SAND2019-6650C

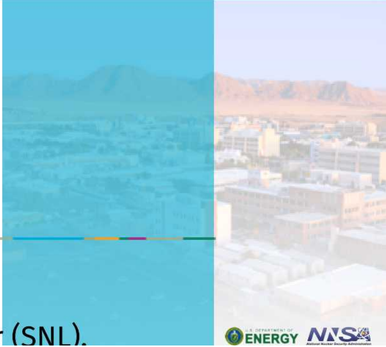
Dealing with nuisance parameters in Bayesian model calibration

PRESENTED BY

Kellin Rumsey[1] & Gabriel Huerta [1]

[1] Department of Statistical Sciences
Sandia National Laboratories

*co-authors: Lauren Hund, Derek Tucker (SNL).



ratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia,
neywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract

subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. SAND NO. 2019-4335 C

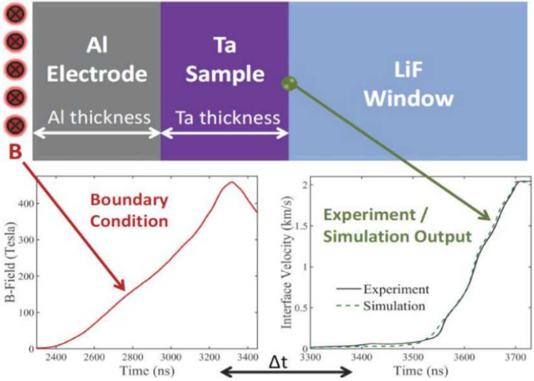
Statement

This work was supported by a Sandia National Laboratories Laboratory Directed Research and Development (LDRD) grant. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND 2019-

Background

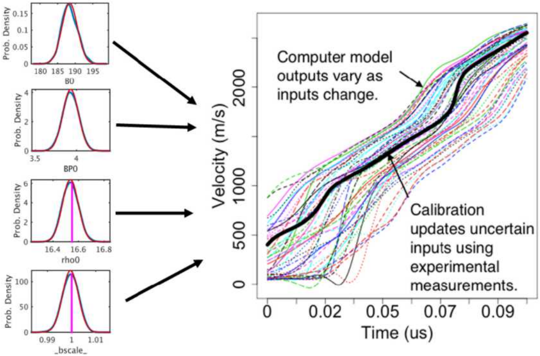
- **Dynamic material properties experiments**: access to the most extreme temperatures and pressures attainable.
- **Sandia National Labs Z-machine**: pulsed power driver that can deliver massive electrical currents over very short timescales (of the order of 60MA over $1\mu s$)).
- **Goal:** Understanding of material models at extreme conditions by coupling computational simulations with experimental data.
- **Parameters of interest are physical**: material properties with "true" value that is of interest.
- **Firstly**: Calibrate a well-understood model - two parameters of the equation of *state of tantalum*.

Experimental setup



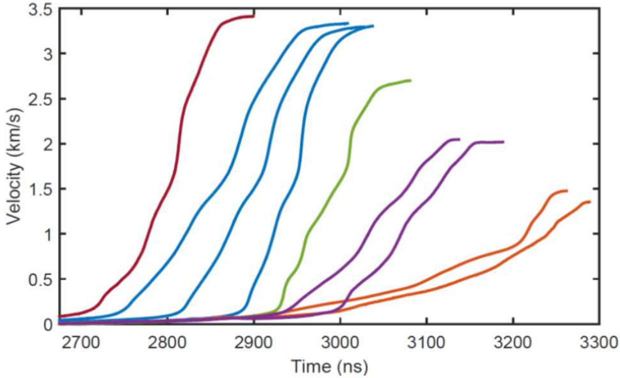
- "By coupling experimental and simulated velocity traces, parameters of the tantalum (Ta) equation of state (EOS) can be estimated".
- Massive electric currents treated as boundary conditions.
- Stress wave propagates thru system.

Calibration



- Uncertain inputs generate velocity curves using a computer model.
- Probability distributions look for ”agreement” of outputs and measurements.
- Bayesian framework is a natural in this context.

Challenge



- How to accurately estimate uncertainties?
- Calibration parameters have physical interpretation.
- Lots of *nuisance* parameters.

Our Framework

- Model the i^{th} observation in the j^{th} experiment as (Kennedy & O'Hagan 2001),

$$y(x_{ij}) = \eta(x_{ij}, \boldsymbol{\alpha}, \boldsymbol{\gamma}_j) + \delta(x_{ij}) + \epsilon_{ij}$$

- $\boldsymbol{\alpha}$ are the (unknown) values of the calibration parameters.
- $\boldsymbol{\gamma}_j$ unknown values of experimental uncertainties for experiment j.
- $y(x_{ij})$ is the observed velocity at time x_{ij}.
- $\eta(x_{ij}, \boldsymbol{\alpha}, \boldsymbol{\gamma}_j)$ is the computer model output at x_{ij}.
- $\delta(x_{ij}) \sim GP(\boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta)$
- ϵ_{ij} are errors at x_{ij}.

Dynamic material property calibration

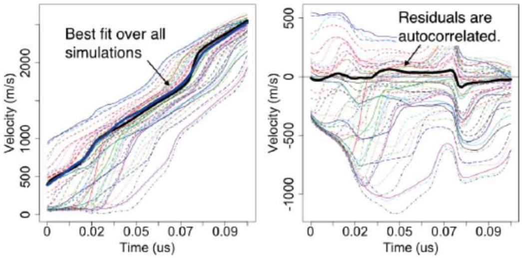
- BMC framework to obtain inference for two material properties of Tantalum.
- B_0 and B_0' are the Bulk modulus of tantalum and its pressure derivative.

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2) = (B_0, B_0')$$

- Four nuisance that may vary across $p = 9$ experiments
  - Tantalum density - γ_1
  - Magnetic field scaling - γ_{2j}, $j = 1, 2, \cdots 9$
  - Aluminum thickness- γ_{3j}, $j = 1, 2, \cdots 9$
  - Tantalum thickness - γ_{4j}, $j = 1, 2, \cdots 9$
- Potential for overfitting and lack of identifiability.

Issues

- Model can fit well to data, solutions far from *true* parameter values.
- Can we diagnose such overfitting? Can we mitigated it?
- **Model discrepancy** can reduce the identifiability of the calibration parameters and lead to systematic bias.



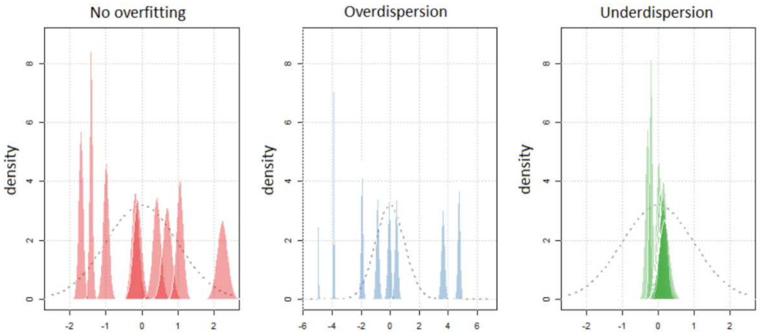
Nuisance parameters and overfitting

- **Aluminum and Tantalum thickness parameters:** These nuisance parameters are measured with a device which we believe to be well registered.
- Measurement error is *exclusive source* of uncertainty. The prior mean and variance of these nuisance parameters are well known.
- Nuisance parameters are standardized (mean 0, variance 1).
- The *standard informative (SI) prior* is:

$$(\gamma_{k1}, \gamma_{k2}, \cdots \gamma_{k9}) \sim N(0, I_9), \; k = 2, 3, 4.$$

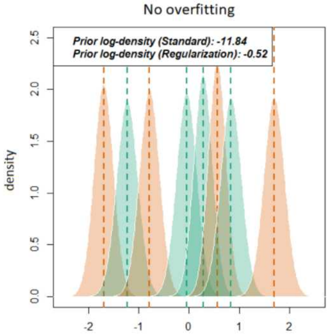
- "True values" are expected to look like a draw from a $N(0, I_9)$ distribution.

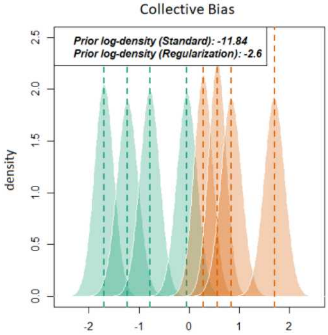
Nuisance parameters and overfitting



- Left panel: agrees with standard prior.
- Middle and left: can lead to overfitting.

Collective Bias for 2 nuisance-sets





- Left: No grouping occurs.
- Right: Collective bias implies systematic overfitting across experiments.
- Standard prior assigns same values.

A metric for overfitting

- We define,

$$M_\gamma = \frac{1}{p}\sum_{j=1}^{p}\gamma_j \qquad V_\gamma = \frac{1}{p-1}\sum_{j=1}^{p}\left(\gamma_j - M_\gamma\right)^2$$

- Prior beliefs about problem structure suggests:

$$M_\gamma \approx 0 \qquad V_\gamma \approx 1$$

- Under standard normal,

$$\pi_{M_\gamma, V_\gamma}(m, v) = N(m \mid 0, 1/p) \times \left[(p-1)\chi^2(v(p-1) \mid p-1)\right]$$

- Reasonable to check that the estimates $\hat{M}_\gamma$ and $\hat{V}_\gamma$ are *coherent* with prior.

A metric for overfitting

- **Definition:** We say that (m, v) is *more coherent with the prior* than (m', v') if

$$\pi_{M_\gamma, V_\gamma}(m, v) > \pi_{M_\gamma, V_\gamma}(m', v')$$

- Define the set of all points which are less coherent with the prior than $(\hat{M}_\gamma, \hat{V}_\gamma)$

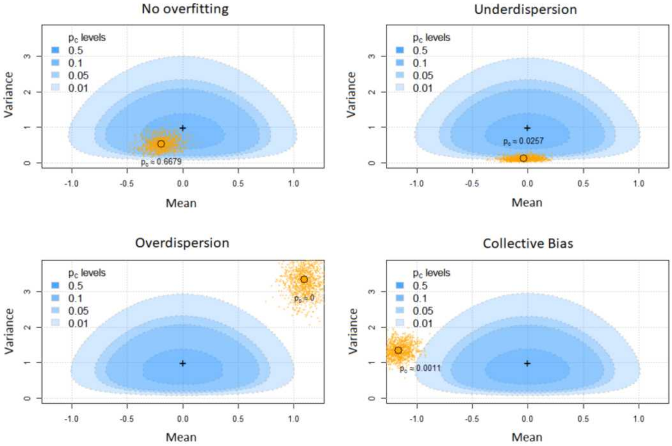
$$\Gamma_{\hat{M}_\gamma, \hat{V}_\gamma} = \left\{ (m, v) \mid \pi_{M_\gamma, V_\gamma}(\hat{M}_\gamma, \hat{V}_\gamma) > \pi_{M_\gamma, V_\gamma}(m, v) \right\}$$

- *Probability of prior coherency* of $(\hat{M}_\gamma, \hat{V}_\gamma)$

$$p_c(\hat{M}_\gamma, \hat{V}_\gamma) = \int_{\Gamma_{\hat{M}_\gamma, \hat{V}_\gamma}} \pi_{M_\gamma, V_\gamma}(m, v) \, dmdv$$

$$\approx \frac{1}{L} \sum_{\ell=1}^{L} \mathbb{1} \left(\pi_{M_\gamma, V_\gamma}(\hat{M}_\gamma, \hat{V}_\gamma) > \pi_{M_\gamma, V_\gamma}(m_\ell, v_\ell) \right)$$

Diagnostic plot for simulated case $p = 10$



- Orange: Point estimates and posterior draws of (M_γ, V_γ)
- Blue: Prior probability contours.

The moment penalization prior

- Overfitting of nuisance parameters leads to $(\hat{M}_\gamma, \hat{V}_\gamma)$ with low prior coherency.
- The *moment penalization (MP) prior* **penalizes** solutions with low prior coherency.
- Using a square loss function,

$$pen_\lambda(\boldsymbol{\gamma}) = \lambda_1(M_\gamma - 0)^2 + \lambda_2(V_\gamma - 1)^2$$

- Prior:

$$\pi_\gamma^{MP}(\boldsymbol{\gamma}) \propto exp\left(-pen_\lambda(\boldsymbol{\gamma})\right)$$

- Tries to encourage solutions with $M_\gamma \approx 0$ and $V_\gamma \approx 1$.

The moment penalization prior

- Therefore,

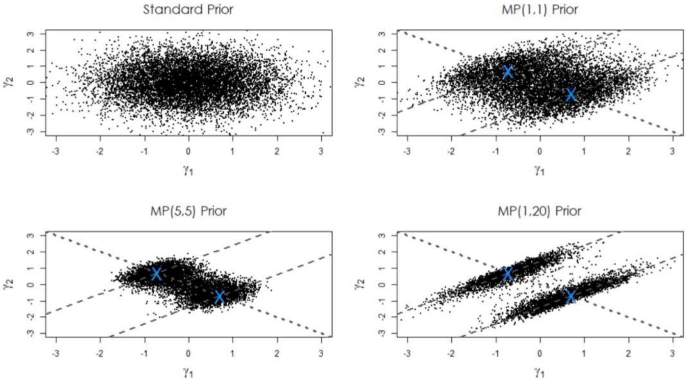
$$\pi_{\gamma}^{MP}(\boldsymbol{\gamma}) \propto \exp\left[-\lambda_1 M_{\gamma}^2\right] \; \exp\left[-\lambda_2 (V_{\gamma} - 1)^2\right]$$

- λ_1 and λ_2 control how strongly we want to enforce constraints.
- *Reparameterize:* $\omega_1 = 2Var(M_{\gamma})\lambda_1$ and $\omega_2 = 2Var(V_{\gamma})\lambda_2$
- Write $\boldsymbol{\gamma} \sim MP(\omega_1, \omega_2)$ to mean that,

$$\pi_{\gamma}^{MP}(\boldsymbol{\gamma}) \propto \exp\left[-\frac{p\omega_1}{2} M_{\gamma}^2\right] \; \exp\left[-\frac{(p-1)\omega_2}{4} (V_{\gamma} - 1)^2\right]$$

- $\boldsymbol{\gamma} \sim MP(1,1)$ is the *standard moment penalization prior.*

Samples from the Standard MP prior



- $10,000$ draws via M-H for $p = 2$.
- As $\omega_1, \omega_2 \to \infty$ all density is placed on $\pm(1/\sqrt{2}, -1/\sqrt{2})$ (Z-regularization).
- As p grows, the induced marginal priors become $N(0,1)$.

Example: The simple machine

- Brynjarsdottir and O'Hagan (2014): The simple machine delivers work

$$\zeta(x) = \frac{E\,x}{1 + x/20}$$

  - x is the amount of *effort* put into the machine.
  - E is the *efficiency* of the machine.
  - Denominator accounts for loss of work due to *friction*.
- The naive simulator introduces model discrepancy

$$\eta(x, E) = Ex$$

Example: The simple machine

- We consider $p = 10$ simple machines, and introduce base efficiency G_j as a machine-dependent nuisance parameter.
- Inputs $x_1, x_2, \cdots x_n$ evenly spaced over $[1, 4]$
- Data generating process:

$$y_{ij} = G_j + \frac{E\, x_i}{1 + x_i/20} + \epsilon_i$$
$$G_j \sim N(0, 0.05^2)$$
$$\epsilon_i \sim N(0, 0.01^2)$$

- Naive simulator:

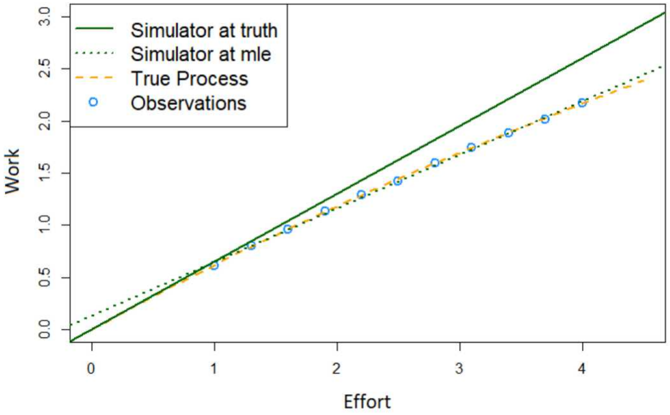
$$\eta(x, E, G) = G + E\, x$$

- True efficiency is $E = 0.65$. Standardize parameters:

$$\alpha = \frac{E - 0.65}{0.3} \sim N(0, 1) \qquad \gamma_k = \frac{G_k - 0}{0.05} \sim N(0, 1)$$

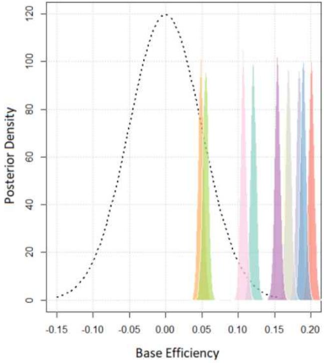
Example: The simple machine

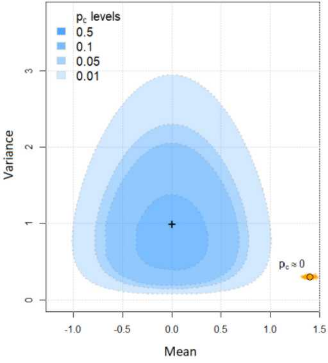
- Model discrepancy leads to *systematic bias.*



Example: The simple machine

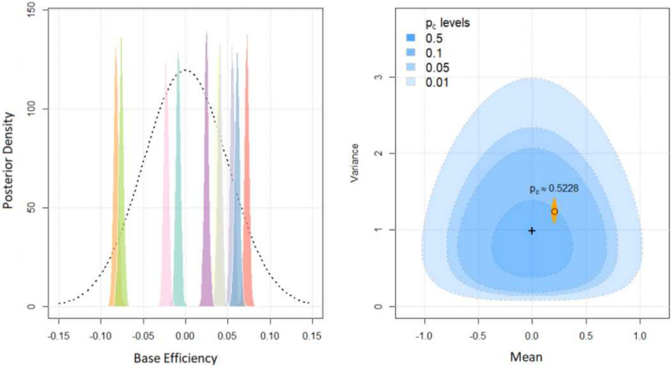
- Under standard informative prior



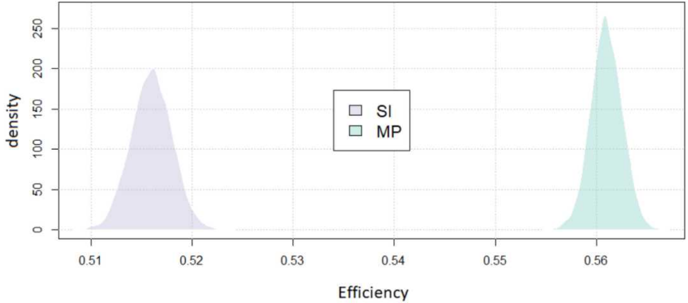


Example: The simple machine

- Under moment penalization prior



Example: The simple machine



- Posterior inference improves under MP, but is still far from truth.
- This is still valuable information! Model discrepancy is leading to biased inference on the parameter of interest.

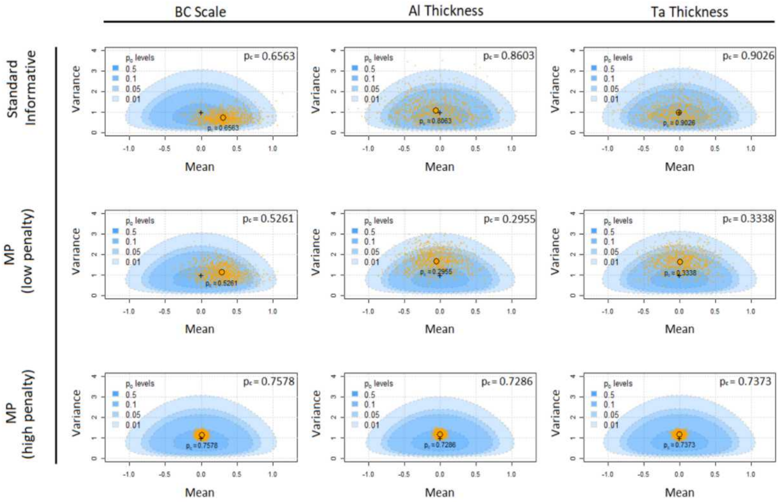
Dynamic material property calibration revisited

- Inference for two material properties of Tantalum.
- B_0 and B_0' are the Bulk modulus of tantalum and its pressure derivative.

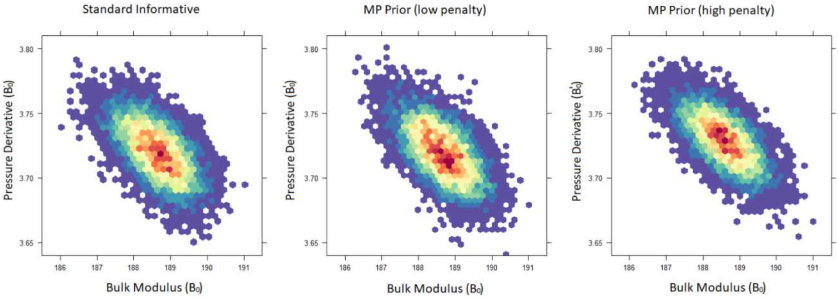
$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2) = (B_0, B_0')$$

- Four nuisance that may vary across $p = 9$ experiments
  - Tantalum density - γ_1
  - Magnetic field scaling - γ_{2j}, $j = 1, 2, \cdots 9$
  - Aluminum thickness- γ_{3j}, $j = 1, 2, \cdots 9$
  - Tantalum thickness - γ_{4j}, $j = 1, 2, \cdots 9$
- Perform BMC for SI, SMP and MP(20, 40) priors.

Dynamic material property calibration
Diagnostic plots



Dynamic material property calibration Physical parameter posteriors



- Similar posterior inference in all cases.
- Indicates that model discrepancy is unlikely to be causing bias in the parameters of interest.

Conclusions

- Overfitting of nuisance parameters leads to systematic bias which is often a symptom of model discrepancy.
- In complex high-dimensional problems, with appropriate problem structure, we can:
    - **Identify:** Probability of prior coherency identifies many types of overfitting, should it occur.
    - **Reduce:** The moment penalization prior reduces the systematic bias of the nuisance parameters.
    - **Diagnose:** Examine the sensitivity of posterior inference in order to diagnose the presence and effect of model discrepancy on the parameters of interest.

References

- Arendt, P.D., Apley, D.W. and Chen, W. (2016) A preposterior analysis to predict identifiability in the experimental calibration of computer models. IIE Trans.
- Bayarri, M. J., et al. (2007a) Computer model validation with functional output. Ann. Statist.,
- Bhattacharya, A., et al. (2015) Dirichlet Laplace Priors for Optimal Shrinkage, JASA.
- Brown, J. L., et al. (2014) Flow strength of antalum under ramp compression to 250 GPa. J. Appl. Phys
- Brown, J. L. and Hund, L. B. (2018) Model calibration via deformation. J.R. Statisti. Soc. C.
- Brynjarsdottir, J. and OHagan, A. (2014) Learning about physical parameters: the importance of model discrepancy. Invrs. Probl.
- Carvahlo, C., Polson, N. and Scott, J. (2010). The horseshoe estimator for sparse signals. Biometrica.
- Kennedy, M. C. and OHagan, A. (2001) Bayesian calibration of computer models (with discussion). J. R. Statist. Soc. B.