Computation of Ground States of the Gross-Pitaevskii Functional via Riemannian Optimization

Bartosz Protas

Department of Mathematics & Statistics
McMaster University, Hamilton, Ontario, Canada
URL: http://www.math.mcmaster.ca/bprotas

24th Midwest Optimization Meeting
The University of Waterloo, October 28-29, 2022

- Collaborator: Ionut Danaila *(Université de Rouen)*

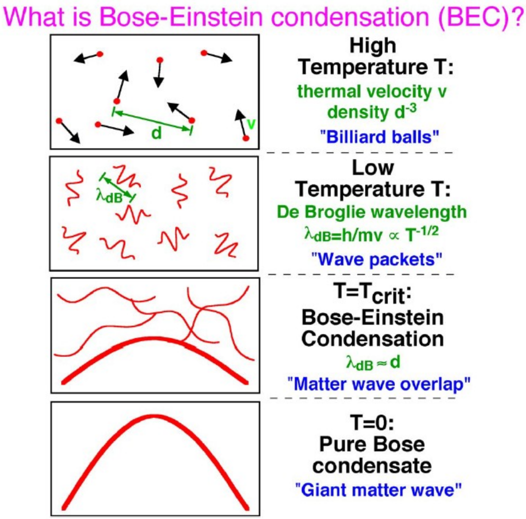
- Acknowledgments:
  - Thematic Program *"Multiscale Scientific Computing: From Quantum Physics and Chemistry to Material Science and Fluid Mechanics"* at the Fields Institute in Toronto (January–April 2016)
  - Visiting Professorship at Université de Rouen in 2017

SIAM J. Sci. Comput.
Vol. 39, No. 6, pp. B1102–B1129

COMPUTATION OF GROUND STATES OF THE GROSS–PITAEVSKII FUNCTIONAL VIA RIEMANNIAN OPTIMIZATION*

IONUT DANAILA† AND BARTOSZ PROTAS‡

Agenda



W. Ketterle, Collége de France, 2005

First experimental realization: Wieman & Cornell (1995; 2001 Nobel Prize)

- Gross-Pitaevskii Free Energy Functional (non-dimensional form)

$$E(u) = \int_{\mathcal{D}} \left[\frac{1}{2} |\nabla u|^2 + C_{\text{trap}} \, |u|^2 + \frac{1}{2} C_g |u|^4 - i C_\Omega \, u^* A^t \cdot \nabla u \right] \, d\mathbf{x},$$

$$\|u\|_2^2 = \int_{\mathcal{D}} |u(\mathbf{x})|^2 \, d\mathbf{x} = 1, \qquad \mathcal{D} \subseteq \mathbb{R}^d$$

where

$$u = \frac{\psi}{\sqrt{N} \, x_s^{-d/2}}, \qquad \psi \text{ — wavefunction}, \quad \psi \, : \, \mathcal{D} \to \mathbb{C}$$

N — number of atoms in the condensate

x_s — characteristic length scale

$A^t = [y, -x, 0]$, $\qquad C_{\text{trap}}(x, y, z)$ — trapping potential

C_g, C_Ω — constants

- C_Ω characterizes the effect of rotation

- Sobolev space $E : H_0^1(\mathcal{D}) \to \mathbb{R}$

$$H_0^1(\mathcal{D}) := \left\{ u : \mathcal{D} \to \mathbb{C} \mid \int_{\mathcal{D}} |u| + |\nabla u|^2 \, d\mathbf{x} < \infty, \quad u = 0 \text{ on } \partial\mathcal{D} \right\}$$

- Variational optimization

$$\min_{u \in H_0^1(\mathcal{D})} E(u)$$
$$\text{subject to } \|u\|_{L_2(\mathcal{D})} = 1$$

- Minimizers constrained to a nonlinear manifold $\mathcal{M}$ in $H_0^1(\mathcal{D})$

$$\mathcal{M} := \left\{ u \in H_0^1(\mathcal{D}) \; : \; \|u\|_{L_2(\mathcal{D})} = 1 \right\}$$

- Computational approaches:
  - Euler-Lagrange equation for $E(u) \implies$ nonlinear eigenvalue problem
  - Direct minimization of $E(u)$ via a gradient method

- Steepest-gradient approach

$$\begin{aligned} u^{(n+1)} &= u^{(n)} - \tau_n \nabla E\big(u^{(u)}\big), && n = 0, 1, \dots, \\ u^{(0)} &= u_0, && \text{(initial guess)}, \end{aligned}$$

where:

$$\tilde{u} = \lim_{n\to\infty} u^{(n)} \qquad \text{— the minimizer ("ground state")}$$

$$\nabla E\big(u^{(u)}\big) \qquad \text{— gradient of } E(u) \text{ at } u^{(n)}$$

$$\tau_n = \operatorname{argmin}_{\tau>0} E\Big(u^{(n)} - \tau \nabla E\big(u^{(u)}\big)\Big) \qquad \text{— optimal step size}$$

- Key issues:
  - Regularity of the minimizers $\tilde{u} \in H_0^1(\mathcal{D}) \implies$ Sobolev gradients
  - Enforcement of the constraint $\tilde{u} \in \mathcal{M} \implies$ Riemannian optimization

- Gâteaux differential of the Gross-Pitaevskii Energy Functional

$$E'(u;v) = \lim_{\epsilon\to 0} \epsilon^{-1}\left[E(u+\epsilon v) - E(u)\right], \qquad u, v \in \mathcal{X}$$

$\mathcal{X}$ — some function space

- Riesz Representation Theorem:

$E'(u;\cdot)$ bounded linear functional on $\mathcal{X}$

$$\Longrightarrow \quad \forall_{v\in\mathcal{X}}\ E'(u;v) = \left\langle \nabla^{\mathcal{X}} E(u), v\right\rangle_{\mathcal{X}}$$

- Relevant inner products (Danaila & Kazemi 2010)

$$\langle u, v\rangle_{L_2} = \int_{\mathcal{D}} \langle u, v\rangle \, d\mathbf{x}, \qquad \text{where } \langle u, v\rangle = uv^*$$

$$\langle u, v\rangle_{H^1} = \int_{\mathcal{D}} \langle u, v\rangle + \langle \nabla u, \nabla v\rangle \, d\mathbf{x}$$

$$\langle u, v\rangle_{H_A} = \int_{\mathcal{D}} \langle u, v\rangle + \langle \nabla_A u, \nabla_A v\rangle \, d\mathbf{x}, \quad \nabla_A = \nabla + iC_\Omega A^t$$

- Different Sobolev gradients ($\mathcal{X} = L_2, H^1, H_A$)

$$E'(u; v) = \Re \left\langle \nabla^{L^2} E(u), v \right\rangle_{L^2} = \Re \left\langle \nabla^{H^1} E(u), v \right\rangle_{H^1} = \Re \left\langle \nabla^{H_A} E(u), v \right\rangle_{H_A}$$

- The L_2 gradient

$$\nabla^{L^2} E(u) = 2 \left(-\frac{1}{2} \nabla^2 u + C_{\text{trap}} u + C_g |u|^2 u - i C_\Omega A^t \cdot \nabla u \right),$$

- The Sobolev gradient $G = \nabla^{H_A} E(u)$ obtained from the L_2 gradient via an elliptic boundary-value problem (Danaila & Kazemi 2010)

$$\forall_{v \in H_0^1(\mathcal{D})} \int_{\mathcal{D}} \left[\left(1 + C_\Omega^2 (x^2 + y^2)\right) G v + \nabla G \cdot \nabla v - 2 i C_\Omega A^t \cdot \nabla G v \right] d\mathbf{x}$$
$$= \int_{\mathcal{D}} \frac{1}{2} \nabla u \cdot \nabla v + \left[C_{\text{trap}} u + C_g |u|^2 u - i C_\Omega A^t \cdot \nabla u \right] v \, d\mathbf{x}$$

- Riemannian Optimization — an "intrinsic" approach with optimization performed directly on the manifold $\mathcal{M}$ without reference to the embedding space $H_0^1(\mathcal{D})$
  - optimization problem becomes *unconstrained*
  - can apply more efficient optimization algorithms (conjugate gradients, Newton's method)
- Riemannian structure at various levels:
  - retraction back to the constraint manifold $\Longleftarrow$
  - vector transport along the constraint manifold $\Longleftarrow$
  - Riemannian metric on the constraint manifold
- Here the formulation made simple by the constraint $\|u\|_{L_2(\mathcal{D})} = 1$
- Reference: P.-A. Absil, R. Mahony and R. Sepulchre, *"Optimization Algorithms on Matrix Manifolds"*, Princeton University Press, (2008).

- Projection of the gradient G on the tangent subspace $\mathcal{T}_u\mathcal{M}$

$$P_{u_n,H_A}G = G - \frac{\Re\left(\langle u_n, G\rangle_{L^2}\right)}{\Re\left(\langle u_n, v_{H_A}\rangle_{L^2}\right)} v_{H_A}, \quad \text{where}$$

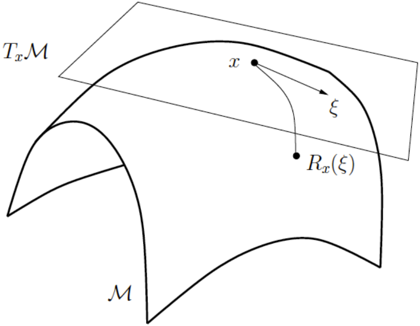
$$\langle v_{H_A}, v\rangle_{H_A} = \langle u_n, v\rangle_{L^2}, \ \forall v \in H_A$$

- There is some freedom in choosing the subtracted field (v_{H_A})
- Approach equivalent to constraint enforcement via Lagrange multipliers
  - Error in constraint satisfaction $\mathcal{O}(\tau_n)$

- **Retraction**

$$\mathcal{R}_u \ : \ \mathcal{T}_u\mathcal{M} \to \mathcal{M}$$

maps a tangent vector $\xi \in \mathcal{T}_u\mathcal{M}$ back to the manifold $\mathcal{M}$



- For our constraint manifold $\mathcal{M}$

$$\mathcal{R}_u(\xi) = \frac{u+\xi}{\|u+\xi\|_{L_2(\mathcal{D})}}$$

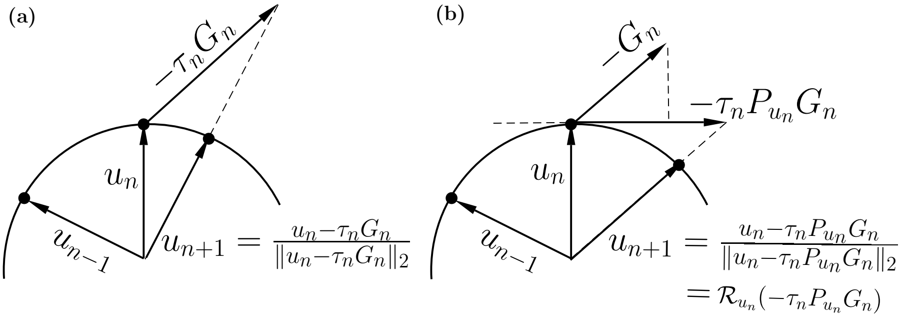
retraction = normalization

- Riemannian steepest descent approach

$$u_{n+1} = \mathcal{R}_{u_n}\left(\tau_n P_{u_n,H_A} G(u_n)\right), \qquad n = 0,1,2,\dots$$
$$u_0 = u^0$$

where

$$\tau_n = \operatorname{argmin}_{\tau>0} E\left(\mathcal{R}_{u_n}(\tau P_{u_n,H_A} G(u_n))\right)$$



(a) The simple ("unprojected") gradient method.

(b) The projected gradient (PG) method.

- Consider $\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x})$, where $f \; : \; \mathbb{R}^N \rightarrow \mathbb{R}$
- Nonlinear Conjugate Gradients Method

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \tau_n \, \mathbf{d}_n, \qquad n = 0, 1, \ldots$$
$$\mathbf{x}_0 = \mathbf{x}^0$$

  - descent direction $\mathbf{d}_n$ is defined as

$$\mathbf{d}_n = -\mathbf{g}_n + \beta_n \, \mathbf{d}_{n-1}, \qquad n = 1, 2, \ldots$$
$$\mathbf{d}_0 = -\mathbf{g}_0, \qquad \mathbf{g}_n = \boldsymbol{\nabla} f(\mathbf{x}_n)$$

  - "momentum" coefficients β_n ensure conjugacy of decent directions

$$\beta_n = \beta_n^{FR} := \frac{\langle \mathbf{g}_n, \mathbf{g}_n \rangle_{\mathcal{X}}}{\langle \mathbf{g}_{n-1}, \mathbf{g}_{n-1} \rangle_{\mathcal{X}}} \qquad \text{(Fletcher-Reeves)},$$
$$\beta_n = \beta_n^{PR} := \frac{\langle \mathbf{g}_n, (\mathbf{g}_n - \mathbf{g}_{n-1}) \rangle_{\mathcal{X}}}{\langle \mathbf{g}_{n-1}, \mathbf{g}_{n-1} \rangle_{\mathcal{X}}} \qquad \text{(Polak-Ribiére)}$$

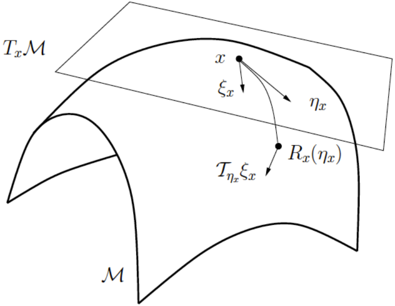
- In the Riemannian setting

$$\mathbf{g}_{n-1}, \mathbf{d}_{n-1} \in \mathcal{T}_{\mathbf{x}_{n-1}} \quad \text{and} \quad \mathbf{g}_n, \mathbf{d}_n \in \mathcal{T}_{\mathbf{x}_n},$$

  hence cannot be added or multiplied ...

- Need a mapping between the tangent spaces $\mathcal{T}_{u_{n-1}}\mathcal{M}$ and $\mathcal{T}_{u_n}\mathcal{M}$

- Vector transport $\mathcal{T}_\eta(\xi) \; : \; \mathcal{TM} \times \mathcal{TM} \to \mathcal{TM}, \quad \xi, \eta \in \mathcal{TM}$ describing how the vector field ξ is transported along the manifold $\mathcal{M}$ by the field η



- For our constraint manifold $\mathcal{M}$:
  - vector transport via differentiated retraction

$$\mathcal{T}_{\eta_x}(\xi_x) = \frac{d}{dt}\mathcal{R}_x(\eta_x + t\xi_x)\Big|_{t=0} = \frac{1}{\|x+\eta_x\|}\left[Id - \frac{(x+\eta_x)(x+\eta_x)^T}{\|x+\eta_x\|^2}\right]\xi_x$$

  - vector transport on Riemannian submanifolds ("parallel" transport)

$$\mathcal{T}_{\eta_x}(\xi_x) = P_{\mathcal{R}_x(\eta_x)}\xi_x = \left[Id - \frac{(x+\eta_x)(x+\eta_x)^T}{\|x+\eta_x\|^2}\right]\xi_x$$

- The two definitions differ by a scalar factor only

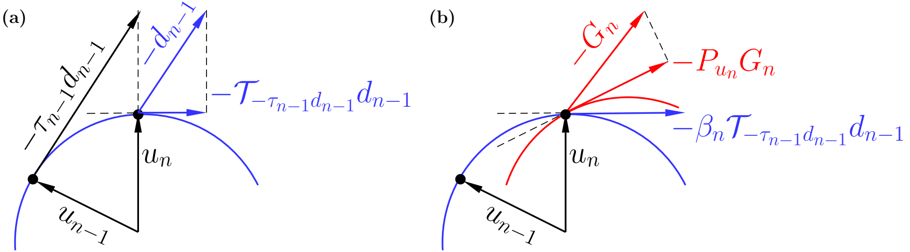
- Riemannian Conjugate Gradients

$$u_{n+1} = \mathcal{R}_{u_n}(\tau_n d_n), \qquad n = 0, 1, \dots$$
$$u_0 = u^0, \qquad \text{where}$$

$$d_n = -P_{u_n,H_A} G(u_n) + \beta_n \mathcal{T}_{-\tau_{n-1} d_{n-1}}(d_{n-1}), \qquad n = 1, 2, \dots$$
$$d_0 = -P_{u_0,H_A} G$$
$$\beta_n = \frac{\left\langle P_{u_n,H_A} G(u_n), \left(P_{u_n,H_A} G(u_n) - \mathcal{T}_{-\tau_{n-1} d_{n-1}} P_{u_n,H_A} G(u_{n-1})\right) \right\rangle_{H_A(\mathcal{D})}}{\left\langle P_{u_n,H_A} G(u_{n-1}), P_{u_n,H_A} G(u_{n-1}) \right\rangle_{H_A(\mathcal{D})}}$$

(Polak-Ribiére)

- Approach straightforward to implement

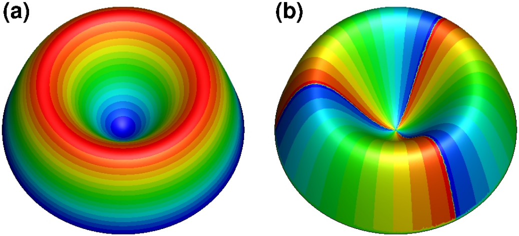


(a) Riemannian vector transport of the anterior conjugate direction d_{n-1}; the transport of the anterior gradient G_{n-1} is performed in a similar way.

(b) Projection of the new Sobolev gradient G_n onto the tangent subspace $\mathcal{T}_{u_n}\mathcal{M}$ resulting in $P_{u_n,H_A}G_n$.

- Implementation in `FreeFEM++`:
  - P^2 (piecewise quadratic) finite elements used to approximate the solution u
  - P^4 (piecewise quartic) finite elements used to represent the nonlinear terms in the gradients
- Discretization of domain $\mathcal{D}$
  - fixed triangulation
    - Mesh I: 24,454 triangles with $h_{min} = 0.0118$
    - Mesh II: 99,329 triangles with $h_{min} = 0.0059$
  - Adaptive mesh refinement (Danaila & Hecht, 2010)
- Arc-search for optimal using Brent's method $\quad \tau_n = \operatorname{argmin}_{\tau>0} E\left(\mathcal{R}_{u_n}(-\tau d_n)\right)$

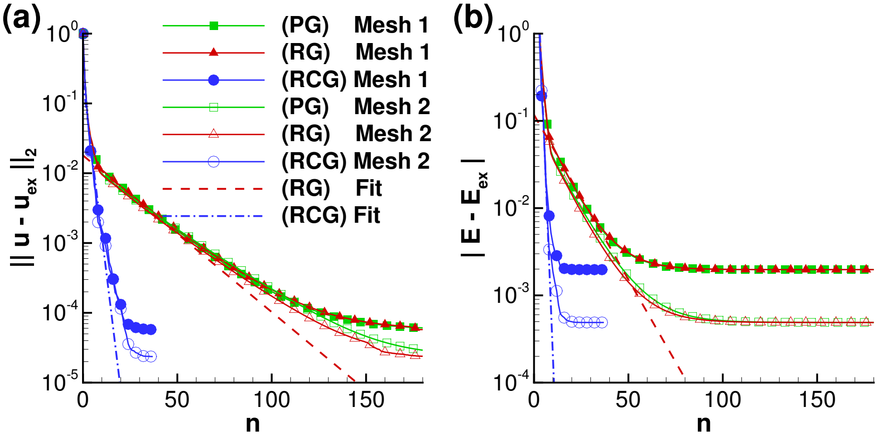
$$u_{ex}(x,y) = U(r)\exp(im\theta), \qquad U(r) = \frac{2\sqrt{21}}{\sqrt{\pi}}\,\frac{r^2\,(R-r)}{R^4}, \quad m \in \mathbb{N}$$



3D-rendering of the modulus $|u_{ex}|$ color-coded with

(a) the modulus itself,

(b) the modulus itself and (b) the phase of the solution for $m = 3$.



Fits: $\|u_n - u_{ex}\|_2 \sim B_u\, A_u^n, \qquad |E_n - E_{ex}| \sim B_e\, A_e^n$

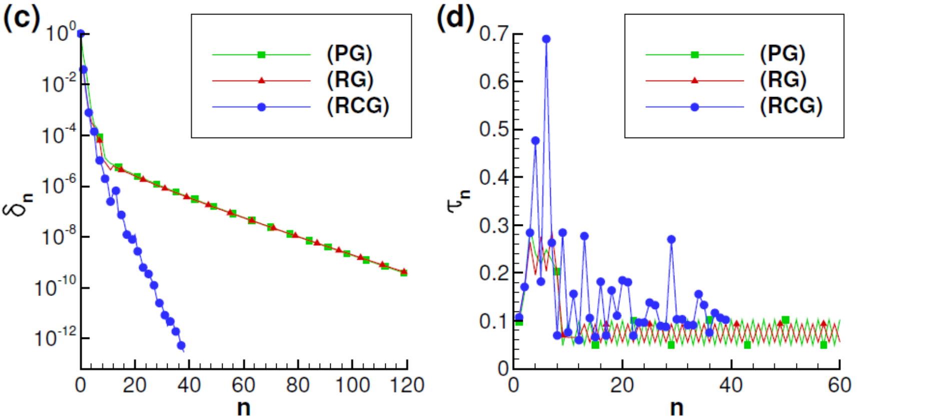
- Constants A_e and A_u $\quad (A_u \approx \sqrt{A_e})$

	Mesh 1			Mesh 2		
	A_e	$\sqrt{A_e}$	A_u	A_e	$\sqrt{A_e}$	A_u
(RG)	0.9167	0.9574	0.9496	0.9268	0.9627	0.9538
(RCG)	0.2909	0.5394	0.5275	0.2924	0.5408	0.5238

- Relation to the “condition number” κ (Euclidean case)
  - simple gradients: $A_u = (\kappa - 1)/(\kappa + 1)$
  - conjugate gradients: $A_u = (\sqrt{\kappa} - 1)/(\sqrt{\kappa} + 1)$

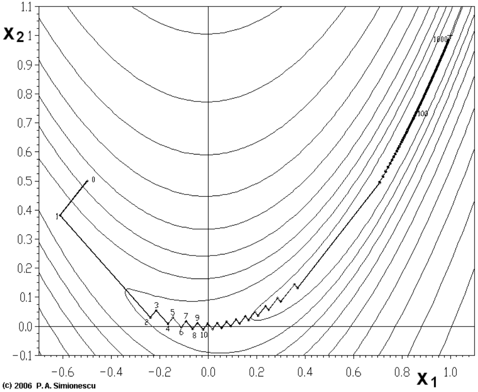
- Estimate κ from A_u
  - RG: $\kappa \approx 42.37$
  - RCG: $\kappa \approx 3.2$

- Speed-up in the Riemannian Conjugate Gradient approach exceeds the theoretical prediction!


(c)
(PG)
(RG)
(RCG)
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
δ_n
0
20
40
60
80
100
120
n
(d)
(PG)
(RG)
(RCG)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
τ_n
0
20
40
60
n

The step size τ_n in the Projected Gradient (PG) and Riemannian Gradient (RG) methods exhibits oscillatory behavior

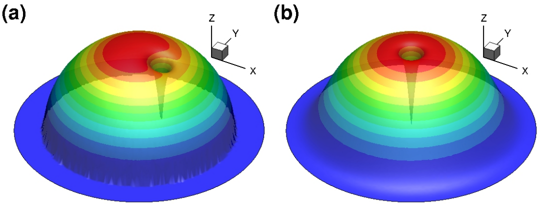
$\Longrightarrow$ iterates u_n trapped in long narrow "valleys"



steepest descent for the "banana function" (from Wikipedia)

BEC trapped in a harmonic potential and rotating at low angular velocities

$$C_{\text{trap}} = r^2/2, \quad C_g = 500, \quad C_\Omega = 0.4$$



3D rendering of the atomic density $\rho = |u|^2$ for:

(a) the initial guess u_0 (Thomas-Fermi approximation)

(b) the converged ground state.

- For comparison, semi-implicit backward Euler (BE) method to solve the normalized gradient flow

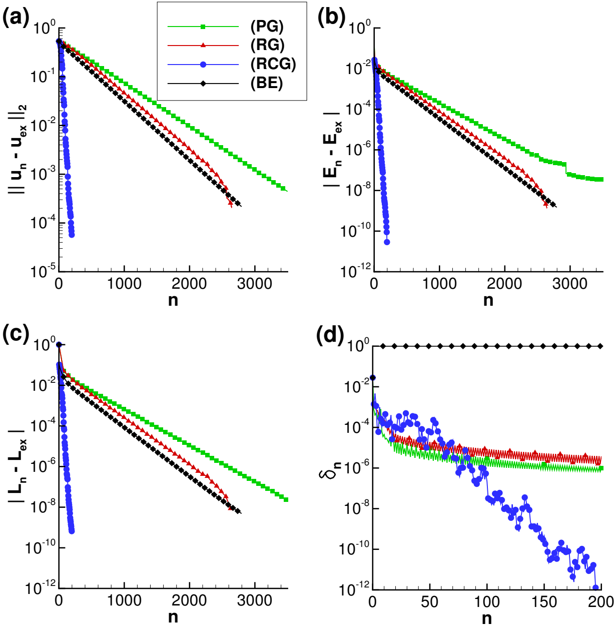
$$\frac{\tilde{u} - u_n}{\delta t} = \frac{1}{2}\nabla^2 \tilde{u} - C_{\text{trap}}\tilde{u} - C_g |u_n|^2 \tilde{u} + i C_\Omega A^t \cdot \nabla \tilde{u}$$

$$u_{n+1} = \frac{\tilde{u}(t_{n+1})}{\|\tilde{u}(t_{n+1})\|_2}.$$

- Additional diagnostic quantities

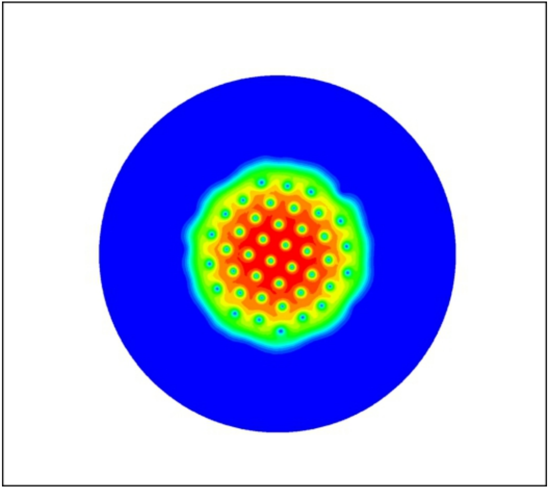
angular momentum: $$L = i \int_{\mathcal{D}} u^* A^t \cdot \nabla u \, d\mathbf{x}$$

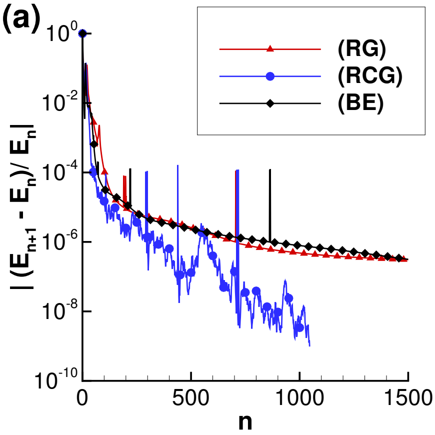
drift away from the constraint manifold: $$\delta_n = \left| 1 - \|\hat{u}_n\|_{L^2(\mathcal{D})} \right|$$

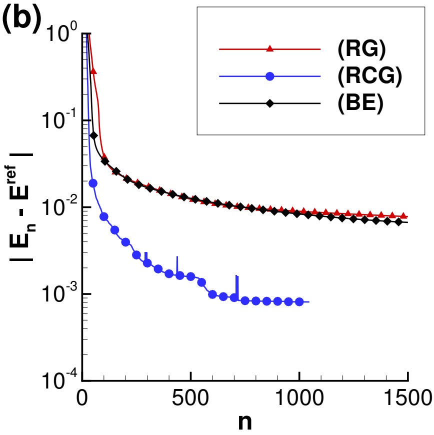

(a)
(b)
(c)
(d)
(PG)
(RG)
(RCG)
(BE)
$|| u_n - u_{ex} ||_2$
$| E_n - E_{ex} |$
$| L_n - L_{ex} |$
δ_n
n
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
0
1000
2000
3000
50
100
150
200

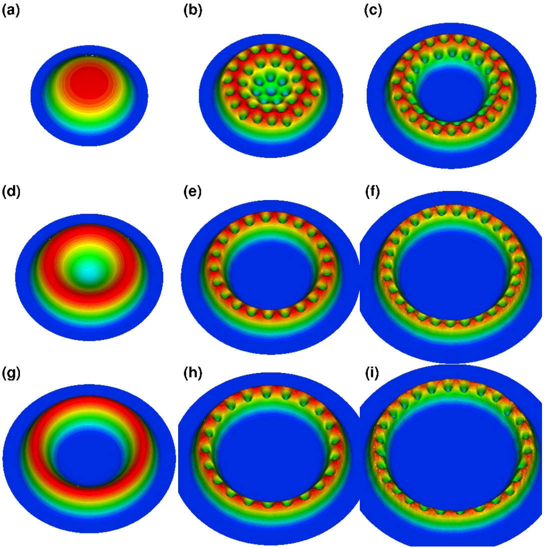
Evolution of $|\phi|$ with iterations

Riemannian Conjugate-Gradient (RCG) Approach with Adaptive Grid Refinement




(a)
(RG)
(RCG)
(BE)
$|(E_{n+1} - E_n)/E_n|$
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
0
500
1000
1500
n


(b)
(RG)
(RCG)
(BE)
$|E_n - E^{ref}|$
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
500
1000
1500
n


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)

Conclusions

- Riemannian approach accelerates solution of equality-constrained optimization problems (computation of ground states in BEC)
  - better performance than other first-order methods
  - comparable performance to some second-order methods (`Ipopt`, which however cannot take advantage of grid adaptation)
- Key enablers for Riemannian Conjugate Gradients:
  - projections onto $\mathcal{T}_{u_n}\mathcal{M}$
  - retractions from $\mathcal{T}_{u_n}\mathcal{M}$ onto $\mathcal{M}$,
  - vector transport between $\mathcal{T}_{u_{n-1}}$ and $\mathcal{T}_{u_n}$
- Ongoing work:
  - Riemannian metric on the constraint manifold
  - Riemannian Newton's method