A Derivative-Free Optimization Method on Hadamard Manifold for Solving Nonlinear Monotone Systems with Application in Image Deblurring

with an Application to Medical Image Restoration

Kazeem Olalekan Aremu

Joint work with Paul K. Imoukhedeme and Ahmed Audu

Department of Mathematics, Usmanu Danfodiyo University Sokoto, Nigeria
Department of Mathematics and Applied Mathematics, Sefako Makgatho Health Sciences University, Pretoria

Mathematical Optimization and Optimal Control Theory and Applications (MOOCTA)
7–10 September 2026 • Manhattan Hotel, Pretoria, South Africa

Outline

1 Introduction

2 Preliminaries

3 Algorithm Formulation

4 Convergence analysis

5 Numerical experiments

6 Application: Restoration of motion-blurred chest X-rays

Find $x^* \in \mathbb{R}^n$ with

$$F(x^*) = 0, \qquad F : \mathbb{R}^n \to \mathbb{R}^n \text{ continuous and monotone: } \langle F(x) - F(y), x - y \rangle \geq 0. \quad (1)$$

- Optimality systems, discretised BVPs, KKT systems of variational inequalities, ℓ_1-regularised signal recovery.
- Newton / quasi-Newton: Jacobian + an $n \times n$ solve per step too costly for large n, nonsmooth or expensive F.

Conjugate gradient (CG): $x_{k+1} = x_k + \alpha_k d_k$, $d_k = -g_k + \beta_k d_{k-1}$: No Jacobian matrix computation.

- $\beta_k^{\text{FR}} = \dfrac{\|g_k\|^2}{\|g_{k-1}\|^2}$: Nice convergence property, but may jam.
  $\beta_k^{\text{PRP}} = \dfrac{\langle g_k, g_k - g_{k-1} \rangle}{\|g_{k-1}\|^2}$: built-in restart, no global convergence in general.
- Hu-Storey hybrid: $\beta_k^{\text{HuS}} = \max\{0, \min\{\beta_k^{\text{PRP}}, \beta_k^{\text{FR}}\}\}$. (2)

Replace g_k by the residual $F_k := F(x_k)$. No merit function but monotonicity suffices: if the trial point $z_k = x_k + \alpha_k d_k$ passes the line search, the hyperplane

$$\mathcal{H}_k = \{x : \langle F(z_k), x - z_k \rangle = 0\} \tag{3}$$

separates x_k from *every* solution. Projecting onto $\mathcal{H}_k$ gives Fejér monotonicity.

- global convergence under monotonicity + Lipschitz continuity.
- Two-term hybrid version (Papp, Rapajić, Ibrahim, Phiangsungnoen 2024): Hu-Storey rule (2) and Cheng's projection $I - F_k F_k^\top / \|F_k\|^2$.
- Limitation: $\beta_k^{\mathrm{PRP}} < 0 \;\Rightarrow\; \beta_k^{\mathrm{HuS}} = 0$-CG memory is discarded.

Convex combination (Djordjević 2016): in $\mathbb{R}^n$

$$\beta_k^{\mathrm{cc}} = \lambda_k \beta_k^{\mathrm{PRP}} + (1-\lambda_k)\beta_k^{\mathrm{FR}}, \quad \lambda_k \in [0,1] \text{ from the Dai-Liao condition } \langle y_{k-1}, d_k \rangle = 0. \tag{4}$$

Never discards the CG memory.

Diffusion tensors, shapes, rotations, hyperbolic embeddings do not live in $\mathbb{R}^n$. For a vector field $F : \mathcal{M} \to T\mathcal{M}$,

$$F(x^*) = 0_{x^*} \in T_{x^*}\mathcal{M}. \tag{5}$$

- Monotone (Németh 1999): $\langle F(x), \exp_x^{-1} y\rangle \leq -\langle F(y), \exp_y^{-1} x\rangle$.
- Existing methods: proximal point (Li–López-Martín-Márquez 2009), extragradient (Ferreira et al. 2005), inertial forward–backward (Abass 2025) resolvent or two evaluations of F per step, Lipschitz constant needed.
- Single-direction projection methods: Zhao et al. 2022, Yao et al. 2024, Sahu–Sharma 2024. Loss of conjugacy under transport makes $\beta_k^{\mathrm{PRP}} < 0$.

Hyperplane → hypersurface. Projection onto (3) is not feasible. So we employ the level set

$$H_k = \{p \in \mathcal{M} : \langle F(z_k), \exp_{z_k}^{-1} p\rangle = 0\} = \partial L_k, \qquad L_k = \{p : \langle F(z_k), \exp_{z_k}^{-1} p\rangle \leq 0\}. \tag{6}$$

$\mathcal{S} \subset L_k$, $x_k \notin L_k$; in constant curvature L_k is geodesically convex, so $P_{\mathcal{S}}(x_k) = P_{H_k}(x_k)$.

Our Contributions

1. Riemannian two-term direction $d_k = -F_k + \beta_k^{\text{cc}} \Pi_k s_{k-1}$ with $\langle F_k, d_k \rangle = -\|F_k\|^2$ always.
2. Convex weight λ_k from transported conjugacy: β_k^{cc} stays between PRP and FR and never degenerates to 0; every Hu-Storey value lies in $\text{conv}\{0, \beta_k^{\text{PRP}}, \beta_k^{\text{FR}}\}$.
3. Fejér monotonicity and convergence of the whole sequence to a zero of F under monotonicity and Lipschitz continuity without derivatives.
4. Up to $n = 50{,}000$: an order of magnitude fewer iterations and evaluations than the Hu-Storey hybrids.

Definition

A Hadamard manifold is a complete, simply connected Riemannian manifold of nonpositive sectional curvature.

Proposition

(*a*) $\exp_x : T_x\mathcal{M} \to \mathcal{M}$ *is a diffeomorphism: geodesics are unique,* $\exp_x^{-1}$ *is well defined,* $\|\exp_x^{-1} y\| = \mathrm{d}(x, y)$.
(*b*) *Distance expansion:* $\mathrm{d}(\exp_x u, \exp_x v) \geq \|u - v\|$, *i.e.* $\|\exp_x^{-1} p - \exp_x^{-1} q\| \leq \mathrm{d}(p, q)$. (7)

Parallel transport $P_{x\to y}$ along $\gamma_{x,y}$ is a linear isometry, continuous in (x, y), with

$$\langle P_{x\to y} u, w\rangle_y = \langle u, P_{y\to x} w\rangle_x, \qquad \exp_y^{-1} x = -P_{x\to y}(\exp_x^{-1} y). \tag{8}$$

Definition (Németh 1999)

F is *monotone* if $\langle P_{y\to x} F(y) - F(x), \exp_x^{-1} y\rangle \geq 0$, *Lipschitz* if $\|P_{y\to x} F(y) - F(x)\| \leq L\ \mathrm{d}(x, y)$.
Then $\mathcal{S} = \{F = 0\}$ is closed and geodesically convex.

Lemma (Projection; Li-López–Martín-Márquez 2009)

For C nonempty, closed, geodesically convex, $P_C(x) = \operatorname{argmin}_{p \in C} \mathrm{d}(x,p)$ is single-valued, nonexpansive, and $\mathrm{d}^2(P_C(x),p) \le \mathrm{d}^2(x,p) - \mathrm{d}^2(x,P_C(x))$ for all $p \in C$.

Lemma (Geodesic Half-space; Sahu-Sharma 2024)

If $\mathcal{M}$ has constant curvature, $L_{q,s} = \{p : \langle s, \exp_q^{-1} p\rangle \le 0\}$ $(0 \neq s \in T_q\mathcal{M})$ is closed and geodesically convex, with boundary $H_{q,s} = \{p : \langle s, \exp_q^{-1} p\rangle = 0\}$.

Constant curvature is used *only* here; $\mathbb{H}^n$, and curvature 0 recovers the Euclidean methods.

Lemma (Fejér monotonicity; Ferreira et al. 2005)

If $\mathrm{d}(x_{k+1},p) \le \mathrm{d}(x_k,p)$ for all $p \in C \neq \emptyset$ and all k, then $\{x_k\}$ is bounded, $\lim_k \mathrm{d}(x_k,p)$ exists, and a cluster point in C is the limit of the whole sequence.

Following from Euclidean space (Cheng 2007; Papp et al. 2024):
$d_k = -F_k + \beta_k (I - F_k F_k^\top / \|F_k\|^2) d_{k-1}$.

Transporting to $T_{x_k}\mathcal{M}$:

- previous *trial step* $s_{k-1} := P_{x_{k-1} \to x_k}(\exp_{x_{k-1}}^{-1} z_{k-1}) = \alpha_{k-1} P_{x_{k-1} \to x_k} d_{k-1}$, so $\|s_{k-1}\| = \mathrm{d}(x_{k-1}, z_{k-1})$;
- residual difference $y_{k-1} := F_k - P_{x_{k-1} \to x_k} F_{k-1}$ — exactly what Lipschitz continuity bounds;
- projection onto $F_k^\perp$: $\Pi_k v := v - \dfrac{\langle F_k, v \rangle}{\|F_k\|^2} F_k$, so $\langle F_k, \Pi_k v \rangle = 0$, $\|\Pi_k v\| \le \|v\|$.

$$\beta_k^{\mathrm{FR}} = \frac{\|F_k\|^2}{\|F_{k-1}\|^2}, \qquad \beta_k^{\mathrm{PRP}} = \frac{\langle F_k, y_{k-1} \rangle}{\|F_{k-1}\|^2}, \qquad d_k = \begin{cases} -F_k, & k = 0, \\ -F_k + \beta_k^{\mathrm{cc}} \Pi_k s_{k-1}, & k \ge 1. \end{cases} \tag{9}$$

Conjugacy to the transported step, $\langle y_{k-1}, d_k\rangle = 0$, with $\theta_k := \langle y_{k-1}, \Pi_k s_{k-1}\rangle$ gives

$$\lambda_k^* = \frac{\langle y_{k-1}, F_k\rangle - \beta_k^{\mathrm{FR}}\theta_k}{(\beta_k^{\mathrm{PRP}} - \beta_k^{\mathrm{FR}})\theta_k}, \qquad \lambda_k = \begin{cases} \min\{1, \max\{0, \lambda_k^*\}\}, & \left|(\beta_k^{\mathrm{PRP}} - \beta_k^{\mathrm{FR}})\theta_k\right| \geq \epsilon_\lambda, \\ \frac{1}{2}, & \text{otherwise.} \end{cases} \tag{10}$$

We use $\epsilon_\lambda = 10^{-12}$; the denominator vanishes only when consecutive residuals are orthogonal after transport or $y_{k-1} \perp \Pi_k s_{k-1}$, where the two parameters are interchangeable. Hence $\min\{\beta_k^{\mathrm{PRP}}, \beta_k^{\mathrm{FR}}\} \leq \beta_k^{\mathrm{cc}} \leq \max\{\beta_k^{\mathrm{PRP}}, \beta_k^{\mathrm{FR}}\}$ and $|\beta_k^{\mathrm{cc}}| \leq \max\{|\beta_k^{\mathrm{PRP}}|, |\beta_k^{\mathrm{FR}}|\}$. (11)

Proposition (Hu–Storey values lie in the convex hull)

$\beta_k^{\mathrm{HuS}} \in \mathrm{conv}\{0, \beta_k^{\mathrm{PRP}}, \beta_k^{\mathrm{FR}}\}$ for every k. If $\min\{\beta_k^{\mathrm{PRP}}, \beta_k^{\mathrm{FR}}\} > 0$ it is an endpoint $\lambda_k \in \{0, 1\}$; if $\beta_k^{\mathrm{PRP}} < 0 < \beta_k^{\mathrm{FR}}$ the value 0 is attained at the interior weight $\lambda_k = \beta_k^{\mathrm{FR}}/(\beta_k^{\mathrm{FR}} - \beta_k^{\mathrm{PRP}})$.

The Hu–Storey rule is a discontinuous choice of λ_k; (10) chooses by conjugacy and is never forced to 0.

Largest $\alpha_k \in \{\bar{\alpha}_k \rho^j\}$ such that $z_k = \exp_{x_k}(\alpha_k d_k)$ satisfies

$$-\langle P_{z_k \to x_k} F(z_k), d_k \rangle \;\geq\; \sigma\, \alpha_k \, \|F(z_k)\| \, \|d_k\|^2, \tag{12}$$

equivalently (by (8)) $\langle F(z_k), \exp_{z_k}^{-1} x_k \rangle \geq \sigma \alpha_k^2 \|F(z_k)\| \|d_k\|^2 > 0$. Derivative-free and scale-invariant; it places x_k strictly outside L_k of (6). Then

$$x_{k+1} = P_{\partial L_k}(x_k) = P_{H_k}(x_k). \tag{13}$$

Remark

- In $\mathbb{R}^n$: $x_{k+1} = x_k - \dfrac{\langle F(z_k), x_k - z_k \rangle}{\|F(z_k)\|^2} F(z_k)$. On $\mathbb{H}^n$: a few Newton/bisection steps along $t \mapsto \exp_{x_k}(-t\, P_{z_k \to x_k} F(z_k))$.
- Initial step: derivative-free secant estimate $\bar{\alpha}_k = \left| \dfrac{\langle F_k, d_k \rangle}{\langle P_{\tilde{z}_k \to x_k} F(\tilde{z}_k) - F_k, d_k \rangle / t} \right|$, $\tilde{z}_k = \exp_{x_k}(t d_k)$, $t = 10^{-8}$, truncated to $[\alpha_{\min}, \alpha_{\max}]$.

Require: $x_0 \in \mathcal{M}$; $\sigma > 0$, $\rho \in (0,1)$, $0 < \alpha_{\min} \le \alpha_{\max}$, $\epsilon_\lambda > 0$; tolerance $\epsilon > 0$; $k_{\max}$.

$k \leftarrow 0$, $F_0 \leftarrow F(x_0)$.
while $\|F_k\| > \epsilon$ and $k < k_{\max}$ **do**
  if $k = 0$ **then** $d_k \leftarrow -F_k$
  else compute $s_{k-1}, y_{k-1}, \Pi_k, \beta_k^{\mathrm{FR}}, \beta_k^{\mathrm{PRP}}$; λ_k by (10); d_k by (9).
  end if
  Choose $\bar{\alpha}_k \in [\alpha_{\min}, \alpha_{\max}]$; backtrack $\alpha_k = \bar{\alpha}_k \rho^j$ until $z_k = \exp_{x_k}(\alpha_k d_k)$ satisfies (12).
  if $\|F(z_k)\| \le \epsilon$ **then return** z_k
  end if
  $x_{k+1} \leftarrow P_{H_k}(x_k)$ by (13); $F_{k+1} \leftarrow F(x_{k+1})$; $k \leftarrow k+1$.
end while
return x_k

Assumptions

A1 F monotone. **A2** F Lipschitz with constant L. **A3** $\mathcal{S} \neq \emptyset$.

$\mathcal{M}$ of constant curvature; sequences generated by CC-DFP with $\epsilon = 0$. A2 is needed only for bounded directions and the final theorem.

Lemma (Sufficient descent)

$\langle F_k, d_k \rangle = -\|F_k\|^2$ and $\|d_k\| \geq \|F_k\|$ for every k; moreover $d_k = \lambda_k d_k^{\mathrm{PRP}} + (1 - \lambda_k) d_k^{\mathrm{FR}}$.

Lemma (Finite line search)

If F is continuous and $F_k \neq 0$, the backtracking (12) terminates; α_k, z_k are well defined.

Lemma (Separation)

If F is continuous and monotone: (a) $\langle F(z_k), \exp_{z_k}^{-1} x_k \rangle \geq \sigma \alpha_k^2 \|F(z_k)\| \|d_k\|^2 > 0$, so $x_k \notin L_k$ and $F(z_k) \neq 0$; (b) $\mathcal{S} \subset L_k$.

Theorem (Fejér monotonicity)

Under A1, A3 and continuity of F, for every $x^* \in \mathcal{S}$ and every k,

$$\mathrm{d}^2(x_{k+1}, x^*) \leq \mathrm{d}^2(x_k, x^*) - \mathrm{d}^2(x_k, x_{k+1}). \tag{14}$$

Hence $\{x_k\}$ is Fejér monotone w.r.t. $\mathcal{S}$ and bounded, $\lim_k \mathrm{d}(x_k, x^*)$ exists, $\sum_k \mathrm{d}^2(x_k, x_{k+1}) < \infty$ and

$$\lim_{k\to\infty} \mathrm{d}(x_k, x_{k+1}) = 0. \tag{15}$$

Lemma (Lower bound on the projection step)

$$\mathrm{d}(x_k, x_{k+1}) \;\geq\; \frac{\langle F(z_k), \exp_{z_k}^{-1} x_k \rangle}{\|F(z_k)\|} \;\geq\; \sigma\, \alpha_k^2 \, \|d_k\|^2, \qquad \textit{so} \quad \mathrm{d}(x_k, z_k) = \alpha_k \|d_k\| \to 0. \tag{16}$$

Lemma (Bounds)

Let A1–A3 hold and $\|F_k\| \geq \varepsilon > 0$ for all k. Then $\|F_k\| \leq \kappa$, $\|F(z_k)\| \leq \bar{\kappa}$, $\|d_k\| \leq M$, and

$$\alpha_k \;\geq\; \min\Big\{\alpha_{\min},\; \frac{\rho\,\varepsilon^2}{(\sigma\bar{\kappa}' + L)M^2}\Big\}, \qquad \bar{\kappa}' = \kappa + L\alpha_{\max}M. \tag{17}$$

Theorem (Global Convergence)

Under A1–A3, either CC-DFP terminates at a point of $\mathcal{S}$, or $\liminf_k \|F_k\| = 0$ and the whole sequence $\{x_k\}$ converges to some $\bar{x} \in \mathcal{S}$; in particular $\|F(x_k)\| \to 0$.

Proof idea: if $\|F_k\| \geq \varepsilon$, (16)–(17) give $\mathrm{d}(x_k, x_{k+1}) \geq \sigma\underline{\alpha}^2\varepsilon^2 > 0$, contradicting (15). A cluster point of a subsequence with $\|F_{k_i}\| \to 0$ lies in $\mathcal{S}$; Fejér monotonicity makes it the limit.

CC-DFP versus the Hu–Storey hybrids of Papp et al. (2024): two-term TT-HuS and original HuS. CC-DFP and TT-HuS differ *only* in the hybridisation rule, (4) versus (2).

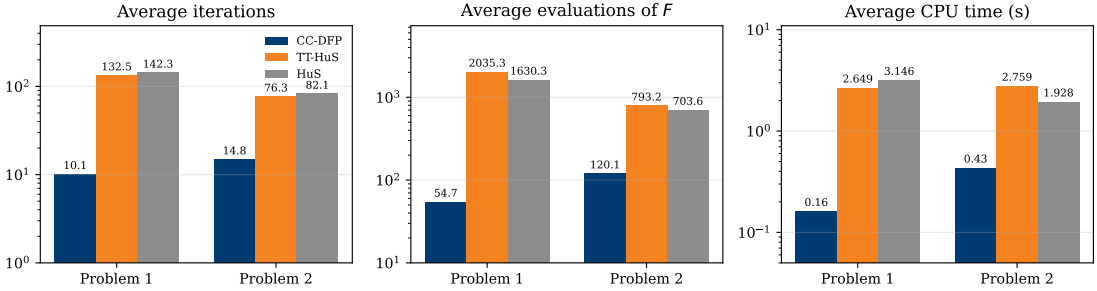
- Julia 1.11.3, Intel Core i3-6100U, 8 GB RAM; $\sigma = 0.3$, $\rho = 0.7$, $t = 10^{-8}$; stop at $\|F_k\| \le 10^{-4}$ or $\|F(z_k)\| \le 10^{-4}$, $k_{\max} = 50{,}000$.
- Ten starting points: $\pm 10 \cdot \mathbf{1}$, $\pm \mathbf{1}$, $0.1 \cdot \mathbf{1}$, $(1, \tfrac{1}{2}, \dots, \tfrac{1}{n})$, $(\tfrac{1}{n}, \dots, 1)$, $(1 - \tfrac{1}{n}, \dots, 0)$, $(\sin 1, \dots, \sin n)$, $\mathcal{N}(0, I_n)$.
- Reported: iterations (IT), evaluations of F (NF), CPU seconds (CT), final residual (Res.).

Test problems

P1 (Yan–Peng–Li 2010) $F_i(x) = 2x_i - \sin|x_i|$, $n \in \{10^3,\ 2\times10^4,\ 5\times10^4\}$; strongly monotone, $L = 3$, $x^* = 0$.
P2 (Nocedal–Wright, modified) $F_i(x) = e^{x_i} - \cos x_i$, $n \in \{10^4,\ 2\times10^4,\ 5\times10^4\}$.
P3 Riemannian centre of mass on $\mathbb{H}^n$: $F(x) = -\sum_j w_j \exp_x^{-1} p_j$, $m = 20$ anchors; singularity = Karcher mean.

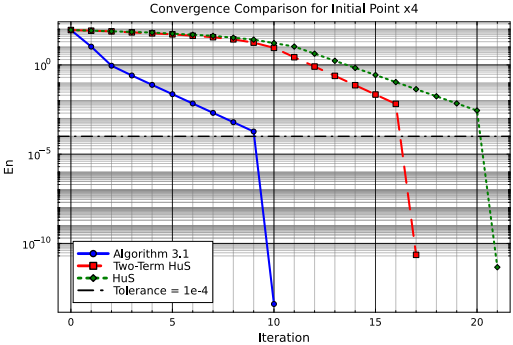
P1–P2 are on $\mathbb{R}^n$ (curvature 0) and test the hybridisation rule; P3 tests the curvature handling.

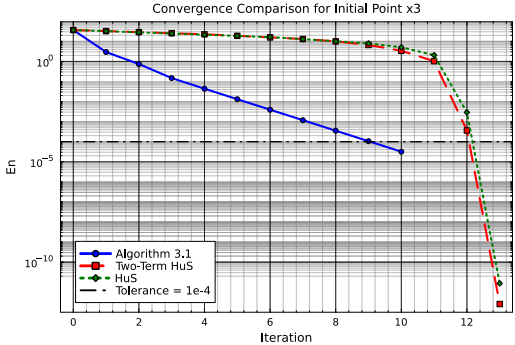
Averages over the 30 runs per problem (Tables 1–2)



- P1: CC-DFP needs 10 iterations / 55 evaluations versus 132 / 2035 for TT-HuS — factors 13 and 37; 16× faster.
- P2: factors 5 (IT), 6.6 (NF), 6.4 (CT). The entire gain is attributable to the convex-combination rule.

Problem 1: residual $\|F_k\|$ against iterations ($n = 1000$)

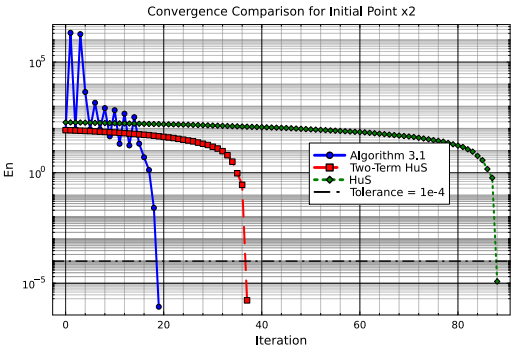


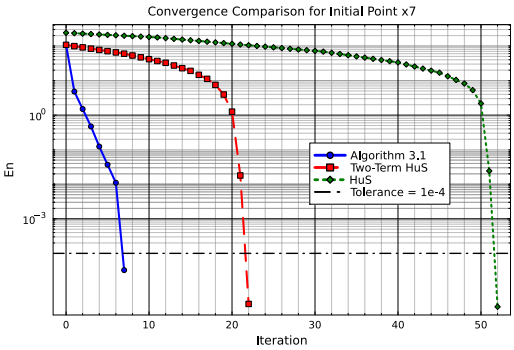


Left: $x^{(4)} = -\mathbf{1}$. Right: $x^{(3)} = \mathbf{1}$. Dash-dotted line: tolerance 10^{-4}.

- CC-DFP (Algorithm 3.1) approaches the tolerance at a roughly constant ratio; the HuS methods stall, then overshoot with a long secant step — hence their much smaller final residuals.

Problem 2: residual $\|F_k\|$ against iterations ($n = 10{,}000$)





Left: $x^{(2)} = -10 \cdot \mathbf{1}$. Right: $x^{(7)} = (\frac{1}{n}, \ldots, 1)$.

- From $x^{(2)}$ the residual of CC-DFP first *rises* and oscillates for ≈ 15 iterations, then decays fast: Fejér monotonicity governs $\mathrm{d}(x_k, \mathcal{S})$, not $\|F_k\|$.
- Far from the solution $\beta_k^{\mathrm{PRP}} < 0$ often, so Hu–Storey sets $\beta_k = 0$; the convex rule keeps the CG memory — iteration counts essentially dimension-independent ($9/13/11$ vs. $115/506/795$ from $x^{(1)}$ on P1).

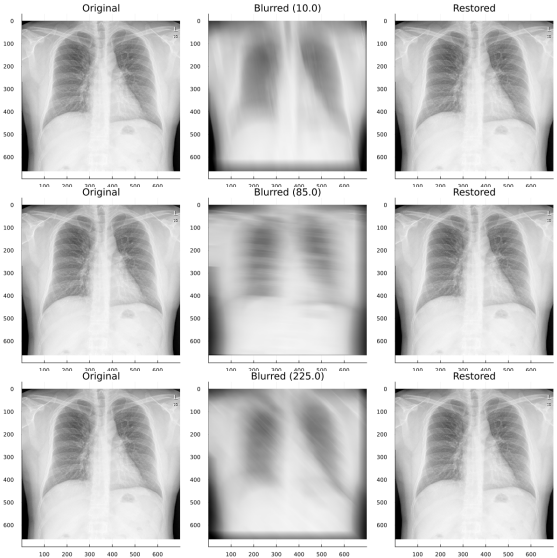
- **Nearby starts** $(0.1 \cdot \mathbf{1}, (1, \frac{1}{2}, \dots))$: HuS variants finish in 2–4 iterations, CC-DFP in 5–9. Near a solution $\beta_k^{\mathrm{PRP}} \approx 0$ and the HuS step is a well-scaled residual step. Cheap remedy: $\lambda_k = 1$ once $\|F_k\| < 10^{-2}$.
- **Final residuals.** HuS often stops at 10^{-11}–10^{-15}, CC-DFP at 10^{-5}–10^{-7}: overshooting secant step versus approach at constant ratio stopping behaviour, but not accuracy.
- **Cost.** NF/IT $\approx$ 5–8 for all methods; the extra cost of CC-DFP is three inner products.

800×864 chest X-ray (Pandey 2019), $n = 691{,}200$; motion blur of length ℓ at angle θ, $b = Kx + \eta$.

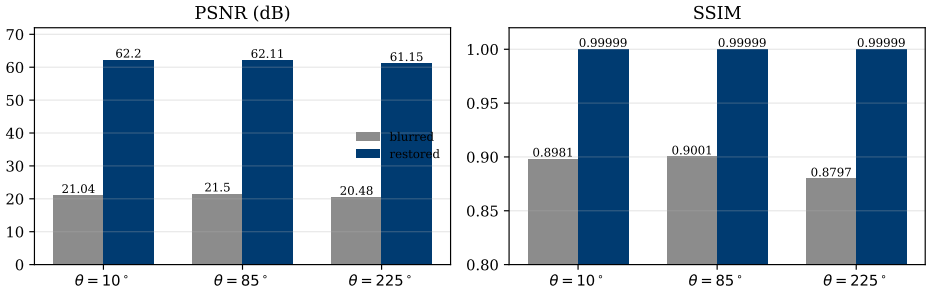
$$\min_x \ \tfrac{1}{2}\|Kx - b\|_2^2 + \tfrac{\lambda}{2}\|x\|_2^2 \quad \Longrightarrow \quad F(x) := K^\top(Kx - b) + \lambda x = 0.$$

- $K^\top K + \lambda I \succ 0$: F strongly monotone (modulus λ), Lipschitz ($\leq 1 + \lambda$) — A1–A3 hold.
- K, $K^\top$ by fast convolution, $O(n \log n)$ per evaluation; never formed as matrices.
- Metrics: MSE, RMSE, NRMSE, PSNR $= 10 \log_{10}(1/\text{MSE})$, SSIM ($11 \times 11$ Gaussian window).
- Three blurs, $\ell = 55$ pixels: $\theta = 10^\circ, 85^\circ, 225^\circ$.

tion


Original
Blurred (10.0)
Restored
Original
Blurred (85.0)
Restored
Original
Blurred (225.0)
Restored

Restoration metrics (Table 5)



- RMSE reduced by more than 99% ($\approx 0.09 \to 0.0008$); PSNR above 61 dB, SSIM indistinguishable from 1; stabilised within 300 iterations.
- Diagonal blur is the hardest case, yet only 1 dB behind after restoration; rib contours and fine bronchovascular markings recovered without ringing.
- Caveats: true kernel and low noise (non-blind); the HuS comparison remains to be run on this $n \approx 7 \times 10^5$ problem.

- Z. Papp, S. Rapajić, A. H. Ibrahim, S. Phiangsungnoen, Hybrid Hu–Storey type methods for large-scale nonlinear monotone systems and signal recovery, *J. Inequal. Appl.* 2024.
- S. S. Djordjević, New hybrid conjugate gradient method as a convex combination of FR and PRP methods, *Filomat* 30 (2016).
- M. V. Solodov, B. F. Svaiter, A globally convergent inexact Newton method for systems of monotone equations, Kluwer, 1998.
- S. Z. Németh, Monotone vector fields, *Publ. Math. Debrecen* 54 (1999).
- C. Li, G. López, V. Martín-Márquez, Monotone vector fields and the proximal point algorithm on Hadamard manifolds, *J. London Math. Soc.* 79 (2009).
- D. R. Sahu, S. Sharma, A derivative-free projection method for the singularities of vector fields with convex constraints on Hadamard manifolds, *Optim. Methods Softw.* (2024).
- T. T. Yao, Z. Zhao, Z. J. Bai, X. Q. Jin, A Riemannian derivative-free Polak–Ribière–Polyak method for tangent vector field, *Numer. Algorithms* (2024).
- O. P. Ferreira, L. R. Lucambio Pérez, S. Z. Németh, Singularities of monotone vector fields and an extragradient-type algorithm, *J. Global Optim.* 31 (2005).
- H. A. Abass, Linear convergence of alternating inertial Tseng-type method for solving inclusion problems on Hadamard manifolds, *Proc. Edinburgh Math. Soc.* 68 (2025).

Thank you for your time and attention

aremukazeemolalekan@gmail.com