Novel Optimization-based Multilevel CVT Construction Algorithm

Zichao Di
Department of Mathematical Sciences
George Mason University

Joint work with:
Maria Emelianenko
Stephen G. Nash

July 22, 2011

Outline

- *Introduction*
  - Centroidal Voronoi tessellation (CVT) formulation
- *Multigrid-based Optimization for CVT construction*
  - Multigrid Optimization (MG/OPT) algorithm: background
  - Applying MG/OPT to the CVT Formulation
- Numerical Experiments
  - 1-dimensional examples
  - 2-dimensional examples
- Comments and Conclusions

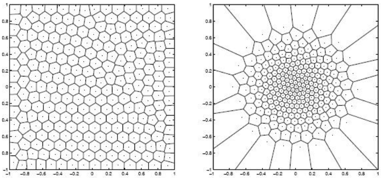
Centroidal Voronoi tessellation (CVT)

- Given
  - a set of points $\{\mathbf{z}_i\}_{i=1}^k$ (generators)
  - a set of regions $\{V_i\}_{i=1}^k$ (Voronoi tessellation)
- Define *the centroids*: $$\mathbf{z}_i^* = \frac{\int_{V_i} \rho(y) y dy}{\int_{V_i} \rho(y) dy}$$
- Define *energy functional*:

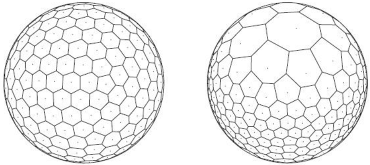
$$\mathcal{G}\left(\{\mathbf{z}_i\}_{i=1}^k\right) = \sum_{i=1}^k \int_{V_i} \rho(\mathbf{y}) |\mathbf{y} - \mathbf{z}_i|^2 \, d\mathbf{y}.$$

- The minimizer of $\mathcal{G}$ forms a CVT with $\mathbf{z}_i = \mathbf{z}_i^*$. This illustrates the optimization property of the CVT

tessellations of a square



tessellations on a sphere



Some Methods to construct CVTs

- Fixed-point iteration methods:
  - · Lloyd method (Lloyd 1960s)
- Probabilistic methods:
  - · random k-means methods (Macqueen 1966)
  - · random 2-clustering method (Imai et.al 1994)
- Newton-based methods:
  - · Newton-Lloyd method (Du/Emelianenko 2008)
  - · quasi-Newton L-BFGS method (Liu et.al 2009)
- Multigrid/multilevel methods:
  - · Lloyd-Multilevel (Du/Emelianenko 2008)
  - · Multilevel Lloyd-Max (Koren/Yavneh 2003,2006)

Introduction of MG/OPT

- MG/OPT is a general framework to accelerate a traditional optimization algorithm labeled "OPT", first introduced in [S.G. Nash 2000].
- Recursively use coarse problems to generate search directions for fine problems
- Inspired by multigrid for elliptic PDEs
  - · highly efficient algorithm

Notations of MG/OPT components

- OPT: a convergent optimization algorithm:

$$\mathbf{z}^{+} \leftarrow \mathrm{Opt}(\mathcal{G}(\cdot), \mathbf{v}, \bar{\mathbf{z}}, k)$$

- h: fine grid; H: coarse grid
- a high-fidelity model $\mathcal{G}_h$
- a low-fidelity model $\mathcal{G}_H$
- a downdate operator I_h^H
- an update operator I_H^h

Multigrid Optimization Algorithm (MG/OPT)

- Given:
  - an initial estimate $\mathbf{v}_h^0$ of the solution $\mathbf{z}_h$ on the fine level
  - Integers k_1 and k_2 satisfying $k_1 + k_2 > 0$
- One iteration of MG/OPT:
  - *Pre-smoothing:* $\bar{\mathbf{z}}_h \leftarrow \mathrm{Opt}(\mathcal{G}_h(\cdot), \mathbf{v}_h, \mathbf{z}_h^j, k_1)$
  - *Recursion:*
    - Compute

$$\begin{aligned} \bar{\mathbf{z}}_H &= I_h^H \bar{\mathbf{z}}_h \\ \bar{\mathbf{v}} &= \hat{I}_h^H \mathbf{v}_h + \nabla \mathcal{G}_H(\bar{\mathbf{z}}_H) - \hat{I}_h^H \nabla \mathcal{G}_h(\bar{\mathbf{z}}_h) \end{aligned}$$

    - Apply MG/Opt recursively to the surrogate model:

$$\mathbf{z}_H^+ \leftarrow \mathrm{MG/Opt}(\mathcal{G}_H(\cdot), \bar{\mathbf{v}}, \bar{\mathbf{z}}_H)$$

    - Compute the search directions $e_H = \mathbf{z}_H^+ - \bar{\mathbf{z}}_H$ and $e_h = I_H^h e_H$.
    - Use a line search to determine $\mathbf{z}_h^+ = \bar{\mathbf{z}}_h + \alpha e_h$ satisfying $\mathcal{G}_h(\mathbf{z}_h^+) \leq \mathcal{G}_h(\bar{\mathbf{z}}_h)$.
  - *Post-smoothing:* $\mathbf{z}_h^{j+1} \leftarrow \mathrm{Opt}(\mathcal{G}_h(\cdot), v_h, \mathbf{z}_h^+, k_2)$

Properties of MG/OPT

- Not dependent on a specific (single-grid) optimization algorithm
- Isolates optimization algorithm, differential equation solver, grid generation
- Convergence & descent can be guaranteed
- Excellent performance at initial iteration

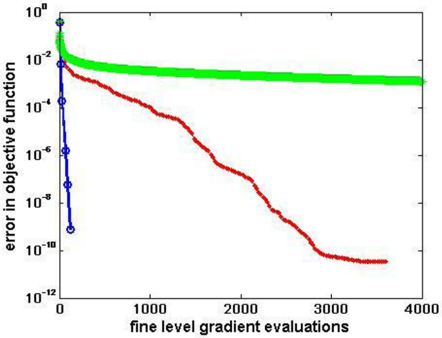
MG/OPT setup

- We have chosen OPT to be the truncated-Newton algorithm in our experiments
- 1-D case:
  - Downdate from finer to coarser grid:
    - Solution restriction (injecton): $[I_h^H v_h]^i = v_h^{2i}, \quad i = 1, 2, \ldots, k/2$
    - Gradient restriction (scaled full weighting): $[\hat{I}_h^H v_h]^i = \frac{1}{2} v_h^{2i-1} + v_h^{2i} + \frac{1}{2} v_h^{2i+1}, \quad i = 1, 2, \ldots, k/2$
  - Update from coarser to finer grid: $I_H^h = (\hat{I}_h^H)^T$
- 2-D case:
  - Downdate from finer to coarser grid:
    - Solution restriction is performed by injection
    - Gradient restriction: $[\hat{I}_h^H v_h]^i = \sum_j \alpha_j^i v_h^j$, where $\alpha_i^i = 1$ and $\alpha_j^i = \frac{1}{2}$ for any j s.t. z_j is a fine node sharing an edge with z_i in the fine level triangulation
  - Update from coarser to finer grid: $I_H^h = 4(\hat{I}_h^H)^T$

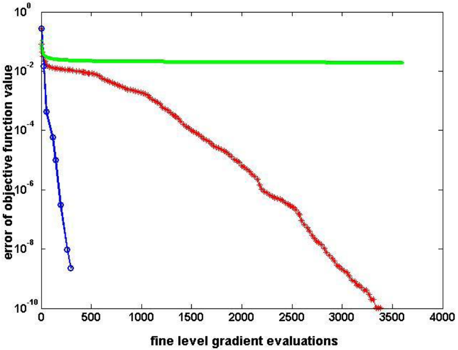
Convergence Result on 1-D CVT

Blue: MG/Opt; Red: OPT; Green: Lloyd

$\rho(y) = 1$

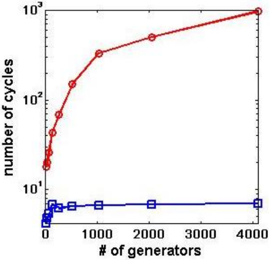


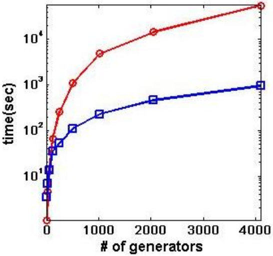
$\rho(y) = 6y^2 e^{-2y^3}$

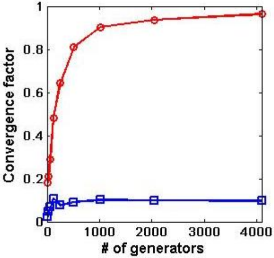


Convergence Result on 1-D CVT (cont'd)

Solving problems of increasing size, MG/Opt versus OPT ($\rho(y) = 1$). Blue: MG/Opt, Red: OPT;

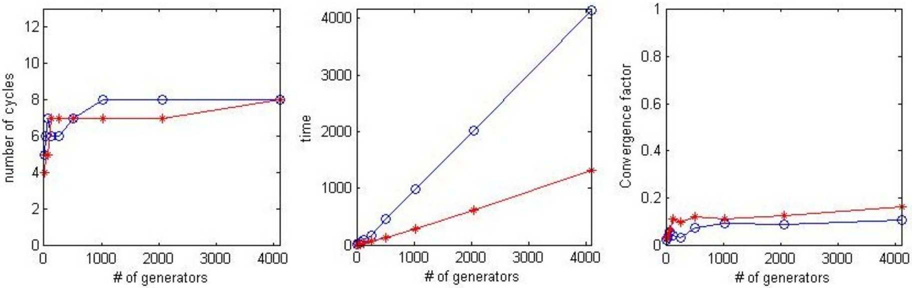






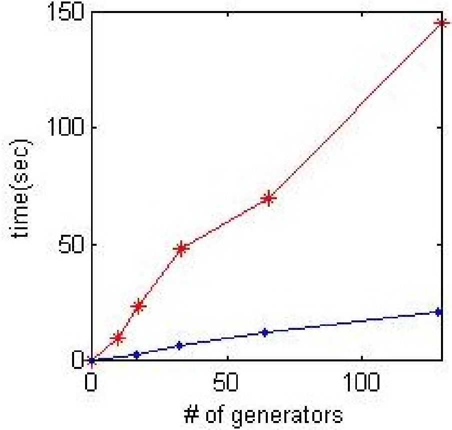
Convergence Result on 1-D CVT (cont'd)

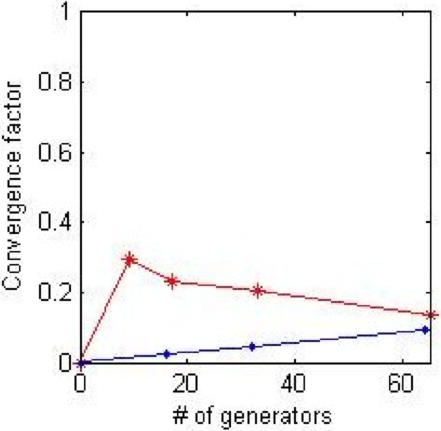
Performance of MG/Opt with different density functions. Blue: $\rho(y) = 1 + 0.1y$. Red: $\rho(y) = 6y^2 e^{-2y^3}$



Comparison with Multilevel-Lloyd

Convergence factors for MG/Opt vs. Multilevel-Lloyd $\rho(y) = 1$; Blue: MG/Opt; Red: Multilevel-Lloyd





Comparison with Multilevel Lloyd-Max

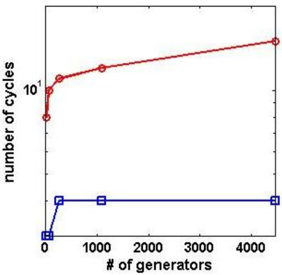
	convergence factor			
k	MG/Opt^2	MG/Opt^1	MLM	OPT
16	0.0350	0.0023	0.1725	0.8060
32	0.0314	0.0177	0.1782	0.9024
64	0.0673	0.0551	0.1847	0.9381
128	0.1120	0.0673	0.1962	0.9736

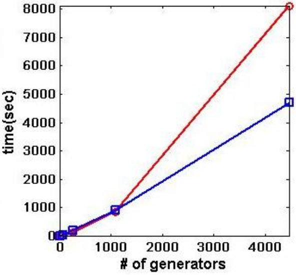
Table: MG/OPT v.s. Multilevel Lloyd-Max v.s. OPT; density: $y = 6x^2 e^{-2x^3}$

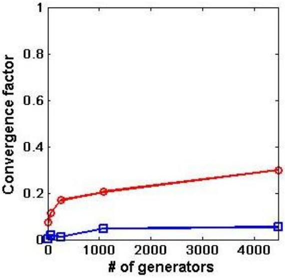
$$C^1 = \frac{|\mathcal{G}(\mathbf{z}^{k+1}) - \mathcal{G}(\mathbf{z}^*)|}{|\mathcal{G}(\mathbf{z}^k) - \mathcal{G}(\mathbf{z}^*)|};\ C^2 = \left(\frac{|\mathcal{G}(\mathbf{z}^{k+1}) - \mathcal{G}(\mathbf{z}^*)|}{|\mathcal{G}(\mathbf{z}^1) - \mathcal{G}(\mathbf{z}^*)|}\right)^{\frac{1}{k+1}}$$

Convergence Result on 2-D CVT based on triangular domain

Red: Opt; Blue: MG/OPT; $\rho(x) = 1$







Di/Emelianenko/Nash, to appear in NMTMA, 2011

Results and challenges

- CVT is in the heart of many applications and the number is growing: computer science, physics, social sciences, biology, engineering ...
- The main advantage of MG/OPT is its superior convergence speed, and the fact that it preserves low convergence factor regardless of the problem size.
- The simplicity of its design and the results of preliminary tests suggest that the method is generalizable to higher dimensions, which is the subject of further investigations
- Future work also includes application of this technique to various scientific and engineering applications, including image analysis and grid generation.

Results and challenges

- CVT is in the heart of many applications and the number is growing: computer science, physics, social sciences, biology, engineering ...
- The main advantage of MG/OPT is its superior convergence speed, and the fact that it preserves low convergence factor regardless of the problem size.
- The simplicity of its design and the results of preliminary tests suggest that the method is generalizable to higher dimensions, which is the subject of further investigations
- Future work also includes application of this technique to various scientific and engineering applications, including image analysis and grid generation.

THANKS!