Using Hamiltonian Monte Carlo to design clinical trials with longitudinal data

Florence Loingeville, Thu Thuy Nguyen, Marie-Karelle Riviere, Giulia Lestini, Sebastian Ueckert, France Mentré

IAME, UMR 1137 INSERM - University Paris Diderot, Paris, France

IDeAl Webinar - WP5, October 27, 2016
Contents

1 Introduction

2 New methods for computation of FIM and Applications
   - Methods
   - Evaluation by CTS
   - Illustration in D-optimal designs for binary and count data

3 Extension of methods for Robust designs and Applications
   - Methods
   - Illustration in Robust optimal designs for count data

4 Discussion
## Contents

1. Introduction

2. New methods for computation of FIM and Applications
   - Methods Evaluation by CTS
   - Illustration in D-optimal designs for binary and count data

3. Extension of methods for Robust designs and Applications
   - Methods
   - Illustration in Robust optimal designs for count data

4. Discussion
Designs in pharmacometrics

- Last decades: several methods/software for **maximum likelihood estimation** of population parameters from **longitudinal data** using **nonlinear mixed effect models** (NLMEM)

- Problem beforehand: **choice of "population" design**
  - To obtain precise estimates / adequate power
    - number of individuals (N) ?
    - number of sampling times/individual (n)?
    - allocation of sampling times?
    - other design variables (doses, etc.)

- **Clinical trial simulation** (CTS): time consuming

- Asymptotic theory: **expected Fisher Information Matrix** \(^1\) (FIM)

---

Fisher Information Matrix in NLMEM

- **From FIM**
  - Derive predicted Relative Standard Errors (RSE) and/or power
  - Compare and/or optimise designs

- **Analytical expression for FIM in NLMEM**
  - Current approach in PFIM\(^2\) and other design software programs\(^3\): first order linearisation of model around the expectation of random effects (FO)
    - Only for continuous data
    - Performs well but has limitations in case of complex nonlinear models and/or large variability

- **New approaches needed for computation of FIM**
  - Without model linearisation
  - For both continuous and discrete data
    - Monte Carlo - Adaptive Gaussian Quadrature (MC-AGQ)\(^4\)\(^,\)\(^5\)
    - Monte Carlo - Hamiltonian Monte Carlo (MC-HMC)\(^6\)

---

Parameter and model uncertainty in designs

- **Optimal design depends on knowledge on model and parameters**
  - Local planification: given the model $m$ and parameter values $\Psi^*_m$
  - Widely used criterion: D-optimality

- **Alternative: Robust designs**
  - Taking into account uncertainty on parameters
  - Across a set of candidate models
Contents

1 Introduction

2 New methods for computation of FIM and Applications
   - Methods
   - Evaluation by CTS
   - Illustration in D-optimal designs for binary and count data

3 Extension of methods for Robust designs and Applications
   - Methods
   - Illustration in Robust optimal designs for count data

4 Discussion
Contents

1. Introduction

2. New methods for computation of FIM and Applications
   - Methods
   - Evaluation by CTS
   - Illustration in D-optimal designs for binary and count data

3. Extension of methods for Robust designs and Applications
   - Methods
   - Illustration in Robust optimal designs for count data

4. Discussion
New methods for computation of FIM in NLMEM

**Population FIM** for one group design: \( M(\Psi, \Xi) = N \times M(\Psi, \xi) \)

Population design \( \Xi = \{\xi, N\} \) with identical elementary design \( \xi \) in all \( N \) subjects

**Elementary FIM:** \( M(\Psi, \xi) = E_y \left( \frac{\partial \log(L(y,\Psi))}{\partial \Psi} \frac{\partial \log(L(y,\Psi))}{\partial \Psi}^T \right) \)

with the **likelihood:**

\[ L(y, \Psi) = \int p(y|b, \Psi) p(b|\Psi) \, db \]

where  
\( p(y|b, \Psi) \): pdf of observations \( y \) given random effects \( b \)
\( p(b|\Psi) \): pdf of \( b \)

⇒ **Two integrals to compute:** w.r.t. \( y \) and w.r.t. \( b \)

- Use of **MC** and **AGQ**
- Use of **MC** and **HMC (in Stan)**

⇒ Both approaches **evaluated by CTS** on several examples (from)

---

Evaluation by CTS: Example of binary response

Logistic model for repeated binary response at several time points with treatment increasing the slope of the logit of the response with time $^5,^6,^9$

$$\text{logit}(\pi) = \beta_1 + \beta_2 (1 + \mu_3 \delta)t,$$

where

- $\pi$ is the probability of success
- $\beta_p = \mu_p + b_p$; $b_p \sim N(0, \omega_p^2)$
- $t$: time among 13 points equally spaced between 0 and 12 months
- 2 treatment groups ($\delta = 0$ & $\delta = 1$)
- $N = 50$ subjects/treatment group

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\Psi^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>-2</td>
</tr>
<tr>
<td>$\mu_2$ (month$^{-1}$)</td>
<td>0.09</td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>5</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>0.70</td>
</tr>
<tr>
<td>$\omega_2$ (month$^{-1}$)</td>
<td>0.17</td>
</tr>
</tbody>
</table>


Evaluation by CTS: Example of count response

Poisson model for repeated count response at several dose levels with a full Imax model describing the relationship between log(λ) and dose \(^5,6\)

\[
P(y = k|b) = \frac{\lambda^k \exp(-\lambda)}{k!}
\]

with log(\(\lambda\)) = \(\beta_1\left(1 - \frac{d}{d + \beta_2}\right)\)

- \(\beta_p = \mu_p \exp(b_p); \ b_p \sim \mathcal{N}(0, \omega_p^2)\)
- \(d\): dose among 3 levels (0, 0.4, 0.7)
- \(N = 20\) subjects, \(n_{\text{rep}} = 30\) replications/subject/dose

<table>
<thead>
<tr>
<th>Parameters</th>
<th>(\Psi^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu_1)</td>
<td>1</td>
</tr>
<tr>
<td>(\mu_2)</td>
<td>0.5</td>
</tr>
<tr>
<td>(\omega_1)</td>
<td>0.3</td>
</tr>
<tr>
<td>(\omega_2)</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Evaluation by CTS: Methods

Comparison of several approaches for evaluation of FIM:

- MC-HMC implemented in R package *MIXFIM* available on CRAN
  - 1000 MC / 200 HMC with 500 burn
  - 1000 MC / 1000 HMC with 1000 burn
  - 5000 MC / 200 HMC with 500 burn
  - 5000 MC / 1000 HMC with 1000 burn

- MC-AGQ implemented in R: 5000 MC / 10 AGQ nodes

- Laplace approximation (LA): 5000 MC / 1 AGQ node

with clinical trial simulations (CTS):

- Simulation of 1000 datasets with \( \Psi = \Psi^* \) using R

- For each dataset: estimate \( \hat{\Psi} \) using Monolix 4.3

in terms of:

- observed RSE and RRMSE from CTS

- *versus* predicted RSE from expected FIM

---

Evaluation by CTS: Results for binary example
Evaluation by CTS: Results for count example
## D-optimal designs for discrete data: Methods

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Binary example</th>
<th>Count example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed design variables</td>
<td>100 subjects, 1 replication, 4 times</td>
<td>60 subjects, 10 replications, 3 doses</td>
</tr>
<tr>
<td>Optimised design variables</td>
<td>(t_2, t_3 \text{ from 1 to 11 (step = 1, no repetition)})</td>
<td>(d_2, d_3 \text{ from 0.1 to 1 (step = 0.1, no repetition)})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Optimisation method</th>
<th>Evaluation of FIM for all possible designs</th>
<th>500 Quasi MC (^{11}), 3 AGQ nodes</th>
<th>5000 MC, 200 HMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-optimality criterion (\Phi_D)</td>
<td>(\det(\mathcal{M}(\Psi^*, \Xi))^{1/P})</td>
<td>(\det(\mathcal{M}(\Psi^*, \Xi))^{1/P})</td>
<td></td>
</tr>
</tbody>
</table>

D-optimal designs for discrete data: Results

Optimal times: $\xi_D = (0, 2, 3, 12)$.
D-optimal designs for discrete data: Results

Optimal doses: $\xi_D = (0, 0.4, 0.5)$. 
New methods developed for computation of FIM avoiding FO

- MC-AGQ and MC-HMC based methods
  - adapted for continuous and discrete NLMEM
  - high agreement with CTS
  - new tool for designs using MC-HMC: R package MIXFIM on CRAN
- Enable first applications to design optimisation for binary and count data
Contents

1. Introduction

2. New methods for computation of FIM and Applications
   - Methods
   - Evaluation by CTS
   - Illustration in D-optimal designs for binary and count data

3. Extension of methods for Robust designs and Applications
   - Methods
   - Illustration in Robust optimal designs for count data

4. Discussion
Robustness w.r.t. parameters of a given model

- Robust FIM, assuming a distribution $p(\Psi)$ on the parameters

$$\mathcal{M}_R(\Xi) = E_{\Psi}(\mathcal{M}(\Psi, \Xi))$$

- two integrals w.r.t. $y$ and w.r.t. $b$ for evaluation of $\mathcal{M}(\Psi, \Xi)$
- one supplementary integral w.r.t. $\Psi$ for evaluation of $\mathcal{M}_R(\Xi)$

- Evaluation by MC-HMC using Stan (drawing jointly $\Psi$ and $y$ by MC)
- DE-criterion for optimisation of robust design $\Xi_{DE}$

$$\Phi_{DE}(\Xi) = \det(\mathcal{M}_R(\Xi))^{1/P}$$

with $P$, number of population parameters of the model
Methods for Robust designs (2)

Robustness w.r.t. a set of $M$ candidate models

- D-criterion for optimisation of design $\Xi_{D,m}$ for each model $m$ given population parameter values $\Psi_m^*$

$$\Phi_{D,m}(\Xi) = \det(M(\Psi_m^*, \Xi))^{1/P_m}$$

with $P_m$, number of population parameters of model $m$

- Compound D-criterion $^{12, 13}$ for optimisation of common design $\Xi_{CD}$

$$\Phi_{CD}(\Xi) = \prod_{m=1}^{M} \Phi_{D,m}(\Xi)^{\alpha_m} = \prod_{m=1}^{M} \left(\det(M(\Psi_m^*, \Xi))\right)^{\alpha_m/P_m}$$

with $\alpha_m$, weight quantifying the balance between $M$ models, $\sum_m \alpha_m = 1$

Implementation in R

- Extension of MIXFIM for evaluation of robust FIM using MC-HMC
- Use of compound optimality criterion to combine several models

---


$^{13}$ Nguyen et al. *Pharm Stat*, 2016.
Application to design optimisation in the previous count example

- **Robust optimal design accounting for uncertainty on parameters**
  - Using robust FIM (5000 MC - 200 HMC) and DE-optimality criterion
  - Comparison between $\Xi_D$ and $\Xi_{DE}$ in terms of
    - Allocation of optimal doses
    - Relative efficiencies of a design $\Xi$ w.r.t. an optimal design
    \[
    D\text{-eff}(\Xi) = \frac{\Phi_D(\Xi)}{\Phi_D(\Xi_D)} \quad \text{and} \quad \text{DE-eff}(\Xi) = \frac{\Phi_{DE}(\Xi)}{\Phi_{DE}(\Xi_{DE})}
    \]

- **Robust optimal design across $M$ candidate models**
  - Using FIM by MC-HMC (5000 MC - 200 HMC) and compound D-optimality ($\alpha_m = 1/M$)
  - Comparison between different $\Xi_D,m$ and $\Xi_{CD}$ in terms of
    - Allocation of optimal doses
    - Relative efficiencies of a design $\Xi$ w.r.t. an optimal design
    \[
    D\text{-eff}_m(\Xi) = \frac{\Phi_{D,m}(\Xi)}{\Phi_{D,m}(\Xi_{D,m})} \quad \text{and} \quad \text{CD-eff}(\Xi) = \frac{\Phi_{CD}(\Xi)}{\Phi_{CD}(\Xi_{CD})}
    \]
Robust design for count data: uncertainty on parameters

Poisson model for repeated count response at several dose levels with a full Imax model describing the relationship between $\log(\lambda)$ and dose

$$P(y = k|b) = \frac{\lambda^k \exp(-\lambda)}{k!}$$

with $\log(\lambda) = \beta_1 \left( 1 - \frac{d}{d + \beta_2} \right)$

- $\beta_p = \mu_p \exp(b_p); b_p \sim \mathcal{N}(0, \omega^2_p)$
- Assuming uncertainty on parameters $\mu_2$ and $\omega_2$

<table>
<thead>
<tr>
<th>$\Psi^*$</th>
<th>$p(\Psi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>1</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>$\mathcal{LN}(-0.89, 0.63)$</td>
</tr>
<tr>
<td></td>
<td>$E(\mu_2) = 0.5; CV(\mu_2) = 70%$</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>0.3</td>
</tr>
<tr>
<td>$\omega_2$</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>$\mathcal{LN}(-1.50, 0.77)$</td>
</tr>
<tr>
<td></td>
<td>$E(\omega_2) = 0.3; CV(\omega_2) = 90%$</td>
</tr>
</tbody>
</table>

- Optimisation of 3 doses with $N = 60, n_{\text{rep}} = 10$
  - fixing $d_1 = 0$
  - choosing $d_2$ and $d_3$ from 0 to 1
Robust design for count data: uncertainty on parameters

Optimal doses: $\xi_D = (0, 0.4, 0.5)$.

Optimal doses: $\xi_{DE} = (0, 0.2, 0.4)$.

Efficiencies

<table>
<thead>
<tr>
<th>Design $\Xi$</th>
<th>D-eff($\Xi$)</th>
<th>DE-eff($\Xi$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Xi_D$</td>
<td>100%</td>
<td>94.1%</td>
</tr>
<tr>
<td>{N = 60, $\xi = (0, 0.4, 0.5)$}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Xi_{DE}$</td>
<td>93.3%</td>
<td>100%</td>
</tr>
<tr>
<td>{N = 60, $\xi = (0, 0.2, 0.4)$}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Robust design for count data: 4 candidate models

- Fixed effects $\mu_1, \mu_2$ for M2, M3, M4 chosen to have similar mean value of $\log(\lambda)$ as for M1 at dose 0 and at dose 1
- Variability $\omega_1 = \omega_2 = 0.3$

$$M1: \log(\lambda) = \beta_1 (1 - \frac{d}{d + \beta_2})$$

$$M2: \log(\lambda) = \beta_1 (1 - \beta_2 d)$$

$$M3: \log(\lambda) = \beta_1 (1 - \beta_2 \log(d + 1))$$

$$M4: \log(\lambda) = \beta_1 (1 - \frac{\beta_3 d}{d + \beta_2})$$
Robust design for count data: 4 candidate models

- Optimal doses: $\xi_{D,M1} = (0, 0.4, 0.5)$.
- Optimal doses: $\xi_{D,M2} = (0, 0.9, 1)$.
- Optimal doses: $\xi_{D,M3} = (0, 0.9, 1)$.
- Optimal doses: $\xi_{D,M4} = (0, 0.2, 1)$.
Robust design for count data: 4 candidate models

\[ D\text{-eff}_m(\Xi) = \frac{\Phi_{D,m}(\Xi)}{\Phi_{D,m}(\Xi_{D,m})} \]

<table>
<thead>
<tr>
<th>Design $\Xi$</th>
<th>$D\text{-eff}_{M1} (\Xi)$</th>
<th>$D\text{-eff}_{M2} (\Xi)$</th>
<th>$D\text{-eff}_{M3} (\Xi)$</th>
<th>$D\text{-eff}_{M4} (\Xi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Xi_{D,M1}$ {(N = 60, \xi = (0, 0.4, 0.5)}}</td>
<td>100%</td>
<td>60.8%</td>
<td>68.9%</td>
<td>50.3%</td>
</tr>
<tr>
<td>$\Xi_{D,M2}$ {(N = 60, \xi = (0, 0.9, 1)}}</td>
<td>87.0%</td>
<td>100%</td>
<td>100%</td>
<td>30.8%</td>
</tr>
<tr>
<td>$\Xi_{D,M3}$ {(N = 60, \xi = (0, 0.9, 1)}}</td>
<td>87.0%</td>
<td>100%</td>
<td>100%</td>
<td>30.8%</td>
</tr>
<tr>
<td>$\Xi_{D,M4}$ {(N = 60, \xi = (0, 0.2, 1)}}</td>
<td>88.4%</td>
<td>85.7%</td>
<td>85.4%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Important loss of efficiency in some scenarios where the model is not correctly pre-specified.
Robust design for count data: 4 candidate models

\[ D\text{-efficiencies} \]

\[ D\text{-eff}_m(\Xi) = \frac{\Phi_{D,m}(\Xi)}{\Phi_{D,m}(\Xi_{D,m})} \]

<table>
<thead>
<tr>
<th>Design ( \Xi )</th>
<th>D-eff(_{M1}) (( \Xi ))</th>
<th>D-eff(_{M2}) (( \Xi ))</th>
<th>D-eff(_{M3}) (( \Xi ))</th>
<th>D-eff(_{M4}) (( \Xi ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Xi_{D,M1} ) {( N = 60, \xi = (0, 0.4, 0.5) )}</td>
<td>100%</td>
<td>60.8%</td>
<td>68.9%</td>
<td>50.3%</td>
</tr>
<tr>
<td>( \Xi_{D,M2} ) {( N = 60, \xi = (0, 0.9, 1) )}</td>
<td>87.0%</td>
<td>100%</td>
<td>100%</td>
<td>30.8%</td>
</tr>
<tr>
<td>( \Xi_{D,M3} ) {( N = 60, \xi = (0, 0.9, 1) )}</td>
<td>87.0%</td>
<td>100%</td>
<td>100%</td>
<td>30.8%</td>
</tr>
<tr>
<td>( \Xi_{D,M4} ) {( N = 60, \xi = (0, 0.2, 1) )}</td>
<td>88.4%</td>
<td>85.7%</td>
<td>85.4%</td>
<td>100%</td>
</tr>
</tbody>
</table>

- Important loss of efficiency in some scenarios where the model is not correctly pre-specified
Robust design for count data: 4 candidate models

**Compound D-optimal design:** $\xi_{CD} = (0, 0.2, 1)$. 

Combination of 4 models
Robust design for count data: 4 candidate models

\[
D\text{-efficiencies} \quad D\text{-eff}_m(\Xi) = \frac{\Phi_{D,m}(\Xi)}{\Phi_{D,m}(\Xi_{D,m})}
\]

\[
CD\text{-efficiencies} \quad CD\text{-eff}(\Xi) = \frac{\Phi_{CD}(\Xi)}{\Phi_{CD}(\Xi_{CD})}
\]

<table>
<thead>
<tr>
<th>Design $\Xi$</th>
<th>$D\text{-eff}_{M1}(\Xi)$</th>
<th>$D\text{-eff}_{M2}(\Xi)$</th>
<th>$D\text{-eff}_{M3}(\Xi)$</th>
<th>$D\text{-eff}_{M4}(\Xi)$</th>
<th>$CD\text{-eff}(\Xi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Xi_{D,M1}$</td>
<td>100%</td>
<td>60.8%</td>
<td>68.9%</td>
<td>50.3%</td>
<td>75.5%</td>
</tr>
<tr>
<td>${N = 60, \xi = (0, 0.4, 0.5)}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Xi_{D,M2}$</td>
<td>87.0%</td>
<td>100%</td>
<td>100%</td>
<td>30.8%</td>
<td>80.2%</td>
</tr>
<tr>
<td>${N = 60, \xi = (0, 0.9, 1)}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Xi_{D,M3}$</td>
<td>87.0%</td>
<td>100%</td>
<td>100%</td>
<td>30.8%</td>
<td>80.2%</td>
</tr>
<tr>
<td>${N = 60, \xi = (0, 0.9, 1)}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Xi_{D,M4}$</td>
<td>88.4%</td>
<td>85.7%</td>
<td>85.4%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>${N = 60, \xi = (0, 0.2, 1)}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Xi_{CD}$</td>
<td>88.4%</td>
<td>85.7%</td>
<td>85.4%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>${N = 60, \xi = (0, 0.2, 1)}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Good performance of the compound D-optimal design
Conclusion (2)

- **Proposed methods for Robust designs**
  - Extension of R package *MIXFIM* to compute the DE-optimality criterion from robust FIM
  - Use of compound optimality criterion to combine several candidate models

- **MC-HMC, relevant approach allowing for the first time robust design optimisation for repeated count data**
  - Robustness w.r.t. parameters: different optimal designs with versus without uncertainty on parameters
  - Robustness w.r.t. models: compound D-optimal design providing a good compromise for different candidate models

- **Ongoing work**
  - Robustness w.r.t. parameters AND models: use of robust FIM in the compound optimality criterion
Contents

1 Introduction

2 New methods for computation of FIM and Applications
   - Methods
   - Evaluation by CTS
   - Illustration in D-optimal designs for binary and count data

3 Extension of methods for Robust designs and Applications
   - Methods
   - Illustration in Robust optimal designs for count data

4 Discussion


Discussion

Summary

- New methods developed for computation of FIM avoiding FO
  - MC-AGQ and MC-HMC: relevant methods for designs
  - New tool for designs using MC-HMC: R package MIXFIM on CRAN
  - Computationally challenging, much slower than FO approach

- Extension of these methods to propose robust optimal designs accounting for uncertainty w.r.t. parameters and/or models

Perspectives

- Replacement of MC in MC-HMC by more efficient approach: quasi-random sampling

- Evaluation of two-stage designs
  - Approaches already proposed and evaluated on continuous data
  - To be evaluated in models for discrete data, accounting for uncertainty w.r.t. parameters and/or models

Individual and Population Bayesian information matrix

Back-up
Introduction

Methods for computation of FIM and Applications

Methods for Robust designs and Applications

Discussion

NLMEM: Notations

For continuous data:

\[ y_i = f(g(\mu, b_i), \xi_i) + \epsilon_i \]

For discrete data:

\[ p(y_i|b_i) = \prod_{j=1}^{n_i} h(y_{ij}, g(\mu, b_i), \xi_i) \]

with

\[ y_i = (y_{i1}, \ldots, y_{in_i})^T \] response for individual \( i (i = 1, \ldots, N) \)

\( f, h \) structural model

\( \xi_i \) elementary design for subject \( i \)

\( \beta_i = g(\mu, b_i) \) individual parameters vector

\( \mu \) vector of fixed effects

\( b_i \) vector of random effects for individual \( i \), \( b_i \sim \mathcal{N}(0, \Omega) \)

\( \epsilon_i \) vector of residual errors, \( \epsilon_i \sim \mathcal{N}(0, \Sigma) \) and \( \Sigma \) diagonal matrix

\( \Psi \): Population parameters \( (\mu, \omega, \sigma) \)

\[ p(y_i|b_i) = \mathcal{N}(f, \Sigma) \]
Fisher Information Matrix (FIM)

Population FIM for one group design: $\mathcal{M}(\Psi, \Xi) = N \times \mathcal{M}(\Psi, \xi)$
Population design $\Xi = \{\xi, N\}$ with identical elementary design $\xi$ in all $N$ subjects

Elementary FIM:

$\mathcal{M}(\psi, \xi) = E_y \left( \frac{\partial \log(L(y,\psi))}{\partial \psi} \frac{\partial \log(L(y,\psi))}{\partial \psi}^T \right)$

with the likelihood:

$L(y, \psi) = \int p(y|b, \psi) p(b|\psi) db$

where $p(y|b, \psi)$: pdf of $y$ given the random effects $b$
$p(b|\psi)$: pdf of $b$
MC-HMC method for FIM evaluation

\[ M(\psi, \xi) = E_y \left( \frac{\partial \log(L(y,\psi))}{\partial \psi} \frac{\partial \log(L(y,\psi))}{\partial \psi}^T \right) \]
MC-HMC method for FIM evaluation

\[ M(\psi, \xi) = E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi} \frac{\partial \log(L(y, \psi))}{\partial \psi}^T \right) \]

\[ M(\psi, \xi)_{k,l} = E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi_k} \frac{\partial \log(L(y, \psi))}{\partial \psi_l}^T \right) \begin{bmatrix} D_y \end{bmatrix} \]

Monte Carlo - MC
MC-HMC method for FIM evaluation

\[
\mathcal{M}(\psi, \xi) = E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi} \frac{\partial \log(L(y, \psi))}{\partial \psi}^T \right)
\]

\[
\mathcal{M}(\psi, \xi)_{k,l} = E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi_k} \frac{\partial \log(L(y, \psi))}{\partial \psi_l}^T \right)_{D_y}
\]

Monte Carlo - MC

After calculation... \( D_y \iff \)

\[
\int_{b_1} \frac{\partial \left( \log(p(y|b_1, \psi)p(b_1|\psi)) \right)}{\partial \psi_k} \frac{p(y|b_1, \psi)p(b_1|\psi)}{\int p(y|b, \psi)p(b|\psi) db} db_1 \cdot \int_{b_2} \frac{\partial \left( \log(p(y|b_2, \psi)p(b_2|\psi)) \right)}{\partial \psi_l} \frac{p(y|b_2, \psi)p(b_2|\psi)}{\int p(y|b, \psi)p(b|\psi) db} db_2
\]
MC-HMC method for FIM evaluation

\[
\mathcal{M}(\psi, \xi) = E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi} \frac{\partial \log(L(y, \psi))}{\partial \psi}^T \right)
\]

\[
\mathcal{M}(\psi, \xi)_{k,l} = E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi_k} \frac{\partial \log(L(y, \psi))}{\partial \psi_l}^T \right)
\]

Monte Carlo - MC

After calculation... \( D_y \leftrightarrow \)

\[
\int_{b_1} \frac{\partial (\log(p(y|b_1, \psi)p(b_1|\psi)))}{\partial \psi_k} \frac{p(y|b_1, \psi)p(b_1|\psi)}{\int p(y|b, \psi)p(b|\psi)db} \cdot \int_{b_2} \frac{\partial (\log(p(y|b_2, \psi)p(b_2|\psi)))}{\partial \psi_l} \frac{p(y|b_2, \psi)p(b_2|\psi)}{\int p(y|b, \psi)p(b|\psi)db} \]

conditional density of \( b \) given \( y \)

\[
\int_{b_1} \frac{\partial (\log(p(y|b_1, \psi)p(b_1|\psi)))}{\partial \psi_k} \frac{p(y|b_1, \psi)p(b_1|\psi)}{\int p(y|b, \psi)p(b|\psi)db} \cdot \int_{b_2} \frac{\partial (\log(p(y|b_2, \psi)p(b_2|\psi)))}{\partial \psi_l} \frac{p(y|b_2, \psi)p(b_2|\psi)}{\int p(y|b, \psi)p(b|\psi)db} \]

conditional density of \( b \) given \( y \)
MC-HMC method for FIM evaluation

\[ \mathcal{M}(\psi, \xi) = E_y \left( \frac{\partial \log(L(y,\psi))}{\partial \psi} \frac{\partial \log(L(y,\psi))}{\partial \psi}^T \right) \]

\[ \mathcal{M}(\psi, \xi)_{k,l} = E_y \left( \frac{\partial \log(L(y,\psi))}{\partial \psi_k} \frac{\partial \log(L(y,\psi))}{\partial \psi_l}^T \right) \]

Monte Carlo - MC

After calculation... \( D_y \leftrightarrow \)

\[ \int_{b_1} \frac{\partial (\log(p(y|b_1,\psi)p(b_1|\psi)))}{\partial \psi_k} \frac{p(y|b_1,\psi)p(b_1|\psi)}{\int p(y|b,\psi)p(b|\psi)db} \, db_1 \cdot \int_{b_2} \frac{\partial (\log(p(y|b_2,\psi)p(b_2|\psi)))}{\partial \psi_l} \frac{p(y|b_2,\psi)p(b_2|\psi)}{\int p(y|b,\psi)p(b|\psi)db} \, db_2 \]

conditional density of \( b \) given \( y \)

\[ E \left( \frac{\partial (\log(p(y|b,\psi)p(b|\psi)))}{\partial \psi_k} \right| Y \), \( E \left( \frac{\partial (\log(p(y|b,\psi)p(b|\psi)))}{\partial \psi_l} \right| Y \)

Markov Chains Monte Carlo - MCMC
The \((k, l)\) term of the FIM estimated as:

\[
\tilde{\mathcal{M}}(\psi,\xi)_{k,l} = \frac{1}{R} \sum_{r=1}^{R} A_{k,r}^{(1)} A_{l,r}^{(2)}
\]

with

\[
A_{k,r}^{(1)} = \frac{1}{M} \sum_{m=1}^{M} \frac{\partial}{\partial \psi_k} \left( \log(p(y_r|b_{m,r}^{(1)}, \psi) p(b_{m,r}^{(1)})) \right)
\]

\[
A_{l,r}^{(2)} = \frac{1}{M} \sum_{m=1}^{M} \frac{\partial}{\partial \psi_l} \left( \log(p(y_r|b_{m,r}^{(2)}, \psi) p(b_{m,r}^{(2)})) \right)
\]

where

- \((y_r)_{r=1,...,R}\) is a \(R\)-sample of the marginal distribution of \(y\) \((MC)\)
- \((b_{m,r}^{(1)})_{m=1,...,M}\) and \((b_{m,r}^{(2)})_{m=1,...,M}\) are \(2RM\)-samples of the conditional density of \(b\) given \(y_r\) \((HMC)\)

To be symmetric \(\Rightarrow \hat{\mathcal{M}}(\psi,\xi) = \frac{\mathcal{M}(\psi,\xi)+\mathcal{M}(\psi,\xi)^T}{2}\)
MC-HMC method for Robust FIM evaluation

Robust FIM: $\mathcal{M}_R(\xi) = E_\psi (\mathcal{M}(\Psi, \xi))$

$\mathcal{M}_R(\xi) = E_\psi \left( E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi} \frac{\partial \log(L(y, \psi))}{\partial \psi}^T \right) \right)$

$\mathcal{M}_R(\xi)_{k,l} = E_\psi \left( E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi_k} \frac{\partial \log(L(y, \psi))}{\partial \psi_l}^T \right) \right)
\left( \begin{array}{c} \partial \psi_k \\ \partial \psi_l \end{array} \right)
D_y$

Monte Carlo - MC - joint sampling of $\psi$ and $y$
MC-HMC method for Robust FIM evaluation

\[ M_R(\xi) = E_{\psi} \left( E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi} \frac{\partial \log(L(y, \psi))}{\partial \psi}^T \right) \right) \]

\[ M_R(\xi)_{k,l} = E_{\psi} \left( E_y \left( \frac{\partial \log(L(y, \psi))}{\partial \psi_k} \frac{\partial \log(L(y, \psi))}{\partial \psi_l}^T \right) \right) \]

Monte Carlo - MC - joint sampling of \( \psi \) and \( y \)

\[ D_y \iff E \left( \frac{\partial (\log(p(y|b, \psi)p(b|\psi)))}{\partial \psi_k} \mid Y \right) . E \left( \frac{\partial (\log(p(y|b, \psi)p(b|\psi)))}{\partial \psi_l} \mid Y \right) \]

Markov Chains Monte Carlo - MCMC
The \((k, l)\) term of the FIM estimated as:

\[
\hat{\mathcal{M}}_R(\xi)_{k,l} = \frac{1}{R} \sum_{r=1}^{R} B^{(1)}_{k,r} B^{(2)}_{l,r}
\]

with

\[
B^{(1)}_{k,r} = \frac{1}{M} \sum_{m=1}^{M} \frac{\partial}{\partial \psi_k} \left( \log(p(y_r|b^{(1)}_{m,r}, \psi_r) p(b^{(1)}_{m,r}, \psi_r)) \right)
\]

\[
B^{(2)}_{l,r} = \frac{1}{M} \sum_{m=1}^{M} \frac{\partial}{\partial \psi_l} \left( \log(p(y_r|b^{(2)}_{m,r}, \psi_r) p(b^{(2)}_{m,r}, \psi_r)) \right)
\]

where

- \((\Psi_r, y_r)_{r=1,...,R}\) is a \(R\)-sample of the joint distribution of \((\Psi, y)\) \((MC)\)
- \((b^{(1)}_{m,r})_{m=1,...,M}\) and \((b^{(2)}_{m,r})_{m=1,...,M}\) are \(2R M\)-samples of the conditional density of \(b\) given \(y_r\) \((HMC)\)

To be symmetric \(\Rightarrow \hat{\mathcal{M}}_R(\xi) = \frac{\hat{\bar{\mathcal{M}}}_R(\xi)+\hat{\bar{\mathcal{M}}}_R(\xi)^T}{2}\)