MANCHESTER 1824

The University of Manchester

M\cr NA
Manchester Numerical Analysis

MCMC for the Laser Flash Experiment

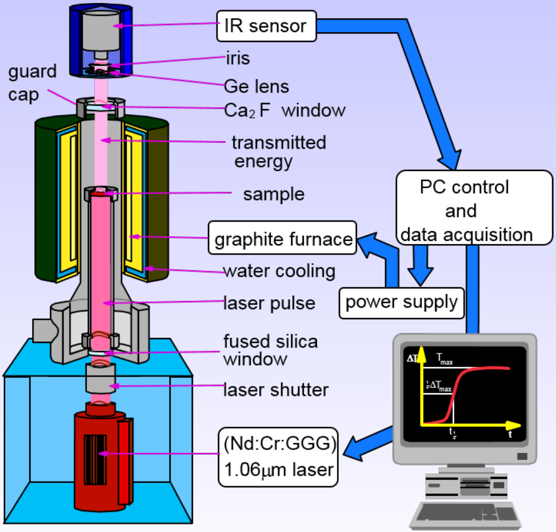
James Rynn
School of Mathematics
The University of Manchester

james.rynn@postgrad.manchester.ac.uk

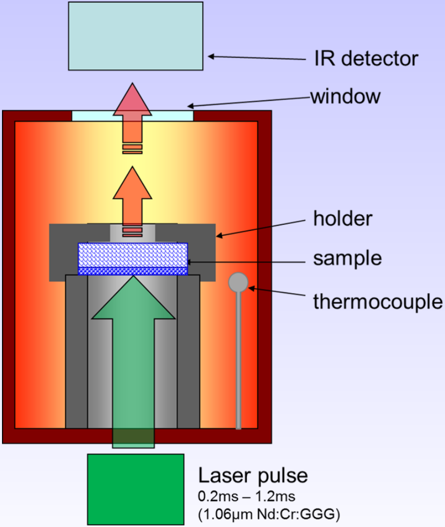
NA-UQ Reading Group, 09/09/16

Outline

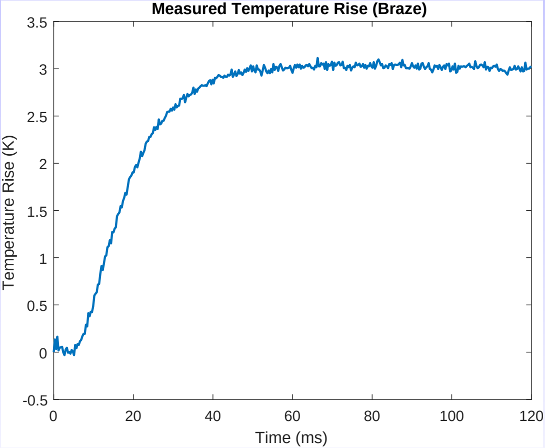
The Laser Flash Experiment



The Laser Flash Experiment



The Laser Flash Experiment



The Laser Flash Experiment

$$\rho c_p \frac{\partial u}{\partial t}(x,t) = \lambda \frac{\partial^2 u}{\partial x^2}(x,t) + Q(x,t), \qquad x \in [0,H],\ t \in [0,T],$$

where

- ρ = density,
- c_p = specific heat capacity,
- λ = thermal conductivity,

are scalar constants.
Forcing term $Q(x,t)$ (representing the laser flash) given by

$$Q(x,t) := \begin{cases} I, & x \in [0,x_f],\ t \in [0,t_f], \\ 0, & \text{otherwise}, \end{cases}$$

The Laser Flash Experiment

$$\frac{\partial u}{\partial t}(x,t) = \alpha \frac{\partial^2 u}{\partial x^2}(x,t) + \frac{Q(x,t)}{\rho c_p}, \qquad x \in [0,H],\ t \in [0,T],$$

where

$$\alpha := \frac{\lambda}{\rho c_p},$$

is the thermal diffusivity.

The Laser Flash Experiment

$$\frac{\partial u}{\partial t}(x,t) = \alpha \frac{\partial^2 u}{\partial x^2}(x,t) + \frac{Q(x,t)}{\rho c_p}, \qquad x \in [0,H],\ t \in [0,T],$$

where

$$\alpha := \frac{\lambda}{\rho c_p},$$

is the thermal diffusivity.

Initial Condition:

$$u(x,0) = T_a \quad \forall\, x \in [0,H].$$

$$\frac{\partial u}{\partial t}(x,t) = \alpha \frac{\partial^2 u}{\partial x^2}(x,t) + \frac{Q(x,t)}{\rho c_p}, \qquad x \in [0,H],\ t \in [0,T],$$

where

$$\alpha := \frac{\lambda}{\rho c_p},$$

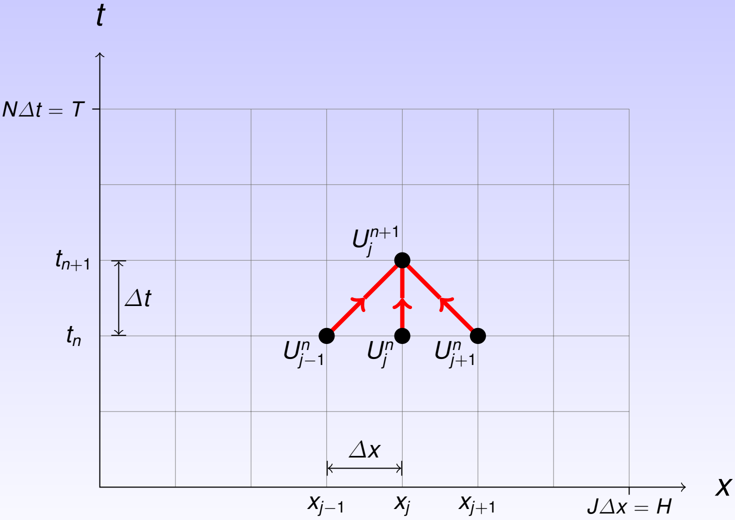
is the thermal diffusivity.

Initial Condition:

$$u(x,0) = T_a \quad \forall\, x \in [0,H].$$

Boundary Conditions: See later.

Finite Differences



Finite Differences

Explicit scheme:

$$U_j^{n+1} = U_j^n + \mu(U_{j+1}^n - 2U_j^n + U_{j-1}^n) + \frac{\Delta t Q_j^n}{\rho c_p},$$

for

$$j = 1, 2, \ldots, J-1, \quad n = 0, 1, \ldots, N-1,$$

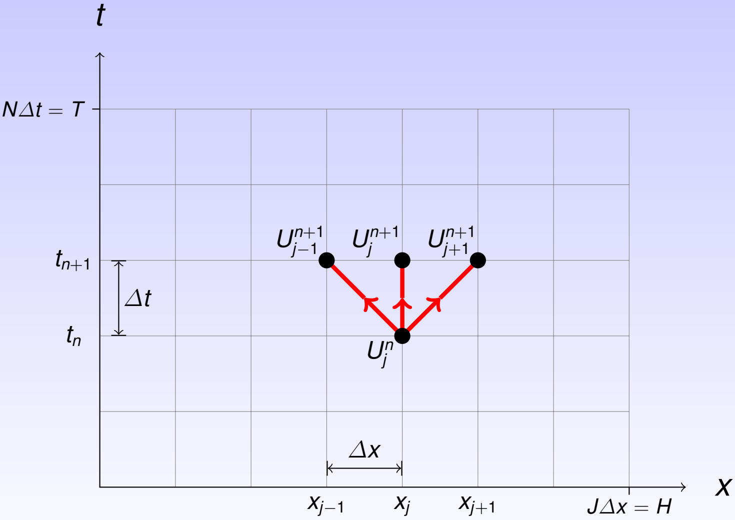
where

$$\mu := \frac{\alpha \Delta t}{(\Delta x)^2}.$$

Explicit method - stable if $\mu \leq \frac{1}{2}$.

What if we don't know α?

Finite Differences



Finite Differences

Use an implicit method. e.g., the Crank–Nicolson scheme:

$$
\begin{aligned}
-\frac{1}{2}\mu U_{j-1}^{n+1} + (1+\mu)U_j^{n+1} - \frac{1}{2}\mu U_{j+1}^{n+1} \\
= \frac{1}{2}\mu U_{j-1}^{n} + (1-\mu)U_j^{n} + \frac{1}{2}\mu U_{j+1}^{n} + \frac{\Delta t Q_j^n}{\rho c_p},
\end{aligned}
$$

for

$$
j = 1, 2, \ldots, J-1, \quad n = 0, 1, \ldots, N-1.
$$

Solve a tridiagonal system at each time step. Currently *under-determined*, we have $J-1$ equations in $J+1$ unknowns.

Stable for **all** values of μ.

Case 1: Uniform Material, Zero Heat Loss

Assume no heat is lost through the faces of the material. Then,

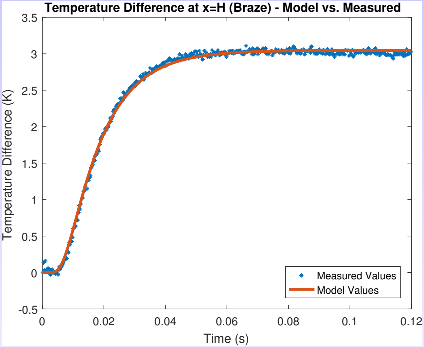
$$\frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(H, t) = 0 \quad \forall \ t \in [0, T].$$

Approximate boundary conditions to give two more equations, then linear system can be solved.

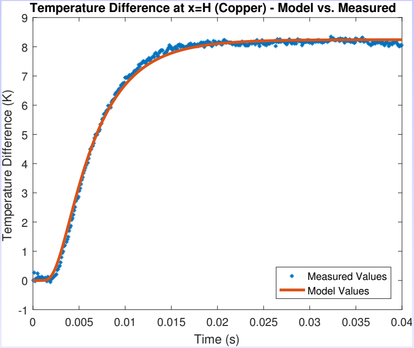
Optimise the values of λ and l using least squares fit to experimental data.
i.e., minimise:

$$(\boldsymbol{f}(\lambda, l) - \boldsymbol{y})^T(\boldsymbol{f}(\lambda, l) - \boldsymbol{y}),$$

Case 1: Uniform Material, Zero Heat Loss

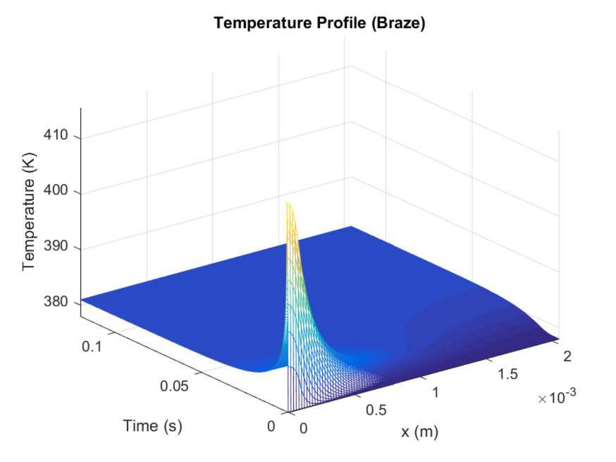


(a) Braze



(b) Copper

Case 1: Uniform Material, Zero Heat Loss



Verification: $T^N - T_a = \frac{t_f x_f I}{\rho c_p H^2}.$

Case 2: Uniform Material, Heat Flux BCs

Now assume that heat is lost at the faces through convection, then

$$\frac{\partial u}{\partial x}(0,t) = k(u(0,t) - T_a) \quad \forall\, t \in [0,T],$$

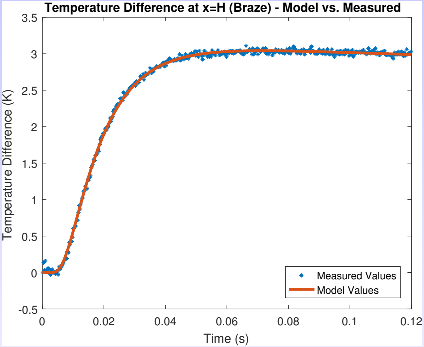
and

$$\frac{\partial u}{\partial x}(H,t) = -k(u(H,t) - T_a) \quad \forall\, t \in [0,T].$$

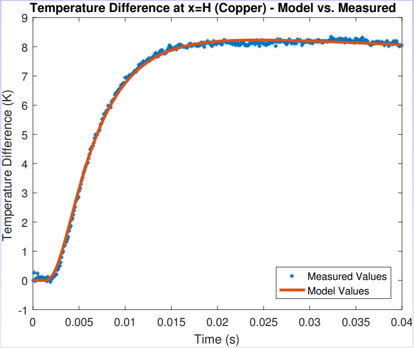
Again approximate boundary conditions to ‘complete’ our linear system.

New unknown k introduced.

Case 2: Uniform Material, Heat Flux BCs



(c) Braze



(d) Copper

Case 2: Uniform Material, Heat Flux BCs

The values found using the optimisation routine were

	λ ($Wm^{-1}K^{-1}$)	I (MWm^{-2})	k ($Wm^{-2}K^{-1}$)
Braze	113	738	14.1
Copper	363	3467	17.7

Case 3: Layered Material, Heat Flux BCs

Now suppose that the material consists of two layers, each with different material properties. So

$$\lambda(x) = \begin{cases} \lambda_1, & 0 \leq x \leq H_1, \\ \lambda_2, & H_1 < x \leq H_1 + H_2 =: H, \end{cases}$$

where

- λ_1 = thermal conductivity of braze,
- λ_2 = thermal conductivity of copper,
- H_1 = thickness of bottom (braze) layer,
- H_2 = thickness of top (copper) layer,
- $H := H_1 + H_2$ = sample thickness.

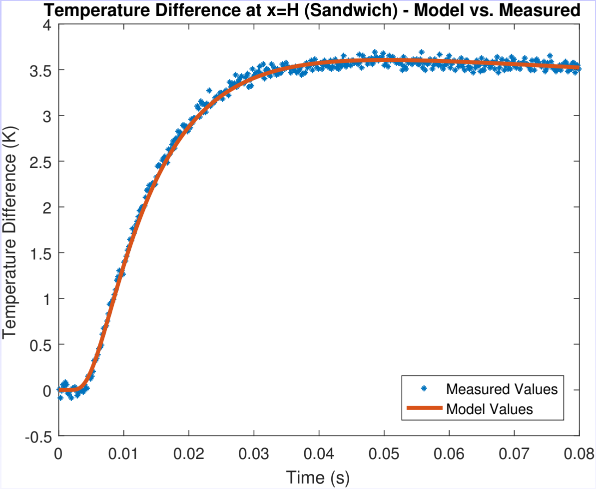
Similarly for the density ρ and specific heat capacity c_p.

Case 3: Layered Material, Heat Flux BCs

Must be careful to adjust CN method to account for the new form of λ (and therefore μ).

Optimisation using data from experiment 3 alone is an ill-posed problem. Must use data from all three experiments to optimise across all 6 unknowns simultaneously.

Case 3: Layered Material, Heat Flux BCs



Case 3: Layered Material, Heat Flux BCs

The values found using the optimisation routine were:

$$\lambda_1 = 115,\ \lambda_2 = 364,$$
$$l_1 = 744,\ l_2 = 3465,\ l_3 = 1412,$$
$$k = 17.6.$$

Problems with Optimisation

- In each experiment we've found similar but different values for each parameter. Which value is (more likely to be) correct?

- How sure of an answer can we be?

- What effect has the noise had on our answers? We haven't accounted for it in our model.

Uncertainty Quantification

We wish to quantify the uncertainty surrounding our solutions.

Do this by finding probability distributions for the uncertain parameters, given (indirect) observations of the data.

Achieved through sampling from the posterior distribution of the parameters given the observations using Markov chain Monte Carlo (MCMC) methods.

Metropolis–Hastings (MH) Algorithm

Used to generate a Markov chain with stationary density, π, that of the distribution we wish to sample from.

Main Idea:

1. assume the current state is $X_j = x$,
2. generate proposed value for the next state of the chain, $Y_{j+1} = y$ (in a clever way)
3. compute the ***acceptance probability***,

$$\alpha(x, y) := \min\left(\frac{\pi(y)\nu(y, x)}{\pi(x)\nu(x, y)}, 1\right),$$

4. accept the proposed value with probability α, or reject and stay at the current state,
5. repeat until enough states (samples) have been generated.

Acceptance Probability

Why choose $\alpha(x, y) = \min\left(\frac{\pi(y)\nu(y,x)}{\pi(x)\nu(x,y)}, 1\right)$?

We want α to represent a probability, so it must be bounded above by 1.

It can be shown that this choice of α results in a chain with stationary density π.

Key Observation: We only need to know π up to a constant of proportionality!

Random Walk Metropolis–Hastings Algorithm

Specific case of the MH algorithm, where the proposals Y_{j+1} are constructed as

$$Y_{j+1} = X_j + \epsilon_{j+1},$$

where the ϵ_j are chosen to be i.i.d. with a symmetric distribution.

We will use

$$\epsilon_j = \beta w,$$

where w is a draw from a Gaussian distribution with covariance matrix reflecting our prior beliefs and β is a parameter to be chosen to maximise the efficiency of the algorithm.

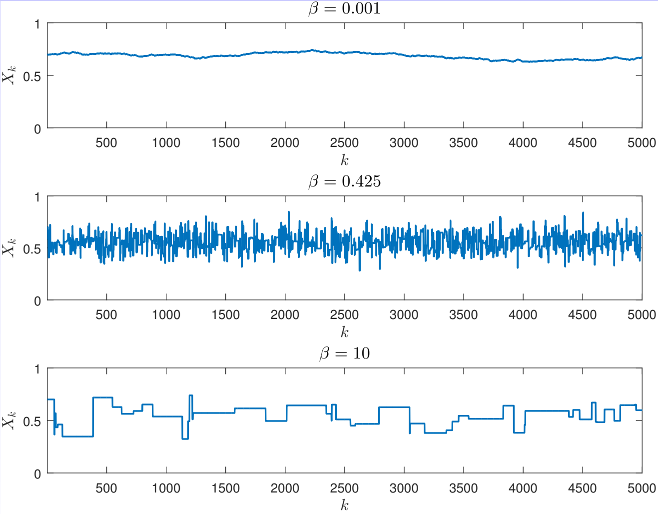
The choice of proposal variance β^2 is very important, the optimal value is problem dependent.

Random Walk Metropolis–Hastings Algorithm

Random Walk Metropolis–Hastings Algorithm:

```
Set initial state X_0
for j = 1, 2, 3, ..., N
    generate w
    let Y_j ← X_{j-1} + βw
    generate U_j ~ U[0, 1]
    if U_j ≤ α(X_{j-1}, Y_j)
        X_j ← Y_j
    else
        X_j ← X_{j-1}
    end if
    output X_j
end for
```

Optimal Value of β



Single Material

We know that λ is a positive scalar value. We therefore model this parameter as

$$\lambda = \exp(\theta_1),$$

where

$$\theta_1 \sim \mathcal{N}(m_1, \gamma_1^2),$$

for some $m_1, \gamma_1 \in \mathbb{R}^+$. That is, we model λ as a *log-normal* random variable.

Similarly we take

$$l = \exp(\theta_2) \qquad k = \exp(\theta_3),$$

for $\theta_2 \sim \mathcal{N}(m_2, \gamma_2^2)$ and $\theta_2 \sim \mathcal{N}(m_2, \gamma_2^2)$.

Assume that λ, l and k are independent.

Bayesian Inverse Problem Setup

We suppose that our data is of the form

$$\boldsymbol{y} = \mathcal{G}(\theta_1, \theta_2, \theta_3) + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \Sigma), \qquad (\star)$$

where

- $\boldsymbol{y}$ is a vector of observed temperatures,

$$\boldsymbol{y} = [u(H, t_{n_0}), u(H, t_{n_1}), \ldots, u(H, t_{n_{400}})]^T \in \mathbb{R}^{401},$$

- $\mathcal{G}$ is the forward operator, mapping values of the unknowns into data. In this case giving the temperature at $x = H$ at the measurement times t_{n_i},
- η is a vector of mean zero Gaussian noise.

Likelihood

Suppose that the random variable $\boldsymbol{y}|\theta_1, \theta_2, \theta_3$ is given by the relation ($\star$), then we have that it has likelihood (Lebesgue density) $L(\boldsymbol{y}|\theta_1, \theta_2, \theta_3)$ satisfying

$$L(\boldsymbol{y}|\theta_1, \theta_2, \theta_3) \propto \exp\left(-\frac{1}{2}\|\boldsymbol{y} - \mathcal{G}(\theta_1, \theta_2, \theta_3)\|_{\Sigma}^2\right),$$

where

$$\|\boldsymbol{x}\|_{\Sigma} := \boldsymbol{x}^T \Sigma^{-1} \boldsymbol{x}, \quad \forall\ \boldsymbol{x} \in \mathbb{R}^{401}$$

is the *covariance weighted norm*.
This is often expressed using the *potential*

$$\Phi(\boldsymbol{y}; \theta_1, \theta_2, \theta_3) := \frac{1}{2}\|\boldsymbol{y} - \mathcal{G}(\theta_1, \theta_2, \theta_3)\|_{\Sigma}^2.$$

Bayes' Theorem

Theorem (Bayes' Theorem)

Assume that

$$Z := \int_{\mathbb{R}^3} L(\boldsymbol{y}|\theta_1, \theta_2, \theta_3)\pi_0(\theta_1, \theta_2, \theta_3) d\theta_1 d\theta_2 d\theta_3 > 0.$$

Then, $\theta_1, \theta_2, \theta_3|\boldsymbol{y}$ *is a random variable with (Lebesgue) density* $\pi(\theta_1, \theta_2, \theta_3|\boldsymbol{y})$ *given by*

$$\pi(\theta_1, \theta_2, \theta_3|\boldsymbol{y}) = \frac{1}{Z} L(\boldsymbol{y}|\theta_1, \theta_2, \theta_3)\pi_0(\theta_1, \theta_2, \theta_3).$$

Bayes' Theorem

Theorem (Bayes' Theorem)

Assume that

$$Z := \int_{\mathbb{R}^3} L(\boldsymbol{y}|\theta_1, \theta_2, \theta_3)\pi_0(\theta_1, \theta_2, \theta_3)d\theta_1 d\theta_2 d\theta_3 > 0.$$

Then, $\theta_1, \theta_2, \theta_3|\boldsymbol{y}$ *is a random variable with (Lebesgue) density* $\pi(\theta_1, \theta_2, \theta_3|\boldsymbol{y})$ *given by*

$$\pi(\theta_1, \theta_2, \theta_3|\boldsymbol{y}) = \frac{1}{Z}L(\boldsymbol{y}|\theta_1, \theta_2, \theta_3)\pi_0(\theta_1, \theta_2, \theta_3).$$

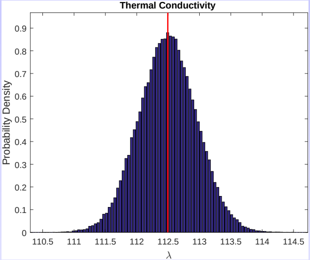
- $\pi(\theta_1, \theta_2, \theta_3|\boldsymbol{y})$ is called the ***posterior density***.

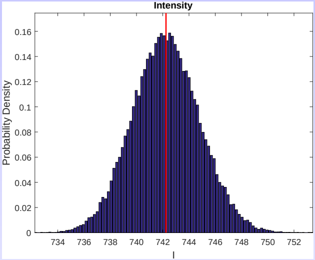
Posterior

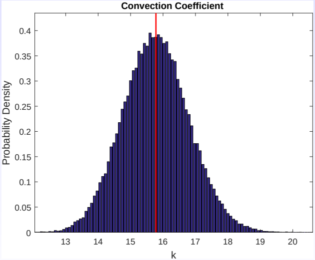
So, by independence,

$$\begin{aligned}\pi(\theta_1, \theta_2, \theta_3 | \boldsymbol{y}) &\propto L(\boldsymbol{y} | \theta_1, \theta_2, \theta_3) \pi_0(\theta_1, \theta_2, \theta_3) \\ &\propto \exp\left(-\frac{1}{2\sigma^2} \| \boldsymbol{y} - \mathcal{G}(\theta_1, \theta_2, \theta_3) \|_2^2 \right) \\ &\quad \times \mathcal{N}(\theta_1; m_1, \gamma_1^2) \cdot \mathcal{N}(\theta_2; m_2, \gamma_2^2) \cdot \mathcal{N}(\theta_3; m_3, \gamma_3^2)\end{aligned}$$

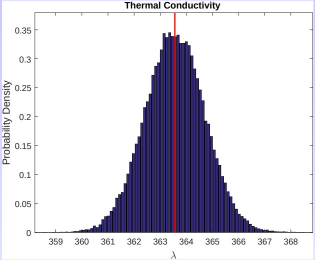
Braze

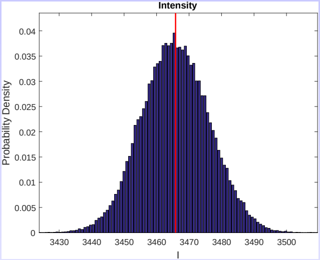

Thermal Conductivity
Probability Density
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
110.5
111
111.5
112
112.5
113
113.5
114
114.5
λ

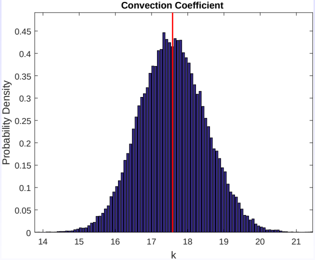

Intensity
Probability Density
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
734
736
738
740
742
744
746
748
750
752
I


Convection Coefficient
Probability Density
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
13
14
15
16
17
18
19
20
k

Copper


Thermal Conductivity
Probability Density
0 0.05 0.1 0.15 0.2 0.25 0.3 0.35
359 360 361 362 363 364 365 366 367 368
λ


Intensity
Probability Density
0 0.005 0.01 0.015 0.02 0.025 0.03 0.035 0.04
3430 3440 3450 3460 3470 3480 3490 3500
I


Convection Coefficient
Probability Density
0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45
14 15 16 17 18 19 20 21
k

Braze

	Optimisation	RWMH Mean	RWMH Variance
λ	113	112	0.214
l	738	742	6.52
k	14.1	15.8	1.04

Copper

	Optimisation	RWMH Mean	RWMH Variance
λ	363	364	1.31
l	3467	3466	108
k	17.7	17.6	0.840

Dual Layered Material

As with the optimisation, we must use all three sets of data simultaneously.
We investigate the posterior of the 6 unknowns given the three data sets.
Let

$$\begin{aligned}
\lambda_1 &= \exp(\theta_1),\\
\lambda_2 &= \exp(\theta_2),\\
l_1 &= \exp(\theta_3),\\
l_2 &= \exp(\theta_4),\\
l_3 &= \exp(\theta_5),\\
k &= \exp(\theta_6),
\end{aligned}$$

and

$$\boldsymbol{\theta} := (\theta_1, \theta_2, \ldots, \theta_6).$$

Bayesian Inverse Problem

In this case,

$$\boldsymbol{y} = \mathcal{G}(\boldsymbol{\theta}) + \boldsymbol{\eta}, \qquad \boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{0}, \Sigma)$$

where

$$\begin{aligned}\boldsymbol{y} = \big[u(H, t_{n_0^{(1)}}), \ldots, u(H, t_{n_{400}^{(1)}}),\\ u(H, t_{n_0^{(2)}}), \ldots, u(H, t_{n_{400}^{(2)}}),\\ u(H, t_{n_0^{(3)}}), \ldots, u(H, t_{n_{400}^{(3)}})\big]^T \in \mathbb{R}^{1203},\end{aligned}$$

and

$$\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_1^2, \sigma_2^2, \ldots, \sigma_2^2, \sigma_3^2, \ldots, \sigma_3^2)$$

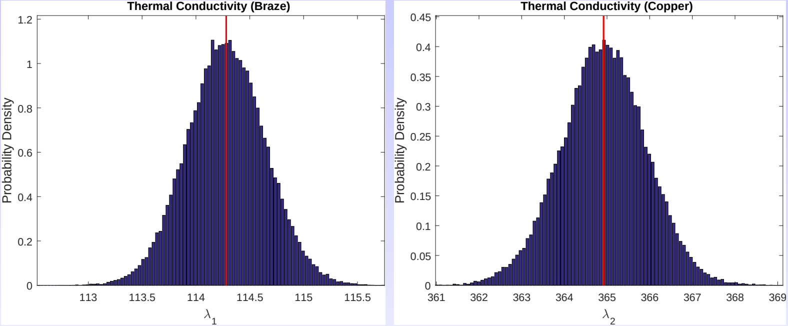
$$\implies \|\boldsymbol{x}\|_\Sigma^2 = \sigma_1^2 \sum_{i=1}^{i=401} x_i^2 + \sigma_2^2 \sum_{i=402}^{i=803} x_i^2 + \sigma_3^2 \sum_{i=804}^{i=1203} x_i^2 \quad \forall\, \boldsymbol{x} \in \mathbb{R}^{1203}.$$

Posterior

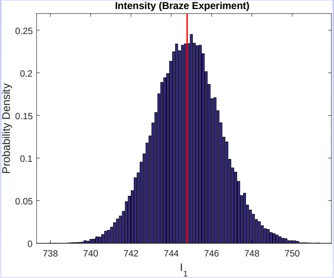
Again using independence,

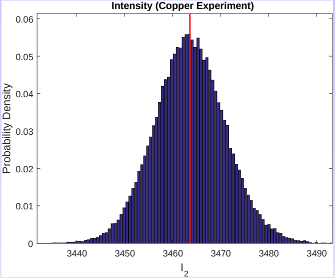
$$\begin{aligned}\pi(\boldsymbol{\theta}|\boldsymbol{y}) &\propto L(\boldsymbol{y}|\boldsymbol{\theta})\pi_0(\boldsymbol{\theta}) \\ &\propto \exp\left(-\frac{1}{2}\|\boldsymbol{y} - \mathcal{G}(\boldsymbol{\theta})\|_{\Sigma}^2\right)\prod_{i=1}^{6}\mathcal{N}(\theta_i; m_i, \gamma_i^2)\end{aligned}$$

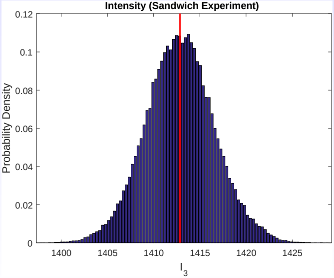
λ_1, λ_2

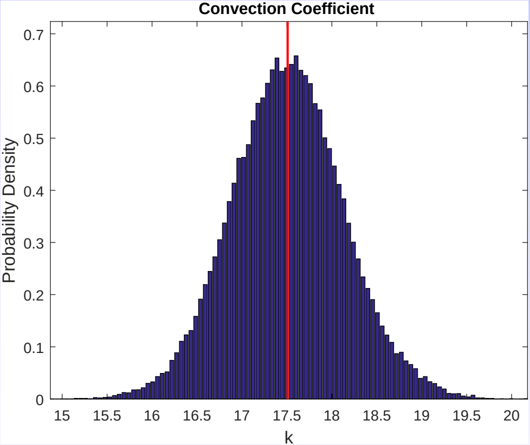


I_1, I_2, I_3


Intensity (Braze Experiment)
Probability Density
0.25
0.2
0.15
0.1
0.05
0
738
740
742
744
746
748
750
I1


Intensity (Copper Experiment)
Probability Density
0.06
0.05
0.04
0.03
0.02
0.01
0
3440
3450
3460
3470
3480
3490
I2


Intensity (Sandwich Experiment)
Probability Density
0.12
0.1
0.08
0.06
0.04
0.02
0
1400
1405
1410
1415
1420
1425
I3


Convection Coefficient
Probability Density
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
15
15.5
16
16.5
17
17.5
18
18.5
19
19.5
20
k

Braze

	Optimisation	RWMH Mean	RWMH Variance
λ_1	115	114	0.133
λ_2	364	365	0.993
l_1	744	745	2.83
l_2	3465	3463	54.5
l_3	1412	1413	13.9
k	17.6	17.5	0.392

Future Work

- Investigate mesh width sensitivity,
- Use data from experiments 1 and 3 to replicate results from experiment 2 (which are unavailable in real life),

Thank you for listening.