Design of Population Studies of HIV Dynamics

Kathryn Chaloner
Department of Biostatistics
Department of Statistics & Actuarial Science
University of Iowa

joint work with
Cong Han PhD, *TAP Pharmaceuticals*

PODE, Sandwich, UK, May 5th 2007

Outline

- Example: HIV Dynamics
- Pragmatic comparison of designs by simulation.
- Use of an approximate model for design.
- Exact calculations for design.

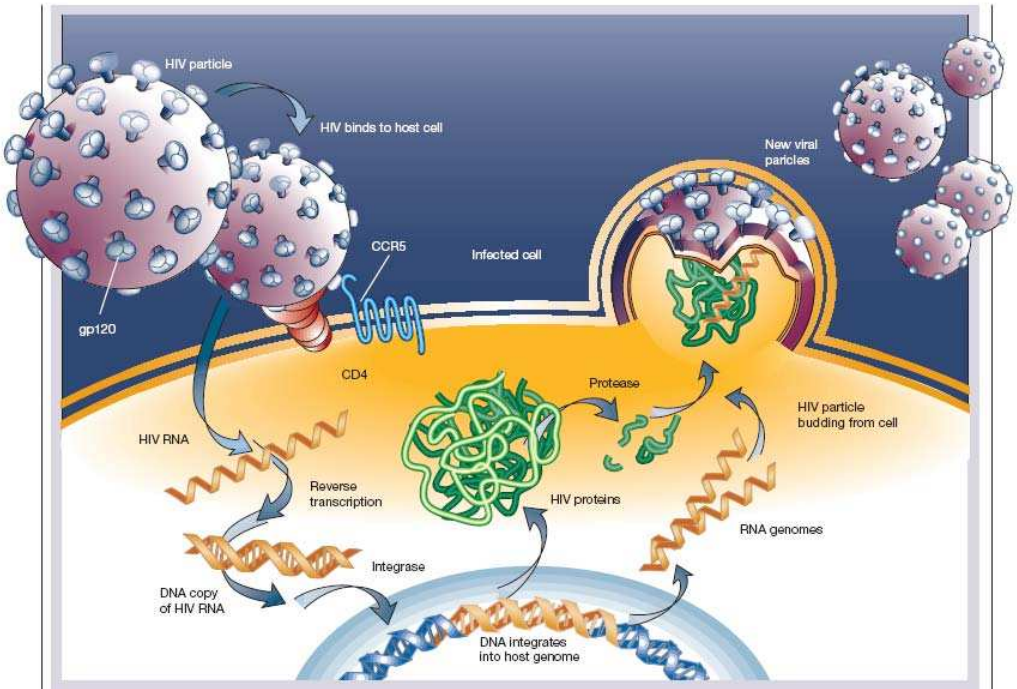


Figure 1 The human immunodeficiency virus. Replication cycle of HIV showing multiple steps available for attack by antiretroviral drugs.

Weiss RA, 2001, Nature, 410:963.

Example: Perelson *et al.*, 1996, *Science* **271**: 1582-1586.

Ritonavir administered to five patients;
15-16 plasma HIV RNA measurements within one week.

Newly generated HIV virions are non-infectious.

Non-infectious virions gradually replace the infectious ones.

Non-infectious virions gradually die out.

Interest in c the rate of decay of free virus and δ the rate of decay of infected cells.

$$
\begin{aligned}
V_0 &= \texttt{baseline viral load (copies/ml)} \\
V_I &= \texttt{infectious HIV viruses(copies/ml)} \\
V_{NI} &= \texttt{noninfectious HIV viruses(copies/ml)} \\
T &= \texttt{uninfected CD4+ cells (cells/ml)} \\
T^* &= \texttt{infected CD4+ cells (cells/ml)} \\
\delta &= \texttt{rate of loss of infected cells per day} \\
c &= \texttt{rate of virion clearance} \\
N &= \texttt{expected number of viruses produced} \\
&\ \texttt{by one infected cell}
\end{aligned}
$$

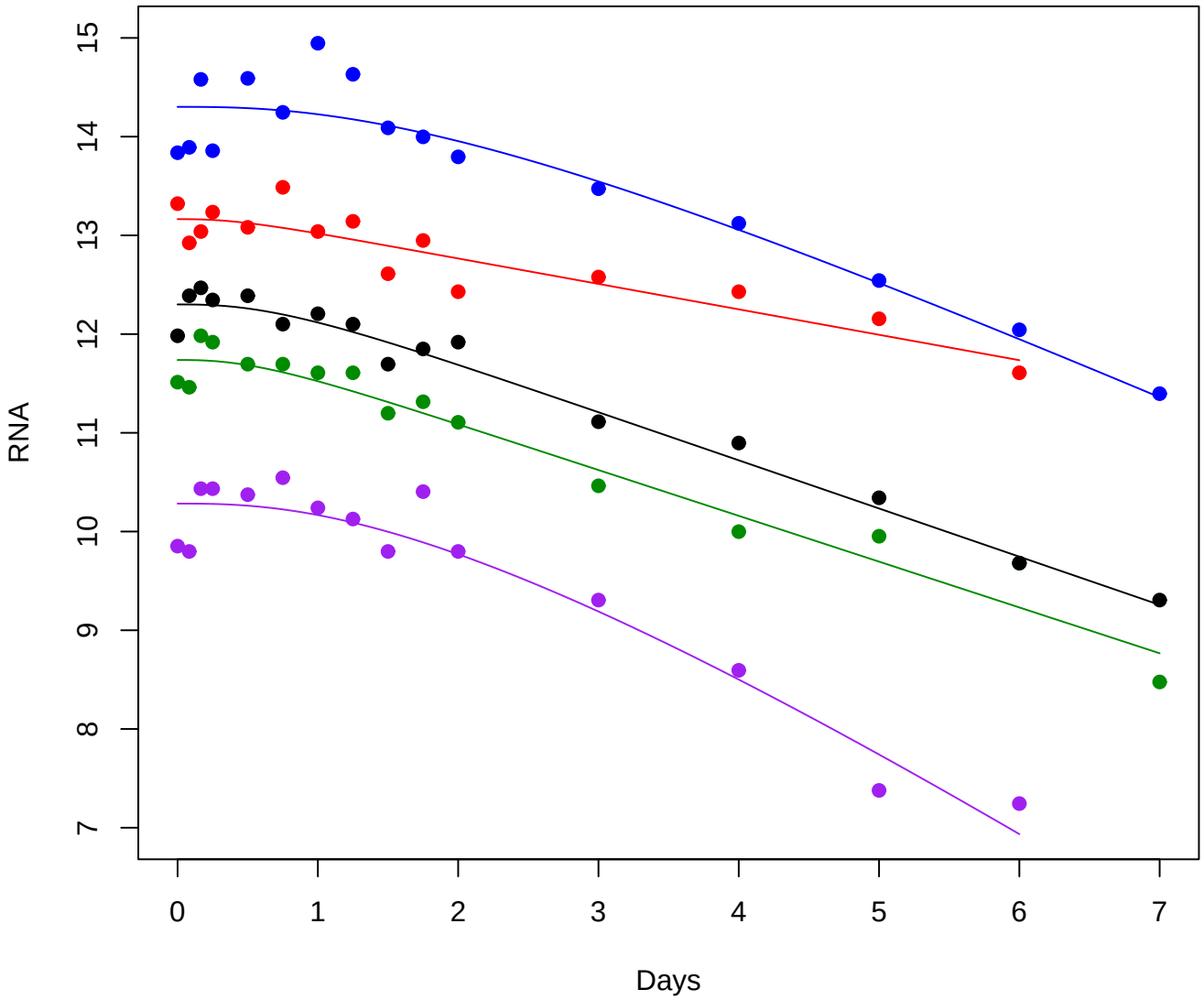
Assume V_0, T_0, T_0^* are approximately constant prior to therapy.

Assume protease inhibitor is 100% effective in producing non-infectious virus.

$$
\begin{aligned}
\frac{dT^*}{dt} &= kV_I - \delta T^* \\
\frac{dV_I}{dt} &= -cV_I \\
\frac{dV_{NI}}{dt} &= N\delta T^* - cV_{NI}
\end{aligned}
$$

Solve with $V(t) = V_I(t) + V_{NI}(t)$:

$$
V(t) = V_0 \left[\frac{c^2}{(c-\delta)^2} e^{-\delta t} - \frac{c^2 - (c-\delta)^2}{(c-\delta)^2} e^{-ct} - \frac{c\delta}{c-\delta} t e^{-ct} \right]
$$


RNA
15
14
13
12
11
10
9
8
7
0
1
2
3
4
5
6
7
Days

The experiment in *Science* provided an estimate of the half life of HIV virus circulating in plasma. Confirmed the rapid replication rate of HIV and led to understanding of resistance to drugs.

Design issues for nonlinear mixed-effects models:

1. At what time to sample subjects and measure HIV RNA?

2. How many measurements of HIV RNA?

3. How many subjects?

Assume the same sampling scheme of m time-points for each of n subjects.

1. Marginal likelihood has no closed-form:

$$
\begin{aligned}
&= L(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \sigma^2; \mathbf{y}_i) \\
&= \int \left[p(\boldsymbol{\theta}_i | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \prod_{j=1}^{m} p(y_{ij} | \boldsymbol{\theta}_i, \sigma^2) \right] d\boldsymbol{\theta}_i \\
&= f(\mathbf{y}_i)
\end{aligned}
$$

2. Evaluation of the information matrix is hard:

$$
\int_{\mathbb{R}^m} \frac{\nabla f(\mathbf{y}_i) \nabla^\top f(\mathbf{y}_i)}{f(\mathbf{y}_i)} d\mathbf{y}_i
$$

   Problematic when number of measurements per person (m) is large. Have done examples where $m = 4$.

Choose a sampling schedule to achieve good parameter estimation but usual methods are hard to implement.

A pragmatic approach

Prior distribution π to be used for data analysis:

$$\begin{aligned} \sigma^{-2} &\sim G(\alpha, \beta) \\ \boldsymbol{\mu} &\sim N(\boldsymbol{\eta}, \boldsymbol{\Lambda}) \\ \boldsymbol{\Sigma}^{-1} &\sim W(\boldsymbol{\Omega}, \nu) \end{aligned}$$

Interested in estimating $\boldsymbol{\mu}$, the population mean parameters (specifically c and δ).

Quadratic loss function in estimating $\boldsymbol{\mu}$:

Loss is the posterior risk.

$$\sum_k \mathrm{Var}^{\pi}(\mu_k|\mathbf{y}) \text{ or } \sum_k d_{kk}\mathrm{Var}^{\pi}(\mu_k|\mathbf{y}).$$

Find a design to minimize preposterior risk

$$\int \sum_k \mathrm{Var}^{\pi}(\mu_k|\mathbf{y}) dG_{\xi}(\mathbf{y}) \text{ or } \int \sum_k d_{kk} \mathrm{Var}^{\pi}(\mu_k|\mathbf{y}) dG_{\xi}(\mathbf{y})$$

Prior distribution, ω:

$$\begin{aligned} \sigma^{-2} &\sim G(\gamma, \delta) \\ \boldsymbol{\mu} &\sim N(\boldsymbol{\zeta}, \mathbf{K}) \\ \boldsymbol{\Sigma}^{-1} &\sim W(\boldsymbol{\Psi}, \chi) \end{aligned}$$

Under ω, a design ξ induces a predictive distribution G_{ξ} on $\mathbf{y}$.

Han and Chaloner, 2004, *Biometrics*, give conditions for the posterior loss and its expectation over the marginal distribution of the data to be finite.

Based on the Perelson experiment and population HIV dynamic model (Han, Chaloner, and Perelson, 2002) a finite set of alternative designs is examined for a particular prior distribution.

The prior distribution was derived from other data available at the time the original data was collected. We use a Gamma distribution for σ^2, a normal distribution for $(V_0, c, \delta)^T$ and a Wishart distribution for Σ.

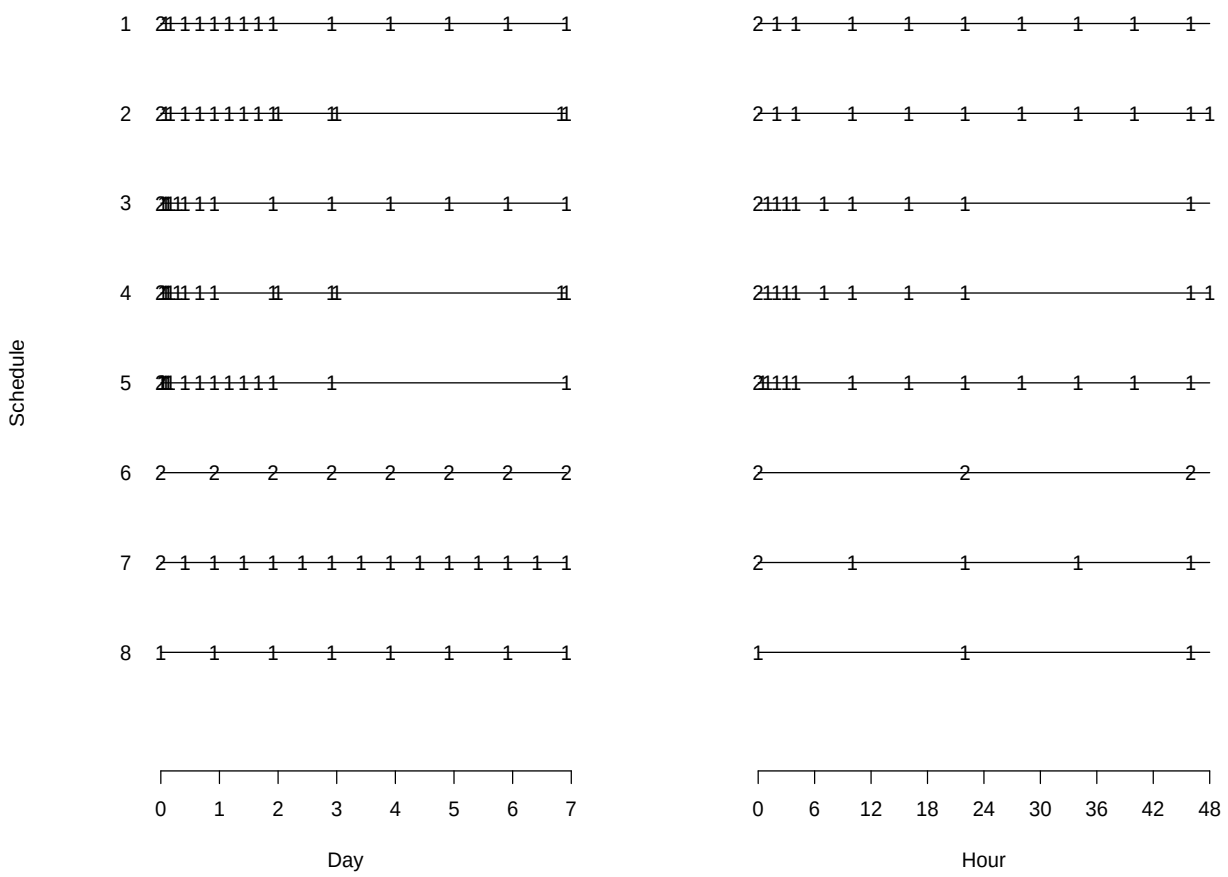
HIV dynamic model:

$$y_{ij}|V_{0i}, c_i, \delta_i, \sigma^2 \overset{ind}{\sim} N\left(\log V_i(t_j), \sigma^2\right),$$

where

$$V(t) = V_0 \left[\frac{c^2}{(c-\delta)^2} e^{-\delta t} - \frac{c^2 - (c-\delta)^2}{(c-\delta)^2} e^{-ct} - \frac{c\delta}{c-\delta} t e^{-ct}\right]$$

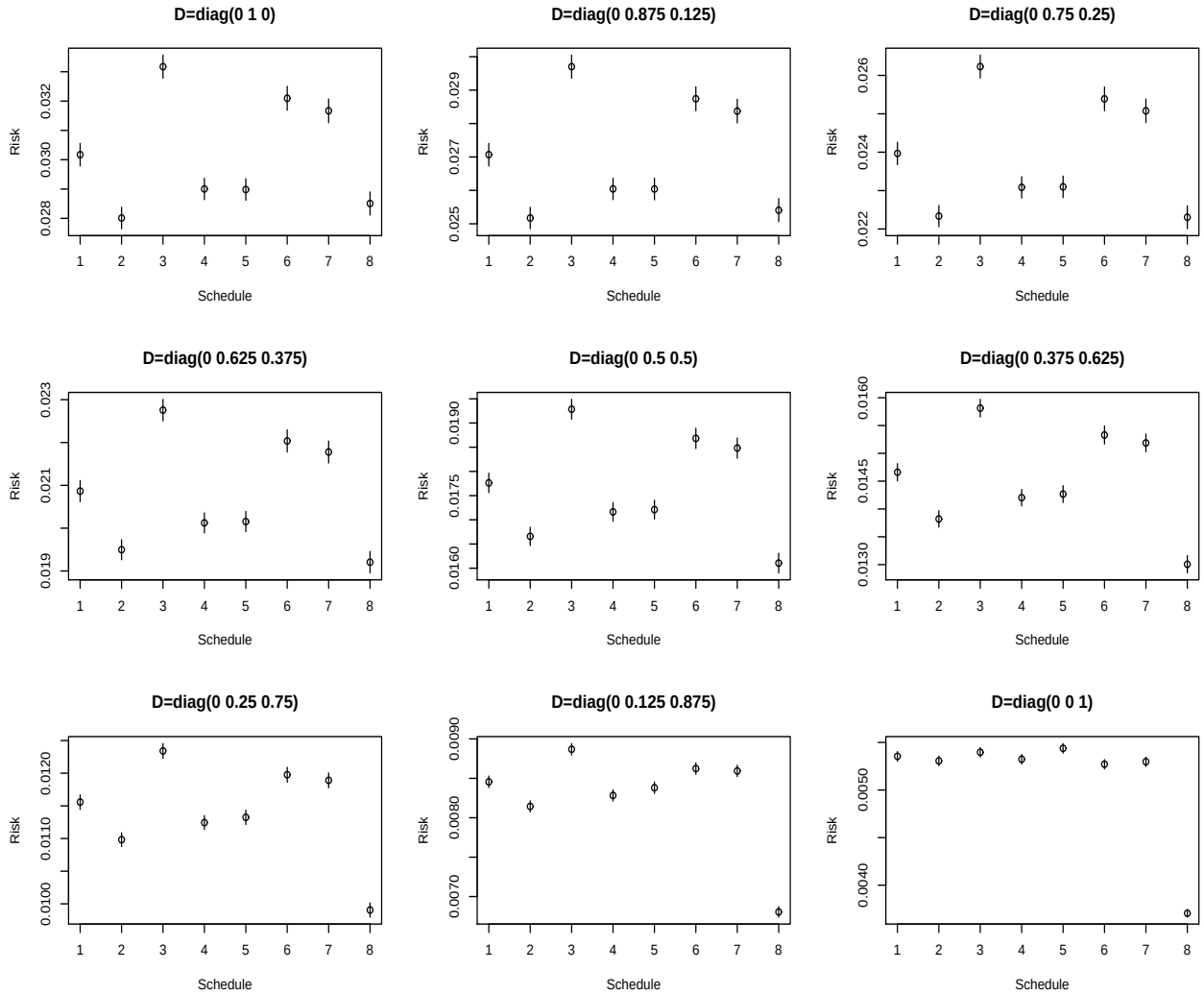
The 8 sampling schedules for a potential HIV dynamic experiment:



Assume 240 measurements are allowed: 15 subjects if schedules 1-7 are used, 30 if schedule 8 is used.

Ignore the overhead cost associated with recruiting subjects.

Preposterior risk $\int \sum_k d_{kk} \mathrm{Var}^{\pi}(\mu_k|\mathbf{y}) dG_{\xi}(\mathbf{y})$.



- Using this pragmatic computational approach different sampling schemes can be compared using an MCMC estimate of the predictive distribution of the posterior variance.

- The approach is computationally intensive and does not find an "optimal" design.

- The criterion does not require computation of the information matrix or any approximation to it.

- The only approximation is through the MCMC to analyse the data and the simulation of the data.

Can we find an optimal design using either

1. the exact calcuation of the information matrix or

2. an approximation to it?

Han and Chaloner (2005, *Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention*, eds. W. Tan and H. Wu) examine both these approaches.

Mentre, Mallet and Baccar (1997) use a first order Taylor expansion of the mean function of the response, compute the information matrix for that model, and use it for design. There are subsequent examples in the pharmacokinetics literature by them and others, including Fedorov, Gagnon and Leonov (2002). See also Wolfinger, 2000.

$$y_{ij}|\boldsymbol{\theta}_i \overset{ind}{\sim} N\left(r(\boldsymbol{\theta}_i, t_j), \sigma^2\right), \ \boldsymbol{\theta}_i \overset{iid}{\sim} N(\boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

and

$$r(\boldsymbol{\theta}_i, t_j) \approx s(\boldsymbol{\mu}, t_j) + (\boldsymbol{\theta}_i - \boldsymbol{\mu})^\top \nabla r(\boldsymbol{\mu}, t_j)$$

where ∇ denotes the gradient with respect to $\boldsymbol{\mu}$.

The approximation is exact if the mean function is linear in the random effects. The approximation works better if the variance of the random effects are small or the nonlinearity is mild. With a fixed number of measurements per subject the estimators can be inconsistent.

The Fisher information matrix of this approximate model can be computed in closed form. It is fairly complicated and each term depends on the time points at which the measurements are taken (Han 2002, Retout and Mentre 2003).

Example, consider the model of Wu and Ding (1999) from HIV dynamics:

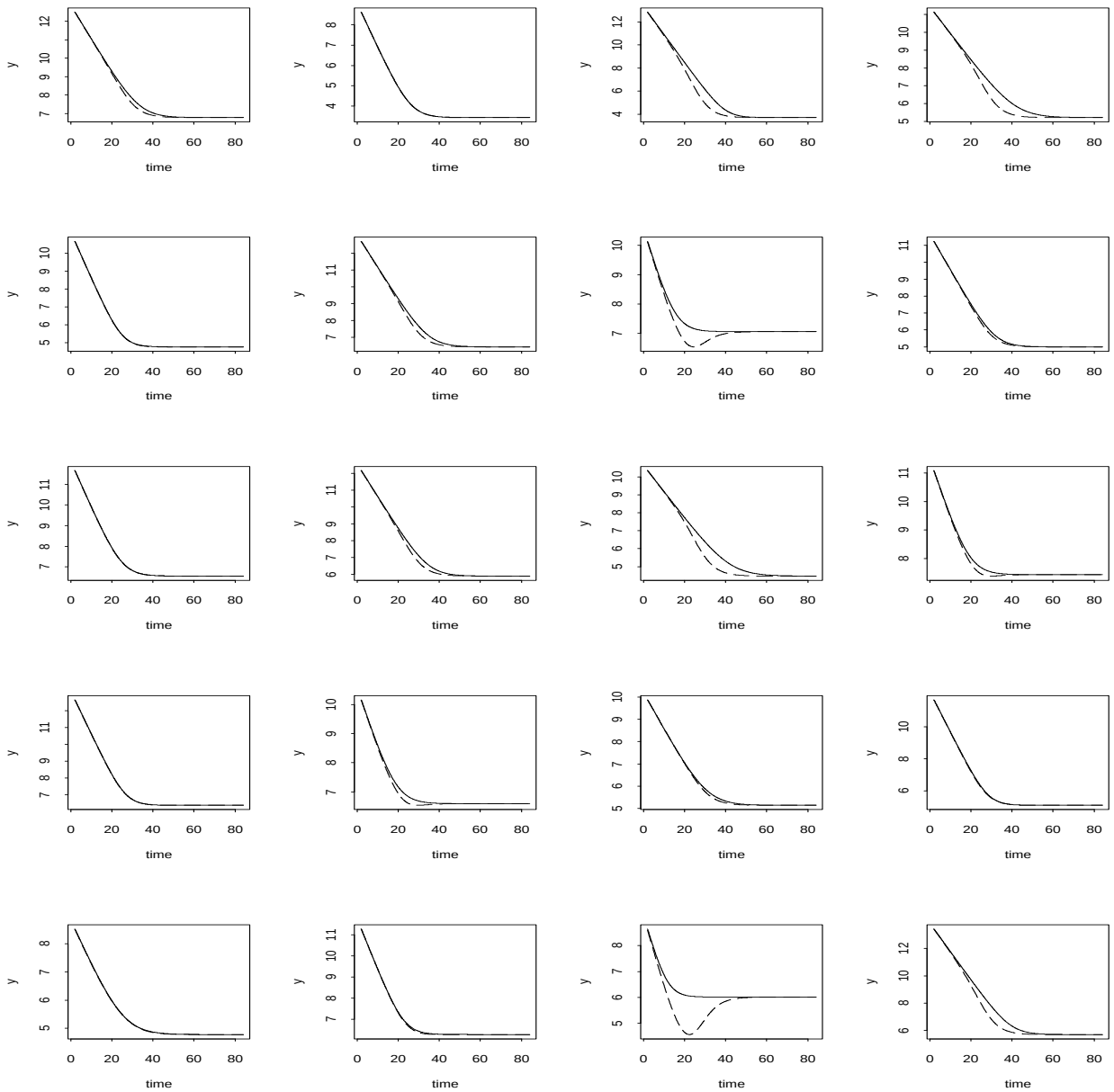
$$r(\boldsymbol{\theta}_i, t_j) = \log[exp(\theta_{i1}) + exp(\theta_{i2} - \theta_{i3} t_j)]$$

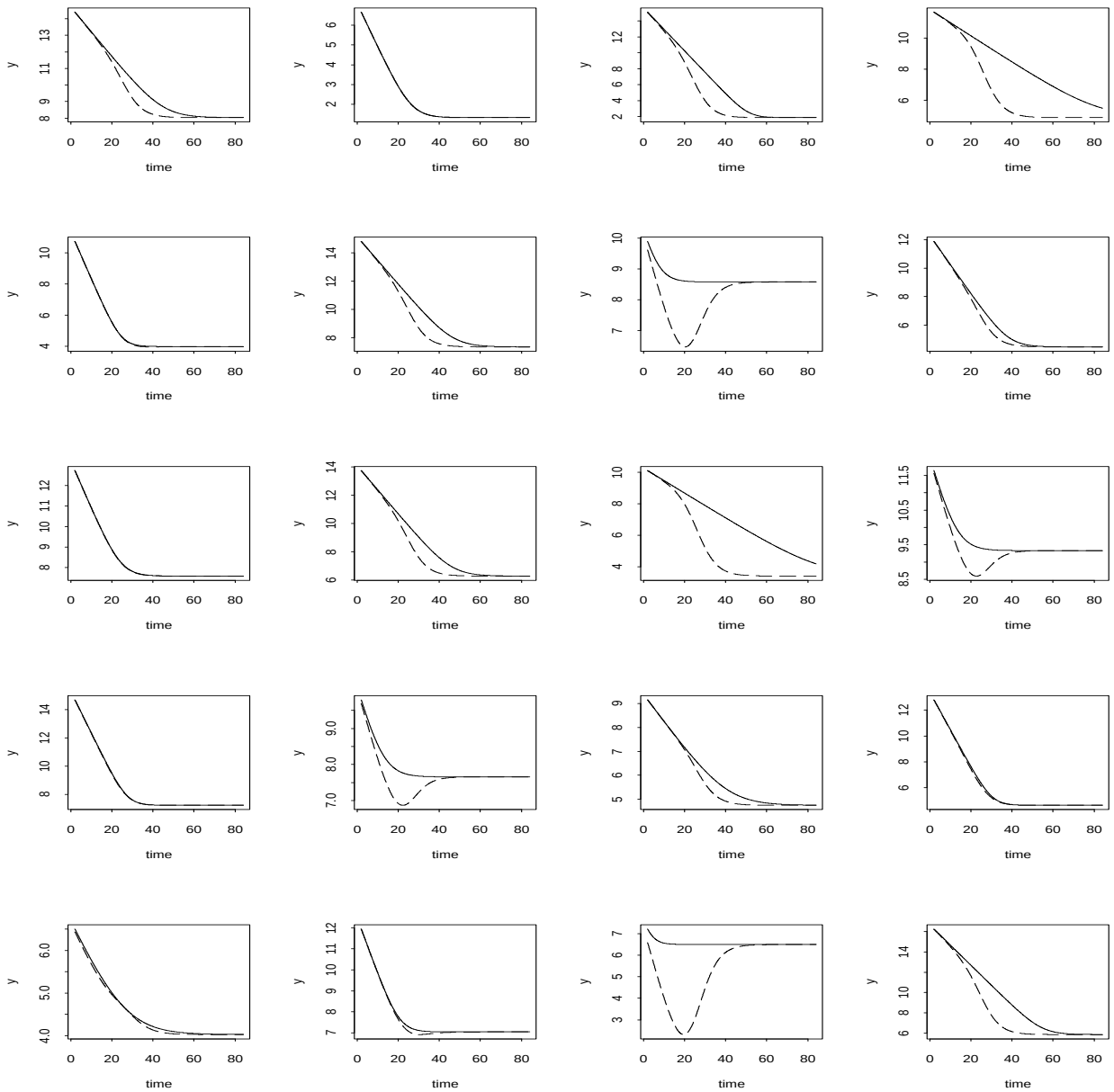
with within subject errors $N(0, \sigma^2)$ and $\boldsymbol{\theta}_i \overset{iid}{\sim} N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Furthermore, let $\boldsymbol{\mu} = (5.530\ 11.023\ 0.216)^T$ and $\boldsymbol{\Sigma} = \text{diag}(1.100^2\ 1.284^2\ 0.035^2)$ and $\sigma^2 = 0.281^2$ (Wu and Ding).

Assume that observations can be taken anywhere between 2 and 84 days and so $\mathcal{T} = [2, 84]$.

The following figures show the true mean response for a 20 random subjects (solid line) and their pseudo-random effects approximations (dotted line) for $\boldsymbol{\Sigma}$ (Fig 1) and $4\boldsymbol{\Sigma}$ (Fig 2).


y
time
0 20 40 60 80


y
time
0 20 40 60 80

Measurements from the same subject are not independent and the information matrix cannot be expressed as a sum of the information matrices for each measurement (or a weighted sum). Exact designs with integer numbers of support points must therefore be examined, not approximate designs.

Let Θ_m be the set of designs supported on m points from $\mathcal{T}$ and take one observation at each point. Let $M(\xi)$ be the Fisher information matrix of the pseudo-random effects model. For values of

$$\begin{aligned} \boldsymbol{\mu} &= (5.530\ 11.023\ 0.216)^T \\ \boldsymbol{\Sigma} &= \text{diag}(1.100^2\ 1.284^2\ 0.035^2) \\ \sigma^2 &= 0.381 \end{aligned}$$

the best value of ξ that maximizes log det $M(\xi)$ can be found for each value of m. This can be called ϕ_m^D-optimality.

For k unknown parameters, define ϕ^D-optimality to be the best choice of ϕ_m^D-optimal design that maximizes log det $M(\xi) - k \log m$ as an indication of "information per measurement".

The following table is for Example 1:

m	Support Points	log det
3	2 22.17 29.40	14.53
4	2 19.70 26.30 84	18.06
5	2 16.90 25.76 25.76 84	19.15
6	2 2 16.77 25.57 25.57 84	19.94
7	2, 2, 16.90 26.05 26.05 84 84	20.50
8	2 2 17.14 17.14 26.60 26.60 84 84	20.96

Note that the value of the criterion ϕ_m^D at the optimal design increases with m, as it should.

Also consider n, the number of subjects that are required to give an information matrix with the log determinant of 40, and the value of $N = nm$.

m	Support Points	log det	n	N
3	2 22.17 29.40	14.53	38.03	114.10
4	2 19.70 26.30 84	18.06	22.96	91.84
5	2 16.90 25.76 25.76 84	19.15	19.65	98.24
6	2 2 16.77 25.57 25.57 84	19.94	17.57	105.43
7	2, 2, 16.90 26.05 26.05 84 84	20.50	16.21	113.48
8	2 2 17.14 17.14 26.60 26.60 84 84	20.96	15.13	121.03

$m = 4$ is the best use of resources if measurements are expensive.

How well does the information for the approximate model approximate the actual value of the information matrix?

Using adaptive Gaussian quadrature we can calculate the information matrix – for simple models only.

Consider a very simple model

$$r(\boldsymbol{\theta}_i, t_j) = \exp(-\theta_i t_j)$$

So

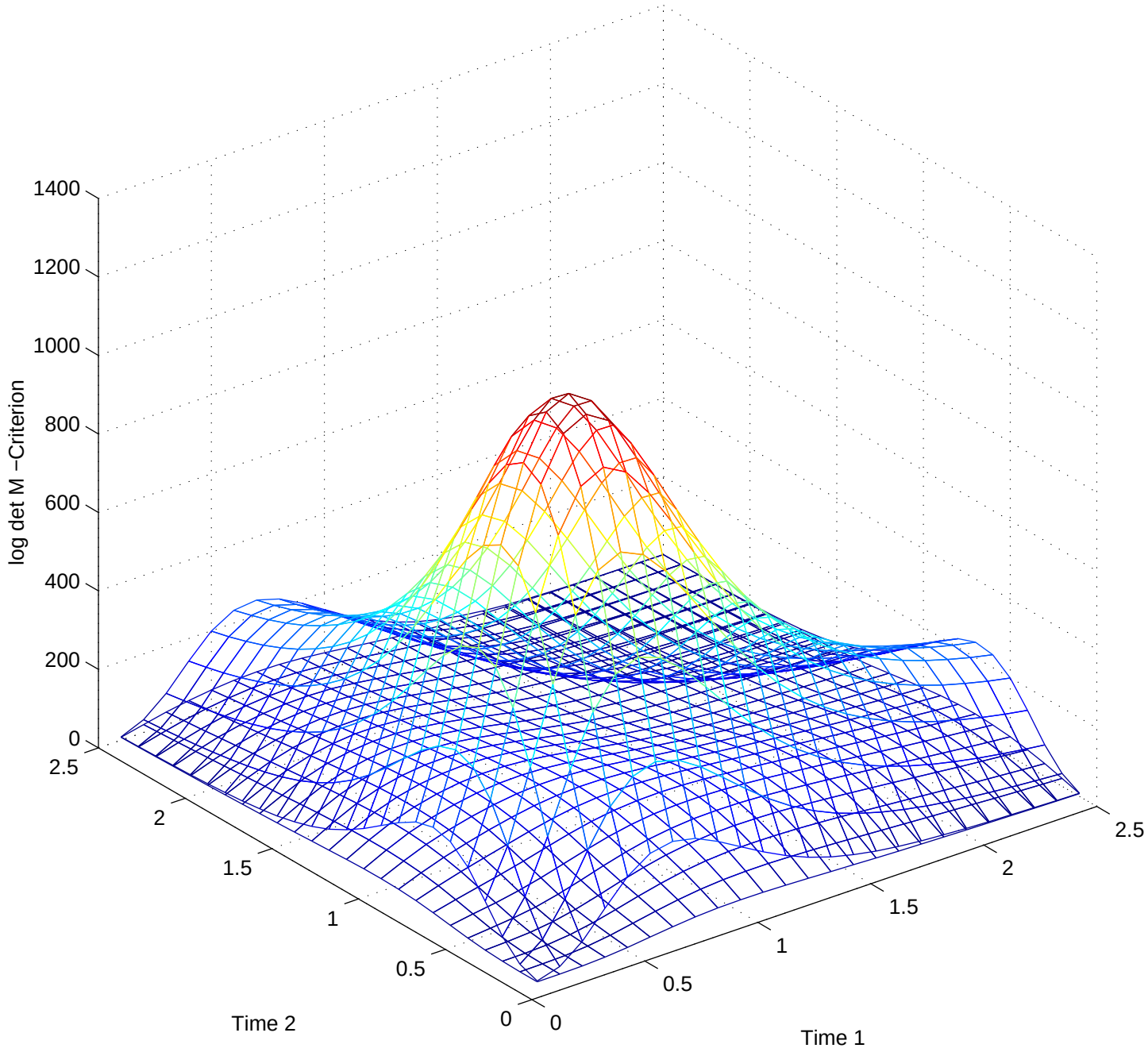
$$y_{ij}|\theta_i \overset{iid}{\sim} N(\exp(\theta_i t_j, \sigma^2)$$

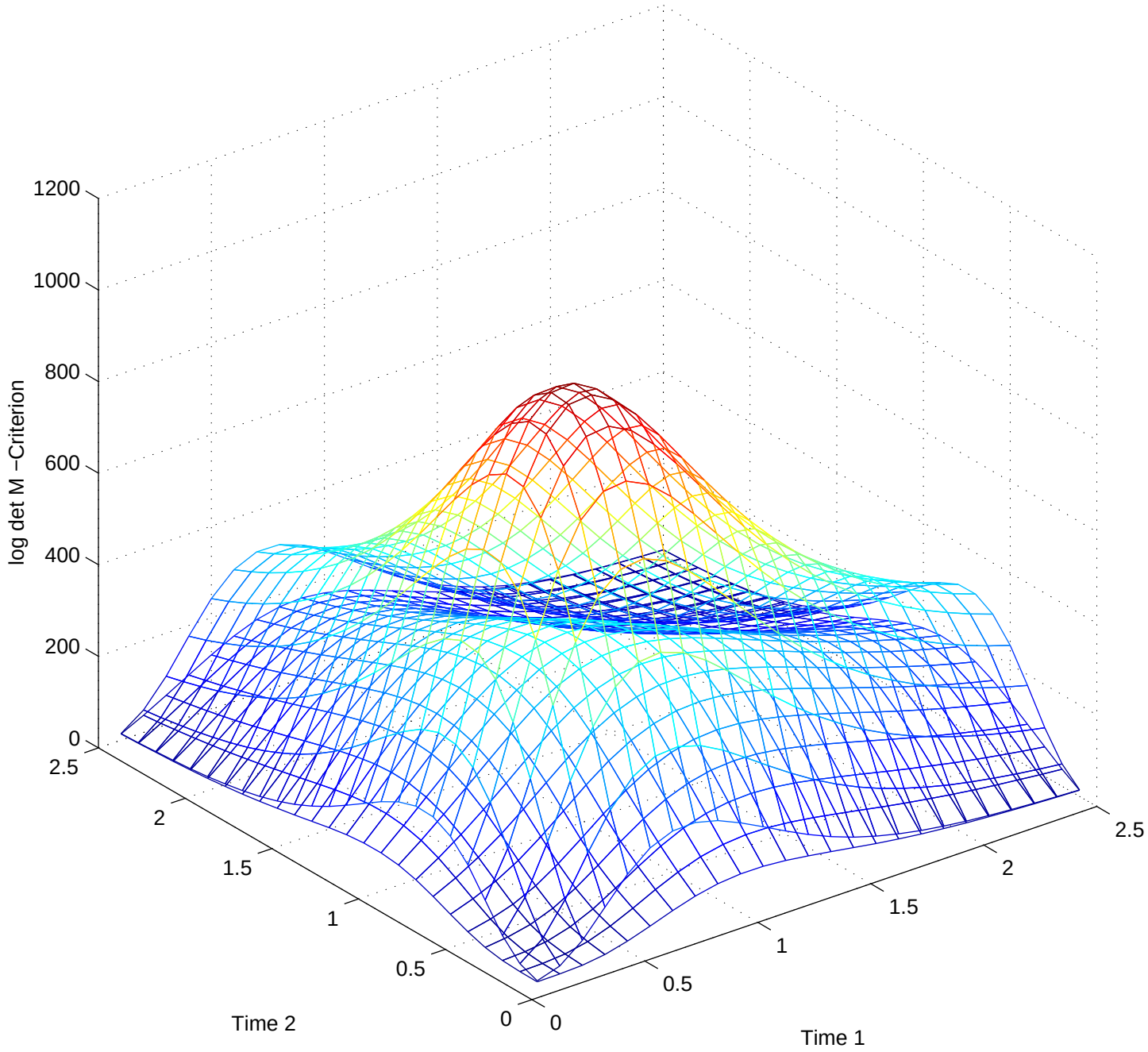
and

$$\theta_i \overset{iid}{\sim} N(\mu, \tau^2).$$

Let $\mu = 1.5$, $\sigma^2 = 0.05$ and $\tau^2 = 0.03$ or 0.09. Consider $m = 2$ and observations lie in $[0.1, 2.5]$. on a grid with mesh 0.1. The Fisher information matrix is 3×3 and can be computed for each design on the grid. Compute log det(M_{11}) or log det(M) or $\log(M^{-1})^{-1}_{11}$.

The following two figures show the true value and the difference between the value computed from the pseudo random effects model, $\tau^2 = 0.03$ (Fig 3) and $\tau^2 = 0.09$ (Fig 4).


1400
1200
1000
800
600
400
200
0
log det M −Criterion
2.5
2
1.5
1
0.5
0
Time 2
0
0.5
1
1.5
2
2.5
Time 1


1200
1000
800
600
400
200
0
log det M −Criterion
2.5
2
1.5
1
0.5
0
Time 2
0
0.5
1
1.5
2
2.5
Time 1

In this and in another example (Merle and Tod, 2001) the design that maximizes the true criterion is almost identical to that which maximizes the criterion based on the approximate model for fixed m. However the difference between them can be large and choosing the best m may be a problem using the approximation.

1. The pseudo-random effects model and information matrix may not well approximate the true.

2. The examples examined with exact calculations are very simple, more realistic models cannot be computed. Values $m > 4$ is challenging.

3. The pseudo-random effects model may provide some candidate designs for exact m for further consideration.

4. The Bayesian methods based on simulations seem a reasonable and pragmatic approach to select between a small set of candidate designs.

Our conclusion is that evaluating the designs by simulating data from a predictive distribution and examining the distribution of the posterior loss from a candidate set of designs is a feasible and pragmatic solution.

See Nowak and May (2000). *Virus Dynamics*, Oxford.

Tan and Wu (2005). *Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention*, World Scientific.