Global Optimization:
For Some Problems, There’s HOPE

Daniel M. Dunlavy
University of Maryland, College Park
Applied Mathematics and Scientific Computation
Outline

• Problem and Existing Methods
• Homotopy Optimization Methods
• Protein Structure Prediction Problem
• Numerical Experiments
• Conclusions/Future Directions
Problem

• Solve the unconstrained minimization problem

\[ f(x^*) = \min_{x \in \mathbb{R}^n} f(x) \quad (f : \mathbb{R}^n \to \mathbb{R}) \]

• Function Characteristics
  – Solution exists, smooth (\( f \in C^2(\mathbb{R}^n, \mathbb{R}) \))
  – Complicated (multiple minima, deep local minima)
  – Good starting points unknown/difficult to compute

• Challenges
  – Finding solution in reasonable amount of time
  – Knowing when solution has been found
Some Existing Methods

• Exhaustive/enumerative search
• Stochastic search [Spall, 2003]; adaptive [Zabinsky, 2003]
• “Globalized” local search [Pinter, 1996]
• Branch and bound [Horst and Tuy, 1996]
• Genetic/evolutionary [Voss, 1999]

• Smoothing methods [Piela, 2002]
• Simulated annealing [Salamon, et al., 2002]
• Homotopy/continuation methods [Watson, 2000]
Outline

• Problem and Existing Methods
• **Homotopy Optimization Methods**
• Protein Structure Prediction Problem
• Numerical Experiments
• Conclusions/Future Directions
Homotopy Methods for Solving Nonlinear Equations

• Goal
  – Solve complicated nonlinear target system
  \[ f_1(x) = 0, \quad (f_1 : \mathbb{R}^n \rightarrow \mathbb{R}^n) \]

• Steps to solution
  – Easy template system: \( f_0(x^0) = 0 \) (\( x^0 \) known)
  – Define a continuous homotopy function:
    \[ h(x, \lambda) = \begin{cases} 
    f_0(x), & \text{if } \lambda = 0 \\
    f_1(x), & \text{if } \lambda = 1 
    \end{cases} \]
  – Example (convex): \( h(x, \lambda) = (1 - \lambda)f_0(x) + \lambda f_1(x) \)
  – Trace path of \( h(x, \lambda) = 0 \) from \( \lambda = 0 \) to \( \lambda = 1 \)
Homotopy Optimization Methods (HOM)

• Goal
  – Minimize complicated nonlinear target function
    \[ \min_{x \in \mathbb{R}^n} f_1(x), \quad (f_1 : \mathbb{R}^n \to \mathbb{R}) \]

• Steps to solution
  – Easy template function: \( f_0(x^0) = \min_{x \in \mathbb{R}^n} f_0(x) \)
  – Define a continuous homotopy function:
    \[ h(x, \lambda) = \begin{cases} 
    f_0(x), & \text{if } \lambda = 0 \\
    f_1(x), & \text{if } \lambda = 1 
    \end{cases} \]
    • Example (convex): \( h(x, \lambda) = (1 - \lambda)f_0(x) + \lambda f_1(x) \)
  – Produce sequence of minimizers of \( h(x, \lambda) \) w.r.t. \( x \)
    starting at \( \lambda = 0 \) and ending at \( \lambda = 1 \)
Illustration of HOM

\[ x^* = \min_{x \in \mathbb{R}} f_1(x) \quad h(x, \lambda) = (1 - \lambda) f_0(x) + \lambda f_1(x) \]

\[ \nabla_x h = 0 \; ; \; \nabla^2_x h > 0 \]
Homotopy Optimization using Perturbations & Ensembles (HOPE)

• Improvements over HOM
  – Produces ensemble of sequences of local minimizers of $h(x, \lambda)$ by perturbing intermediate results
  – Increases likelihood of predicting global minimizer

• Algorithmic considerations
  – Maximum ensemble size
  – Determining ensemble members
Illustration of HOPE

Constraint: ensemble size $\equiv |E| \leq 2$

$x^* = \min_{x \in \mathbb{R}} f_1(x) \quad h(x, \lambda) = (1-\lambda)f_0(x) + \lambda f_1(x)$

$\nabla_x h = 0 ; \quad \nabla^2_x h > 0$
Convergence of HOPE

Goal: \( m \in \mathbb{Z}^+ \text{ s.t. } \mathbb{P}(\exists x \in \mathbb{E}_m \mid x \in S_N) > \rho \)
**Convergence of HOPE**

\[ f_0(x) = -\sin(x) + \sin(Nx) \]

\[ f_1(x) = \sin(x) + \sin(Nx) \]
Convergence of HOPE

\[ \xi(x) = x + \delta ; \ \delta \in \mathcal{U} \left[ \frac{-3\pi}{N}, \frac{3\pi}{N} \right] \Rightarrow P = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ \cdots & \cdots & \cdots \\ 1 & 1 & 1 \end{bmatrix} \]

\[ \mathbb{E}_0 = \{ x^0 \} ; \ \mathbb{E}_k = \mathbb{E}_{k-1} \cup \{ \text{perturbed versions of } \mathbb{E}_{k-1} \} \]

No constraints on ensemble size: \[ |\mathbb{E}_m| = 2^m \]

\[ \mathcal{P}(\exists x \in \mathbb{E}_m : x \in S_N) = 1 - \prod_{k=0}^{m} \left( 1 - e_i^T P^k e_N \right)^{\binom{m}{k}} \]

\[ \geq 1 - \prod_{k=\kappa}^{m} \left( 1 - P^k_{N/2,N} \right)^{\binom{m}{k}} \quad (\kappa = \min\{i, N - i\}) \]

\[ = 1 - \prod_{k=\kappa}^{m} \left( 1 - \frac{1}{3N} \sum_{l=0}^{N-1} (-1)^l \left( 1 + 2 \cos \left( \frac{2\pi l}{N} \right) \right) \right)^{\binom{m}{k}} \]
Outline

• Problem and Existing Methods
• Homotopy Optimization Methods
• **Protein Structure Prediction Problem**
• Numerical Experiments
• Conclusions/Future Directions
Protein Structure Prediction

Given the amino acid sequence of a protein (1D), is it possible to predict its native structure (3D)?

Amino Acid Sequence

Protein Structure
Protein Structure Prediction

- **Given:**
  - Protein model
    - Molecular properties
    - Potential energy function (force field)

- **Goal:**
  - Predict lowest energy conformation
    - Native structure [Anfinsen, 1973]
  - Develop hybrid method, combining:
    - Energy minimization [numerical optimization]
    - Comparative modeling [bioinformatics]
      - Use **template** (known structure) to predict **target** structure
Protein Model: Particle Properties

• **Backbone model**
  – Single chain of particles with residue attributes
  – Particles model $C_\alpha$ atoms in proteins

  ![Backbone model diagram]

• **Properties of particles**
  – Hydrophobic, Hydrophilic, Neutral
  – Diverse hydrophobic-hydrophobic interactions

Protein Model: Energy Function

\[ E(X) = E_{bl}(X) + E_{ba}(X) + E_{dih}(X) + E_{non}(X) \]

\[ E_{bl}(X) = \sum_{i=1}^{n-1} \frac{k_r}{2} (r_{i,i+1} - \bar{r})^2 \]

\[ E_{ba}(X) = \sum_{i=1}^{n-2} \frac{k_\theta}{2} (\theta_i - \bar{\theta})^2 \]

\[ E_{dih}(X) = \sum_{i=1}^{n-3} \left[ A_i (1 + \cos \phi_i) + B_i (1 + \cos 3\phi_i) \right] \]

\[ E_{non}(X) = \sum_{i=1}^{n-3} \sum_{j=i+3}^{n} \gamma_{ij} \left\{ \alpha_{ij} \left( \frac{r}{r_{ij}} \right)^{12} - \beta_{ij} \left( \frac{r}{r_{ij}} \right)^6 \right\} \]
Homotopy Optimization Method for Proteins

• Goal
  – Minimize energy function of target protein
    \[ \min_{X \in \mathbb{R}^{3n}} E^1(X), \quad (E^1 : \mathbb{R}^{3n} \rightarrow \mathbb{R}) \]

• Steps to solution
  – Energy of template protein: \( E^0(X^0) = \min_{X \in \mathbb{R}^{3n}} E^0(X) \)
  – Define a homotopy function:
    \[ H(X, \lambda) = \rho^0(\lambda)E^0(X) + \rho^1(\lambda)E^1(X) \]
    • Deforms template protein into target protein
  – Produce sequence of minimizers of \( H(X, \lambda) \)
    starting at \( \lambda = 0 \) and ending at \( \lambda = 1 \)
Outline

• Problem and Existing Methods
• Homotopy Optimization Methods
• Protein Structure Prediction Problem
• **Numerical Experiments**
• Conclusions/Future Directions
Numerical Experiments

9 chains (22 particles) with known structure

Loop Region

<table>
<thead>
<tr>
<th>Hydrophobic</th>
<th>Hydrophilic</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>E</td>
<td>F</td>
</tr>
<tr>
<td>G</td>
<td>H</td>
<td>I</td>
</tr>
</tbody>
</table>

Sequence Matching (%)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>77</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td>86</td>
<td>91</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td>91</td>
<td>86</td>
<td>77</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>73</td>
<td>82</td>
<td>73</td>
<td>82</td>
<td>100</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68</td>
<td>68</td>
<td>59</td>
<td>77</td>
</tr>
<tr>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68</td>
<td>68</td>
<td>59</td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68</td>
<td>59</td>
</tr>
<tr>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>73</td>
</tr>
</tbody>
</table>
Numerical Experiments
Numerical Experiments

• 62 template-target pairs
  – 10 pairs had identical native structures

• Methods
  – HOM vs. Newton’s method w/trust region (N-TR)
  – HOPE vs. simulated annealing (SA)
    • Different ensemble sizes (2,4,8,16)
    • Averaged over 10 runs
    • Perturbations where sequences differ

• Measuring success
  – Structural overlap function: $0 \leq \chi \leq 1$
    • Percentage of interparticle distances off by more than 20% of the average bond length ($\bar{r}$)
  – Root mean-squared deviation (RMSD)
## Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Ensemble Size</th>
<th>$\chi = 0$</th>
<th>Success</th>
<th>Mean $\chi$</th>
<th>Mean RMSD</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOM</td>
<td>2</td>
<td>33.40</td>
<td>0.54</td>
<td>0.14</td>
<td>0.17</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>43.10</td>
<td>0.70</td>
<td>0.08</td>
<td>0.11</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>54.60</td>
<td>0.88</td>
<td>0.03</td>
<td>0.04</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>59.00</td>
<td>0.95</td>
<td>0.01</td>
<td>0.02</td>
<td>200</td>
</tr>
<tr>
<td>N-TR</td>
<td>4</td>
<td>13.10</td>
<td>0.21</td>
<td>0.27</td>
<td>0.36</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>28.50</td>
<td>0.46</td>
<td>0.13</td>
<td>0.19</td>
<td>229</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>40.20</td>
<td>0.65</td>
<td>0.08</td>
<td>0.12</td>
<td>434</td>
</tr>
</tbody>
</table>
Results

Success of HOPE and SA with ensembles of size 16 for each template-target pair. The size of each circle represents the percentage of successful predictions over the 10 runs.
Outline

• Problem and Existing Methods
• Homotopy Optimization Methods
• Protein Structure Prediction Problem
• Numerical Experiments
• Conclusions/Future Directions
Conclusions

• Homotopy optimization methods
  – More successful than standard minimizers

• HOPE
  – For problems with $f^0, x^0 (E^0, X^0)$ readily available
  – Solves protein structure prediction problem
  – Outperforms ensemble-based simulated annealing
    • SA parameters not optimal
HOPEful Directions

• **Protein structure prediction**
  – Protein Data Bank (templates)
  – Different size chains for template/target

• **HOPE for other minimization problems**
  – Standard test problems
  – Probabilistic convergence analysis \((\mathbb{R}^n)\)

• **HOPE for large-scale problems**
  – Inherently parallelizable
  – Communication: enforce maximum ensemble size
Other Work/Interests

• **Optimization**
  – Surrogate models in APPSPACK (pattern search)

• **Linear Algebra**
  – Structure preserving eigensolvers
    • Quaternion-based Jacobi-like methods
  – RF circuit design – efficient DAE solvers
    • Preconditioners, harmonic-balance methods

• **Information processing/extraction**
  – Entity recognition/disambiguation
    • Persons, locations, organization
  – Querying, clustering and summarizing documents
Acknowledgements

• Dianne O’Leary (UM)
  – Advisor
• Dev Thirumalai (UM), Dmitri Klimov (GMU)
  – Model, suggestions
• Ron Unger (Bar-Ilan)
  – Problem formulation
• National Library of Medicine (NLM)
  – Grant: F37-LM008162
Thank You

Daniel Dunlavy – HOPE

http://www.math.umd.edu/~ddunlavy
ddunlavy@math.umd.edu
Homotopy Parameter Functions

• Split low/high frequency dihedral terms
  \[ E_{dih}(X) = E_{dih1}(X) + E_{dih2}(X) \]
  \[ = \sum_{i=1}^{n-3} A_i(1 + \cos \phi_i) + \sum_{i=1}^{n-3} B_i(1 + \cos 3\phi_i) \]

• Homotopy parameter functions for each term
  \[ H(X, \lambda) = E_{bl}^0(X) + E_{ba}^0(X) \]
  \[ + (1 - \rho_{dih1}(\lambda))E_{dih1}^0(X) + \rho_{dih1}(\lambda)E_{dih1}^1(X) \]
  \[ + (1 - \rho_{dih2}(\lambda))E_{dih2}^0(X) + \rho_{dih2}(\lambda)E_{dih2}^1(X) \]
  \[ + (1 - \rho_{non}(\lambda))E_{non}^0(X) + \rho_{non}(\lambda)E_{non}^1(X) \]

\[ \rho(\lambda) = \begin{cases} 
0, & \text{if } \lambda = 0 \\
1, & \text{if } \lambda = 1 
\end{cases} \]
Homotopy Function for Proteins

Different $\rho(\lambda)$ for individual energy terms

Template $\rho^0(\lambda)$

Target $\rho^1(\lambda)$
HOPE Algorithm

Given: $f_0, f_1, h, x^0$  
$N_E$ (max ensemble size), $N_\xi$ (# of perturbations)

Initialize: $E_0 = x^0$, $\lambda = \Delta \lambda$, $k = 1$

while ($\lambda \leq 1$)
  for $j = 1 : \text{length}(E_{k-1})$
    for $i = 1 : N_\xi$
      $X(j,i) = \min_x h(x, \lambda)$, using $\xi(E_{k-1,j})$ as starting point
    end
  end
  $X(j,0) = \min_x h(x, \lambda)$, using $E_{k-1,j}$ as starting point
end

save at most $N_E$ ”best” points in $X$ in $E_k$

$\lambda = \lambda + \Delta \lambda$, $k = k + 1$
end
Structural Overlap Function

\[
\chi = 1 - \frac{2}{n^2 - 5n + 6} \sum_{i=1}^{n-3} \sum_{j=i+3}^{n} \Theta \left( 0.2\bar{r} - |r_{ij} - r_{ij}^*| \right)
\]
RMSD

Measures the distance between corresponding particles in the predicted and lowest energy conformations when they are optimally superimposed.

\[ RMSD(X) = \min_{S(X)} \sqrt{\frac{1}{n} \sum_{i=1}^{n} \| X_i - X_i^* \|^2} \]

where \( S(X) \) is a rotation and translation of \( X \).