A Sobolev trust-region method for numerical solution of the Ginzburg-Landau equations

Robert J. Renka Parimah Kazemi

Department of Computer Science & Engineering
University of North Texas

June 6, 2012

Outline of Talk

- Sobolev Gradient
- Nonlinear Least Squares
- Preconditioned Gradient Descent
- Sobolev Trust Region Method
- Ginzburg-Landau Equations and Test Results

Sobolev Gradient

Suppose ϕ is a real-valued C^1 function on a Hilbert space H so that for each $u \in H$ the Fréchet derivative of ϕ at u is a bounded linear functional on H. Then, by the Riesz Representation theorem, there is a unique element of H, termed the Sobolev gradient $\nabla_S \phi(u)$ such that

$$\phi'(u)\delta u = \langle \delta u, \nabla_S \phi(u) \rangle_H \quad \forall \delta u \in H.$$

Note that

$$\|\nabla_S \phi(u)\|_H = \sup_{\|\delta u\|_H = 1} |\phi'(u)\delta u|, \quad u \in H.$$

Gradient System

THEOREM: Suppose ϕ is a nonnegative C^1 function with a locally Lipschitz continuous gradient on a Hilbert space H. Then for each $x \in H$ there is a unique function $z : [0, \infty) \to H$ such that

$$z(0) = x, \quad z'(t) = -\nabla_S \phi(z(t)), \ t \geq 0.$$

Also, if $u = \lim_{t\to\infty} z(t)$ exists, then $\nabla_S \phi(u) = 0$.

J. W. Neuberger, Sobolev Gradients and Differential Equations, Second ed., Springer Lecture Notes in Mathematics #1670, Springer, 2010.

Finite Difference Discretization

Suppose $D : \mathbb{R}^n \to \mathbb{R}^{3m}$ is a finite difference approximation of the differential operator $\begin{pmatrix} I \\ \nabla \end{pmatrix} : H^1(\Omega) \to L^2(\Omega)^3$ for $\Omega \subset \mathbb{R}^2$. Then the discretized Sobolev gradient satisfies

$$\begin{aligned} \phi'(u)\delta u &= \langle \delta u, \nabla_S \phi(u) \rangle_S = \langle D\delta u, D\nabla_S \phi(u) \rangle_{\mathbb{R}^{3m}} \\ &= \langle \delta u, D^t D \nabla_S \phi(u) \rangle_{\mathbb{R}^n} \quad \forall \delta u \in \mathbb{R}^n. \end{aligned}$$

The Fréchet derivative is also represented by a Euclidean gradient $\nabla\phi(u)$ in the Euclidean inner product:

$$\phi'(u)\delta u = \langle \delta u, \nabla \phi(u) \rangle_{\mathbb{R}^n}.$$

The gradients are thus related by the preconditioner $D^t D = I - \Delta$:

$$\nabla_S \phi(u) = (D^t D)^{-1} \nabla \phi(u).$$

Finite Element Discretization

For $\Omega \subset \mathbb{R}^2$ define $D : H^1(\Omega) \to L^2(\Omega)^3$ by $Du = \binom{u}{\nabla u}$. Let $u^h = \sum u_j \psi_j$, where $\{\psi_j\}_{j=1}^n$ is a basis for an n-dimensional subspace $V^h \subset H^1(\Omega)$. The Fréchet derivative $\phi'(u^h)$ is usually not continuous in the L^2 norm, and ϕ therefore has no L^2 gradient. The components of the Euclidean gradient $\mathbf{e} \in \mathbb{R}^n$, however, are projections in $H^1(\Omega)$ of the Sobolev gradient $\nabla_S \phi(u^h) = \sum g_j \psi_j$ onto the basis functions: for $i = 1, \ldots, n$,

$$\begin{aligned} e_i &= \frac{\partial \phi}{\partial u_i} = \phi'(u^h)\psi_i = \langle \psi_i, \nabla_S \phi(u^h) \rangle_{H^1(\Omega)} \\ &= \langle D\psi_i, D \sum_j g_j \psi_j \rangle_{L^2(\Omega)^3} = \sum_j \langle D\psi_i, D\psi_j \rangle_{L^2(\Omega)^3} g_j. \end{aligned}$$

Thus the coefficients of the Sobolev gradient are computed from $L\mathbf{g} = \mathbf{e}$ for stiffness matrix

$$L_{ij} = \langle D\psi_i, D\psi_j \rangle_{L^2(\Omega)^3} = \int_\Omega (\psi_i \psi_j + \nabla \psi_i \cdot \nabla \psi_j)$$

Nonlinear Least Squares Problem

Suppose D is a discretized differential operator.

Energy Functional:

$$\phi(u) = F(Du) = \frac{1}{2}\langle r(Du), r(Du)\rangle$$

Euclidean Gradient:

$$\phi'(u)h = F'(Du)Dh = \langle r(Du), r'(Du)Dh\rangle \quad \Rightarrow$$

$$\nabla\phi(u) = D^t\nabla F(Du) = D^t r'(Du)^t r(Du)$$

Hessian:

$$\nabla^2\phi(u) = D^t[\nabla^2 F(Du)]D = D^t[r'(Du)^t r'(Du) + \sum_i r_i(Du) r_i''(Du)]D$$

Linear Least Squares Problem

Energy Functional:

$$\phi(u) = \frac{1}{2}\|Au - b\|^2$$

Gradient:

$$\nabla\phi(u) = A^t(Au - b)$$

Normal Equations:

$$\nabla\phi(u) = A^tAu - A^tb = 0$$

Hessian:

$$\nabla^2\phi(u) = A^tA$$

Advantages of Least Squares over Galerkin

- Symmetric positive definite efficiently solved linear systems
- No restriction on choice of finite elements
- Universal type-independent formulation
- Robustness without upwinding or artificial viscosity required for stability
- Functional serves as a built-in error indicator used both to monitor convergence and for adaptive mesh refinement
- Treats multiphysics, and allows essential boundary conditions and more general side conditions to be treated as residuals

Preconditioned Gradient Descent Iteration

Iteration:

$$u_{n+1} = u_n - \alpha_n C_n^{-1} \nabla \phi(u_n),$$

where $C_n^{-1}\nabla\phi$ is the gradient of ϕ with respect to the inner product $\langle v, w \rangle_{C_n} = \langle v, C_n w \rangle$.

1. $C_n = I$: Standard steepest descent
2. $C_n = \mathrm{diag}(\nabla^2 \phi(u_n))$: Jacobi preconditioning
3. $C_n = D^t D$: Standard Sobolev gradient descent
4. $C_n = D^t [r'(Du_n)^t r'(Du_n)] D$: Gauss-Newton
5. $C_n = \nabla^2 \phi(u_n)$: Damped Newton iteration

Preconditioner Tradeoffs

- I and D^tD are independent of u_n. The other three choices correspond to variable metric methods.
- The last three are associated with Sobolev norms, and preserve the regularity of u_n. If $r(Du) = Du - b$ they are identical.
- The Newton and Gauss-Newton iterations can achieve ultimate quadratic convergence rates. Each iterate can be computed by PCG with D^tD as preconditioner.
- All except Gauss-Newton apply to a more general energy functional

Sobolev Preconditioning of Newton Iterations

For $D : H \to K$ and $\phi(u) = F(Du)$, consider D^*D as a preconditioning operator for $g_0'(u) = D^* g_F'(Du)D$, where g_0 and g_F are L^2 gradients of ϕ and F, respectively. Denote the Sobolev gradient of ϕ by g_1. Then

$$\phi'(u)h = F'(Du)Dh = \langle Dh, g_F(Du)\rangle_K = \langle h, g_1(u)\rangle_H.$$

Hence

$$\phi''(u)kh = \langle Dh, g_F'(Du)Dk\rangle_K = \langle h, g_1'(u)k\rangle_H = \langle Dh, Dg_1'(u)k\rangle_K.$$

Suppose $g_F'(Du)$ is coercive and bounded on $R(D)$ with spectral bounds $0 < m \leq M < \infty$:

$$m \leq \frac{\langle Dh, g_F'(Du)Dh\rangle_K}{\langle Dh, Dh\rangle_K} = \frac{\langle Dh, Dg_1'(u)h\rangle_K}{\langle Dh, Dh\rangle_K} \leq M \quad \forall h \in H.$$

Sobolev Preconditioning of Newton Iterations continued

The above expression is a weak form of

$$m \leq \frac{\langle h, D^* D g_1'(u) h \rangle_{L^2}}{\langle h, D^* D h \rangle_{L^2}} \leq M,$$

where $(D^*D)g_1'(u) = g_0'(u)$. Hence $g_0'(u)$ and D^*D are spectrally equivalent and define equivalent Sobolev norms. Using a discretization of D^*D as preconditioner for Newton steps in PCG results in convergence rate

$$r = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}$$

for spectral condition number $\kappa \leq M/m$ in the Sobolev norm:

$$\|e_n\| / \|e_0\| \leq 2r^n.$$

Standard Method for Least Squares Treatment of PDE's

Rather than treating the nonlinear least squares problem of computing critical points of ϕ, the usual approach is to first linearize $r(Du)$ and then solve a linear least squares problem for the perturbation $\Delta u = u_{n+1} - u_n$. Newton's method for the linearization leads to the functional

$$\psi(\Delta u) = \frac{1}{2}\|r(Du_n) + r'(Du_n)D\Delta u\|^2.$$

The linear system $\nabla\psi(\Delta u) = 0$ is a Gauss-Newton step for a zero of $\nabla\phi$; i.e., *Newton for* $r = 0$ *is GN for* $\nabla\phi = 0$.

Our approach is more flexible. In the case of large residuals, a Newton iteration is usually required to achieve a quadratic rate of convergence.

Trust region subproblem

$$\min q(d) = \phi(u_n) + d^t \nabla \phi(u_n) + \frac{1}{2} d^t H_n d$$

$$\text{subject to } \|d\|_{C_n} = \sqrt{d^t C_n d} \leq \Delta,$$

where $H_n \approx \nabla^2 \phi(u_n)$ and Δ is a radius which is adjusted according to the ratio of actual to predicted reduction in ϕ:

$$\rho = \frac{\phi(u_n) - \phi(u_n + d)}{q(0) - q(d)}.$$

Levenberg-Marquardt Method

THEOREM: The trust region subproblem has global solution d^* iff d^* is feasible, and there exists $\lambda \geq 0$ such that

$$(H_n + \lambda C_n)d^* = -\nabla\phi(u_n)$$

and either $\lambda = 0$ or $\|d^*\|_{C_n} = \Delta$.

The trust region method is thus equivalent to a blend of a Newton-like method and a steepest descent method with λ implicitly defined by Δ.

Trust Region Norms

The shape of the trust region is defined by the norm $\|d\|_{C_n} = \sqrt{d^t C_n d}$.

- **Euclidean norm** $C_n = I$: Levenberg method
- **Hyperellipsoidal norm** $C_n = \mathrm{diag}(H_n)$: Levenberg-Marquardt method
- **Sobolev norm** $C_n = D^t D$: Levenberg-Marquardt-Neuberger (LMN) method

The other two choices for C_n are not useful. Only $C_n = D^t D$ restricts the trust region to the Sobolev space — a *Sobolev trust region*.

Ginzburg-Landau Energy Functional in 2D

$$G(\psi, A) = \frac{1}{2}\int_{\Omega} |\nabla\psi - iA\psi|^2 + |\nabla \times A - H_0|^2 + \frac{\kappa^2}{2}(|\psi|^2 - 1)^2$$

- $G =$ Gibbs free energy
- $\psi =$ complex-valued order parameter; $|\psi|^2 =$ density of superconducting electrons
- $A =$ vector potential of induced magnetic field $H = \nabla \times A$
- $H_0 =$ external magnetic field (normal to plane of Ω)
- $\kappa =$ Ginzburg-Landau parameter; $\kappa > 1/\sqrt{2} \Rightarrow$ Type II mixed state

Superconductor States

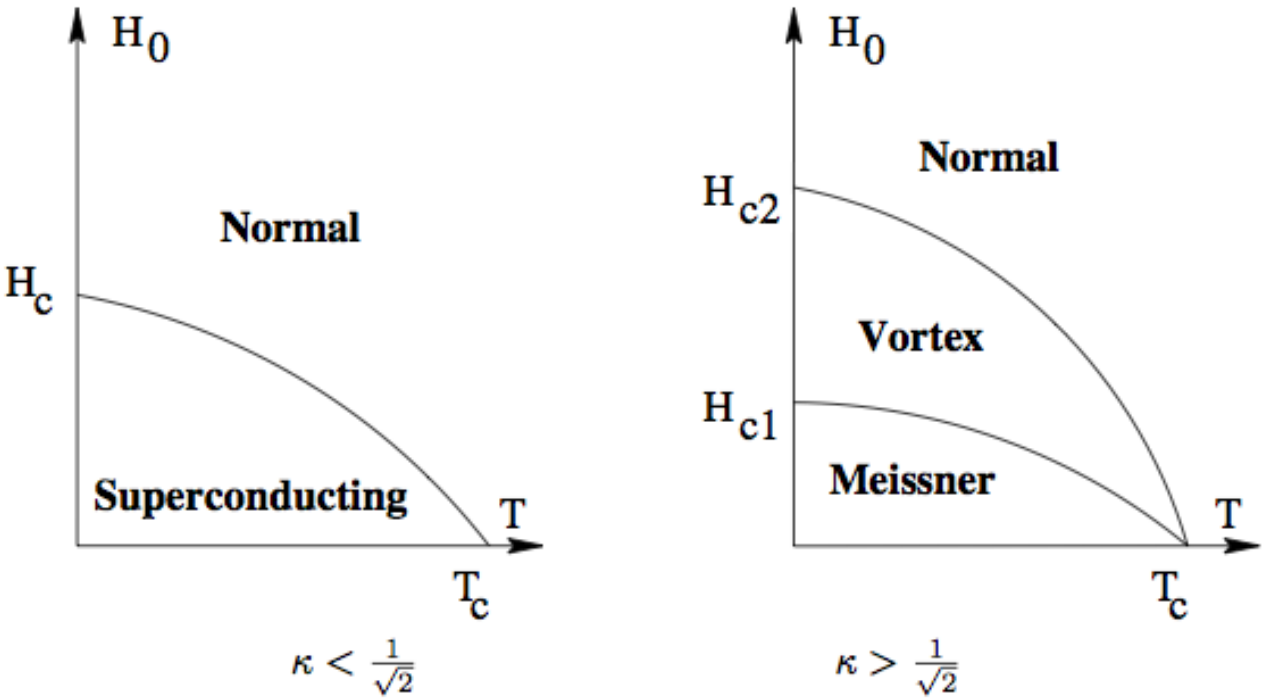


FIG. 1. *The various states of superconductors.*

Nonlinear Natural Boundary Conditions

Integration by parts applied to the first variation of G gives the natural boundary conditions

$$H - H_0 = 0, \quad (\nabla\psi - iA\psi) \cdot \mathbf{n} = 0.$$

It follows from the second condition that

$$J \cdot \mathbf{n} = 0 \text{ and } \nabla(|\psi|) \cdot \mathbf{n} = 0$$

on the boundary.

Gauge Transformation

The functional G, electron density $|\psi|^2$, magnetic field $\nabla \times A$, and current density $J = |\psi|^2(\nabla\theta - A)$ are invariant under a gauge transformation

$$\psi \to \psi e^{i\chi}, \quad A \to A + \nabla\chi$$

for a real-valued function$\chi \in H^2(\Omega)$. The Coulomb gauge constraint is $\nabla \cdot A = 0 \in \Omega$ and $A \cdot \mathbf{n} = 0$ on the boundary.

Least Squares Formulation, Real Functions

Let $H = [H^1(\Omega)]^4$, $L_1 = [L^2(\Omega)]^{12}$, and $L_2 = [L^2(\Omega)]^7$. For $(\psi, A) \in H$, with $\psi = p + iq$ and $A = \binom{a}{b}$, define $D : H \to L_1$ by $D(\psi, A) = \left(\binom{p}{\nabla p}, \binom{q}{\nabla q}, \binom{a}{\nabla a}, \binom{b}{\nabla b} \right)$, and define $r : \mathbb{R}^{12} \to \mathbb{R}^7$ by

$$r(D(\psi, A)) = \begin{pmatrix} p_1 + aq \\ q_1 - ap \\ p_2 + bq \\ q_2 - bp \\ b_1 - a_2 - H_0 \\ \frac{\kappa}{\sqrt{2}}(p^2 + q^2 - 1) \\ a_1 + b_2 \end{pmatrix},$$

Then the energy functional is

$$E(\psi, A) = \frac{1}{2} \| r(D(\psi, A)) \|_{L_2}^2.$$

Numerical Methods for Ginzburg-Landau Equations

- Discretize the Euler-Lagrange equations and boundary conditions, linearize, and solve.
- Discretize the time-dependent equations and boundary conditions, and evolve the system to steady-state.
- Discretize the energy functional and compute a zero of the Euclidean gradient.

The Sobolev gradient method is NOT a preconditioned Euler iteration for evolving the time-dependent equations. We use a Sobolev trust-region dogleg method with the Hessian modified by the Coulomb gauge constraint $A \cdot \mathbf{n} = 0$ on the boundary.

Numerical Experiments

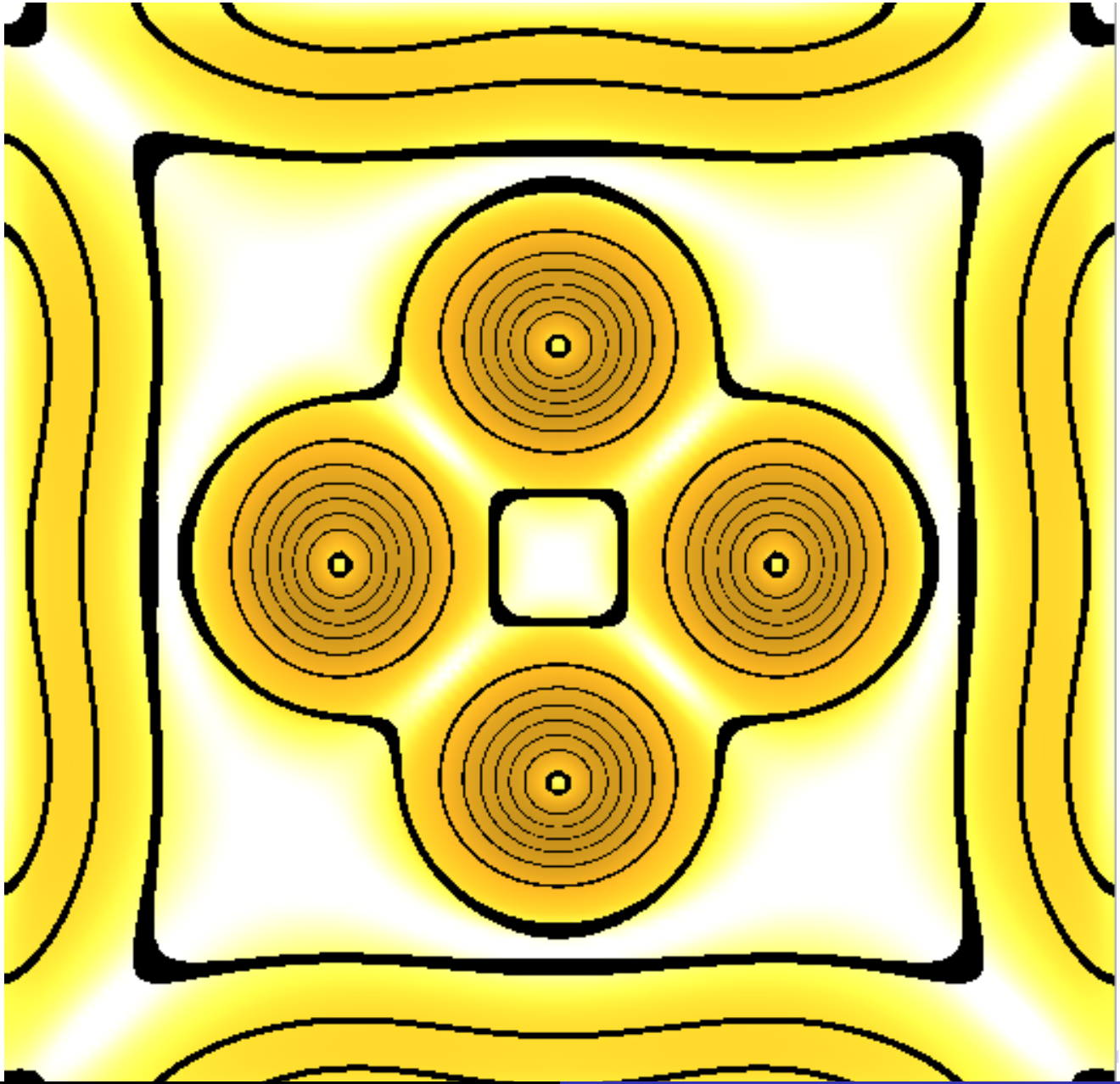
- $\Omega = [0, 5] \times [0, 5]$
- **Coarse Grid**: 65 by 65, **Fine Grid**: 129 by 129
- Initial estimate: constant functions $\psi = 1, A = 0$
- Convergence defined by upper bounds of .5e-13 on the mean absolute Euclidean gradient component and 1.e-15 on the squared trust-region radius

Iteration Counts

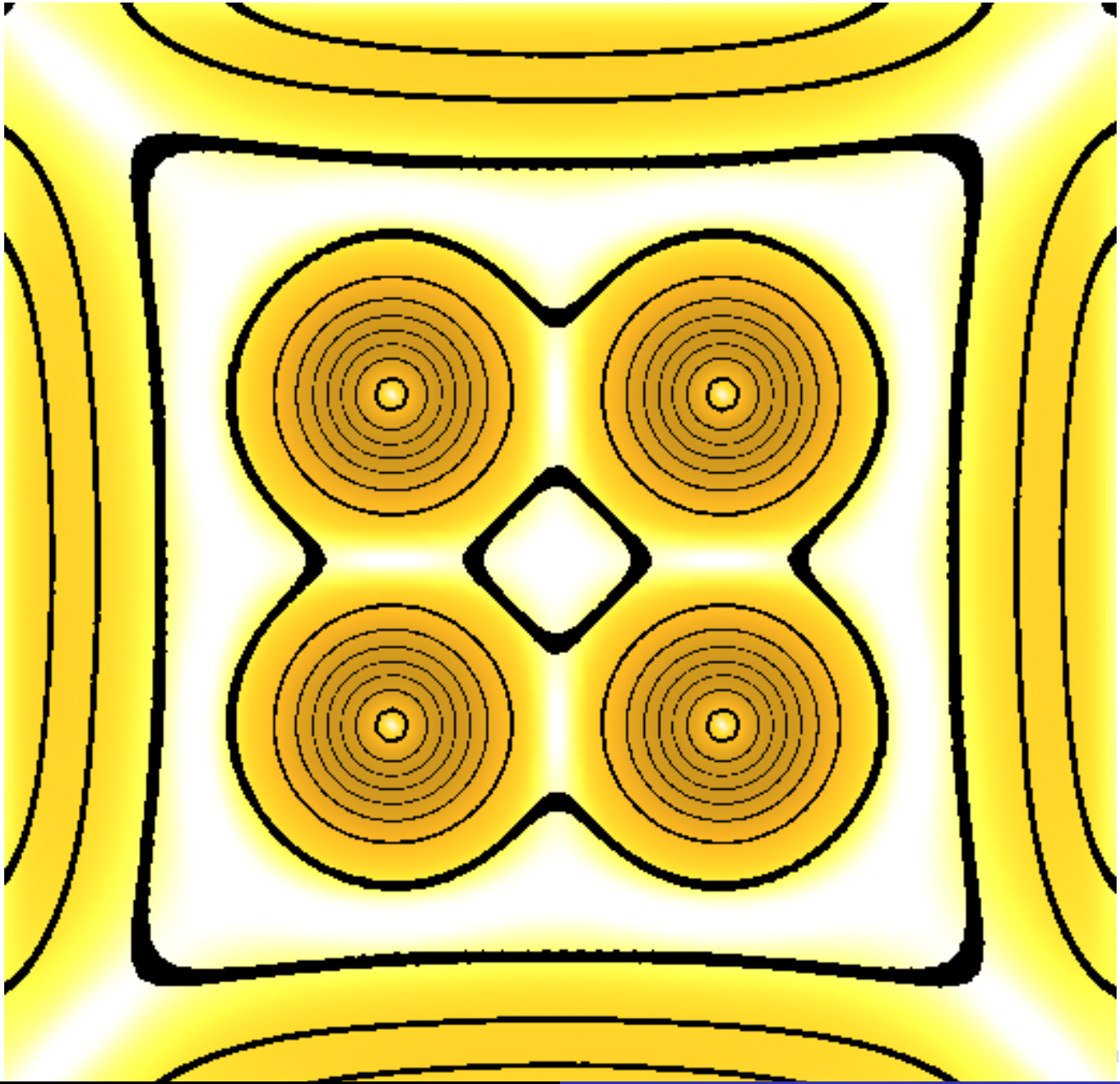
Grid	H_0	Iterations	E	$\|\nabla E\|$
Coarse	4	23	44.677	4.3e-14
Coarse	6	22	55.946	1.0e-13
Coarse	8	63	67.255	4.2e-16
Fine	4	87	44.046	2.0e-16
Fine	6	28	55.845	9.2e-11
Fine	8	88	65.858	6.4e-15

Table: Iteration counts, energy values, and errors for $\kappa = 4$.

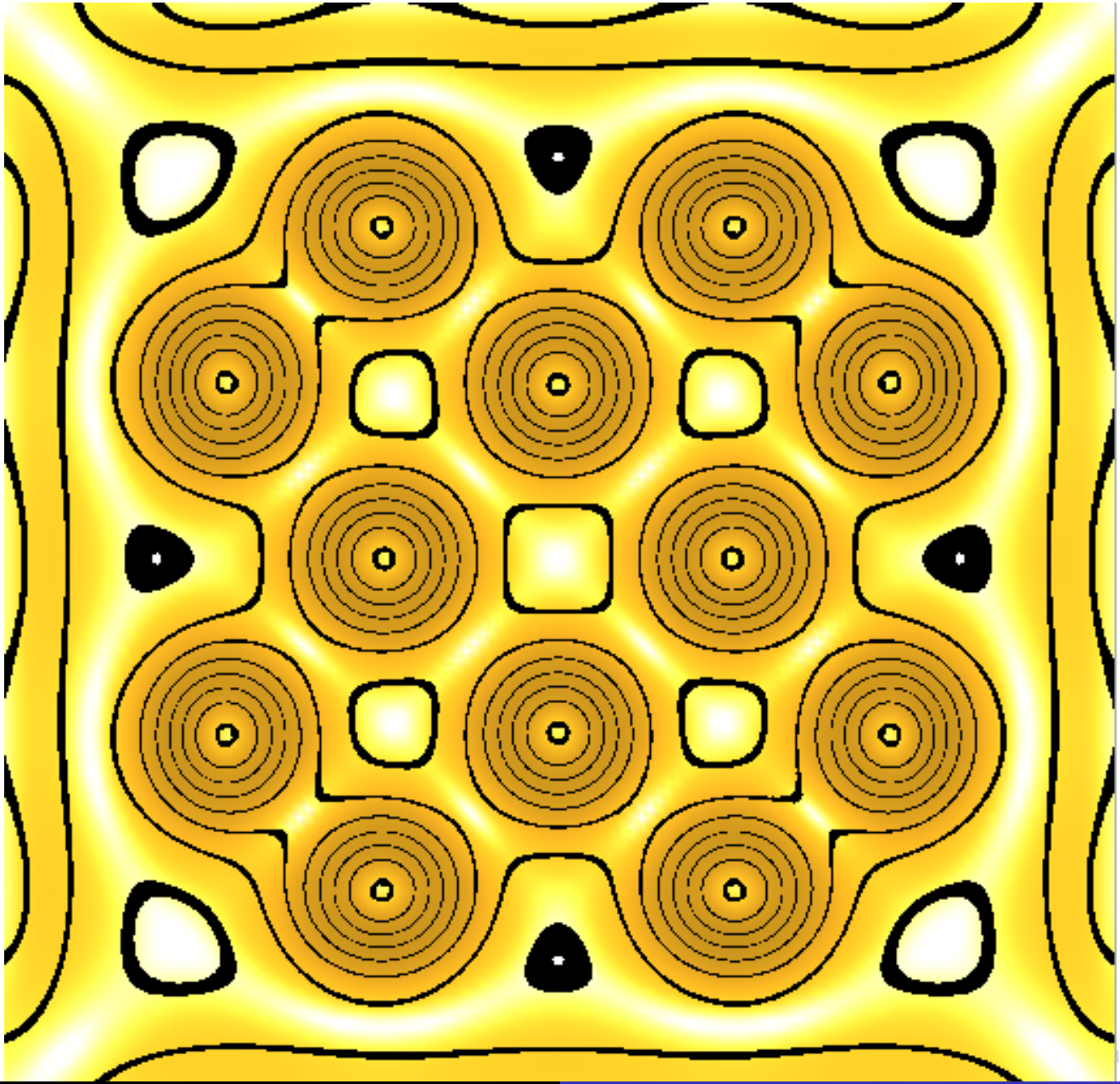
Electron Density on Coarse Grid, $\kappa = 4, H_0 = 4$



Electron Density on Fine Grid, $\kappa = 4, H_0 = 4$



Electron Density on Fine Grid, $\kappa = 4, H_0 = 6$



Electron Density on Fine Grid, $\kappa = 4, H_0 = 8$