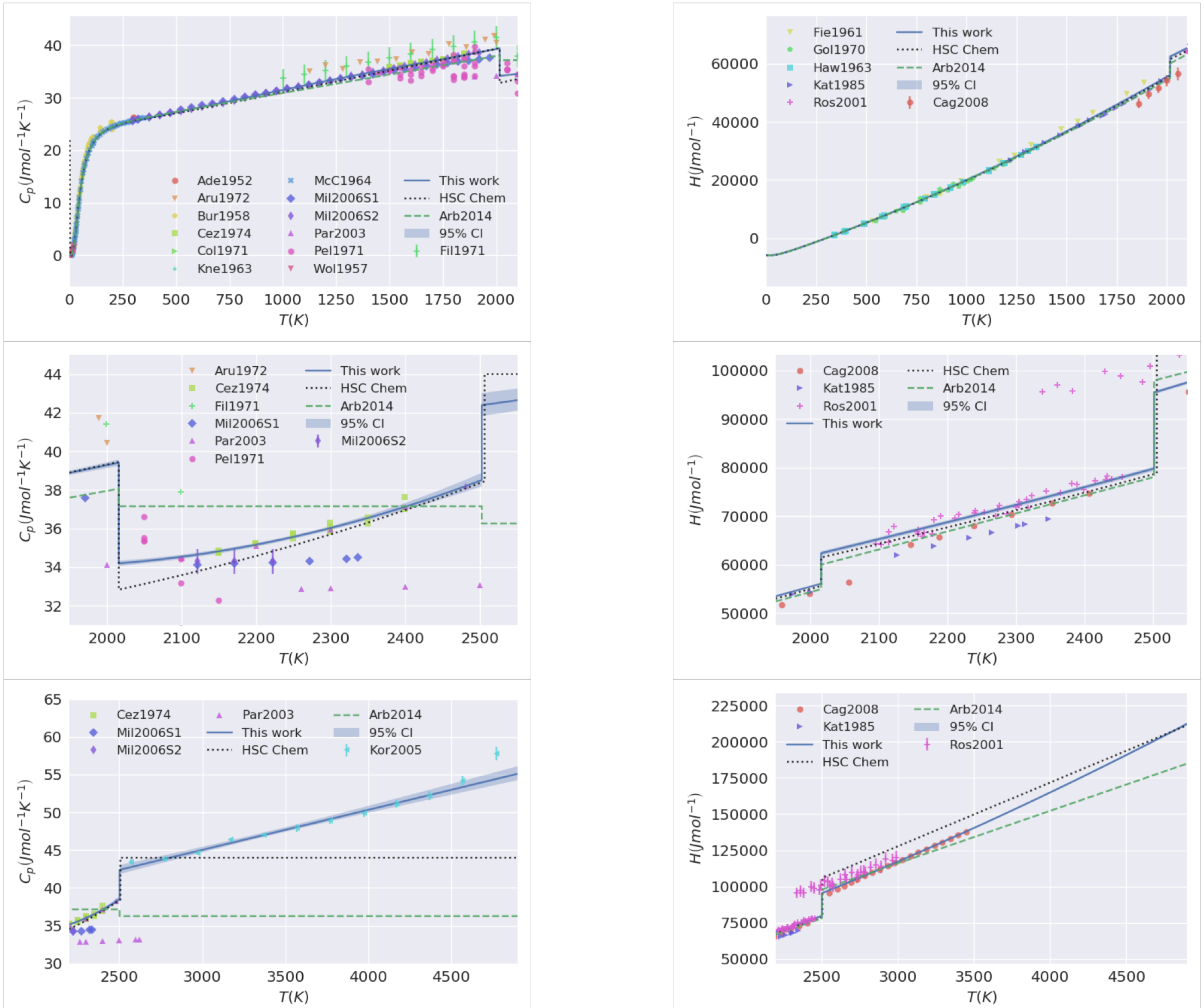
Likelihood:
- Student's t-distribution for robustness to outliers
- hyperparameters to rescale reported errors
- simultaneous regression for H and C_p

Prior Definition:
- Stage A: define broad uniform priors
- Stage B: narrow priors to 5-sigma Stage A – posterior

Model Selection:
marginal Likelihood

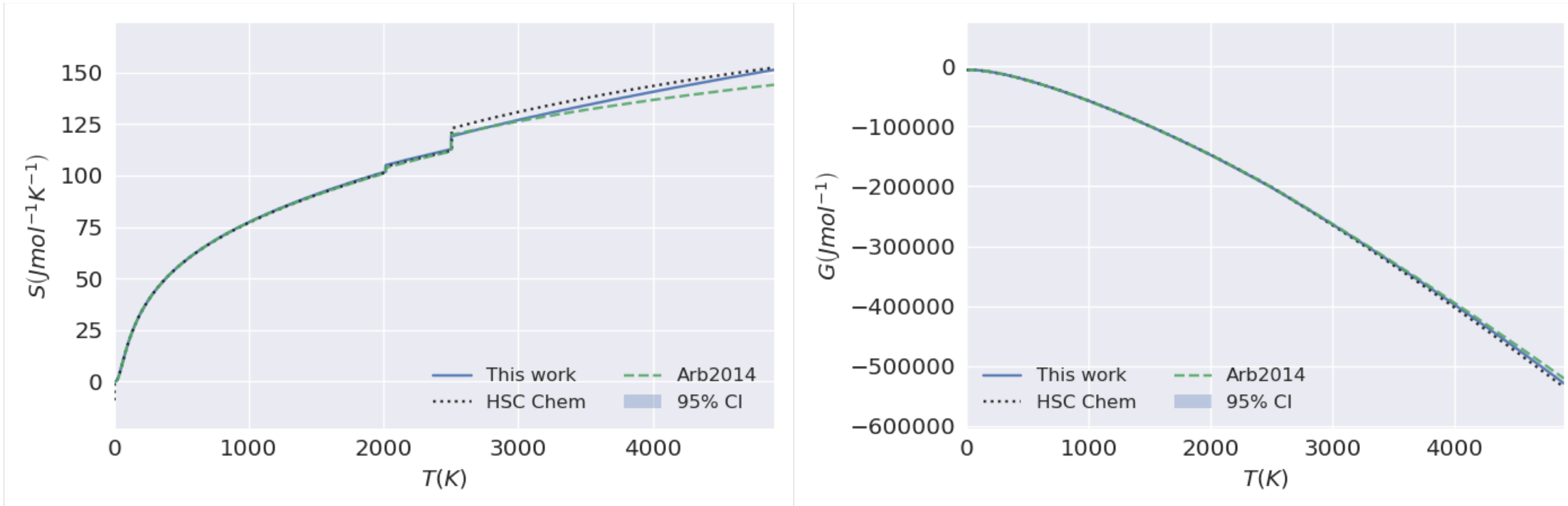


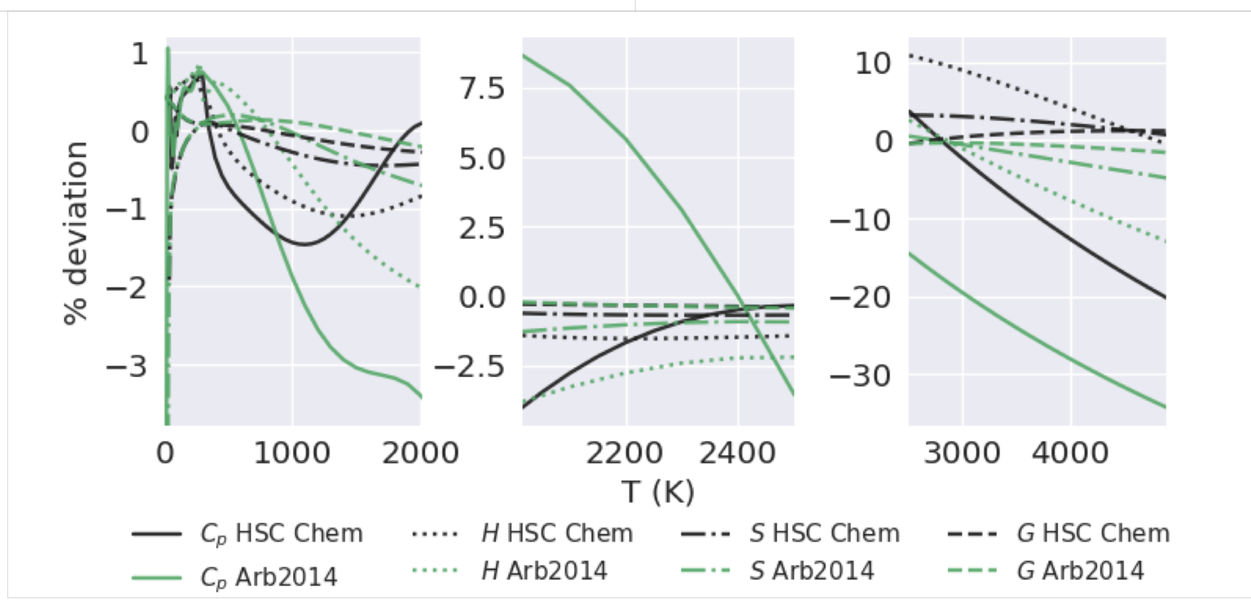
CASE STUDY FINAL MODEL



Arblaster 2014, Paulson 2018

CASE STUDY **FINAL MODEL**





Paulson 2018

GENERAL BAYESIAN REGRESSION

- Analytic Bayesian methods are difficult for non-linear models and non-conjugate priors
- *Markov chain Monte Carlo* (MCMC) methods can accurately sample posterior
- Simplest algorithm: Metropolis-Hastings (M-H) (1953).

Propose random step in parameter space from $\boldsymbol{z}^{(\tau)}$ to $\boldsymbol{z}^*$ according to proposal distribution $q_k\left(\boldsymbol{z}\middle|\boldsymbol{z}^{(\tau)}\right)$

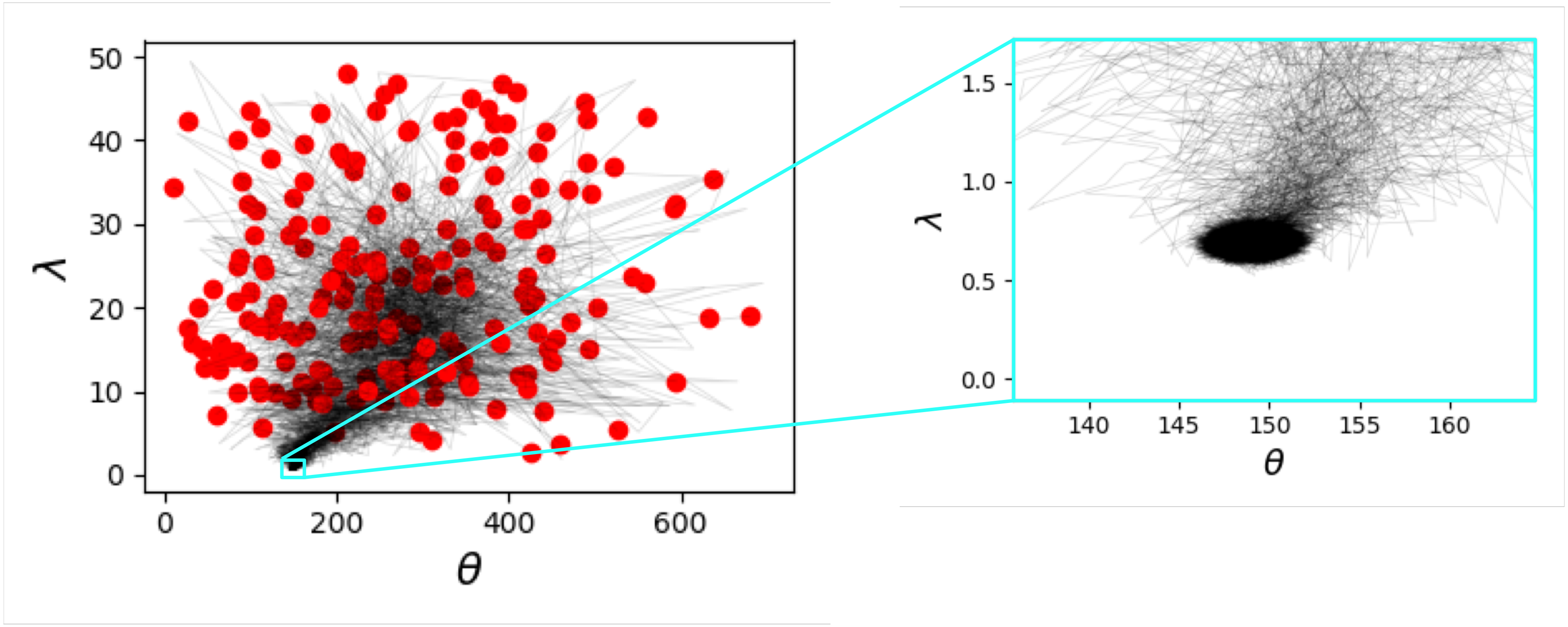
If $A_k\left(\boldsymbol{z}^*, \boldsymbol{z}^{(\tau)}\right) > \mathrm{U}[0,1]$ then $\boldsymbol{z}^{(\tau+1)} = \boldsymbol{z}^*$,
else $\boldsymbol{z}^{(\tau+1)} = \boldsymbol{z}^{(\tau)}$

$$A_k\left(\boldsymbol{z}^*, \boldsymbol{z}^{(\tau)}\right) = \frac{\tilde{p}(\boldsymbol{z}^*)}{\tilde{p}(\boldsymbol{z}^{(\tau)})}$$

- Interactive visualization: https://chi-feng.github.io/mcmc-demo/app.html#RandomWalkMH,standard

MCMC IN PYTHON: EMCEE

- Python implementation of *Affine Invariant Ensemble Sampler* (Goodman, 2010).
- *Affine Invariant*: addresses inefficiencies in MCMC sampling of posteriors with large covariances
- *Ensemble Sampler*: large set of *walkers* simultaneously explore posterior
  - Positions of other walkers make proposal distribution



BACKGROUND
LINEAR REGRESSION

Equation for linear model :

$$y(\mathbf{x}, \mathbf{w}) = \sum_{j=0}^{M-1} w_j \phi_j(\boldsymbol{x}) = \mathbf{w}^{\mathrm{T}} \boldsymbol{\phi}\,(\mathbf{x})$$

$\boldsymbol{x}$: feature vector of interest
$\boldsymbol{w}$: vector of weights controlling form of model
$\boldsymbol{\phi}(\mathbf{x})$: vector of basis functions, may include polynomial, harmonic, sigmoidal, Gaussian functions

LINEAR REGRESSION FREQUENTIST APPROACH

target variable: $t = y(\mathbf{x}, \mathbf{w}) + \varepsilon$

likelihood function: $p(\mathbf{t}|\mathbf{X}, \mathbf{w}, \beta) = \prod_{n=1}^{N} N\left(t_n \middle| \mathbf{w}^{\mathrm{T}} \boldsymbol{\phi}(\mathbf{x}_n), \beta^{-1}\right)$

$\boldsymbol{x}_n, t_n$: input vector, response for data point n
$\mathbf{t}$: vector of responses for all data points
$\mathbf{X}$: array of input vectors as follows; $\mathbf{X} \equiv [\mathbf{x}_1 \ldots \mathbf{x}_n \ldots \mathbf{x}_N]$

maximum likelihood estimate: $\mathbf{w}_{\mathrm{ML}} = \left(\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Phi}^{\mathrm{T}}$

$\boldsymbol{\Phi}$: data matrix

$$\boldsymbol{\Phi} = \begin{pmatrix} \phi_0(\mathrm{x}_1) & \cdots & \phi_{\mathrm{M}-1}(\mathrm{x}_1) \\ \vdots & \ddots & \vdots \\ \phi_0(\mathrm{x}_\mathrm{N}) & \cdots & \phi_{\mathrm{M}-1}(\mathrm{x}_\mathrm{N}) \end{pmatrix}$$