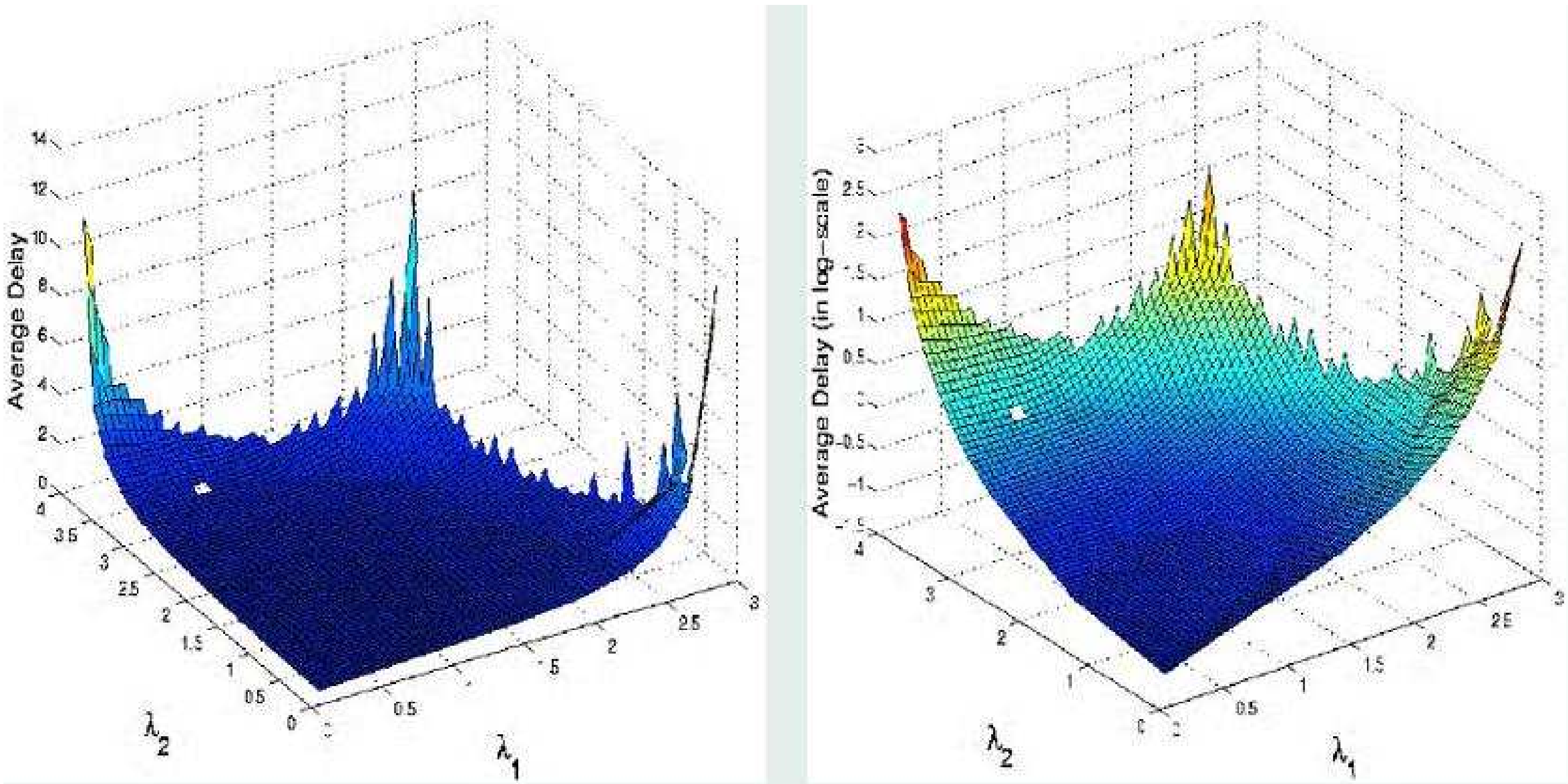
Inverse Problem for Complex Computer Codes

Pritam Ranjan

Simon Fraser University

Application

- Dell - IBM example



Inverse problem

Inverse problem

- Different than standard inverse problem?

Inverse problem

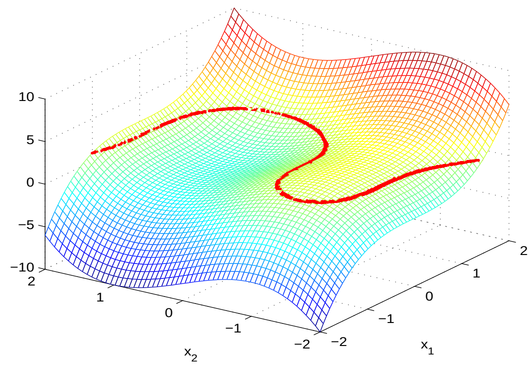
- Different than standard inverse problem?
- Simulator is expensive

Inverse problem

- Different than standard inverse problem?
- Simulator is expensive
- Budget is fixed (say n)

Formulation of the problem

- Interested in only a contour $S(a) = \{x : y(x) = a\}$



- Choice of design points

Steps of the solution

Steps of the solution

1. Start with a good initial design of size $n_0 < n$

Steps of the solution

1. Start with a good initial design of size $n_0 < n$
2. Fit a surrogate response surface

Steps of the solution

1. Start with a good initial design of size $n_0 < n$
2. Fit a surrogate response surface
3. Select a new trial aiming to improve upon the information of the underlying contour

Steps of the solution

1. Start with a good initial design of size $n_0 < n$
2. Fit a surrogate response surface
3. Select a new trial aiming to improve upon the information of the underlying contour
4. Run the trial, get augmented data

Steps of the solution

1. Start with a good initial design of size $n_0 < n$
2. Fit a surrogate response surface
3. Select a new trial aiming to improve upon the information of the underlying contour
4. Run the trial, get augmented data
5. Repeat Step 2 - Step 4 until budget allows

Steps of the solution

1. Start with a good initial design of size $n_0 < n$
2. Fit a surrogate response surface
3. Select a new trial aiming to improve upon the information of the underlying contour
4. Run the trial, get augmented data
5. Repeat Step 2 - Step 4 until budget allows
6. Extraction of the contour $S(a)$, where, $S(a) = \{x : y(x) = a\}$

Overview

- Response surface
- Selection of new trials
- Contour extraction
- Example
- Summary
- Concluding remarks
- Current work

Surrogate model

- Deterministic simulators

$$y(x_i) = \mu + z(x_i); \qquad i = 1, \ldots, n,$$

Surrogate model

- Deterministic simulators

$$y(x_i) = \mu + z(x_i); \qquad i = 1, \ldots, n,$$

- Use Gaussian Process model

$$z(\mathbf{x}) \sim N(\mathbf{0}, \sigma^2 \mathbf{R})$$

Surrogate model

- Deterministic simulators

$$y(x_i) = \mu + z(x_i); \qquad i = 1, \ldots, n,$$

- Use Gaussian Process model

$$z(\mathbf{x}) \sim N(\mathbf{0}, \sigma^2 \mathbf{R})$$

- Separable correlation structure

$$corr(z(x_i), z(x_j)) = R_{ij} = \prod_{k=1}^{d} \exp\left\{-\theta_k (x_{ik} - x_{jk})^2\right\}$$

Predicted surface

- Model fitting:
  (a) Relatively few parameters to estimate $(\theta_1, ..., \theta_d, \mu, \sigma^2)$
  (b) Bayesian or Likelihood method of estimation
  (c) Easy to modify the model (smoothness)

Predicted surface

- Model fitting:

  (a) Relatively few parameters to estimate ($\theta_1, ..., \theta_d, \mu, \sigma^2$)

  (b) Bayesian or Likelihood method of estimation

  (c) Easy to modify the model (smoothness)

- Predicted surface

$$\hat{y}(x) = \hat{\mu} + r'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{1_n}\hat{\mu})$$

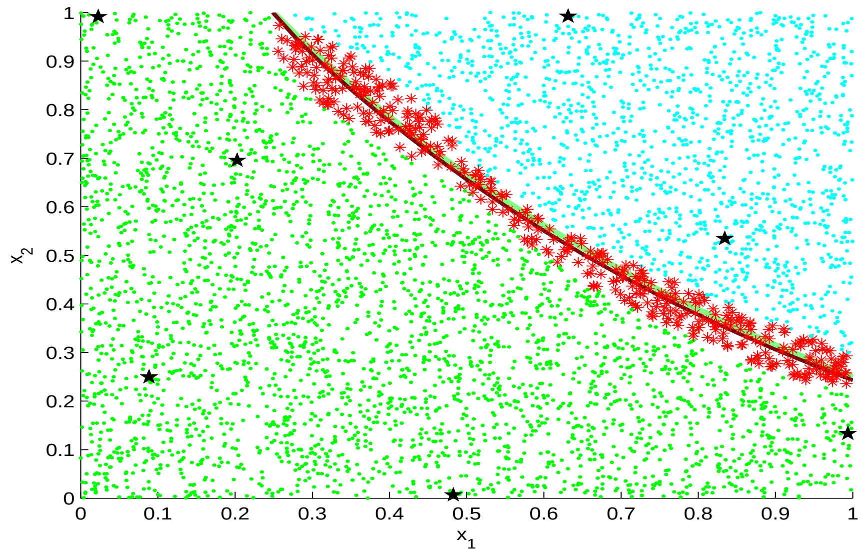
where $r = (r_1(x), ..., r_n(x))', \ \& \ r_i(x) = corr(z(x), z(x_i))$

Selection of new trials

- Choose $x \in \chi$ s.t. $\hat{y}(x) \in (a - \epsilon, a + \epsilon)$

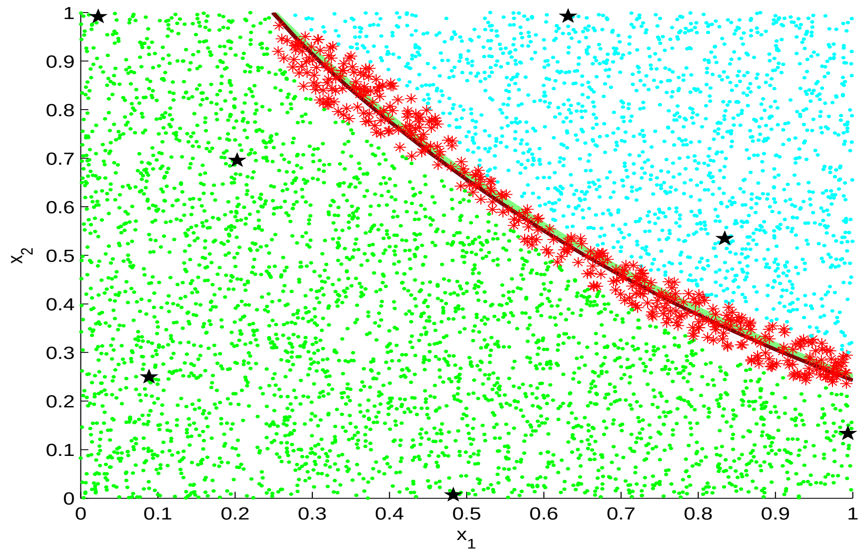
Selection of new trials

- Choose $x \in \chi$ s.t. $\hat{y}(x) \in (a - \epsilon, a + \epsilon)$



Selection of new trials

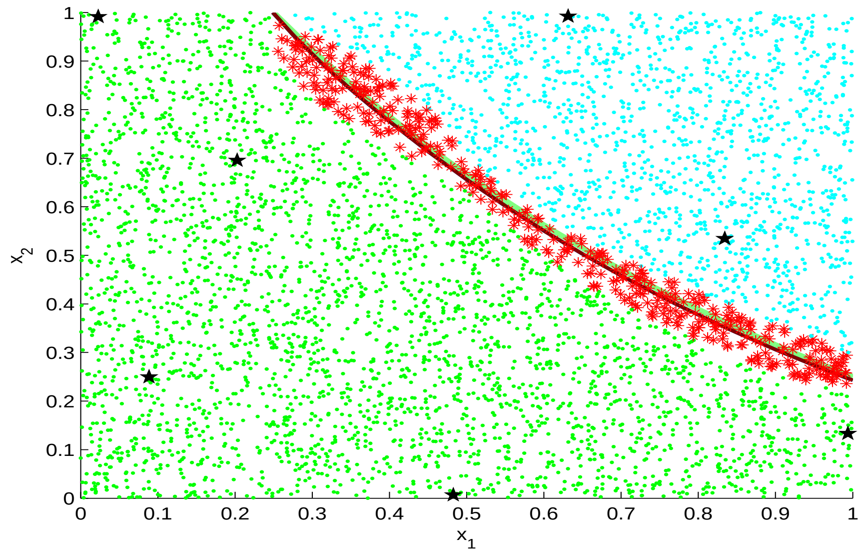
- Choose $x \in \chi$ s.t. $\hat{y}(x) \in (a - \epsilon, a + \epsilon)$



$$
\begin{aligned}
I(x) &= 0 & \text{if } y(x) \leq a - \epsilon(x) \\
&= \epsilon^2(x) - (y(x) - a)^2 & \text{if } a - \epsilon(x) < y(x) \leq a + \epsilon(x) \\
&= 0 & \text{if } a + \epsilon(x) < y(x)
\end{aligned}
$$

Selection of new trials

- Choose $x \in \chi$ s.t. $\hat{y}(x) \in (a - \epsilon, a + \epsilon)$



$$
\begin{aligned}
I(x) &= 0 && \text{if } y(x) \le a - \epsilon(x) \\
&= \epsilon^2(x) - (y(x) - a)^2 && \text{if } a - \epsilon(x) < y(x) \le a + \epsilon(x) \\
&= 0 && \text{if } a + \epsilon(x) < y(x)
\end{aligned}
$$

- Hence, define $I(x) = \epsilon^2(x) - min\{(y(x) - a)^2, \epsilon^2(x)\}$ and $\epsilon(x) = \alpha s(x)$

Selection of new trials (contd.)

- Temptation : maximize the improvement function

Selection of new trials (contd.)

- Temptation : maximize the improvement function
- Possibility : segment of the contour was not detected

Selection of new trials (contd.)

- Temptation : maximize the improvement function
- Possibility : segment of the contour was not detected
- Expected Improvement

$$\begin{aligned} E[I(x)] &= [\alpha^2 s^2(x) - (\hat{y}(x) - a)^2]\left[\Phi\left(\frac{a-\hat{y}(x)}{s(x)} + \alpha\right) - \Phi\left(\frac{a-\hat{y}(x)}{s(x)} - \alpha\right)\right] \\ &+ 2(\hat{y}(x) - a)s^2(x)\left[\phi\left(\frac{a-\hat{y}(x)}{s(x)} + \alpha\right) - \phi\left(\frac{a-\hat{y}(x)}{s(x)} - \alpha\right)\right] \\ &- \int_{a-\alpha s(x)}^{a+\alpha s(x)} (\hat{y}(x) - y)^2 \phi\left(\frac{\hat{y}(x) - y}{s(x)}\right) dy \end{aligned}$$

Selection of new trials (contd.)

- Temptation : maximize the improvement function
- Possibility : segment of the contour was not detected
- Expected Improvement

$$
\begin{aligned}
E[I(x)] \quad &= \quad [\alpha^2 s^2(x) - (\hat{y}(x) - a)^2] \left[\Phi\left(\frac{a - \hat{y}(x)}{s(x)} + \alpha\right) - \Phi\left(\frac{a - \hat{y}(x)}{s(x)} - \alpha\right) \right] \\
&+ \quad 2(\hat{y}(x) - a) s^2(x) \left[\phi\left(\frac{a - \hat{y}(x)}{s(x)} + \alpha\right) - \phi\left(\frac{a - \hat{y}(x)}{s(x)} - \alpha\right) \right] \\
&- \quad \int_{a - \alpha s(x)}^{a + \alpha s(x)} (\hat{y}(x) - y)^2 \phi\left(\frac{\hat{y}(x) - y}{s(x)}\right) dy
\end{aligned}
$$

- Selects trial from $(a - \epsilon, a + \epsilon)$

Selection of new trials (contd.)

- Temptation : maximize the improvement function
- Possibility : segment of the contour was not detected
- Expected Improvement

$$\begin{aligned} E[I(x)] &= [\alpha^2 s^2(x) - (\hat{y}(x) - a)^2]\left[\Phi\left(\frac{a-\hat{y}(x)}{s(x)} + \alpha\right) - \Phi\left(\frac{a-\hat{y}(x)}{s(x)} - \alpha\right)\right] \\ &+ 2(\hat{y}(x) - a)s^2(x)\left[\phi\left(\frac{a-\hat{y}(x)}{s(x)} + \alpha\right) - \phi\left(\frac{a-\hat{y}(x)}{s(x)} - \alpha\right)\right] \\ &- \int_{a-\alpha s(x)}^{a+\alpha s(x)} (\hat{y}(x) - y)^2 \phi\left(\frac{\hat{y}(x) - y}{s(x)}\right) dy \end{aligned}$$

- Selects trial from $(a - \epsilon, a + \epsilon)$
- Minimizes the overall variability

Selection of new trials (contd.)

- Temptation : maximize the improvement function
- Possibility : segment of the contour was not detected
- Expected Improvement

$$\begin{aligned} E[I(x)] &= [\alpha^2 s^2(x) - (\hat{y}(x) - a)^2]\left[\Phi\left(\frac{a - \hat{y}(x)}{s(x)} + \alpha\right) - \Phi\left(\frac{a - \hat{y}(x)}{s(x)} - \alpha\right)\right] \\ &+ 2(\hat{y}(x) - a)s^2(x)\left[\phi\left(\frac{a - \hat{y}(x)}{s(x)} + \alpha\right) - \phi\left(\frac{a - \hat{y}(x)}{s(x)} - \alpha\right)\right] \\ &- \int_{a-\alpha s(x)}^{a+\alpha s(x)} (\hat{y}(x) - y)^2 \phi\left(\frac{\hat{y}(x) - y}{s(x)}\right) dy \end{aligned}$$

- Selects trial from $(a - \epsilon, a + \epsilon)$
- Minimizes the overall variability
- Optimization issues

Contour extraction

- Need an implicit function to model the contour

Contour extraction

- Need an implicit function to model the contour
- Convenient feature of GASP model, recall that the predictor is $\hat{y}(x) = \hat{\mu} + r'\mathbf{R}^{-1}(Y - \hat{\mu}\mathbf{1_n})$ therefore,

$$a = \hat{\mu} + r'\mathbf{R}^{-1}(Y - \hat{\mu}\mathbf{1_n})$$

Contour extraction

- Need an implicit function to model the contour
- Convenient feature of GASP model, recall that the predictor is $\hat{y}(x) = \hat{\mu} + r'\mathbf{R}^{-1}(Y - \hat{\mu}\mathbf{1_n})$ therefore,

$$a = \hat{\mu} + r'\mathbf{R}^{-1}(Y - \hat{\mu}\mathbf{1_n})$$

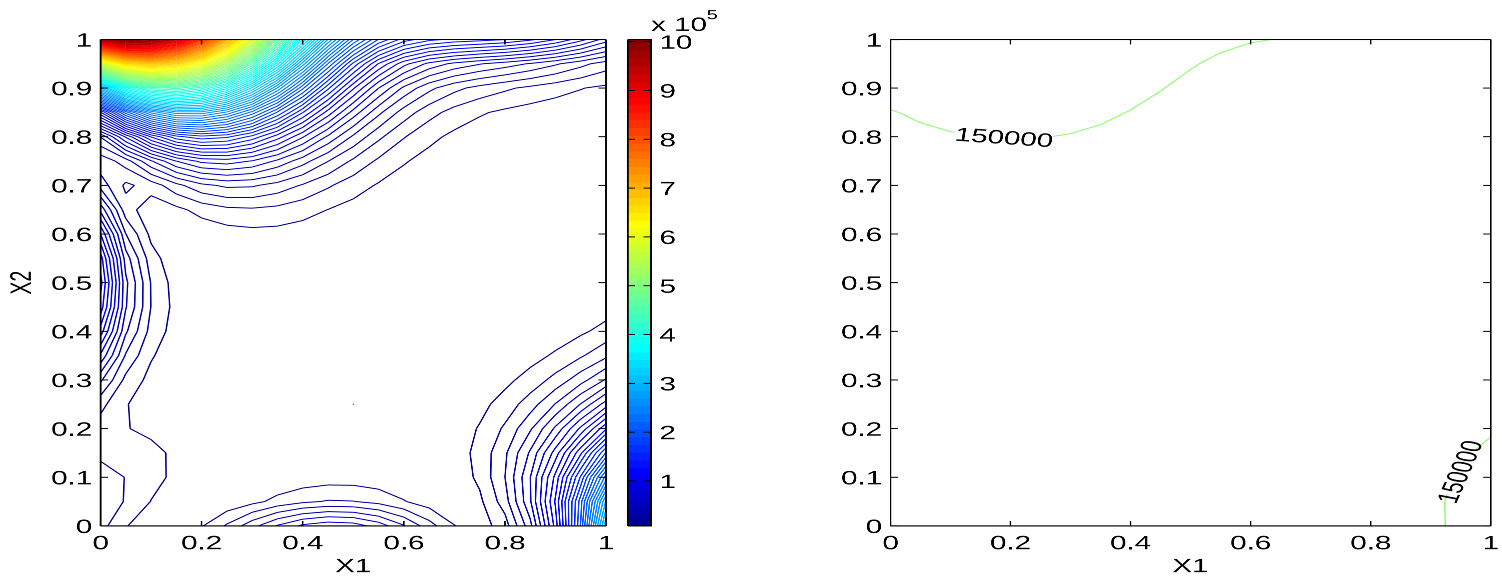
- No other values of response should be used for contour extraction

Example

Illustration / Example

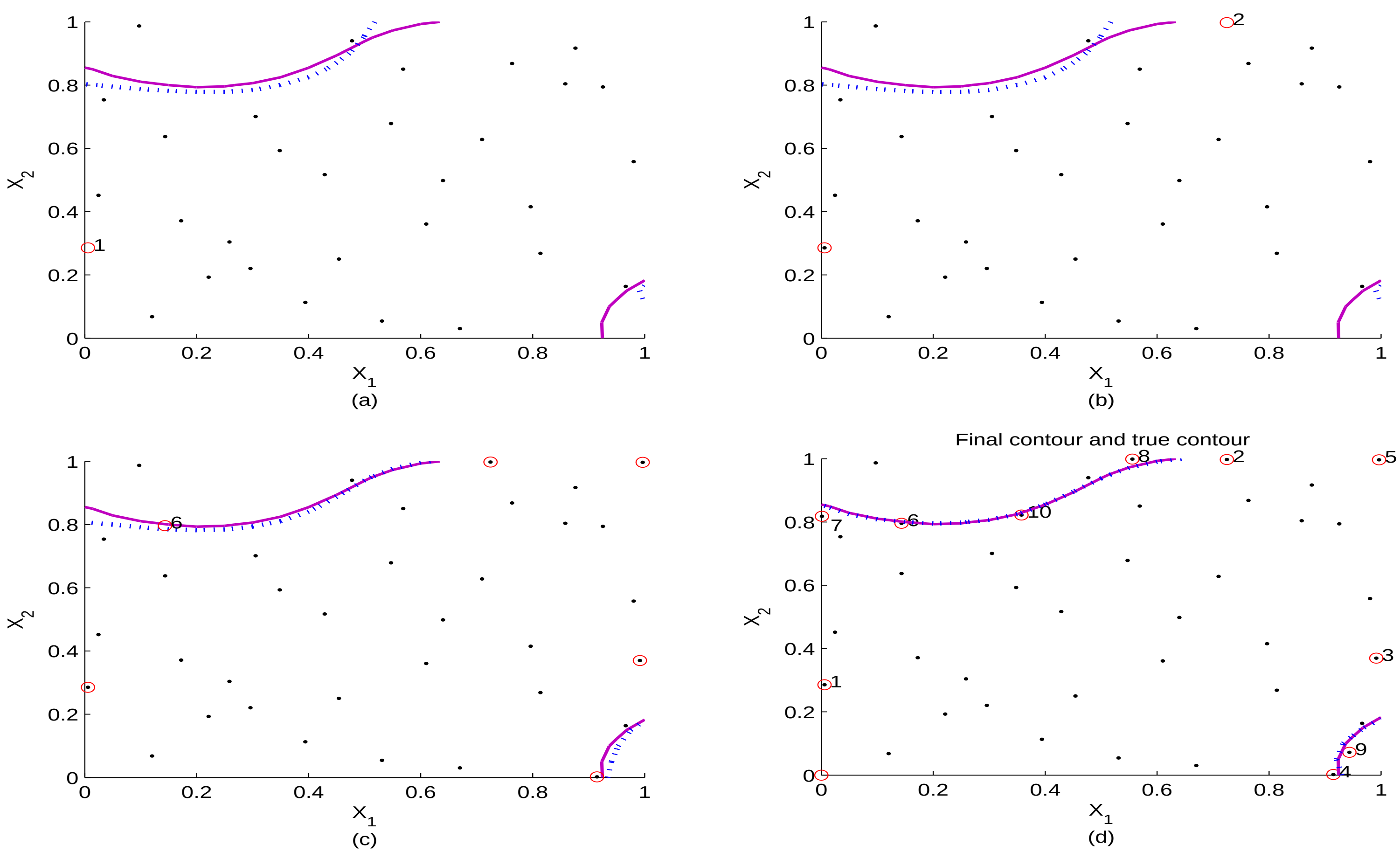
- Goldprice function

$$
\begin{aligned}
f(x,y) &= [1+(x+y+1)^2(19-14x+3x^2-14y+6xy+3y^2)] \\
&\quad *[30+(2x-3y)^2(18-32x+12x^2+48y-36xy+27y^2)]
\end{aligned}
$$



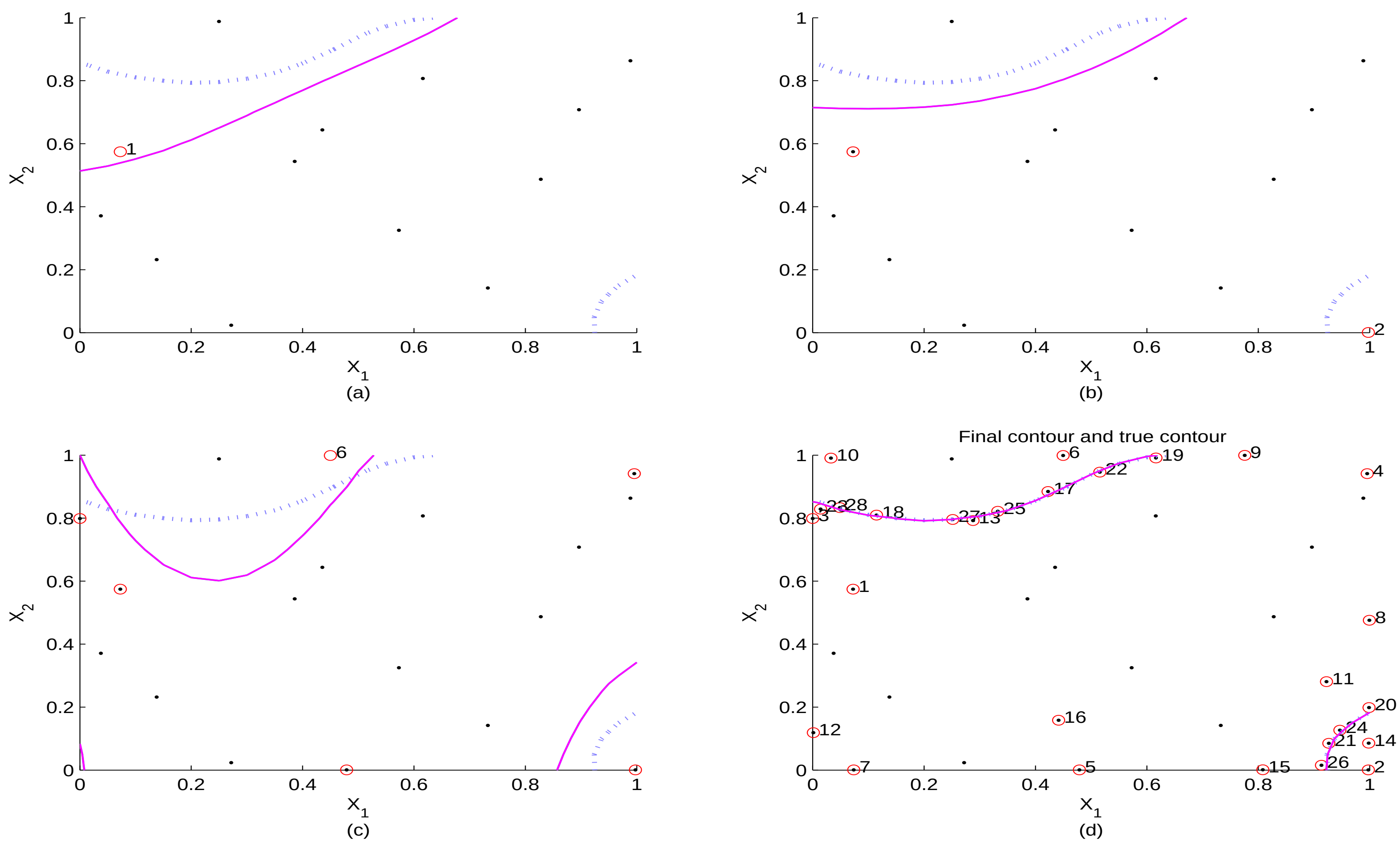
Example (contd.)

- $n_0 = 30$ and $n = 40$



Example (contd.)

- $n_0 = 12$ and $n = 40$



Aside: Goodness of fit measures

- Lack in Correlation between $C_{n_0,k}$ and truth

$$M_1 = \sum_{l=1}^{d} (1 - corr(x^{lk}, x^{lt}))$$

Aside: Goodness of fit measures

- Lack in Correlation between $C_{n_0,k}$ and truth

$$M_1 = \sum_{l=1}^{d} (1 - corr(x^{lk}, x^{lt}))$$

- Average L_2 distance between $C_{n_0,k}$ and truth

$$M_2 = \frac{1}{|C_{n_0,k}|} \sum_{x \in C_{n_0,k}} d(x, C_t)$$

where, $d(x, C_t) = min\{||x - y||_2 : y \in C_t\}$.

Aside: Goodness of fit measures

- Lack in Correlation between $C_{n_0,k}$ and truth

$$M_1 = \sum_{l=1}^{d} (1 - corr(x^{lk}, x^{lt}))$$

- Average L_2 distance between $C_{n_0,k}$ and truth

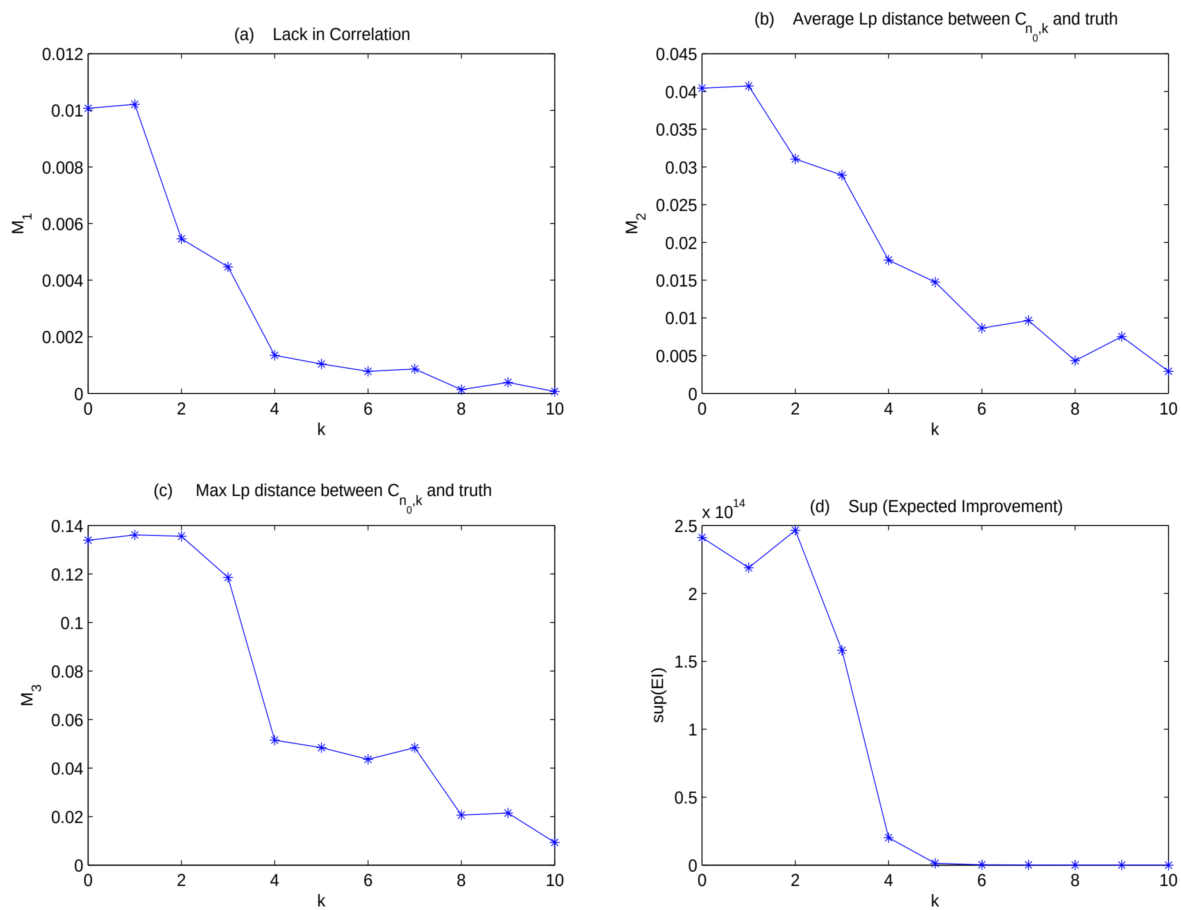
$$M_2 = \frac{1}{|C_{n_0,k}|} \sum_{x \in C_{n_0,k}} d(x, C_t)$$

where, $d(x, C_t) = min\{||x - y||_2 : y \in C_t\}$.

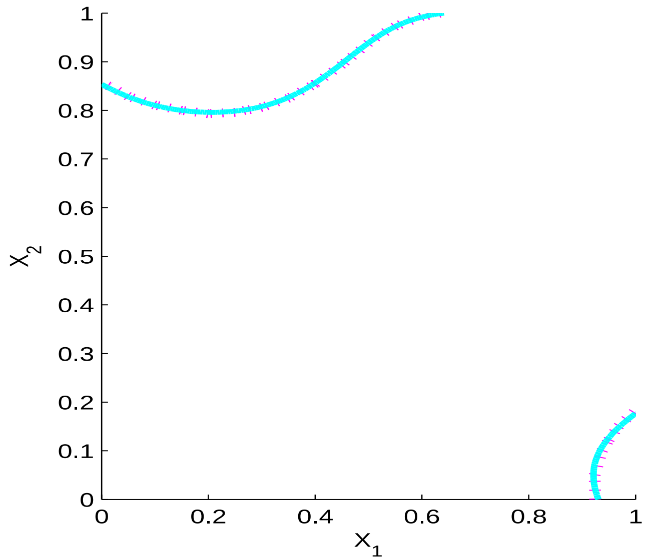
- Maximum L_2 distance between $C_{n_0,k}$ and truth

$$M_3 = max\{d(x, C_t) : x \in C_{n_0,k}\}$$

Example (contd.)



Example (contd.)



True contour and the implicit plot using GASP

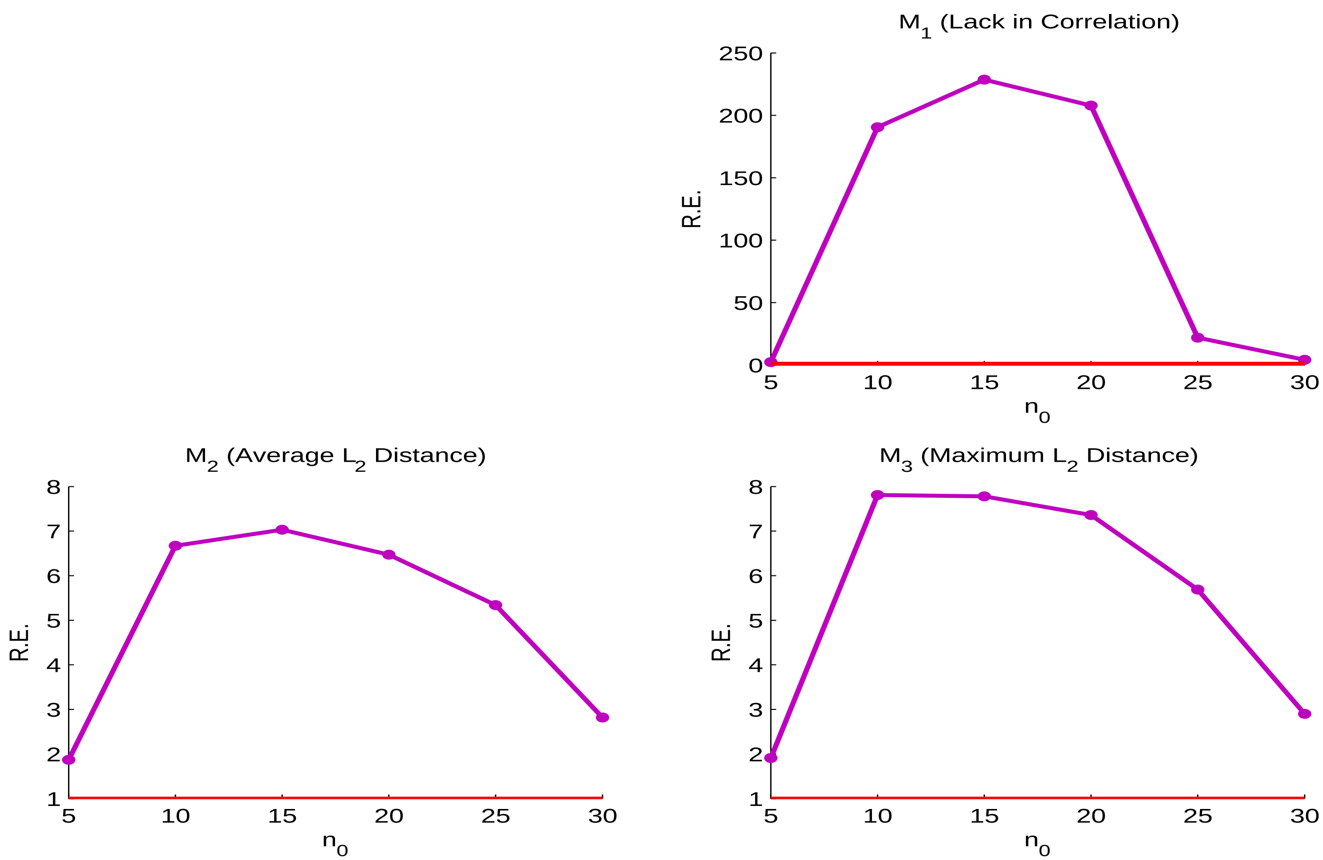
Example (contd.)

- Simulation study

	$n_0 = 30,\ k = 5\ (n = 35)$			$n_0 = 30,\ k = 10\ (n = 40)$		
measures	LHC	Sequential	R.E.	LHC	Sequential	R.E.
M_1	0.4573	0.1108	4.13	0.1968	3.14e-04	625.51
M_2	0.1048	0.0371	2.82	0.0514	0.0053	9.70
M_3	0.4468	0.1540	2.90	0.2303	0.0246	9.36

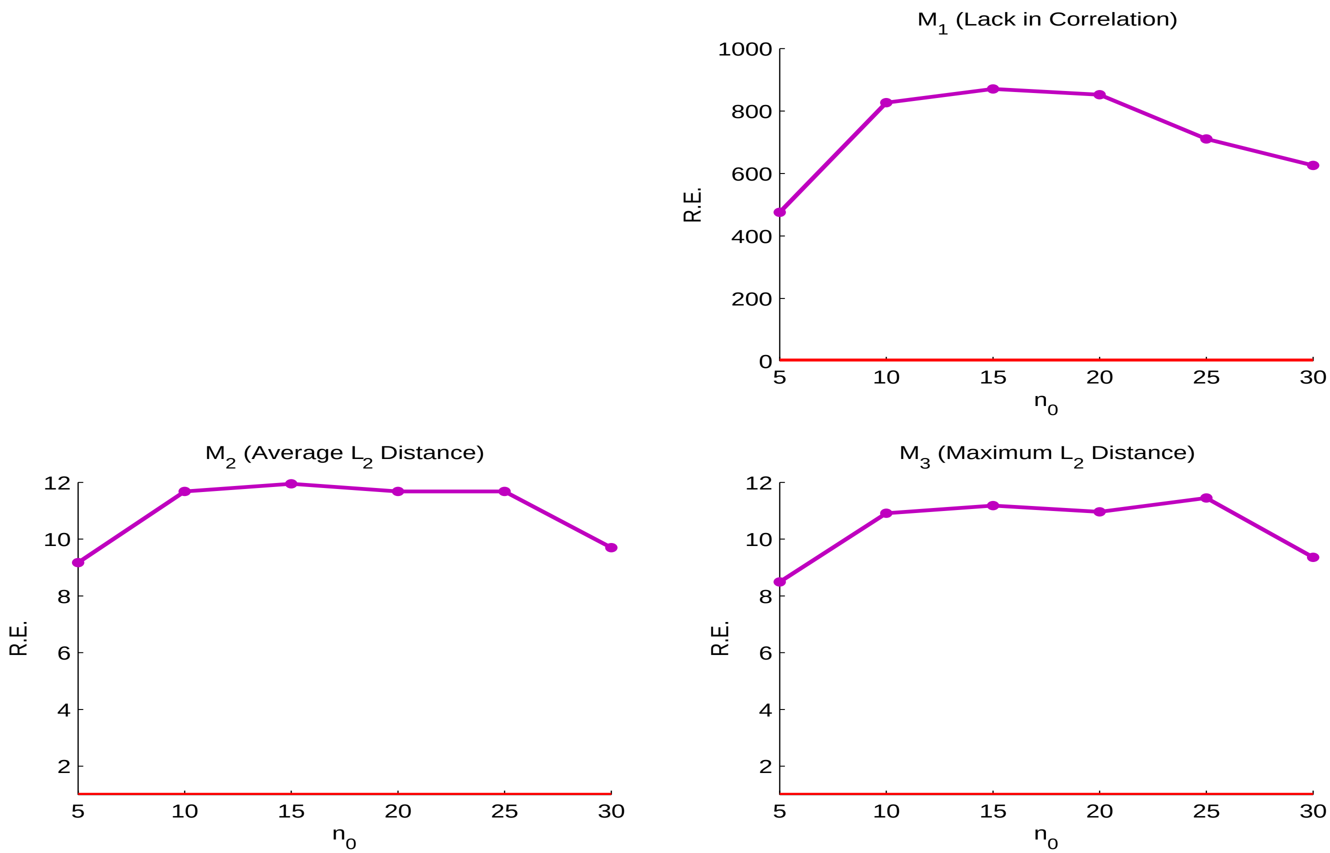
Example (contd.)

- Simulation study (contd):
  here, $n_0 + k = 35$.



Example (contd.)

- Simulation study (contd):
  here, $n_0 + k = 40$.



Summary

1. Started with $n_0 (< n)$ design points

Summary

1. Started with $n_0 (< n)$ design points
2. Fitted a GASP surface based on this design

Summary

1. Started with $n_0(< n)$ design points
2. Fitted a GASP surface based on this design
3. Selected new trials using *Expected Improvement function*

Summary

1. Started with $n_0(< n)$ design points
2. Fitted a GASP surface based on this design
3. Selected new trials using *Expected Improvement function*
4. Updated the design

Summary

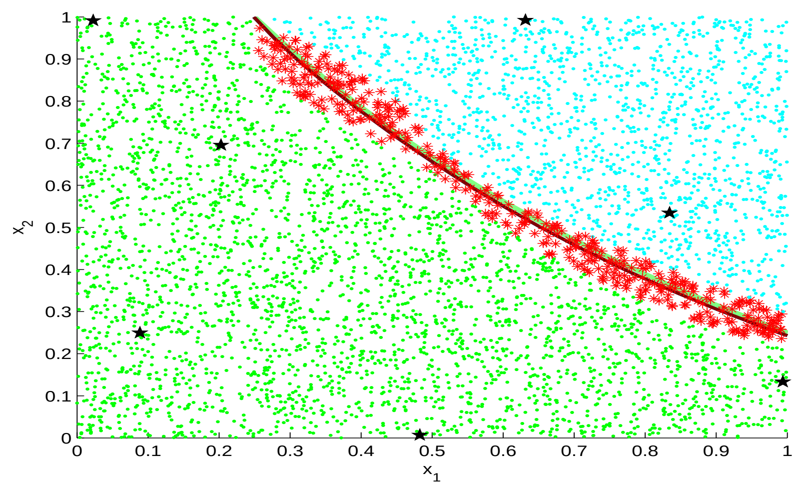
1. Started with $n_0 (< n)$ design points
2. Fitted a GASP surface based on this design
3. Selected new trials using *Expected Improvement function*
4. Updated the design
5. Repeat step-2 to step-4, until budget allows

Summary

1. Started with $n_0 (< n)$ design points
2. Fitted a GASP surface based on this design
3. Selected new trials using *Expected Improvement function*
4. Updated the design
5. Repeat step-2 to step-4, until budget allows
6. Extract the contour from the final surface

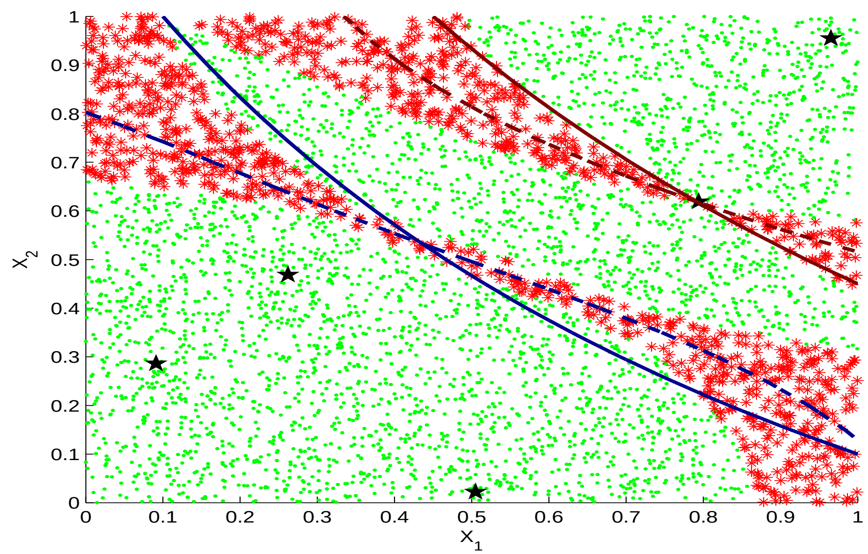
Concluding remarks

- Choice of n_0
- Goodness of fit measures (C_t versus C_{k+1})
- Choice of band radius (α)
  Recall that the ϵ-band is $(a - \alpha s(x), a + \alpha s(x))$
  and $y(x) \sim N(\hat{y}(x), s^2(x))$



Related work

- Multiple contours



Without loss of generality : $a_1 < a_2 < \ldots < a_k$

Related work

- Group-sequential approach for trials selection

Related work

- Group-sequential approach for trials selection
- Inverse problem for expensive simulators with functional response (nuclear waste dumping)

Thank you!