Gaussian Mixture ESMDA

Preliminary conclusions

www.ltrace.com.br

Leandro P. de Figueiredo
www.ltrace.com.br
leandro@ltrace.com.br

Summary

- The idea of generalizing ESMDA for a Gaussian Mixture prior models
- Advantages:
  - Partly overcome the Gaussian unimodal assumption
  - In the particular case where the modes are interpreted as different facies, the facies estimation becomes an integral part of the solution
  - It might be useful to test different assumptions of the reservoir, e.g. depositional trends
- However, it may rise numerical and computation issues such as
  - Evaluation of PDF with high dimension
  - The number of models in the ensemble is probably higher

Literature

- Only one reference was found in the literature by searching for keywords in google scholar

Contents lists available at ScienceDirect

ELSEVIER

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Iterative ensemble smoothers in the annealed importance sampling framework

Andreas S. Stordal[a,*], Ahmed H. Elsheikh[b]

[a] *IRIS, P.O. Box 8046, Stavanger 4068, Norway*
[b] *Heriot-Watt University, Institute of Petroleum Engineering, Edinburgh EH14 4AS, United Kingdom*

- However, the context of the use of Gaussian mixture differs from our proposal

Stordal and Elsheikh, 2015

- It is introduced in the context of "importance sampling"
- The mixture is used to weigh each model of the ensemble

Set $j = 0$
Draw X_0^i from the prior $p(\cdot)$ for $i = 1, \dots N$
Propagate the sample to obtain $\mathcal{H}(X_0^i)$ for $i = 1, \dots N$
Draw X_1^i from X_0^i using $K_1(\cdot|X_0^i)$, $i = 1, \dots, N$
Evaluate the weights

$$\{w_1(X_1^i)\} = \frac{p(X_1^i)\ell(X_1^i)^{\alpha_1}T_1(X_0^i|X_1^i)}{p(X_0^i)K_1(X_1^i|X_0^i)}$$

Normalize the weights to obtain $\{w_1^i\}_{i=1}^N$
while -- **Sampling** -- **do**
  Set $j = j + 1$
  Draw X_j^i from $K_j(\cdot|X_{j-1}^i)$, for $i = 1, \dots N$
  Calculate $\{w_j(X_j^i)\}$ from Eq. (2.19), which is the weight correctionterm for the transition form X_{j-1}^i to X_j^i
  if -- **Resampling** -- **then**
    Sample $\{i^k\}_{k=1}^N$ with replacement from $(1, \dots, N)$ with$P(i^k = \ell) = w_j^\ell$
    Set $X_j^k = X_j^{i^k}$
    Set $w(X_j^i) = N^{-1}$, $i = 1, \dots, N$
  end
end

$$w_j \propto w_{j-1}\frac{p_j(x_j|y)T_j(x_{j-1}|x_j)}{p_{j-1}(x_{j-1}|y)K_j(x_j|x_{j-1})}$$

Algorithm 1: Weighted ESMDA.

Set $j = 0$
Draw X_0^i from the prior $p(\cdot)$ for $i = 1, \dots N$
Propagate the sample to obtain $\mathcal{H}(X_0^i)$ for $i = 1, \dots N$
Draw X_1^i from X_0^i defined by Eq. (4.3) Evaluate the weights $\{w_1(X_1^i)\} = \Phi(y|\mathcal{H}(X_{j-1}), h^2\mathrm{Cov}\{\mathcal{H}(X_{j-1})\} + \alpha_j^{-1}\mathbb{R})$
Normalize the weights to obtain $\{w_1^i\}_{i=1}^N$
while -- **Sampling** -- **do**
  Set $j = j + 1$
  Draw X_j^i from Eq. (4.3), for $i = 1, \dots N$
  Calculate $\{w_j(X_j^i) \propto w_{j-1}^i\Phi(y|\mathcal{H}(X_{j-1}), h^2\mathrm{Cov}\{\mathcal{H}(X_{j-1})\} + \alpha_j^{-1}\mathbb{R})\}$
  Calculate τ from Eq. (4.2)
  Set $w_j = \tau w_j + (1-\tau)N^{-1}$
  if -- **Resampling** -- **then**
    Sample $\{i^k\}_{k=1}^N$ with replacement from $(1, \dots, N)$ with $P(i^k = \ell) = w_j^\ell$
    Set $X_j^k = X_j^{i^k}$
    Set $w_j(X_j^i) = N^{-1}$, $i = 1, \dots, N$
  end
end

Algorithm 2: Gaussian Mixture MDA.

Posterior distribution:

$$\widehat{p}_j(x|y) \propto \sum_{i=1}^N w(X_j^i)\Phi(x|X_j^i, \mathbb{P}_j),$$

Stordal and Elsheikh, 2015

- It is introduced in the context of “importance sampling”
- The mixture is used to weigh each model of the ensemble

Set $j = 0$
Draw X_0^i from the prior $p(\cdot)$ for $i = 1, \dots N$
Propagate the sample to obtain $\mathcal{H}(X_0^i)$ for $i = 1, \dots N$
Draw X_1^i from X_0^i using $K_1(\cdot|X_0^i)$, $i = 1, \dots, N$
Evaluate the weights

$$\{w_1(X_1^i)\} = \frac{p(X_1^i)\ell(X_1^i)^{\alpha_1} T_1(X_0^i|X_1^i)}{p(X_0^i)K_1(X_1^i|X_0^i)}$$

Normalize the weights to obtain $\{w_1^i\}_{i=1}^N$
while -- Sampling -- do
  Set $j = j + 1$
  Draw X_j^i from $K_j(\cdot|X_{j-1}^i)$, for $i = 1, \dots N$
  Calculate $\{w_j(X_j^i)\}$ from Eq. (2.19), which is the weight correctionterm for the transition form X_{j-1}^i to X_j^i
  if -- Resampling -- then
    Sample $\{i^k\}_{k=1}^N$ with replacement from $(1, \dots, N)$ with$P(i^k = \ell) = w_j^\ell$
    Set $X_j^k = X_j^{i^k}$
    Set $w(X_j^i) = N^{-1}$, $i = 1, \dots, N$
  end
end

Algorithm 1: Weighted ESMDA.

$$w_j \propto w_{j-1} \frac{p_j(x_j|y)T_j(x_{j-1}|x_j)}{p_{j-1}(x_{j-1}|y)K_j(x_j|x_{j-1})}$$

Set $j = 0$
Draw X_0^i from the prior $p(\cdot)$ for $i = 1, \dots N$
Propagate the sample to obtain $\mathcal{H}(X_0^i)$ for $i = 1, \dots N$
Draw X_1^i from X_0^i defined by Eq. (4.3) Evaluate the weights $\{w_1(X_1^i)\} = \Phi(y|\mathcal{H}(X_{j-1}), h^2\mathrm{Cov}\{\mathcal{H}(X_{j-1})\} + \alpha_j^{-1}\mathbb{R})$
Normalize the weights to obtain $\{w_1^i\}_{i=1}^N$
while -- Sampling -- do
  Set $j = j + 1$
  Draw X_j^i from Eq. (4.3), for $i = 1, \dots N$
  Calculate
  $\{w_j(X_j^i) \propto w_{j-1}^i \Phi(y|\mathcal{H}(X_{j-1}), h^2\mathrm{Cov}\{\mathcal{H}(X_{j-1})\} + \alpha_j^{-1}\mathbb{R})\}$
  Calculate τ from Eq. (4.2)
  Set $w_j = \tau w_j + (1 - \tau)N^{-1}$
  if -- Resampling -- then
    Sample $\{i^k\}_{k=1}^N$ with replacement from $(1, \dots, N)$ with $P(i^k = \ell) = w_j^\ell$
    Set $X_j^k = X_j^{i^k}$
    Set $w_j(X_j^i) = N^{-1}$, $i = 1, \dots, N$
  end
end

Algorithm 2: Gaussian Mixture MDA.

Posterior distribution:

$$\hat{p}_j(x|y) \propto \sum_{i=1}^N w(X_j^i)\Phi(x|X_j^i, \mathbb{P}_j),$$

“weights according to (2.19) an adaptive weight shrinkage is applied”

$$w_{\mathrm{new}} = \tau w_{\mathrm{old}} + (1 - \tau)N^{-1},$$

where

$$\tau = \left(\sum_{j=1}^N (w_{\mathrm{old}}^i)^2\right)^{-1}.$$

Gaussian Mixture BLI

- Forward model with Gaussian error:

$$\boldsymbol{d} = \boldsymbol{G}\boldsymbol{m} + \boldsymbol{e}$$

- Gaussian mixture prior:

$$\boldsymbol{m} \sim \sum_{k=1}^{F} \lambda_k N\left(\boldsymbol{m}; \boldsymbol{\mu}_{m|k}, \boldsymbol{C}_{m|k}\right)$$

- The posterior distribution is also a Gaussian mixture and it is analytically obtained:

$$\boldsymbol{m}|\boldsymbol{d} \sim \sum_{k=1}^{F} \lambda_{k|d} N\left(\boldsymbol{m}; \boldsymbol{\mu}_{m|d,k}, \boldsymbol{C}_{m|d,k}\right)$$

$$\boldsymbol{\mu}_{r|m,k} = \boldsymbol{\mu}_{r|k} + \boldsymbol{\Sigma}_{r|k}\mathbf{G}^T\left(\mathbf{G}\boldsymbol{\Sigma}_{r|k}\mathbf{G}^T + \boldsymbol{\Sigma}_e\right)^{-1}\left(\boldsymbol{m} - \mathbf{G}\boldsymbol{\mu}_{r|k}\right)$$

$$\boldsymbol{\Sigma}_{r|m,k} = \boldsymbol{\Sigma}_{r|k} - \boldsymbol{\Sigma}_{r|k}\mathbf{G}^T\left(\mathbf{G}\boldsymbol{\Sigma}_{r|k}\mathbf{G}^T + \boldsymbol{\Sigma}_e\right)^{-1}\mathbf{G}\boldsymbol{\Sigma}_{r|k}$$

$g(\boldsymbol{\mu}_{r|k})$

$\boldsymbol{C}_{dd|k} + \boldsymbol{C}_e$

$$\lambda_{k|d} \propto \lambda_k N\left(\boldsymbol{d}; \boldsymbol{G}\boldsymbol{\mu}_{m|,k}, \boldsymbol{G}\boldsymbol{C}_{m|k}\boldsymbol{G}^T + \boldsymbol{C}_e\right)$$

General comments

- Since the distributions in the ESMDA are indirectly given by the multiple realizations in the ensemble, in our implementation we define a Gaussian mixture prior by drawing multiple realizations from each Gaussian component
- In practice, it is like we have multiple ensembles
- The posterior is then computed from the ensemble after the updates

GM-ESMDA - First approach

- Straightforward application of the Gaussian Mixture BLI
- Prior considering 2 modes:

$$\boldsymbol{m} \sim \sum_{k=1}^{2} \omega_k N(\boldsymbol{m}; \boldsymbol{\mu}_{m|k}, \boldsymbol{C}_{m|k})$$

- GM-ESMDA:

$$\boldsymbol{m}_1 \sim N(\boldsymbol{m}; \boldsymbol{\mu}_{m|1}, \boldsymbol{C}_{m|1}) \qquad \boldsymbol{m}_2 \sim N(\boldsymbol{m}; \boldsymbol{\mu}_{m|2}, \boldsymbol{C}_{m|2})$$

$$\boldsymbol{m}_1^{i+1} = \boldsymbol{m}_1^i + \boldsymbol{C}_{md|1}^i \left(\boldsymbol{C}_{dd|1}^i + \alpha \boldsymbol{C}_e\right)^{-1} (\boldsymbol{d}_{obs}^i - g(\boldsymbol{m}_1^i)) \qquad \boldsymbol{m}_2^{i+1} = \boldsymbol{m}_2^i + \boldsymbol{C}_{md|2}^i \left(\boldsymbol{C}_{dd|2}^i + \alpha \boldsymbol{C}_e\right)^{-1} (\boldsymbol{d}_{obs}^i - g(\boldsymbol{m}_2^i))$$

$$\lambda_k^{i+1} \propto \lambda_k^i N(\boldsymbol{d}_{obs}; g(\boldsymbol{\mu}_{m_k}), \boldsymbol{C}_{dd|k} + \boldsymbol{C}_d)$$

$$p(\boldsymbol{m}|\boldsymbol{d}) \approx \lambda_1 kdensity\,(\boldsymbol{m}_1) + \lambda_2 kdensity\,(\boldsymbol{m}_2)$$

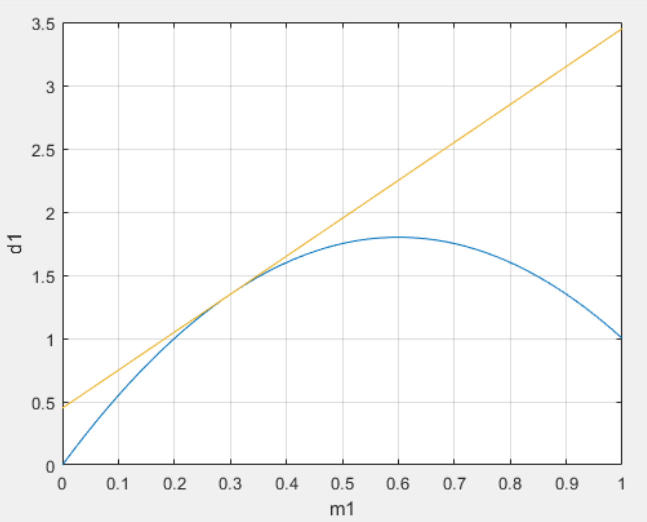
Ultra Toy Problem

- Bivariate problem with a 2nd order polynomial forward model

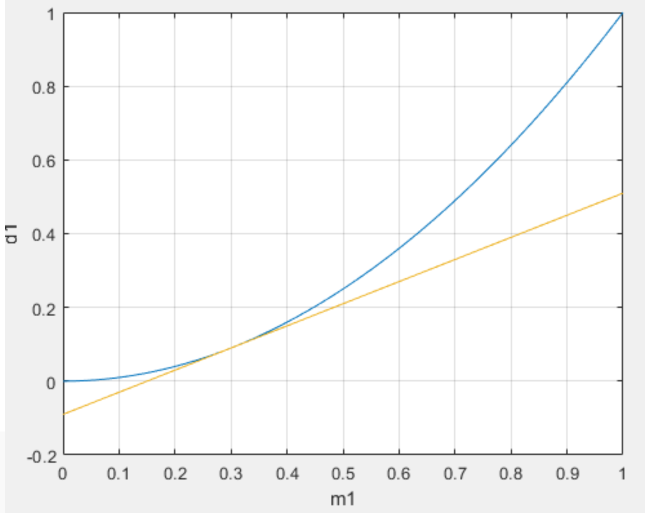
$$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} g(m_1) \\ g(m_2) \end{pmatrix}$$

- It allows us to compare the algorithm with Gaussian Mix BLI
- Additional parameter to control de non-linearity

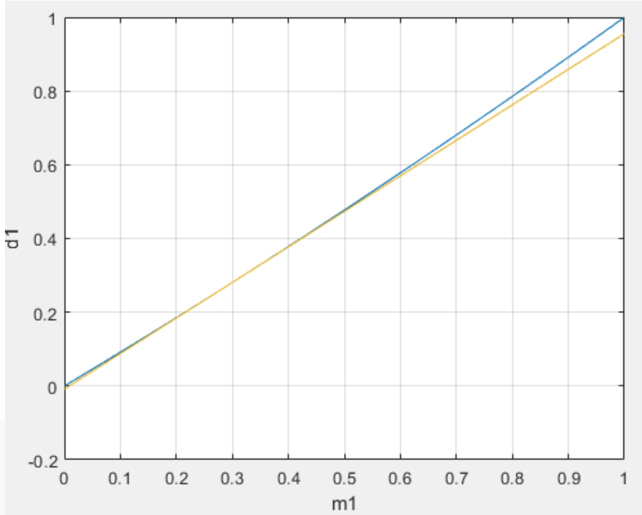
Linear factor = - 0.6



Linear factor = 0

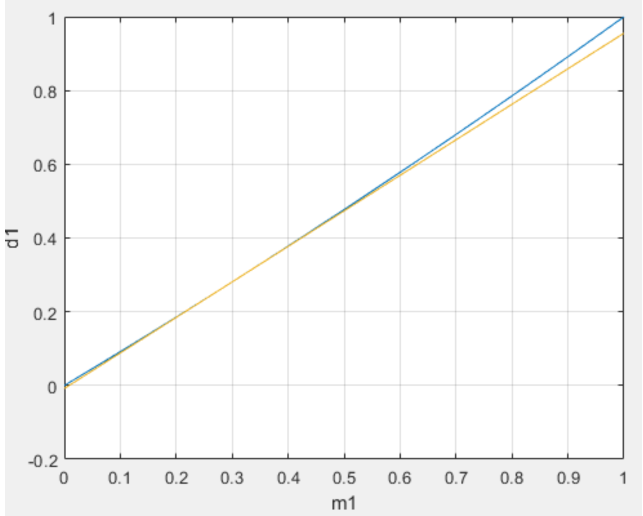


Linear factor = 5



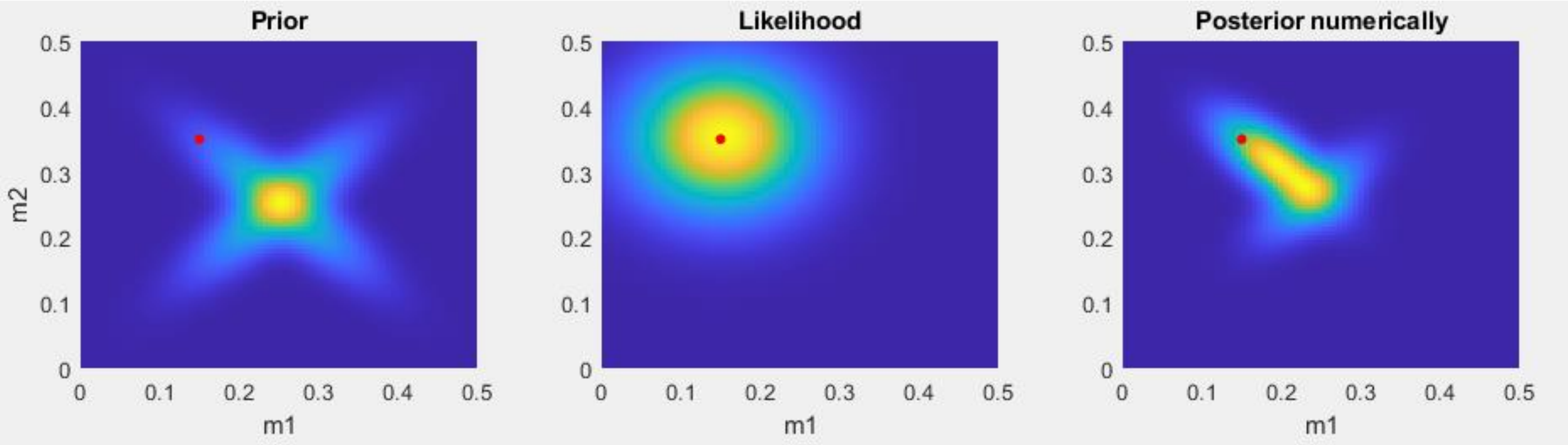
Experiment 1
Bivariate linear case
Priors with the same mean and variance but different correlations

Linear factor = 10



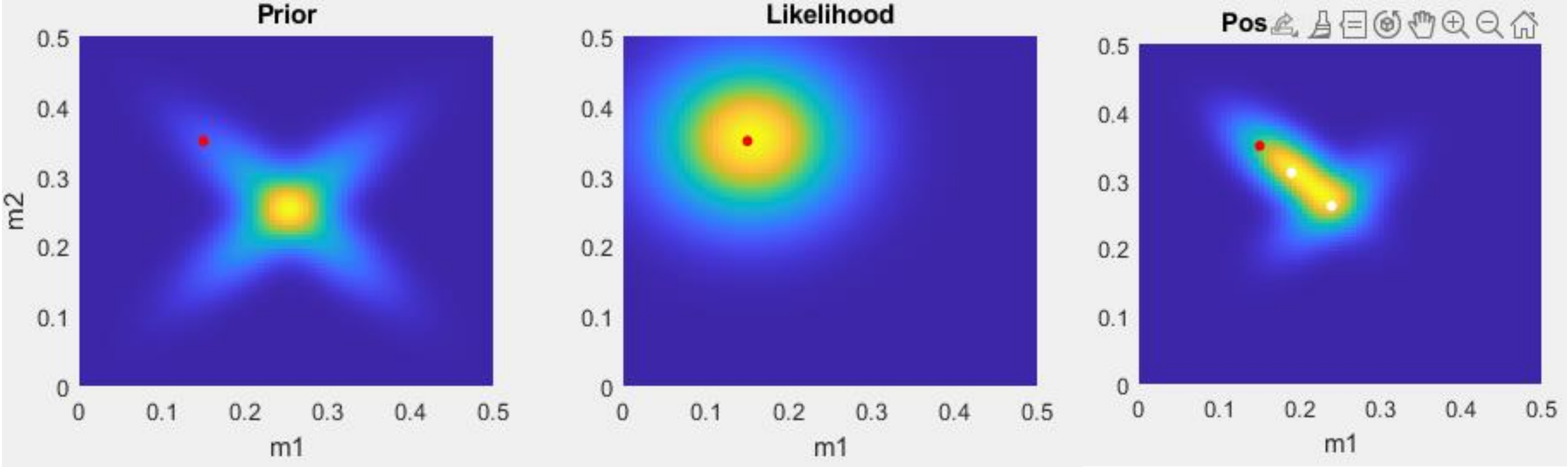
Reference solution – Linear case

- Extensively numerically evaluation of the model space



Reference solution – Linear case

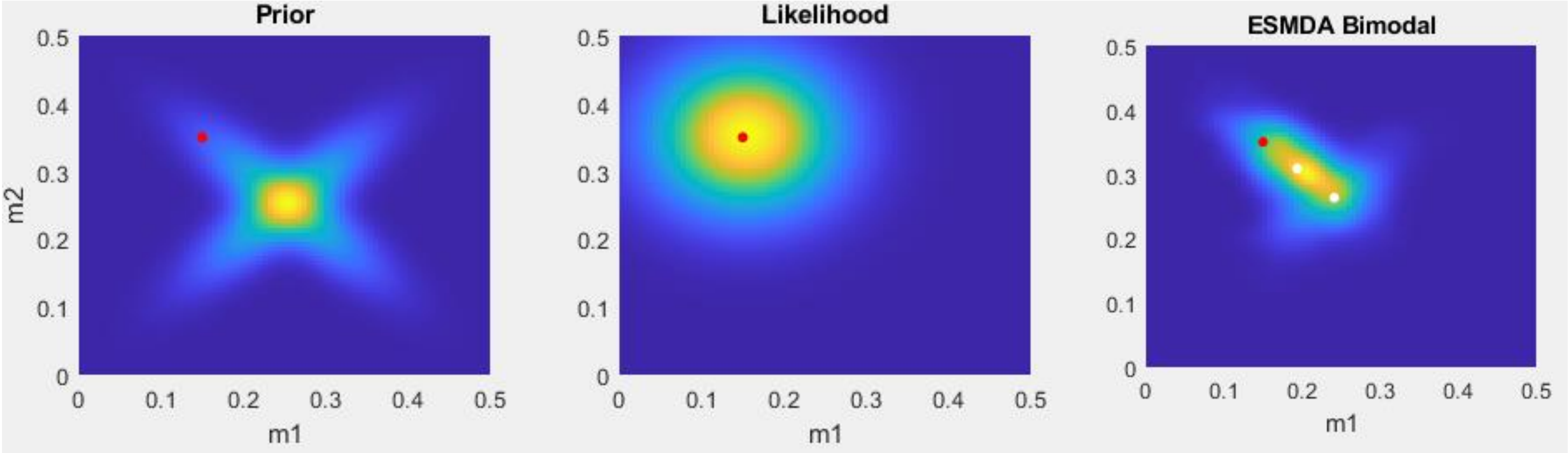
- Analytical Gaussian mixture BLI



$$\lambda_1 = 0.3 \qquad \lambda_2 = 0.7$$

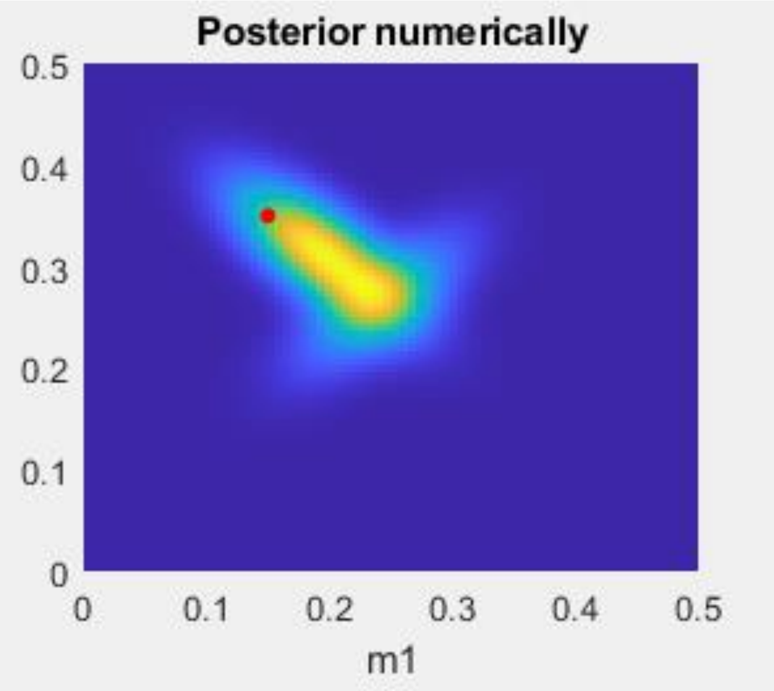
Gaussian Mix ESMDA - Linear case

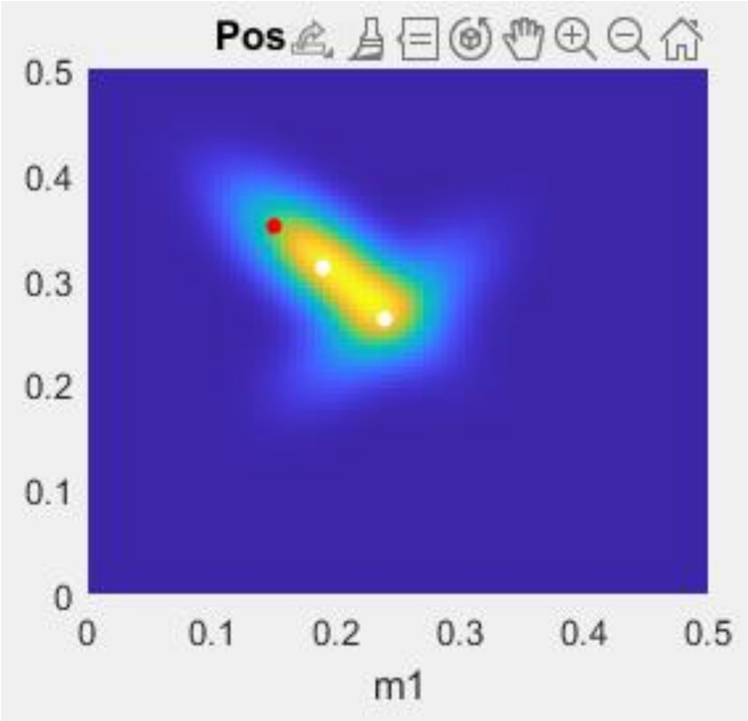
- First approach

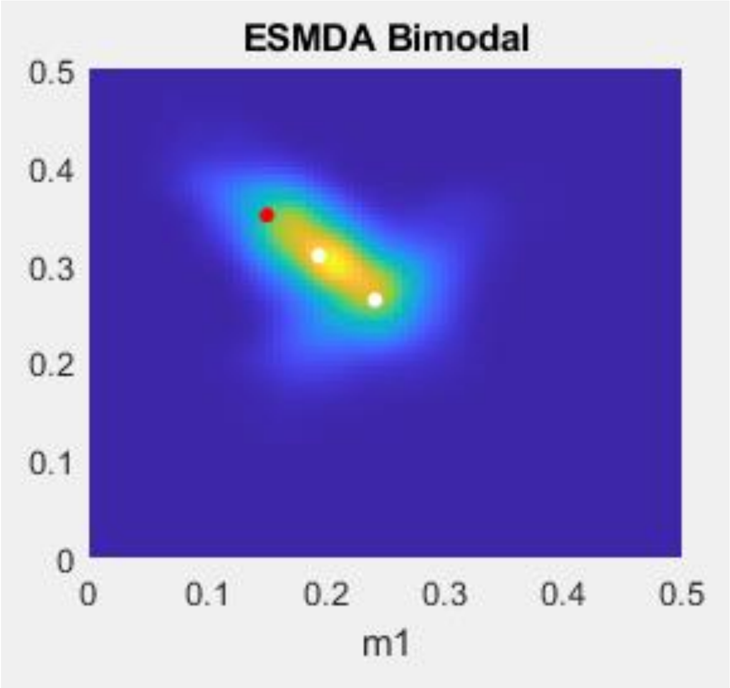


$$\lambda_1 = 0.26 \qquad \lambda_2 = 0.74$$

Posterior estimates – Linear case





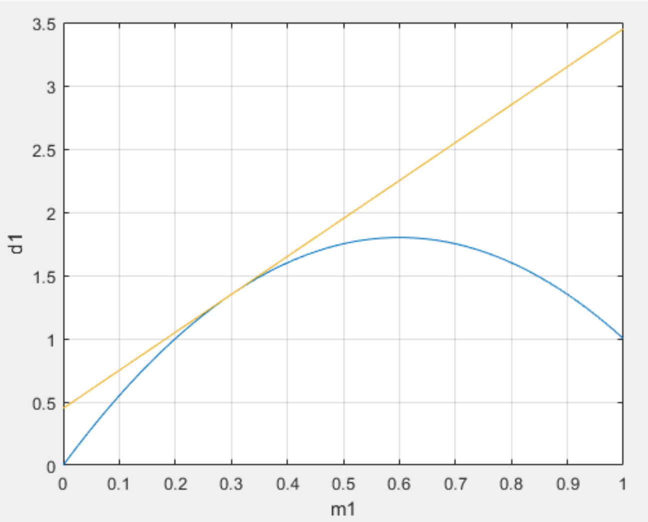


$\lambda_1 = 0.3$ $\lambda_2 = 0.7$

$\lambda_1 = 0.26$ $\lambda_2 = 0.74$

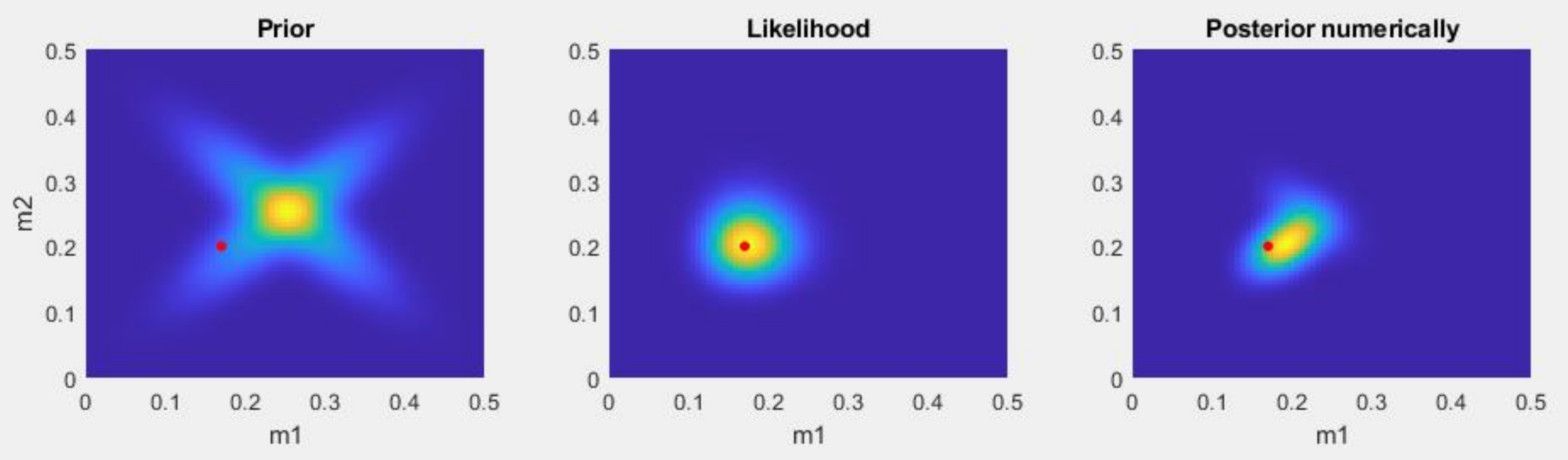
Experiment 2
Bivariate NON linear case
Priors with the same mean and variance but different correlations

Linear factor = - 0.6



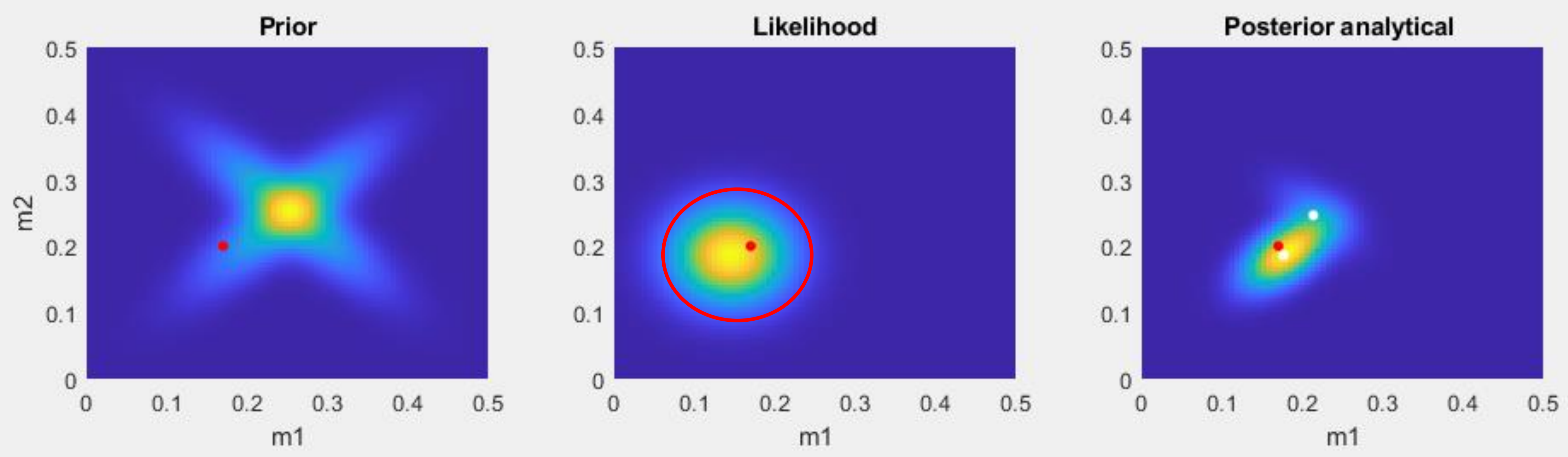
Reference solution – Non-Linear case

- Extensively numerically evaluation of the model space



Reference solution – Non-Linear case

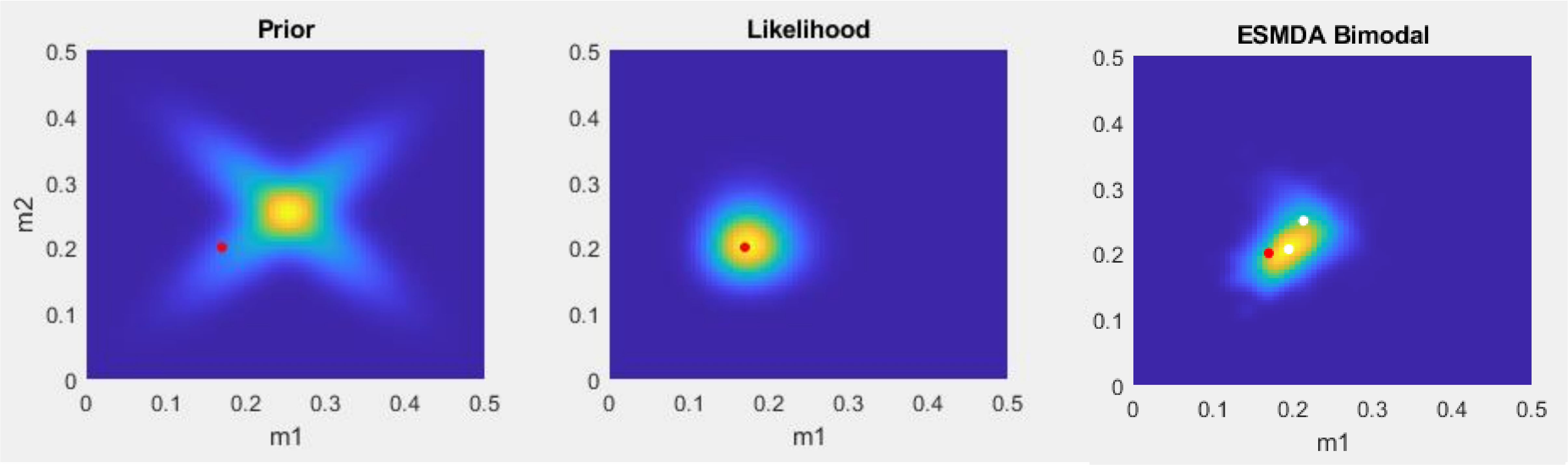
- Analytical Gaussian mixture BLI



$\lambda_1 = 0.83 \qquad \lambda_2 = 0.17$

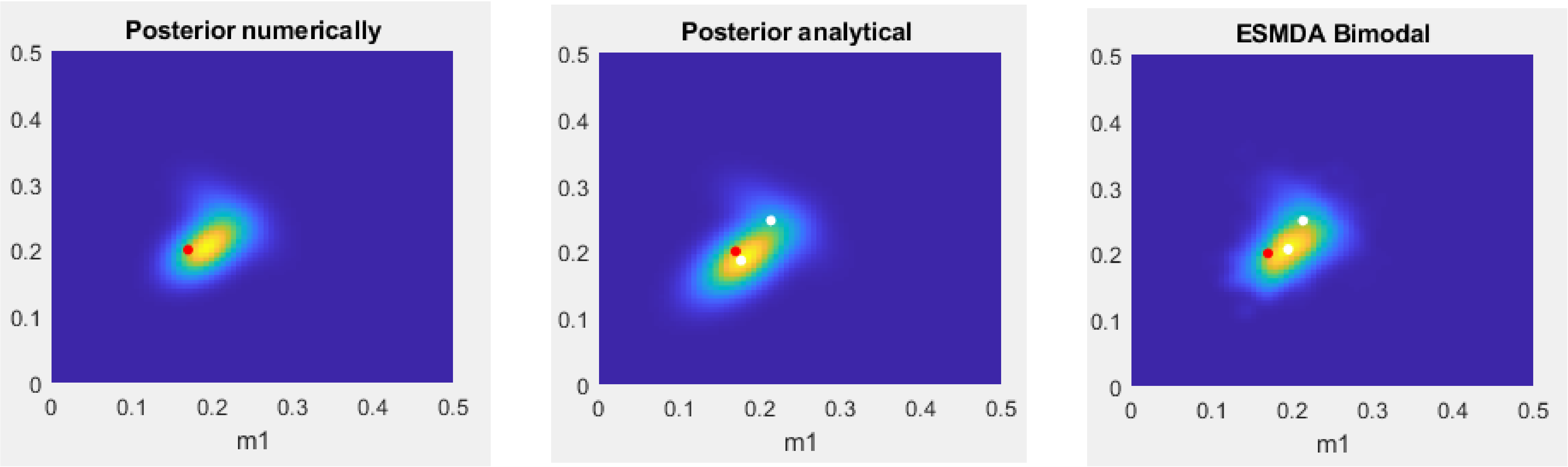
Gaussian Mix ESMDA – Non-Linear case

- First approach



$\lambda_1 = 0.72 \qquad \lambda_2 = 0.28$

Posterior estimates – Non linear case



$\lambda_1 = 0.83$ $\lambda_2 = 0.17$

$\lambda_1 = 0.72$ $\lambda_2 = 0.28$

Conclusions

- For the linear model, the proposed GM-ESMDA shows similar results to the GM-BLI and numerical results

- For the non-linear model, the proposed GM-ESMDA shows similar results to the numerical results and it differs from the GM-BLI, which is expected since the forward model was linearized.

Future work

- Experiments with different prior means
- Compute the statistical moments of the posterior distributions for a quantitative comparison
- Investigate further Gibbs algorithm approach
- Compute Kullback–Leibler divergence between the reference posterior and the estimated posterior of the methods
- Application in a pointwise petrophysical inversion from elastic properties and non-linear rock-physics data – Well log of Volve data set as reference petrophysical data
    - It is still a bi/tri-variate problem
    - “Easy” and fast publication
- Meanwhile, we think in a way for handling the lambda evaluation for high-dimensional problems

LTrace
Geophysical Solutions

PETROBRAS

Thank you

www.ltrace.com.br

Leandro P. de Figueiredo

leandro@ltrace.com.br