

VII. COMPUTATIONAL MULTI-POINT BME AND BME CONFIDENCE SETS

Until now we have considered the estimation of the value of a S/TRF at one point at the time, independently of the estimated values at other estimation points, i.e. a single-point mapping approach. The Bayesian Maximum Entropy (BME; Christakos, 1990, 1992) method offers a framework where the prediction of a S/TRF at several estimation points may be considered jointly, which is referred to as a multi-point mapping approach. Whereas single-point mapping uses an univariate pdf to describe each estimated value separately, the multi-point mapping approach uses a more informative multivariate pdf describing all the estimated values jointly. Multi-point estimation leads in some cases to maps that are considerably different than that obtained with the single-point approach, thus providing a different representation that might be a considerable improvement. Additionally the hallmark of any stochastic estimation method is its ability to provide an assessment of the uncertainty associated to the estimated map, and BME provides an accurate and exhaustive assessment of uncertainty by means of BME confidence sets, which are a generalization of the idea of confidence intervals, for the case of multi-point mapping.

7.1 The Multi-Point Mapping Approach

In the multi-point approach we consider a set of estimation points \mathbf{p}_{k_ℓ} , $\ell = 1, \dots, \rho$, jointly, and we let the vector of random variables $\boldsymbol{\chi}_k = [\chi_{k_1} \dots \chi_{k_\rho}]^T$ represent the S/TRF $X(\mathbf{p})$ at these estimation points. The goal of the multi-point mapping approach is to obtain a multivariate (posterior) pdf $f_K(\boldsymbol{\chi}_k)$ given the general and specificatory knowledge

available. The multivariate pdf $f_K(\boldsymbol{\chi}_k)$ describes the values estimated jointly, which is more informative than the univariate pdf's $f_K(\chi_{k_\ell})$, $\ell = 1, \dots, \rho$, obtained from single-point mapping.

Using the usual notation we let $\boldsymbol{\chi}_{\text{hard}} = [\chi_1 \dots \chi_{m_h}]^T$ and $\boldsymbol{\chi}_{\text{soft}} = [\chi_{m_h+1} \dots \chi_m]^T$ represent the hard and soft data for the S/TRF $X(\boldsymbol{p})$ at m space/time points, and we let $\boldsymbol{\chi}_{\text{data}}^T = [\boldsymbol{\chi}_{\text{hard}}^T \boldsymbol{\chi}_{\text{soft}}^T]$ and $\boldsymbol{\chi}_{\text{map}}^T = [\boldsymbol{\chi}_{\text{data}}^T \boldsymbol{\chi}_k^T]$. Note that this is the same notation as for the single-point mapping, except that we have replaced the scalar χ_k with the vector $\boldsymbol{\chi}_k = [\chi_{k_1} \dots \chi_{k_\rho}]^T$, and the vector $\boldsymbol{\chi}_{\text{map}}^T = [\boldsymbol{\chi}_{\text{data}}^T \chi_k]$ with the vector $\boldsymbol{\chi}_{\text{map}}^T = [\boldsymbol{\chi}_{\text{data}}^T \boldsymbol{\chi}_k]$. Using this extended notation, the basic BME equations will lead, in principle, to the multi-point formulation.

The general knowledge is expressed in terms of constraints of the form

$$\overline{g_\alpha} = \int d\boldsymbol{\chi}_{\text{map}} f_G(\boldsymbol{\chi}_{\text{map}}) g_\alpha(\boldsymbol{\chi}_{\text{map}}) \quad (7.1)$$

where $\alpha = 0, 1, \dots, N_c$ and the $g_\alpha(\boldsymbol{\chi}_{\text{map}})$ are given functions of χ_i 's. The prior pdf $f_G(\boldsymbol{\chi}_{\text{map}})$ is obtained by maximizing the entropy function $S[f_G] = -\int d\boldsymbol{\chi}_{\text{map}} f_G(\boldsymbol{\chi}_{\text{map}}) \log f_G(\boldsymbol{\chi}_{\text{map}})$ subject to the constraints (7.1). The solution to this constrained maximization problem is the prior pdf given by

$$f_G(\boldsymbol{\chi}_{\text{map}}) = Z^{-1} \exp\left[\sum_{\alpha=1}^{N_c} \mu_\alpha g_\alpha(\boldsymbol{\chi}_{\text{map}})\right] \quad (7.2)$$

where the N_c unknown coefficients μ_α are calculated from the N_c equations of Eqs. (7.1).

The multi-point (multivariate, ρ -points) posterior pdf $f_K(\boldsymbol{\chi}_k)$ is obtained using specificatory knowledge containing hard and soft data. For soft data of the interval type as defined by Eq. (2.26), the multi-point posterior pdf is

$$f_K(\boldsymbol{\chi}_k) = A^{-1} \int_l^u d\boldsymbol{\chi}_{\text{soft}} f_G(\boldsymbol{\chi}_{\text{map}}), \quad (7.3)$$

where $A = \int_l^u d\boldsymbol{\chi}_{\text{soft}} f_G(\boldsymbol{\chi}_{\text{data}})$ and $f_G(\boldsymbol{\chi}_{\text{data}}) = \int d\boldsymbol{\chi}_k f_G(\boldsymbol{\chi}_{\text{map}})$. For soft data of the probabilistic type as defined by Eq. (2.27), the multi-point posterior pdf is given by

$$f_K(\boldsymbol{\chi}_k) = A^{-1} \int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) f_G(\boldsymbol{\chi}_{\text{map}}), \quad (7.4)$$

where $A = \int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) f_G(\boldsymbol{\chi}_{\text{data}})$.

I propose next a formulation which accounts for general knowledge consisting of the mean and the covariance function of the S/TRF $X(\boldsymbol{p})$. This formulation is similar to that of Chapters IV and V, extended to the case of multi-point mapping, and it will lead to an efficient numerical implementation.

7.2. A proposed Formulation of the Posterior PDF for Efficient Computation

Let's assume that the Space/Time Random Field $X(\boldsymbol{p})$ has a known mean $m_x(\boldsymbol{p}) = \overline{X(\boldsymbol{p})}$, as well as a known covariance function $c_x(\boldsymbol{p}, \boldsymbol{p}')$, usually obtained from fitting to experimental data. Let $\boldsymbol{x}_{\text{map}}$ be the vector of random variables taking values $\boldsymbol{\chi}_{\text{map}}$. The covariance matrix $\boldsymbol{C}_{\text{map}} = \overline{(\boldsymbol{x}_{\text{map}} - \boldsymbol{m}_{\text{map}})(\boldsymbol{x}_{\text{map}} - \boldsymbol{m}_{\text{map}})^T}$ associated with the vector of random variables $\boldsymbol{x}_{\text{map}}$ is equal to

$$\mathbf{C}_{\text{map}} = \begin{bmatrix} c_x(\mathbf{p}_1, \mathbf{p}_1) & \dots & c_x(\mathbf{p}_1, \mathbf{p}_m) & c_x(\mathbf{p}_1, \mathbf{p}_{k_1}) & \dots & c_x(\mathbf{p}_1, \mathbf{p}_{k_\rho}) \\ \vdots & & \vdots & \vdots & & \vdots \\ c_x(\mathbf{p}_m, \mathbf{p}_1) & \dots & c_x(\mathbf{p}_m, \mathbf{p}_m) & c_x(\mathbf{p}_m, \mathbf{p}_{k_1}) & \dots & c_x(\mathbf{p}_m, \mathbf{p}_{k_\rho}) \\ c_x(\mathbf{p}_{k_1}, \mathbf{p}_1) & \dots & c_x(\mathbf{p}_{k_1}, \mathbf{p}_m) & c_x(\mathbf{p}_{k_1}, \mathbf{p}_{k_1}) & \dots & c_x(\mathbf{p}_{k_1}, \mathbf{p}_{k_\rho}) \\ \vdots & & \vdots & \vdots & & \vdots \\ c_x(\mathbf{p}_{k_\rho}, \mathbf{p}_1) & \dots & c_x(\mathbf{p}_{k_\rho}, \mathbf{p}_m) & c_x(\mathbf{p}_{k_\rho}, \mathbf{p}_{k_1}) & \dots & c_x(\mathbf{p}_{k_\rho}, \mathbf{p}_{k_\rho}) \end{bmatrix}. \quad (7.5)$$

For this type of general knowledge, the prior pdf $f_G(\boldsymbol{\chi}_{\text{map}})$ of Eq. (7.2) can be expressed as a multivariate Gaussian pdf with mean \mathbf{m}_{map} and covariance matrix \mathbf{C}_{map} . Assuming without loss of generality that $\mathbf{m}_{\text{map}} = \mathbf{0}$ (or equivalently that the mean has been subtracted from the hard and soft data), we may write the prior pdf as

$$f_G(\boldsymbol{\chi}_{\text{map}}) = \phi(\boldsymbol{\chi}_{\text{map}}; \mathbf{0}, \mathbf{C}_{\text{map}}) \quad (7.6)$$

where $\phi(\mathbf{x}; \bar{\mathbf{x}}, \mathbf{C})$ denote the n -point Gaussian pdf of the random vector \mathbf{x} with mean $\bar{\mathbf{x}}$ and covariance matrix \mathbf{C} as defined in Eq. (4.4).

In the following, it is convenient to define the partitioned matrices

$$\mathbf{C}_{\text{map}} = \begin{bmatrix} \mathbf{C}_{hs,hs} & \mathbf{C}_{hs,k} \\ \mathbf{C}_{k,hs} & \mathbf{C}_{k,k} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{h,h} & \mathbf{C}_{h,s} & \mathbf{C}_{h,k} \\ \mathbf{C}_{s,h} & \mathbf{C}_{s,s} & \mathbf{C}_{s,k} \\ \mathbf{C}_{k,h} & \mathbf{C}_{k,s} & \mathbf{C}_{k,k} \end{bmatrix}, \quad (7.7)$$

and

$$\mathbf{C}_{kh,kh} = \begin{bmatrix} \mathbf{C}_{k,k} & \mathbf{C}_{k,h} \\ \mathbf{C}_{h,k} & \mathbf{C}_{h,h} \end{bmatrix}, \quad (7.8)$$

where the subscripts h , s , and k denote hard data points, soft data points and estimation points, respectively.

The posterior pdf for soft data of the interval type is given by Eq. (7.3). Substituting for the prior pdf $f_G(\boldsymbol{\chi}_{\text{map}}) = \phi(\boldsymbol{\chi}_{\text{map}}; \boldsymbol{\theta}, \mathbf{C}_{\text{map}})$ and after some manipulations we obtain

$$f_K(\boldsymbol{\chi}_k) = A'^{-1} \phi(\boldsymbol{\chi}_k; \mathbf{B}_{k|h} \boldsymbol{\chi}_{\text{hard}}, \mathbf{C}_{k|h}) \int_{l-B_{s|kh} \boldsymbol{\chi}_{kh}}^{u-B_{s|kh} \boldsymbol{\chi}_{kh}} d\boldsymbol{\chi}_{\text{soft}} \phi(\boldsymbol{\chi}_{\text{soft}}; \boldsymbol{\theta}, \mathbf{C}_{s|kh}), \quad (7.9)$$

where $\boldsymbol{\chi}_{kh}^T = [\boldsymbol{\chi}_k^T \ \boldsymbol{\chi}_{\text{hard}}^T]$, $\mathbf{B}_{k|h} = \mathbf{C}_{k,h} \mathbf{C}_{h,h}^{-1}$, $\mathbf{C}_{k|h} = \mathbf{C}_{k,k} - \mathbf{B}_{k|h} \mathbf{C}_{h,k}$, $\mathbf{B}_{s|kh} = \mathbf{C}_{s,kh} \mathbf{C}_{kh,kh}^{-1}$, $\mathbf{C}_{s|kh} = \mathbf{C}_{s,s} - \mathbf{B}_{s|kh} \mathbf{C}_{kh,s}$, and $A' = \int_l^u d\boldsymbol{\chi}_{\text{soft}} \phi(\boldsymbol{\chi}_{\text{soft}}; \mathbf{B}_{s|h} \boldsymbol{\chi}_{\text{hard}}, \mathbf{C}_{s|h})$.

In the case of soft data of the probabilistic type we obtain

$$f_K(\boldsymbol{\chi}_k) = A'^{-1} \phi(\boldsymbol{\chi}_k; \mathbf{B}_{k|h} \boldsymbol{\chi}_{\text{hard}}, \mathbf{C}_{k|h}) \int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) \phi(\boldsymbol{\chi}_{\text{soft}}; \mathbf{B}_{s|kh} \boldsymbol{\chi}_{kh}, \mathbf{C}_{s|kh}), \quad (7.10)$$

where $A' = \int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) \phi(\boldsymbol{\chi}_{\text{soft}}; \mathbf{B}_{s|h} \boldsymbol{\chi}_{\text{hard}}, \mathbf{C}_{s|h})$.

7.3. Multi-point BME Mode Estimate

The multi-point BME mode estimate $\hat{\boldsymbol{\chi}}_k^{MP} = [\hat{\chi}_{k_1}^{MP} \ \dots \ \hat{\chi}_{k_\rho}^{MP}]$ is the most probable value for the S/TRF $X(\mathbf{p})$ at the estimation point \mathbf{p}_{k_j} , $j = 1, \dots, \rho$, with respect to the multi-point posterior pdf $f_K(\boldsymbol{\chi}_k)$. The MP superscript emphasizes that $\hat{\boldsymbol{\chi}}_k^{MP}$ is obtained using the multi-point approach. Note that in the single-point approach, we obtain a single-point (univariate) posterior pdf $f_K^{SP}(\chi_{k_\ell})$ for each of the estimation points, $\ell = 1, \dots, \rho$, considered independent one from another. The single-point BME mode for each of the

estimation point is denoted as $\hat{\chi}_{k_j}^{SP}$, $\ell = 1, \dots, \rho$, where the SP superscript emphasizes that $\hat{\chi}_{k_j}^{SP}$ is obtained using the single-point approach.

The multi-point mode $\hat{\chi}_k^{MP}$ is obtained by finding the vector which maximizes the multi-point posterior pdf, i.e.

$$\left. \begin{array}{l} \nabla_{\boldsymbol{\chi}_k} f_K^{MP}(\boldsymbol{\chi}_k) \Big|_{\boldsymbol{\chi}_k = \hat{\boldsymbol{\chi}}_k^{MP}} = \mathbf{0} \\ \text{or} \\ \frac{\partial}{\partial \chi_{k_\ell}} f_K^{MP}(\boldsymbol{\chi}_k) \Big|_{\boldsymbol{\chi}_k = \hat{\boldsymbol{\chi}}_k^{MP}} = 0 \quad \ell = 1, \dots, \rho \end{array} \right\} \quad (7.11)$$

Note that in general the ℓ -Th. component $\hat{\chi}_{k_\ell}^{MP}$ of the multi-point mode vector $\hat{\boldsymbol{\chi}}_k^{MP}$ is different from the single-point mode $\hat{\chi}_{k_\ell}^{SP}$ obtained with Eq. (4.8).

In order to calculate the multi-point mode we could easily derive an equation similar to that of Eq. (4.10) for interval soft data, or Eq. (5.3) for probabilistic soft data. However such an equation would require an iterative numerical scheme that is not very efficient. A better approach is to use a maximization numerical algorithm which finds the maximum of the multivariate posterior pdf. This approach was found to be more efficient and it is the approach used in the numerical implementation.

7.4 Moments of the multi-point BME posterior pdf

The mean $\bar{\boldsymbol{x}}_{k|K}^{MP} = [\bar{x}_{k_1|K}^{MP} \dots \bar{x}_{k_\rho|K}^{MP}]$ of the multi-point posterior pdf $f_K^{MP}(\boldsymbol{\chi}_k)$ has components $\bar{x}_{k_\ell|K}^{MP} = \int d\boldsymbol{\chi}_k \chi_{k_\ell} f_K^{MP}(\boldsymbol{\chi}_k)$, $\ell = 1, \dots, \rho$. The MP superscript emphasizes that $\bar{x}_{k_\ell|K}^{MP}$ is the mean of the multi-point posterior pdf, to distinguish it from the mean of the single-point posterior pdf, $\bar{x}_{k_\ell|K}^{SP}$, given by Eq. (4.13) for interval soft data, and by Eq.

(5.5) for probabilistic soft data. For soft data of the probabilistic type the mean of the multi-point posterior pdf is obtained by writing

$$\begin{aligned}
\bar{x}_{k_\ell|K}^{MP} &= \int d\boldsymbol{\chi}_k \chi_{k_\ell} f_K^{MP}(\boldsymbol{\chi}_k) \\
&= A^{-1} \int_{-\infty}^{+\infty} d\chi_{k_\ell} \chi_{k_\ell} \int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) \underbrace{\int_{-\infty}^{+\infty} d\chi_{k_1} \cdots \int_{-\infty}^{+\infty} d\chi_{k_{\ell-1}} \int_{-\infty}^{+\infty} d\chi_{k_{\ell+1}} \cdots \int_{-\infty}^{+\infty} d\chi_{k_\rho} \phi(\boldsymbol{\chi}_{\text{map}}; \boldsymbol{\theta}, \mathbf{C}_{\text{map}})}_{\phi(\boldsymbol{\chi}_{\text{map}}^{\text{SP},\ell}; \boldsymbol{\theta}, \mathbf{C}_{\text{map}}^{\text{SP},\ell})} \\
&= \int_{-\infty}^{+\infty} d\chi_{k_\ell} \chi_{k_\ell} f_K^{\text{SP}}(\chi_{k_\ell}) \\
&= \bar{x}_{k_\ell|K}^{\text{SP}}
\end{aligned} \tag{7.12}$$

where $\boldsymbol{\chi}_{\text{map}}^{\text{SP},\ell} = [\boldsymbol{\chi}_{\text{hard}}^{\text{T}} \boldsymbol{\chi}_{\text{soft}}^{\text{T}} \chi_{k_\ell}]^{\text{T}}$, $\mathbf{C}_{\text{map}}^{\text{SP},\ell} = \begin{bmatrix} \mathbf{C}_{h,h} & \mathbf{C}_{h,s} & \mathbf{C}_{h,k_\ell} \\ \mathbf{C}_{s,h} & \mathbf{C}_{s,s} & \mathbf{C}_{s,k_\ell} \\ \mathbf{C}_{k_\ell,h} & \mathbf{C}_{k_\ell,s} & \mathbf{C}_{k_\ell,k_\ell} \end{bmatrix}$, $f_K^{\text{SP}}(\chi_{k_\ell})$ is the single-point posterior pdf for x_{k_ℓ} , and thus $\bar{x}_{k_\ell|K}^{MP}$ is equal to $\bar{x}_{k_\ell|K}^{\text{SP}}$.

In other words the BME mean estimator at a point \mathbf{p}_{k_ℓ} is the same whether considering several estimation points \mathbf{p}_{k_ℓ} , $\ell = 1, \dots, \ell, \dots, \rho$ jointly (multi-point approach), or whether considering only the point of interest \mathbf{p}_{k_ℓ} by itself (single point approach). This result is explained by the fact that the single-point posterior pdf $f_K^{\text{SP}}(\chi_{k_\ell})$ is the marginal pdf of the multi-point posterior pdf $f_K(\boldsymbol{\chi}_k)$ with respect to χ_{k_ℓ} . Hence the estimator minimizing the error variance must be the same whether using the single-point or multi-point approach. The practical implication of this finding is that when using the BME mean estimator, one can perform a multi-point estimation at the numerical of the single-point estimation.

Other statistical moments of interest for the multi-point pdf are the moments of order 2. In the case of soft data of probabilistic type, the moment of order 2 between points \mathbf{p}_{k_ℓ} and $\mathbf{p}_{k_{\ell'}}$ is the covariance $C_{k_\ell, k_{\ell'}|K}$ given by

$$\begin{aligned}
C_{k_\ell, k_{\ell'}|K} &= \overline{(x_{k_\ell} - \bar{x}_{k_\ell|K})(x_{k_{\ell'}} - \bar{x}_{k_{\ell'}|K})} \\
&= A^{-1} \int d\boldsymbol{\chi}_k (\chi_{k_\ell} - \bar{x}_{k_\ell|K})(\chi_{k_{\ell'}} - \bar{x}_{k_{\ell'}|K}) \int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) f(\boldsymbol{\chi}_{\text{map}}) \\
&= A^{-1} \phi(\boldsymbol{\chi}_{\text{hard}}, \boldsymbol{\theta}, \mathbf{C}_{h,h}) \int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) \phi(\boldsymbol{\chi}_{\text{soft}}; \mathbf{B}_{s|h} \boldsymbol{\chi}_h, \mathbf{C}_{s|h}) \\
&\quad \int d\boldsymbol{\chi}_k (\chi_{k_\ell} - \bar{x}_{k_\ell|K})(\chi_{k_{\ell'}} - \bar{x}_{k_{\ell'}|K}) \phi(\boldsymbol{\chi}_k; \mathbf{B}_{k|hs} \boldsymbol{\chi}_{hs}, \mathbf{C}_{k|hs}). \tag{7.13}
\end{aligned}$$

But we can write

$$\begin{aligned}
&\int d\boldsymbol{\chi}_k (\chi_{k_\ell} - \bar{x}_{k_\ell|K})(\chi_{k_{\ell'}} - \bar{x}_{k_{\ell'}|K}) \phi(\boldsymbol{\chi}_k; \mathbf{B}_{k|hs} \boldsymbol{\chi}_{\text{data}}, \mathbf{C}_{k|hs}) \\
&= \int d\boldsymbol{\chi}_k (\chi_{k_\ell} - \mathbf{B}_{k_\ell|hs} \boldsymbol{\chi}_{\text{data}})(\chi_{k_{\ell'}} - \mathbf{B}_{k_{\ell'}|hs} \boldsymbol{\chi}_{\text{data}}) \phi(\boldsymbol{\chi}_k; \mathbf{B}_{k|hs} \boldsymbol{\chi}_{\text{data}}, \mathbf{C}_{k|hs}) \\
&\quad + \underbrace{\int d\boldsymbol{\chi}_k (\chi_{k_\ell} - \mathbf{B}_{k_\ell|hs} \boldsymbol{\chi}_{\text{data}})(\mathbf{B}_{k_{\ell'}|hs} \boldsymbol{\chi}_{\text{data}} - \bar{x}_{k_{\ell'}|K}) \phi(\boldsymbol{\chi}_k; \mathbf{B}_{k|hs} \boldsymbol{\chi}_{\text{data}}, \mathbf{C}_{k|hs})}_0 \\
&\quad + \underbrace{\int d\boldsymbol{\chi}_k (\mathbf{B}_{k_\ell|hs} \boldsymbol{\chi}_{\text{data}} - \bar{x}_{k_\ell|K})(\chi_{k_{\ell'}} - \mathbf{B}_{k_{\ell'}|hs} \boldsymbol{\chi}_{\text{data}}) \phi(\boldsymbol{\chi}_k; \mathbf{B}_{k|hs} \boldsymbol{\chi}_{\text{data}}, \mathbf{C}_{k|hs})}_0 \\
&\quad + \int d\boldsymbol{\chi}_k (\mathbf{B}_{k_\ell|hs} \boldsymbol{\chi}_{\text{data}} - \bar{x}_{k_\ell|K})(\mathbf{B}_{k_{\ell'}|hs} \boldsymbol{\chi}_{\text{data}} - \bar{x}_{k_{\ell'}|K}) \phi(\boldsymbol{\chi}_k; \mathbf{B}_{k|hs} \boldsymbol{\chi}_{\text{data}}, \mathbf{C}_{k|hs}) \\
&= C_{k_\ell, k_{\ell'}|hs} + (\mathbf{B}_{k_\ell|hs} \boldsymbol{\chi}_{\text{data}} - \bar{x}_{k_\ell|K})(\mathbf{B}_{k_{\ell'}|hs} \boldsymbol{\chi}_{\text{data}} - \bar{x}_{k_{\ell'}|K}), \tag{7.14}
\end{aligned}$$

where $\mathbf{B}_{k_\ell|hs} = \mathbf{C}_{k_\ell, hs} \mathbf{C}_{hs, hs}^{-1}$, $\mathbf{B}_{k_{\ell'}|hs} = \mathbf{C}_{k_{\ell'}, hs} \mathbf{C}_{hs, hs}^{-1}$ and $C_{k_\ell, k_{\ell'}|hs} = C_{k_\ell, k_{\ell'}} - \mathbf{C}_{k_\ell, hs} \mathbf{C}_{hs, hs}^{-1} \mathbf{C}_{hs, k_{\ell'}}$.

Combining (7.13) and (7.14) we get

$$\begin{aligned}
C_{k_\ell, k_{\ell'}|K} &= C_{k_\ell, k_{\ell'}|hs} \\
&\quad + \frac{\int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) (\mathbf{B}_{k_\ell|hs} \boldsymbol{\chi}_{\text{data}} - \bar{x}_{k_\ell|K})(\mathbf{B}_{k_{\ell'}|hs} \boldsymbol{\chi}_{\text{data}} - \bar{x}_{k_{\ell'}|K}) \phi(\boldsymbol{\chi}_{\text{soft}}; \mathbf{B}_{s|h} \boldsymbol{\chi}_{\text{hard}}, \mathbf{C}_{s|h})}{\int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) \phi(\boldsymbol{\chi}_{\text{soft}}; \mathbf{B}_{s|h} \boldsymbol{\chi}_{\text{hard}}, \mathbf{C}_{s|h})}. \tag{7.15}
\end{aligned}$$

From this we derive the coefficient of correlation between any two estimation points given the physical knowledge available as

$$\rho_{k_\ell, k_{\ell'}|K} = \frac{C_{k_\ell, k_{\ell'}|K}}{\sqrt{C_{k_\ell, k_\ell|K} C_{k_{\ell'}, k_{\ell'}|K}}}. \quad (7.16)$$

This correlation coefficient is a feature of the multi-point approach. By considering several estimation points jointly, the multi-point approach provides not only estimated values, but also describe how estimated values at different location are correlated given the physical knowledge available.

Note that when $\ell = \ell'$ we obtain the variance $(\sigma_{k_\ell|K})^2$, i.e. $(\sigma_{k_\ell|K})^2 = C_{k_\ell, k_\ell|K}$.

Using Eq. (7.15) we get

$$\begin{aligned} (\sigma_{k_\ell|K})^2 &= C_{k_\ell, k_\ell|hs} + \frac{\int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) (\mathbf{B}_{k_\ell|hs} \boldsymbol{\chi}_{\text{data}} - \bar{x}_{k_\ell|K})^2 \phi(\boldsymbol{\chi}_{\text{soft}}; \mathbf{B}_{s|h} \boldsymbol{\chi}_{\text{hard}}, \mathbf{C}_{s|h})}{\int d\boldsymbol{\chi}_{\text{soft}} f_S(\boldsymbol{\chi}_{\text{soft}}) \phi(\boldsymbol{\chi}_{\text{soft}}; \mathbf{B}_{s|h} \boldsymbol{\chi}_{\text{hard}}, \mathbf{C}_{s|h})} \\ &= (\sigma_{k_\ell|K}^{\text{SP}})^2 \end{aligned} \quad (7.16)$$

where $(\sigma_{k_\ell|K}^{\text{SP}})^2$ is the variance at the estimation point \mathbf{p}_{k_ℓ} with respect to the single-point BME posterior pdf, as given in Eq. (5.7). In other words the error variance $\sigma_{k_\ell|K}^2$ at an estimation point \mathbf{p}_{k_ℓ} is the same whether using the single-point approach or multi-point approach.

By way of summary we have found that multi-point BME mapping produces the same mean estimator $\bar{x}_{k|K}$ and error variance $\sigma_{k_\ell|K}^2$ as in single-point BME mapping, and it provides additionally the coefficient $\rho_{k_\ell, k_{\ell'}|K}$ describing the correlation between any two estimated values given the physical knowledge available. However the true power of multi-point BME mapping resides in its ability to provide an error assessment of the joint

estimation of a S/TRF at several estimation points considered simultaneously. This is done using a unique feature of multi-point BME, called the BME confidence sets, which is presented next.

7.5 Uncertainty Assessment: Going Beyond the Confidence Intervals- The BME Confidence Sets

Traditionally in spatiotemporal mapping each estimation point is considered independently from other estimation points, what we refer to as the single-point approach. In this case a confidence interval provides a useful tool to assess the uncertainty associated with the estimated value. Typically one specifies a confidence level that is a function of the mapping situation, and the single-point BME posterior pdf is used to determine the bounds of an interval containing the estimated value within the required confidence level (or probability).

A more realistic assessment of the mapping error in the case of a multi-point mapping may be achieved using the concept of BME confidence sets. The latter is an extension of the classical confidence interval. Consider multi-point mapping where, in general, the multi-point posterior pdf $f_K(\boldsymbol{\chi}_k)$, $\boldsymbol{\chi}_k = [\chi_{k_1} \dots \chi_{k_p}]^T$, is non-Gaussian. A confidence set Φ_η is a set of values for $\boldsymbol{\chi}_k$ such that $P[\boldsymbol{\chi}_k \in \Phi_\eta] = \eta$. The confidence level η ($0 \leq \eta \leq 1$) depends on the required reliability of the mapping situation. Intuitively, as the confidence level increases so does the "size" of the corresponding confidence set. The size of the confidence set is defined as

$$\|\Phi_\eta\| = \int_{\boldsymbol{\chi}_k \in \Phi_\eta} d\boldsymbol{\chi}_k. \quad (7.17)$$

If, e.g., $\rho = 1$ (single-point analysis) and the confidence set is the interval $\Phi_\eta = [\chi_{k,l}, \chi_{k,u}]$, then its size is simply the interval length $\|\Phi_\eta\| = \chi_{k,u} - \chi_{k,l}$. Clearly,

for a given confidence level η there are several confidence sets Φ_η such that $P[\boldsymbol{\chi}_k \in \Phi_\eta] = \eta$. We define the *BME confidence set* as the confidence set with the smallest size. In our example ($\rho = 1$), the BME confidence set is simply the confidence interval $\Phi_\eta = [\chi_{k,l}, \chi_{k,u}]$ with the smallest length $\chi_{k,u} - \chi_{k,l}$ such that $P[\boldsymbol{\chi}_k \in \Phi_\eta] = \eta$. In this context, the following result is useful in determining BME confidence sets. The high probability density set defined as

$$\Lambda_\beta = \{\boldsymbol{\chi}_k : f_K(\boldsymbol{\chi}_k) / f_K(\hat{\boldsymbol{\chi}}_k) \geq 1 - \beta\} \quad (7.18)$$

is a BME confidence set $\Phi_{\eta(\beta)}$ with confidence level given by $\eta(\beta) = P[\boldsymbol{\chi}_k \in \Lambda_\beta]$.

The derivation of Eq. (7.18) is given in Appendix G.

High probability density sets are easily constructed from the posterior pdf, in which case Eq. (7.18) provides the corresponding BME confidence set. Note that, if the inverse function for $\eta(\beta)$ exists, then one can directly obtain a BME confidence set Φ_η for a selected confidence level η . In this case, the Φ_η is

$$\Phi_\eta = \Lambda_\beta : P[\boldsymbol{\chi}_k \in \Lambda_\beta] = \eta. \quad (7.19)$$

A useful corollary of the analysis above for $\rho = 1$ is when the BME confidence set is a single interval. In this case, the

$$\Phi_\eta = [\chi_{k,l}, \chi_{k,u}] \quad (7.20)$$

is a *BME confidence interval* such that $f_K(\chi_{k,l}) = f_K(\chi_{k,u})$ and $P[\chi_{k,l} \leq \chi_k \leq \chi_{k,u}] = \eta$.

Traditionally, confidence intervals are taken such that the probability of the estimated value falling on the left part of the interval is equal to the probability of falling on the right part. In contrast, the BME confidence intervals are as small as possible. BME confidence sets are a generalization of the confidence intervals for multi-point analysis, where several estimated values are considered simultaneously. The same concepts can be easily extended to the vector estimation of several random fields jointly.

7.6 Simulated Case Studies

Three simulated case studies are presented to illustrate BME mapping with the multi-point approach. The first two case studies illustrate the difference between multi-point mapping and single-point mapping when using the mode estimator. The third case study focuses on BME confidence sets.

7.6.1 Multi-point Mapping versus single-point Mapping: Test Case 1

This first simulated test case illustrates the difference between multi-point and single-point approaches using two estimation points. Using the stochastic simulation method described in Chapter II, a realization of the S/TRF is generated, with $m_h=10$ hard data points, 3 soft data points of the interval type, and 2 estimation points. The location of the estimation points, hard data points, and soft data points is shown in Fig. 7.1. Also shown are the value of the S/TRF $X(\mathbf{p})$ at these points obtained by LU simulation using a Gaussian covariance model given by $c(r) = c_o \exp(-r^2 / a_r^2)$, where the range parameter is $a_r=0.4$ and the sill is $c_o=1.0$. At the soft data points we only consider the soft knowledge $x_{m_h+i} \in [l_i, u_i]$, $i=m_h+1, \dots, m$, (the simulated values are ignored). The intervals are constructed such that they contain the simulated values, and such that their length is 1.5. As usual the simulated values at the estimation are considered "true" value, and they are (re-)estimated using only hard and soft (interval) data. The two estimation points of Fig.

7.1 are numbered 1 and 2, such that estimation point 1 is the point on the left and estimation point 2 is the point on the right in Fig. 7.1.

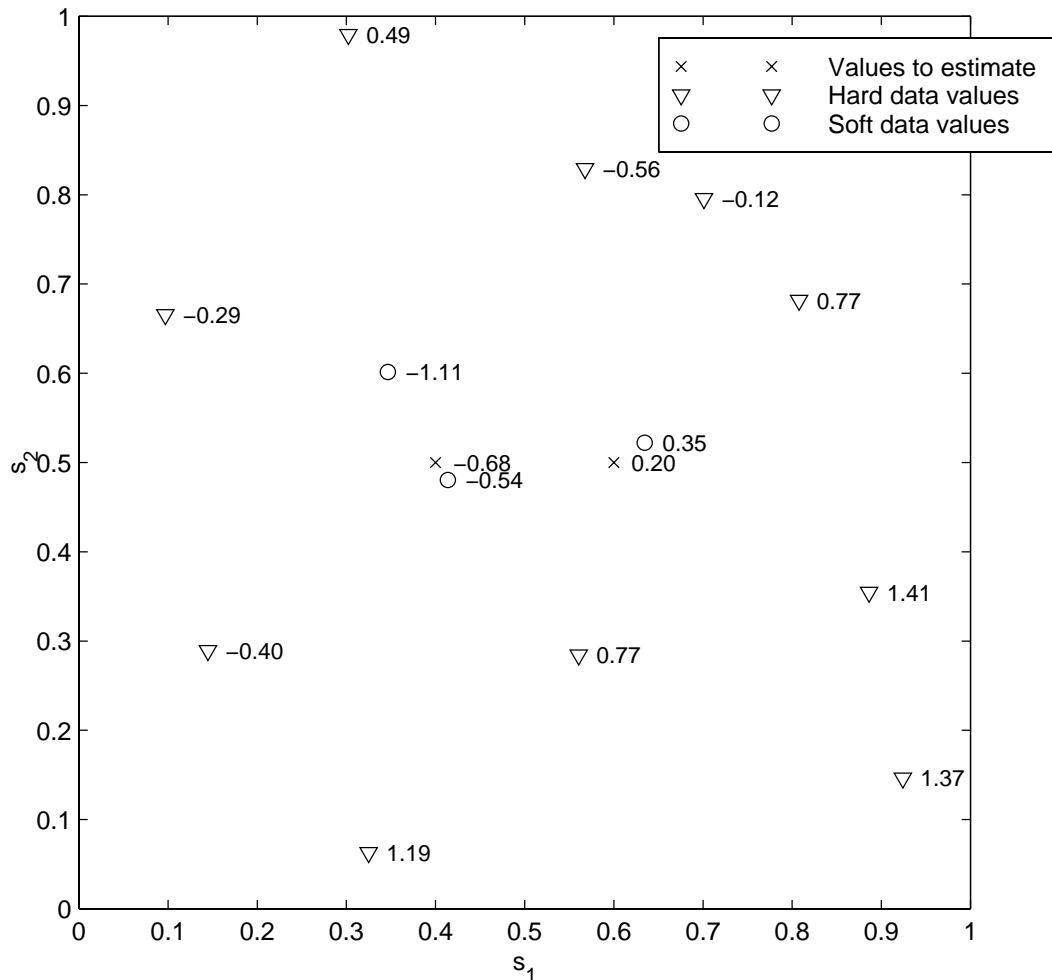


Figure 7.1: A realization of the S/TRF $X(\mathbf{p})$. The location of the estimation points, hard data points and soft data points are shown with crosses, triangles, and circles, respectively. The simulated (true) value the S/TRF is written next to each point.

Using the single-point BME approach we may calculate the posterior BME pdf for each of the two estimation points considered independently. The single-point posterior pdf $f_K^{SP}(\chi_{k_1})$ for estimation point 1 is shown in Fig. 7.2a, and $f_K^{SP}(\chi_{k_2})$ for estimation point 2 is shown in Fig. 7.2b. Also shown in Fig. 7.2 are the mode values of the single-point posterior pdf's, which are $\hat{\chi}_{k_1}^{SP} = -0.54$ for first (left) estimation point, and $\hat{\chi}_{k_2}^{SP} = 0.14$ for

second (right) estimation point.. These values are the most probable values for each estimation points considered separately given the hard data and soft data, and they are in good agreement of the "true" (simulated) values shown in Fig. 7.1, i.e. $\chi_{k_1} = -0.68$ and $\chi_{k_2} = 0.20$.

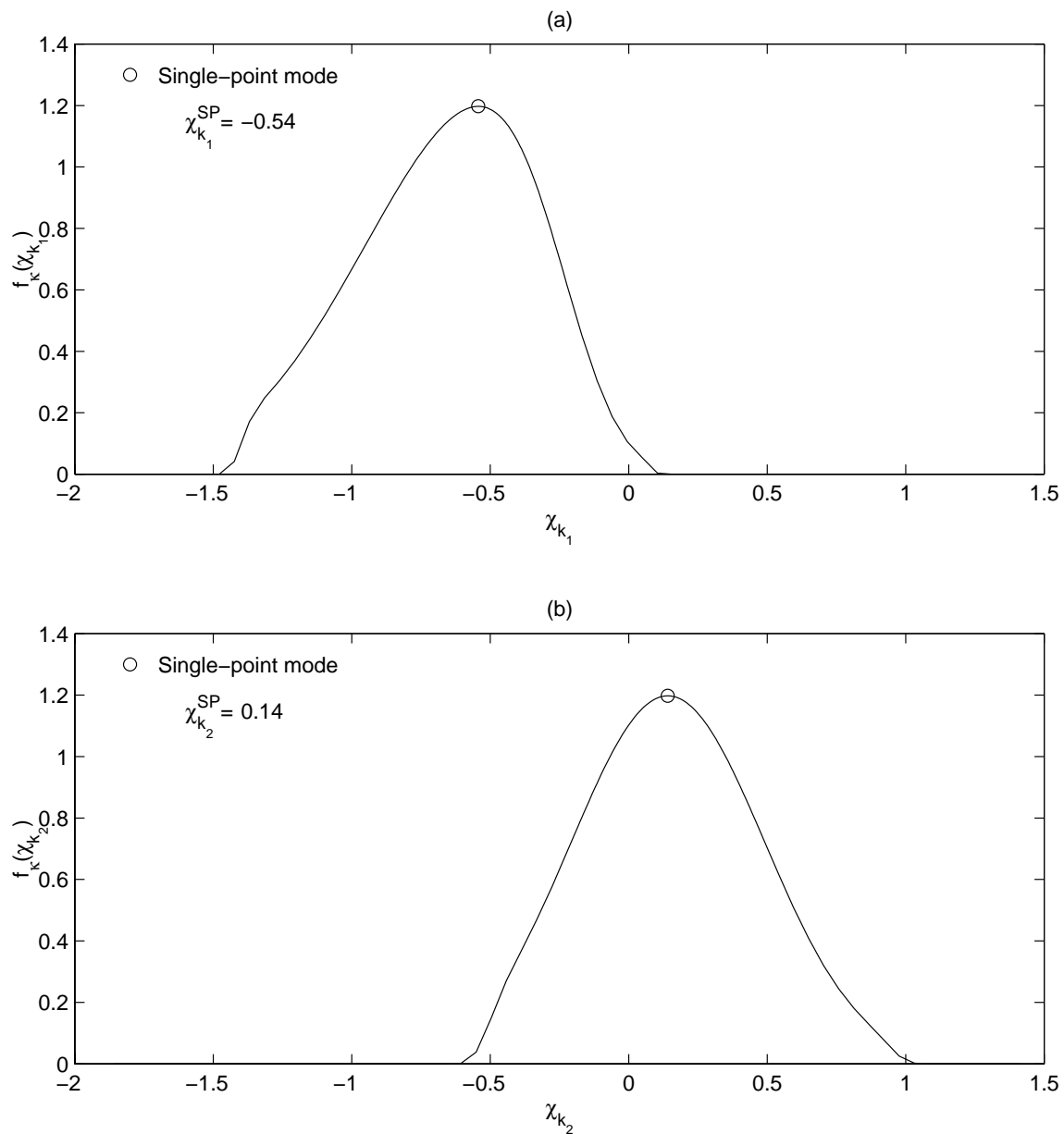


Figure 7.2: Single-point posterior pdf for (a) estimation point 1 on the left, $f_K^{SP}(\chi_{k_1})$, and (b) estimation point 2 on the right, $f_K^{SP}(\chi_{k_2})$. The single-point mode values, shown with circles, are (a) $\hat{\chi}_{k_1}^{SP} = -0.54$ for first (left) estimation point, and (b) $\hat{\chi}_{k_2}^{SP} = 0.14$ for second (right) estimation point.

In the multi-point approach the estimation points are considered jointly, allowing to calculate the multivariate posterior pdf given by Eq. (7.9). For the two estimation points considered in this test case, we obtain the bivariate posterior pdf $f_K^{MP}(\chi_{k_1}, \chi_{k_2})$ shown in Fig. 7.3. In the figure $f_K^{MP}(\chi_{k_1}, \chi_{k_2})$ is normalized by its maximum value, $\max(f_K^{MP})$, i.e. the contour lines of equal value of $f_K^{MP}(\chi_{k_1}, \chi_{k_2})/\max(f_K^{MP})$ are shown, with values of $0.1+0.1k$, $k=0, \dots, 9$. It is clear from the figures that the multi-point posterior pdf $f_K^{MP}(\chi_{k_1}, \chi_{k_2})$ of Fig. 7.3 is much informative than the single-point posterior pdf's of Fig 7.2: While the multi-point posterior pdf's provides any single-point pdf as a marginal pdf, it also describes the interdependence at the two estimation points. Also shown in Fig. 7.3 with a circle is the location of the mode vector $\hat{\chi}_k^{MP} = [\hat{\chi}_{k_1}^{MP} \ \hat{\chi}_{k_2}^{MP}]$, with values of $\hat{\chi}_{k_1}^{MP} = -0.54$ and $\hat{\chi}_{k_2}^{MP} = 0.19$. These values are the most probable values at the estimation point considered jointly. We note (for this realization of the S/TRF) that while at the first estimation point the multi-point and single-point estimates are the same, at the second estimation point the error dropped from 0.06 for the single-point estimate to only 0.01 for the multi-point estimate. This test case illustrates the difference between the multi-point mapping and single-point mapping based on only two estimation point. Next we consider a case involving 25 estimation points.

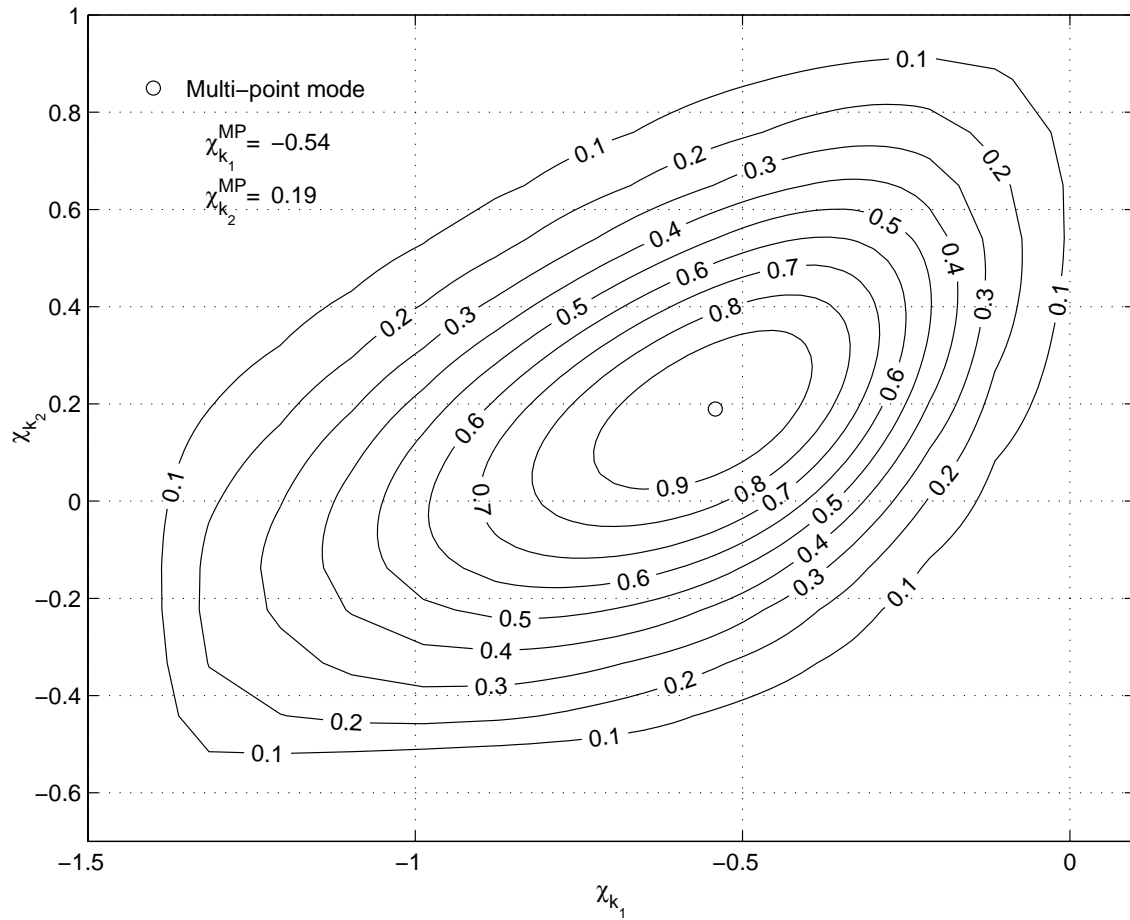


Figure 7.3: Multi point posterior pdf $f_K^{MP}(\chi_{k_1}, \chi_{k_2})$. The contour lines of equal value of $f_K^{MP}(\chi_{k_1}, \chi_{k_2}) / \max(f_K^{MP})$ are shown, with values of $0.1 + 0.1k$, $k=0, \dots, 9$. The multi-point mode (○) is $\hat{\chi}_{k_1}^{MP} = -0.54$ and $\hat{\chi}_{k_2}^{MP} = 0.19$, such that $f_K^{MP}(\hat{\chi}_{k_1}^{MP}, \hat{\chi}_{k_2}^{MP}) / \max(f_K^{MP}) = 1$

7.6.2 Multi-point Mapping Versus Single-Point Mapping: Test Case 2

In this second test case we consider a mapping situation involving 25 estimation points located on a grid in the neighborhood of $m_h=2$ hard data points, and $m - m_h=5$ points with soft data of the interval type. The location of the 2 hard data points, the 5 soft data points and the grid of estimation points are shown in Fig 7.4. Also shown are the values of a two-dimensional Spatial Random Field (SRF) generated using the LU decomposition method with a zero mean and the Gaussian covariance function $c_x(r) = c_o \exp[-r^2 / a_r^2]$,

where $r = |s - s'|$, $a_r = 0.6$ and $c_o = 1.0$. At the soft data points we only consider the soft knowledge $x_{m_h+i} \in [l_i, u_i]$, $i = m_h+1, \dots, m$, (the simulated values are ignored).

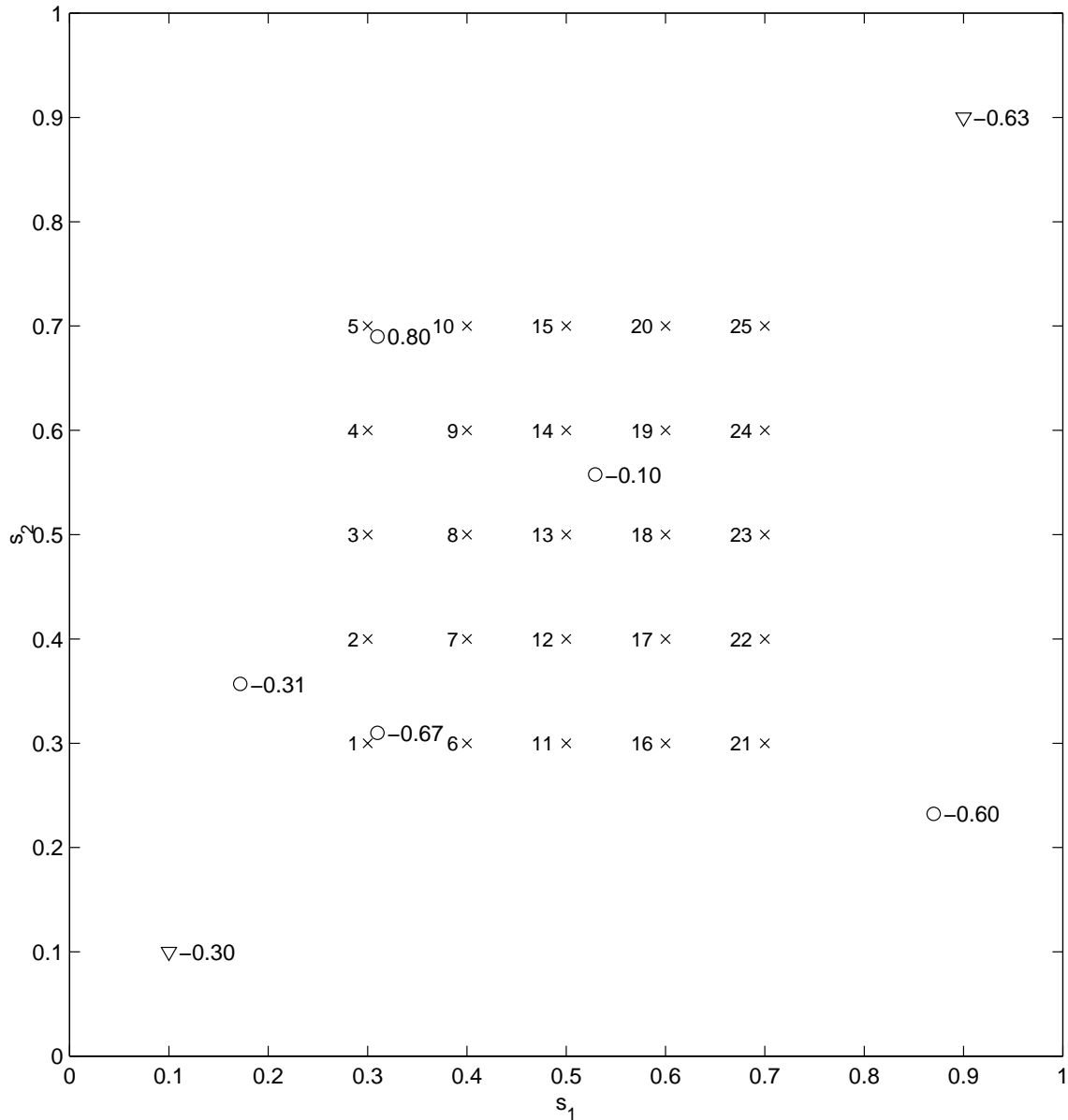


Figure 7.4: Location of points with estimation point numbers (\times), hard data values (\circ) and soft data values (Δ).

Using hard data values and the soft data intervals, BME produces multi-point estimates at the 25 estimation point. The resulting multi-point based estimates at these 25 points is used to draw the contour lines of equal values of estimated values shown in Fig.

7.5 using plain line. Also shown in dashed line are the contour lines of the estimated values using single-point BME. The estimation contour map obtained using multi-point BME is noticeably different than the one obtained using single-point BME. Indeed, the multi-point approach offers a description that accounts for the inter-dependence between estimation points.

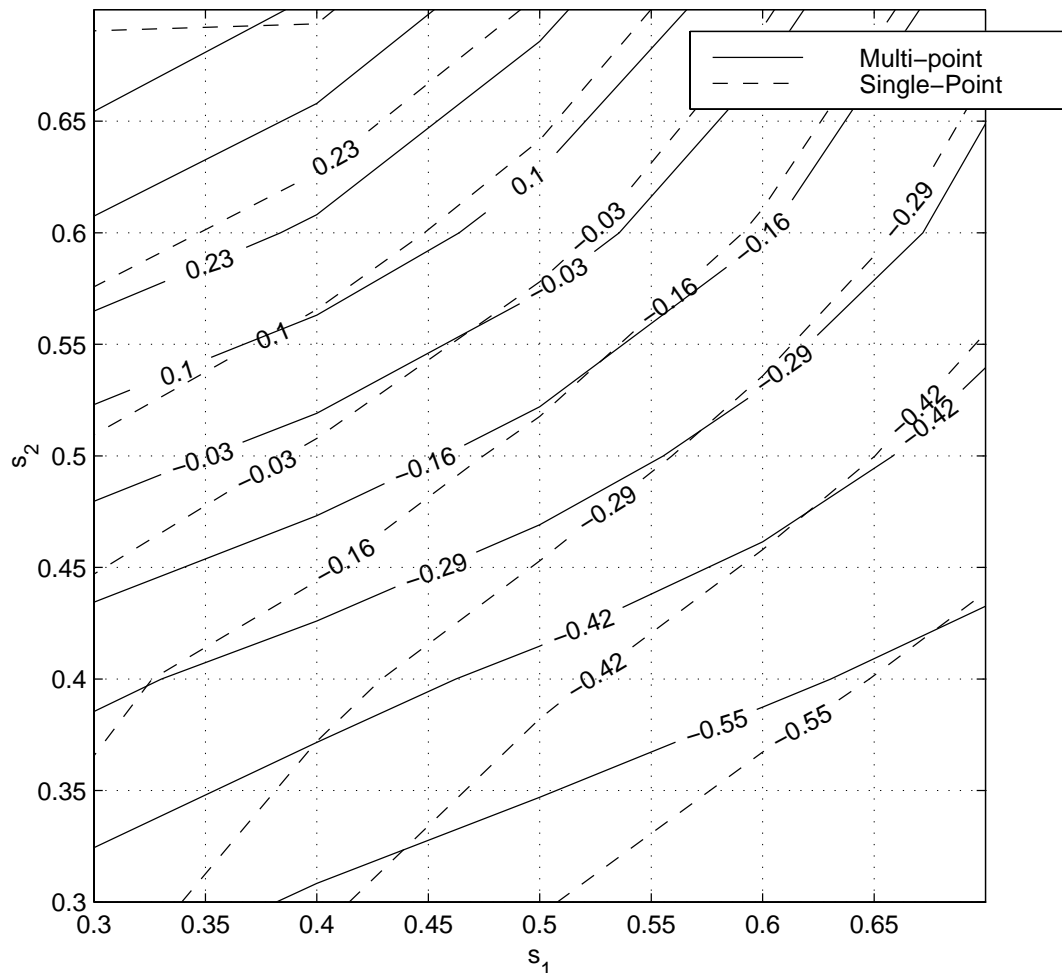


Figure 7.5: Map of estimated values at the 25 estimation points. Note the area covers only part of the original domain. The contour lines of equal estimated values are shown using (—) for multi-point BME (- -) for single point BME.

As in the previous test-case, the inter-dependence between two specific estimation points may be further examined by looking at their bivariate pdf. Consider for example the estimation points 1 and 6 (located at the bottom left of the grid). The bivariate pdf

$f_K^{MP}(\chi_{k_1}, \chi_{k_6})$ for these two estimation points, shown in Fig. 7.6, provides a good description of their interdependence.

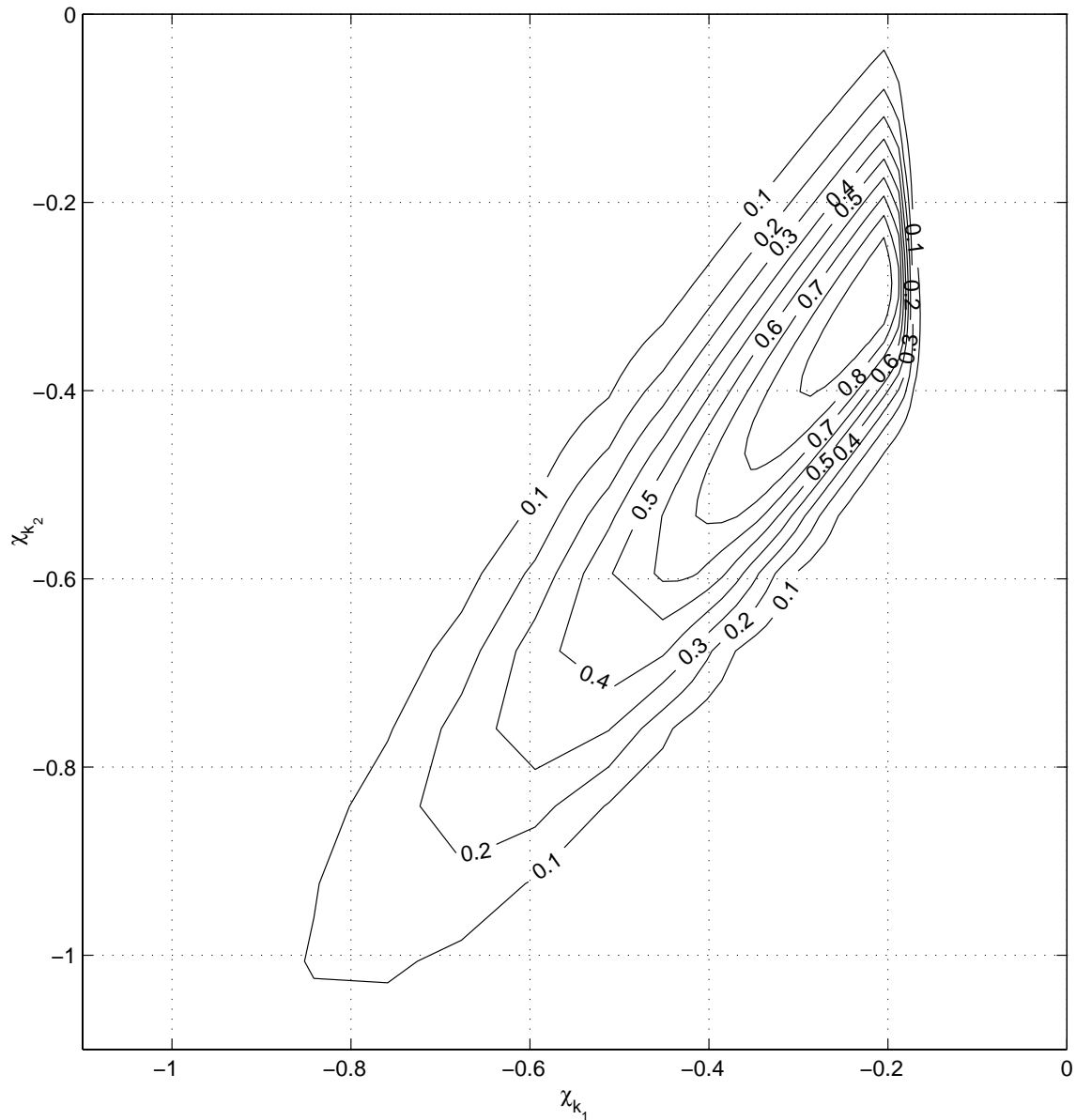


Figure 7.6: Bivariate posterior pdf $f_K^{MP}(\chi_{k_1}, \chi_{k_6})$ for estimation point 1 and 6. The contour lines of equal value of $f_K^{MP}(\chi_{k_1}, \chi_{k_6})/\max(f_K^{MP})$ are shown, with values of $0.1+0.1k$, $k=0, \dots, 9$. The mode is at $\hat{\chi}_{k_1} = -0.204$ and $\hat{\chi}_{k_6} = -0.286$.

A stochastic estimation method produces a set of estimated values as well as an assessment of estimation error. In this test case we compared the estimated values produced by multi-point mapping and single point mapping, but we did not compare the error assessment provided by these two approaches. This is done in the next test-case, where we look at the BME confidence sets provided by the multi-point mapping approach using two estimation points.

7.6.3 BME Confidence Sets in Action

In this case study we consider the joint estimation of a Spatial Random Field (SRF) $X(\mathbf{p})$ at 2 estimation points, given 2 hard data points and 2 soft data points. The location of the estimation points, hard data points and soft data points are shown in Fig. 7.7.

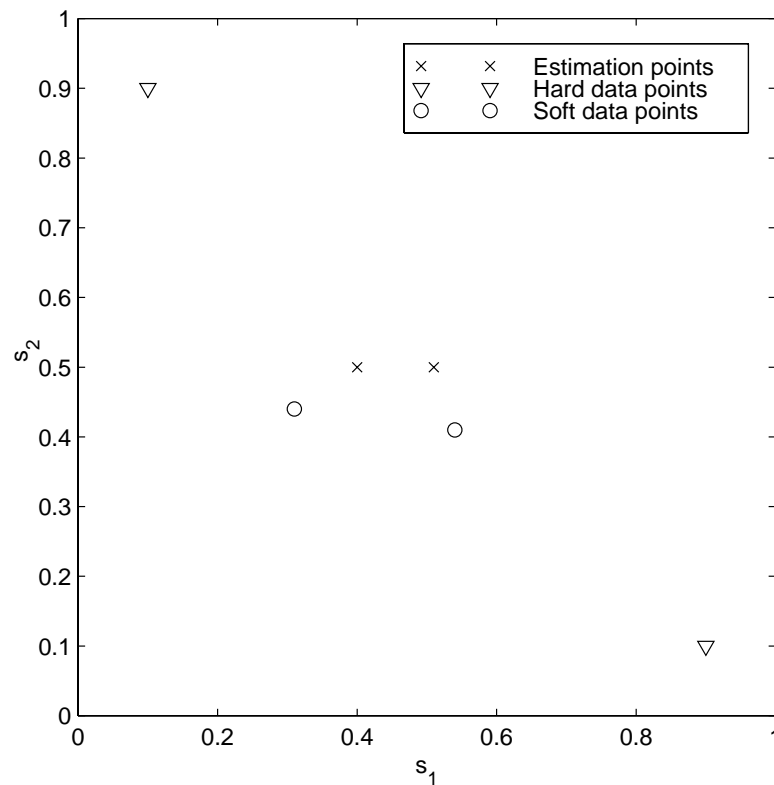


Figure 7.7: Location of the estimation, hard and soft data points.

The soft data considered is of the probability type, corresponding to the generator pdf $f_v(\vartheta)$ shown in Fig. 7.8. Using the stochastic simulation method described in Chapter II and using the generator pdf $f_v(\vartheta)$ of Fig. 7.8, we obtain a realization of the SRF $X(\mathbf{p})$ at the estimation points, hard data points and soft data points. As way of summary, the stochastic simulation method involves using the LU decomposition method to generate values of $X(\mathbf{p})$ at the estimation points and hard data points, and then using the generator pdf $f_v(\vartheta)$ to produce the soft data of probabilistic type. As usual, the simulated value of $X(\mathbf{p})$ at the estimation point are considered the "true" values. The knowledge of these true values is ignored during the estimation process, so that they are (re-)estimated using only the hard and soft data. Then the difference between true value and estimated value forms the estimation error.

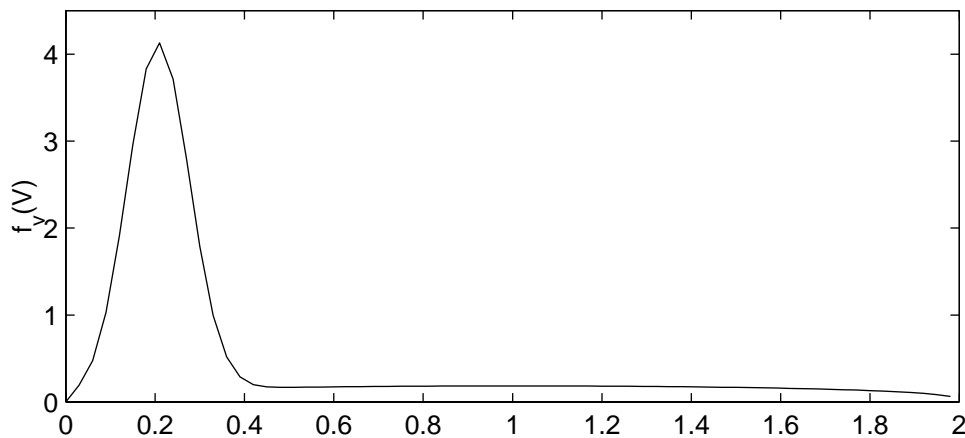


Figure 7.8: Probability distribution function $f_v(\vartheta)$ used to generate the soft pdf realizations.

Let the random variables x_{k_1} and x_{k_2} represent the SRF $X(\mathbf{p})$ at the estimation points \mathbf{p}_{k_1} and \mathbf{p}_{k_2} , where \mathbf{p}_{k_1} is the estimation point on the left in Fig. 7.7 and \mathbf{p}_{k_2} is on the right. The simulated value for $X(\mathbf{p})$ is $\chi_{k_1} = -1.887$ at \mathbf{p}_{k_1} and $\chi_{k_2} = -1.777$ at \mathbf{p}_{k_2} . Using the multi-point BME approach we calculate the two-point posterior pdf $f_K(\chi_{k_1}, \chi_{k_2})$. The mode of this two-point pdf provides the multi-point estimate, which we

calculate to be $\hat{\chi}_{k_1}^{MP} = -1.610$ and $\hat{\chi}_{k_2}^{MP} = -1.668$. For comparison purpose, we also calculate $f_K(\chi_{k_1})$ and $f_K(\chi_{k_2})$ using single point analysis, and we find the corresponding single-point mode estimate to be $\hat{\chi}_{k_1}^{SP} = -1.605$ and $\hat{\chi}_{k_2}^{SP} = -1.655$. We note that the multi-point mode estimate provides some improvements over the single-point mode estimate, however what we are interested in this case study is the error assessment for the joint estimate of x_{k_1} and x_{k_2} . This error assessment is best described using BME confidence sets.

The BME confidence sets Φ_η for multi-point mapping are a generalization of the confidence interval idea of single-point mapping. The BME confidence sets obtained using Eq. (7.19) for x_{k_1} and x_{k_2} are shown in Fig. 7.9a. The lines delineate the contour of the confidence sets, and the labels indicate the confidence levels (ranging from 0.1 to 0.9 with increments of 0.1). Each contour represents the smallest set of values for x_{k_1} and x_{k_2} for the confidence probability noted in the label. These sets are small because they take into consideration the correlation between x_{k_1} and x_{k_2} .

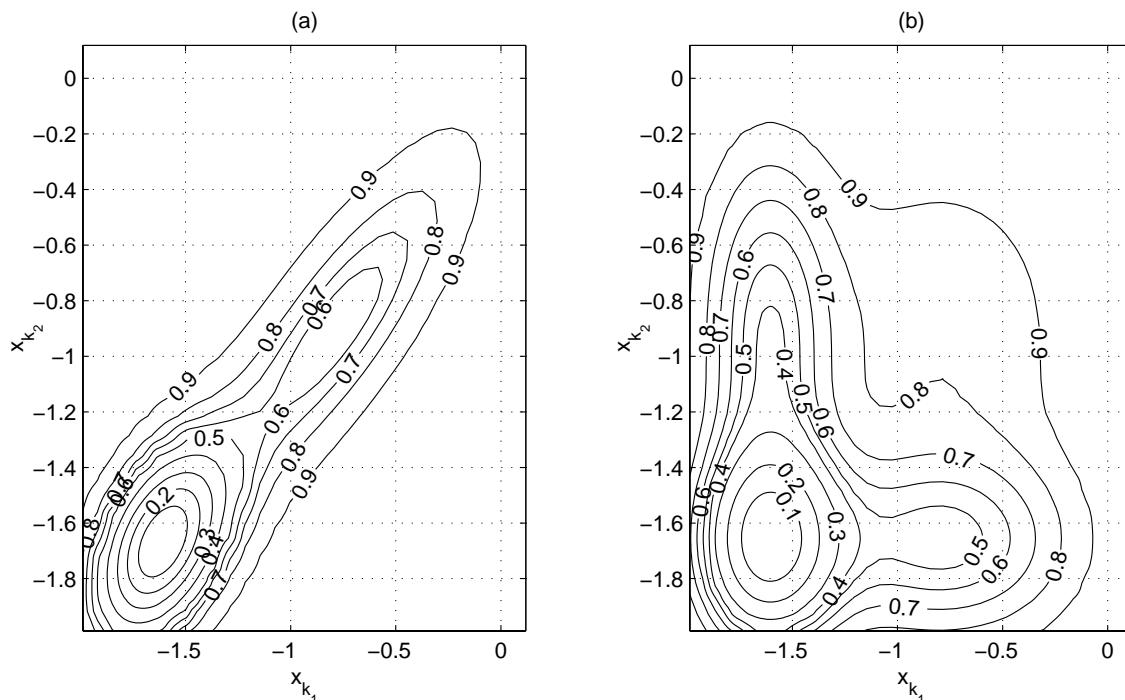


Figure 7.9: BME confidence sets Φ_η for x_{k_1} and x_{k_2} constructed from (a) multi-point analysis, and (b) single-point analysis using the BME confidence intervals.

For comparison purposes, in Fig. 7.9b we plot the confidence sets which are obtained if we were to replace the bivariate pdf $f_K(\mathcal{X}_{k_1}, \mathcal{X}_{k_2})$ of the two-point analysis with the product of univariate pdf $f_K(\mathcal{X}_{k_1})$ and $f_K(\mathcal{X}_{k_2})$ of the single-point analysis --which corresponds to (incorrectly) assuming that the variables x_{k_1} and x_{k_2} are independent. It is evident from Figs. 7.9a and 7.9b that the shapes of the confidence sets are significantly different, which means that Fig. 7.9b provides an incorrect representation of the real situation. The BME confidence set size $\|\Phi_\eta\|$ --which is equal to the area surrounded by the corresponding confidence probability η -contour in Fig. 7.9 --is plotted in Fig. 7.10 as a function of η for both the multi-point case of Fig. 7.9a and the single-point case of Fig. 7.9b. From Fig. 7.10 we note that for any η the size of the corresponding multi-point confidence set of Fig. 7.9a is consistently smaller than the size of the single-point confidence set of Fig. 7.9b. Since a stochastic estimation method is considered more accurate when it produces smaller confidence sets at a given confidence level, this simulation study demonstrate that the multi-point confidence sets provide a considerable improvement over single-point confidence sets.

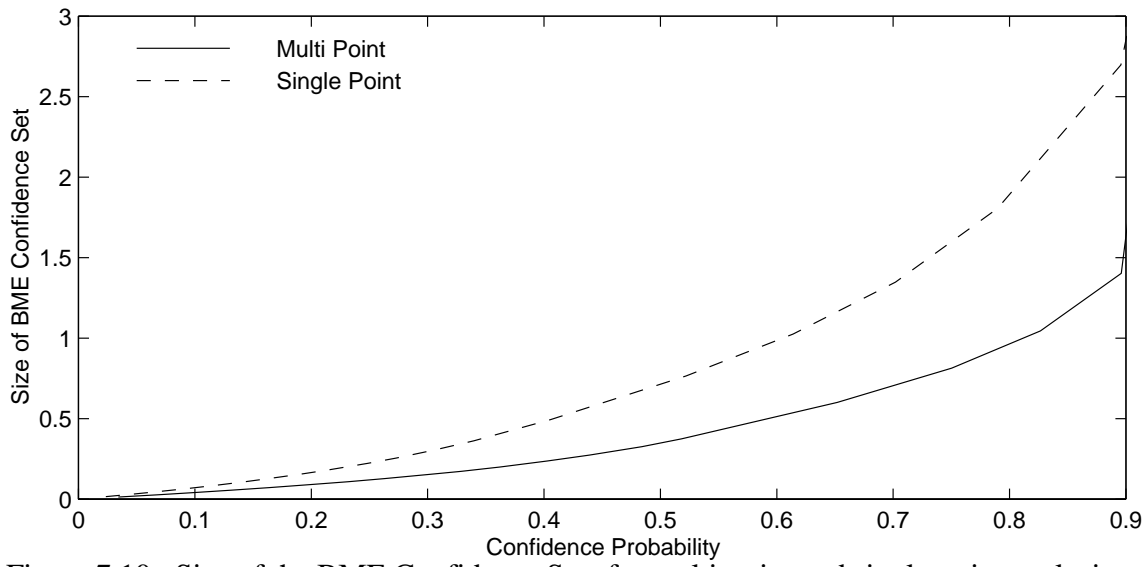


Figure 7.10: Size of the BME Confidence Sets for multi-point and single-point analysis