Nano-second Bayesian inference

Goal: Type I Error proof-by-simulation

Tools: Bounds + INLA + GPUs

Me: T. Ben Thompson
Confirm Solutions team: Michael Sklar, James Yang
October 12th, 2022

Confirm Solutions

Goal: Solve complex trial design:

1. Rigorous Type I Error proof-by-simulation.
2. Faster regulatory approval for wider classes of designs.

What we do:

1. Exciting *theory*: Continuous simulation extension.
2. Exciting *computation*: Trillions of simulations. ← me!

Confirm Solutions

What we really do

Allow YOU to do more cool Bayesian trial designs!

Continuous simulation extension

Proof by simulation

Question: “How many different combinations of parameters to explore?”

Answer: All of them.

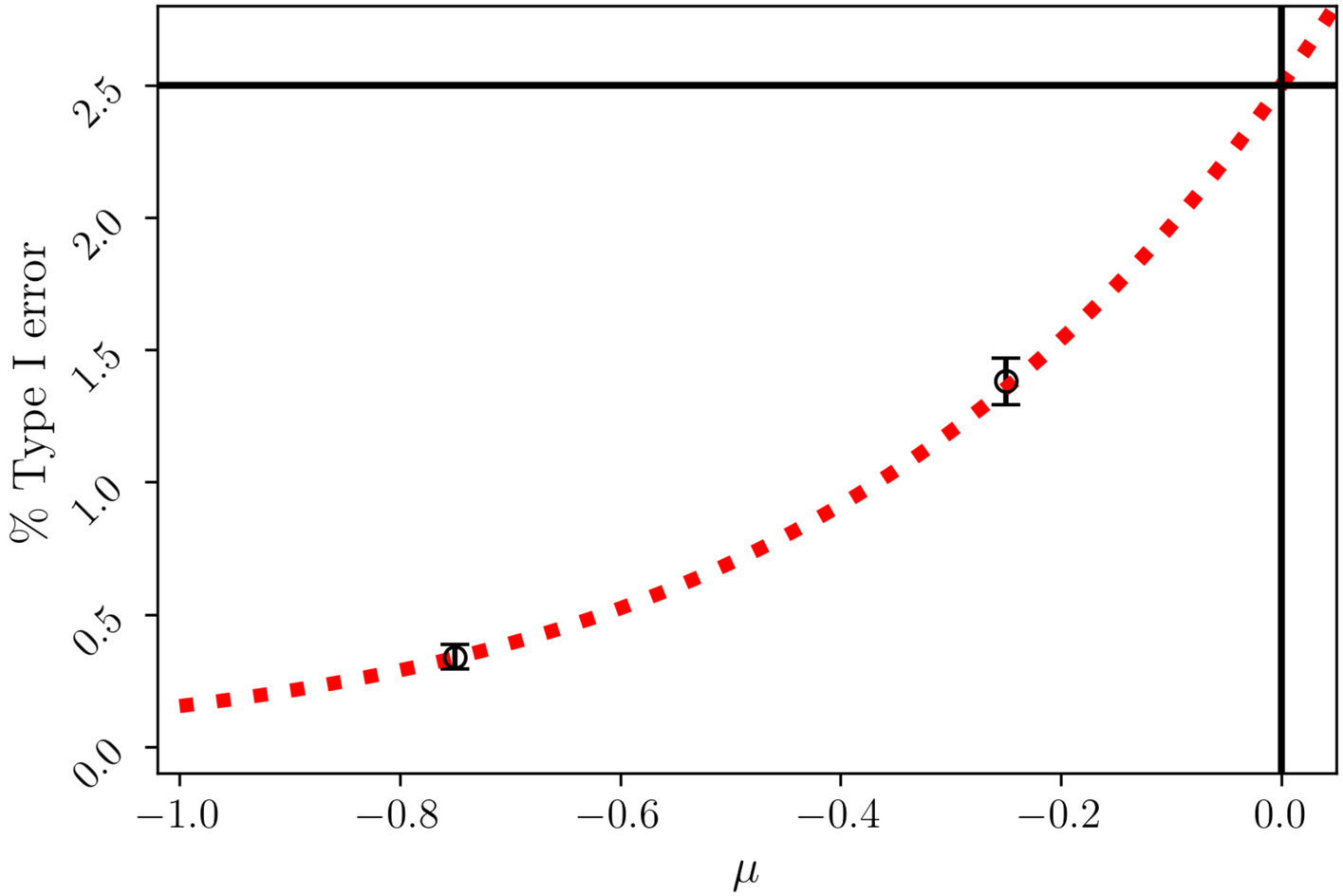
Upper bounds on Type I error

One-sided z-test

$X \sim \mathcal{N}(\mu, 1)$

$H_0 : \mu \leq 0$

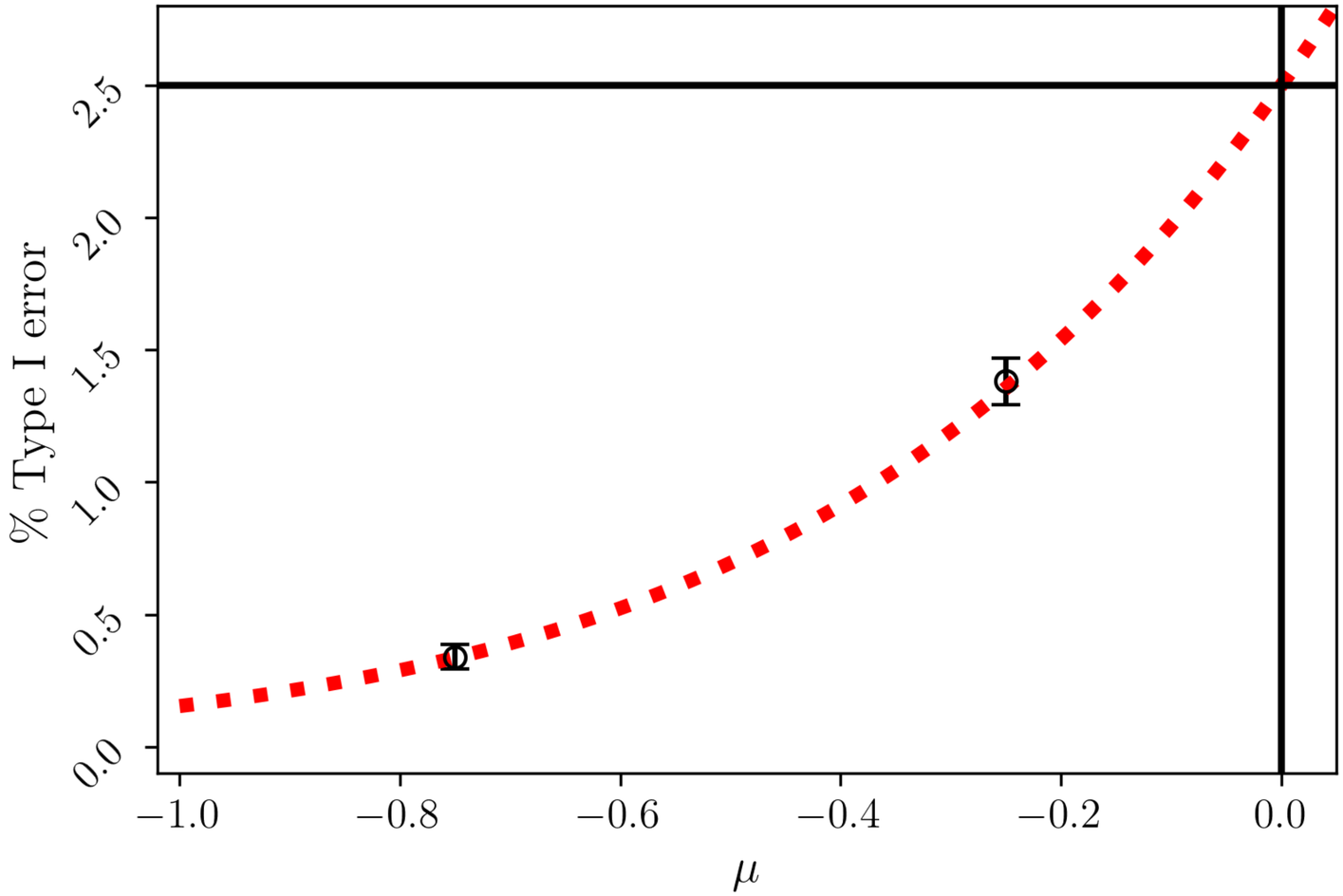
Reject if $X > 1.96$



Upper bounds on Type I error

One-sided z-test

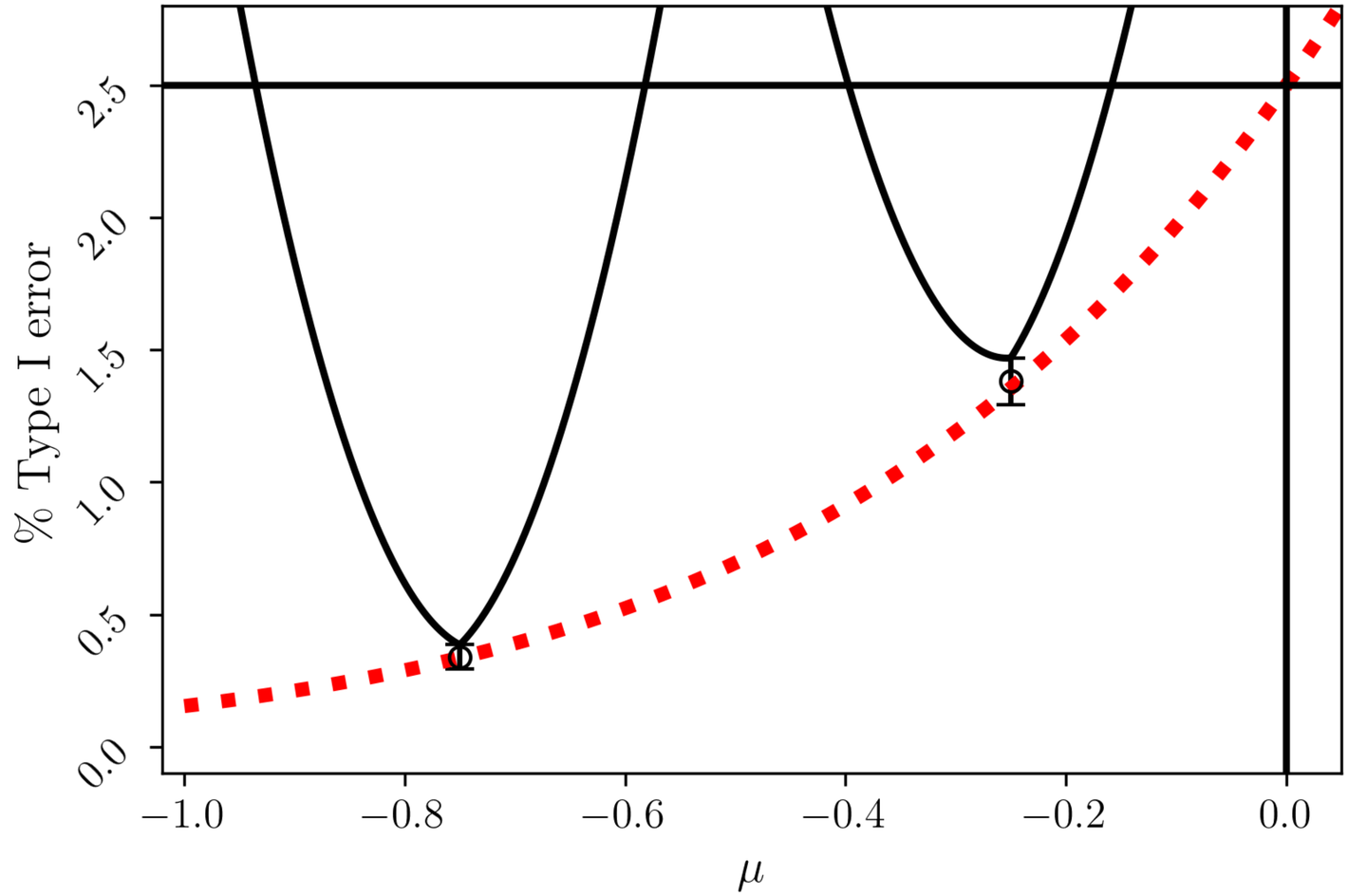
Taylor expand and bound the terms!



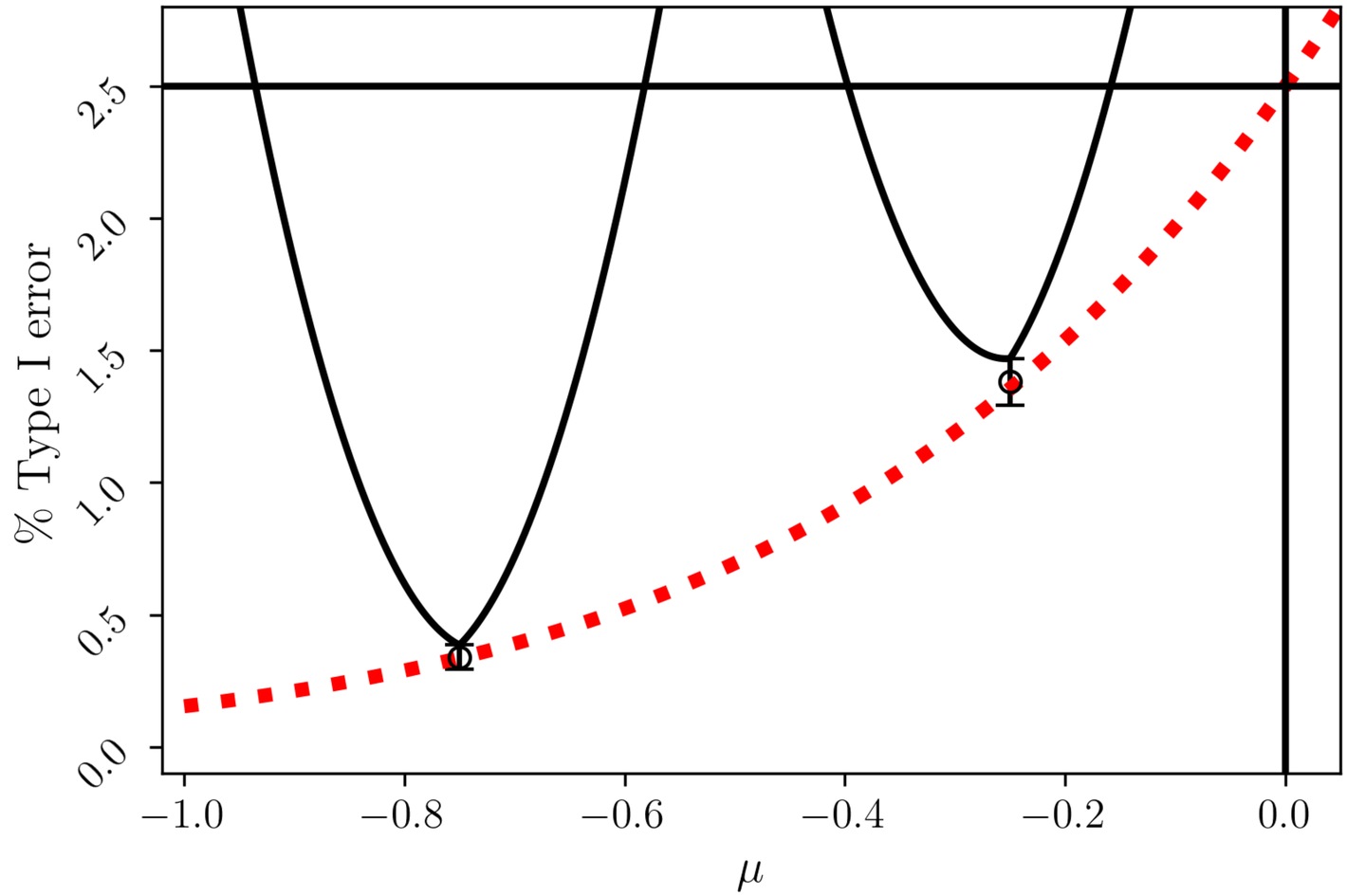
Upper bounds on Type I error

One-sided z-test

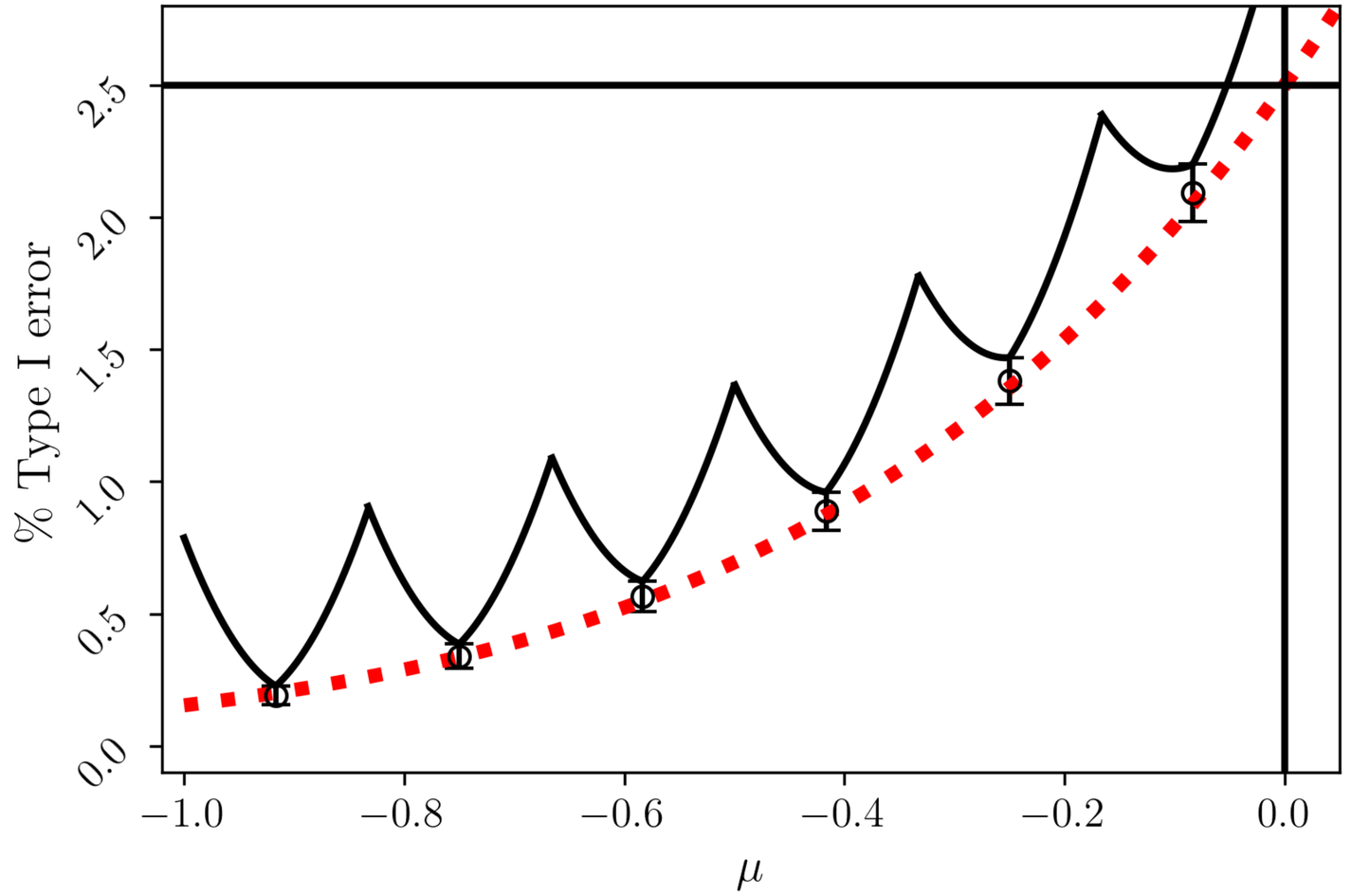
Taylor expand and bound the terms!



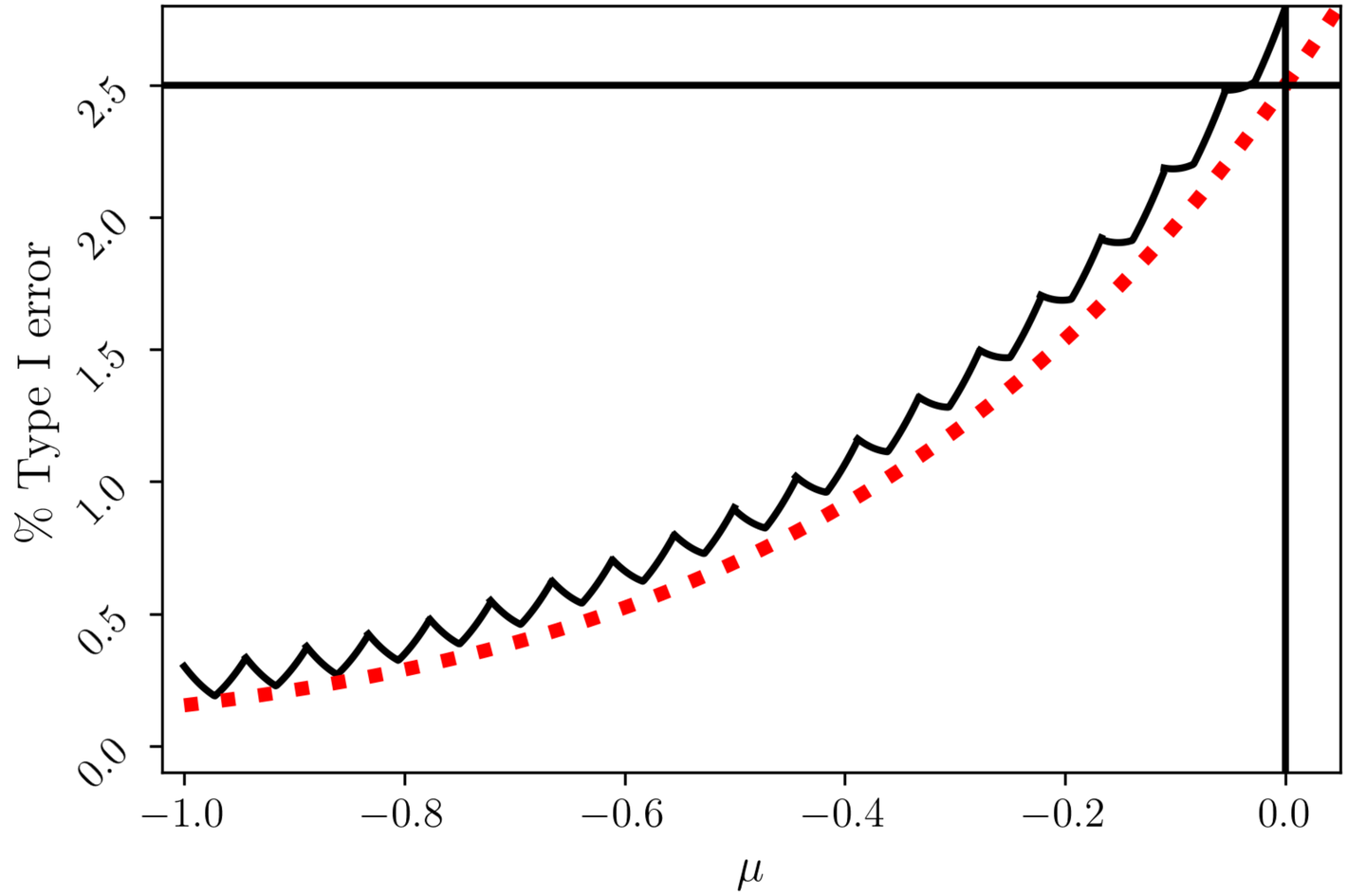
Too loose



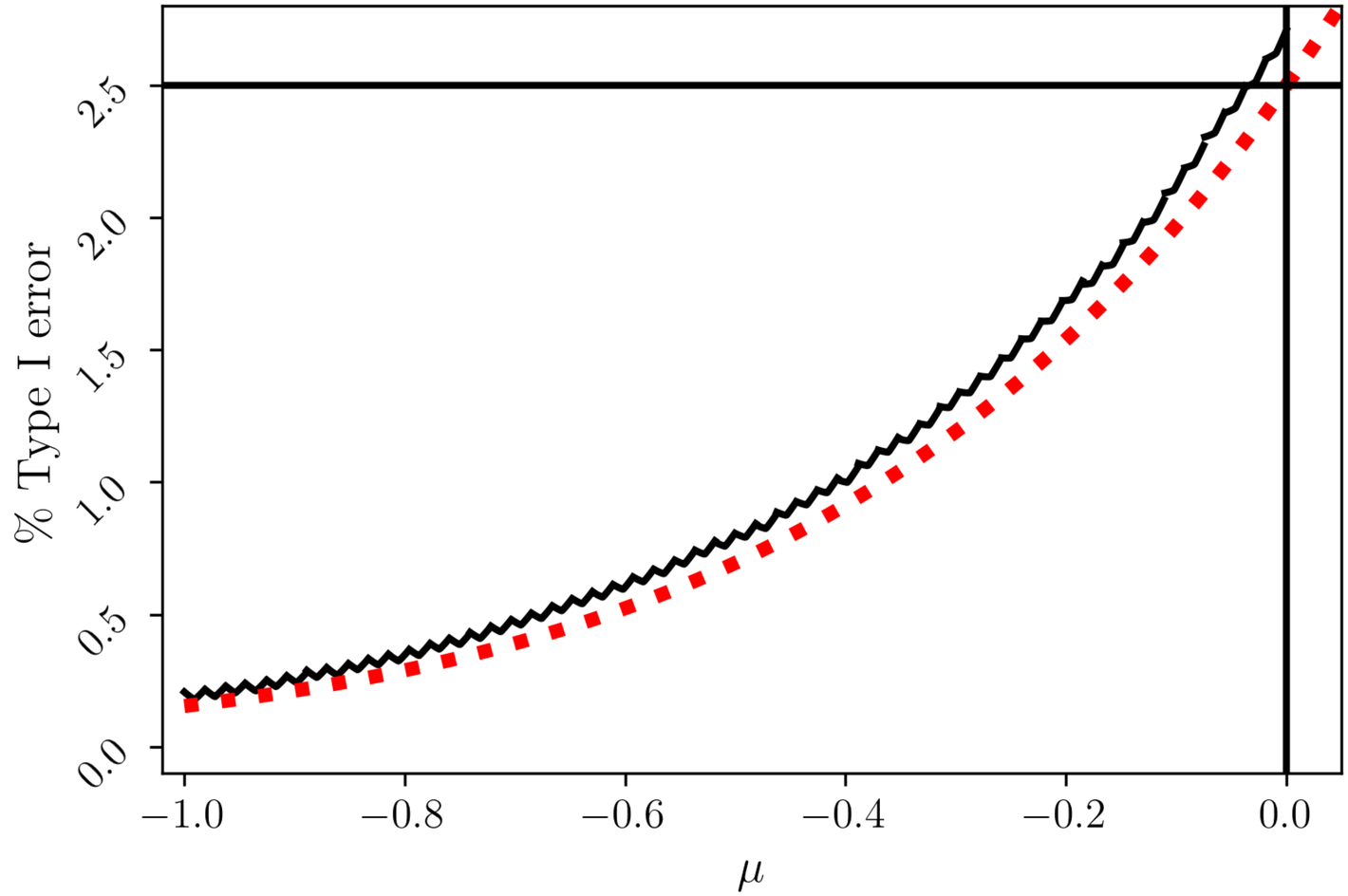
Add more points!



And more points

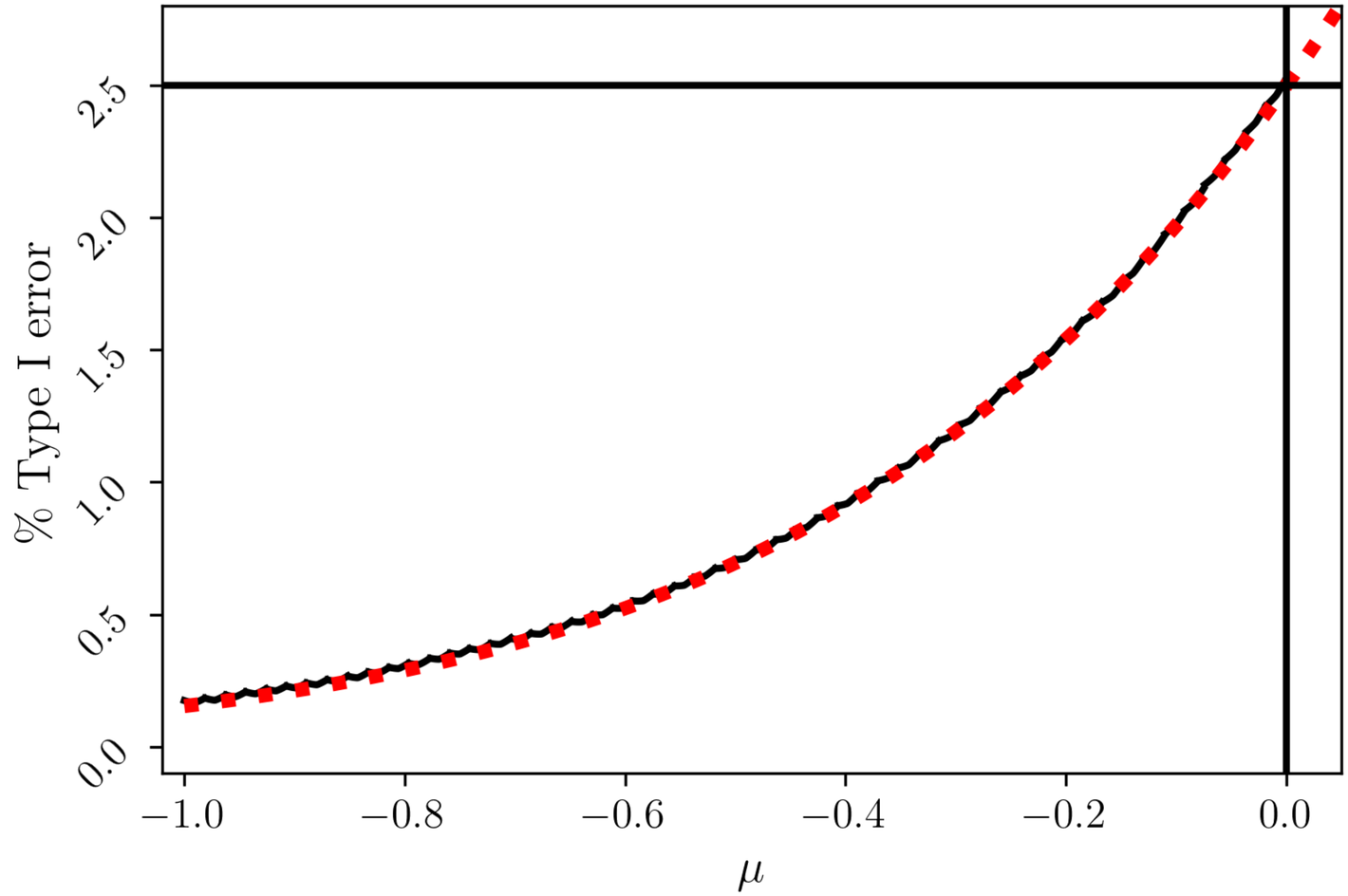


Even more points…

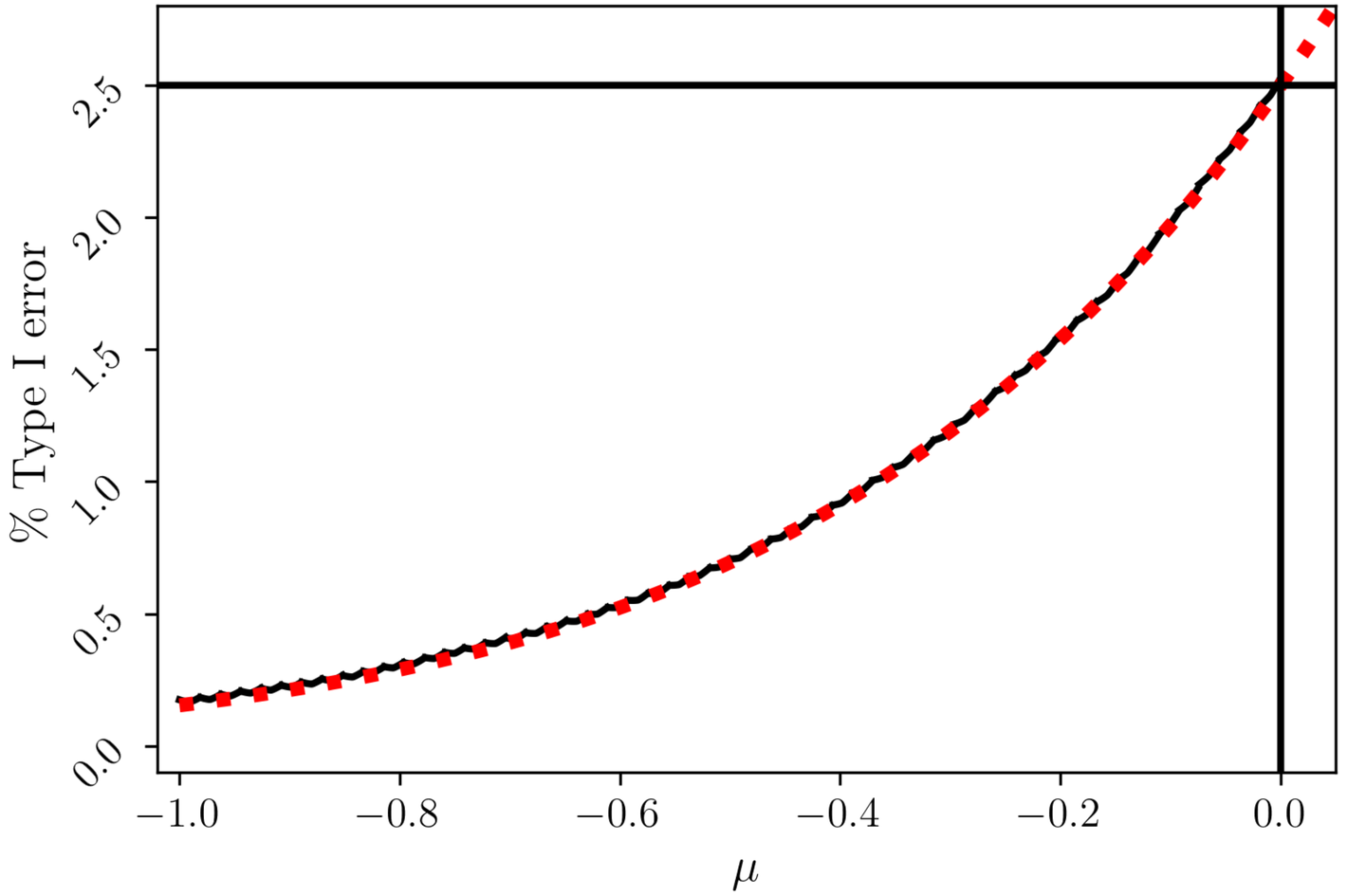


Finally, add more simulations

Tighten the confidence interval at the simulation points.



This is continuous simulation extension!



Bayesian Phase II/III selection (personal comm, N. 4/2/22)

Bayesian Phase II/III selection (personal comm, N. 4/2/22)

Idealized three arm (two treatment, one control) Phase II/III selection.

- Hierarchical Bayesian model based on Berry et al. 2013

Bayesian Phase II/III selection (personal comm, N. 4/2/22)

Idealized three arm (two treatment, one control) Phase II/III selection.

- Hierarchical Bayesian model based on Berry et al. 2013

Phase II:

- 1:1:1 assignment to arms.
- Up to 3 interim analyses with selection for Phase III and futility dropping based on the probability of being the best regimen (PrBest) and predictive probability of success (PPS).

Bayesian Phase II/III selection (personal comm, N. 4/2/22)

Idealized three arm* (two treatment, one control) Phase II/III selection.

- Hierarchical Bayesian model based on Berry et al. 2013

Phase II:

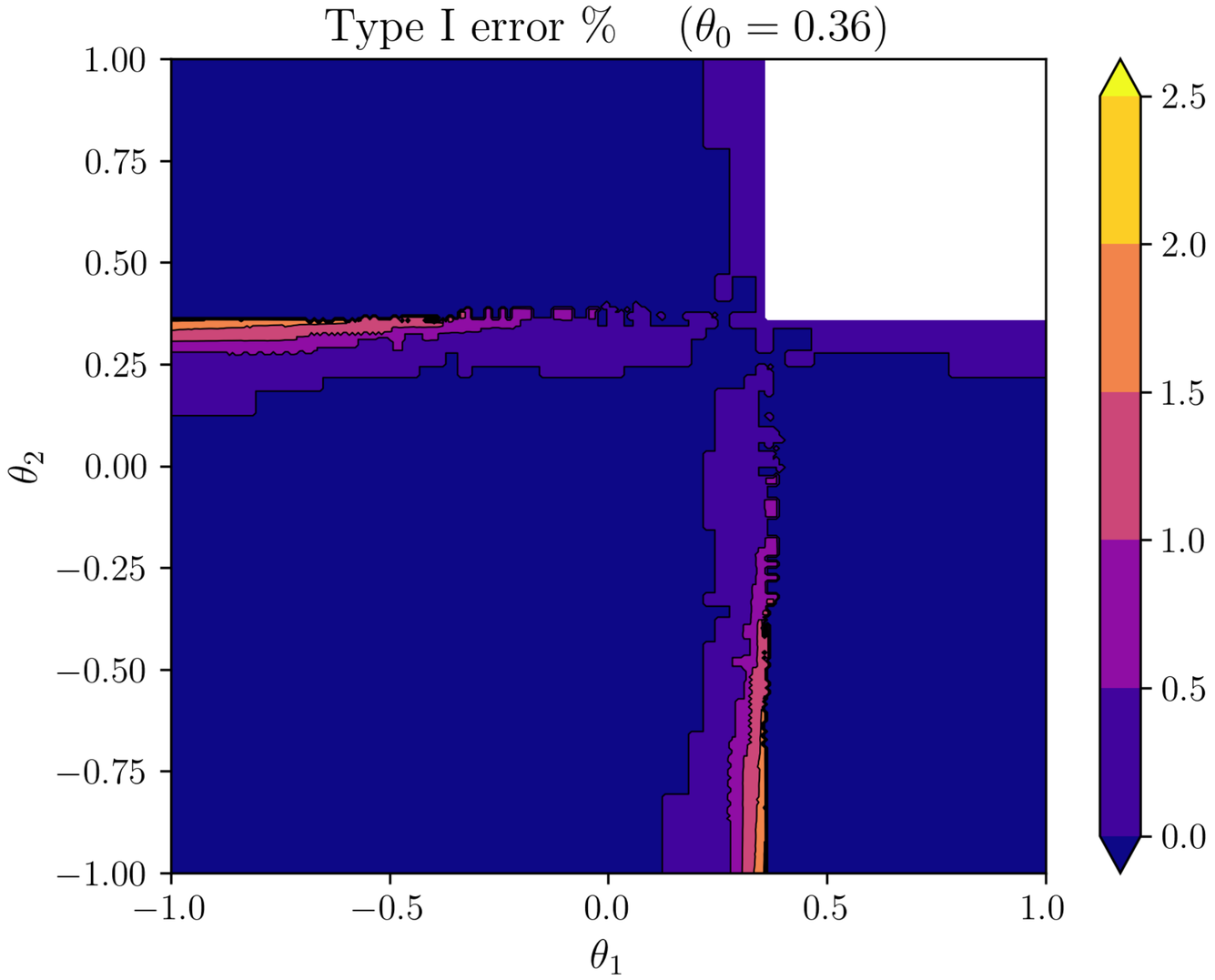
- 1:1:1 assignment to arms.
- Up to 3 interim analyses with selection for Phase III and futility dropping based on the probability of being the best regimen (PrBest) and predictive probability of success (PPS).

Phase III:

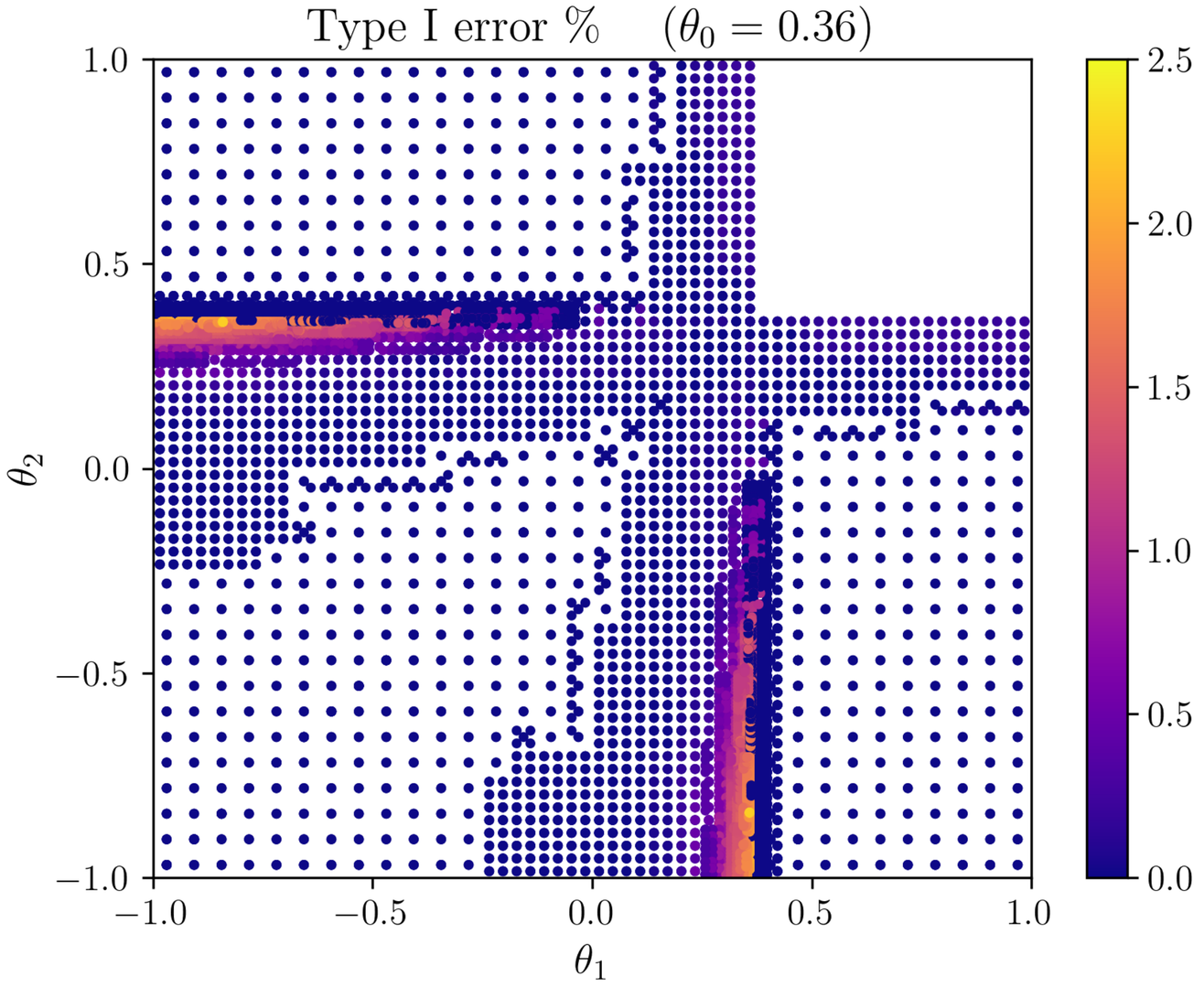
- Selected treatment vs control, one interim plus final analysis based on posterior probability of meaningful improvement over the control.

* Four arms coming soon!

Bayesian Phase II/III selection



Adaptive Gridding



Made possible by fast Bayesian inference

Integrated Nested Laplace Approximation (INLA)

Made possible by fast Bayesian inference

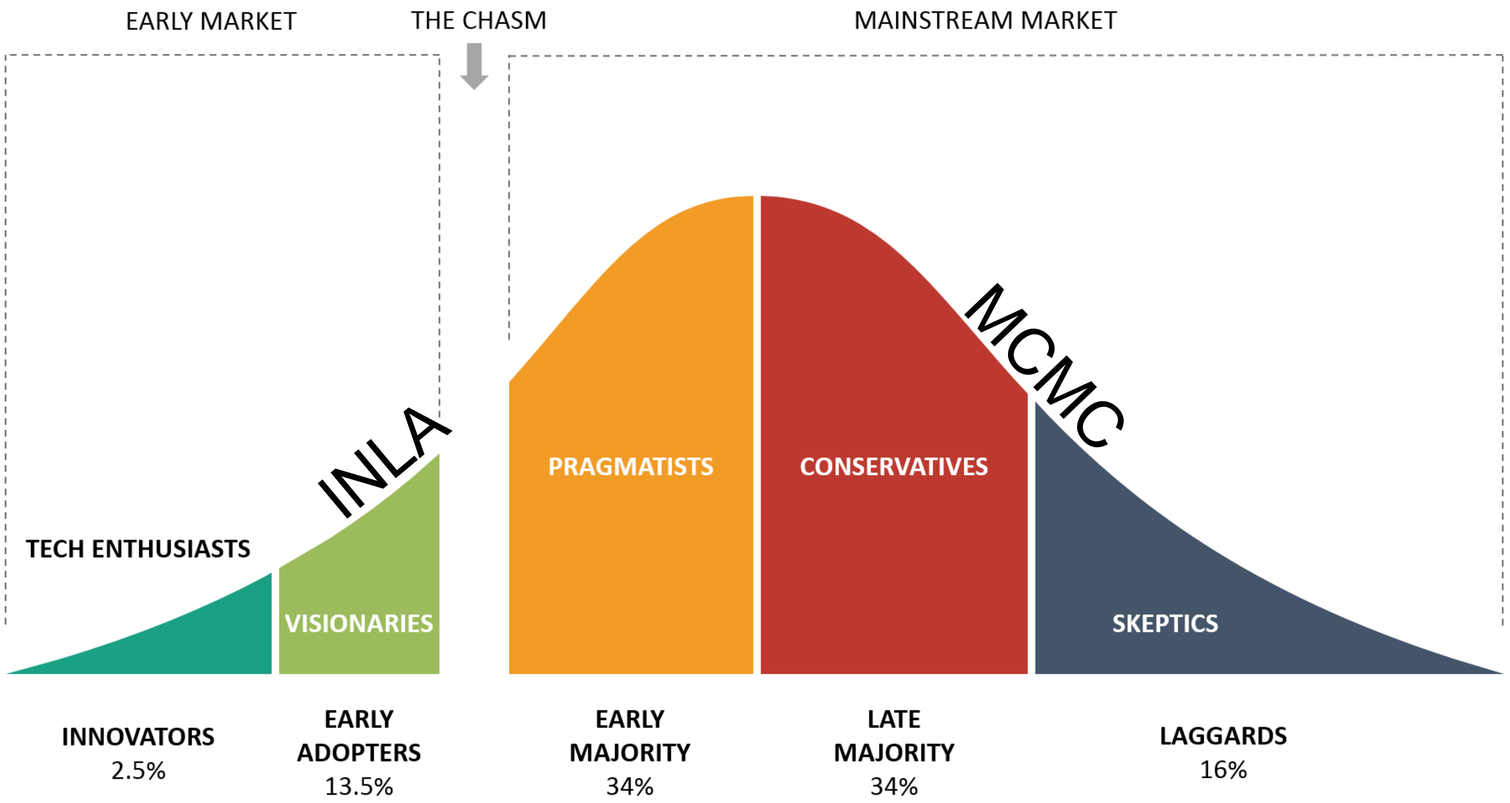
Integrated Nested Laplace Approximation (INLA)

Design type	Number of sims	Number of null points	Total simulations	Hardware	Time
Exponential hazard	100,000	16,384	1.6b	CPU	1.6 min
Thompson sampling	100,000	16,384	1.6b	CPU	4.0 min
Binomial selection	100,000	262,144	26b	CPU	2.4 min
Berry et al. 2013	10,000	5.3m	53b	CPU	2.5 hrs
Berry et al. 2013	500,000	16m	**8.3t**	**GPU**	4 hrs
Bayesian Phase II/III	variable	10m	**1.2t**	**GPU**	3 hrs

Bayesian Phase II/III: 5 years with optimized MCMC on the GPU*

* Numpyro + JAX

Why isn’t INLA more popular?



INLA is only 13 years old! (Rue et al. 2009).

Metropolis-Hastings is over 50!

Hierarchical Bayesian Basket Trial

Based on Berry et al. 2013

$$Y_j \sim Binom(n_j, p_j) \quad j = 1, \ldots, 4$$
$$p_j = \text{expit}(\theta_j + \text{logit}(q_j))$$

Hierarchical Bayesian Basket Trial

Based on Berry et al. 2013

$$
\begin{aligned}
Y_j &\sim Binom(n_j, p_j) \quad j = 1, \ldots, 4 \\
p_j &= \text{expit}(\theta_j + \text{logit}(q_j)) \\
\theta_j &\sim \mathcal{N}\left(\mu, \sigma^2\right)
\end{aligned}
$$

Hierarchical Bayesian Basket Trial

Based on Berry et al. 2013

$$
\begin{aligned}
Y_j &\sim Binom(n_j, p_j) \quad j = 1, \ldots, 4 \\
p_j &= \text{expit}(\theta_j + \text{logit}(q_j)) \\
\theta_j &\sim \mathcal{N}\left(\mu, \sigma^2\right) \\
\mu &\sim \mathcal{N}\left(\mu_0, \sigma_0^2\right)
\end{aligned}
$$

Hierarchical Bayesian Basket Trial

Based on Berry et al. 2013.

$$
\begin{aligned}
Y_j &\sim Binom(n_j, p_j) \quad j = 1, \ldots, 4 \\
p_j &= \text{expit}(\theta_j + \text{logit}(q_j)) \\
\theta_j &\sim \mathcal{N}\left(\mu, \sigma^2\right) \\
\mu &\sim \mathcal{N}\left(\mu_0, \sigma_0^2\right) \\
\sigma^2 &\sim \Gamma^{-1}(\alpha, \beta)
\end{aligned}
$$

Hierarchical Bayesian Basket Trial

Based on Berry et al. 2013.

$$
\begin{aligned}
Y_j &\sim Binom(n_j, p_j) \quad j = 1, \ldots, 4 \\
p_j &= \text{expit}(\theta_j + \text{logit}(q_j)) \\
\theta_j &\sim \mathcal{N}\left(\mu, \sigma^2\right) \\
\mu &\sim \mathcal{N}\left(\mu_0, \sigma_0^2\right) \\
\sigma^2 &\sim \Gamma^{-1}(\alpha, \beta)
\end{aligned}
$$

Reject if $\mathbb{P}[p_i > p_0 | Y] > c_i$ **for some arm** i

Computational goal: $\mathbb{P}\left[\theta_j \middle| Y\right]$

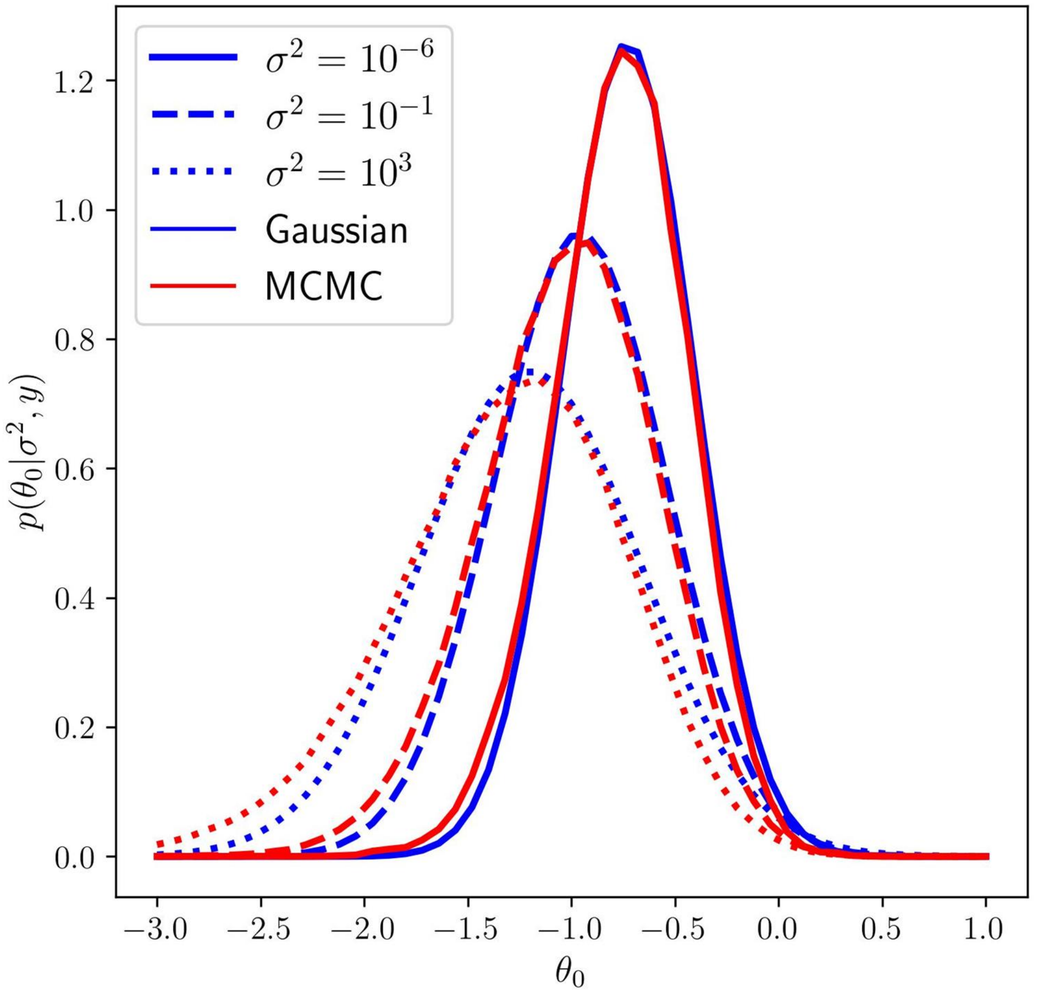
INLA: Gaussian slices

1. Fix the hyper-parameter σ^2

$$\theta_j \sim \mathcal{N}\left(\mu, \sigma^2\right)$$
$$\mu \sim \mathcal{N}\left(\mu_0, \sigma_0^2\right)$$
$$\sigma^2 \sim \Gamma^{-1}(\alpha, \beta)$$

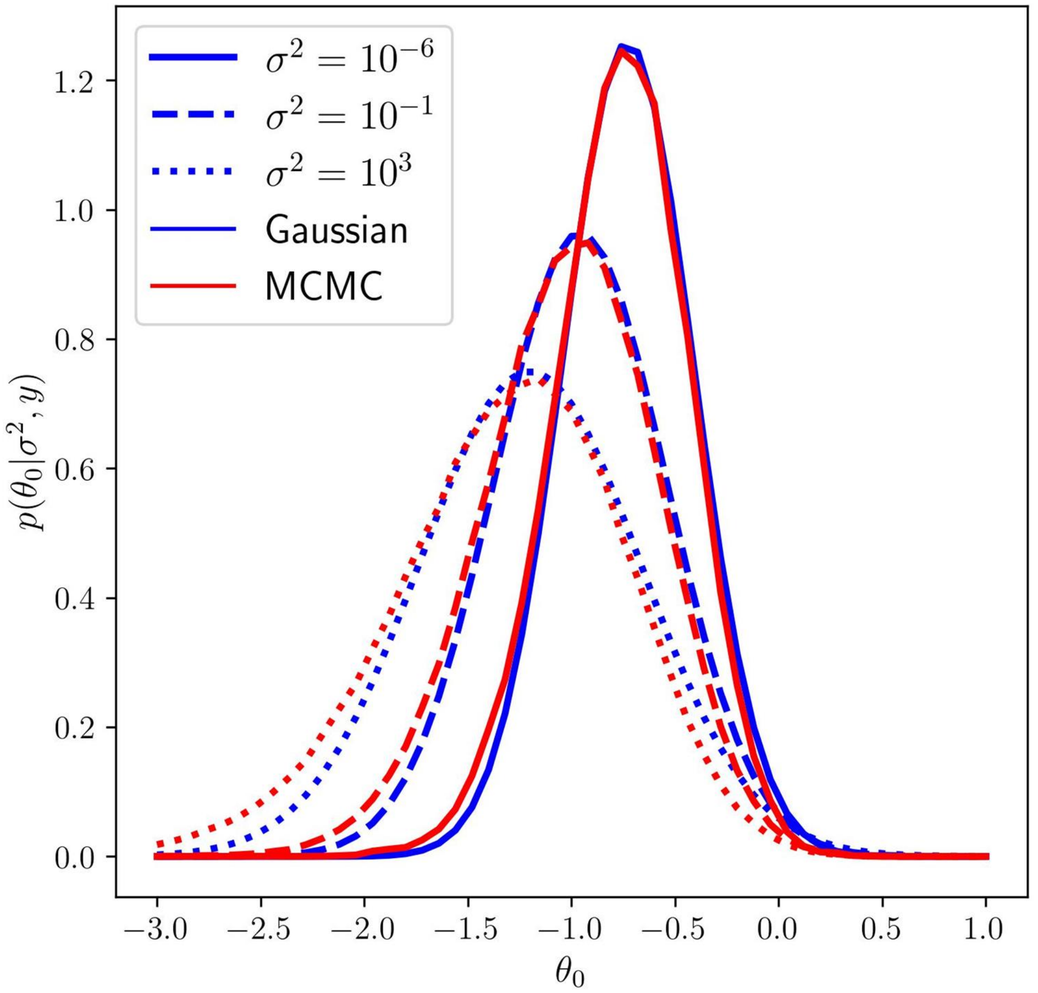
INLA: Gaussian slices

1. Fix the hyper-parameter σ^2
2. Treat the conditional posterior $\mathbb{P}[\theta_i|\sigma^2, Y]$ as approximately Gaussian



INLA: Gaussian slices

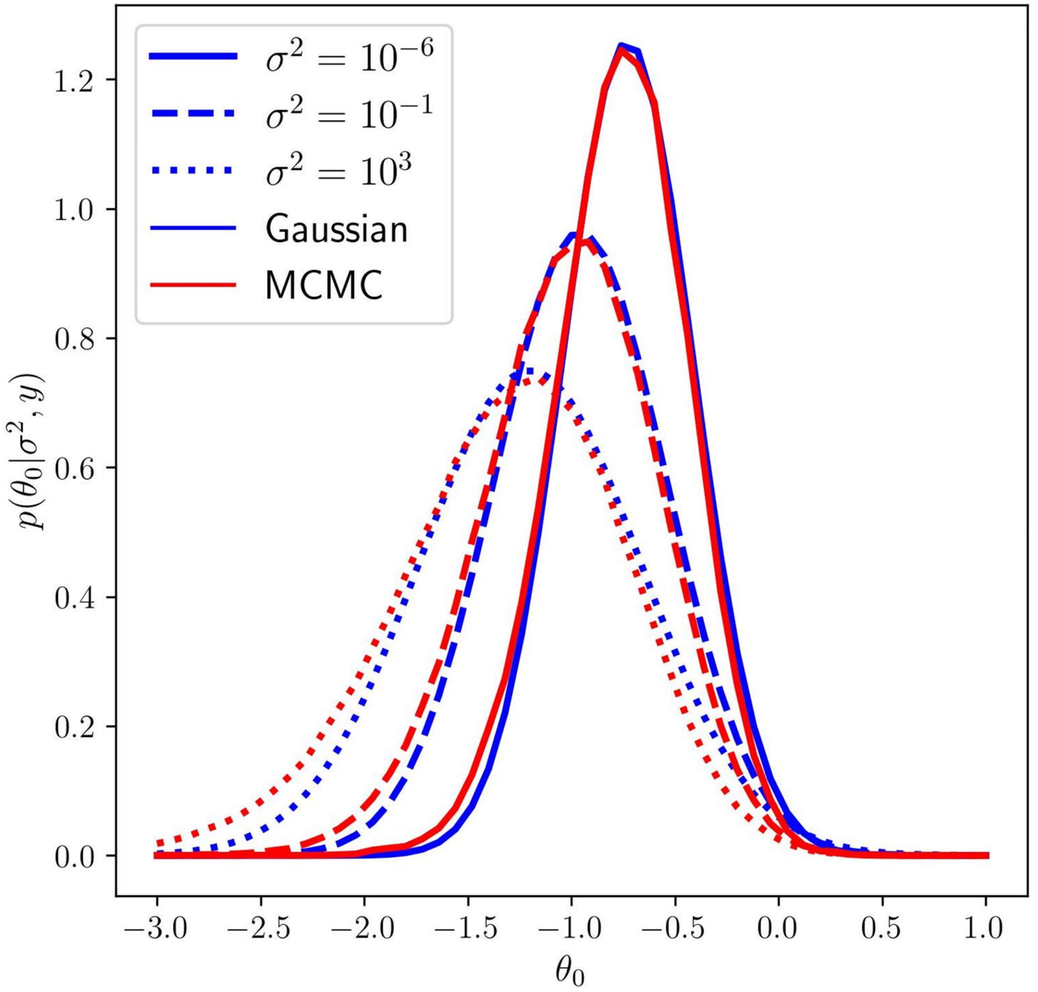
1. Fix the hyper-parameter σ^2
2. Treat the conditional posterior $\mathbb{P}[\theta_i|\sigma^2, Y]$ as approximately Gaussian
3. ND analytical integral to get: $\mathbb{P}[\sigma^2|Y]$



INLA: Gaussian slices

1. Fix the hyper-parameter σ^2
2. Treat the conditional posterior $\mathbb{P}[\theta_i|\sigma^2, Y]$ as approximately Gaussian
3. ND analytical integral to get: $\mathbb{P}[\sigma^2|Y]$
4. 1D numerical integral to get: $\mathbb{P}[\theta_i|Y]$

We call this “Gaussian INLA”

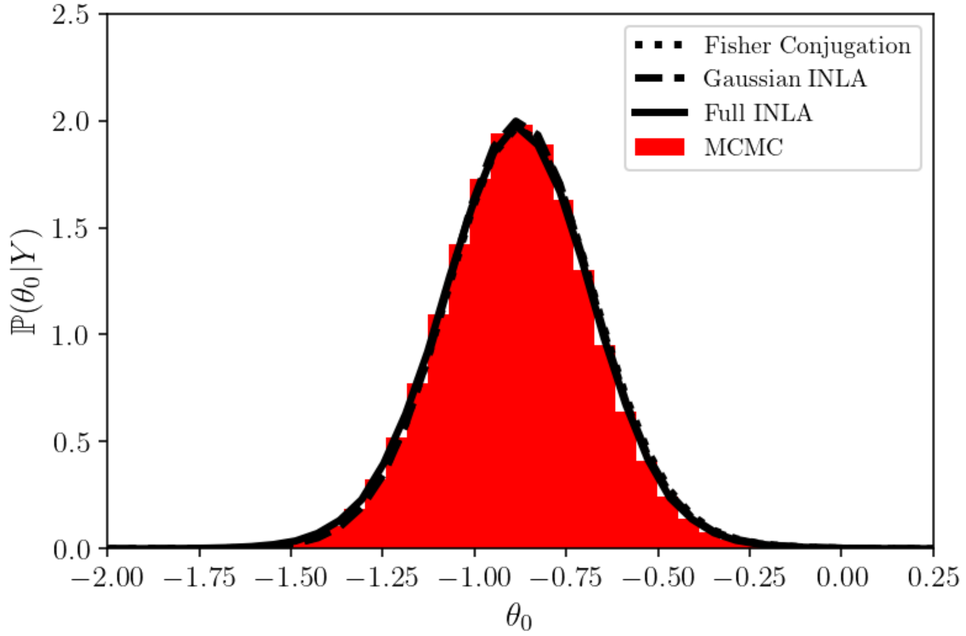


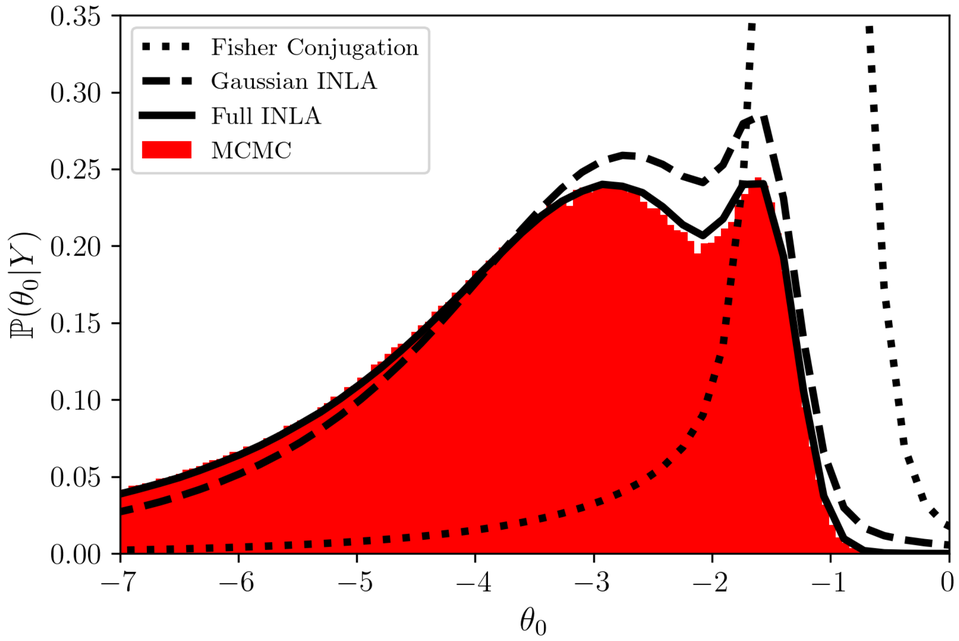
Adaptivity with INLA

Problem: INLA doesn't converge because it's not a sequence

Our approach:

1. Use a hierarchy of approximations.
2. Use K-L divergence to decide when to stop.





Why is our INLA code fast?

Optimized for small dimension

Why is our INLA code fast?

Optimized for small dimension

	Runtime	Speedup
R-INLA*	511 ms	1X
small-d	1.09 us	469,000X
GPU	41 ns	12,460,000X

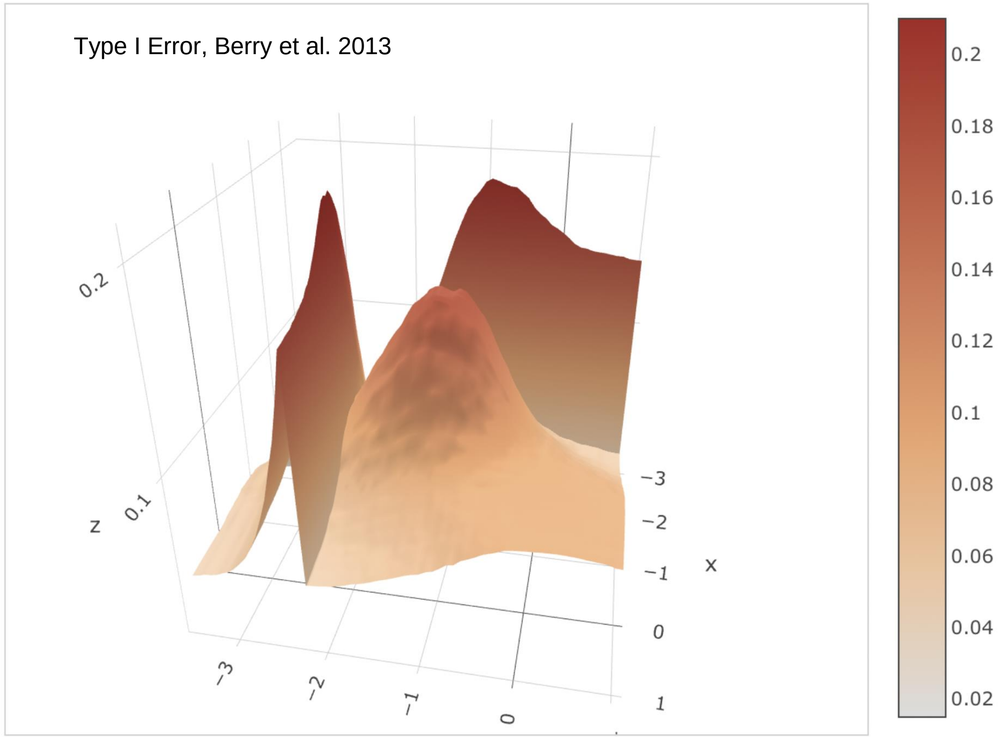
* R-INLA is fast for high dimensional problems.

Confirm Solutions

When your simulations run MILLIONS of times faster, it opens a lot of doors.

Continuous simulation extension:

Type I Error proofs for complex trial designs



Please share trial designs that you’d like us to validate!