

# Méthodes d'optimisation sans gradient cas des fonctions partiellement séparables

**Laurent DUMAS**

avec D. Ding (IFPEN), B. Marteau (doctorant IFPEN)

Laboratoire de Mathématiques de Versailles (LMV)  
Université de Versailles Saint Quentin en Yvelines (UVSQ)

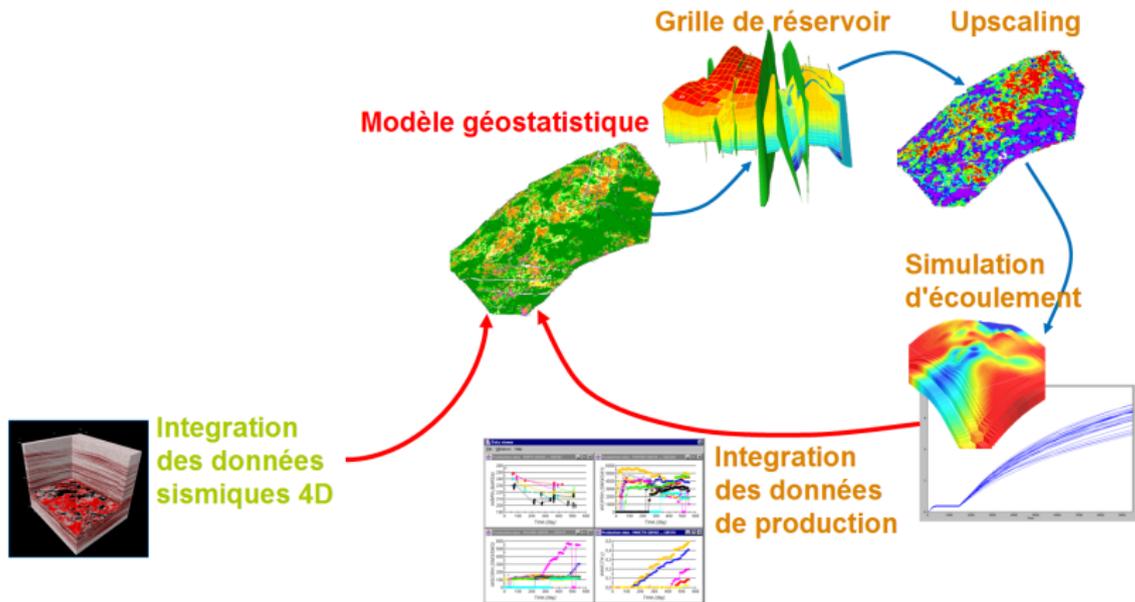
10 février 2014

# Plan de l'exposé

- 1 Problèmes inverses en ingénierie pétrolière
- 2 Méthodes de type région de confiance
- 3 Adaptation au cas des fonctions partiellement séparables
- 4 Résultats numériques
- 5 Un résultat de convergence

- 1 Problèmes inverses en ingénierie pétrolière
  - Le calage d'historique
  - Caractéristiques du problème inverse
  - Séparabilité partielle de la fonction objectif
  - Objectifs
- 2 Méthodes de type région de confiance
- 3 Adaptation au cas des fonctions partiellement séparables
- 4 Résultats numériques
- 5 Un résultat de convergence

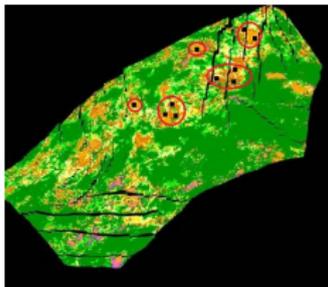
# Le calage d'historique



# Caractéristiques du problème inverse

- L'évaluation de la fonction objectif  $(x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n)$  à optimiser nécessite la simulation d'écoulements complexes dans un modèle de réservoir, processus très coûteux.
- Les caractéristiques principales du problème à résoudre sont :
  - Dépendance en de **nombreux paramètres**
  - **Evaluation coûteuse** de la fonction objectif
  - Fonction objectif **partiellement séparable**

# Séparabilité partielle de la fonction objectif



La fonction objectif peut s'écrire sous la forme :

$$\begin{aligned}
 f(x_1, \dots, x_n) &= \frac{1}{2} \sum_{i=1}^{n_1} \frac{\omega_i^P}{N_P(i)} \sum_{j=1}^{N_P(i)} \left( \frac{P_{i,j}^{obs}(x) - P_{i,j}^{sim}(x)}{\sigma_{i,j}^P} \right)^2 \\
 &= \sum_{i=1}^P f_i(x_1, \dots, x_n) \\
 &\approx \sum_{i=1}^P f_i(x_{1_i}, \dots, x_{n_i})
 \end{aligned}$$

avec  $\forall i, n_i \leq n$

# Objectifs

- Le problème général à résoudre peut se formuler de la manière suivante : *obtenir le meilleur calage d'historique à partir d'un nombre maximal fixé de simulations d'écoulement.*
- Plusieurs questions se posent alors :
  - **Paramétrisation** : déterminer les paramètres les plus pertinents.
  - **Initialisation du problème** : proposer la meilleure initialisation possible.
  - **Optimisation** : exploiter au mieux les connaissances sur la fonction.

# Objectifs

- De nombreux tests ont été réalisés dans la littérature et à l'IFPEN pour le domaine pétrolier.
- Il en ressort que les **méthodes sans gradient** sont en général à privilégier en l'absence d'estimation simple du gradient.
- Quatre grandes familles de méthodes d'optimisation sans gradient sont disponibles :
  - Méthodes directes de type simplexe (Nelder Mead)
  - Méthodes de type évolutionnaires (recuit simulé, algorithmes génétiques, CMA-ES)
  - Méthodes de type surfaces de réponse (RBF, krigeage)
  - Méthodes de type région de confiance

# Objectifs

- De nombreux tests ont été réalisés dans la littérature et à l'IFPEN pour le domaine pétrolier.
- Il en ressort que les **méthodes sans gradient** sont en général à privilégier en l'absence d'estimation simple du gradient.
- Quatre grandes familles de méthodes d'optimisation sans gradient sont disponibles :
  - Méthodes directes de type simplexe (Nelder Mead)
  - Méthodes de type évolutionnaires (recuit simulé, algorithmes génétiques, CMA-ES)
  - Méthodes de type surfaces de réponse (RBF, krigeage)
  - **Méthodes de type région de confiance**

- 1 Problèmes inverses en ingénierie pétrolière
- 2 Méthodes de type région de confiance
  - Description
  - Méthode NEWUOA
  - Auto-correction de la géométrie
- 3 Adaptation au cas des fonctions partiellement séparables
- 4 Résultats numériques
- 5 Un résultat de convergence

# Description d'une méthode de type région de confiance

## ALGORITHME SIMPLE

- 1 **Initialisation** Construire un modèle quadratique initial  $m_0$  de la fonction sur  $B(x_0, \Delta_0)$ .
- 2 **Itération  $k$** 
  - Calculer  $x_k^+$  le minimum de  $m_k$  sur  $B(x_k, \Delta_k)$ .
  - Remplacer le point d'interpolation le plus éloigné du point courant par  $x_k^+$
  - Si  $x_k^+$  est bon alors  $x_{k+1} := x_k^+$  et augmenter la région de confiance, sinon  $x_{k+1} := x_k$  et réduire la région de confiance
  - Construire le nouveau modèle  $m_k$  de la fonction.
- 3 **Condition d'arrêt.** S'arrêter lorsque le gradient du modèle devient trop petit.

## Critère d'acceptation du nouveau point

- Le critère d'acceptation/rejet du point  $x_k^+$  est basé sur le calcul du rapport :

$$\rho_k = \frac{m_k(x_k^+) - f(x_k)}{f(x_k^+) - f(x_k)}$$

- Le modèle est supposé bon lorsque  $\rho_k \geq \nu$  avec  $\rho \in ]0, 1[$ .

# Construction du modèle quadratique

- Dans le cas d'une méthode sans gradient, le modèle quadratique sur la région de confiance peut-être obtenu par **interpolation de Lagrange** à partir de  $p = \frac{(n+1)(n+2)}{2}$  points.
- Il est possible aussi d'utiliser **moins de points d'interpolation** et de déterminer le reste des paramètres du modèle quadratique en résolvant un problème de type moindres carrés.
- On choisit de travailler ici sur une méthode basée sur cette deuxième approche, proche de la méthode NEWUOA de Powell (2004)

# Calcul du modèle quadratique approché dans NEWUOA

- Dans la méthode NEWUOA, l'obtention du modèle quadratique  $m_k$  utilise un ensemble de points d'interpolation  $Y = \{y_1, \dots, y_m\}$  comprenant **au minimum  $n + 1$  points**.
- Le modèle quadratique approché  $m_k$  est obtenu en résolvant le problème de minimisation suivant :

$$\min(\|\alpha_Q\|_F^2)$$

sous la contrainte :  $m_k(y_j) = f(y_j)$  pour tout  $j \in \{1, \dots, m\}$  où

$$m_k(x) = f(x_k) + \langle \alpha_L, x \rangle + \frac{1}{2} \langle x, \alpha_Q x \rangle$$

# Description de la méthode NEWUOA

## Algorithme NEWUOA

- 1 **Initialisation** : construire un modèle quadratique initial  $m_0$  de la fonction (à partir de  $m$  points où  $m \geq n + 1$ )
- 2 **Itération  $k$**  :
  - Si le modèle  $m_k$  n'est pas assez "bon", remplacer un point d'interpolation par un point améliorant la géométrie du modèle
  - Sinon, calculer  $x_k^+$  le minimum de  $m_k$  sur  $B(x_k, \Delta_k)$ .
  - Remplacer le point d'interpolation le plus éloigné du point courant par  $x_k^+$
  - Si  $x_k^+$  est bon alors  $x_{k+1} := x_k^+$  et augmenter la région de confiance, sinon  $x_{k+1} := x_k$  et réduire la région de confiance
- 3 **Condition d'arrêt** : s'arreter lorsque le gradient du modèle devient trop petit

# Etape d'initialisation dans NEWUOA

- De bons résultats sont en général obtenus par la méthode NEWUOA avec  $m = 2n + 1$  points.
- Dans l'étape d'initialisation, les points d'interpolation sont choisis à partir du point courant  $x_0$  et d'un réel  $\rho$  :

$$\forall i \in \{1, 2, \dots, n\} \begin{cases} y_{i+1} = x_0 + \rho e_i \\ y_{i+n+1} = x_0 - \rho e_i \end{cases}$$

# Etape de vérification de la géométrie dans NEWUOA

- Cette méthode possède une **étape indispensable** de vérification et éventuellement d' **amélioration de la géométrie** des points utilisés dans la construction du modèle quadratique.
- Ces points sont choisis de telle sorte que l'ensemble soit unisolvant : on dit que l'ensemble  $Y = \{y_1, \dots, y_m\}$  est  **$\Lambda$  unisolvant sur  $B$**  ssi la famille des polynômes de Lagrange associée à  $Y$  est telle que

$$\Lambda \geq \max_{j=1, \dots, m} \max_{x \in B} |l_j(x)|$$

# Etape de vérification de la géométrie dans NEWUOA

L'amélioration de la géométrie s'effectue en utilisant le Lemme suivant :

## Lemma

*Soit  $Y$  un ensemble de points d'interpolation dans  $B$  et soit  $\Lambda > 1$ . On considère  $x \in B$  tel que  $|I_j(x)| \geq \Lambda$ . La procédure consistant à remplacer un des points de  $Y$  par  $x$  s'arrête après un nombre fini d'étape et permet d'obtenir un ensemble d'interpolation  $\Lambda$  unisolvant.*

# Auto-correction de la géométrie

- Un **principe d'autocorrection de la géométrie** a été introduit dans l'algorithme précédent (*Scheinberg, Toint, 2009*)
- Plus précisément, il repose sur le lemme suivant :

## Lemma

Soit  $\beta \in (0, 1)$ . Pour tout  $\Lambda > 1$ , il existe  $k_\Lambda$  tel que si l'itération  $k$  n'est pas réussie, c'est à dire  $\rho_k < \nu$  et que :

$$F_k := \{y_{k,j} \in Y_k \text{ tel que } \|y_{k,j} - x_k\| > \beta \Delta_k \text{ et } l_{k,j}(x_k^+) \neq 0\} = \emptyset$$

avec  $\Delta_k < k_\Lambda \|\nabla m_k(x_k)\|$ , alors l'ensemble

$$C_k := \{y_{k,j} \in Y_k \setminus \{x_k\} \text{ tel que } \|y_{k,j} - x_k\| \leq \beta \Delta_k \text{ et } l_{k,j}(x_k^+) \geq \Lambda\}$$

est non vide.

# Auto-correction de la géométrie

## Algorithme TR avec auto-correction de la géométrie

### 1 Initialisation

### 2 Test de criticité

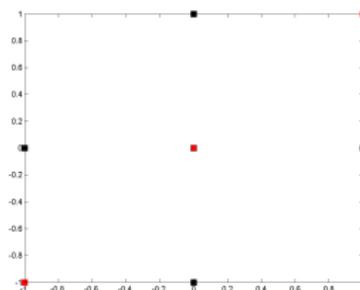
### 3 Itération $k$

- Calculer  $x_k^+$  le minimum de  $m_k$  sur  $B(x_k, \Delta_k)$ .
- Si  $x_k^+$  est bon, remplacer  $x_k$  par  $x_k^+$ , augmenter  $\Delta_k$  et remplacer le point d'interpolation le plus éloigné par  $x_k^+$ .
- Si  $x_k^+$  n'est pas bon, conserver  $x_k$  et  $\Delta_k$  et remplacer l'un des points d'interpolation dans  $F_k$  (loin de  $x_k^+$ ) ou dans  $C_k$  (près de  $x_k^+$ ) par  $x_k^+$ , si l'un des deux ensemble est non vide, sinon réduire  $\Delta_k$ .
- Construire le nouveau modèle  $m_k$  de la fonction.

- 1 Problèmes inverses en ingénierie pétrolière
- 2 Méthodes de type région de confiance
- 3 Adaptation au cas des fonctions partiellement séparables**
  - Adaptation de l'initialisation
  - Adaptation du modèle quadratique
  - Autocorrection de la géométrie
  - Méthode PSOF
- 4 Résultats numériques
- 5 Un résultat de convergence

## Adaptation de l'initialisation

- Dans le cas où  $f$  s'écrit :  $f(x_1, x_2) = f_1(x_1) + f_2(x_2)$ , l'initialisation se fait grâce à trois points et non cinq points :



- De manière générale, l'initialisation se fait avec un **nombre réduit de points** exploitant l'indépendance partielle des variables (nombre de couleurs d'un graphe).

# Adaptation du modèle quadratique

- L'idée est de créer un **modèle quadratique  $m_i$**  pour chaque **sous fonction objectif**. Le modèle quadratique global  $m(x)$  s'écrit ainsi :

$$m(x_1, \dots, x_n) = \sum_{i=1}^p m_i(x_{1_i}, \dots, x_{n_i})$$

- On espère ainsi :
  - Obtenir des **modèles plus précis**.
  - Construire des modèles avec **moins de points d'interpolation**.
- Un problème apparaît pour l'étape d'amélioration des modèles : sans modifications, il serait nécessaire d'améliorer tous les sous modèles, ce qui aurait un coût prohibitif.

# Notion de sous modèle dominant

- On définit la notion de **sous modèle dominant** : il s'agit du sous-modèle maximisant la valeur

$$\rho_k^j = \frac{m_k^j(x_k^+) - f_j(x_k)}{f(x_k^+) - f(x_k)}$$

## Autocorrection de la géométrie

Le principe d' **autocorrection de la géométrie** s'étend aux fonctions partiellement séparables, plus précisément au **sous-modèle dominant** :

### Lemma

*On note  $F_k^i$  et  $C_k^i$  les analogues de  $F_k$  et  $C_k$  du lemme précédent pour le  $i$ ème sous modèle. A l'iteration  $k$ , pour tout  $\Lambda > 1$ , si :*

$$\left\{ \begin{array}{l} j = \arg \max_i (m_k^i(x_k^+) - m_k^i(x_k)) \\ \Delta_k \leq k_\Lambda \|g_k\| \\ \rho_k^j < \nu \\ F_k^j = \emptyset \end{array} \right.$$

*alors  $C_k^j \neq \emptyset$ .*

# Description de la méthode PSOF

## Algorithme PSOF

### 1 Initialisation

### 2 Test de criticité

### 3 Itération $k$

- Calculer  $x_k^+$  le minimum de  $m_k$  sur une région de confiance puis calculer  $\rho_k$  et chaque  $\rho_k^i$ .
- Traitement de chaque sous modèle suivant 3 cas :
  - $\rho_k > \nu$ ,
  - $\rho_k < \nu$  et  $\rho_k^i > \nu$
  - $\rho_k < \nu$  et  $\rho_k^i < \nu$  avec les ensembles  $F_k^i$  et  $C_k^i$
- Amélioration d'un moins un sous modèle non dominant (cas  $\rho_k < \nu$  et  $\rho_k^i > \nu$  et aucune amélioration à l'étape précédente)
- Mise à jour de  $\Delta_k$ .
- Construire le nouveau modèle  $m_k$  et les modèles  $m_k^i$ .

- 1 Problèmes inverses en ingénierie pétrolière
- 2 Méthodes de type région de confiance
- 3 Adaptation au cas des fonctions partiellement séparables
- 4 Résultats numériques**
  - Fonctions tests
  - Comparaison avec NEWUOA
  - Cas tests en ingénierie de réservoir
- 5 Un résultat de convergence

# Fonctions tests

Des tests de la nouvelle méthode ont été réalisés sur des fonctions analytiques partiellement séparables : pour  $x = (x_1, \dots, x_n)$

$$DQDRTIC(x) = \sum_{i=1}^{n-2} (x_i^2 + x_{i+1}^2 + x_{i+2}^2)$$

$$LIARWHD(x) = \sum_{i=1}^n (4(x_i^2 - x_1)^2 + (x_i - 1)^2)$$

$$BDQRTIC(x) = \sum_{i=1}^{n-4} ((-4x_i + 3)^2 + (x_i^2 + x_{i+1}^2 + x_{i+2}^2 + x_{i+3}^2)) + 5x_n^2$$

$$ARWHEAD(x) = \sum_{i=1}^{n-1} ((x_i^2 + x_n^2)^2 - 4x_i + 3)$$

$$ROSENBROCK(x) = \sum_{i=1}^{n-1} (100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2)$$

## Résultats numériques, comparaison avec NEWUOA

- Le tableau suivant compare la minimisation des fonctions tests précédentes réalisées une méthode de type NEWUOA ou avec la nouvelle méthode PSOF.
- Les mêmes paramètres de départ sont pris pour les 2 méthodes, on compare dans ce tableau le nombre de simulations pour arriver à convergence

Fonction	10 param.		50 param.	
	NEWUOA	PSOF	NEWUOA	PSOF
DQDRTIC	204	21	+1300	20
LIARWHD	174	51	1215	66
BDQRTIC	231	149	+2000	169
ARWHEAD	368	43	+1300	45
ROSENBROCK	244	128	+1600	233

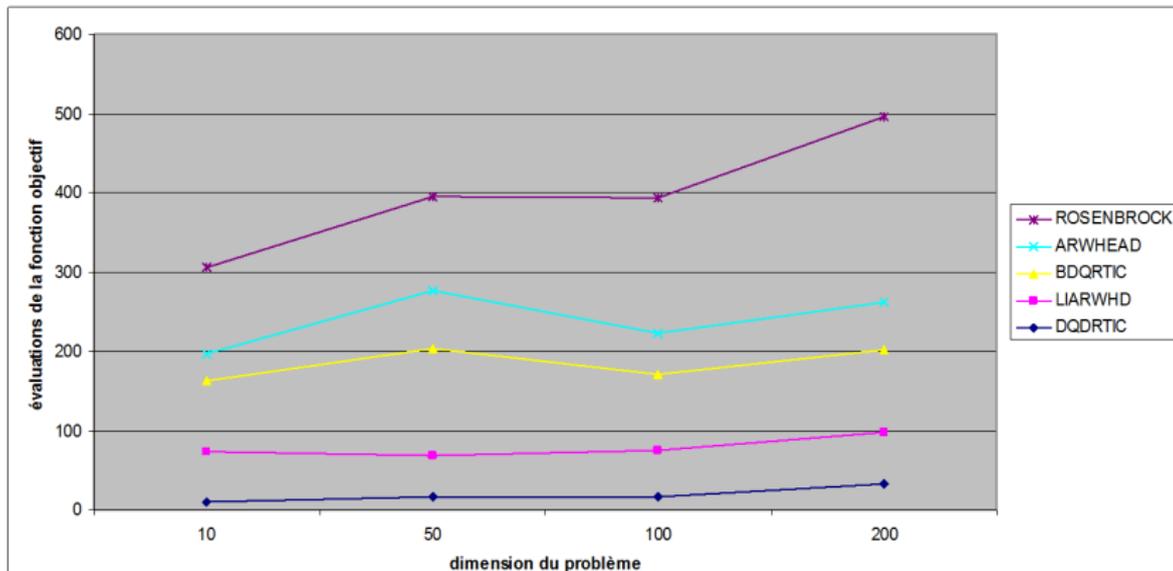
## Résultats numériques, comparaison avec NEWUOA

Ce tableau compare la valeur de la fonction objectif à l'arrêt des algorithmes :

Fonction	10 param.		50 param.	
	NEWUOA	PSOF	NEWUOA	PSOF
DQDRTIC	$1.9 * 10^{-4}$	$8.0 * 10^{-17}$	13000	$9.3 * 10^{-18}$
LIARWHD	0.01	$6.0 * 10^{-9}$	0.01	$3.6 * 10^{-7}$
BDQRTIC	37.9	18.5	312	178.9
ARWHEAD	2.77	$7.7 * 10^{-9}$	2.9	$1.8 * 10^{-8}$
ROSENBROCK	1.8	$6.0 * 10^{-5}$	129	0.04

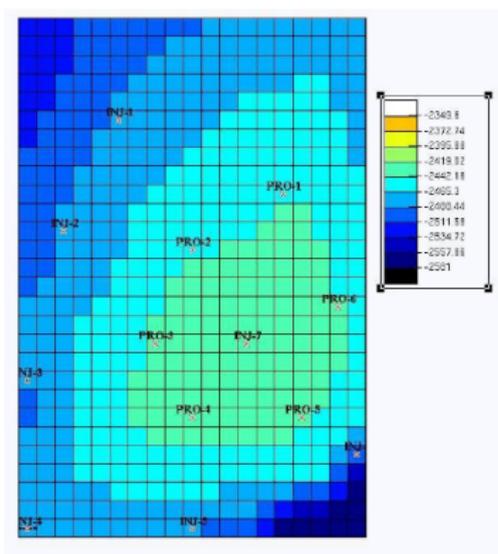
## Dépendance au nombre de variables d'optimisation

Les résultats de la méthode PSOF sont quasiment indépendants du nombre de variables :



## Cas test en ingénierie de réservoir

- Le cas test Punq, issu d'un modèle de terrain, est considéré comme un modèle représentatif de réservoir de petite taille.



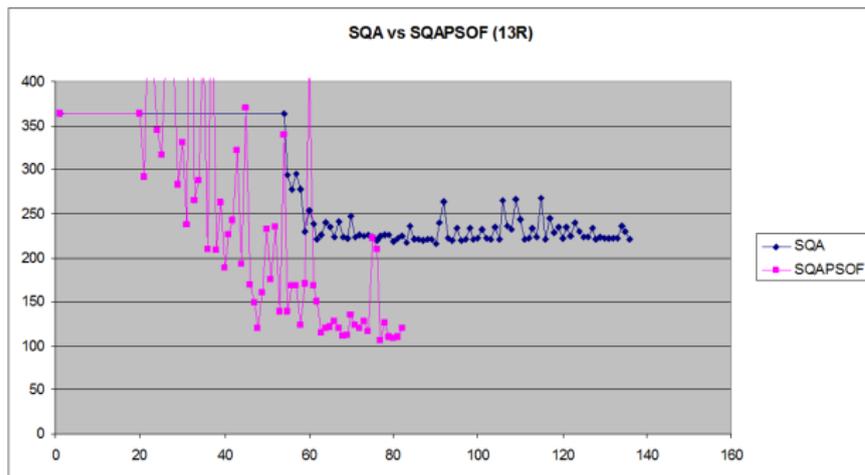
## Cas test en ingénierie de réservoir

- Ce cas test comprend 6 puits producteurs et 7 puits injecteurs.
- La séparabilité partielle de la fonction objectif est seulement ici approchée et vérifiée a posteriori :

$f_{PRO-1}$	←	$X_{PRO-1}$	$X_{INJ-1}$	$X_{PRO-2}$	$X_{INJ-6}$	$X_{INJ-7}$	
$f_{INJ-1}$	←	$X_{PRO-1}$	$X_{INJ-1}$	$X_{PRO-2}$	$X_{INJ-2}$		
$f_{PRO-2}$	←	$X_{PRO-1}$	$X_{INJ-1}$	$X_{PRO-2}$	$X_{INJ-2}$	$X_{PRO-2}$	$X_{INJ-7}$
$f_{INJ-2}$	←	$X_{INJ-1}$	$X_{INJ-2}$	$X_{PRO-2}$	$X_{INJ-3}$	$X_{PRO-3}$	
$f_{PRO-3}$	←	$X_{INJ-2}$	$X_{PRO-3}$	$X_{INJ-3}$	$X_{PRO-4}$	$X_{INJ-4}$	$X_{INJ-7}$
$f_{INJ-3}$	←	$X_{PRO-2}$	$X_{INJ-3}$	$X_{PRO-3}$			
$f_{PRO-4}$	←	$X_{INJ-3}$	$X_{PRO-4}$	$X_{INJ-4}$	$X_{INJ-5}$	$X_{INJ-7}$	
$f_{INJ-4}$	←	$X_{PRO-2}$	$X_{PRO-3}$	$X_{INJ-4}$			
$f_{PRO-5}$	←	$X_{PRO-4}$	$X_{PRO-5}$	$X_{INJ-5}$	$X_{INJ-6}$	$X_{INJ-7}$	
$f_{INJ-5}$	←	$X_{PRO-4}$	$X_{PRO-5}$	$X_{INJ-5}$	$X_{INJ-6}$		
$f_{PRO-5}$	←	$X_{PRO-1}$	$X_{PRO-5}$	$X_{PRO-6}$	$X_{INJ-6}$	$X_{INJ-7}$	
$f_{INJ-6}$	←	$X_{PRO-5}$	$X_{PRO-6}$	$X_{INJ-6}$			
$f_{PRO-5}$	←	$X_{PRO-2}$	$X_{PRO-3}$	$X_{PRO-4}$	$X_{PRO-5}$	$X_{PRO-6}$	$X_{INJ-7}$

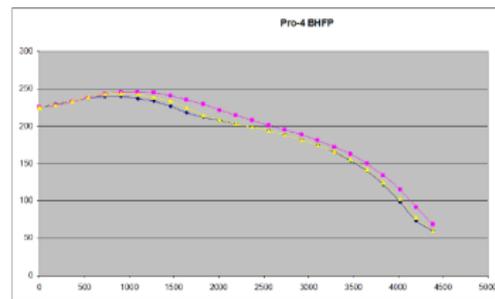
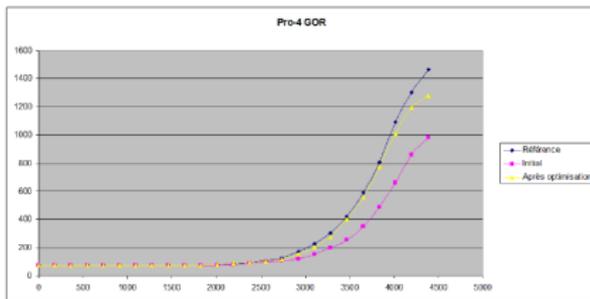
## Cas test en ingénierie de réservoir

- Les résultats d'optimisation montrent l'amélioration obtenue en utilisant PSOF par rapport à une méthode de type SQA classique :



## Cas test en ingénierie de réservoir

- L'amélioration du calage d'historique sur l'exemple du puits 4 est présenté ici :



- 1 Problèmes inverses en ingénierie pétrolière
- 2 Méthodes de type région de confiance
- 3 Adaptation au cas des fonctions partiellement séparables
- 4 Résultats numériques
- 5 Un résultat de convergence**
  - Théorème de convergence
  - Outils de la preuve

# Théorème de convergence

## Theorem

*On suppose que :*

*(i) la fonction  $f$  est différentiable et  $\nabla f$  continue et Lipschitzienne sur un ensemble  $\mathcal{V}$  contenant toutes les itérations de l'algorithme.*

*(ii)  $f$  est minorée sur  $\mathcal{V}$ .*

*(iii) pour tout  $k \in \mathbb{N}$ ,  $\|H_k\| \leq C$*

*Alors, la suite d'itérations  $(x_k)_{k \in \mathbb{N}}$  de la méthode PSOF précédente est telle que*

$$\lim_{k \rightarrow +\infty} \inf \nabla f(x_k) = 0$$

## Autocorrection de la géométrie

Le théorème de convergence repose sur le résultat d'autocorrection de la géométrie du sous-modèle dominant :

### Lemma

On note  $F_k^i$  et  $C_k^i$  les analogues de  $F_k$  et  $C_k$  du lemme précédent pour le  $i$ ème sous modèle. A l'iteration  $k$ , pour tout  $\Lambda > 1$ , si :

$$\left\{ \begin{array}{l} j = \arg \max_i (m_k^i(x_k^+) - m_k^i(x_k)) \\ \Delta_k \leq k_\Lambda \|g_k\| \\ \rho_k^j < \nu \\ F_k^j = \emptyset \end{array} \right.$$

alors  $C_k^j \neq \emptyset$ .

## Autocorrection de la géométrie

Pour sa démonstration, on utilise le résultat suivant sur l'erreur d'approximation avec un ensemble  $\Lambda$  unisolvant :

### Lemma

*Etant donné une boule fermée  $B(x, \Delta)$  et un ensemble d'interpolation  $Y = \{y_1, \dots, y_m\}$   $\Lambda$ -unisolvant, il existe des constantes  $K_1$  et  $K_2$  telle que*

$$\|f(y) - m(y)\| \leq K_1 \sum_{j=1}^m \|y_j - y\|^2 |l_j(y)|$$

et

$$\|\nabla f(y) - \nabla m(y)\| \leq K_2 \Lambda \Delta$$

## Principales étapes de la preuve

Outre le résultat essentiel d'autocorrection de la géométrie pour le sous-modèle dominant, les étapes de la preuve sont les suivantes :

- Le rayon de la région de confiance ne peut tendre vers 0 loin d'un point critique.
- Convergence vers un point critique dans le cas d'un nombre fini de succès.
- Convergence vers un point critique dans le cas d'un nombre fini de succès.

## Conclusion et perspectives

- Une **nouvelle méthode d'optimisation sans gradient** a été développée pour résoudre des problèmes inverses dans le domaine pétrolier (ingénierie de réservoir).
- Cette méthode consiste à **adapter un procédé de type région de confiance au cas de fonctions partiellement séparables**.
- Cette approche permet d'améliorer les méthodes existantes, en particulier dans le cas d'un grand nombre de variables à optimiser.
- Travail en cours : test de séparabilité partielle de la fonction, traitement des contraintes, etc...