Newton-CG for Nonlinear Inverse Problems

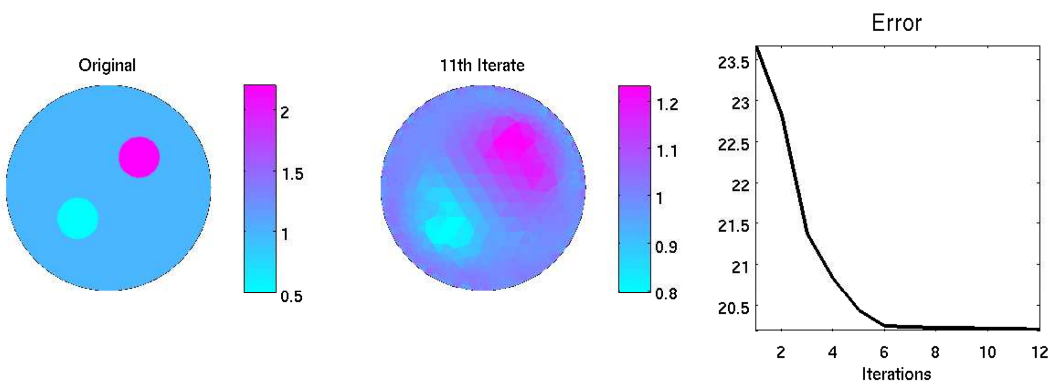
Andreas Rieder

UNIVERSITÄT KARLSRUHE (TH)

Institut für Wissenschaftliches Rechnen
und Mathematische Modellbildung

and

Institut für Angewandte und Numerische Mathematik



Overview

- Regularizations of Newton type for nonlinear problems
- Inexact Newton-CG regularizations
  - Convergence analysis
  - Aspects of implementation
- Numerical experiments: impedance tomography
- References
- Conclusion: What to remember from this talk

Newton regularizations

$$F : \mathrm{D}(F) \subset X \to Y, \quad X, Y \text{ Hilbert spaces}$$

$$F(x) = y^\delta$$

where $\|y - y^\delta\| \le \delta$, $y = F(x^+)$, and $F(x) = y$ locally ill-posed in x^+.

Let x_n be an approximation to x^+: $\quad x_{n+1} = x_n + s_n$

The exact Newton step $s_n^{\mathrm{e}} = x^+ - x_n$ satisfies ($A_n := F'(x_n)$)

$$A_n s_n^{\mathrm{e}} = y - F(x_n) - E(x^+, x_n)$$

$\Longrightarrow$ Determine s_n as regularized solution of

$$A_n s = b_n^\delta, \quad b_n^\delta := y^\delta - F(x_n)$$

Newton regularizations (continued)

$$s_n = s_{n,\ell} = g_\ell(A_n^* A_n) A_n^* b_n^\delta + \big(I - g_\ell(A_n^* A_n) A_n^* A_n\big) s_{\star,n}$$

$\{g_\ell\}$ regularizing filter:

$$\lim_{\ell\to\infty} s_{n,\ell} = \begin{cases} A_n^+ b_n^\delta + P_{\mathsf{N}(A_n)} s_{\star,n} & : \; b_n^\delta \in \mathsf{R}(A_n) \oplus \mathsf{R}(A_n)^\perp \\ \infty & : \; \text{otherwise} \end{cases}$$

$$x_{n+1} = x_n + g_{i_n}(A_n^* A_n) A_n^* b_n^\delta + \big(I - g_{i_n}(A_n^* A_n) A_n^* A_n\big) s_{\star,n}$$

Tasks:

1. choice of $\{g_\ell\}$
2. choice of $s_{\star,n}$
3. choice of i_n
4. stopping of the Newton iteration

Examples

$$x_{n+1} = x_n + g_{i_n}(A_n^* A_n) A_n^* b_n^\delta + \big(I - g_{i_n}(A_n^* A_n) A_n^* A_n\big) s_{\star,n}$$

- nonlinear Landweber (Hanke, Neubauer, Scherzer 95)

$$x_{n+1} = x_n + A_n^* b_n^\delta, \qquad \|A_n\| \le 1,$$

$g_\ell(t) = \sum_{j=0}^{\ell-1}(1-t)^j$, $i_n = 1$, $s_{\star,n} = 0$, discrepancy principle

Examples

$$x_{n+1} = x_n + g_{i_n}(A_n^* A_n) A_n^* b_n^\delta + \big(I - g_{i_n}(A_n^* A_n) A_n^* A_n\big) s_{\star,n}$$

- nonlinear Landweber (Hanke, Neubauer, Scherzer 95)

$$x_{n+1} = x_n + A_n^* b_n^\delta, \qquad \|A_n\| \leq 1,$$

$g_\ell(t) = \sum_{j=0}^{\ell-1} (1-t)^j$, $i_n = 1$, $s_{\star,n} = 0$, discrepancy principle

- nonlinear gradient decent (Scherzer 96)

$$x_{n+1} = x_n + \lambda_n A_n^* b_n^\delta, \qquad \lambda_n = \|A_n^* b_n^\delta\|^2 / \|A_n A_n^* b_n^\delta\|^2,$$

$i_n = 1$, $s_{\star,n} = 0$, discrepancy principle

Examples (continued)

- iteratively regularized Gauss-Newton methods (Bakushinsky 92, Blaschke(Kaltenbacher), Neubauer & Scherzer 97, Kaltenbacher 98, ...)

$$x_{n+1} = x_0 + g_{i_n}(A_n^* A_n) A_n^* \big(b_n^\delta - A_n(x_0 - x_n)\big)$$

$g_\ell(t) = 1/(t + \alpha_\ell)$, $\lim_{\ell\to\infty} \alpha_\ell = 0$ (strongly monotone),
$\{i_n\}$ chosen a priori (strongly increasing),
$s_{\star,n} = x_0 - x_n$, discrepancy principle/a priori

Textbooks (on regularizations for nonlinear problems)

- A. B. Bakushinsky, M. Yu. Kokurin
  Iterative Methods for Approximate Solution of Inverse Problems
  Springer, 2004

- Y. Alber, I. Ryazantseva
  Nonlinear Ill-posed Problems of Monotone Type
  Springer, 2006

- B. Kaltenbacher, A. Neubauer, O. Scherzer
  Iterative Regularization Methods for Nonlinear Ill-posed Problems
  Springer, 2006 (to appear??)

- A. Rieder
  Keine Probleme mit Inversen Problemen (Chapter 7)
  Vieweg, 2003

Inexact Newton-CG scheme

$$x_{n+1} = x_n + g_{i_n}(A_n^* A_n) A_n^* b_n^\delta + \big(I - g_{i_n}(A_n^* A_n) A_n^* A_n\big) s_{\star,n}$$

Guiding principles

1. $s_{\star,n} = 0$ (because $s_n \approx s_n^{\mathsf{e}} = x^+ - x_n \to 0$)
2. Adapt i_n dynamically according to the local degree of ill-posedness
3. Choose the most efficient regularization scheme for $\{g_\ell\}$
4. Stop the iteration by an a-posteriori principle.

Inexact Newton-CG scheme (continued)

CG-REGINN$(x, R, \{\mu_k\})$

$n = 0$, $x_0 = x$

while $\|b_n^\delta\| > R\,\delta$ do

$\{ \quad i_n = 0$

repeat

$i_n = i_n + 1$

$s_{i_n} = g_{i_n}^{\mathrm{cg}}\big(A_n^* A_n, b_n^\delta\big) A_n^* b_n^\delta$

until $\dfrac{\|A_n\, s_{i_n} - b_n^\delta\|}{\|b_n^\delta\|} < \mu_n$

$x_{n+1} = x_n + s_{i_n}$

$n = n + 1 \quad \}$

$x = x_n$

Termination of CG-REGINN

Theorem: Assume the tangential cone condition

$$\|E(v,u)\| = \|F(v) - F(u) - F'(u)(v-u)\| \le \omega \|F(v) - F(u)\|, \quad v, u \in B_\rho(x^+)$$

Moreover: ω, R, $\{\mu_k\}$ "suitable", $x_0 \in B_\rho(x^+)$

Then, all iterates $\{x_1, \ldots, x_{N(\delta)}\}$ are well defined in $B_\rho(x^+)$. Additionally,

$$\|x^+ - x_n\|_X < \|x^+ - x_{n-1}\|_X, \quad n = 1, \ldots, N(\delta).$$

Furthermore,

$$\|y^\delta - F(x_{N(\delta)})\|_Y \le R\delta < \|y^\delta - F(x_n)\|_Y, \quad n = 0, \ldots, N(\delta) - 1$$

and

$$\frac{\|y^\delta - F(x_{n+1})\|_Y}{\|y^\delta - F(x_n)\|_Y} < \mu_n + \frac{\omega}{1-\omega} \le \Lambda < 1, \quad n = 0, \ldots, N(\delta) - 1.$$

Convergence of CG-REGINN

Corollary: Let F be weakly sequentially closed.

Then, any subsequence of $\{x_{N(\delta)}\}_{0<\delta\le\delta_{\max}}$ contains a subsequence which converges weakly to a solution of $F(x) = y$.

Convergence of CG-REGINN

Corollary: Let F be weakly sequentially closed.

Then, any subsequence of $\{x_{N(\delta)}\}_{0<\delta\leq\delta_{\max}}$ contains a subsequence which converges weakly to a solution of $F(x) = y$.

Theorem: Assume the factorization

$$\left.\begin{aligned} F'(v) &= Q(v,w)F'(w) \\ \|I - Q(v,w)\| &\leq C_Q \, \|v - w\| \end{aligned}\right\} \forall v, w \in B_\rho(x^+)$$

Moreover: $C_Q\rho$, R, $\{\mu_k\}$ “suitable’

Then, there is a $\kappa_{\text{min}} < 1$ such that $x_0 \in B_\rho(x^+)$ with

$$x^+ - x_0 = |F'(x^+)|^\kappa w \text{ for one } \kappa \in]\kappa_{\text{min}}, 1]$$

and $\|w\| \, \|y^\delta - F(x_0)\|$ sufficiently small, imply that

$$\|x^+ - x_{N(\delta)}\| \leq C_\kappa \, \delta^{\frac{\kappa - \kappa_{\text{min}}}{1+\kappa}} \, \|w\|^{\frac{1}{1+\kappa}} \quad \text{as } \delta \to 0.$$

Aspects of implementation: dynamic choice of tolerances

Goal: μ_n as small as possible; since

$$\frac{\|F(x_{n+1}) - y^\delta\|}{\|F(x_n) - y^\delta\|} \approx \mu_n$$

Aspects of implementation: dynamic choice of tolerances

Goal: μ_n as small as possible; since

$$\frac{\|F(x_{n+1}) - y^\delta\|}{\|F(x_n) - y^\delta\|} \approx \mu_n$$

Initialize $\mu_{\text{start}} \in \,]0,1[, \quad \gamma \in \,]0,1], \quad \mu_{\text{max}} \in \,]\mu_{\text{start}}, 1[, \quad \widetilde{\mu}_0 = \widetilde{\mu}_1 := \mu_{\text{start}}$

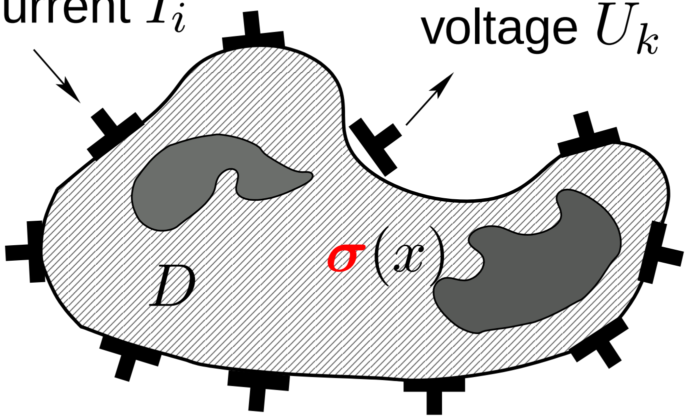
Choose

$$\mu_n := \mu_{\text{max}} \max\left\{R\,\delta / \|F(x_n) - y^\delta\|,\, \widetilde{\mu}_n\right\}$$

where

$$\widetilde{\mu}_n := \begin{cases} 1 - \frac{i_{n-2}}{i_{n-1}}\,(1 - \mu_{n-1}) & : \quad i_{n-1} > i_{n-2} \\ \gamma\, \mu_{n-1} & : \quad \text{otherwise} \end{cases} \qquad n \geq 2$$

Numerical experiments: Impedance tomography



$$\operatorname{div}(\sigma \nabla u) = 0 \quad \text{in } D,$$

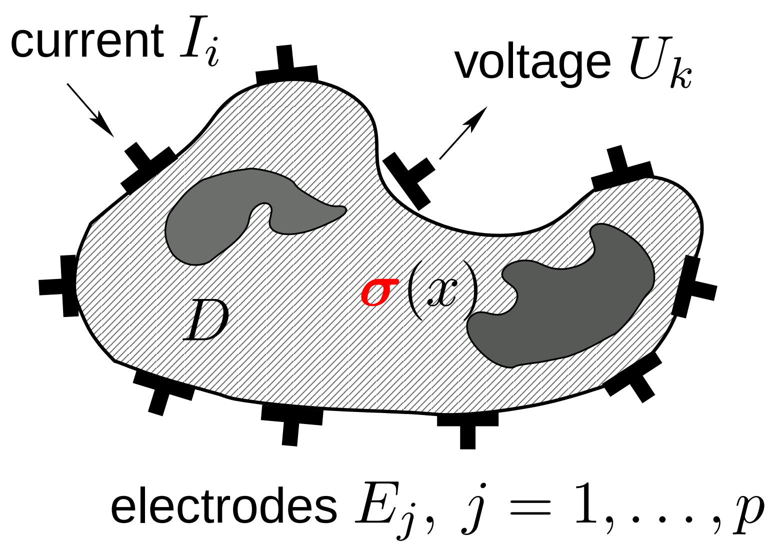
$$u + z_j\, \sigma \partial_{\mathbf{n}} u = U_j \quad \text{on } E_j$$

$$\sigma \partial_{\mathbf{n}} u = 0 \quad \text{on } \partial D \setminus \cup_j E_j.$$

$$\int_{E_j} \sigma \partial_{\mathbf{n}} u \, \mathrm{d}S = I_j$$

$F : \sigma \mapsto U = (U_1, \dots, U_p)^\top \in \mathbb{R}^p_\diamond$ for a fixed current pattern $I \in \mathbb{R}^p_\diamond$

Numerical experiments: Impedance tomography



$$\begin{aligned} \operatorname{div}(\sigma \nabla u) &= 0 \quad \text{in } D, \\ u + z_j\, \sigma \partial_{\mathbf{n}} u &= U_j \quad \text{on } E_j \\ \sigma \partial_{\mathbf{n}} u &= 0 \quad \text{on } \partial D \setminus \cup_j E_j. \end{aligned}$$

$$\int_{E_j} \sigma \partial_{\mathbf{n}} u \,\mathrm{d}S = I_j$$

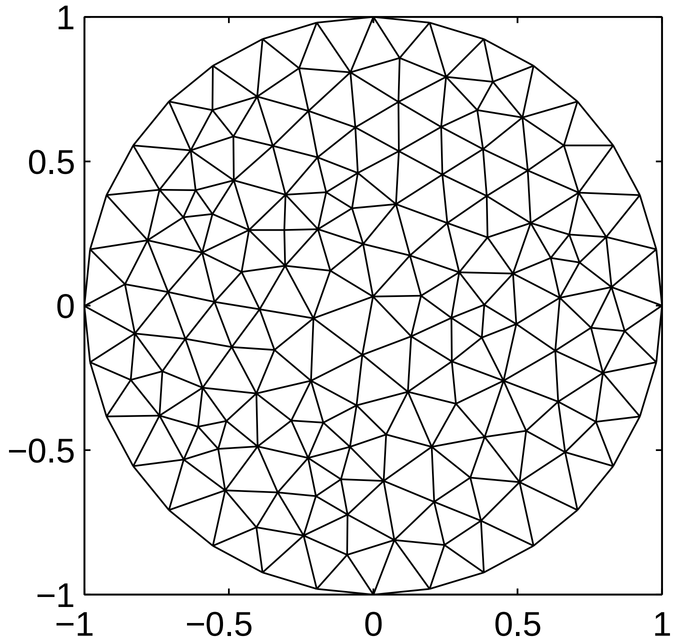
$F : \sigma \mapsto U = (U_1, \ldots, U_p)^\top \in \mathbb{R}^p_\diamond$ for a fixed current pattern $I \in \mathbb{R}^p_\diamond$

Modification: ℓ different current patterns are applied, that is,

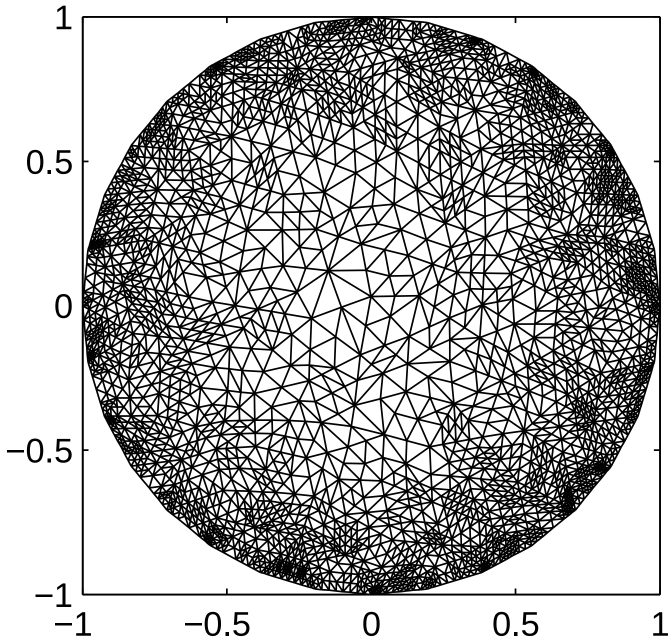
$\boldsymbol{F} : \sigma \mapsto \boldsymbol{U} \in \mathbb{R}^{p\cdot\ell}_\diamond$ for $\boldsymbol{I} \in \mathbb{R}^{p\cdot\ell}_\diamond$

In all numerical experiments: $\ell = p$

Numerical experiments: FE meshes for 16 electrodes



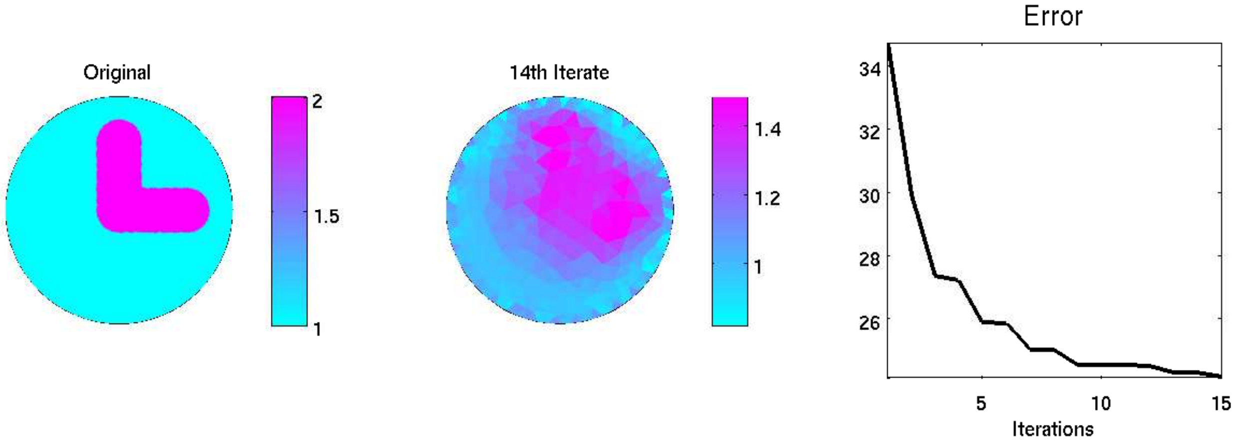
Finite element mesh
for reconstruction



Finite element mesh for
computing the Jacobian

Numerical experiments: Results

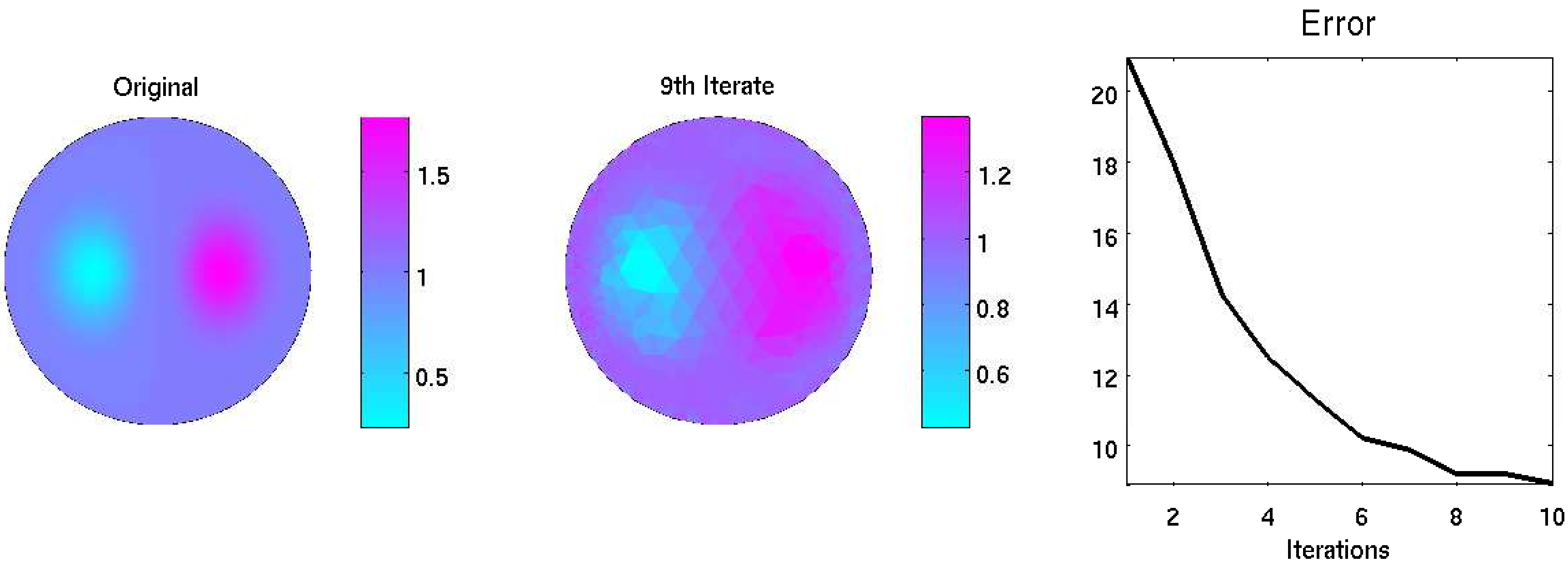
32 electrodes, 0.5% artificial noise



n	i_n	μ_n	error in %
0	0	—	34.72
1	2	0.799	29.92
2	4	0.799	27.35
3	3	0.899	27.20
4	6	0.871	25.88
5	2	0.935	25.84
6	6	0.906	25.01
7	1	0.968	25.00
8	6	0.938	24.52
9	1	0.989	24.52
10	2	0.958	24.50
11	1	0.978	24.50
12	5	0.948	24.30
13	1	0.989	24.29
14	5	0.958	24.15
15	1	0.991	24.15

Numerical experiments: Results (continued)

64 electrodes, 1% artificial noise



References

1. M. Hanke
   Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems
   Numer. Funct. Anal. Optim. 18 (1998), 971-993.
2. A. Rieder
   Inexact Newton regularization using conjugate gradients as inner iteration
   SIAM J. Numer. Anal. 43 (2005), 604-622.
3. A. Lechleiter, A. Rieder
   Newton regularizations for impedance tomography: a numerical study
   Inverse Problems (2006), to appear
4. A. Lechleiter, A. Rieder
   A convergence analysis of the Newton-type regularization CG-REGINN
   work in progress

Conclusion: What to remember from this talk

- The vital part of any Newton-like regularization is the stable computation of the Newton step from the locally linearized system. As the degree of ill-posedness of this system may change during the iteration a careful selection of the level of regularization is indispensable.
  Surprisingly, this is not the case for most Newton methods.

- **`CG-REGINN`** selects the level of regularization of the locally linearized system incorporating information on the local degree of ill-posedness gained during the iteration. To this end the numerical effort for computing the Newton steps is monitored. An increase (decrease) of this numerical effort indicates an increase (decrease) of the local degree of ill-posedness. Accordingly the magnitude of regularization for the following linearized system is adjusted.