Numerical Optimization for Large Scale Problems Report

Matteo Aldovardi (student ID: 360426)
Andrea Nocchiero (student ID: 361794)

January 11, 2026

Contents

1 Introduction

2 Implemented methods

Both the implemented methods use a `switch/case` to handle the three options for derivative evaluation supported:

1. exact gradient and exact Hessian (Mode = exact)
2. finite-difference (FD) approximations of both gradient and Hessian (Mode = full fd)
3. exact gradient and finite-difference approximation of the Hessian using the gradient (Mode = mixed fd)

2.1 Modified Newton method with Backtracking

The implemented algorithm is a Modified Newton method that applies a Hessian correction to enforce positive definiteness before computing the search direction, thereby ensuring that the resulting direction is a descent direction. Once the direction is obtained, an inexact line search is carried out via backtracking until the Armijo condition is satisfied.

We now outline the algorithmic steps performed at each iteration and highlight the role of the Hessian modification and the Armijo backtracking line search.

2.1.1 Hessian computation and symmetrization

The outer loop starts by computing the Hessian (or Hessian approximation) at x_k, denoted by $H_k \approx \nabla^2 f(x_k)$. Since FD-based evaluations may introduce numerical asymmetries, symmetry is enforced by updating $H_k \leftarrow \frac{1}{2}(H_k + H_k^\top)$.

2.1.2 Positive definiteness enforcement via Cholesky and diagonal shift

Subsequently, τ is initialized and updated until a Cholesky factorization succeeds. First, m is computed as the minimum diagonal entry of H_k, i.e. $m \leftarrow \min\big(\mathrm{diag}(H_k)\big)$. If $m > 0$ then $\tau \leftarrow 0$, otherwise $\tau \leftarrow \max(0, \beta - m)$. This is done because positive diagonal entries are a necessary condition for positive definiteness, but since it is not a sufficient condition, positive definiteness must be explicitly verified. A failure (reported by the flag output of `chol`) indicates that the current matrix is not positive definite and must be corrected through a diagonal shift. Hence, a Cholesky factorization is attempted on $H_k + \tau I$ and, if it fails, τ is increased according to $\tau \leftarrow \max(2\tau, \beta)$ and the factorization is retried. The process continues until Cholesky factorization succeeds, which means $H_k + \tau I$ has been perturbed enough to be positive definite.

2.1.3 Search direction and Armijo backtracking

Once Cholesky succeeds, the factor R is used to compute the descent direction p_k, and backtracking is performed until the Armijo condition is satisfied.

2.1.4 Stopping criterion and safeguards

The outer loop repeats until the stopping criterion is satisfied or the maximum number of iterations is reached (safeguard). The stopping criterion implemented is $||\nabla f(x_k)|| <$ `tol`, where `tol` is the given tolerance.

2.2 Truncated Newton method

The implemented algorithm is a Truncated Newton (Newton-CG) method. At each outer iteration k, the search direction p_k is computed by approximately solving the Newton system $H_k p_k = -g_k$ via the Conjugate Gradient (CG) method. This avoids the explicit computation and factorization of the Hessian, making the method suitable for large-scale problems.

2.2.1 Inner CG iteration and forcing sequence

The inner CG loop is terminated when the residual $r_i = H_k p_i + g_k$ satisfies the condition $\|r_i\| \leq \eta_k \|g_k\|$, where the forcing sequence is chosen as $\eta_k = \min(0.5, \sqrt{\|g_k\|})$. This ensures superlinear convergence near the solution.

2.2.2 Negative curvature handling

Since the CG method is designed for positive definite systems, negative curvature in H_k (detected if $d^T H_k d \leq 0$ during CG) requires termination of the inner loop. If negative curvature is encountered at the first inner iteration, the steepest descent direction $-g_k$ is used. Otherwise, the direction computed so far is returned.

2.2.3 Line search

After determining the search direction p_k, a backtracking line search is performed to satisfy the Armijo condition, ensuring global convergence.

3 Test problems

3.1 Problem 31 Broyden Tridiagonal

The objective function for the Broyden tridiagonal function, formulated as a nonlinear least-squares problem, is given by:

$$F(x) = \frac{1}{2} \sum_{k=1}^{n} f_k(x)^2$$

where the residual functions $f_k(x)$ are defined as:

$$f_k(x) = (3 - 2x_k)x_k - x_{k-1} - 2x_{k+1} + 1$$

This function is subject to the following boundary conditions:

$$x_0 = 0 \quad \text{and} \quad x_{n+1} = 0$$

The standard starting point for optimization algorithms is $x_i = -1$ for all $i = 1, \ldots, n$.

Gradient of the Objective Function

The gradient of the objective function is required for derivative-based optimization methods. The factor of $\frac{1}{2}$ in the objective function simplifies the expression.

$$\frac{\partial F}{\partial x_j} = \sum_{k=1}^{n} f_k(x) \frac{\partial f_k}{\partial x_j}$$

Using the Kronecker delta, δ_{ij}, the partial derivative of the residual function $f_k(x)$ with respect to x_j is:

$$\frac{\partial f_k}{\partial x_j} = (3 - 4x_k)\delta_{kj} - \delta_{k,j+1} - 2\delta_{k,j-1}$$

Substituting this into the gradient expression and simplifying the sum by using the sifting property of the Kronecker delta, we arrive at a single compact formula. By defining $f_0 = 0$ and $f_{n+1} = 0$ to handle the boundaries, the gradient components are:

$$\frac{\partial F}{\partial x_j} = f_j(x)(3 - 4x_j) - f_{j+1}(x) - 2f_{j-1}(x)$$

This single expression is valid for all $j = 1, \ldots, n$.

Hessian of the Objective Function

For second-order optimization methods, the Hessian of $F(x) = \frac{1}{2}\mathbf{f}(x)^T\mathbf{f}(x)$ is given by:

$$\nabla^2 F(x) = J(x)^T J(x) + \sum_{k=1}^{n} f_k(x)\nabla^2 f_k(x)$$

where $J(x)$ is the Jacobian matrix of the vector of residuals $\mathbf{f}(x) = [f_1(x), \ldots, f_n(x)]^T$.

The Jacobian matrix, $J(x)$, has a tridiagonal structure based on the partial derivatives of the residuals:

$$J(x) = \begin{pmatrix} 3 - 4x_1 & -2 & 0 & \dots & 0 \\ -1 & 3 - 4x_2 & -2 & \dots & 0 \\ 0 & -1 & 3 - 4x_3 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -1 & 3 - 4x_n \end{pmatrix}$$

The second term involves the Hessians of the individual residual functions, $\nabla^2 f_k(x)$. The only non-zero second partial derivative of $f_k(x)$ is:

$$\frac{\partial^2 f_k}{\partial x_k^2} = -4$$

This means the summation term simplifies to a diagonal matrix:

$$\sum_{k=1}^{n} f_k(x)\nabla^2 f_k(x) = \operatorname{diag}(-4f_1(x), -4f_2(x), \ldots, -4f_n(x))$$

Combining the terms, the full Hessian of the objective function is:

$$\nabla^2 F(x) = J(x)^T J(x) - 4 \cdot \operatorname{diag}(f_1(x), f_2(x), \ldots, f_n(x))$$

3.2 Problem 83 Lukšan and Vlček

The objective function for Problem 83, formulated as a nonlinear least-squares problem, is given by:

$$F(x) = \frac{1}{2}\sum_{k=1}^{n} f_k(x)^2$$

where the residual functions $f_k(x)$ are defined as:

$$f_k(x) = 2x_k + h^2(x_k + \sin(x_k)) - x_{k-1} - x_{k+1}$$

with the constant $h = 1/(n+1)$. The function is subject to the following boundary conditions:

$$x_0 = 0 \quad \text{and} \quad x_{n+1} = 1$$

The standard starting point for optimization algorithms is $x_i = 1$ for all $i = 1, \ldots, n$.

Gradient of the Objective Function

The gradient of the objective function is found using the chain rule, simplified by the $\frac{1}{2}$ factor in the objective function:

$$\frac{\partial F}{\partial x_j} = \sum_{k=1}^{n} f_k(x) \frac{\partial f_k}{\partial x_j}$$

The partial derivative of the residual function $f_k(x)$ with respect to x_j can be expressed compactly using the Kronecker delta, δ_{ij}:

$$\frac{\partial f_k}{\partial x_j} = (2 + h^2 + h^2 \cos(x_k))\delta_{kj} - \delta_{k,j+1} - \delta_{k,j-1}$$

By defining $f_0 = 0$ and $f_{n+1} = 0$ to handle the boundaries, substituting this into the gradient expression and simplifying the sum yields a single general formula for the gradient components:

$$\frac{\partial F}{\partial x_j} = f_j(x)(2 + h^2 + h^2 \cos(x_j)) - f_{j+1}(x) - f_{j-1}(x)$$

This single expression is valid for all $j = 1, \dots, n$.

Hessian of the Objective Function

For second-order optimization methods, the Hessian of $F(x) = \frac{1}{2}\mathbf{f}(x)^T \mathbf{f}(x)$ is given by:

$$\nabla^2 F(x) = J(x)^T J(x) + \sum_{k=1}^{n} f_k(x) \nabla^2 f_k(x)$$

where $J(x)$ is the Jacobian matrix of the vector of residuals $\mathbf{f}(x)$.

The Jacobian matrix, $J(x)$, is symmetric and tridiagonal:

$$J(x) = \begin{pmatrix} \alpha_1 & -1 & 0 & \dots \\ -1 & \alpha_2 & -1 & \dots \\ 0 & -1 & \ddots & \dots \\ \vdots & \ddots & \ddots & -1 \\ 0 & \dots & -1 & \alpha_n \end{pmatrix}$$

where $\alpha_k = 2 + h^2 + h^2 \cos(x_k)$.

The second term in the Hessian expression involves the Hessians of the individual residual functions, $\nabla^2 f_k(x)$. The only non-zero second partial derivative of $f_k(x)$ is:

$$\frac{\partial^2 f_k}{\partial x_k^2} = -h^2 \sin(x_k)$$

This simplifies the summation term to a diagonal matrix:

$$\sum_{k=1}^{n} f_k(x) \nabla^2 f_k(x) = \mathrm{diag}(-h^2 f_1(x) \sin(x_1), \dots, -h^2 f_n(x) \sin(x_n))$$

Combining the terms, the full Hessian of the objective function is:

$$\nabla^2 F(x) = J(x)^T J(x) - h^2 \cdot \mathrm{diag}(f_1(x) \sin(x_1), \dots, f_n(x) \sin(x_n))$$

4 Finite Differences implementation

First and second order finite difference approximations were implemented by exploiting the problem structure and sparsity pattern, in order to reduce computational cost and improve performance. When perturbing a single component, only the locally affected terms are updated.

4.1 FD Problem 31

4.1.1 FD Gradient Problem 31

For this specific problem, each residual f_k depends only on (x_{k-1}, x_k, x_{k+1}). This is directly exploited in the computation of the centered finite differences (preferred over forwards/backward because error is $O(h^2)$ instead of $O(h)$):

$$[\nabla F(x)]_i \;\approx\; \frac{F(x+he_i) - F(x-he_i)}{2h} \tag{1}$$

Since perturbing a single component x_i can only affect the residuals f_{i-1}, f_i, f_{i+1}, all other terms in the sum defining F remain unchanged and cancel out in the difference $F(x+he_i)-F(x-he_i)$. Therefore, defining:

$$S(i) := \{i-1, i, i+1\} \cap \{1,\dots,n\}, \qquad E_{\mathrm{loc}}(x;i) := \frac{1}{2}\sum_{k\in S(i)} f_k(x)^2,$$

we can compute the same centered FD quotient via the local contribution only:

$$[\nabla F(x)]_i \;\approx\; \frac{E_{\mathrm{loc}}(x+he_i;i) - E_{\mathrm{loc}}(x-he_i;i)}{2h}.$$

This drastically reduces the cost to compute FD approximation of the gradient. For example, naive FD implementation for $n = 10^3$ requires $2 \cdot 10^6$ residual evaluations, while our efficient implementation requires only 5996 residual evaluations, without sacrificing accuracy.

4.1.2 FD Hessian Problem 31

The same locality property can be exploited to approximate the Hessian entries by centered finite differences. For the diagonal terms, we use the standard second-order centered formula:

$$H_{ii}(x) = \frac{\partial^2 F(x)}{\partial x_i^2} \;\approx\; \frac{F(x+h_ie_i) - 2F(x) + F(x-h_ie_i)}{h_i^2}. \tag{2}$$

Since perturbing x_i affects only the residuals f_{i-1}, f_i, f_{i+1}, we can replace full evaluations of F with the local energy $E_{\mathrm{loc}}(\cdot;i)$ defined on $S(i) = \{i-1,i,i+1\} \cap \{1,\dots,n\}$, obtaining:

$$H_{ii}(x) \;\approx\; \frac{E_{\mathrm{loc}}(x+h_ie_i;i) - 2E_{\mathrm{loc}}(x;i) + E_{\mathrm{loc}}(x-h_ie_i;i)}{h_i^2}. \tag{3}$$

For off-diagonal entries ($i \neq j$), we approximate mixed second derivatives with the centered cross difference formula:

$$H_{ij}(x) \approx \frac{F(x+h_ie_i+h_je_j) - F(x+h_ie_i-h_je_j) - F(x-h_ie_i+h_je_j) + F(x-h_ie_i-h_je_j)}{4h_ih_j}. \tag{4}$$

Due to the problem structure, H_{ij} is nonzero only for $|i-j| \le 2$, hence the Hessian is pentadiagonal. When perturbing (x_i, x_{i+1}) only the residuals with indices $S(i,i+1) = \{i-1, i, i+1, i+2\} \cap \{1,\dots,n\}$ can change, and when perturbing (x_i, x_{i+2}) only $S(i,i+2) =$

$\{i-1, i, i+1, i+2, i+3\} \cap \{1, \dots, n\}$ can change. Therefore, each $F(\cdot)$ appearing in the mixed difference formula can be replaced by the corresponding local energy evaluation, leading to an efficient construction of the five relevant diagonals $(0, \pm 1, \pm 2)$ and assembly of the sparse pentadiagonal Hessian in i and j of the second order finite difference $H_{i,j}$.

In terms of residual evaluations, for $n = 10^3$ a naive Hessian FD implementation requires $\approx 10^7$ residual evaluations, whereas our implementation requires only $\approx 4.5 \cdot 10^4$ residual evaluations, without sacrificing accuracy.

4.2 FD Problem 83

4.2.1 FD Gradient Problem 83

As in (1), we approximate each component of the gradient via centered finite differences. Similarly as in problem 31 each residual f_k depends only on (x_{k-1}, x_k, x_{k+1}), so perturbing x_i affects only f_{i-1},f_i and f_{i+1}. Therefore, in the difference $F(x + h_i e_i) - F(x - h_i e_i)$ all the terms $f_k(x)^2$ with $k \notin \{i-1, i, i+1\}$ remain unchanged and cancel out. Defining

$$S(i) := \{i-1, i, i+1\} \cap \{1, \dots, n\}, \qquad E_{\text{loc}}(x; i) := \frac{1}{2} \sum_{k \in S(i)} f_k(x)^2,$$

we can compute the same centered FD quotient via the local contribution only:

$$[\nabla F(x)]_i \;\approx\; \frac{E_{\text{loc}}(x + h_i e_i; i) - E_{\text{loc}}(x - h_i e_i; i)}{2h_i}.$$

Here the residuals f_k are evaluated consistently with the boundary conditions $x_0 = 0$ and $x_{n+1} = 1$ (in particular, the last residual includes the constant term -1).

In terms of residual evaluations the same computational advantages already discussed at the end of 4.1.1 are obtained.

4.2.2 FD Hessian Problem 83

The same locality property can be exploited to approximate the Hessian entries by centered finite differences. For the diagonal terms, we use the standard second-order centered formula (2). Since perturbing x_i affects only the residuals f_{i-1}, f_i, f_{i+1}, we can replace full evaluations of F with the local energy $E_{\text{loc}}(\cdot; i)$ defined on $S(i) = \{i-1, i, i+1\} \cap \{1, \dots, n\}$, obtaining (3).

For first off-diagonal entries ($j = i+1$), we approximate mixed second derivatives with the centered cross difference formula (4). Here, perturbing (x_i, x_{i+1}) affects only the residuals with indices $S(i, i+1) = \{i-1, i, i+1, i+2\} \cap \{1, \dots, n\}$, hence each $F(\cdot)$ term in the formula above is replaced by the corresponding local energy evaluation on $S(i, i+1)$.

Due to the problem structure, H_{ij} is nonzero only for $|i - j| \leq 2$, hence the Hessian is pentadiagonal. It's important to note that for Problem 83 the second off-diagonal entries ($j = i+2$) are known analytically and are equal to 1, so they are inserted directly without finite differences. The resulting sparse pentadiagonal Hessian is then assembled by constructing the five relevant diagonals $(0, \pm 1, \pm 2)$ and exploiting symmetry of the formula defining $H_{i,j}$. Boundary conditions $x_0 = 0$ and $x_{n+1} = 1$ are handled consistently in the residual evaluations.

5 Experimental setup

We use a common protocol for both test problems and for both solvers (Modified Newton and Truncated Newton).

Dimensions.

$$n \in \{2,\ 10^3,\ 10^4,\ 10^5\}.$$

Seed. All randomized runs are reproducible with

$$\texttt{seed} = \min\{360426, 361794\}.$$

Initialization. For each dimension n, we use as baseline initial guess the point suggested in the reference PDF, denoted by $\bar{x} \in \mathbb{R}^n$. In addition, we generate 5 perturbed starting points by sampling uniformly in the hypercube centered at $\bar{x}$, i.e.

$$x_0 \sim \mathcal{U}\Big([\bar{x}_1 - 1, \bar{x}_1 + 1] \times \cdots \times [\bar{x}_n - 1, \bar{x}_n + 1]\Big),$$

so that each component $(x_0)_i$ is drawn independently from $\mathcal{U}\big([\bar{x}_i - 1, \bar{x}_i + 1]\big)$.

Derivative regimes and finite differences. We run three derivative settings:

- **Exact:** exact gradient and exact Hessian;
- **Mixed:** exact gradient and FD Hessian
- **Full FD:** FD gradient and FD Hessian.

In the FD-based experiments we test the increments $h = 10^{-k}$ with

$$k \in \{4,\ 8,\ 12\},$$

and we consider both a *constant* step h and a *component-wise* step h_i proportional to the current evaluation point.

Success/failure criteria. A run is successful if the stopping test is met within $k_{\max}$; otherwise it is a failure. For all experiments the stopping criteria is $\|\nabla f(x^k)\|_2 < tol$

Logged metrics. For each run we store the final outputs

$$\Big(x_k,\ \|\nabla f(x_k)\|_2,\ \texttt{iters} = k,\ \texttt{btflag},\ p_{\text{exp}},\ t_{\text{run}}\Big),$$

together with the diagnostic sequences

$$\Big(\{x_j\}_{j=0}^{k},\ \{\texttt{bt}_j\}_{j=0}^{k-1},\ \{\texttt{mod}_j\}_{j=0}^{k-1},\ \{\tau_j\}_{j=0}^{k-1}\Big),$$

where $\texttt{bt}_j$ is the number of backtracking reductions, $\texttt{mod}_j$ the number of Hessian modifications, and τ_j the final damping parameter at iteration j.

6 Parameter Tuning

Parameter tuning – Problem 31 (Modified Newton)						
Derivatives	$\mathbf{k}_{\max}$	**tol**	$\mathbf{c_1}$	$\boldsymbol{\rho}$	$\mathbf{bt}_{\max}$	$\boldsymbol{\beta}$
Exact	`1000`	`1e-8`	`1e-4`	`0.5`	`100`	`1e-3`
Mixed	...	...	...	...	...	...
Full FD	...	...	...	...	...	...

7 Results - Exact derivatives

7.1 Problem 31

7.1.1 Modified Newton

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
31	2	exact	$\overline{x}$	2.61×10^{-15}	6/1000	yes		1.98	0.0056s
			1	1.92×10^{-15}	7/1000	yes		2.01	0.0012s
			2	2.16×10^{-9}	6/1000	yes		1.99	0.0012s
			3	8.94×10^{-10}	6/1000	yes		1.99	0.0009s
			4	2.67×10^{-10}	5/1000	yes		1.99	0.0008s
			5	1.04×10^{-12}	8/1000	yes		2.00	0.0013s
			Avg (successes)	5.55×10^{-10}	6.3/1000	6/6		1.99	0.0018s

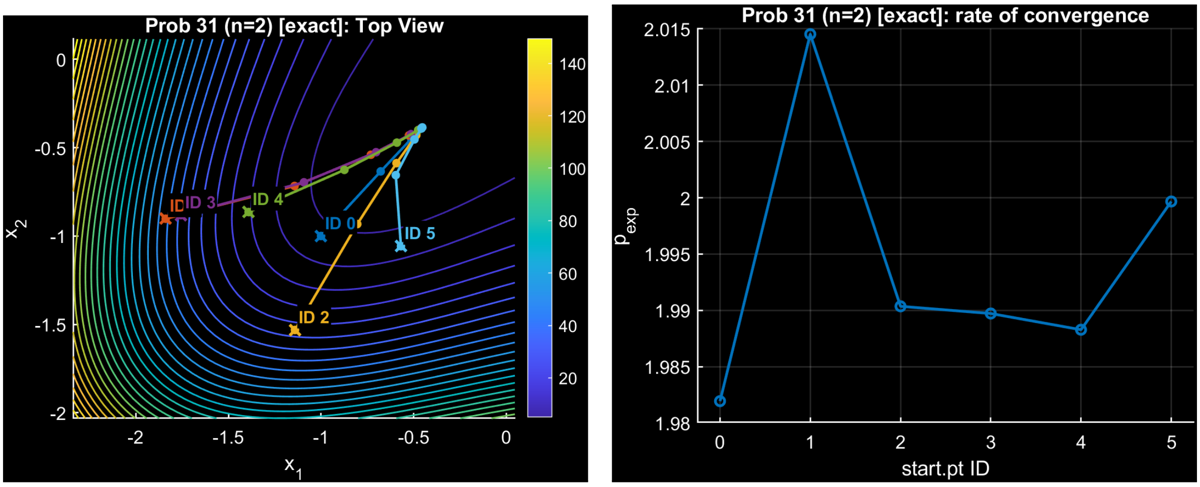


Figure 1: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
31	1000	exact	$\overline{x}$	1.94×10^{-14}	6/1000	yes		1.37	0.0052s
			1	4.69×10^{-13}	10/1000	yes		1.98	0.0075s
			2	2.91×10^{-14}	10/1000	yes		1.83	0.0068s
			3	1.07×10^{-13}	9/1000	yes		1.98	0.0071s
			4	6.62×10^{-14}	9/1000	yes		1.96	0.0073s
			5	1.43×10^{-12}	9/1000	yes		1.99	0.0073s
			Avg (successes)	3.53×10^{-13}	8.8/1000	6/6		1.85	0.0069s

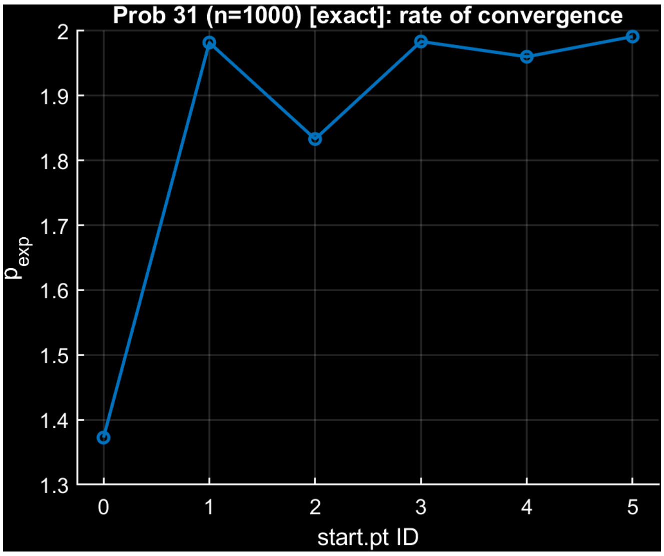


Figure 2: Problem 31, Modified Newton, exact derivatives ($n = 1000$): experimental convergence rate p_{exp} for converged runs.

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
31	10000	exact	$\overline{x}$	1.94×10^{-14}	6/1000	yes		1.35	0.0285s
			1	9.46×10^{-12}	10/1000	yes		2.00	0.0583s
			2	3.12×10^{-9}	9/1000	yes		1.95	0.0552s
			3	2.38×10^{-14}	11/1000	yes		1.79	0.0727s
			4	7.10×10^{-14}	9/1000	yes		1.81	0.0637s
			5	3.85×10^{-14}	11/1000	yes		1.97	0.0604s
			Avg (successes)	5.22×10^{-10}	9.3/1000	6/6		1.81	0.0565s

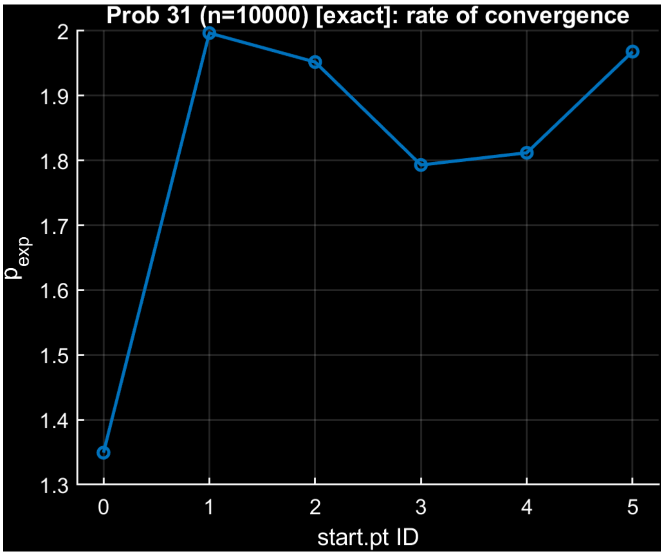


Figure 3: Problem 31, Modified Newton, exact derivatives ($n = 10000$): experimental convergence rate p_{exp} for converged runs.

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
31	100000	exact	$\overline{x}$	1.94×10^{-14}	6/1000	yes		1.26	0.4384s
			1	1.70×10^{-13}	9/1000	yes		1.53	0.9005s
			2	2.44×10^{-14}	13/1000	yes		1.46	1.32s
			3	3.45×10^{-11}	10/1000	yes		2.00	1.18s
			4	1.09×10^{-13}	11/1000	yes		1.94	1.47s
			5	1.23×10^{-9}	10/1000	yes		2.00	1.60s
			Avg (successes)	2.10×10^{-10}	9.8/1000	6/6		1.70	1.15s

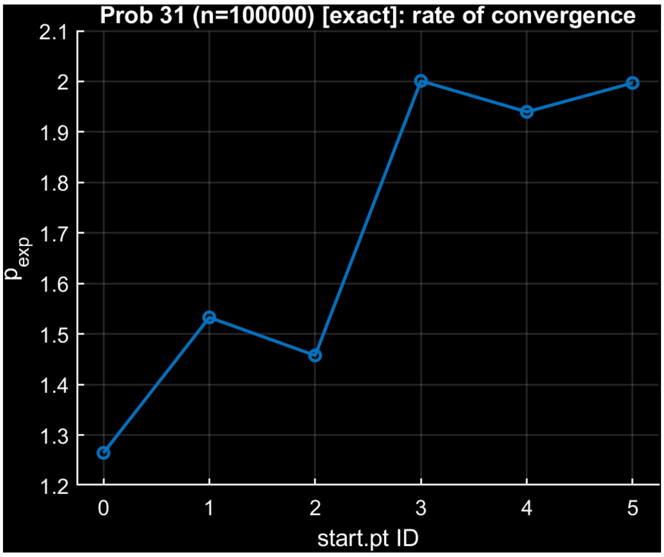


Figure 4: Problem 31, Modified Newton, exact derivatives ($n = 100000$): experimental convergence rate p_{exp} for converged runs.

7.1.2 Truncated Newton

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
31	2	exact	$\overline{x}$	2.61×10^{-15}	6/1000	yes		1.98	0.0056s
			1	1.92×10^{-15}	7/1000	yes		2.01	0.0012s
			2	2.16×10^{-9}	6/1000	yes		1.99	0.0012s
			3	8.94×10^{-10}	6/1000	yes		1.99	0.0009s
			4	2.67×10^{-10}	5/1000	yes		1.99	0.0008s
			5	1.04×10^{-12}	8/1000	yes		2.00	0.0013s
			Avg (successes)	5.55×10^{-10}	6.3/1000	6/6		1.99	0.0018s

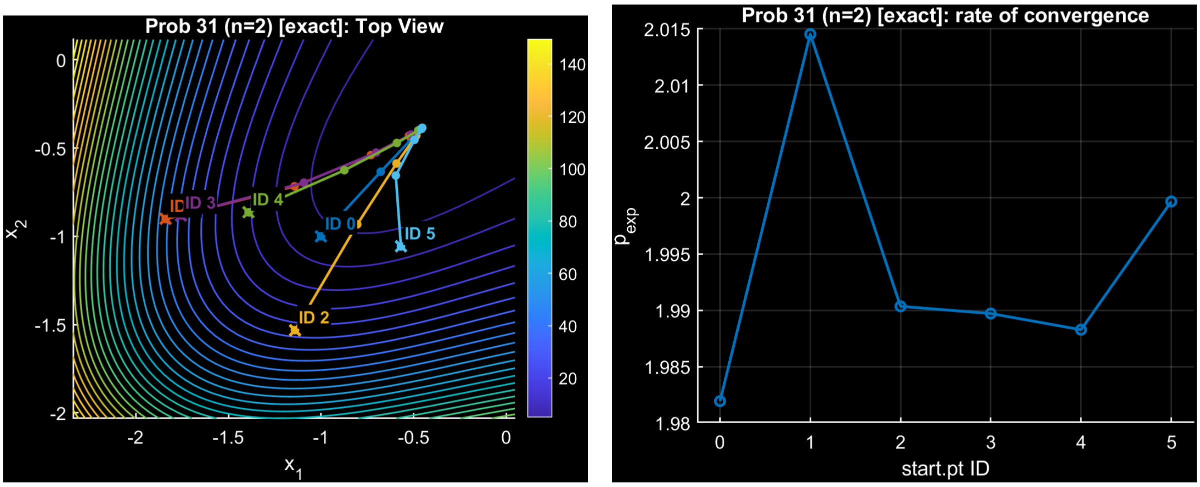


Figure 5: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
31	1000	exact	$\overline{x}$	1.94×10^{-14}	6/1000	yes		1.37	0.0052s
			1	4.69×10^{-13}	10/1000	yes		1.98	0.0075s
			2	2.91×10^{-14}	10/1000	yes		1.83	0.0068s
			3	1.07×10^{-13}	9/1000	yes		1.98	0.0071s
			4	6.62×10^{-14}	9/1000	yes		1.96	0.0073s
			5	1.43×10^{-12}	9/1000	yes		1.99	0.0073s
			Avg (successes)	3.53×10^{-13}	8.8/1000	6/6		1.85	0.0069s

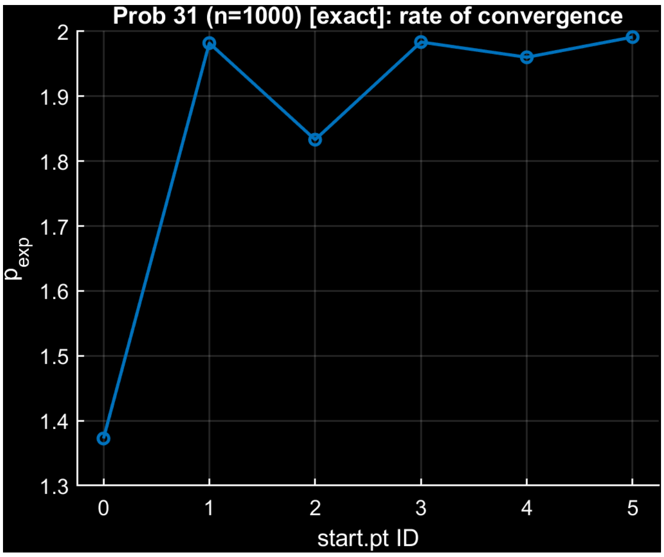


Figure 6: Problem 31, Truncated Newton, exact derivatives ($n = 1000$): experimental convergence rate p_{exp} for converged runs.

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
31	10000	exact	$\overline{x}$	1.94×10^{-14}	6/1000	yes		1.35	0.0285s
			1	9.46×10^{-12}	10/1000	yes		2.00	0.0583s
			2	3.12×10^{-9}	9/1000	yes		1.95	0.0552s
			3	2.38×10^{-14}	11/1000	yes		1.79	0.0727s
			4	7.10×10^{-14}	9/1000	yes		1.81	0.0637s
			5	3.85×10^{-14}	11/1000	yes		1.97	0.0604s
			Avg (successes)	5.22×10^{-10}	9.3/1000	6/6		1.81	0.0565s

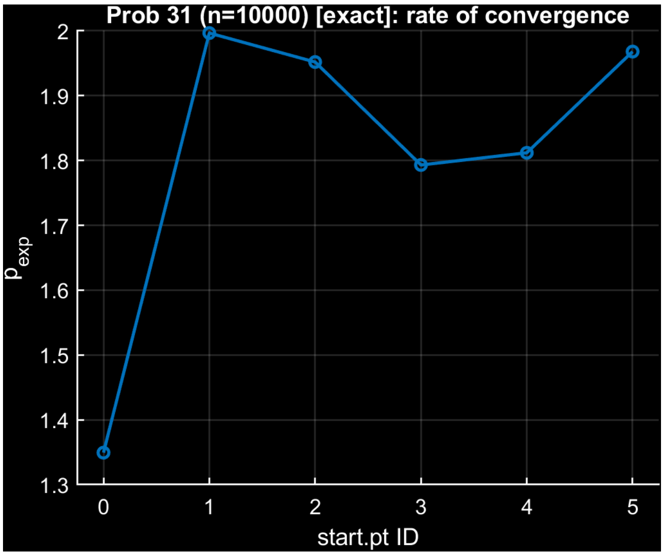


Figure 7: Problem 31, Truncated Newton, exact derivatives ($n = 10000$): experimental convergence rate p_{exp} for converged runs.

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
31	100000	exact	$\overline{x}$	1.94×10^{-14}	6/1000	yes		1.26	0.4384s
			1	1.70×10^{-13}	9/1000	yes		1.53	0.9005s
			2	2.44×10^{-14}	13/1000	yes		1.46	1.32s
			3	3.45×10^{-11}	10/1000	yes		2.00	1.18s
			4	1.09×10^{-13}	11/1000	yes		1.94	1.47s
			5	1.23×10^{-9}	10/1000	yes		2.00	1.60s
			Avg (successes)	2.10×10^{-10}	9.8/1000	6/6		1.70	1.15s

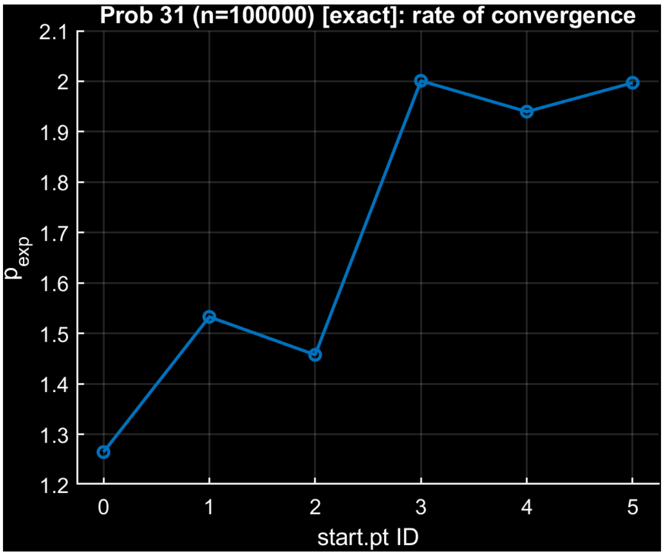


Figure 8: Problem 31, Truncated Newton, exact derivatives ($n = 100000$): experimental convergence rate p_{exp} for converged runs.

7.2 Problem 83

7.2.1 Modified Newton

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
83	2	exact	$\overline{x}$	1.34×10^{-9}	3/1000	yes		1.84	0.0136s
			1	7.49×10^{-10}	3/1000	yes		1.97	0.0027s
			2	4.39×10^{-15}	4/1000	yes		1.98	0.0014s
			3	2.58×10^{-9}	3/1000	yes		1.96	0.0018s
			4	2.70×10^{-16}	4/1000	yes		1.95	0.0032s
			5	9.34×10^{-15}	4/1000	yes		1.90	0.0009s
			Avg (successes)	7.78×10^{-10}	3.5/1000	6/6		1.93	0.0040s

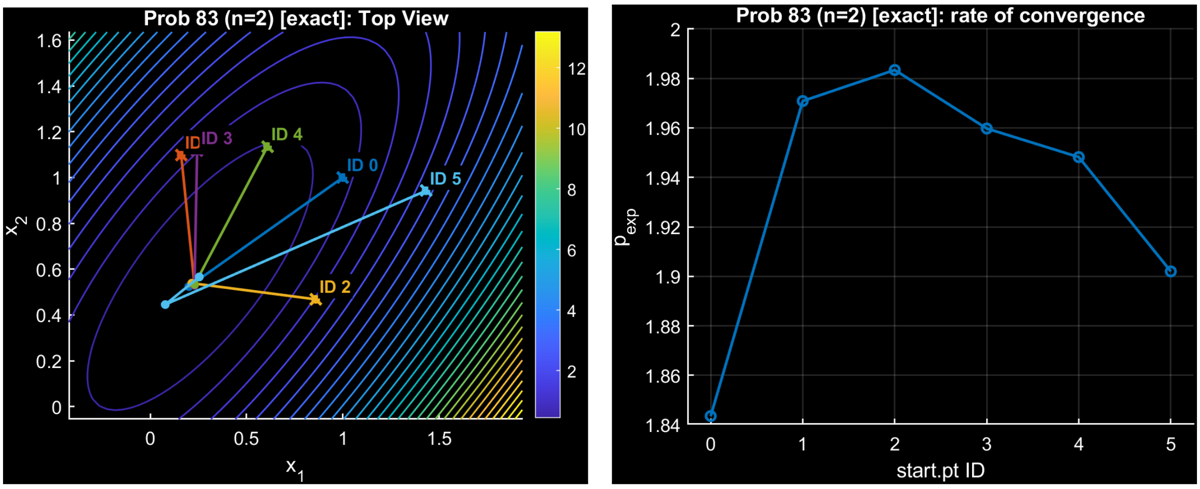


Figure 9: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
83	1000	exact	$\overline{x}$	5.24×10^{-13}	2/1000	yes		NaN	0.0058s
			1	3.91×10^{-10}	3/1000	yes		0.48	0.0041s
			2	4.18×10^{-11}	3/1000	yes		0.72	0.0037s
			3	9.06×10^{-10}	3/1000	yes		0.40	0.0038s
			4	7.95×10^{-10}	3/1000	yes		0.42	0.0037s
			5	1.32×10^{-9}	3/1000	yes		0.33	0.0033s
			Avg (successes)	5.76×10^{-10}	2.8/1000	6/6		0.47	0.0041s

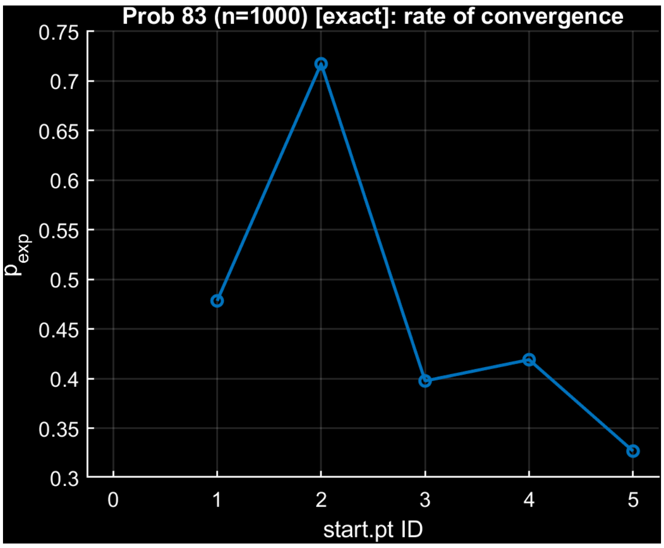


Figure 10: Problem 83, Modified Newton, exact derivatives ($n = 1000$): experimental convergence rate p_{exp} for converged runs.

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
83	10000	exact	$\overline{x}$	2.48×10^{-14}	2/1000	yes		NaN	0.0125s
			1	2.32×10^{-10}	3/1000	yes		0.94	0.0208s
			2	7.81×10^{-10}	4/1000	yes		5.08	0.0279s
			3	5.63×10^{-11}	3/1000	yes		1.02	0.0205s
			4	4.56×10^{-10}	2/1000	yes		NaN	0.0147s
			5	2.77×10^{-10}	2/1000	yes		NaN	0.0145s
			Avg (successes)	3.00×10^{-10}	2.7/1000	6/6		2.35	0.0185s

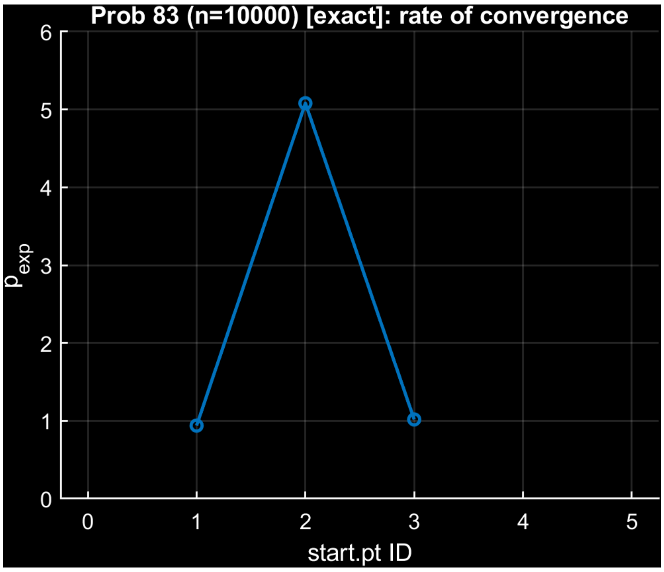


Figure 11: Problem 83, Modified Newton, exact derivatives ($n = 10000$): experimental convergence rate p_{exp} for converged runs.

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
83	100000	exact	$\overline{x}$	1.14×10^{-10}	1/1000	yes		NaN	0.0854s
			1	1.69×10^{-13}	2/1000	yes		NaN	0.1599s
			2	1.86×10^{-13}	2/1000	yes		NaN	0.1494s
			3	1.94×10^{-13}	2/1000	yes		NaN	0.1597s
			4	2.04×10^{-13}	2/1000	yes		NaN	0.1490s
			5	2.74×10^{-9}	2/1000	yes		NaN	0.1882s
			Avg (successes)	4.76×10^{-10}	1.8/1000	6/6		NaN	0.1486s

7.2.2 Truncated Newton

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
83	2	exact	$\overline{x}$	1.34×10^{-9}	3/1000	yes		1.84	0.0136s
			1	7.49×10^{-10}	3/1000	yes		1.97	0.0027s
			2	4.39×10^{-15}	4/1000	yes		1.98	0.0014s
			3	2.58×10^{-9}	3/1000	yes		1.96	0.0018s
			4	2.70×10^{-16}	4/1000	yes		1.95	0.0032s
			5	9.34×10^{-15}	4/1000	yes		1.90	0.0009s
			Avg (successes)	7.78×10^{-10}	3.5/1000	6/6		1.93	0.0040s

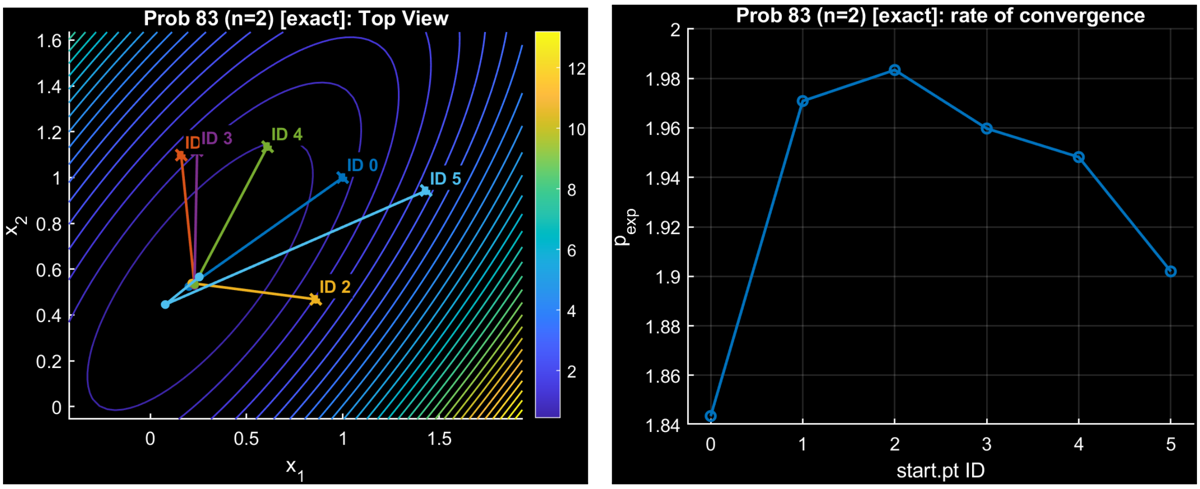


Figure 12: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
83	1000	exact	$\overline{x}$	5.24×10^{-13}	2/1000	yes		NaN	0.0058s
			1	3.91×10^{-10}	3/1000	yes		0.48	0.0041s
			2	4.18×10^{-11}	3/1000	yes		0.72	0.0037s
			3	9.06×10^{-10}	3/1000	yes		0.40	0.0038s
			4	7.95×10^{-10}	3/1000	yes		0.42	0.0037s
			5	1.32×10^{-9}	3/1000	yes		0.33	0.0033s
			Avg (successes)	5.76×10^{-10}	2.8/1000	6/6		0.47	0.0041s

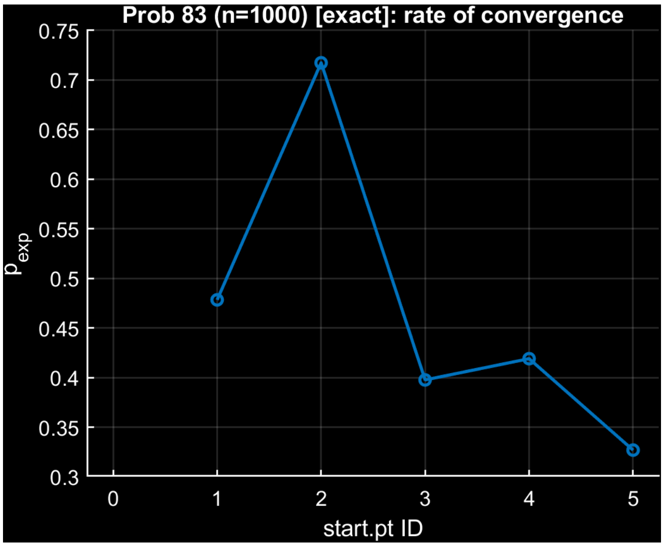


Figure 13: Problem 83, Truncated Newton, exact derivatives ($n = 1000$): experimental convergence rate p_{exp} for converged runs.

Prob.	Dim.	Mode	start.pt ID	grad.norm	iters/max	succ.	flag	rate	time
83	10000	exact	$\overline{x}$	2.48×10^{-14}	2/1000	yes		NaN	0.0125s
			1	2.32×10^{-10}	3/1000	yes		0.94	0.0208s
			2	7.81×10^{-10}	4/1000	yes		5.08	0.0279s
			3	5.63×10^{-11}	3/1000	yes		1.02	0.0205s
			4	4.56×10^{-10}	2/1000	yes		NaN	0.0147s
			5	2.77×10^{-10}	2/1000	yes		NaN	0.0145s
			Avg (successes)	3.00×10^{-10}	2.7/1000	6/6		2.35	0.0185s

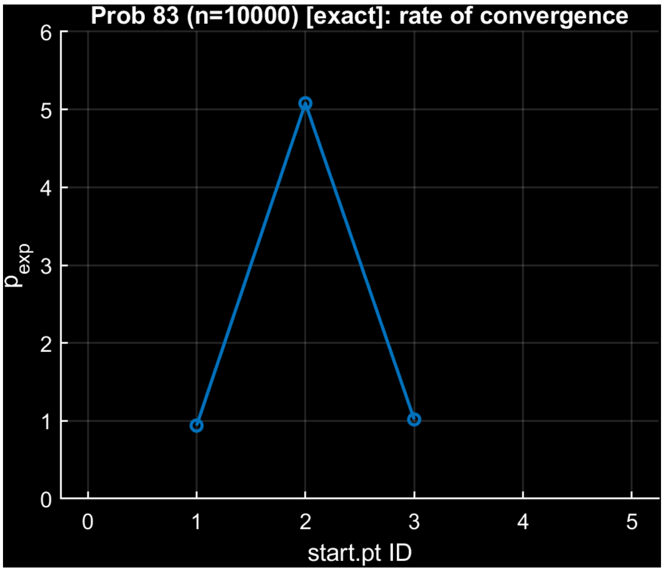


Figure 14: Problem 83, Truncated Newton, exact derivatives ($n = 10000$): experimental convergence rate p_{exp} for converged runs.

Prob.	**Dim.**	**Mode**	**start.pt ID**	**grad.norm**	**iters/max**	**succ.**	**flag**	**rate**	**time**
83	100000	exact	$\overline{x}$	1.14×10^{-10}	1/1000	yes		NaN	0.0854s
			1	1.69×10^{-13}	2/1000	yes		NaN	0.1599s
			2	1.86×10^{-13}	2/1000	yes		NaN	0.1494s
			3	1.94×10^{-13}	2/1000	yes		NaN	0.1597s
			4	2.04×10^{-13}	2/1000	yes		NaN	0.1490s
			5	2.74×10^{-9}	2/1000	yes		NaN	0.1882s
			Avg (successes)	4.76×10^{-10}	1.8/1000	6/6		NaN	0.1486s

8 Results - Mixed (constant h)

8.1 Problem 31

8.1.1 Modified Newton

Mixed FD (constant h), $k = 4$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	mixed fd	$\overline{x}$	4	2.14×10^{-7}	5/1000	yes		1.89	0.0013s
			1	4	2.17×10^{-7}	6/1000	yes		1.89	0.0008s
			2	4	1.22×10^{-8}	6/1000	yes		1.69	0.0008s
			3	4	7.20×10^{-9}	6/1000	yes		1.64	0.0008s
			4	4	3.70×10^{-9}	5/1000	yes		1.57	0.0008s
			5	4	2.55×10^{-10}	8/1000	yes		1.30	0.0014s
			Avg (successes)	4	7.56×10^{-8}	6.0/1000	6/6		1.66	0.0010s

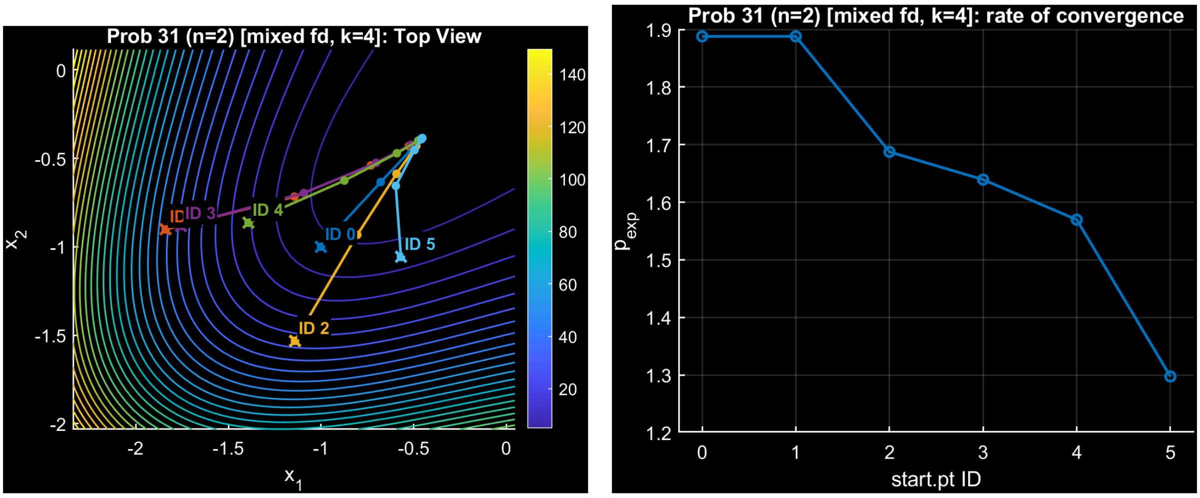


Figure 15: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	mixed fd	$\overline{x}$	4	5.08×10^{-8}	5/1000	yes		1.58	0.0053s
			1	4	1.91×10^{-10}	10/1000	yes		1.24	0.0123s
			2	4	4.62×10^{-7}	9/1000	yes		1.88	0.0090s
			3	4	9.03×10^{-11}	9/1000	yes		1.21	0.0115s
			4	4	7.51×10^{-11}	9/1000	yes		1.17	0.0113s
			5	4	3.14×10^{-10}	9/1000	yes		1.30	0.0103s
			Avg (successes)	4	8.55×10^{-8}	8.5/1000	6/6		1.40	0.0099s

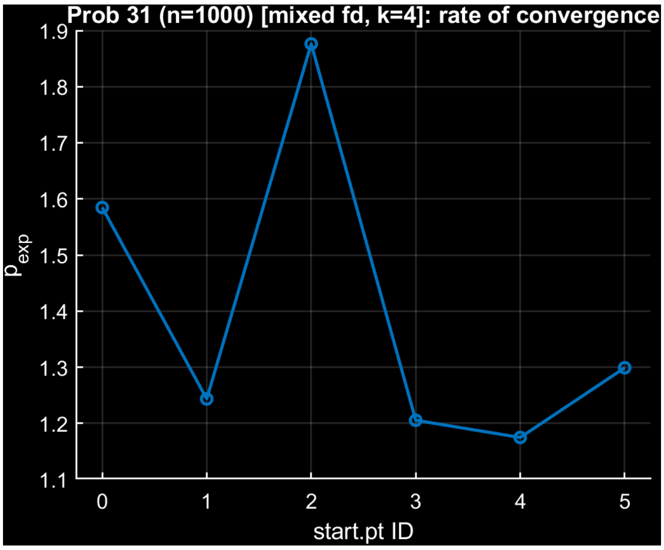


Figure 16: Problem 31, Modified Newton, mixed FD (constant h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	mixed fd	$\overline{x}$	4	5.85×10^{-8}	5/1000	yes		1.40	0.0368s
			1	4	8.75×10^{-10}	10/1000	yes		1.37	0.0833s
			2	4	1.65×10^{-8}	9/1000	yes		1.66	0.0741s
			3	4	2.64×10^{-7}	10/1000	yes		1.90	0.0840s
			4	4	6.23×10^{-7}	8/1000	yes		1.84	0.0683s
			5	4	7.37×10^{-7}	10/1000	yes		1.90	0.0893s
			Avg (successes)	4	2.83×10^{-7}	8.7/1000	6/6		1.68	0.0726s

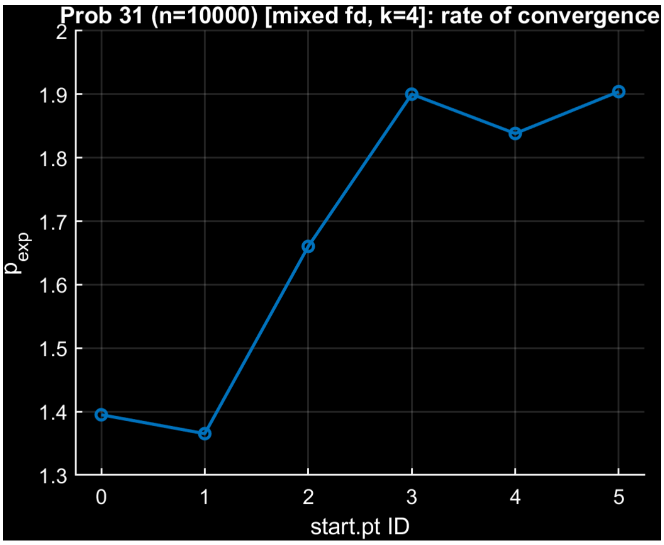


Figure 17: Problem 31, Modified Newton, mixed FD (constant h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	mixed fd	$\overline{x}$	4	1.09×10^{-7}	5/100	yes		1.33	0.4854s
			1	4	2.72×10^{-7}	8/100	yes		1.81	1.04s
			2	4	6.13×10^{-8}	12/100	yes		1.82	1.51s
			3	4	1.56×10^{-9}	10/100	yes		1.45	1.29s
			4	4	1.25×10^{-10}	11/100	yes		1.17	1.54s
			5	4	1.38×10^{-7}	12/100	yes		1.84	1.68s
			Avg (successes)	4	9.69×10^{-8}	9.7/100	6/6		1.57	1.26s

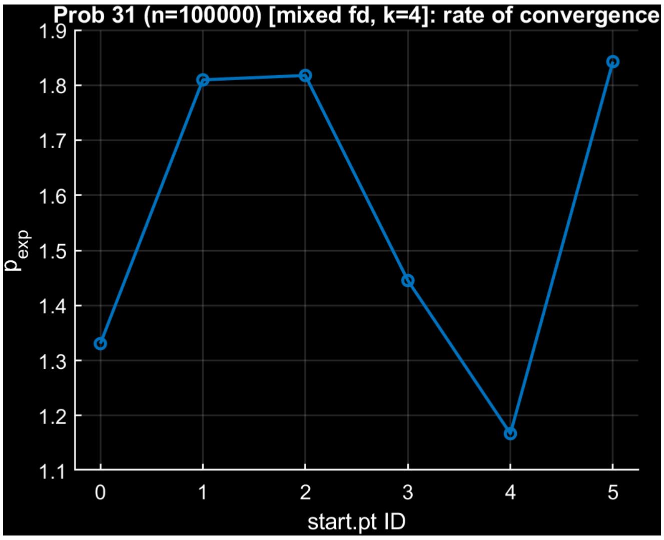
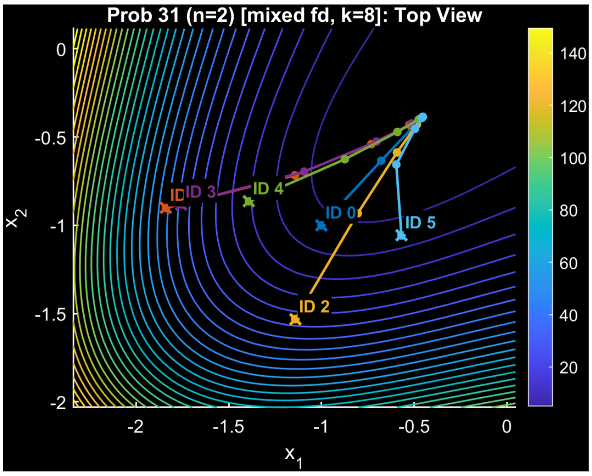


Figure 18: Problem 31, Modified Newton, mixed FD (constant h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Mixed FD (constant h), $k = 8$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	mixed fd	$\overline{x}$	8	1.32×10^{-7}	5/1000	yes		1.99	0.0006s
			1	8	1.40×10^{-7}	6/1000	yes		1.98	0.0008s
			2	8	2.17×10^{-9}	6/1000	yes		1.99	0.0010s
			3	8	8.95×10^{-10}	6/1000	yes		1.99	0.0009s
			4	8	2.68×10^{-10}	5/1000	yes		1.99	0.0007s
			5	8	1.07×10^{-12}	8/1000	yes		2.00	0.0013s
			Avg (successes)	8	4.59×10^{-8}	6.0/1000	6/6		1.99	0.0009s



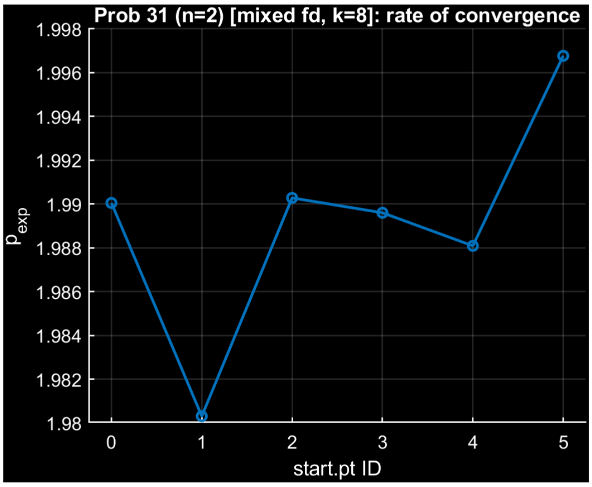


Figure 19: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	mixed fd	$\overline{x}$	8	1.79×10^{-8}	5/1000	yes		1.76	0.0054s
			1	8	4.85×10^{-13}	10/1000	yes		1.98	0.0111s
			2	8	3.30×10^{-7}	9/1000	yes		1.95	0.0085s
			3	8	1.09×10^{-13}	9/1000	yes		1.98	0.0105s
			4	8	8.28×10^{-7}	8/1000	yes		1.93	0.0105s
			5	8	1.46×10^{-12}	9/1000	yes		1.99	0.0107s
			Avg (successes)	8	1.96×10^{-7}	8.3/1000	6/6		1.93	0.0094s

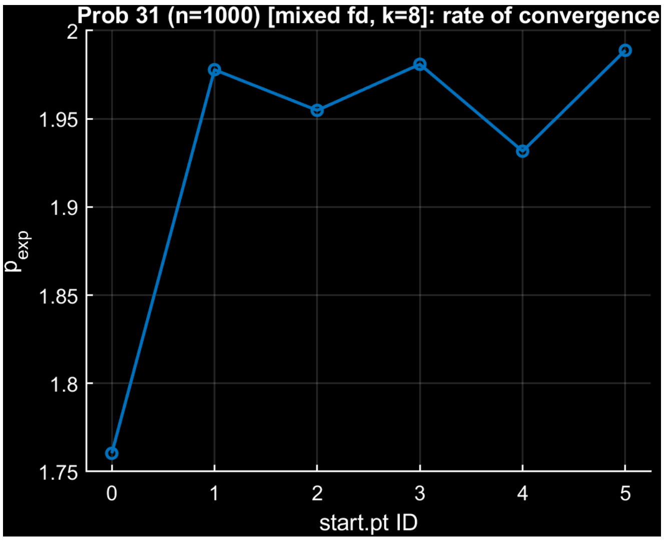


Figure 20: Problem 31, Modified Newton, mixed FD (constant h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	mixed fd	$\bar{x}$	8	1.79×10^{-8}	5/1000	yes		1.56	0.0394s
			1	8	9.52×10^{-12}	10/1000	yes		2.00	0.0833s
			2	8	3.12×10^{-9}	9/1000	yes		1.95	0.0735s
			3	8	2.15×10^{-7}	10/1000	yes		1.99	0.0805s
			4	8	4.62×10^{-7}	8/1000	yes		1.90	0.0672s
			5	8	5.96×10^{-7}	10/1000	yes		1.96	0.0824s
			Avg (successes)	8	2.16×10^{-7}	8.7/1000	6/6		1.89	0.0710s

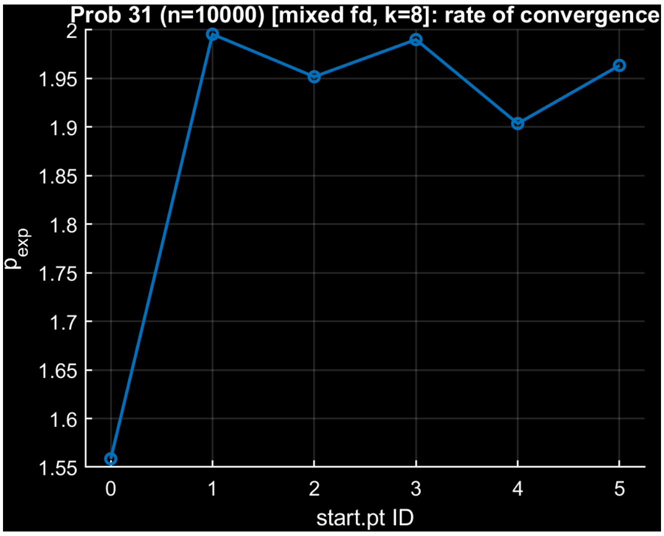


Figure 21: Problem 31, Modified Newton, mixed FD (constant h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	mixed fd	$\bar{x}$	8	1.79×10^{-8}	5/100	yes		1.56	0.5159s
			1	8	1.73×10^{-7}	8/100	yes		1.90	1.05s
			2	8	3.80×10^{-8}	12/100	yes		1.99	1.50s
			3	8	3.46×10^{-11}	10/100	yes		2.00	1.26s
			4	8	1.15×10^{-13}	11/100	yes		1.93	1.56s
			5	8	1.23×10^{-9}	10/100	yes		2.00	1.50s
			Avg (successes)	8	3.84×10^{-8}	9.3/100	6/6		1.90	1.23s

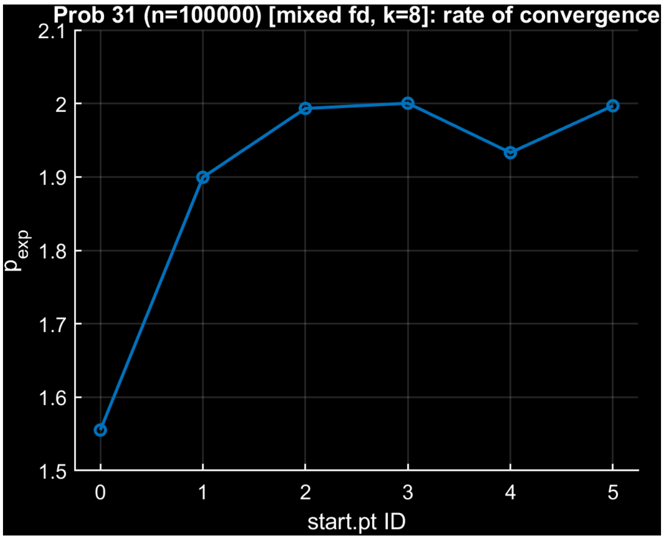


Figure 22: Problem 31, Modified Newton, mixed FD (constant h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Mixed FD (constant h), $k = 12$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	mixed fd	$\overline{x}$	12	1.32×10^{-7}	5/1000	yes		1.99	0.0007s
			1	12	1.40×10^{-7}	6/1000	yes		1.98	0.0013s
			2	12	2.16×10^{-9}	6/1000	yes		1.99	0.0009s
			3	12	8.94×10^{-10}	6/1000	yes		1.99	0.0009s
			4	12	2.67×10^{-10}	5/1000	yes		1.99	0.0009s
			5	12	1.04×10^{-12}	8/1000	yes		2.00	0.0016s
			Avg (successes)	12	4.59×10^{-8}	6.0/1000	6/6		1.99	0.0010s

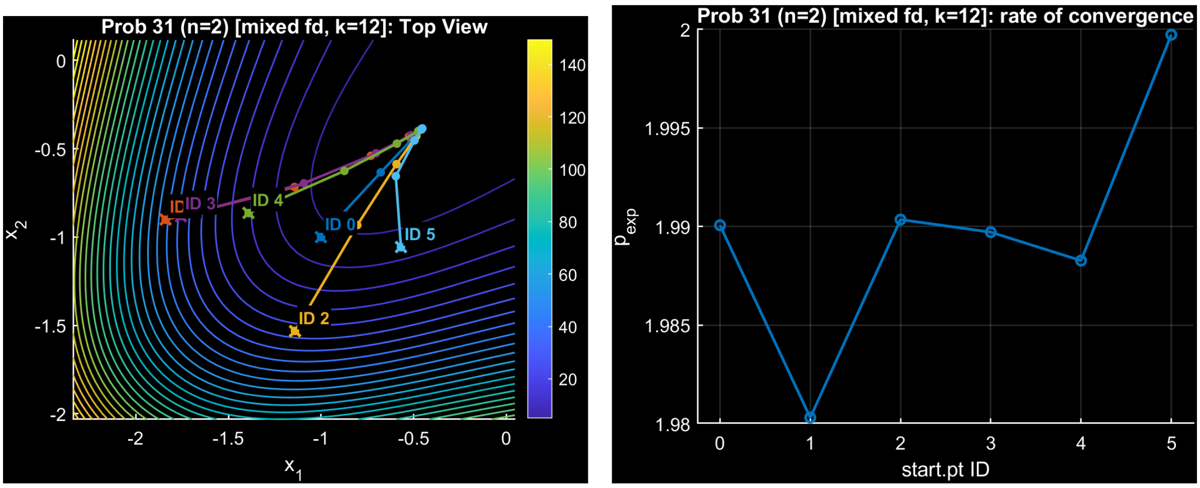


Figure 23: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	mixed fd	$\overline{x}$	12	1.79×10^{-8}	5/1000	yes		1.76	0.0052s
			1	12	4.70×10^{-13}	10/1000	yes		1.98	0.0116s
			2	12	3.30×10^{-7}	9/1000	yes		1.95	0.0090s
			3	12	1.06×10^{-13}	9/1000	yes		1.98	0.0104s
			4	12	8.28×10^{-7}	8/1000	yes		1.93	0.0097s
			5	12	1.43×10^{-12}	9/1000	yes		1.99	0.0111s
			Avg (successes)	12	1.96×10^{-7}	8.3/1000	6/6		1.93	0.0095s

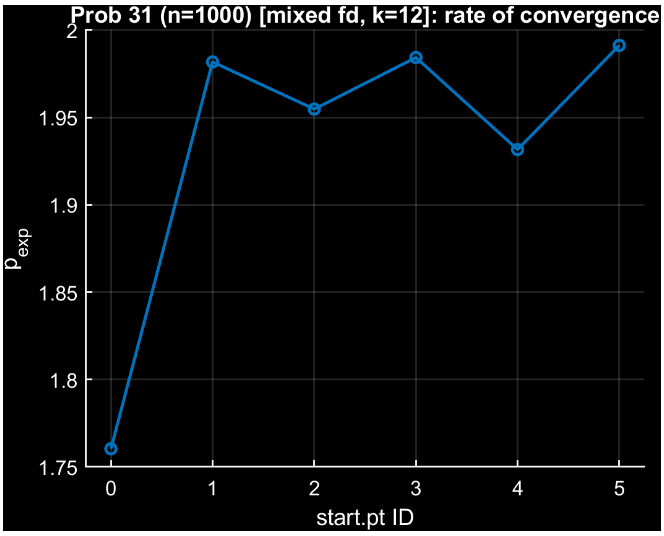


Figure 24: Problem 31, Modified Newton, mixed FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	mixed fd	$\overline{x}$	12	1.79×10^{-8}	5/1000	yes		1.56	0.0349s
			1	12	9.45×10^{-12}	10/1000	yes		2.00	0.0805s
			2	12	3.12×10^{-9}	9/1000	yes		1.95	0.0751s
			3	12	2.15×10^{-7}	10/1000	yes		1.99	0.0834s
			4	12	4.62×10^{-7}	8/1000	yes		1.90	0.0672s
			5	12	5.96×10^{-7}	10/1000	yes		1.96	0.0802s
			Avg (successes)	12	2.16×10^{-7}	8.7/1000	6/6		1.89	0.0702s

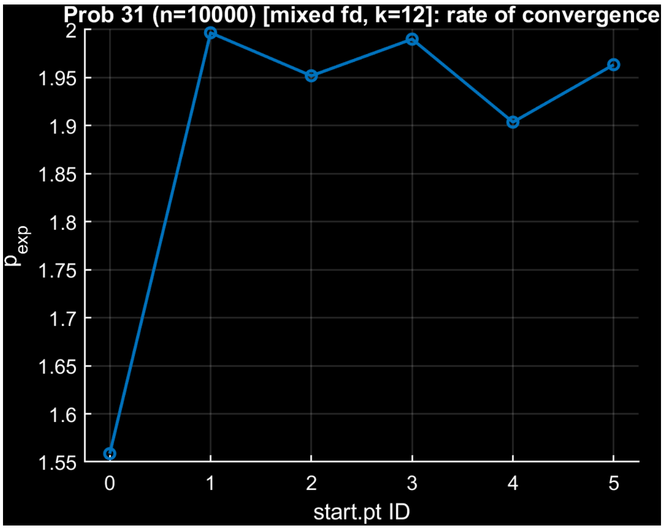


Figure 25: Problem 31, Modified Newton, mixed FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	mixed fd	$\overline{x}$	12	1.79×10^{-8}	5/100	yes		1.56	0.5098s
			1	12	1.73×10^{-7}	8/100	yes		1.90	1.05s
			2	12	3.80×10^{-8}	12/100	yes		1.99	1.46s
			3	12	3.45×10^{-11}	10/100	yes		2.00	1.24s
			4	12	1.08×10^{-13}	11/100	yes		1.94	1.51s
			5	12	1.23×10^{-9}	10/100	yes		2.00	1.43s
			Avg (successes)	12	3.84×10^{-8}	9.3/100	6/6		1.90	1.20s

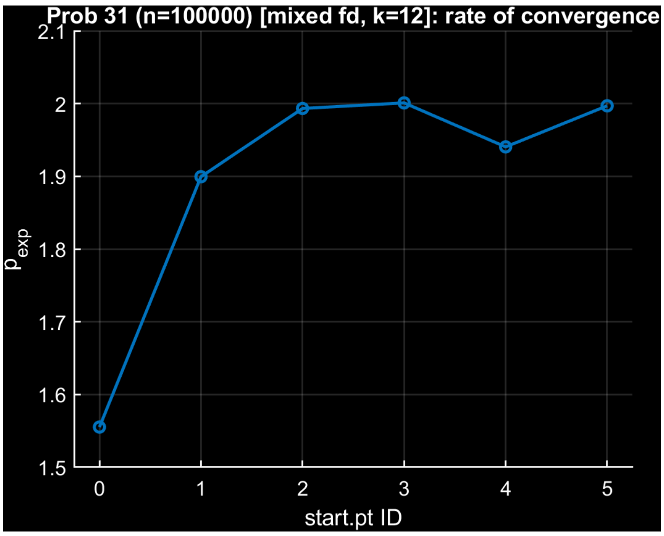


Figure 26: Problem 31, Modified Newton, mixed FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

8.1.2 Truncated Newton

Mixed FD (constant h), $k = 4$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	mixed fd	$\overline{x}$	4	2.14×10^{-7}	5/1000	yes		1.89	0.0013s
			1	4	2.17×10^{-7}	6/1000	yes		1.89	0.0008s
			2	4	1.22×10^{-8}	6/1000	yes		1.69	0.0008s
			3	4	7.20×10^{-9}	6/1000	yes		1.64	0.0008s
			4	4	3.70×10^{-9}	5/1000	yes		1.57	0.0008s
			5	4	2.55×10^{-10}	8/1000	yes		1.30	0.0014s
			Avg (successes)	4	7.56×10^{-8}	6.0/1000	6/6		1.66	0.0010s

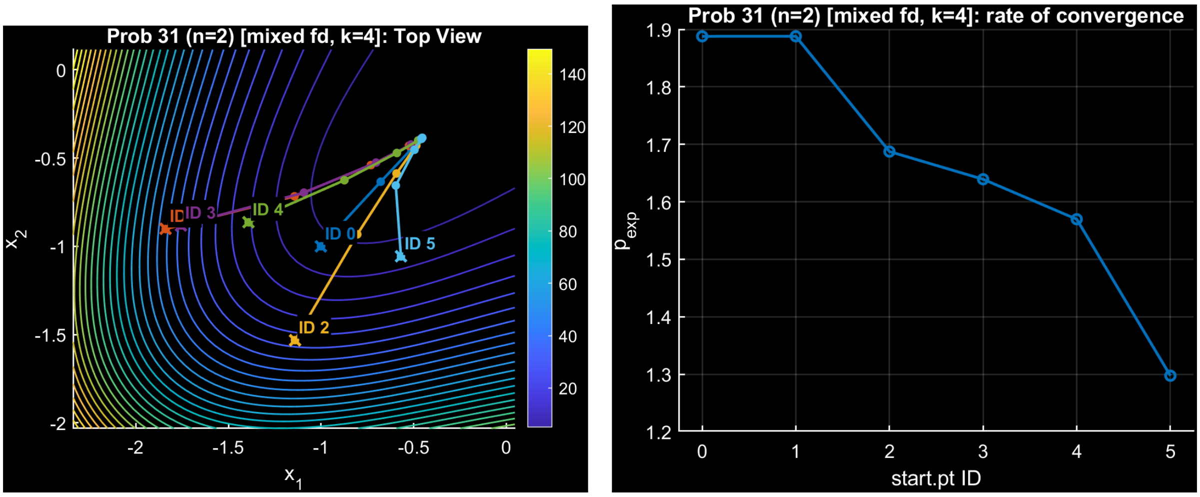


Figure 27: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	mixed fd	$\overline{x}$	4	5.08×10^{-8}	5/1000	yes		1.58	0.0053s
			1	4	1.91×10^{-10}	10/1000	yes		1.24	0.0123s
			2	4	4.62×10^{-7}	9/1000	yes		1.88	0.0090s
			3	4	9.03×10^{-11}	9/1000	yes		1.21	0.0115s
			4	4	7.51×10^{-11}	9/1000	yes		1.17	0.0113s
			5	4	3.14×10^{-10}	9/1000	yes		1.30	0.0103s
			Avg (successes)	4	8.55×10^{-8}	8.5/1000	6/6		1.40	0.0099s

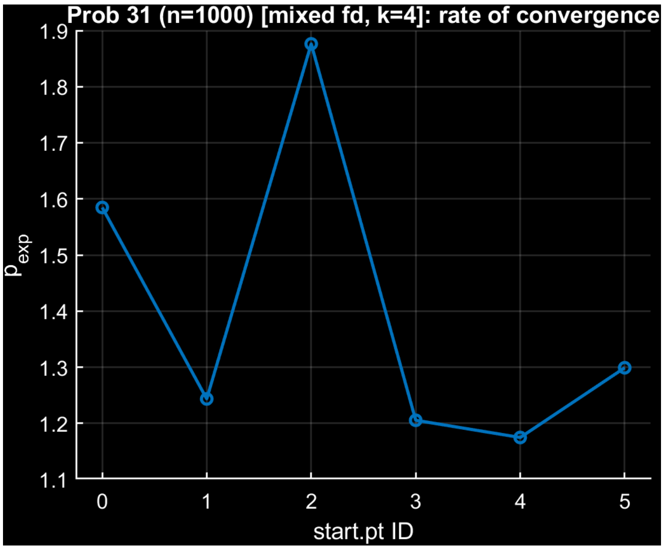


Figure 28: Problem 31, Truncated Newton, mixed FD (constant h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	mixed fd	$\overline{x}$	4	5.85×10^{-8}	5/1000	yes		1.40	0.0368s
			1	4	8.75×10^{-10}	10/1000	yes		1.37	0.0833s
			2	4	1.65×10^{-8}	9/1000	yes		1.66	0.0741s
			3	4	2.64×10^{-7}	10/1000	yes		1.90	0.0840s
			4	4	6.23×10^{-7}	8/1000	yes		1.84	0.0683s
			5	4	7.37×10^{-7}	10/1000	yes		1.90	0.0893s
			Avg (successes)	4	2.83×10^{-7}	8.7/1000	6/6		1.68	0.0726s

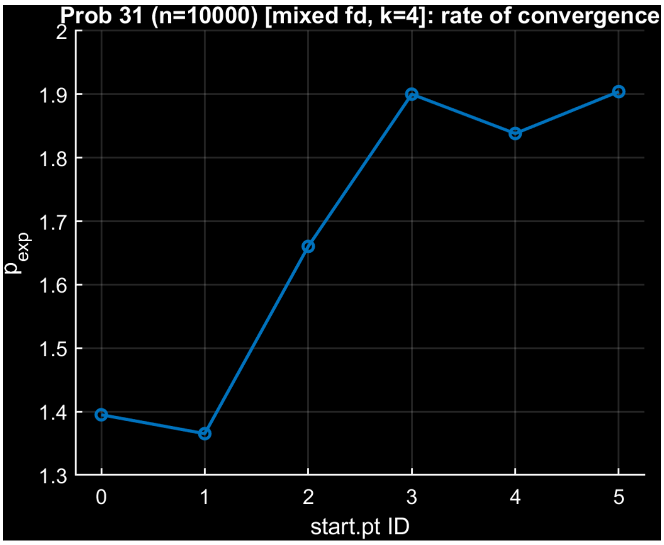


Figure 29: Problem 31, Truncated Newton, mixed FD (constant h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	mixed fd	$\overline{x}$	4	1.09×10^{-7}	5/100	yes		1.33	0.4854s
			1	4	2.72×10^{-7}	8/100	yes		1.81	1.04s
			2	4	6.13×10^{-8}	12/100	yes		1.82	1.51s
			3	4	1.56×10^{-9}	10/100	yes		1.45	1.29s
			4	4	1.25×10^{-10}	11/100	yes		1.17	1.54s
			5	4	1.38×10^{-7}	12/100	yes		1.84	1.68s
			Avg (successes)	4	9.69×10^{-8}	9.7/100	6/6		1.57	1.26s

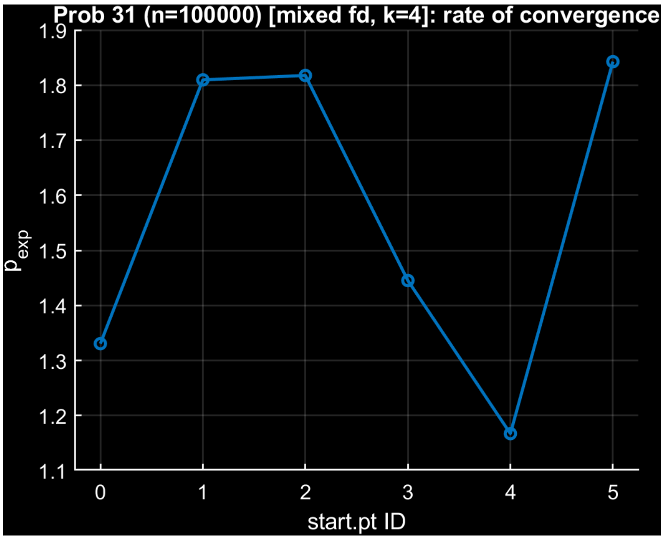


Figure 30: Problem 31, Truncated Newton, mixed FD (constant h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Mixed FD (constant h), $k = 8$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	mixed fd	$\overline{x}$	8	1.32×10^{-7}	5/1000	yes		1.99	0.0006s
			1	8	1.40×10^{-7}	6/1000	yes		1.98	0.0008s
			2	8	2.17×10^{-9}	6/1000	yes		1.99	0.0010s
			3	8	8.95×10^{-10}	6/1000	yes		1.99	0.0009s
			4	8	2.68×10^{-10}	5/1000	yes		1.99	0.0007s
			5	8	1.07×10^{-12}	8/1000	yes		2.00	0.0013s
			Avg (successes)	8	4.59×10^{-8}	6.0/1000	6/6		1.99	0.0009s

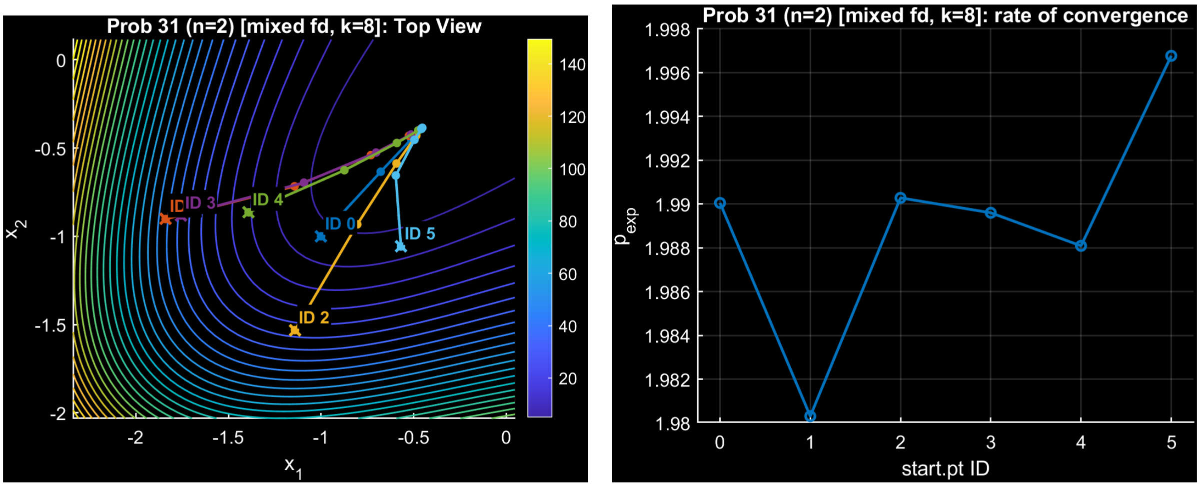


Figure 31: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	mixed fd	$\overline{x}$	8	1.79×10^{-8}	5/1000	yes		1.76	0.0054s
			1	8	4.85×10^{-13}	10/1000	yes		1.98	0.0111s
			2	8	3.30×10^{-7}	9/1000	yes		1.95	0.0085s
			3	8	1.09×10^{-13}	9/1000	yes		1.98	0.0105s
			4	8	8.28×10^{-7}	8/1000	yes		1.93	0.0105s
			5	8	1.46×10^{-12}	9/1000	yes		1.99	0.0107s
			Avg (successes)	8	1.96×10^{-7}	8.3/1000	6/6		1.93	0.0094s

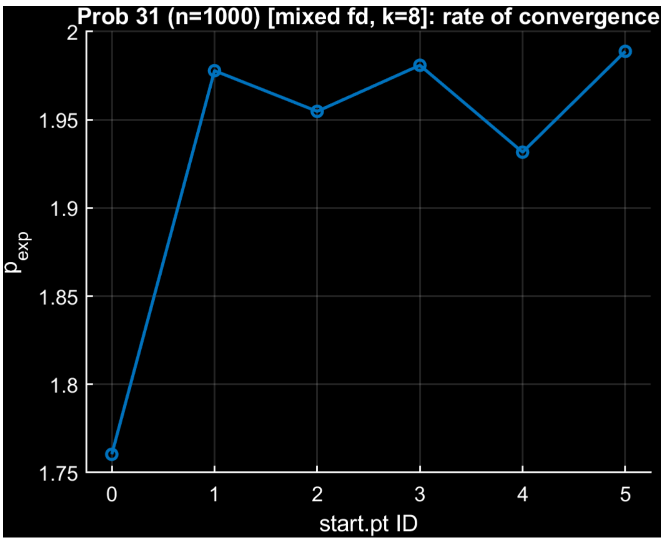


Figure 32: Problem 31, Truncated Newton, mixed FD (constant h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	mixed fd	$\overline{x}$	8	1.79×10^{-8}	5/1000	yes		1.56	0.0394s
			1	8	9.52×10^{-12}	10/1000	yes		2.00	0.0833s
			2	8	3.12×10^{-9}	9/1000	yes		1.95	0.0735s
			3	8	2.15×10^{-7}	10/1000	yes		1.99	0.0805s
			4	8	4.62×10^{-7}	8/1000	yes		1.90	0.0672s
			5	8	5.96×10^{-7}	10/1000	yes		1.96	0.0824s
			Avg (successes)	8	2.16×10^{-7}	8.7/1000	6/6		1.89	0.0710s

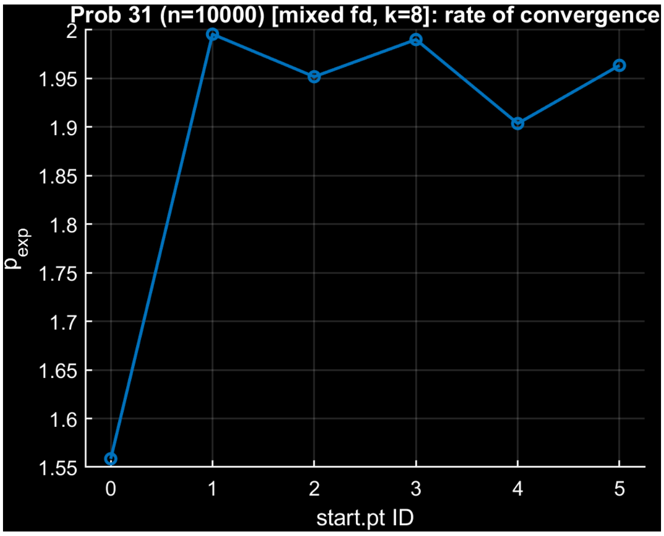


Figure 33: Problem 31, Truncated Newton, mixed FD (constant h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	mixed fd	$\overline{x}$	8	1.79×10^{-8}	5/100	yes		1.56	0.5159s
			1	8	1.73×10^{-7}	8/100	yes		1.90	1.05s
			2	8	3.80×10^{-8}	12/100	yes		1.99	1.50s
			3	8	3.46×10^{-11}	10/100	yes		2.00	1.26s
			4	8	1.15×10^{-13}	11/100	yes		1.93	1.56s
			5	8	1.23×10^{-9}	10/100	yes		2.00	1.50s
			Avg (successes)	8	3.84×10^{-8}	9.3/100	6/6		1.90	1.23s

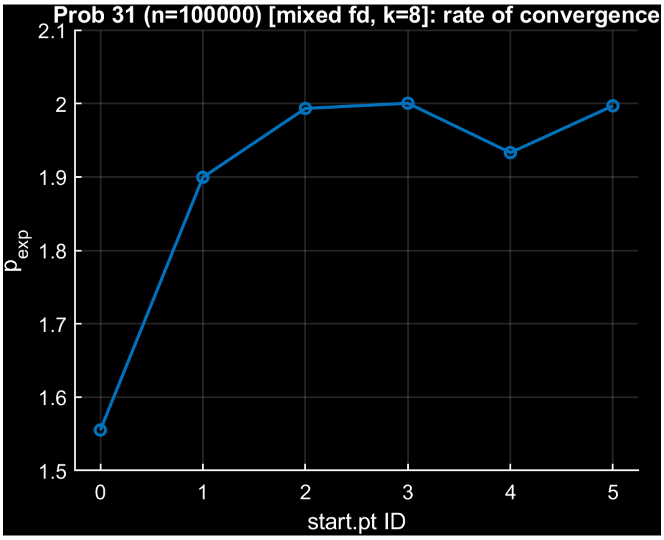


Figure 34: Problem 31, Truncated Newton, mixed FD (constant h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Mixed FD (constant h), $k = 12$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	mixed fd	$\overline{x}$	12	1.32×10^{-7}	5/1000	yes		1.99	0.0007s
			1	12	1.40×10^{-7}	6/1000	yes		1.98	0.0013s
			2	12	2.16×10^{-9}	6/1000	yes		1.99	0.0009s
			3	12	8.94×10^{-10}	6/1000	yes		1.99	0.0009s
			4	12	2.67×10^{-10}	5/1000	yes		1.99	0.0009s
			5	12	1.04×10^{-12}	8/1000	yes		2.00	0.0016s
			Avg (successes)	12	4.59×10^{-8}	6.0/1000	6/6		1.99	0.0010s

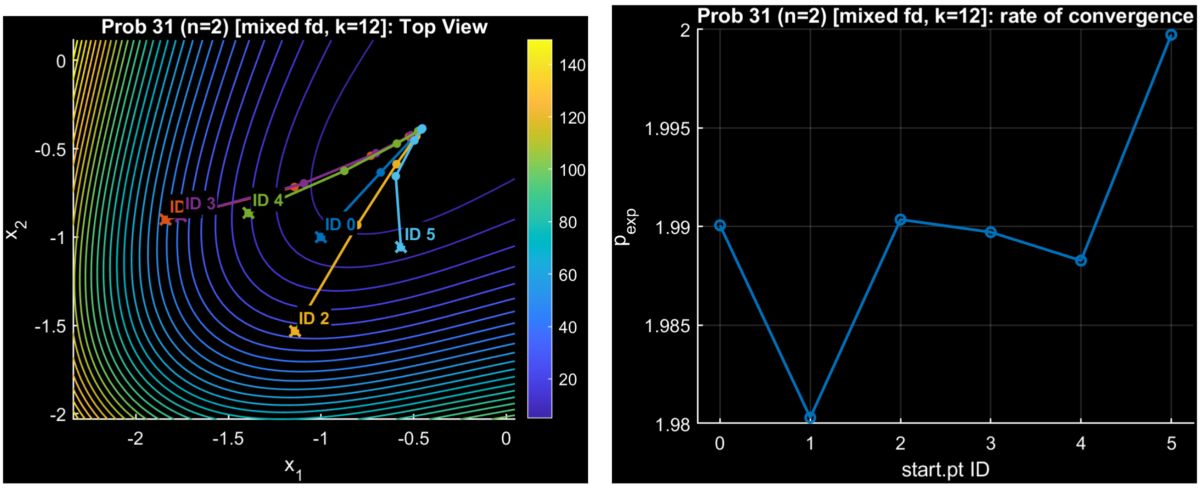


Figure 35: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	mixed fd	$\overline{x}$	12	1.79×10^{-8}	5/1000	yes		1.76	0.0052s
			1	12	4.70×10^{-13}	10/1000	yes		1.98	0.0116s
			2	12	3.30×10^{-7}	9/1000	yes		1.95	0.0090s
			3	12	1.06×10^{-13}	9/1000	yes		1.98	0.0104s
			4	12	8.28×10^{-7}	8/1000	yes		1.93	0.0097s
			5	12	1.43×10^{-12}	9/1000	yes		1.99	0.0111s
			Avg (successes)	12	1.96×10^{-7}	8.3/1000	6/6		1.93	0.0095s

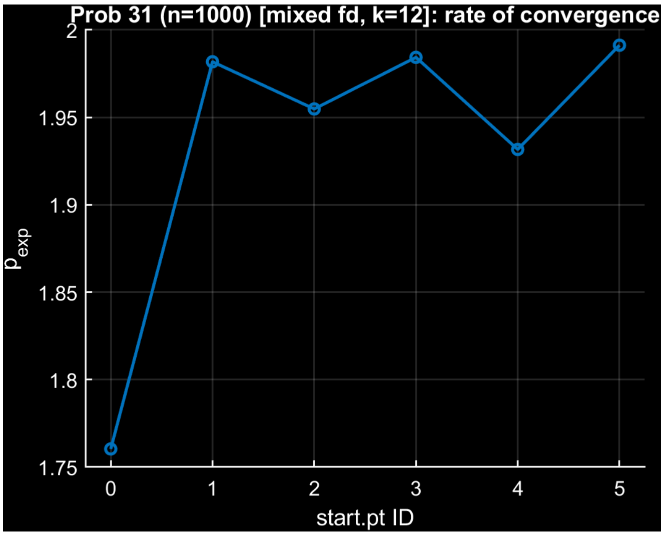


Figure 36: Problem 31, Truncated Newton, mixed FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	mixed fd	$\overline{x}$	12	1.79×10^{-8}	5/1000	yes		1.56	0.0349s
			1	12	9.45×10^{-12}	10/1000	yes		2.00	0.0805s
			2	12	3.12×10^{-9}	9/1000	yes		1.95	0.0751s
			3	12	2.15×10^{-7}	10/1000	yes		1.99	0.0834s
			4	12	4.62×10^{-7}	8/1000	yes		1.90	0.0672s
			5	12	5.96×10^{-7}	10/1000	yes		1.96	0.0802s
			Avg (successes)	12	2.16×10^{-7}	8.7/1000	6/6		1.89	0.0702s

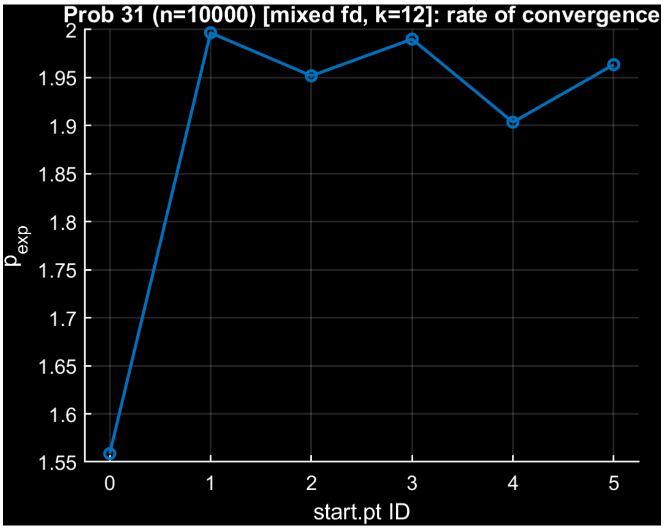


Figure 37: Problem 31, Truncated Newton, mixed FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	mixed fd	$\overline{x}$	12	1.79×10^{-8}	5/100	yes		1.56	0.5098s
			1	12	1.73×10^{-7}	8/100	yes		1.90	1.05s
			2	12	3.80×10^{-8}	12/100	yes		1.99	1.46s
			3	12	3.45×10^{-11}	10/100	yes		2.00	1.24s
			4	12	1.08×10^{-13}	11/100	yes		1.94	1.51s
			5	12	1.23×10^{-9}	10/100	yes		2.00	1.43s
			Avg (successes)	12	3.84×10^{-8}	9.3/100	6/6		1.90	1.20s

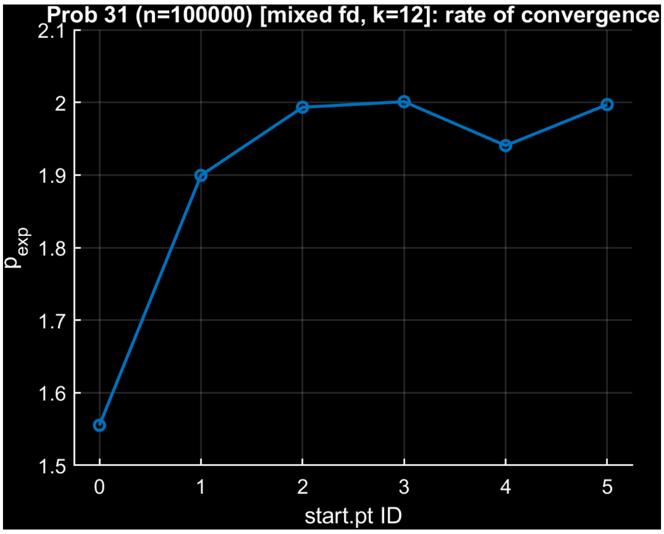


Figure 38: Problem 31, Truncated Newton, mixed FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

8.2 Problem 83

8.2.1 Modified Newton

Mixed FD (constant h), $k = 4$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	mixed fd	$\overline{x}$	4	1.07×10^{-9}	3/1000	yes		1.88	0.0258s
			1	4	7.21×10^{-10}	3/1000	yes		1.98	0.0143s
			2	4	4.13×10^{-7}	3/1000	yes		1.95	0.0137s
			3	4	1.61×10^{-9}	3/1000	yes		2.04	0.0142s
			4	4	7.45×10^{-8}	3/1000	yes		2.03	0.0150s
			5	4	6.41×10^{-12}	4/1000	yes		1.24	0.0194s
			Avg (successes)	4	8.18×10^{-8}	3.2/1000	6/6		1.85	0.0171s

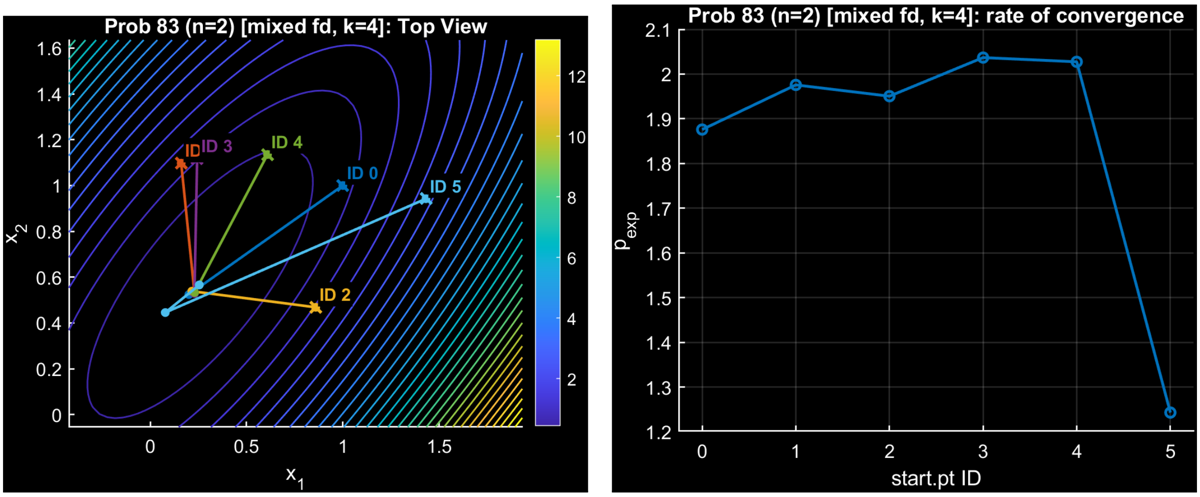


Figure 39: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	mixed fd	$\overline{x}$	4	5.71×10^{-7}	2/1000	yes		NaN	0.0159s
			1	4	3.32×10^{-8}	2/1000	yes		NaN	0.0160s
			2	4	2.02×10^{-8}	2/1000	yes		NaN	0.0149s
			3	4	3.18×10^{-8}	2/1000	yes		NaN	0.0154s
			4	4	2.70×10^{-8}	2/1000	yes		NaN	0.0152s
			5	4	2.73×10^{-8}	2/1000	yes		NaN	0.0148s
			Avg (successes)	4	1.18×10^{-7}	2.0/1000	6/6		NaN	0.0154s

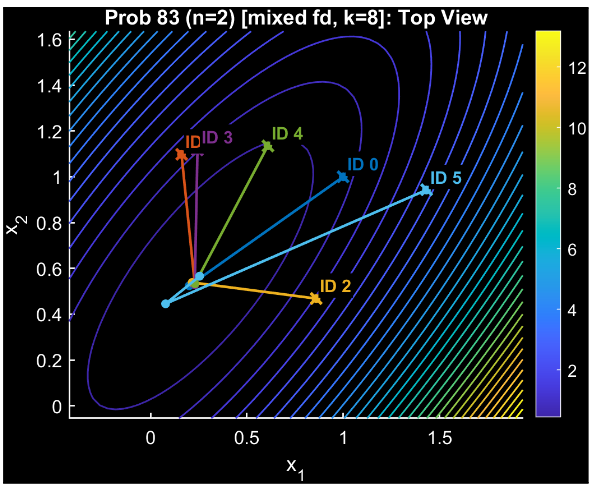
Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	mixed fd	$\bar{x}$	4	1.42×10^{-8}	1/1000	yes		NaN	0.0196s
			1	4	7.05×10^{-10}	2/1000	yes		NaN	0.0407s
			2	4	9.28×10^{-10}	2/1000	yes		NaN	0.0395s
			3	4	2.69×10^{-8}	2/1000	yes		NaN	0.0353s
			4	4	4.10×10^{-10}	2/1000	yes		NaN	0.0328s
			5	4	3.57×10^{-10}	2/1000	yes		NaN	0.0342s
			Avg (successes)	4	7.24×10^{-9}	1.8/1000	6/6		NaN	0.0337s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	mixed fd	$\bar{x}$	4	9.67×10^{-9}	1/100	yes		NaN	0.1631s
			1	4	1.68×10^{-7}	1/100	yes		NaN	0.1678s
			2	4	1.68×10^{-7}	1/100	yes		NaN	0.1656s
			3	4	1.68×10^{-7}	1/100	yes		NaN	0.1669s
			4	4	1.68×10^{-7}	1/100	yes		NaN	0.1638s
			5	4	1.68×10^{-7}	1/100	yes		NaN	0.1726s
			Avg (successes)	4	1.41×10^{-7}	1.0/100	6/6		NaN	0.1666s

Mixed FD (constant h), $k = 8$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	mixed fd	$\bar{x}$	8	1.34×10^{-9}	3/1000	yes		1.84	0.0142s
			1	8	7.49×10^{-10}	3/1000	yes		1.97	0.0149s
			2	8	4.34×10^{-7}	3/1000	yes		1.94	0.0144s
			3	8	2.58×10^{-9}	3/1000	yes		1.96	0.0155s
			4	8	8.39×10^{-8}	3/1000	yes		2.00	0.0158s
			5	8	8.71×10^{-15}	4/1000	yes		1.91	0.0194s
			Avg (successes)	8	8.71×10^{-8}	3.2/1000	6/6		1.94	0.0157s



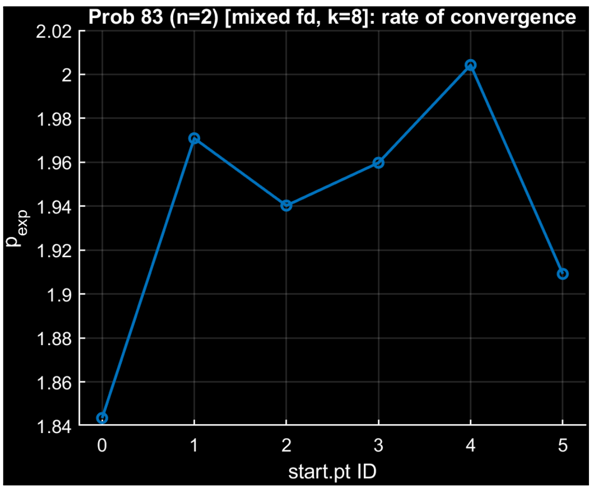


Figure 40: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	mixed fd	$\overline{x}$	8	6.67×10^{-13}	2/1000	yes		NaN	0.0130s
			1	8	2.51×10^{-8}	2/1000	yes		NaN	0.0136s
			2	8	2.51×10^{-8}	2/1000	yes		NaN	0.0146s
			3	8	2.80×10^{-8}	2/1000	yes		NaN	0.0147s
			4	8	2.94×10^{-8}	2/1000	yes		NaN	0.0144s
			5	8	2.36×10^{-8}	2/1000	yes		NaN	0.0151s
			Avg (successes)	8	2.19×10^{-8}	2.0/1000	6/6		NaN	0.0142s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	mixed fd	$\overline{x}$	8	1.14×10^{-8}	1/1000	yes		NaN	0.0171s
			1	8	4.65×10^{-8}	2/1000	yes		NaN	0.0412s
			2	8	3.90×10^{-8}	2/1000	yes		NaN	0.0396s
			3	8	2.68×10^{-8}	2/1000	yes		NaN	0.0405s
			4	8	1.77×10^{-9}	2/1000	yes		NaN	0.0390s
			5	8	1.50×10^{-9}	2/1000	yes		NaN	0.0393s
			Avg (successes)	8	2.12×10^{-8}	1.8/1000	6/6		NaN	0.0361s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	mixed fd	$\overline{x}$	8	1.17×10^{-10}	1/100	yes		NaN	0.1538s
			1	8	1.68×10^{-7}	1/100	yes		NaN	0.1675s
			2	8	1.68×10^{-7}	1/100	yes		NaN	0.1666s
			3	8	1.68×10^{-7}	1/100	yes		NaN	0.1682s
			4	8	1.68×10^{-7}	1/100	yes		NaN	0.1798s
			5	8	1.67×10^{-7}	1/100	yes		NaN	0.1752s
			Avg (successes)	8	1.40×10^{-7}	1.0/100	6/6		NaN	0.1685s

Mixed FD (constant h), $k = 12$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	mixed fd	$\overline{x}$	12	1.34×10^{-9}	3/1000	yes		1.84	0.0137s
			1	12	7.49×10^{-10}	3/1000	yes		1.97	0.0150s
			2	12	4.34×10^{-7}	3/1000	yes		1.94	0.0179s
			3	12	2.58×10^{-9}	3/1000	yes		1.96	0.0145s
			4	12	8.39×10^{-8}	3/1000	yes		2.00	0.0140s
			5	12	9.27×10^{-15}	4/1000	yes		1.90	0.0171s
			Avg (successes)	12	8.71×10^{-8}	3.2/1000	6/6		1.94	0.0154s

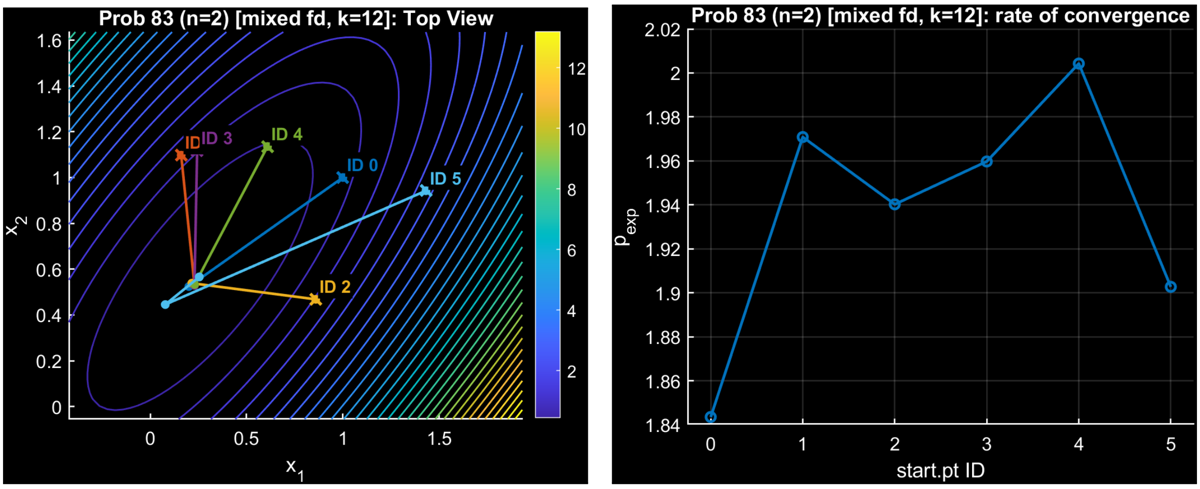


Figure 41: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	mixed fd	$\overline{x}$	12	4.63×10^{-11}	2/1000	yes		NaN	0.0134s
			1	12	2.62×10^{-8}	2/1000	yes		NaN	0.0161s
			2	12	2.89×10^{-8}	2/1000	yes		NaN	0.0156s
			3	12	2.90×10^{-8}	2/1000	yes		NaN	0.0148s
			4	12	3.20×10^{-8}	2/1000	yes		NaN	0.0157s
			5	12	2.64×10^{-8}	2/1000	yes		NaN	0.0145s
			Avg (successes)	12	2.37×10^{-8}	2.0/1000	6/6		NaN	0.0150s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	mixed fd	$\overline{x}$	12	1.13×10^{-8}	1/1000	yes		NaN	0.0180s
			1	12	5.30×10^{-8}	2/1000	yes		NaN	0.0377s
			2	12	4.31×10^{-8}	2/1000	yes		NaN	0.0394s
			3	12	2.90×10^{-8}	2/1000	yes		NaN	0.0385s
			4	12	1.40×10^{-7}	2/1000	yes		NaN	0.0380s
			5	12	1.43×10^{-7}	2/1000	yes		NaN	0.0379s
			Avg (successes)	12	6.99×10^{-8}	1.8/1000	6/6		NaN	0.0349s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	mixed fd	$\overline{x}$	12	3.23×10^{-10}	1/100	yes		NaN	0.1728s
			1	12	1.82×10^{-7}	1/100	yes		NaN	0.1848s
			2	12	1.82×10^{-7}	1/100	yes		NaN	0.1851s
			3	12	1.82×10^{-7}	1/100	yes		NaN	0.1639s
			4	12	1.83×10^{-7}	1/100	yes		NaN	0.1628s
			5	12	1.82×10^{-7}	1/100	yes		NaN	0.1665s
			Avg (successes)	12	1.52×10^{-7}	1.0/100	6/6		NaN	0.1726s

8.2.2 Truncated Newton

9 Results - Mixed (component-wise h)

9.1 Problem 31

9.1.1 Modified Newton

Mixed FD (component-wise h), $k = 4$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	mixed fd h prop	$\overline{x}$	4	1.46×10^{-7}	5/1000	yes		1.97	0.0042s
			1	4	1.55×10^{-7}	6/1000	yes		1.96	0.0009s
			2	4	3.81×10^{-9}	6/1000	yes		1.89	0.0008s
			3	4	2.05×10^{-9}	6/1000	yes		1.85	0.0014s
			4	4	8.96×10^{-10}	5/1000	yes		1.79	0.0007s
			5	4	4.23×10^{-11}	8/1000	yes		1.51	0.0010s
			Avg (successes)	4	5.13×10^{-8}	6.0/1000	6/6		1.83	0.0015s

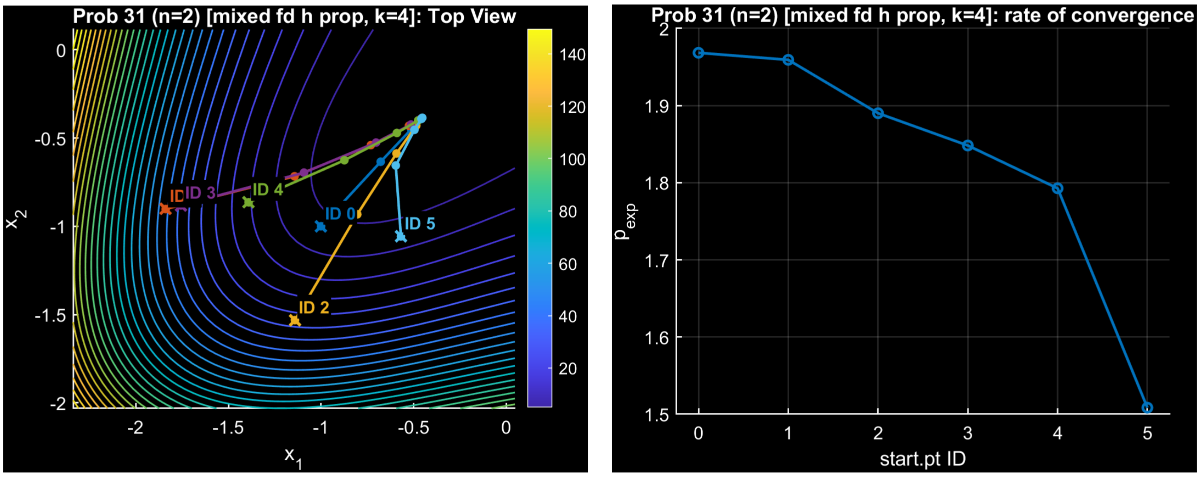


Figure 42: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	mixed fd h prop	$\overline{x}$	4	2.70×10^{-8}	5/1000	yes		1.69	0.0053s
			1	4	9.15×10^{-11}	10/1000	yes		1.33	0.0583s
			2	4	4.02×10^{-7}	9/1000	yes		1.91	0.0093s
			3	4	4.01×10^{-11}	9/1000	yes		1.28	0.0107s
			4	4	9.38×10^{-7}	8/1000	yes		1.90	0.0132s
			5	4	1.49×10^{-10}	9/1000	yes		1.39	0.0128s
			Avg (successes)	4	2.28×10^{-7}	8.3/1000	6/6		1.58	0.0182s

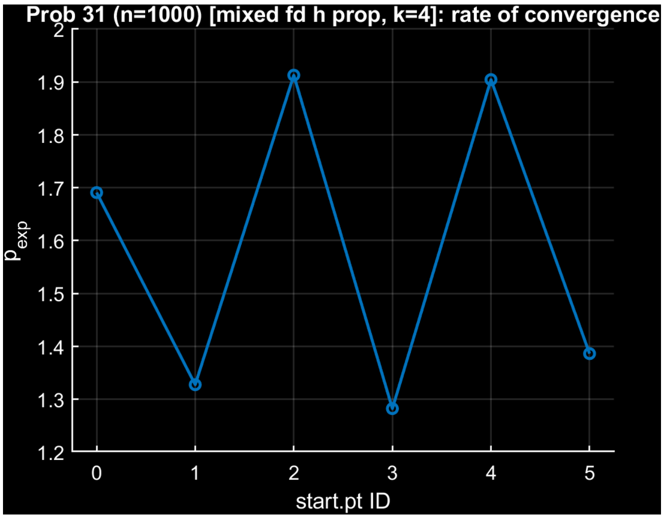


Figure 43: Problem 31, Modified Newton, mixed FD (component-wise h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	mixed fd h prop	$\overline{x}$	4	3.03×10^{-8}	5/1000	yes		1.49	0.0407s
			1	4	4.48×10^{-10}	10/1000	yes		1.46	0.0969s
			2	4	9.84×10^{-9}	9/1000	yes		1.75	0.0850s
			3	4	2.64×10^{-7}	10/1000	yes		1.94	0.0935s
			4	4	5.45×10^{-7}	8/1000	yes		1.87	0.0845s
			5	4	6.85×10^{-7}	10/1000	yes		1.93	0.0979s
			Avg (successes)	4	2.56×10^{-7}	8.7/1000	6/6		1.74	0.0831s

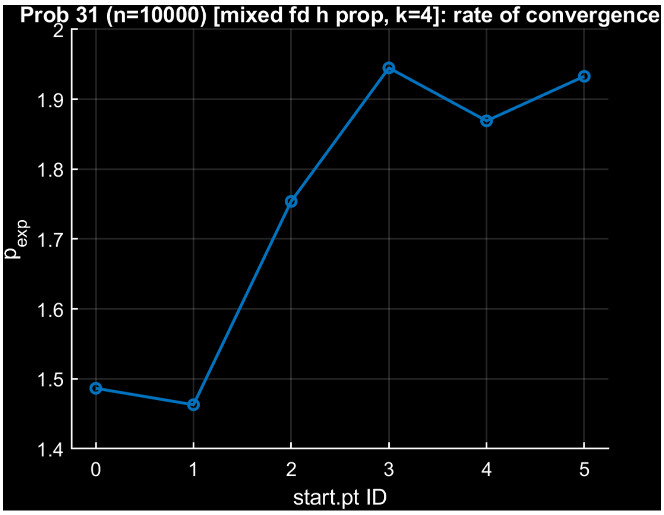


Figure 44: Problem 31, Modified Newton, mixed FD (component-wise h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	mixed fd h prop	$\overline{x}$	4	5.31×10^{-8}	5/100	yes		1.42	0.5230s
			1	4	2.21×10^{-7}	8/100	yes		1.85	1.14s
			2	4	5.96×10^{-8}	12/100	yes		1.90	1.54s
			3	4	7.71×10^{-10}	10/100	yes		1.54	1.40s
			4	4	5.39×10^{-11}	11/100	yes		1.24	1.63s
			5	4	5.16×10^{-9}	10/100	yes		1.74	1.53s
			Avg (successes)	4	5.67×10^{-8}	9.3/100	6/6		1.62	1.29s

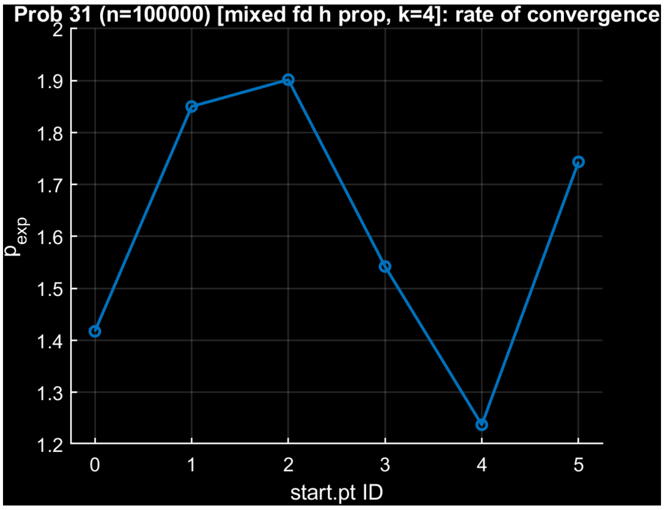
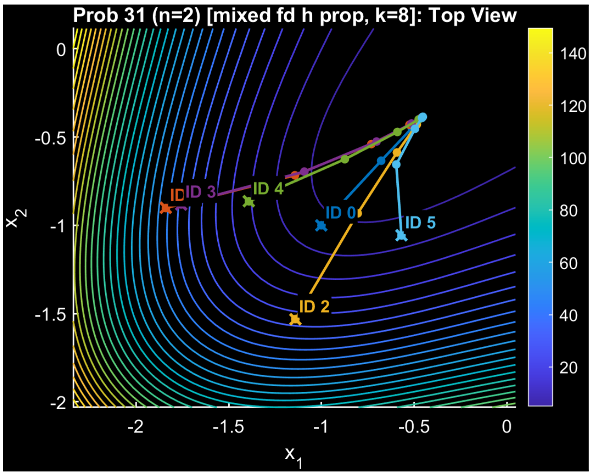


Figure 45: Problem 31, Modified Newton, mixed FD (component-wise h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Mixed FD (component-wise h), $k = 8$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	mixed fd h prop	$\overline{x}$	8	1.32×10^{-7}	5/1000	yes		1.99	0.0005s
			1	8	1.40×10^{-7}	6/1000	yes		1.98	0.0008s
			2	8	2.16×10^{-9}	6/1000	yes		1.99	0.0007s
			3	8	8.94×10^{-10}	6/1000	yes		1.99	0.0008s
			4	8	2.68×10^{-10}	5/1000	yes		1.99	0.0008s
			5	8	1.05×10^{-12}	8/1000	yes		2.00	0.0009s
			Avg (successes)	8	4.59×10^{-8}	6.0/1000	6/6		1.99	0.0008s



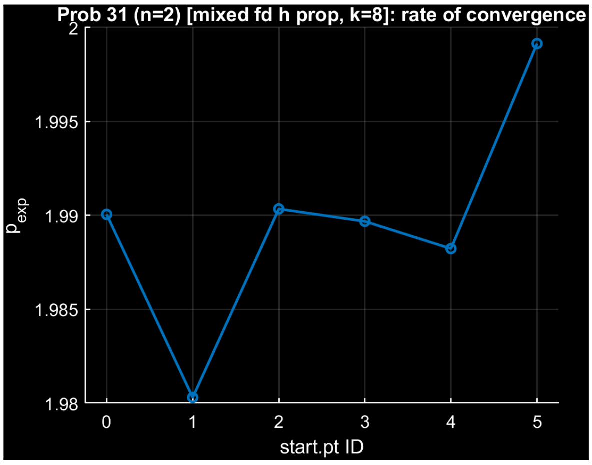


Figure 46: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	mixed fd h prop	$\overline{x}$	8	1.79×10^{-8}	5/1000	yes		1.76	0.0053s
			1	8	4.77×10^{-13}	10/1000	yes		1.98	0.0141s
			2	8	3.30×10^{-7}	9/1000	yes		1.95	0.0298s
			3	8	1.07×10^{-13}	9/1000	yes		1.98	0.0141s
			4	8	8.28×10^{-7}	8/1000	yes		1.93	0.0119s
			5	8	1.44×10^{-12}	9/1000	yes		1.99	0.0149s
			Avg (successes)	8	1.96×10^{-7}	8.3/1000	6/6		1.93	0.0150s

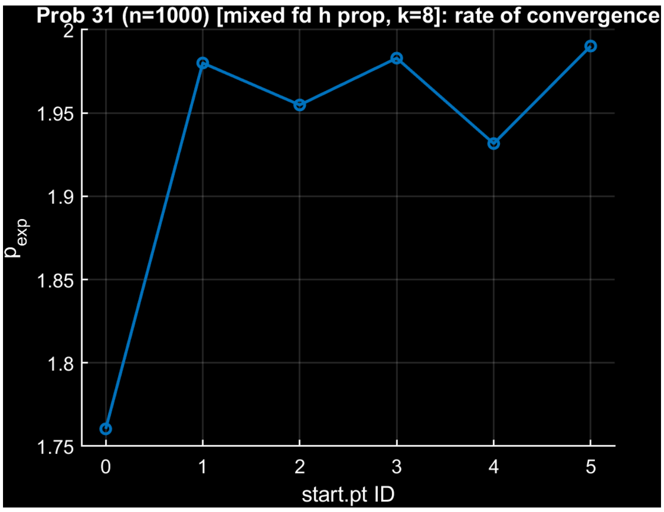


Figure 47: Problem 31, Modified Newton, mixed FD (component-wise h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	mixed fd h prop	$\overline{x}$	8	1.79×10^{-8}	5/1000	yes		1.56	0.0384s
			1	8	9.49×10^{-12}	10/1000	yes		2.00	0.0912s
			2	8	3.12×10^{-9}	9/1000	yes		1.95	0.0961s
			3	8	2.15×10^{-7}	10/1000	yes		1.99	0.0964s
			4	8	4.62×10^{-7}	8/1000	yes		1.90	0.0708s
			5	8	5.96×10^{-7}	10/1000	yes		1.96	0.0930s
			Avg (successes)	8	2.16×10^{-7}	8.7/1000	6/6		1.89	0.0810s

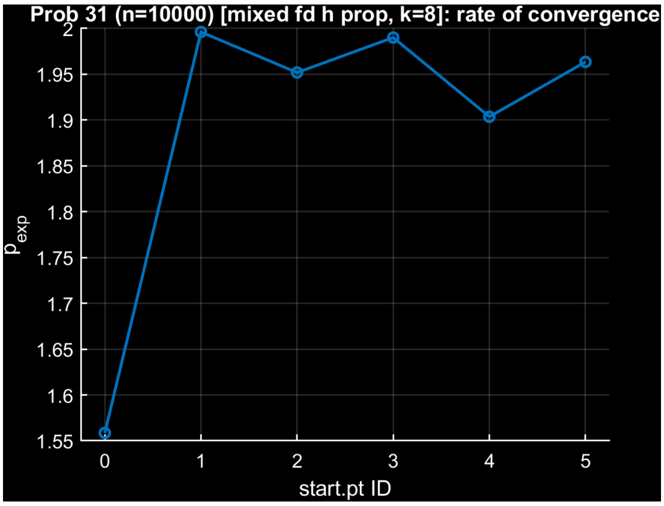


Figure 48: Problem 31, Modified Newton, mixed FD (component-wise h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	mixed fd h prop	$\overline{x}$	8	1.79×10^{-8}	5/100	yes		1.56	0.7282s
			1	8	1.73×10^{-7}	8/100	yes		1.90	1.79s
			2	8	3.80×10^{-8}	12/100	yes		1.99	1.73s
			3	8	3.45×10^{-11}	10/100	yes		2.00	1.70s
			4	8	1.12×10^{-13}	11/100	yes		1.94	2.10s
			5	8	1.23×10^{-9}	10/100	yes		2.00	1.57s
			Avg (successes)	8	3.84×10^{-8}	9.3/100	6/6		1.90	1.60s

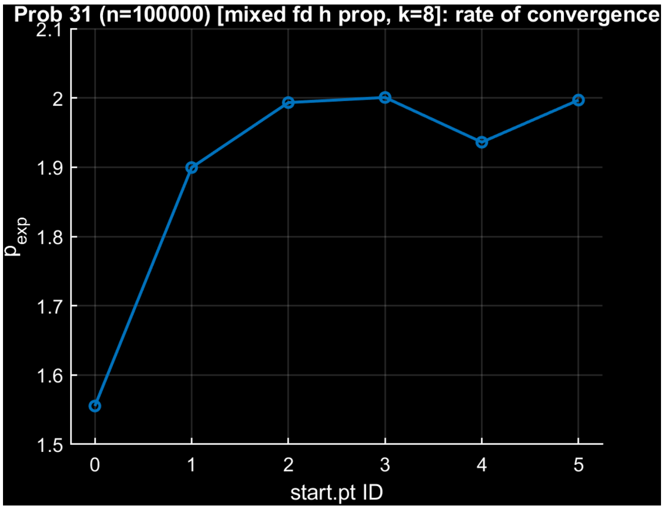


Figure 49: Problem 31, Modified Newton, mixed FD (component-wise h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Mixed FD (component-wise h), $k = 12$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	mixed fd h prop	$\overline{x}$	12	1.32×10^{-7}	5/1000	yes		1.99	0.0005s
			1	12	1.40×10^{-7}	6/1000	yes		1.98	0.0008s
			2	12	2.16×10^{-9}	6/1000	yes		1.99	0.0007s
			3	12	8.94×10^{-10}	6/1000	yes		1.99	0.0007s
			4	12	2.67×10^{-10}	5/1000	yes		1.99	0.0008s
			5	12	1.04×10^{-12}	8/1000	yes		2.00	0.0009s
			Avg (successes)	12	4.59×10^{-8}	6.0/1000	6/6		1.99	0.0008s

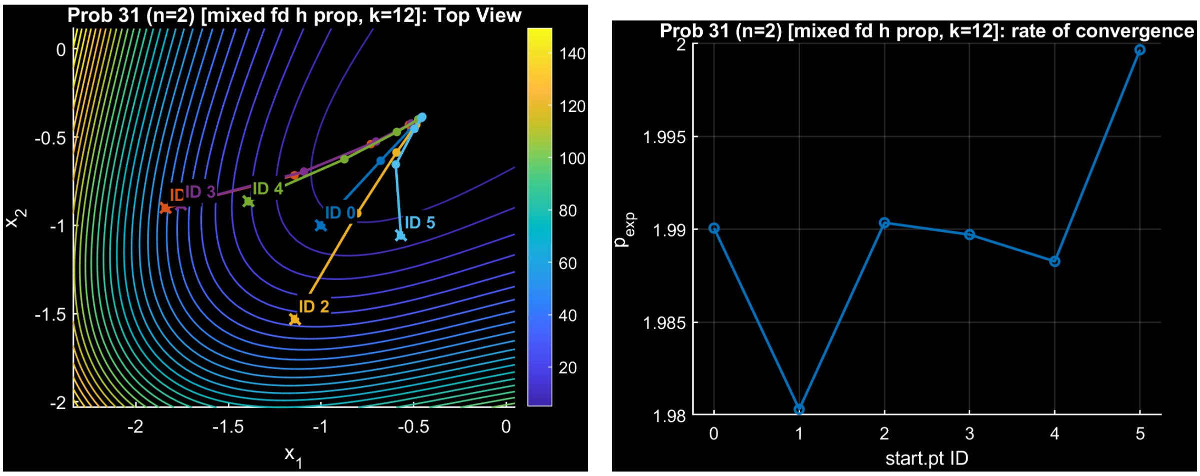


Figure 50: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	mixed fd h prop	$\overline{x}$	12	1.79×10^{-8}	5/1000	yes		1.76	0.0086s
			1	12	4.69×10^{-13}	10/1000	yes		1.98	0.0115s
			2	12	3.30×10^{-7}	9/1000	yes		1.95	0.0101s
			3	12	1.08×10^{-13}	9/1000	yes		1.98	0.0102s
			4	12	8.28×10^{-7}	8/1000	yes		1.93	0.0091s
			5	12	1.43×10^{-12}	9/1000	yes		1.99	0.0101s
			Avg (successes)	12	1.96×10^{-7}	8.3/1000	6/6		1.93	0.0099s

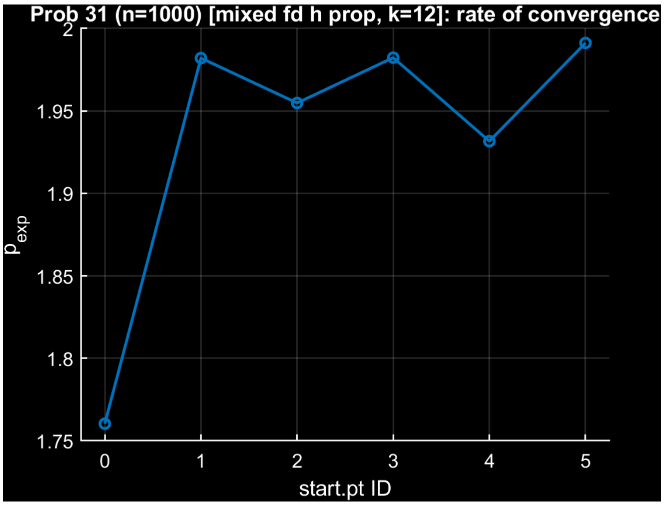


Figure 51: Problem 31, Modified Newton, mixed FD (component-wise h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	mixed fd h prop	$\overline{x}$	12	1.79×10^{-8}	5/1000	yes		1.56	0.0407s
			1	12	9.45×10^{-12}	10/1000	yes		2.00	0.0851s
			2	12	3.12×10^{-9}	9/1000	yes		1.95	0.0805s
			3	12	2.15×10^{-7}	10/1000	yes		1.99	0.1016s
			4	12	4.62×10^{-7}	8/1000	yes		1.90	0.0726s
			5	12	5.96×10^{-7}	10/1000	yes		1.96	0.0850s
			Avg (successes)	12	2.16×10^{-7}	8.7/1000	6/6		1.89	0.0776s

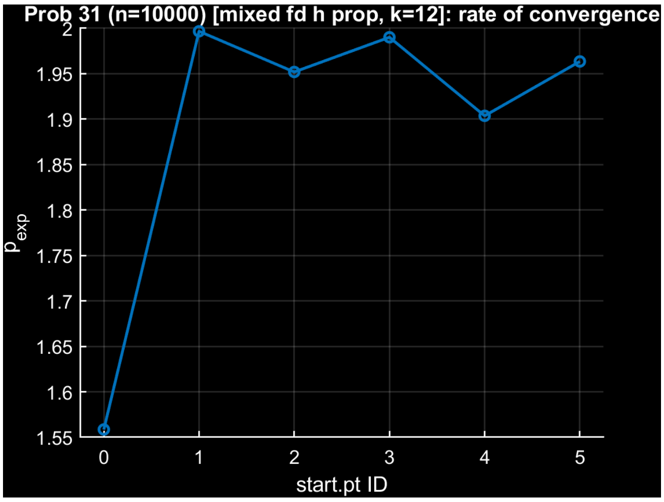


Figure 52: Problem 31, Modified Newton, mixed FD (component-wise h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	mixed fd h prop	$\overline{x}$	12	1.79×10^{-8}	5/100	yes		1.56	0.6381s
			1	12	1.73×10^{-7}	8/100	yes		1.90	1.60s
			2	12	3.80×10^{-8}	12/100	yes		1.99	1.72s
			3	12	3.45×10^{-11}	10/100	yes		2.00	1.45s
			4	12	1.08×10^{-13}	11/100	yes		1.94	1.77s
			5	12	1.23×10^{-9}	10/100	yes		2.00	1.51s
			Avg (successes)	12	3.84×10^{-8}	9.3/100	6/6		1.90	1.45s

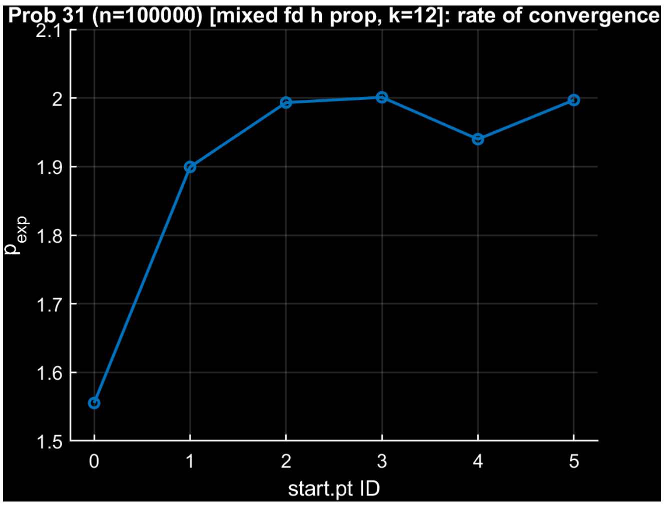


Figure 53: Problem 31, Modified Newton, mixed FD (component-wise h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

9.1.2 Truncated Newton

Mixed FD (component-wise h), $k = 4$

Case $n = 2$

Case $n > 2$

Mixed FD (component-wise h), $k = 8$

Case $n = 2$

Case $n > 2$

Mixed FD (component-wise h), $k = 12$

Case $n = 2$

Case $n > 2$

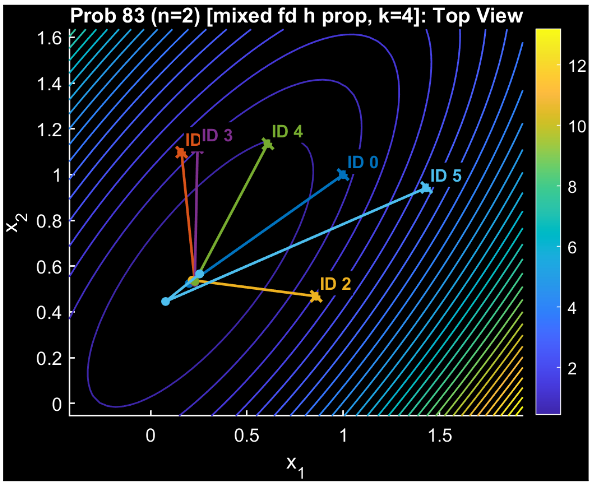
9.2 Problem 83

9.2.1 Modified Newton

Mixed FD (component-wise h), $k = 4$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	mixed fd h prop	$\overline{x}$	4	9.20×10^{-10}	3/1000	yes		1.90	0.0172s
			1	4	4.53×10^{-10}	3/1000	yes		2.05	0.0150s
			2	4	4.27×10^{-7}	3/1000	yes		1.94	0.0140s
			3	4	2.03×10^{-9}	3/1000	yes		2.00	0.0135s
			4	4	8.07×10^{-8}	3/1000	yes		2.01	0.0142s
			5	4	1.03×10^{-12}	4/1000	yes		1.42	0.0171s
			Avg (successes)	4	8.52×10^{-8}	3.2/1000	6/6		1.89	0.0152s



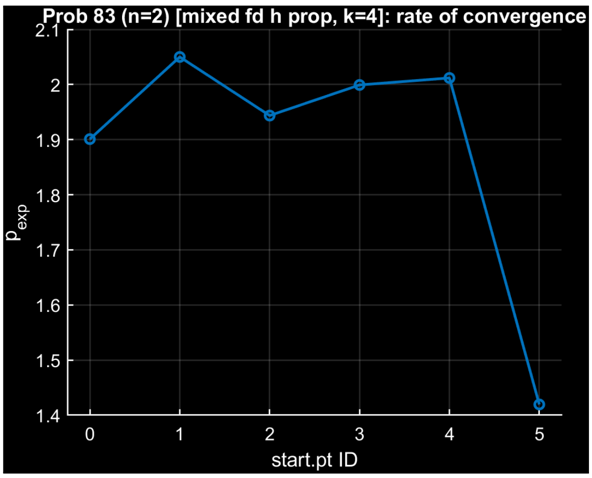


Figure 54: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	mixed fd h prop	$\overline{x}$	4	1.68×10^{-10}	2/1000	yes		NaN	0.0139s
			1	4	3.54×10^{-8}	2/1000	yes		NaN	0.0143s
			2	4	2.03×10^{-8}	2/1000	yes		NaN	0.0145s
			3	4	2.52×10^{-8}	2/1000	yes		NaN	0.0153s
			4	4	1.44×10^{-8}	2/1000	yes		NaN	0.0148s
			5	4	1.35×10^{-8}	2/1000	yes		NaN	0.0143s
			Avg (successes)	4	1.82×10^{-8}	2.0/1000	6/6		NaN	0.0145s

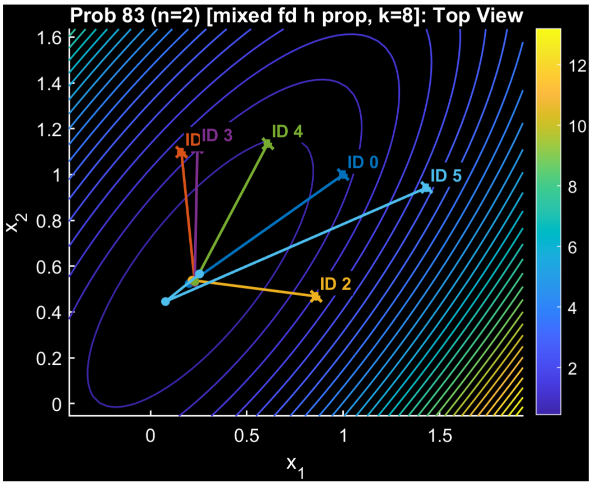
Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	mixed fd h prop	$\overline{x}$	4	1.42×10^{-8}	1/1000	yes		NaN	0.0160s
			1	4	4.58×10^{-8}	2/1000	yes		NaN	0.0366s
			2	4	3.89×10^{-8}	2/1000	yes		NaN	0.0343s
			3	4	2.67×10^{-8}	2/1000	yes		NaN	0.0333s
			4	4	3.20×10^{-8}	2/1000	yes		NaN	0.0356s
			5	4	3.15×10^{-8}	2/1000	yes		NaN	0.0337s
			Avg (successes)	4	3.15×10^{-8}	1.8/1000	6/6		NaN	0.0316s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	mixed fd h prop	$\overline{x}$	4	9.67×10^{-9}	1/100	yes		NaN	0.1614s
			1	4	1.68×10^{-7}	1/100	yes		NaN	0.1675s
			2	4	1.68×10^{-7}	1/100	yes		NaN	0.1637s
			3	4	1.68×10^{-7}	1/100	yes		NaN	0.1637s
			4	4	1.68×10^{-7}	1/100	yes		NaN	0.1689s
			5	4	1.68×10^{-7}	1/100	yes		NaN	0.1631s
			Avg (successes)	4	1.41×10^{-7}	1.0/100	6/6		NaN	0.1647s

Mixed FD (component-wise h), $k = 8$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	mixed fd h prop	$\overline{x}$	8	1.34×10^{-9}	3/1000	yes		1.84	0.0120s
			1	8	7.49×10^{-10}	3/1000	yes		1.97	0.0146s
			2	8	4.34×10^{-7}	3/1000	yes		1.94	0.0135s
			3	8	2.58×10^{-9}	3/1000	yes		1.96	0.0146s
			4	8	8.39×10^{-8}	3/1000	yes		2.00	0.0146s
			5	8	9.22×10^{-15}	4/1000	yes		1.90	0.0169s
			Avg (successes)	8	8.71×10^{-8}	3.2/1000	6/6		1.94	0.0144s



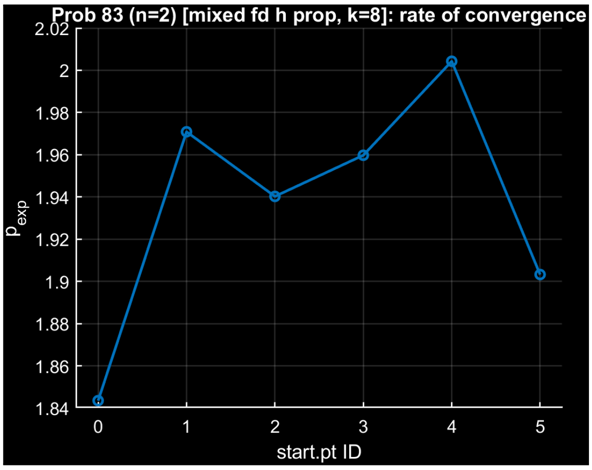


Figure 55: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	mixed fd h prop	$\overline{x}$	8	3.19×10^{-12}	2/1000	yes		NaN	0.0126s
			1	8	2.51×10^{-8}	2/1000	yes		NaN	0.0157s
			2	8	2.51×10^{-8}	2/1000	yes		NaN	0.0131s
			3	8	2.80×10^{-8}	2/1000	yes		NaN	0.0142s
			4	8	2.94×10^{-8}	2/1000	yes		NaN	0.0136s
			5	8	2.36×10^{-8}	2/1000	yes		NaN	0.0146s
			Avg (successes)	8	2.19×10^{-8}	2.0/1000	6/6		NaN	0.0140s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	mixed fd h prop	$\overline{x}$	8	1.14×10^{-8}	1/1000	yes		NaN	0.0151s
			1	8	4.64×10^{-8}	2/1000	yes		NaN	0.0378s
			2	8	3.94×10^{-8}	2/1000	yes		NaN	0.0336s
			3	8	2.75×10^{-8}	2/1000	yes		NaN	0.0372s
			4	8	9.42×10^{-10}	2/1000	yes		NaN	0.0331s
			5	8	6.53×10^{-10}	2/1000	yes		NaN	0.0336s
			Avg (successes)	8	2.11×10^{-8}	1.8/1000	6/6		NaN	0.0317s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	mixed fd h prop	$\overline{x}$	8	1.17×10^{-10}	1/100	yes		NaN	0.1458s
			1	8	2.48×10^{-7}	1/100	yes		NaN	0.1649s
			2	8	2.10×10^{-7}	1/100	yes		NaN	0.1728s
			3	8	7.93×10^{-10}	2/100	yes		NaN	0.3251s
			4	8	1.59×10^{-9}	2/100	yes		NaN	0.2990s
			5	8	2.11×10^{-7}	1/100	yes		NaN	0.1681s
			Avg (successes)	8	1.12×10^{-7}	1.3/100	6/6		NaN	0.2126s

Mixed FD (component-wise h), $k = 12$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	mixed fd h prop	$\overline{x}$	12	1.34×10^{-9}	3/1000	yes		1.84	0.0123s
			1	12	7.49×10^{-10}	3/1000	yes		1.97	0.0144s
			2	12	4.34×10^{-7}	3/1000	yes		1.94	0.0159s
			3	12	2.58×10^{-9}	3/1000	yes		1.96	0.0144s
			4	12	8.39×10^{-8}	3/1000	yes		2.00	0.0141s
			5	12	9.37×10^{-15}	4/1000	yes		1.90	0.0180s
			Avg (successes)	12	8.71×10^{-8}	3.2/1000	6/6		1.94	0.0148s

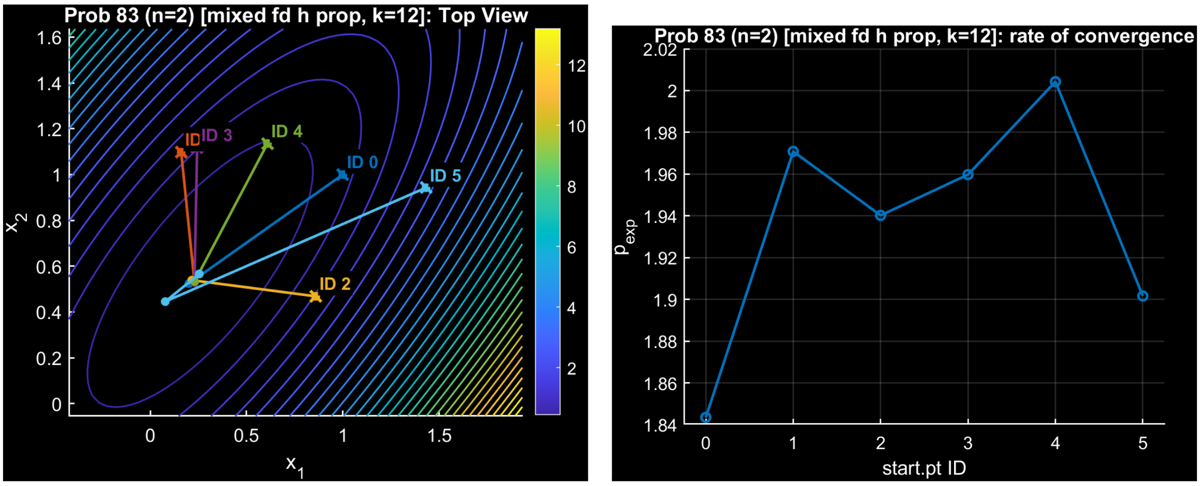


Figure 56: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	mixed fd h prop	$\overline{x}$	12	3.02×10^{-10}	2/1000	yes		NaN	0.0112s
			1	12	2.92×10^{-8}	2/1000	yes		NaN	0.0144s
			2	12	2.98×10^{-8}	2/1000	yes		NaN	0.0149s
			3	12	3.18×10^{-8}	2/1000	yes		NaN	0.0176s
			4	12	3.36×10^{-8}	2/1000	yes		NaN	0.0147s
			5	12	3.01×10^{-8}	2/1000	yes		NaN	0.0144s
			Avg (successes)	12	2.58×10^{-8}	2.0/1000	6/6		NaN	0.0146s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	mixed fd h prop	$\overline{x}$	12	1.13×10^{-8}	1/1000	yes		NaN	0.0153s
			1	12	8.14×10^{-8}	2/1000	yes		NaN	0.0359s
			2	12	8.88×10^{-9}	2/1000	yes		NaN	0.0311s
			3	12	3.68×10^{-8}	2/1000	yes		NaN	0.0357s
			4	12	1.70×10^{-8}	2/1000	yes		NaN	0.0317s
			5	12	1.23×10^{-7}	2/1000	yes		NaN	0.0334s
			Avg (successes)	12	4.65×10^{-8}	1.8/1000	6/6		NaN	0.0305s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	mixed fd h prop	$\overline{x}$	12	3.23×10^{-10}	1/100	yes		NaN	0.1602s
			1	12	7.28×10^{-8}	2/100	yes		NaN	0.4280s
			2	12	2.66×10^{-8}	2/100	yes		NaN	0.3932s
			3	12	1.53×10^{-8}	2/100	yes		NaN	0.4486s
			4	12	1.64×10^{-8}	2/100	yes		NaN	0.5788s
			5	12	8.69×10^{-8}	2/100	yes		NaN	0.4401s
			Avg (successes)	12	3.64×10^{-8}	1.8/100	6/6		NaN	0.4081s

9.2.2 Truncated Newton

Mixed FD (component-wise h), $k = 4$

Case $n = 2$

Case $n > 2$

Mixed FD (component-wise h), $k = 8$

Case $n = 2$

Case $n > 2$

Mixed FD (component-wise h), $k = 12$

Case $n = 2$

Case $n > 2$

10 Results - Full FD (constant h)

10.1 Problem 31

10.1.1 Modified Newton

Full FD (constant h), $k = 4$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	full fd	$\overline{x}$	4	1.73×10^{-7}	5/1000	yes		1.93	0.0066s
			1	4	1.78×10^{-7}	6/1000	yes		1.93	0.0008s
			2	4	7.12×10^{-9}	6/1000	yes		1.78	0.0009s
			3	4	3.99×10^{-9}	6/1000	yes		1.74	0.0007s
			4	4	1.95×10^{-9}	5/1000	yes		1.67	0.0007s
			5	4	1.20×10^{-10}	8/1000	yes		1.38	0.0011s
			Avg (successes)	4	6.06×10^{-8}	6.0/1000	6/6		1.74	0.0018s

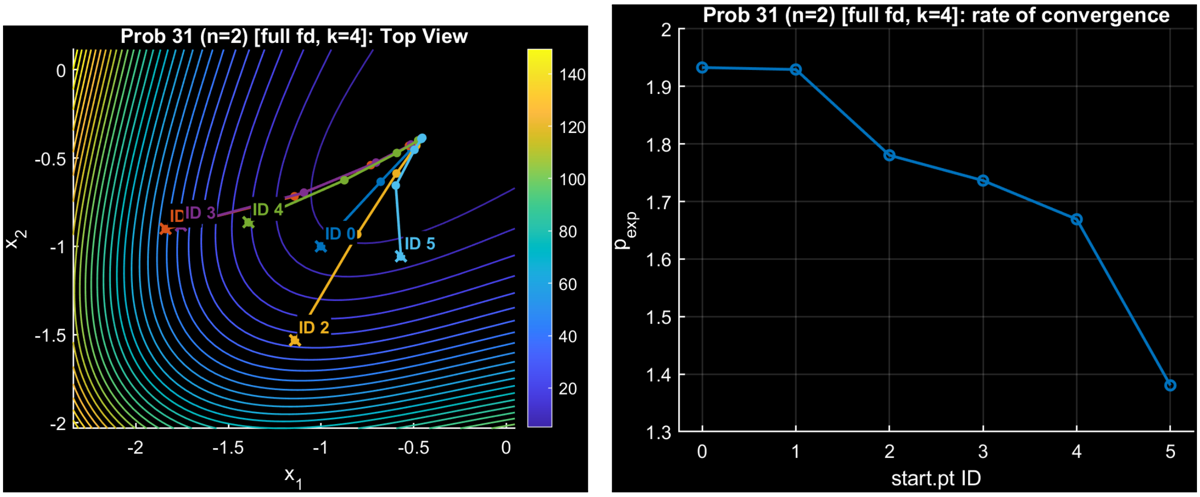


Figure 57: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	full fd	$\overline{x}$	4	3.39×10^{-8}	5/1000	yes		1.65	0.0800s
			1	4	8.94×10^{-11}	10/1000	yes		1.33	0.1612s
			2	4	3.95×10^{-7}	9/1000	yes		1.91	0.1535s
			3	4	4.01×10^{-11}	9/1000	yes		1.28	0.1548s
			4	4	9.29×10^{-7}	8/1000	yes		1.90	0.1424s
			5	4	1.48×10^{-10}	9/1000	yes		1.39	0.1430s
			Avg (successes)	4	2.26×10^{-7}	8.3/1000	6/6		1.58	0.1392s

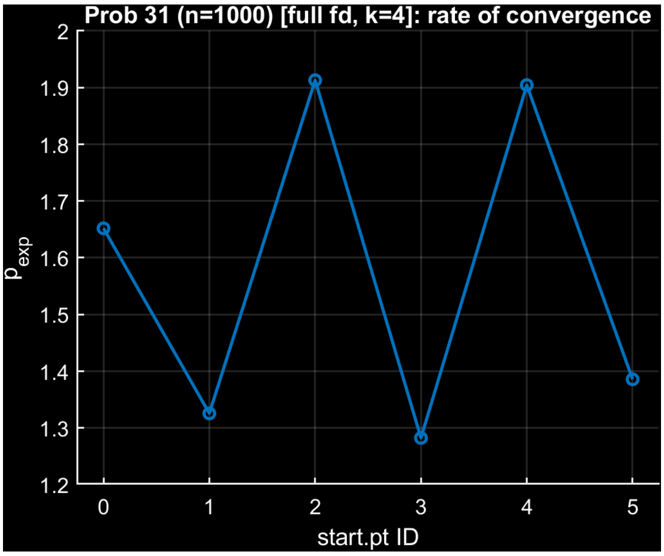


Figure 58: Problem 31, Modified Newton, full FD (constant h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	full fd	$\overline{x}$	4	3.66×10^{-8}	5/1000	yes		1.46	0.9511s
			1	4	4.34×10^{-10}	10/1000	yes		1.46	2.50s
			2	4	9.71×10^{-9}	9/1000	yes		1.75	1.56s
			3	4	2.40×10^{-7}	10/1000	yes		1.94	1.44s
			4	4	5.41×10^{-7}	8/1000	yes		1.87	1.13s
			5	4	6.67×10^{-7}	10/1000	yes		1.93	1.41s
			Avg (successes)	4	2.49×10^{-7}	8.7/1000	6/6		1.74	1.50s

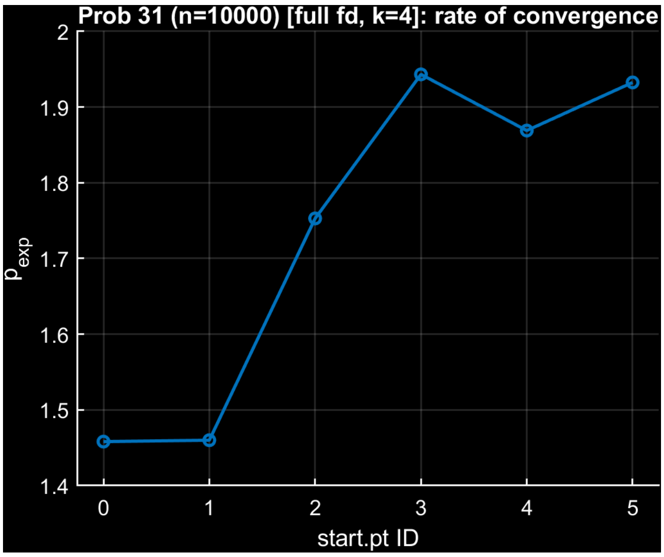


Figure 59: Problem 31, Modified Newton, full FD (constant h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	full fd	$\overline{x}$	4	5.69×10^{-8}	5/100	yes		1.41	7.05s
			1	4	2.22×10^{-7}	8/100	yes		1.85	11.40s
			2	4	5.06×10^{-8}	12/100	yes		1.90	16.74s
			3	4	7.71×10^{-10}	10/100	yes		1.54	14.08s
			4	4	5.39×10^{-11}	11/100	yes		1.24	15.54s
			5	4	1.06×10^{-7}	12/100	yes		1.90	16.89s
			Avg (successes)	4	7.26×10^{-8}	9.7/100	6/6		1.64	13.62s

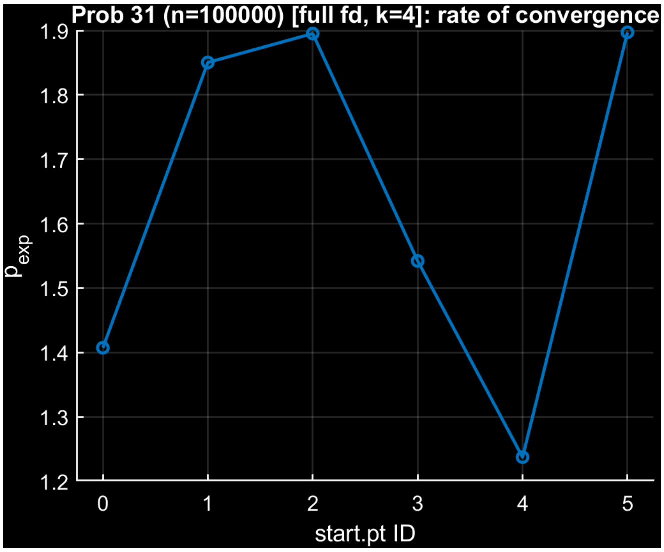


Figure 60: Problem 31, Modified Newton, full FD (constant h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Full FD (constant h), $k = 8$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	full fd	$\overline{x}$	8	1.32×10^{-7}	5/1000	yes		1.99	0.0006s
			1	8	1.40×10^{-7}	6/1000	yes		1.98	0.0010s
			2	8	2.16×10^{-9}	6/1000	yes		1.99	0.0009s
			3	8	8.95×10^{-10}	6/1000	yes		1.99	0.0008s
			4	8	2.68×10^{-10}	5/1000	yes		1.99	0.0006s
			5	8	1.06×10^{-12}	8/1000	yes		2.00	0.0008s
			Avg (successes)	8	4.59×10^{-8}	6.0/1000	6/6		1.99	0.0008s

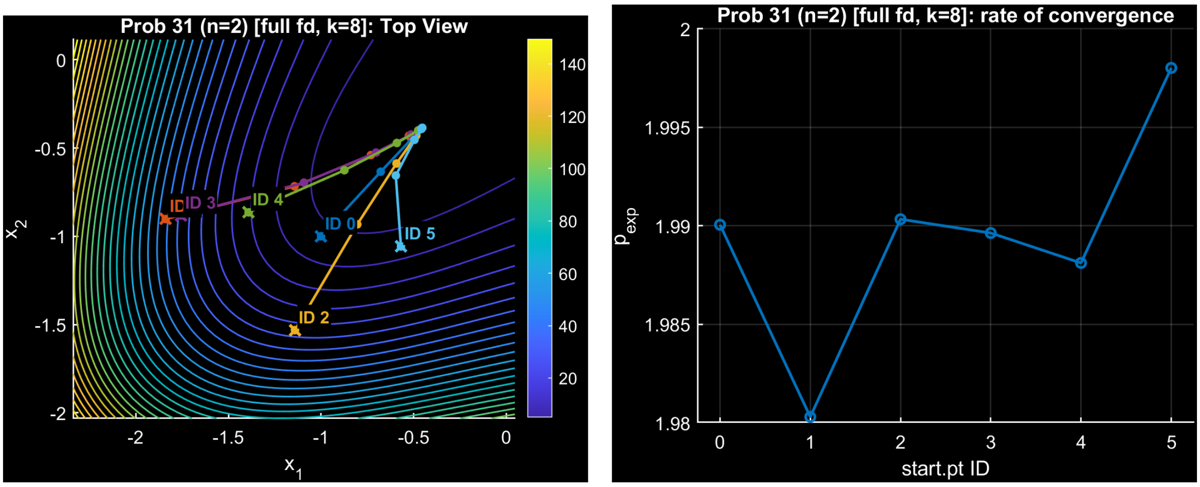


Figure 61: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	full fd	$\overline{x}$	8	1.79×10^{-8}	5/1000	yes		1.76	0.0867s
			1	8	4.57×10^{-13}	10/1000	yes		1.99	0.1477s
			2	8	3.30×10^{-7}	9/1000	yes		1.95	0.1456s
			3	8	1.33×10^{-13}	9/1000	yes		1.96	0.1354s
			4	8	8.28×10^{-7}	8/1000	yes		1.93	0.1241s
			5	8	1.43×10^{-12}	9/1000	yes		1.99	0.1547s
			Avg (successes)	8	1.96×10^{-7}	8.3/1000	6/6		1.93	0.1324s

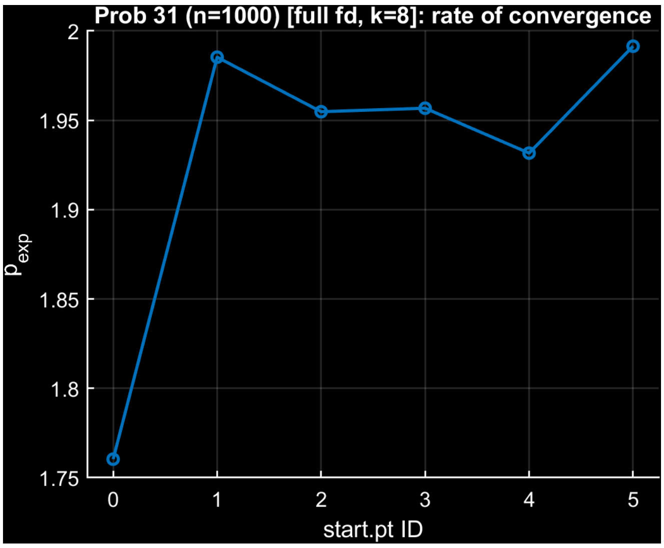


Figure 62: Problem 31, Modified Newton, full FD (constant h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	full fd	$\overline{x}$	8	1.79×10^{-8}	5/1000	yes		1.56	0.7164s
			1	8	9.41×10^{-12}	10/1000	yes		2.00	1.40s
			2	8	3.12×10^{-9}	9/1000	yes		1.95	1.27s
			3	8	2.15×10^{-7}	10/1000	yes		1.99	1.39s
			4	8	4.62×10^{-7}	8/1000	yes		1.90	1.13s
			5	8	5.96×10^{-7}	10/1000	yes		1.96	1.41s
			Avg (successes)	8	2.16×10^{-7}	8.7/1000	6/6		1.89	1.22s

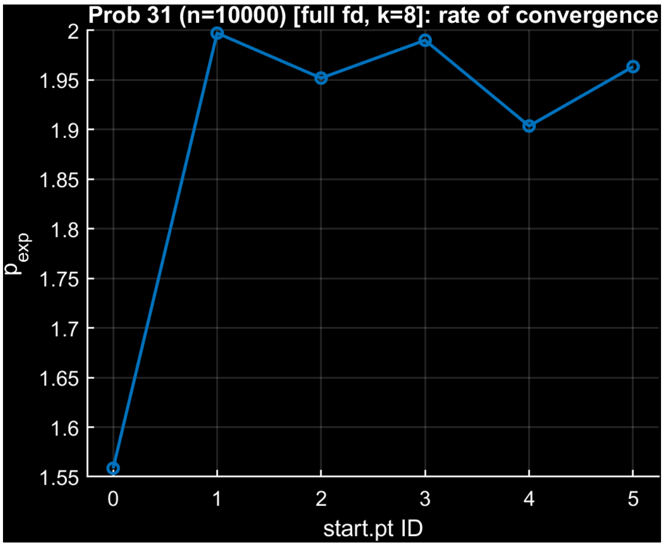


Figure 63: Problem 31, Modified Newton, full FD (constant h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	full fd	$\overline{x}$	8	1.79×10^{-8}	5/100	yes		1.56	7.02s
			1	8	1.73×10^{-7}	8/100	yes		1.90	11.36s
			2	8	3.80×10^{-8}	12/100	yes		1.99	16.69s
			3	8	3.45×10^{-11}	10/100	yes		2.00	13.91s
			4	8	7.23×10^{-13}	11/100	yes		1.71	15.46s
			5	8	1.23×10^{-9}	10/100	yes		2.00	14.19s
			Avg (successes)	8	3.84×10^{-8}	9.3/100	6/6		1.86	13.11s

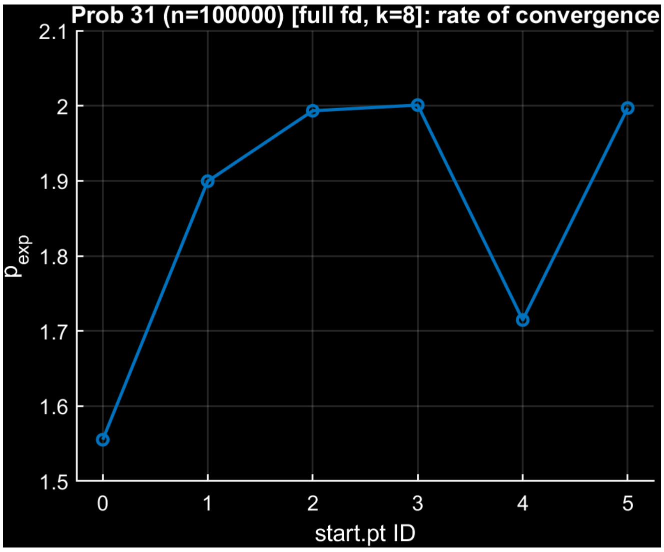
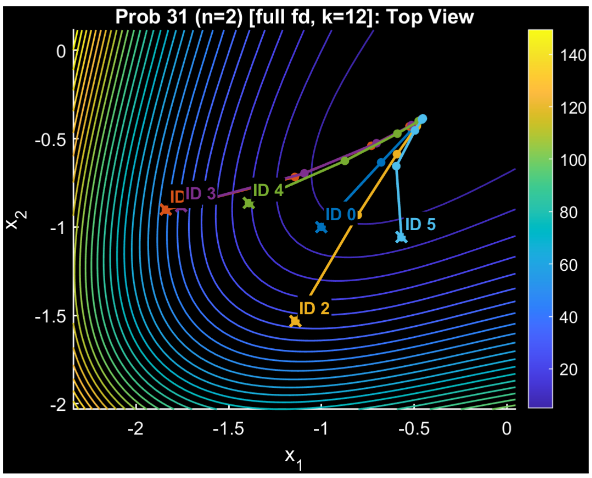


Figure 64: Problem 31, Modified Newton, full FD (constant h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Full FD (constant h), $k = 12$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	full fd	$\overline{x}$	12	1.83×10^{-7}	5/1000	yes		1.92	0.0010s
			1	12	1.92×10^{-7}	6/1000	yes		1.91	0.0008s
			2	12	4.71×10^{-9}	6/1000	yes		1.86	0.0011s
			3	12	3.89×10^{-9}	6/1000	yes		1.74	0.0007s
			4	12	5.00×10^{-10}	5/1000	yes		1.89	0.0006s
			5	12	7.56×10^{-11}	8/1000	yes		1.39	0.0010s
			Avg (successes)	12	6.41×10^{-8}	6.0/1000	6/6		1.78	0.0009s



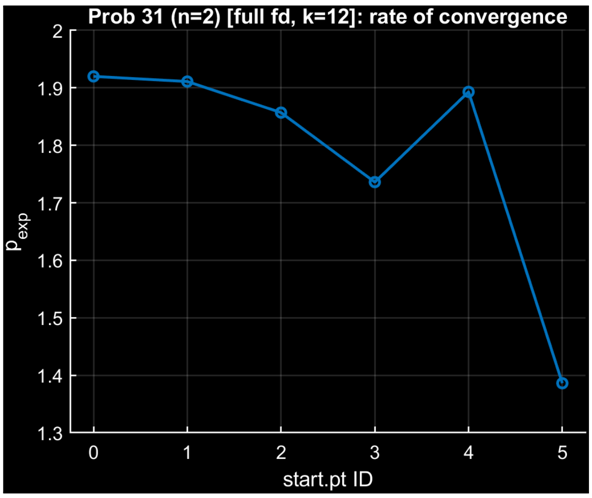


Figure 65: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	full fd	$\bar{x}$	12	3.10×10^{-8}	5/1000	yes		1.67	0.0746s
			1	12	1.40×10^{-10}	10/1000	yes		1.25	0.1717s
			2	12	7.21×10^{-7}	9/1000	yes		1.92	0.1393s
			3	12	2.62×10^{-11}	9/1000	yes		1.29	0.1529s
			4	12	7.52×10^{-7}	8/1000	yes		1.96	0.1346s
			5	12	2.03×10^{-10}	9/1000	yes		1.36	0.1941s
			Avg (successes)	12	2.51×10^{-7}	8.3/1000	6/6		1.57	0.1446s

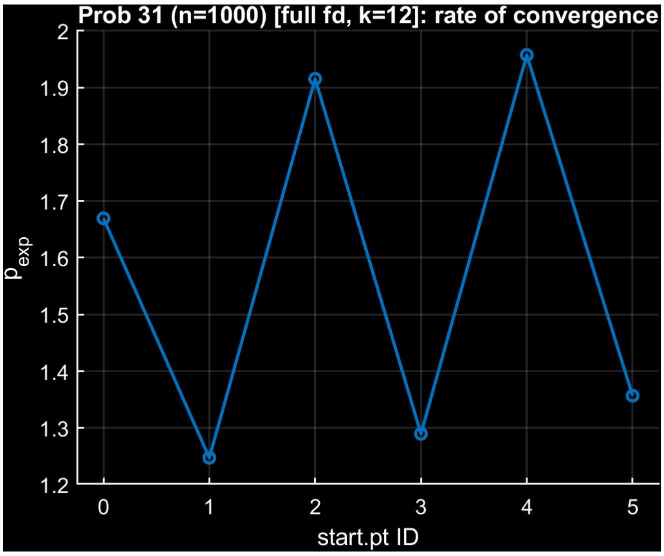


Figure 66: Problem 31, Modified Newton, full FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	full fd	$\bar{x}$	12	3.12×10^{-8}	5/1000	yes		1.49	0.7128s
			1	12	4.94×10^{-10}	10/1000	yes		1.47	1.41s
			2	12	1.57×10^{-8}	9/1000	yes		1.67	1.28s
			3	12	4.10×10^{-7}	10/1000	yes		1.91	1.40s
			4	12	6.31×10^{-7}	8/1000	yes		1.85	1.13s
			5	12	7.05×10^{-7}	10/1000	yes		1.97	1.41s
			Avg (successes)	12	2.99×10^{-7}	8.7/1000	6/6		1.73	1.22s

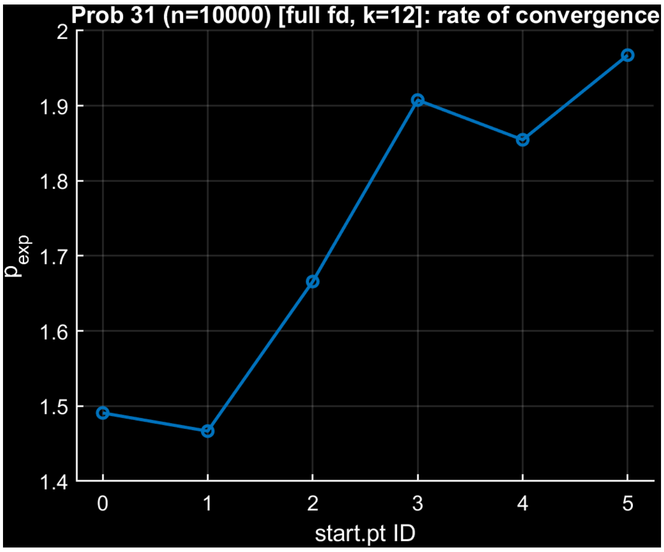


Figure 67: Problem 31, Modified Newton, full FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	full fd	$\overline{x}$	12	3.23×10^{-8}	5/100	yes		1.51	7.23s
			1	12	1.98×10^{-7}	8/100	yes		1.85	13.42s
			2	12	1.70×10^{-8}	11/100	yes		1.73	15.48s
			3	12	1.13×10^{-9}	10/100	yes		1.48	13.58s
			4	12	3.92×10^{-7}	10/100	yes		1.99	13.65s
			5	12	2.15×10^{-10}	11/100	yes		1.09	14.92s
			Avg (successes)	12	1.07×10^{-7}	9.2/100	6/6		1.61	13.05s

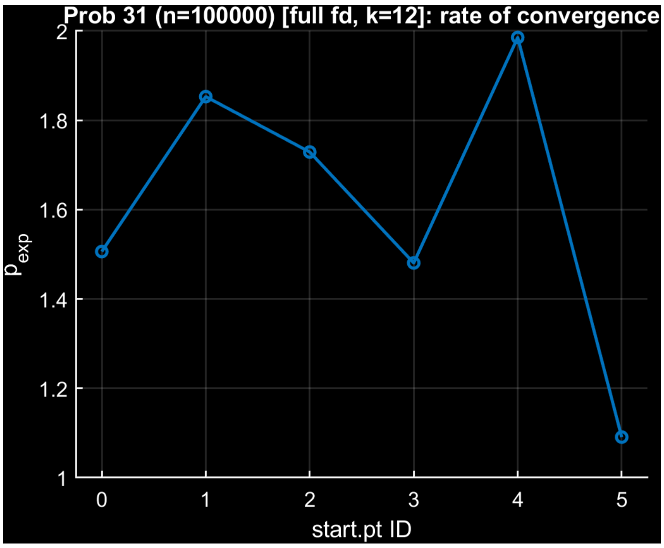


Figure 68: Problem 31, Modified Newton, full FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

10.1.2 Truncated Newton

Full FD (constant h), $k = 4$

Case $n = 2$

Case $n > 2$

Full FD (constant h), $k = 8$

Case $n = 2$

Case $n > 2$

Full FD (constant h), $k = 12$

Case $n = 2$

Case $n > 2$

10.2 Problem 83

10.2.1 Modified Newton

Full FD (constant h), $k = 4$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	full fd	$\overline{x}$	4	9.49×10^{-10}	3/1000	yes		1.90	0.0179s
			1	4	4.70×10^{-10}	3/1000	yes		2.04	0.0007s
			2	4	4.27×10^{-7}	3/1000	yes		1.94	0.0005s
			3	4	2.02×10^{-9}	3/1000	yes		2.00	0.0006s
			4	4	8.05×10^{-8}	3/1000	yes		2.01	0.0005s
			5	4	2.11×10^{-12}	4/1000	yes		1.35	0.0008s
			Avg (successes)	4	8.51×10^{-8}	3.2/1000	6/6		1.87	0.0035s

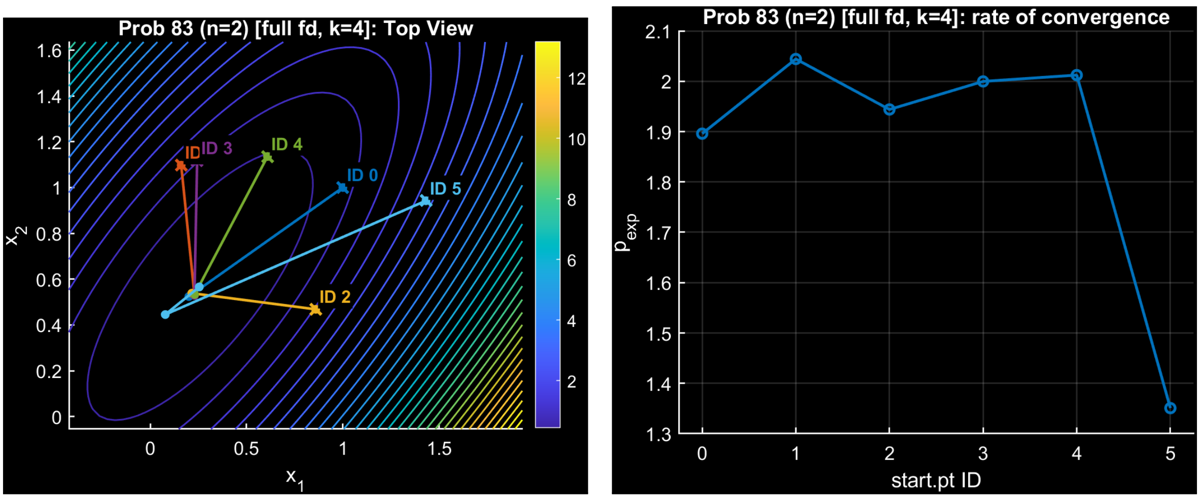


Figure 69: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	full fd	$\overline{x}$	4	5.29×10^{-13}	2/1000	yes		NaN	0.0291s
			1	4	2.51×10^{-8}	2/1000	yes		NaN	0.0339s
			2	4	2.51×10^{-8}	2/1000	yes		NaN	0.0262s
			3	4	2.80×10^{-8}	2/1000	yes		NaN	0.0263s
			4	4	2.94×10^{-8}	2/1000	yes		NaN	0.0253s
			5	4	2.36×10^{-8}	2/1000	yes		NaN	0.0250s
			Avg (successes)	4	2.19×10^{-8}	2.0/1000	6/6		NaN	0.0276s

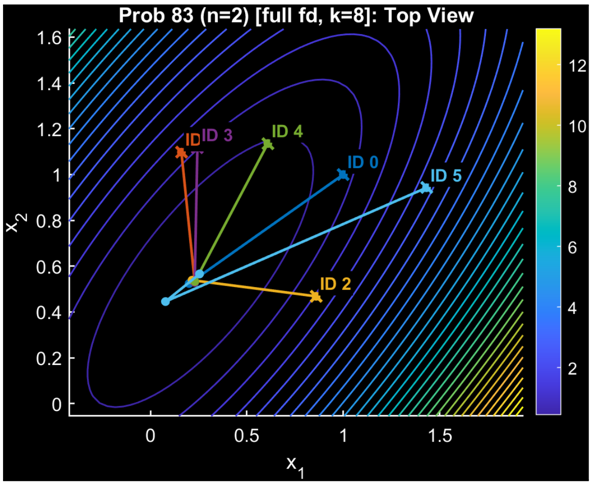
Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	full fd	$\overline{x}$	4	1.14×10^{-8}	1/1000	yes		NaN	0.0971s
			1	4	6.73×10^{-10}	2/1000	yes		NaN	0.1870s
			2	4	5.35×10^{-10}	2/1000	yes		NaN	0.1873s
			3	4	1.56×10^{-10}	2/1000	yes		NaN	0.1873s
			4	4	4.33×10^{-10}	2/1000	yes		NaN	0.1883s
			5	4	2.88×10^{-10}	2/1000	yes		NaN	0.1924s
			Avg (successes)	4	2.25×10^{-9}	1.8/1000	6/6		NaN	0.1732s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	full fd	$\overline{x}$	4	9.67×10^{-9}	1/100	yes		NaN	1.03s
			1	4	1.68×10^{-7}	1/100	yes		NaN	0.9743s
			2	4	1.68×10^{-7}	1/100	yes		NaN	0.9779s
			3	4	1.68×10^{-7}	1/100	yes		NaN	1.01s
			4	4	1.68×10^{-7}	1/100	yes		NaN	0.9861s
			5	4	1.68×10^{-7}	1/100	yes		NaN	0.9875s
			Avg (successes)	4	1.41×10^{-7}	1.0/100	6/6		NaN	0.9943s

Full FD (constant h), $k = 8$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	full fd	$\overline{x}$	8	1.34×10^{-9}	3/1000	yes		1.84	0.0012s
			1	8	7.49×10^{-10}	3/1000	yes		1.97	0.0007s
			2	8	4.34×10^{-7}	3/1000	yes		1.94	0.0007s
			3	8	2.58×10^{-9}	3/1000	yes		1.96	0.0005s
			4	8	8.39×10^{-8}	3/1000	yes		2.00	0.0007s
			5	8	1.02×10^{-14}	4/1000	yes		1.89	0.0006s
			Avg (successes)	8	8.71×10^{-8}	3.2/1000	6/6		1.94	0.0008s



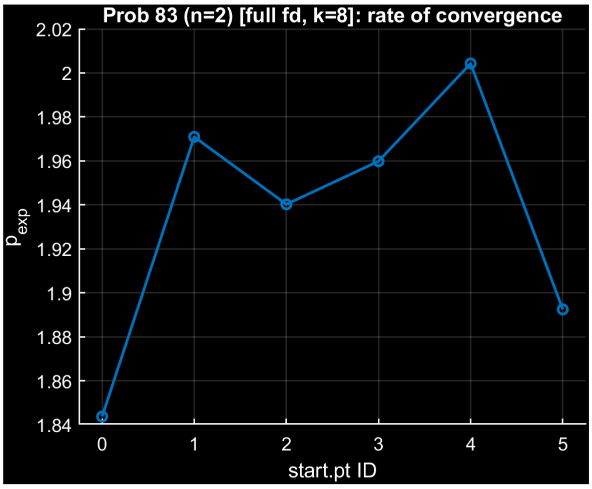


Figure 70: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	full fd	$\overline{x}$	8	1.36×10^{-12}	2/1000	yes		NaN	0.0223s
			1	8	2.09×10^{-8}	2/1000	yes		NaN	0.0226s
			2	8	3.77×10^{-8}	2/1000	yes		NaN	0.0219s
			3	8	2.61×10^{-8}	2/1000	yes		NaN	0.0215s
			4	8	2.70×10^{-8}	2/1000	yes		NaN	0.0218s
			5	8	2.50×10^{-8}	2/1000	yes		NaN	0.0205s
			Avg (successes)	8	2.28×10^{-8}	2.0/1000	6/6		NaN	0.0218s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	full fd	$\overline{x}$	8	3.43×10^{-8}	1/1000	yes		NaN	0.0981s
			1	8	1.32×10^{-9}	2/1000	yes		NaN	0.1870s
			2	8	8.09×10^{-11}	2/1000	yes		NaN	0.1894s
			3	8	1.15×10^{-10}	2/1000	yes		NaN	0.1925s
			4	8	1.96×10^{-8}	2/1000	yes		NaN	0.1961s
			5	8	9.33×10^{-11}	2/1000	yes		NaN	0.2042s
			Avg (successes)	8	9.26×10^{-9}	1.8/1000	6/6		NaN	0.1779s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	full fd	$\overline{x}$	8	2.96×10^{-8}	1/100	yes		NaN	0.9861s
			1	8	4.21×10^{-8}	2/100	yes		NaN	1.94s
			2	8	4.93×10^{-8}	2/100	yes		NaN	1.96s
			3	8	5.21×10^{-8}	2/100	yes		NaN	1.95s
			4	8	1.08×10^{-7}	2/100	yes		NaN	1.94s
			5	8	6.41×10^{-8}	2/100	yes		NaN	1.96s
			Avg (successes)	8	5.75×10^{-8}	1.8/100	6/6		NaN	1.79s

Full FD (constant h), $k = 12$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	full fd	$\overline{x}$	12	6.53×10^{-9}	3/1000	yes		1.60	0.0012s
			1	12	9.82×10^{-9}	3/1000	yes		1.56	0.0007s
			2	12	3.45×10^{-7}	3/1000	yes		1.99	0.0006s
			3	12	2.04×10^{-8}	3/1000	yes		1.62	0.0006s
			4	12	1.04×10^{-7}	3/1000	yes		1.96	0.0005s
			5	12	4.98×10^{-7}	3/1000	yes		2.36	0.0005s
			Avg (successes)	12	1.64×10^{-7}	3.0/1000	6/6		1.85	0.0007s

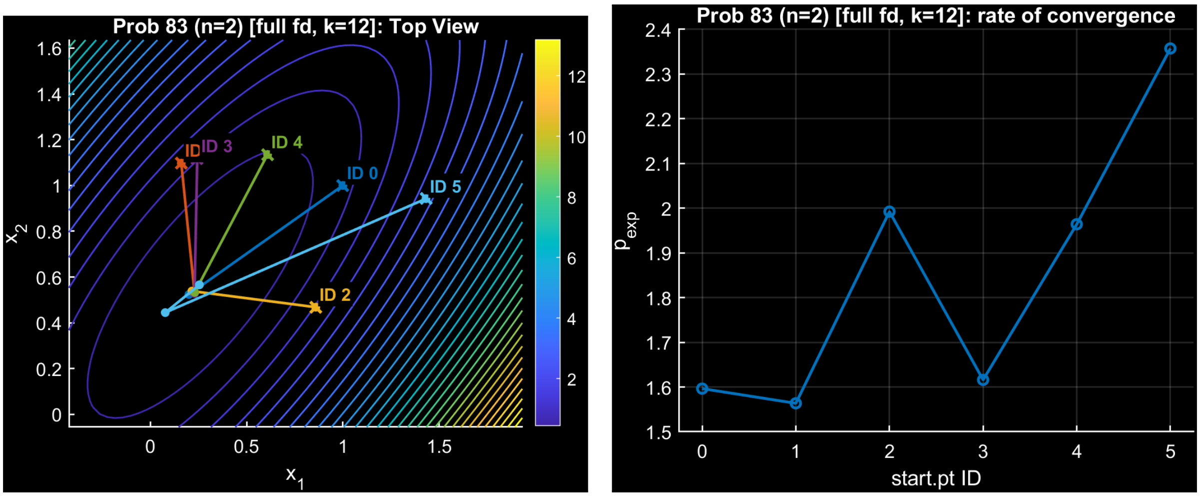


Figure 71: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	full fd	$\overline{x}$	12	4.77×10^{-7}	4/1000	yes		0.29	0.0407s
			1	12	3.91×10^{-7}	9/1000	yes		NaN	0.0901s
			2	12	1.16×10^{-7}	5/1000	yes		2.17	0.0510s
			3	12	1.52×10^{-7}	3/1000	yes		0.69	0.0304s
			4	12	1.47×10^{-8}	4/1000	yes		1.42	0.0395s
			5	12	2.86×10^{-7}	3/1000	yes		0.60	0.0304s
			Avg (successes)	12	2.40×10^{-7}	4.7/1000	6/6		1.03	0.0470s

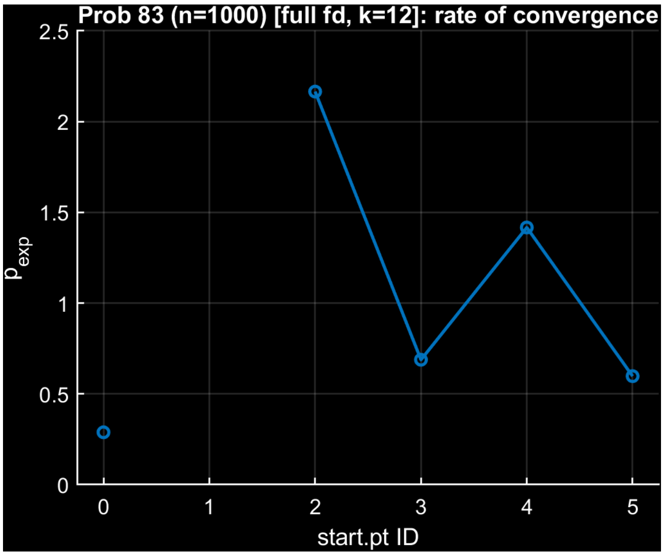


Figure 72: Problem 83, Truncated Newton, full FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	full fd	$\overline{x}$	12	3.26×10^{-7}	2/1000	yes		NaN	0.1895s
			1	12	1.26×10^{-8}	4/1000	yes		1.39	0.3876s
			2	12	1.47×10^{-7}	3/1000	yes		0.66	0.3044s
			3	12	5.91×10^{-7}	5/1000	yes		1.12	0.4783s
			4	12	1.26×10^{-7}	3/1000	yes		0.62	0.3019s
			5	12	1.23×10^{-7}	3/1000	yes		0.66	0.3028s
			Avg (successes)	12	2.21×10^{-7}	3.3/1000	6/6		0.89	0.3274s

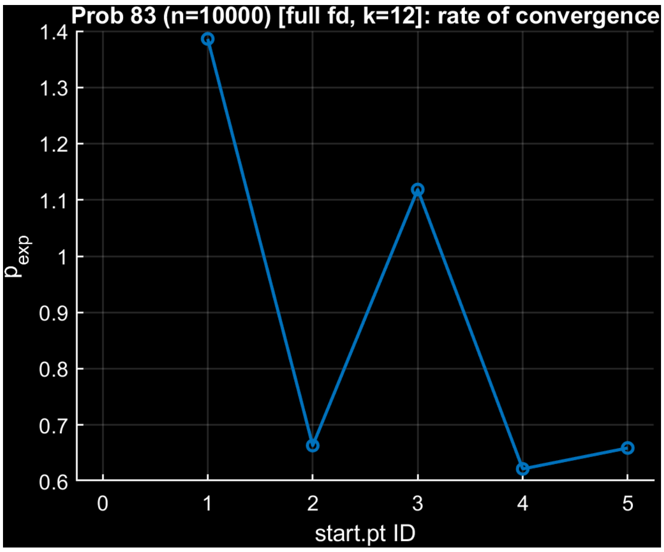


Figure 73: Problem 83, Truncated Newton, full FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	full fd	$\overline{x}$	12	8.31×10^{-7}	6/100	yes		0.90	5.83s
			1	12	3.14×10^{-7}	3/100	yes		0.68	3.28s
			2	12	1.10×10^{-7}	4/100	yes		0.58	4.23s
			3	12	1.83×10^{-8}	5/100	yes		1.00	5.20s
			4	12	3.23×10^{-7}	3/100	yes		0.67	3.25s
			5	12	7.10×10^{-7}	3/100	yes		0.59	3.28s
			Avg (successes)	12	3.84×10^{-7}	4.0/100	6/6		0.74	4.18s

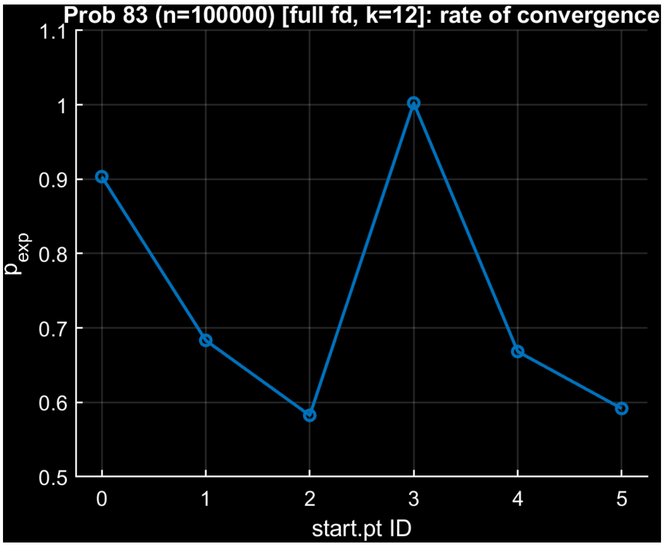


Figure 74: Problem 83, Truncated Newton, full FD (constant h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

10.2.2 Truncated Newton

11 Results - Full FD (component-wise h)

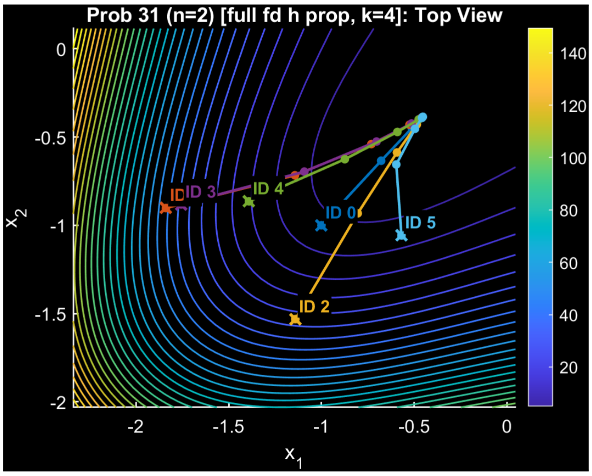
11.1 Problem 31

11.1.1 Modified Newton

Full FD (component-wise h), $k = 4$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	full fd h prop	$\bar{x}$	4	1.39×10^{-7}	5/1000	yes		1.98	0.0010s
			1	4	1.47×10^{-7}	6/1000	yes		1.97	0.0006s
			2	4	2.98×10^{-9}	6/1000	yes		1.93	0.0006s
			3	4	1.47×10^{-9}	6/1000	yes		1.90	0.0007s
			4	4	5.79×10^{-10}	5/1000	yes		1.86	0.0005s
			5	4	2.14×10^{-11}	8/1000	yes		1.60	0.0010s
			Avg (successes)	4	4.86×10^{-8}	6.0/1000	6/6		1.87	0.0008s



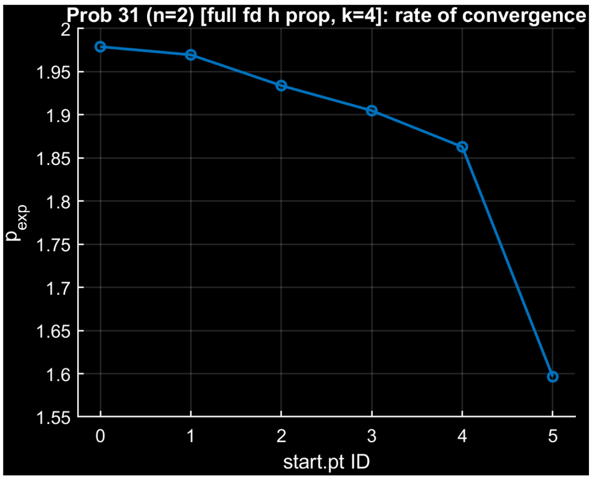


Figure 75: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	full fd h prop	$\bar{x}$	4	2.22×10^{-8}	5/1000	yes		1.72	0.0711s
			1	4	4.39×10^{-11}	10/1000	yes		1.41	0.1375s
			2	4	3.65×10^{-7}	9/1000	yes		1.93	0.1237s
			3	4	1.88×10^{-11}	9/1000	yes		1.36	0.1301s
			4	4	8.83×10^{-7}	8/1000	yes		1.92	0.1120s
			5	4	7.26×10^{-11}	9/1000	yes		1.47	0.1228s
			Avg (successes)	4	2.12×10^{-7}	8.3/1000	6/6		1.64	0.1162s

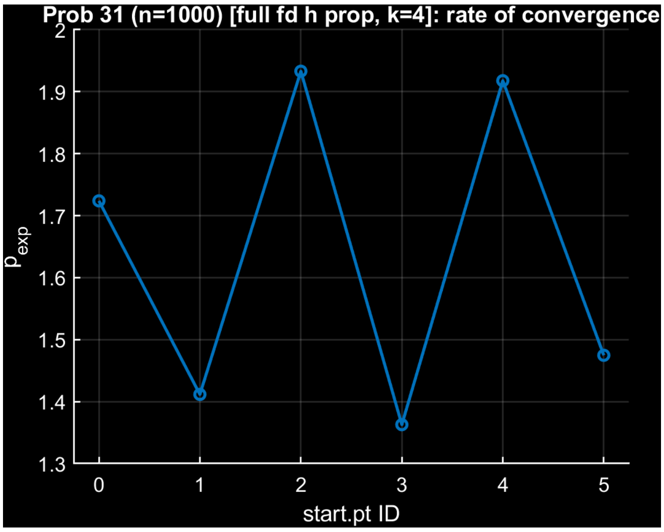


Figure 76: Problem 31, Modified Newton, full FD (component-wise h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	full fd h prop	$\overline{x}$	4	2.32×10^{-8}	5/1000	yes		1.52	0.6792s
			1	4	2.23×10^{-10}	10/1000	yes		1.56	1.39s
			2	4	6.40×10^{-9}	9/1000	yes		1.83	1.25s
			3	4	2.39×10^{-7}	10/1000	yes		1.97	1.37s
			4	4	5.03×10^{-7}	8/1000	yes		1.89	1.11s
			5	4	6.41×10^{-7}	10/1000	yes		1.95	1.36s
			Avg (successes)	4	2.35×10^{-7}	8.7/1000	6/6		1.78	1.19s

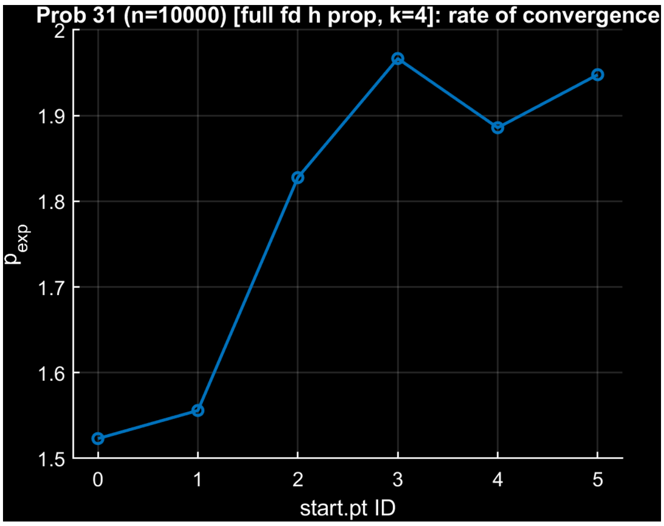


Figure 77: Problem 31, Modified Newton, full FD (component-wise h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	full fd h prop	$\bar{x}$	4	3.16×10^{-8}	5/100	yes		1.48	6.76s
			1	4	1.97×10^{-7}	8/100	yes		1.87	10.80s
			2	4	4.87×10^{-8}	12/100	yes		1.94	15.98s
			3	4	3.95×10^{-10}	10/100	yes		1.64	13.36s
			4	4	2.48×10^{-11}	11/100	yes		1.31	15.00s
			5	4	3.18×10^{-9}	10/100	yes		1.83	13.57s
			Avg (successes)	4	4.68×10^{-8}	9.3/100	6/6		1.68	12.58s

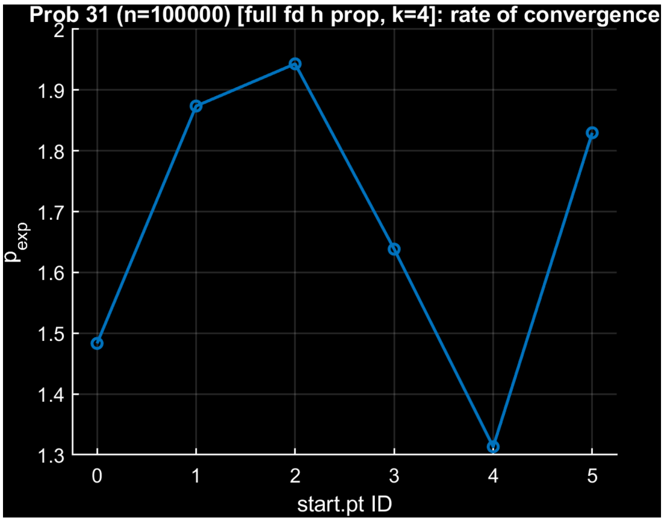


Figure 78: Problem 31, Modified Newton, full FD (component-wise h, $k = 4$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Full FD (component-wise h), $k = 8$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	full fd h prop	$\bar{x}$	8	1.32×10^{-7}	5/1000	yes		1.99	0.0004s
			1	8	1.40×10^{-7}	6/1000	yes		1.98	0.0007s
			2	8	2.16×10^{-9}	6/1000	yes		1.99	0.0006s
			3	8	8.94×10^{-10}	6/1000	yes		1.99	0.0006s
			4	8	2.68×10^{-10}	5/1000	yes		1.99	0.0032s
			5	8	1.06×10^{-12}	8/1000	yes		2.00	0.0008s
			Avg (successes)	8	4.59×10^{-8}	6.0/1000	6/6		1.99	0.0011s

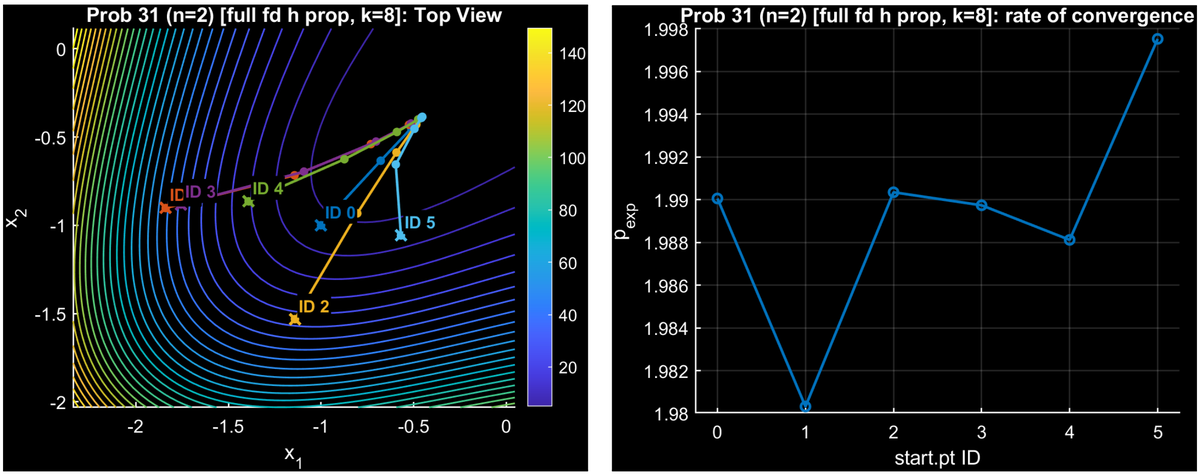


Figure 79: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	full fd h prop	$\overline{x}$	8	1.79×10^{-8}	5/1000	yes		1.76	0.0700s
			1	8	9.30×10^{-14}	10/1000	yes		1.97	0.1400s
			2	8	3.37×10^{-7}	9/1000	yes		1.95	0.1264s
			3	8	2.78×10^{-10}	8/1000	yes		1.98	0.1105s
			4	8	8.28×10^{-7}	8/1000	yes		1.93	0.1092s
			5	8	1.26×10^{-12}	9/1000	yes		2.00	0.1227s
			Avg (successes)	8	1.97×10^{-7}	8.2/1000	6/6		1.93	0.1132s

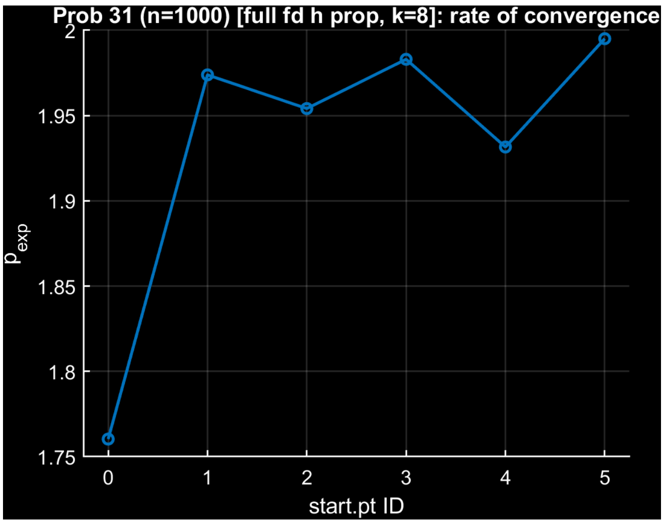


Figure 80: Problem 31, Modified Newton, full FD (component-wise h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	full fd h prop	$\overline{x}$	8	1.79×10^{-8}	5/1000	yes		1.56	0.7011s
			1	8	2.90×10^{-12}	10/1000	yes		2.00	1.39s
			2	8	2.54×10^{-9}	9/1000	yes		1.95	1.25s
			3	8	6.56×10^{-8}	9/1000	yes		1.99	1.23s
			4	8	4.85×10^{-7}	8/1000	yes		1.90	1.11s
			5	8	7.03×10^{-10}	9/1000	yes		2.00	1.23s
			Avg (successes)	8	9.53×10^{-8}	8.3/1000	6/6		1.90	1.15s

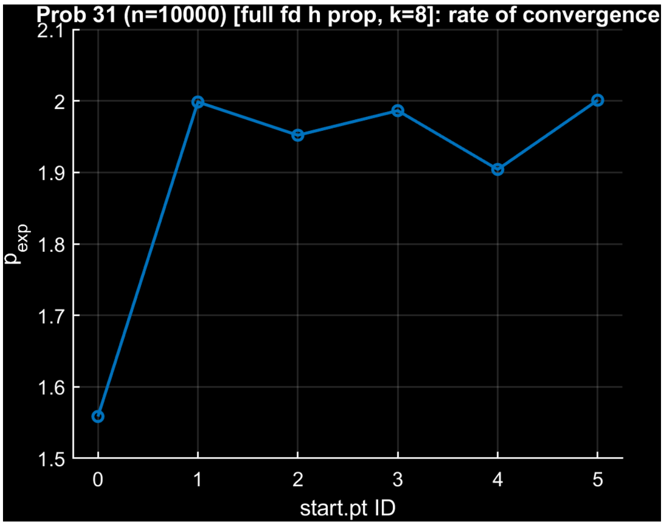


Figure 81: Problem 31, Modified Newton, full FD (component-wise h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	full fd h prop	$\overline{x}$	8	1.79×10^{-8}	5/100	yes		1.56	6.90s
			1	8	6.04×10^{-8}	13/100	yes		2.00	17.45s
			2	8	1.20×10^{-8}	9/100	yes		1.96	12.35s
			3	8	2.14×10^{-11}	11/100	yes		2.00	14.75s
			4	8	4.89×10^{-7}	11/100	yes		1.99	14.86s
			5	8	4.81×10^{-12}	11/100	yes		1.99	14.69s
			Avg (successes)	8	9.65×10^{-8}	10.0/100	6/6		1.92	13.50s

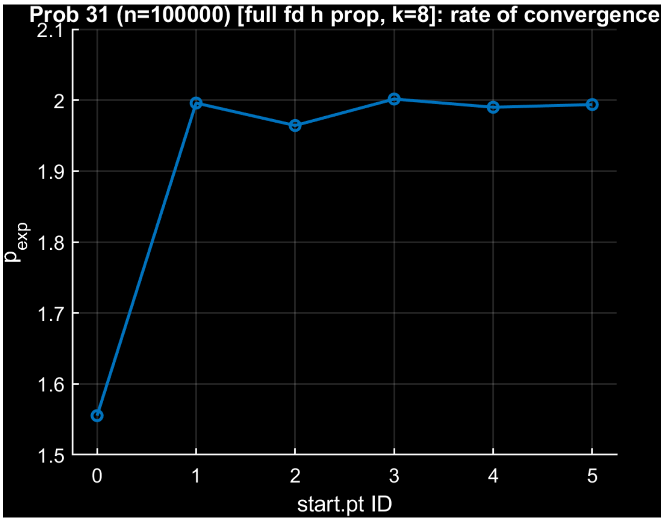


Figure 82: Problem 31, Modified Newton, full FD (component-wise h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Full FD (component-wise h), $k = 12$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	2	full fd h prop	$\overline{x}$	12	1.20×10^{-7}	5/1000	yes		2.01	0.0006s
			1	12	1.16×10^{-7}	6/1000	yes		2.02	0.0007s
			2	12	7.52×10^{-9}	6/1000	yes		1.80	0.0006s
			3	12	7.35×10^{-9}	6/1000	yes		1.63	0.0006s
			4	12	6.63×10^{-10}	5/1000	yes		1.85	0.0005s
			5	12	4.92×10^{-8}	7/1000	yes		1.75	0.0007s
			Avg (successes)	12	5.03×10^{-8}	5.8/1000	6/6		1.84	0.0006s

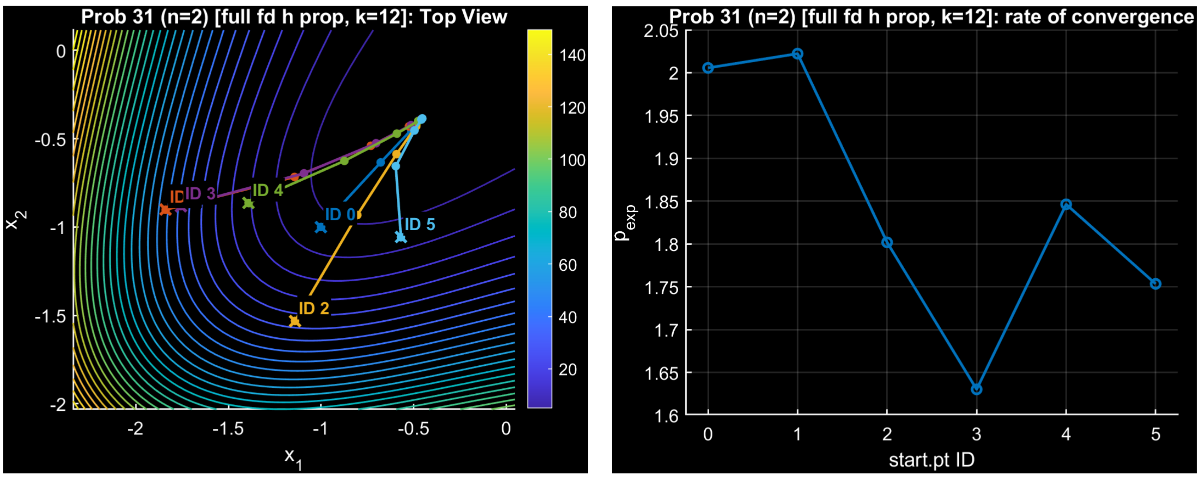


Figure 83: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	1000	full fd h prop	$\overline{x}$	12	3.11×10^{-8}	5/1000	yes		1.66	0.0703s
			1	12	1.12×10^{-9}	9/1000	yes		1.40	0.1218s
			2	12	6.73×10^{-11}	10/1000	yes		1.33	0.1360s
			3	12	2.73×10^{-10}	11/1000	yes		1.26	0.1502s
			4	12	8.41×10^{-9}	9/1000	yes		1.72	0.1226s
			5	12	2.21×10^{-8}	9/1000	yes		1.73	0.1219s
			Avg (successes)	12	1.05×10^{-8}	8.8/1000	6/6		1.52	0.1205s

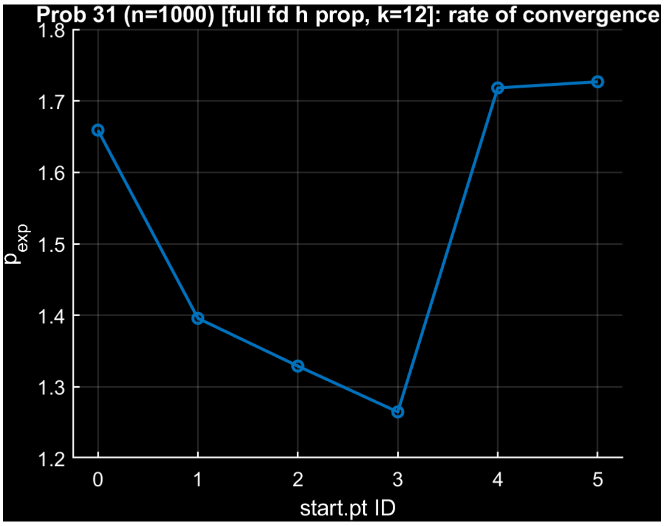


Figure 84: Problem 31, Modified Newton, full FD (component-wise h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	10000	full fd h prop	$\overline{x}$	12	3.57×10^{-8}	5/1000	yes		1.43	0.6985s
			1	12	3.46×10^{-7}	10/1000	yes		1.89	1.38s
			2	12	3.00×10^{-8}	16/1000	yes		1.83	2.16s
			3	12	3.25×10^{-9}	18/1000	yes		1.49	2.41s
			4	12	6.00×10^{-7}	19/1000	yes		1.99	2.51s
			5	12	5.80×10^{-8}	15/1000	yes		1.92	2.00s
			Avg (successes)	12	1.79×10^{-7}	13.8/1000	6/6		1.76	1.86s

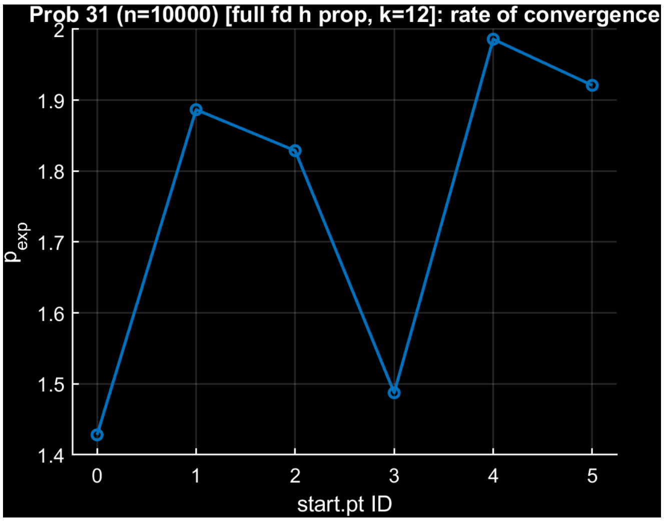


Figure 85: Problem 31, Modified Newton, full FD (component-wise h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
31	100000	full fd h prop	$\overline{x}$	12	6.61×10^{-8}	5/100	yes		1.31	6.71s
			1	12	2.65×10^{-10}	37/100	yes		1.36	49.76s
			2	12	5.45×10^{-10}	53/100	yes		1.24	70.38s
			3	12	4.39×10^{-4}	100/100	no		-1.32	131.49s
			4	12	3.99×10^{-4}	100/100	no		-0.57	137.33s
			5	12	4.02×10^{-4}	100/100	no		NaN	142.25s
			Avg (successes)	12	2.23×10^{-8}	31.7/100	3/6		1.30	42.28s

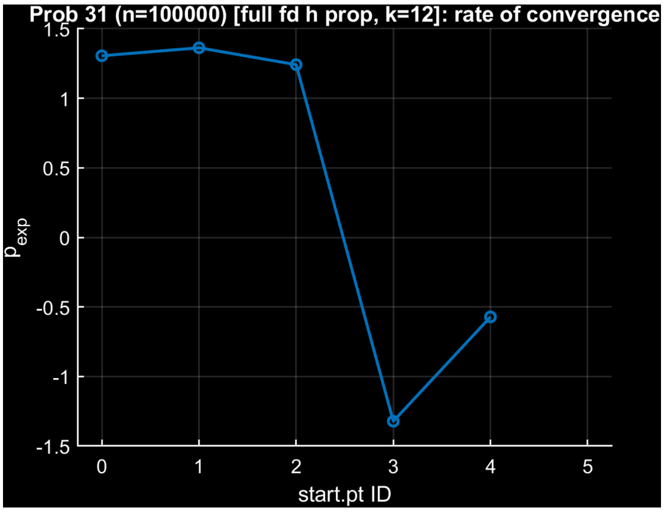


Figure 86: Problem 31, Modified Newton, full FD (component-wise h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

11.2 Problem 83

11.2.1 Modified Newton

Full FD (component-wise h), $k = 4$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	full fd h prop	$\overline{x}$	4	1.18×10^{-9}	3/1000	yes		1.86	0.0019s
			1	4	6.25×10^{-10}	3/1000	yes		2.00	0.0005s
			2	4	4.31×10^{-7}	3/1000	yes		1.94	0.0005s
			3	4	2.35×10^{-9}	3/1000	yes		1.98	0.0005s
			4	4	8.25×10^{-8}	3/1000	yes		2.01	0.0006s
			5	4	3.21×10^{-13}	4/1000	yes		1.54	0.0007s
			Avg (successes)	4	8.63×10^{-8}	3.2/1000	6/6		1.89	0.0008s

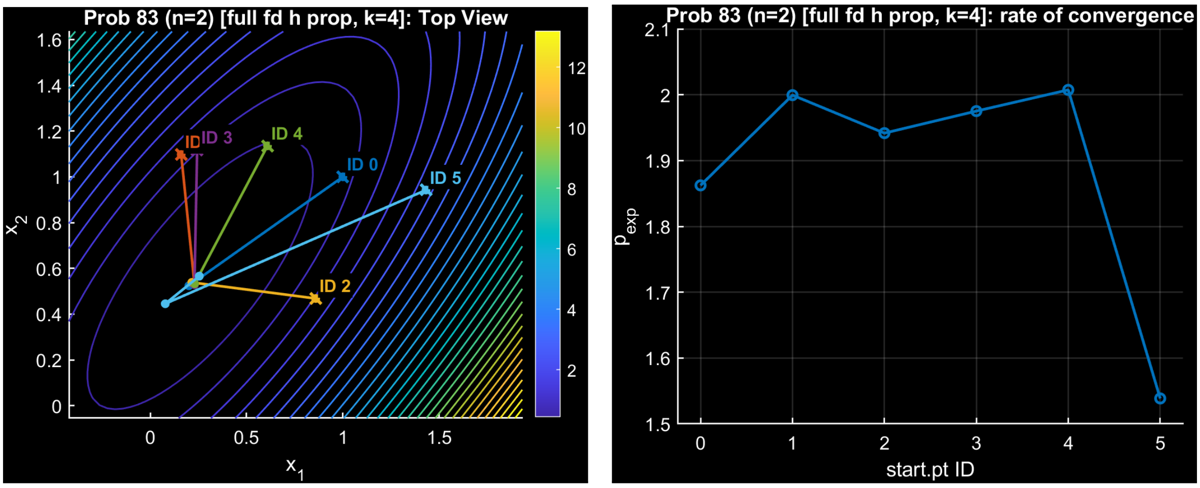


Figure 87: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	full fd h prop	$\overline{x}$	4	5.30×10^{-13}	2/1000	yes		NaN	0.0196s
			1	4	2.84×10^{-8}	2/1000	yes		NaN	0.0200s
			2	4	1.05×10^{-9}	2/1000	yes		NaN	0.0202s
			3	4	2.53×10^{-8}	2/1000	yes		NaN	0.0210s
			4	4	3.13×10^{-8}	2/1000	yes		NaN	0.0205s
			5	4	2.21×10^{-8}	2/1000	yes		NaN	0.0211s
			Avg (successes)	4	1.80×10^{-8}	2.0/1000	6/6		NaN	0.0204s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	full fd h prop	$\overline{x}$	4	1.14×10^{-8}	1/1000	yes		NaN	0.1005s
			1	4	1.85×10^{-8}	2/1000	yes		NaN	0.2006s
			2	4	4.63×10^{-10}	2/1000	yes		NaN	0.2096s
			3	4	2.44×10^{-9}	2/1000	yes		NaN	0.2077s
			4	4	1.82×10^{-8}	2/1000	yes		NaN	0.1907s
			5	4	8.77×10^{-9}	2/1000	yes		NaN	0.1909s
			Avg (successes)	4	9.98×10^{-9}	1.8/1000	6/6		NaN	0.1833s

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	full fd h prop	$\overline{x}$	4	9.67×10^{-9}	1/100	yes		NaN	1.01s
			1	4	4.82×10^{-8}	2/100	yes		NaN	2.27s
			2	4	2.57×10^{-8}	2/100	yes		NaN	2.16s
			3	4	8.04×10^{-8}	2/100	yes		NaN	2.39s
			4	4	7.06×10^{-8}	2/100	yes		NaN	2.26s
			5	4	4.40×10^{-8}	2/100	yes		NaN	2.15s
			Avg (successes)	4	4.64×10^{-8}	1.8/100	6/6		NaN	2.04s

Full FD (component-wise h), $k = 8$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	full fd h prop	$\overline{x}$	8	1.34×10^{-9}	3/1000	yes		1.84	0.0017s
			1	8	7.49×10^{-10}	3/1000	yes		1.97	0.0006s
			2	8	4.34×10^{-7}	3/1000	yes		1.94	0.0006s
			3	8	2.58×10^{-9}	3/1000	yes		1.96	0.0005s
			4	8	8.39×10^{-8}	3/1000	yes		2.00	0.0005s
			5	8	2.06×10^{-14}	4/1000	yes		1.82	0.0007s
			Avg (successes)	8	8.71×10^{-8}	3.2/1000	6/6		1.92	0.0008s

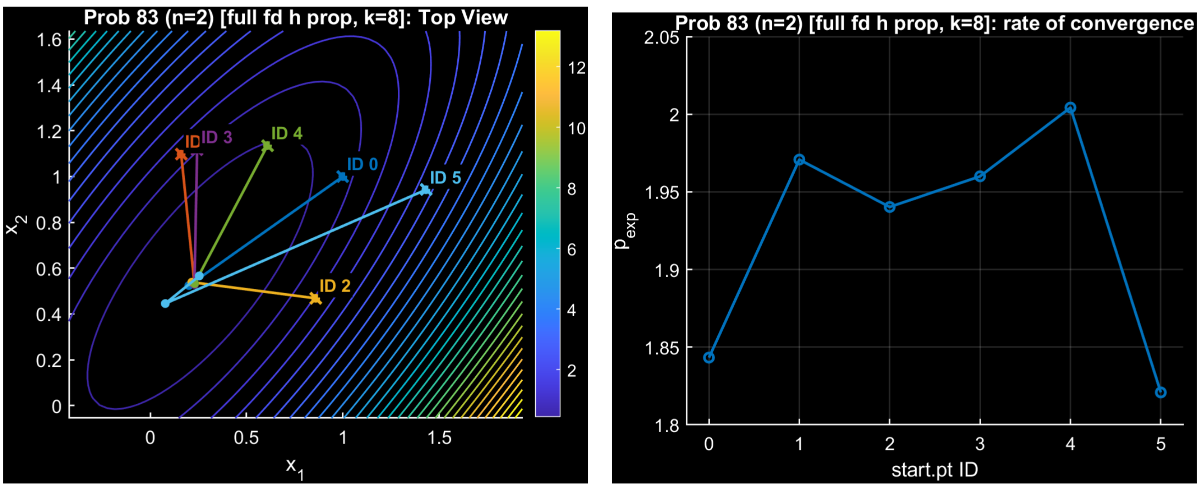


Figure 88: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	full fd h prop	$\overline{x}$	8	5.31×10^{-12}	2/1000	yes		NaN	0.0206s
			1	8	1.67×10^{-7}	3/1000	yes		1.44	0.0324s
			2	8	2.05×10^{-7}	3/1000	yes		0.62	0.0317s
			3	8	4.27×10^{-7}	2/1000	yes		NaN	0.0209s
			4	8	3.71×10^{-9}	3/1000	yes		0.53	0.0323s
			5	8	6.33×10^{-9}	5/1000	yes		9.18	0.0517s
			Avg (successes)	8	1.35×10^{-7}	3.0/1000	6/6		2.94	0.0316s

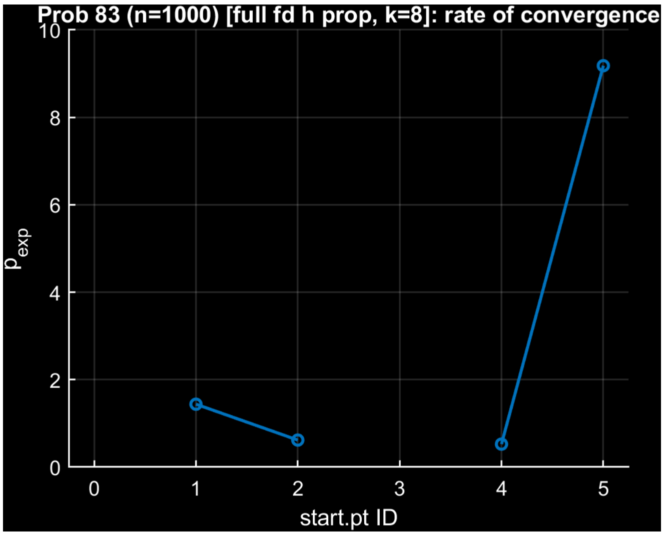


Figure 89: Problem 83, Modified Newton, full FD (component-wise h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	full fd h prop	$\overline{x}$	8	3.43×10^{-8}	1/1000	yes		NaN	0.0970s
			1	8	1.60×10^{-8}	3/1000	yes		0.94	0.2841s
			2	8	4.94×10^{-8}	2/1000	yes		NaN	0.2137s
			3	8	2.70×10^{-7}	2/1000	yes		NaN	0.2128s
			4	8	5.54×10^{-8}	2/1000	yes		NaN	0.2096s
			5	8	2.71×10^{-7}	2/1000	yes		NaN	0.2031s
			Avg (successes)	8	1.16×10^{-7}	2.0/1000	6/6		0.94	0.2034s

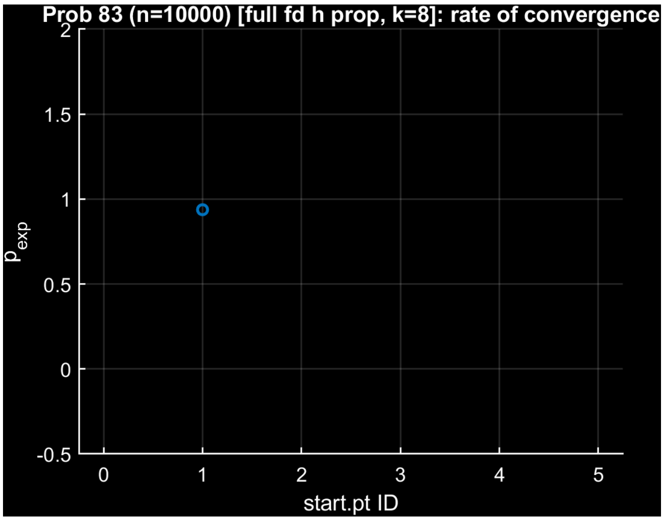


Figure 90: Problem 83, Modified Newton, full FD (component-wise h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	full fd h prop	$\overline{x}$	8	2.96×10^{-8}	1/100	yes		NaN	0.9962s
			1	8	6.67×10^{-8}	3/100	yes		2.80	3.19s
			2	8	4.06×10^{-8}	3/100	yes		1.72	3.11s
			3	8	4.84×10^{-8}	3/100	yes		2.34	3.65s
			4	8	4.49×10^{-7}	5/100	yes		0.28	4.98s
			5	8	2.60×10^{-8}	3/100	yes		0.87	3.48s
			Avg (successes)	8	1.10×10^{-7}	3.0/100	6/6		1.60	3.23s

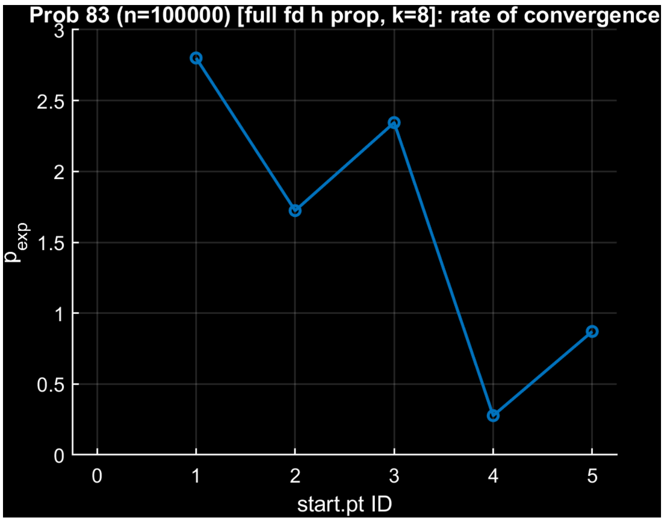


Figure 91: Problem 83, Modified Newton, full FD (component-wise h, $k = 8$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

Full FD (component-wise h), $k = 12$

Case $n = 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	2	full fd h prop	$\overline{x}$	12	2.30×10^{-8}	3/1000	yes		1.41	0.0013s
			1	12	4.72×10^{-9}	3/1000	yes		1.68	0.0005s
			2	12	7.62×10^{-7}	3/1000	yes		1.82	0.0005s
			3	12	3.70×10^{-8}	3/1000	yes		1.50	0.0011s
			4	12	3.27×10^{-7}	3/1000	yes		1.73	0.0005s
			5	12	3.79×10^{-10}	4/1000	yes		1.12	0.0007s
			Avg (successes)	12	1.92×10^{-7}	3.2/1000	6/6		1.54	0.0008s

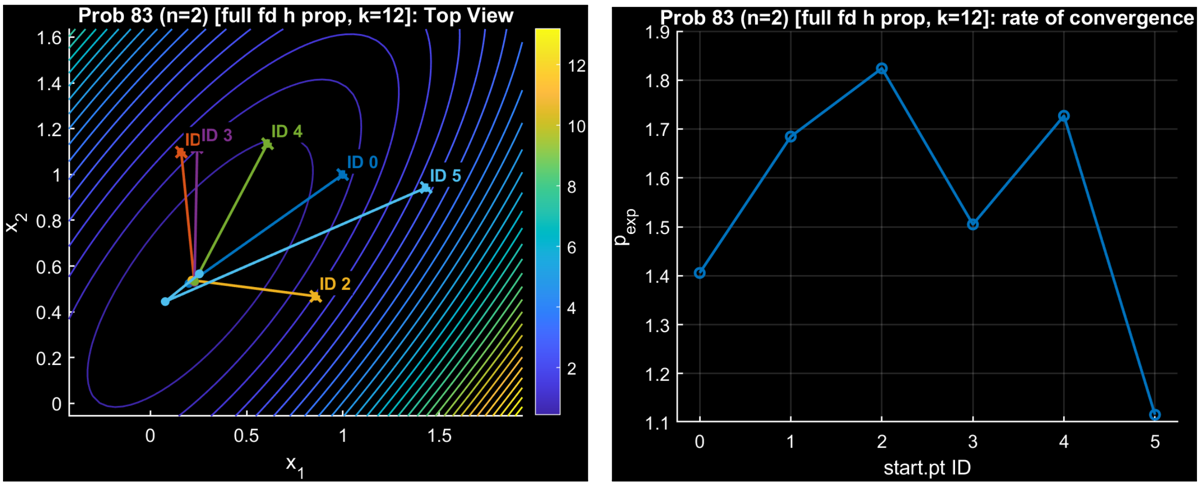


Figure 92: (Left) Top view of the objective level sets with the optimization paths from the different starting points; (right) experimental convergence rate for converged runs.

Case $n > 2$

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	1000	full fd h prop	$\overline{x}$	12	4.50×10^{-7}	2/1000	yes		NaN	0.0214s
			1	12	7.29×10^{-7}	6/1000	yes		NaN	0.0684s
			2	12	4.88×10^{-7}	5/1000	yes		0.47	0.0509s
			3	12	3.10×10^{-7}	5/1000	yes		1.06	0.0567s
			4	12	1.99×10^{-7}	6/1000	yes		0.58	0.0592s
			5	12	9.69×10^{-8}	8/1000	yes		4.04	0.0830s
			Avg (successes)	12	3.79×10^{-7}	5.3/1000	6/6		1.54	0.0566s

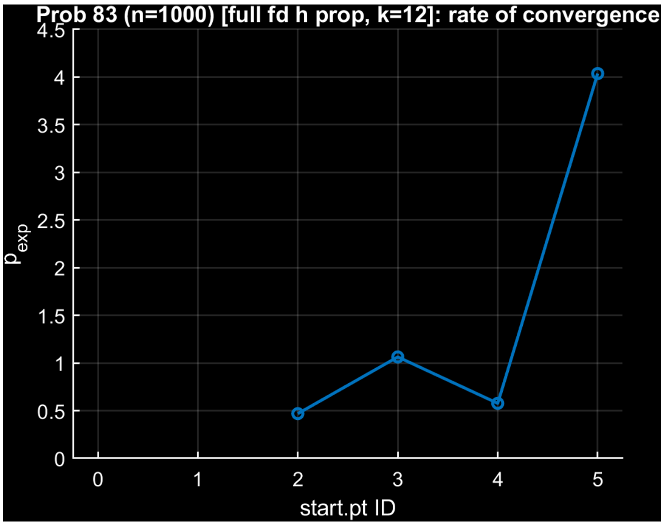


Figure 93: Problem 83, Modified Newton, full FD (component-wise h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 1000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	10000	full fd h prop	$\overline{x}$	12	8.87×10^{-8}	2/1000	yes		NaN	0.1885s
			1	12	2.80×10^{-7}	19/1000	yes		NaN	1.79s
			2	12	3.29×10^{-8}	8/1000	yes		0.83	0.7724s
			3	12	3.74×10^{-7}	9/1000	yes		1.37	0.9001s
			4	12	8.30×10^{-8}	8/1000	yes		0.61	0.7714s
			5	12	6.14×10^{-7}	6/1000	yes		0.65	0.6076s
			Avg (successes)	12	2.45×10^{-7}	8.7/1000	6/6		0.87	0.8385s

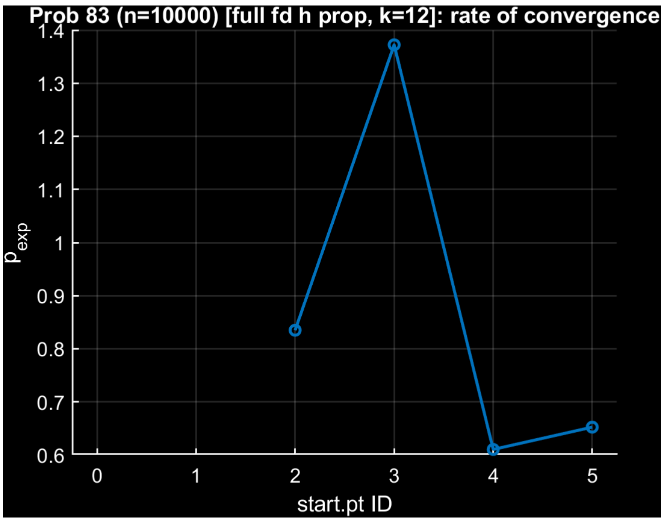


Figure 94: Problem 83, Modified Newton, full FD (component-wise h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 10000$).

Prob.	Dim.	Mode	start.pt ID	k	grad.norm	iters/max	succ.	flag	rate	time
83	100000	full fd h prop	$\overline{x}$	12	5.41×10^{-7}	2/100	yes		NaN	1.91s
			1	12	1.49×10^{-7}	88/100	yes		0.69	79.20s
			2	12	1.53×10^{-7}	68/100	yes		1.88	62.16s
			3	12	4.38×10^{-7}	72/100	yes		0.37	65.51s
			4	12	5.72×10^{-8}	60/100	yes		0.56	56.46s
			5	12	1.84×10^{-8}	98/100	yes		1.24	88.81s
			Avg (successes)	12	2.26×10^{-7}	64.7/100	6/6		0.95	59.01s

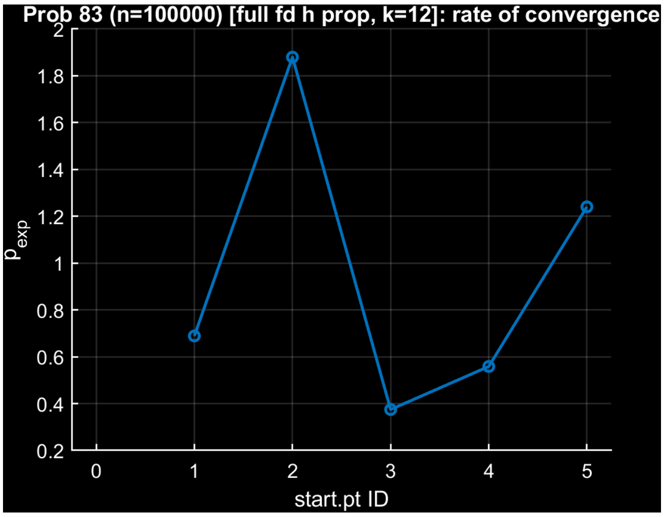


Figure 95: Problem 83, Modified Newton, full FD (component-wise h, $k = 12$): experimental convergence rate p_{exp} for converged runs ($n = 100000$).

12 Analysis of the results