## Quantifying the Closeness to a Set of Random Curves via the Mean Marginal Likelihood

C. Rommel<sup>1,2</sup>, J. F. Bonnans<sup>1</sup>, B. Gregorutti<sup>2</sup> and P. Martinon<sup>1</sup>

CMAP Ecole Polytechnique - INRIA<sup>1</sup> Safety Line<sup>2</sup>

COMPSTAT - August 28th 2018







イロト イポト イヨト イヨト 二日

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 1 / 27



$$\dot{x}(t) = g(u(t), x(t)) + \varepsilon(t)$$

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood



$$\dot{x}(t) = g(u(t), x(t)) + \varepsilon(t)$$

Optimal Control Problem

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt,$$
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = g(\mathbf{u}(t), \mathbf{x}(t)) + \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$
(OCP)

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood



$$\dot{x}(t) = g(u(t), x(t)) + \varepsilon(t)$$

Optimal Control Problem

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt,$$
  
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = g(\mathbf{u}(t), \mathbf{x}(t)) + \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$
(OCP)

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood



$$\dot{x}(t) = g(u(t), x(t)) + \varepsilon(t)$$

Optimal Control Problem

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt,$$
  
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = \hat{g}(\mathbf{u}(t), \mathbf{x}(t)) + \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$
(OCP)

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood



$$\dot{x}(t) = g(u(t), x(t)) + \varepsilon(t)$$

《曰》《聞》《臣》《臣》 [] 臣.

August 28<sup>th</sup> 2018

2/27

Optimal Control Problem

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt,$$
  
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = \hat{g}(\mathbf{u}(t), \mathbf{x}(t)) + \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$
(OCP)

Use of past data to learn how to control a system efficiently

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood



$$\dot{x}(t) = g(u(t), x(t)) + \varepsilon(t)$$

Optimal Control Problem

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt,$$
  
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = \hat{g}(\mathbf{u}(t), \mathbf{x}(t)) + \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$
(OCP)

Use of past data to learn how to control a system efficiently

#### "Model-based reinforcement learning" - [Recht, 2018]

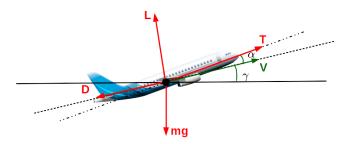
(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 2 / 27

◆□ > ◆□ > ◆三 > ◆三 > ・ 三 · のへの

## FLIGHT OPTIMIZATION

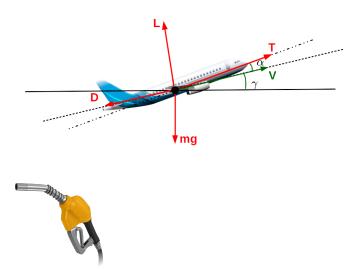


(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

<ロト < 回 ト < 三 ト < 三 ト 三 の へ () August 28<sup>th</sup> 2018 3 / 27

## FLIGHT OPTIMIZATION

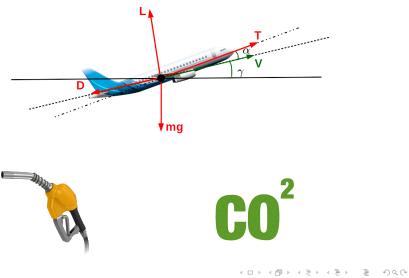


(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

<ロト < 回 ト < 三 ト < 三 ト 三 の へ () August 28<sup>th</sup> 2018 3 / 27

### FLIGHT OPTIMIZATION



(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 3 / 27

#### Dynamics are learned from QAR data

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

<ロト < 回 ト < 画 ト < 三 ト < 三 ト 三 のへで August 28<sup>th</sup> 2018 4 / 27

## Dynamics are learned from QAR data



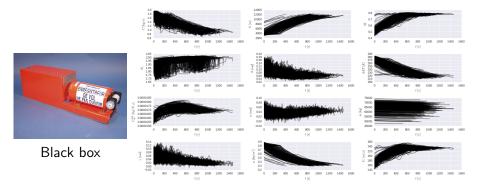
Black box

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

イロト イ団ト イヨト モー つへで August 28<sup>th</sup> 2018 4 / 27

## Dynamics are learned from QAR data



#### Recorded flights = functional data

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

$$\begin{aligned} \min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt, \\ \hat{\mathbf{x}}(t) = \hat{g}(\mathbf{u}(t), \mathbf{x}(t)) = \varepsilon(t), \quad \text{for a.e. } t \in [0, t_{f}], \end{aligned} \tag{OCP}$$
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = \hat{g}(\mathbf{u}(t), \mathbf{x}(t)) = \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

<ロ > < 部 > < 画 > < 画 > < 画 > < 画 > < 画 > < 画 > < つ へ (~ August 28<sup>th</sup> 2018 5 / 27

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_f} C(\mathbf{u}(t), \mathbf{x}(t)) dt,$$
 (OCP)  
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = \hat{g}(\mathbf{u}(t), \mathbf{x}(t)) = \varepsilon(t), & \text{for a.e. } t \in [0, t_f], \\ \text{Other constraints...} \end{cases}$$

$$\Rightarrow \hat{\pmb{z}} = (\hat{\pmb{x}}, \hat{\pmb{u}})$$
 solution of (OCP).

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

< □ > < □ > < □ > < Ξ > < Ξ > < Ξ > Ξ の () August 28<sup>th</sup> 2018 5 / 27

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt,$$
 (OCP)  
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = \hat{g}(\mathbf{u}(t), \mathbf{x}(t)) = \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$

$$\Rightarrow \hat{\pmb{z}} = (\hat{\pmb{x}}, \hat{\pmb{u}})$$
 solution of (OCP).

#### Is $\hat{z}$ inside the validity region of the dynamics model $\hat{g}$ ?

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

<ロト < 部 、 < 注 ト < 注 ト 注 の < で August 28<sup>th</sup> 2018 5 / 27

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt,$$
 (OCP)  
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = \hat{g}(\mathbf{u}(t), \mathbf{x}(t)) = \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$

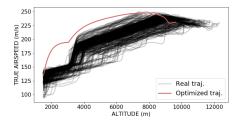
$$\Rightarrow \hat{z} = (\hat{x}, \hat{u})$$
 solution of (OCP).

#### Is $\hat{z}$ inside the validity region of the dynamics model $\hat{g}$ ?

#### Does it look like a real trajectory ?

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

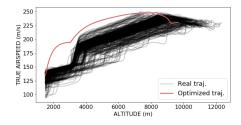


<sup>1</sup>NATS UK air traffic control

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

(ロ ト イ 戸 ト イ 三 ト イ 三 ト 三 今 Q ペ
 August 28<sup>th</sup> 2018 6 / 27





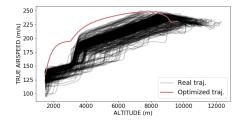
#### Pilots acceptance

<sup>1</sup>NATS UK air traffic control

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

イロト イ団ト イヨト イヨト ヨー のへで August 28<sup>th</sup> 2018 6 / 27





Pilots acceptance

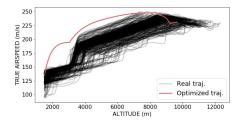


Air Traffic Control<sup>1</sup>

<sup>1</sup>NATS UK air traffic control

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood







#### Pilots acceptance Air Traffic Control<sup>1</sup> How can we quantify the closeness from the optimized trajectory to the set of real flights?

<sup>1</sup>NATS UK air traffic control (CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

- $X \sim f_{\theta^*}$ ,
- x: observation of X,
- Likelihood function of  $\theta$ , given x:

$$\mathcal{L}( heta|x) = f_{ heta}(x).$$

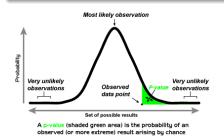
(1)

 <sup>0</sup>Picture source: wikipedia, P-Value, author: Repapetilto: CC. (IP) · (E) · (E) · (E) · (E) · (CMAP, INRIA, SAFETY LINE)
 MEAN MARGINAL LIKELIHOOD
 August 28<sup>th</sup> 2018
 7 / 27

- $X \sim f_{ heta^*}$ ,
- x: observation of X,
- Likelihood function of  $\theta$ , given x:

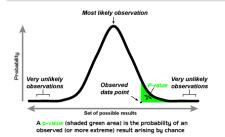
$$\mathcal{L}( heta|x) = f_{ heta}(x).$$

(1)



- $X \sim f_{ heta^*}$ ,
- x: observation of X,
- Likelihood function of  $\theta$ , given x:

$$\mathcal{L}(\theta|x) = f_{\theta}(x).$$



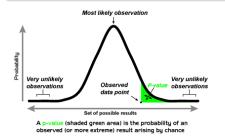
#### In our case:

 the optimized trajectory plays the role of θ, (1)

 <sup>0</sup>Picture source: wikipedia, P-Value, author: Repapetilton CC. → (E) (CMAP, INRIA, SAFETY LINE)
 MEAN MARGINAL LIKELIHOOD
 August 28<sup>th</sup> 2018
 7 / 27

- $X \sim f_{ heta^*}$ ,
- x: observation of X,
- Likelihood function of  $\theta$ , given x:

$$\mathcal{L}(\theta|x) = f_{\theta}(x).$$



#### In our case:

- the optimized trajectory plays the role of θ,
- the set of real flights plays the role of *x*,

(1)

 <sup>0</sup>Picture source: wikipedia, P-Value, author: Repapetilto: CC. The set of the

**Assumption:** We suppose that the real flights are observations of the same functional random variable  $Z = (Z_t)$  valued in  $\mathcal{C}(\mathbb{T}, E)$ , with E compact subset of  $\mathbb{R}^d$  and  $\mathbb{T} = [0, t_f]$ .

Mean Marginal Likelihood

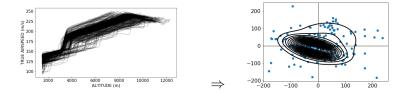
**Assumption:** We suppose that the real flights are observations of the same functional random variable  $Z = (Z_t)$  valued in  $\mathcal{C}(\mathbb{T}, E)$ , with E compact subset of  $\mathbb{R}^d$  and  $\mathbb{T} = [0, t_f]$ .

**Problem:** Computation of probability densities in infinite dimensional space is untractable...

**Assumption:** We suppose that the real flights are observations of the same functional random variable  $Z = (Z_t)$  valued in  $\mathcal{C}(\mathbb{T}, E)$ , with E compact subset of  $\mathbb{R}^d$  and  $\mathbb{T} = [0, t_f]$ .

**Problem:** Computation of probability densities in infinite dimensional space is untractable...

 Standard approach FDA: use FPCA to decompose the data in a small number of coefficients



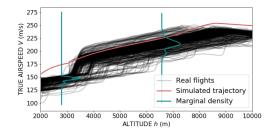
Mean Marginal Likelihood

4 ロ ト 4 日 ト 4 王 ト 王 - つへで August 28<sup>th</sup> 2018 8 / 27

**Assumption:** We suppose that the real flights are observations of the same functional random variable  $Z = (Z_t)$  valued in  $\mathcal{C}(\mathbb{T}, E)$ , with E compact subset of  $\mathbb{R}^d$  and  $\mathbb{T} = [0, t_f]$ .

**Problem:** Computation of probability densities in infinite dimensional space is untractable...

- Standard approach FDA: use FPCA to decompose the data in a small number of coefficients
- Or: we can aggregate the marginal densities



(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

## How do we aggregate the marginal LIKELIHOODS?

(CMAP, INRIA, SAFETY LINE) MEAN MARGINAL LIKELIHOOD

August 28<sup>th</sup> 2018 9 / 27

•  $f_t$  marginal density of Z, i.e. probability density function of  $Z_t$ ,

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

- $f_t$  marginal density of Z, i.e. probability density function of  $Z_t$ ,
- y new trajectory,

- $f_t$  marginal density of Z, i.e. probability density function of  $Z_t$ ,
- y new trajectory,
- $f_t(\mathbf{y}(t))$  marginal likelihood of  $\mathbf{y}$  at t, i.e. likelihood of observing  $Z_t = \mathbf{y}(t)$ .

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □

- $f_t$  marginal density of Z, i.e. probability density function of  $Z_t$ ,
- y new trajectory,
- $f_t(\mathbf{y}(t))$  marginal likelihood of  $\mathbf{y}$  at t, i.e. likelihood of observing  $Z_t = \mathbf{y}(t)$ .

Why not average over time ?...

$$\frac{1}{t_f}\int_0^{t_f}f_t(\boldsymbol{y}(t))dt$$

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 9 / 27

<ロト < 目 > < 目 > < 目 > < 目 > < 目 > < 0 < 0</p>

- $f_t$  marginal density of Z, i.e. probability density function of  $Z_t$ ,
- y new trajectory,
- $f_t(\mathbf{y}(t))$  marginal likelihood of  $\mathbf{y}$  at t, i.e. likelihood of observing  $Z_t = \mathbf{y}(t)$ .

Why not average over time ?...

$$\frac{1}{t_f}\int_0^{t_f}f_t(\boldsymbol{y}(t))dt$$

#### Marginal densities may have really different shapes

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 9 / 27

◆□ > ◆□ > ◆三 > ◆三 > ・ 三 · のへの

- $f_t$  marginal density of Z, i.e. probability density function of  $Z_t$ ,
- y new trajectory,
- $f_t(\mathbf{y}(t))$  marginal likelihood of  $\mathbf{y}$  at t, i.e. likelihood of observing  $Z_t = \mathbf{y}(t)$ .

Mean marginal likelihood

$$\mathsf{MML}(Z, \mathbf{y}) = rac{1}{t_f} \int_0^{t_f} \psi[f_t, \mathbf{y}(t)] dt,$$

where  $\psi: L^1(E, \mathbb{R}_+) \times \mathbb{R} \to [0; 1]$  is a continuous scaling map.

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

4 ロ ト イ 日 ト イ 王 ト イ 王 ト 王 今 Q (で August 28<sup>th</sup> 2018 9 / 27

# How do we aggregate the marginal likelihoods?

Possible scalings are the normalized density

$$\psi[f_t, \mathbf{y}(t)] := rac{\mathbf{y}(t)}{\max_{z \in E} f_t(z)},$$

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

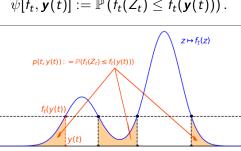
イロト イラト イヨト モークへで August 28<sup>th</sup> 2018 10 / 27

## How do we aggregate the marginal LIKELIHOODS?

Possible scalings are the normalized density

$$\psi[f_t, \boldsymbol{y}(t)] := rac{\boldsymbol{y}(t)}{\displaystyle\max_{z \in E} f_t(z)},$$

or the confidence level



$$\psi[f_t, \mathbf{y}(t)] := \mathbb{P}\left(f_t(Z_t) \leq f_t(\mathbf{y}(t))\right).$$

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

◆□▶ ◆□▶ ◆三▶ ◆三▶ ● ○ ○ ○ August 28<sup>th</sup> 2018 10 / 27

#### How do we deal with sampled curves?

(CMAP, INRIA, SAFETY LINE) MEAN MARGINAL LIKELIHOOD

▲□▶ ▲□▶ ▲三▶ ▲三▶ ▲□ ● ● ● August 28<sup>th</sup> 2018 11 / 27

#### How do we deal with sampled curves?

In practice, the m trajectories are sampled at variable discrete times:

$$\begin{aligned} \mathcal{T}^{D} &:= \{ (t_{j}^{r}, z_{j}^{r}) \}_{\substack{1 \leq j \leq n \\ 1 \leq r \leq m}} \subset \mathbb{T} \times E, \qquad \qquad z_{j}^{r} &:= \mathbf{z}^{r}(t_{j}^{r}), \\ \mathcal{Y} &:= \{ (\tilde{t}_{j}, y_{j}) \}_{j=1}^{\tilde{n}} \subset \mathbb{T} \times E, \qquad \qquad y_{j} &:= \mathbf{y}(\tilde{t}_{j}). \end{aligned}$$

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

イロト イラト イヨト ヨークへで August 28<sup>th</sup> 2018 11 / 27

#### How do we deal with sampled curves?

In practice, the m trajectories are sampled at variable discrete times:

$$\begin{aligned} \mathcal{T}^{D} &:= \{ (t_{j}^{r}, z_{j}^{r}) \}_{\substack{1 \leq j \leq n \\ 1 \leq r \leq m}} \subset \mathbb{T} \times E, \qquad \qquad z_{j}^{r} &:= \mathbf{z}^{r}(t_{j}^{r}), \\ \mathcal{Y} &:= \{ (\tilde{t}_{j}, y_{j}) \}_{j=1}^{\tilde{n}} \subset \mathbb{T} \times E, \qquad \qquad y_{j} &:= \mathbf{y}(\tilde{t}_{j}). \end{aligned}$$

Hence, we approximate the MML using a Riemann sum which aggregates consistent estimators  $\hat{f}_{\tilde{t}_i}^m$  of the marginal densities  $f_{\tilde{t}_j}$ :

$$\mathsf{EMML}_m(\mathcal{T}^D,\mathcal{Y}) := rac{1}{t_f}\sum_{j=1}^{\tilde{n}}\psi[\hat{f}^m_{\tilde{t}_j},y_j]\Delta\tilde{t}_j.$$

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

4 ロ ト 4 部 ト 4 注 ト 注 の 9 で August 28<sup>th</sup> 2018 11 / 27

Suppose that sampling times {t<sub>j</sub><sup>r</sup> : j = 1,..., n; r = 1,..., m} are i.i.d. sampled from r.v. T, indep. Z;

Mean Marginal Likelihood

4 ロ ト 4 日 ト 4 王 ト 4 王 ト 王 今 9 0 0 August 28<sup>th</sup> 2018 12 / 27

- Suppose that sampling times {t<sub>j</sub><sup>r</sup> : j = 1,..., n; r = 1,..., m} are i.i.d. sampled from r.v. T, indep. Z;
- $f_t$  is the density of  $Z_t = (Z_T | T = t) = (Y | X);$

Mean Marginal Likelihood

4 ロ ト 4 回 ト 4 三 ト 4 三 ト 三 からで AUGUST 28<sup>th</sup> 2018 12 / 27

- Suppose that sampling times {t<sub>j</sub><sup>r</sup> : j = 1,..., n; r = 1,..., m} are i.i.d. sampled from r.v. T, indep. Z;
- $f_t$  is the density of  $Z_t = (Z_T | T = t) = (Y | X)$ ;
- Our problem can be seen as a conditional probability density learning problem with  $(X, Y) = (T, Z_T)$ .

<ロト < 目 > < 目 > < 目 > < 目 > < 目 > < 0 < 0</p>

- Suppose that sampling times {t<sub>j</sub><sup>r</sup> : j = 1,..., n; r = 1,..., m} are i.i.d. sampled from r.v. T, indep. Z;
- $f_t$  is the density of  $Z_t = (Z_T | T = t) = (Y | X)$ ;
- Our problem can be seen as a conditional probability density learning problem with  $(X, Y) = (T, Z_T)$ .

 $\Rightarrow$  We could apply SOA conditional density estimation techniques, such as LS-CDE [Sugiyama et al., 2010],

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

#### How can we estimate marginal densities?

- Suppose that sampling times {t<sub>j</sub><sup>r</sup> : j = 1,..., n; r = 1,..., m} are i.i.d. sampled from r.v. T, indep. Z;
- $f_t$  is the density of  $Z_t = (Z_T | T = t) = (Y | X)$ ;
- Our problem can be seen as a conditional probability density learning problem with  $(X, Y) = (T, Z_T)$ .

 $\Rightarrow$  We could apply SOA conditional density estimation techniques, such as LS-CDE [Sugiyama et al., 2010],

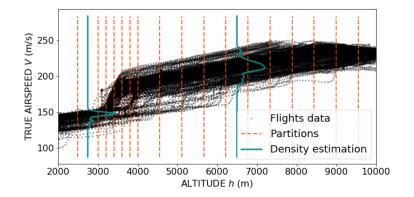
 $\Rightarrow$  Instead, we choose to use a fine partitioning of the time domain.

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

AUGUST 28<sup>th</sup> 2018 12 / 27

#### PARTITION BASED MARGINAL DENSITY ESTIMATION



(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

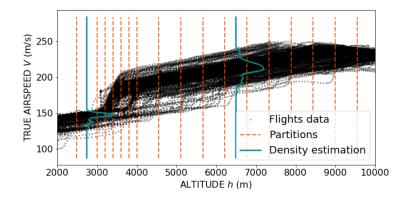
August 28<sup>th</sup> 2018 13 / 27

Э

500

< D > < D >

#### PARTITION BASED MARGINAL DENSITY ESTIMATION



Idea: to average in time the marginal densities over small bins by applying classical multivariate density estimation techniques to each subset.

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 13 / 27

San

We denote by:

- $\Theta: \mathcal{S} \to L^1(E, \mathbb{R}_+)$  multivariate density estimation statistic,
- $S = \{(z_k)_{k=1}^N \in E^N : N \in \mathbb{N}^*\}$  set of finite sequences,

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 14 / 27

<ロ> <同> < 三> < 三> < 三> < 三> < ○<</p>

We denote by:

- $\Theta: \mathcal{S} \to L^1(\mathcal{E}, \mathbb{R}_+)$  multivariate density estimation statistic,
- $\mathcal{S} = \{(z_k)_{k=1}^N \in E^N : N \in \mathbb{N}^*\}$  set of finite sequences,
- *m* the number of random curves;
- $\mathcal{T}_t^m$  subset of data points whose sampling times fall in the bin containing t;

<ロト < 目 > < 目 > < 目 > < 目 > < 目 > < 0 < 0</p>

We denote by:

- $\Theta: \mathcal{S} \to L^1(\mathcal{E}, \mathbb{R}_+)$  multivariate density estimation statistic,
- $\mathcal{S} = \{(z_k)_{k=1}^N \in E^N : N \in \mathbb{N}^*\}$  set of finite sequences,
- *m* the number of random curves;
- $\mathcal{T}_t^m$  subset of data points whose sampling times fall in the bin containing t;
- $\hat{f}_t^m := \Theta[\mathcal{T}_t^m]$  estimator trained using  $\mathcal{T}_t^m$ .

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

<ロト < 部ト < 国ト < 国ト < 国ト 国 の Q (\*) August 28<sup>th</sup> 2018 14 / 27

Assumption 1 - Positive time density  $\nu \in L^{\infty}(E, \mathbb{R}_+)$  density function of *T*, s.t.

$$u_+ := \mathop{\mathrm{ess\,sup}}_{t\in\mathbb{T}} 
u(t) < \infty, \qquad 
u_- := \mathop{\mathrm{ess\,inf}}_{t\in\mathbb{T}} 
u(t) > 0.$$

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

4 ロ ト 4 日 ト 4 目 ト 4 目 ト 目 今 9 0 August 28<sup>th</sup> 2018 15 / 27

Assumption 1 - Positive time density  $\nu \in L^{\infty}(E, \mathbb{R}_+)$  density function of *T*, s.t.

$$u_+ := \mathop{\mathrm{ess\,sup}}_{t\in\mathbb{T}} \nu(t) < \infty, \qquad 
u_- := \mathop{\mathrm{ess\,inf}}_{t\in\mathbb{T}} \nu(t) > 0.$$

ASSUMTION 2 - LIPSCHITZ IN TIME Function  $(t, z) \in \mathbb{T} \times E \mapsto f_t(z)$  is continuous and  $|f_{t_1}(z) - f_{t_2}(z)| \le L|t_1 - t_2|, \qquad L > 0.$ 

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 15 / 27

<ロト < 目 > < 目 > < 目 > < 目 > < 目 > < 0 < 0</p>

Assumption 1 - Positive time density  $\nu \in L^{\infty}(E, \mathbb{R}_+)$  density function of T, s.t.

$$u_+ := \mathop{\mathrm{ess\,sup}}_{t\in\mathbb{T}} 
u(t) < \infty, \qquad 
u_- := \mathop{\mathrm{ess\,inf}}_{t\in\mathbb{T}} 
u(t) > 0.$$

Assumption 2 - Lipschitz in time Function  $(t, z) \in \mathbb{T} \times E \mapsto f_t(z)$  is continuous and  $|f_{t_1}(z) - f_{t_2}(z)| \le L|t_1 - t_2|,$ L > 0.

Assumption 3 - Shrinking Bins The homogeneous partition  $\{B_{\ell}^m\}_{\ell=1}^{q_m}$  of  $[0; t_f]$ , with binsize  $b_m$ , is s.t.

$$\lim_{m\to\infty}b_m=0,\qquad \lim_{m\to\infty}mb_m=\infty.$$

(CMAP, INRIA, SAFETY LINE)

MEAN MARGINAL LIKELIHOOD

・ロト ・ 日 ・ ・ 日 ・ ・ 日 ・ - 日 ・ August 28<sup>th</sup> 2018 15 / 27

nan

Assumption 4 - I.I.D. Consistency

- ${\mathcal G}$  arbitrary family of probability density functions on E,  $ho \in {\mathcal G}$ ,
- $S_{\rho}^{N}$  <u>i.i.d</u> sample of size N drawn from  $\rho$  valued in S.

The estimator obtained by applying  $\Theta$  to  $S_{\rho}^{N}$ , denoted by

$$\hat{\rho}^{N} := \Theta[S^{N}_{\rho}] \in L^{1}(E, \mathbb{R}_{+}),$$

is a (pointwise) consistent density estimator, uniformly in  $\rho$ :

For all  $z \in E, \varepsilon > 0, \alpha_1 > 0$ , there is  $N_{\varepsilon,\alpha_1} > 0$  such that, for any  $\rho \in \mathcal{G}$ ,  $N \ge N_{\varepsilon,\alpha_1} \Rightarrow \mathbb{P}\left(\left|\hat{\rho}^N(z) - \rho(z)\right| < \varepsilon\right) > 1 - \alpha_1.$ 

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

4 ロ ト 4 日 ト 4 王 ト 4 王 ト 王 今 9 0 AUGUST 28<sup>th</sup> 2018 16 / 27

Theorem 1 - [Rommel et al., 2018]

Under assumptions 1 to 4, for any  $z \in E$  and  $t \in \mathbb{T}$ ,  $\hat{f}_{\ell^m(t)}^m(z)$  consistently approximates the marginal density  $f_t(z)$  as the number of curves *m* grows:

$$\forall \varepsilon > 0, \quad \lim_{m \to \infty} \mathbb{P}\left( |\hat{f}_t^m(z) - f_t(z)| < \varepsilon \right) = 1.$$

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 17 / 27

(4日) (1日) (日) (日) (日)

Theorem 1 - [Rommel et al., 2018]

Under assumptions 1 to 4, for any  $z \in E$  and  $t \in \mathbb{T}$ ,  $\hat{f}_{\ell^m(t)}^m(z)$  consistently approximates the marginal density  $f_t(z)$  as the number of curves *m* grows:

$$\forall \varepsilon > 0, \quad \lim_{m \to \infty} \mathbb{P}\left( |\hat{f}_t^m(z) - f_t(z)| < \varepsilon \right) = 1.$$

Note that:

•  $m \to \infty \neq N \to \infty$ ,

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

・ロト ・ 日 ・ ・ 日 ・ ・ 日 ・ - 日 ・

August 28<sup>th</sup> 2018 17 / 27

#### Theorem 1 - [Rommel et al., 2018]

Under assumptions 1 to 4, for any  $z \in E$  and  $t \in \mathbb{T}$ ,  $\hat{f}_{\ell^m(t)}^m(z)$  consistently approximates the marginal density  $f_t(z)$  as the number of curves *m* grows:

$$\forall \varepsilon > 0, \quad \lim_{m \to \infty} \mathbb{P}\left( |\hat{f}_t^m(z) - f_t(z)| < \varepsilon \right) = 1.$$

#### Note that:

• 
$$m \to \infty \neq N \to \infty$$
,

• Number of samples = random.

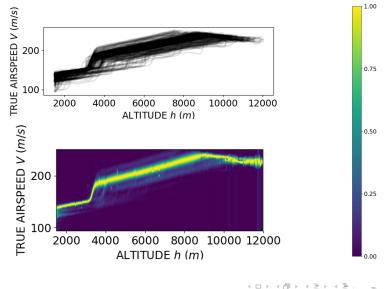
(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 17 / 27

- 不同 と そうと 不良 とう き

#### MARGINAL DENSITY ESTIMATION RESULTS



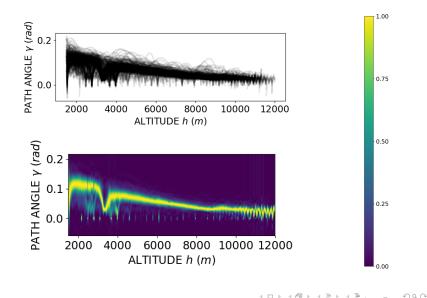
(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 18 / 27

San

#### MARGINAL DENSITY ESTIMATION RESULTS

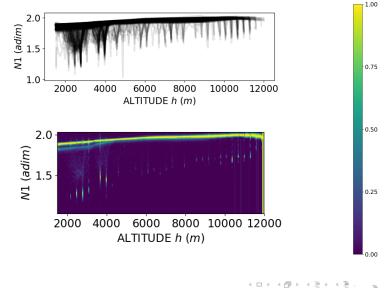


(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 19 / 27

### MARGINAL DENSITY ESTIMATION RESULTS



(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 20 / 27

San

(CMAP, INRIA, SAFETY LINE) MEAN MARGINAL LIKELIHOOD

◆□ > ◆□ > ◆臣 > ◆臣 > □臣 = のへで August 28<sup>th</sup> 2018 21 / 27

• Training set of m = 424 flights  $\simeq 334$  531 point observations,

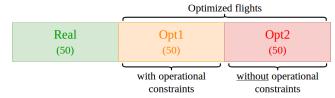
(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

4 ロ ト 4 日 ト 4 王 ト モ ク 4 で
August 28<sup>th</sup> 2018 21 / 27

- Training set of m = 424 flights  $\simeq 334$  531 point observations,
- Test set of 150 flights

• Training set of m = 424 flights  $\simeq 334$  531 point observations,



• Test set of 150 flights

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

- Training set of m = 424 flights  $\simeq 334$  531 point observations,
  - Optimized flights

     Real
     Opt1
     Opt2

     (50)
     (50)
     (50)

     with operational constraints
     without operational constraints
- Discrimination power comparison with (gmm-)FPCA and (integrated) LS-CDE:

(CMAP, INRIA, SAFETY LINE)

Test set of 150 flights

Mean Marginal Likelihood

(4 回下) イヨト イヨト

AUGUST 28<sup>th</sup> 2018

=

200

21 / 27

- Training set of m = 424 flights  $\simeq 334$  531 point observations,
  - Real (50)
     Opt1 (50)
     Opt2 (50)

     with operational constraints
     without operational constraints
- Discrimination power comparison with (gmm-)FPCA and (integrated) LS-CDE:

VAR.	Estimated Likelihoods		
	Real	Opt1	Opt2
MML	$\textbf{0.63} \pm \textbf{0.07}$	$\textbf{0.43} \pm \textbf{0.08}$	$\textbf{0.13}\pm\textbf{0.02}$
FPCA	$0.16\pm0.12$	$6.4\text{E-}03 \pm 3.8\text{E-}03$	$3.6e-03 \pm 5.4e-03$
LS-CDE	$0.77\pm0.05$	$0.68 \pm 0.04$	$0.49 \pm 0.06$

(CMAP, INRIA, SAFETY LINE)

Test set of 150 flights

Mean Marginal Likelihood

▲ □ → < 三 → < 三 → 三 < つ へ ○ AUGUST 28<sup>th</sup> 2018 21 / 27

#### $MML \ {\tt penalty}$

The MML can be used not only to assess the optimization solutions, but also to penalize the optimization itself:

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt$$
 (OCP)  
s.t. 
$$\begin{cases} \dot{\mathbf{x}}(t) = g(\mathbf{u}(t), \mathbf{x}(t)) + \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 22 / 27

◆□ > ◆□ > ◆臣 > ◆臣 > □臣 = のへで

#### $MML \ {\tt penalty}$

The MML can be used not only to assess the optimization solutions, but also to penalize the optimization itself:

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt - \lambda \operatorname{MML}(\mathbf{Z}, \mathbf{x}),$$

$$\sup_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{Z} \times \mathbb{U} \\ \text{Other constraints...}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt - \lambda \operatorname{MML}(\mathbf{Z}, \mathbf{x}),$$

$$(OCP)$$

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 22 / 27

◆□ > ◆□ > ◆臣 > ◆臣 > □臣 = のへで

### MML penalty

The MML can be used not only to assess the optimization solutions, but also to penalize the optimization itself:

$$\min_{\substack{(\mathbf{x}, \mathbf{u}) \in \mathbb{X} \times \mathbb{U} \\ \text{s.t.}}} \int_{0}^{t_{f}} C(\mathbf{u}(t), \mathbf{x}(t)) dt - \lambda \operatorname{MML}(\mathbf{Z}, \mathbf{x}),$$

$$\text{s.t.} \begin{cases} \dot{\mathbf{x}}(t) = g(\mathbf{u}(t), \mathbf{x}(t)) + \varepsilon(t), & \text{for a.e. } t \in [0, t_{f}], \\ \text{Other constraints...} \end{cases}$$

$$(OCP)$$

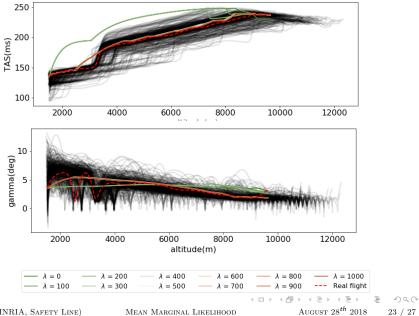
 $\bullet~\lambda$  sets trade-off between a fuel minimization and a likelihood maximization,

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

4 ロ ト 4 団 ト 4 国 ト 4 国 ト 国 の Q C AUGUST 28<sup>th</sup> 2018 22 / 27

#### PENALTY EFFECT



(CMAP, INRIA, SAFETY LINE)

23 / 27

#### CONSUMPTION X ACCEPTABILITY TRADE-OFF

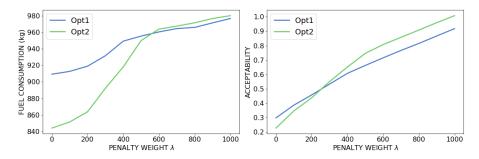


FIGURE: Average over 20 flights of the fuel consumption and MML score (called acceptability here) of optimized trajectories with varying MML-penalty weight  $\lambda$ .

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 24 / 27

General probabilistic criterion for quantifying the closeness between a curve and a set random trajectories,

Mean Marginal Likelihood

AUGUST 28<sup>th</sup> 2018 25 / 27

- General probabilistic criterion for quantifying the closeness between a curve and a set random trajectories,
- ② Class of consistent plug-in estimators, based on "histogram" of multivariate density estimators,

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 25 / 27

<ロト < 目 > < 目 > < 目 > < 目 > < 目 > < 0 < 0</p>

- General probabilistic criterion for quantifying the closeness between a curve and a set random trajectories,
- ② Class of consistent plug-in estimators, based on "histogram" of multivariate density estimators,
- ③ Applicable to the case of aircraft climb trajectories,

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

August 28<sup>th</sup> 2018 25 / 27

<ロト < 目 > < 目 > < 目 > < 目 > < 目 > < 0 < 0</p>

- General probabilistic criterion for quantifying the closeness between a curve and a set random trajectories,
- ② Class of consistent plug-in estimators, based on "histogram" of multivariate density estimators,
- ③ Applicable to the case of aircraft climb trajectories,
  - Competitive with other well-established SOA approaches,

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

《日》 《圖》 《문》 《문》 三臣

August 28<sup>th</sup> 2018 25 / 27

- General probabilistic criterion for quantifying the closeness between a curve and a set random trajectories,
- Class of consistent plug-in estimators, based on "histogram" of multivariate density estimators,
- ③ Applicable to the case of aircraft climb trajectories,
  - Competitive with other well-established SOA approaches,
- Showed that it can be used in optimal control problems to obtain solutions close to optimal, and still realistic.

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

イロト 不得下 イヨト イヨト 二日

August 28<sup>th</sup> 2018

25 / 27

#### THANK YOU FOR YOUR ATTENTION

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

イロト イラト イヨト モークへで AUGUST 28<sup>th</sup> 2018 26 / 27

#### References

- Recht, B. (2018). A tour of reinforcement learning: The view from continuous control. arXiv preprint arXiv:1806.09460.
- Rommel, C., Bonnans, J. F., Gregorutti, B., and Martinon, P. (2018). Quantifying the closeness to a set of random curves via the mean marginal likelihood. HAL report hal-01816407.
- Sugiyama, M., Takeuchi, I., Suzuki, T., Kanamori, T., Hachiya, H., and Okanohara, D. (2010). Conditional density estimation via least-squares density ratio estimation. In <u>Proceedings of the Thirteenth International</u> <u>Conference on Artificial Intelligence and Statistics</u>, pages 781–788.

(CMAP, INRIA, SAFETY LINE)

Mean Marginal Likelihood

<ロト < 部ト < 主ト < 主ト 主 のので August 28<sup>th</sup> 2018 27 / 27