Ideal Point Distributions, Best Mode Selections and Optimal Spatial Partitions via Centroidal Voronoi Tessellations *

Qiang Du † Maria Emelianenko ‡ Hyung-Chun Lee § Xiaoqiang Wang ¶

Abstract

There are many new applications of the centroidal Voronoi tessellations that come to life in recent years, along with more mathematical understandings and new algorithmic advances in their efficient computation. Some examples are presented in this paper as an illustration with an emphasis on the construction of ideal point distributions, best mode selections and optimal spatial partitions.

1 Introduction

A centroidal Voronoi tessellation (CVT) is a Voronoi tessellation of a given set such that the associated generating points are centroids, i.e., the centers of mass with respect to a given density function, of the corresponding Voronoi regions [13].

By introducing different metrics and distances in the definition of Voronoi regions and different notions of centroids, CVTs may be used to find ideal distributions of generators or representatives, best selections of modes in model reduction, and optimal partitions or clusterings for various practical purposes. We hereby briefly review some of the optimal properties associated with the CVTs defined with a standard Euclidean metric and also present some variants with more general metrics. Recent algorithmic advances in the computation of CVTs are also discussed. Several examples from different application fields are provided, with an emphasis on the use of the CVTs for constructing ideal point distributions, the best mode selections and optimal spatial tessellations. These constructions are often needed in problems ranging from data and image processing, geometric rendering and representation, model simplifications and reduction, to resource distribution and allocation. It is concluded that CVTs, due to their optimality, simplicity and universality, are becoming favorite concepts to be explored in many applications.

2 Centroidal Voronoi Tessellations

A Voronoi tessellation refers to a tessellation of a given domain $\Omega \in \mathbb{R}^N$ by the Voronoi regions $\{V_i\}_{i=1}^k$ associated with a set of given *generating points* or *generators* $\{\mathbf{z}_i\}_{i=1}^k \subset \Omega$ [39]. For each i, $\{V_i\}_{i=1}^k$ consists of all points in the domain Ω that are closer to $\mathbf{z}_i$ than to all the other generating points.

For a given density function ρ defined on Ω, we may define the centroids, or mass centers, of regions $\{V_i\}_{i=1}^k$ by

$$\mathbf{z}_i^* = \int_{V_i} \mathbf{y}\rho(\mathbf{y})\, d\mathbf{y} \Big/ \int_{V_i} \rho(\mathbf{y})\, d\mathbf{y}\,.$$

Then, a *centroidal Voronoi tessellation* (CVT) refers to a Voronoi tessellation for which the generators themselves are the centroids of their respective Voronoi regions, that is, $\mathbf{z}_i = \mathbf{z}_i^*$ for all i. We refer to [13] for a more comprehensive review of the mathematical theory

*Proceedings of the 2nd International Symposium on Voronoi Diagrams in Science and Engineering (VD2005), pp.325-333, Seoul, Korea, Oct, 2005. The research is supported in part by NSF-DMS 0409297, NSF-CCF 0430349 and NSF-ITR 0205232. The research of Hyung-Chun Lee is supported in part by KOSEF R01-2000-00008 and KOSEF Sundo fund.

†Department of Mathematics, Pennsylvania State University, University Park, PA 16802 (qdu@math.psu.edu).

‡Department of Mathematics, Pennsylvania State University, University Park, PA 16802. Current address: Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213 (masha@cmu.edu).

§Department of Mathematics, Ajou University, Suwon, Korea 443-749 (hclee@ajou.ac.kr).

¶Department of Mathematics, Pennsylvania State University, University Park, PA 16802. Current address: Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 44544 (wang@ima.umn.edu).

and diverse applications of CVTs.

In the simple geometric setting of Euclidean spaces, CVTs often give very special and elegant tessellations. In figure 1, a two dimensional example using the standard Euclidean distance and the constant density is shown.

CVT is not merely a simple geometric concept, it can be extended to more general settings, including abstract sets and spaces, and more general metrics. They have a variety of applications including data compression, image compression, optimal allocations of resources, territorial behavior of animals, optimal sensor and actuator location, and numerical analysis including both grid-based and meshfree algorithms for interpolation, multi-dimensional integration, and partial differential equations; see [5, 4, 6, 7, 8, 10, 12, 13, 14, 15, 16, 18, 20, 21, 22, 23, 26, 27, 30, 33, 34, 38, 39, 43, 45, 46].

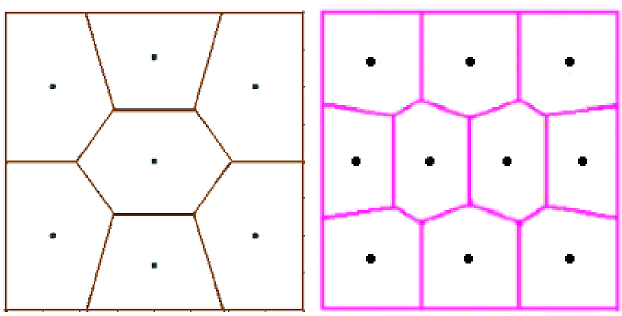

Figure 1: An illustration of CVTs (having 7 and 10 clusters respectively) in a square with a constant density.

Given a discrete set of points $W = \{\mathbf{x}_j\}_{j=1}^n$ belonging to $\mathbb{R}^N$, we define the error with respect to a tessellation $\{V_i\}_{i=1}^k$ of W and a set of points $\{\mathbf{z}_i\}_{i=1}^k$ belonging to W or, more generally, belonging to $\mathbb{R}^N$ by

$$\mathcal{F}\big((\mathbf{z}_i, V_i), i = 1, \ldots, k\big) = \sum_{i=1}^{k} \sum_{\mathbf{y} \in V_i} \rho(\mathbf{y}) |\mathbf{y} - \mathbf{z}_i|^2 .$$

It can be shown that a necessary condition for the error $\mathcal{F}$ to be minimized is that the pair $\{\mathbf{z}_i, V_i\}_{i=1}^k$ form a CVT of W. We note that the above error is also often referred to as the *variance, cost, distortion error,* or *mean square error.*

CVTs need not be the global energy minimizers of the error functional as some may actually be saddle points. In our discussion, we largely are only concerned with optimal CVTs which refer to the minimizers of the error functional $\mathcal{F}$. The optimality may often be translated into superior properties of the tessellation of the space and the spatial distributions of the generators, thus, optimal CVTs are becoming favorites in many applications.

The optimal CVTs of discrete sets are closely related to optimal *k-means clusters* and Voronoi regions and centroids are referred to as *clusters* and *cluster centers*, respectively. Clustering analysis provides a selection of a finite collection of templates that well represent, in some sense, a large collection of data as illustrated in [13]. It can be shown that, using the variance-based criteria to define optimality, the optimal clustering corresponds to a centroidal Voronoi tessellation.

3 Variants of CVTs

A number of variants of CVTs have been studied in recent years, which were used in different applications.

One particular variant is given by the constrained CVTs where the generators and the Voronoi regions are confined to a general Riemannian manifold M embedded in the Euclidean space R^d. The metric, however, is still taken as the standard metric of R^d. Systematic development of such constrained CVTs as well as algorithms for the constructions have been given in [16]. Extending the concept to general metric, in [22], anisotropic CVTs have been defined corresponding to a Riemannian metric tensor. The key is to define a one-sided distance function in the definition of the Voronoi regions, which allows their simple computation. In more detail, given a Riemannian matrix metric tensor M, the *directional distance* from a point $\vec{Q}$ to another point $\vec{P}$:

$$d_P(\vec{P}, \vec{Q}) = \sqrt{\overrightarrow{PQ}^T M(\vec{P}) \overrightarrow{PQ}} .$$

The use of such a distances bypasses the need of computing the geodesic distance corresponding to the

Riemannian metric. The *anisotropic Voronoi region* of a point $\mathbf{z}_i$ (with respect to the generator set $\{\mathbf{z}_j\}$ in Ω) is defined as:

$$V_i = \{\mathbf{x} \in \Omega \mid d_x(\mathbf{x}, \mathbf{z}_i) < d_x(\mathbf{x}, \mathbf{z}_j) \ \forall \ \mathbf{z}_j \neq \mathbf{z}_i\}.$$

The mass centroids are obtained straightforwardly through the minimization of a quadratic error function in V_i:

$$F(\mathbf{y}) = \int_{V_i} d^2_{\mathbf{x}}(\mathbf{x}, \mathbf{y}) d\mathbf{x} \ .$$

Once the AVRs and their mass centers are defined, the *anisotropic centroidal Voronoi tessellation* (ACVT) can then be defined [22], similar to the conventional CVTs.

Another variant of the CVTs, the mixture model based CVT, has been considered by coupling CVT with the classical EM algorithm in [26]. A distinctive feature of mixture model based CVTs lies in the automated construction of the anisotropic metric tensor through the available data points. This metric tensor defines the mixture model that describes the underlying sample distributions. Similar extensions were made to define CVTs for both continuous and discrete vector fields defined in an Euclidean domain or on a general Riemannian manifold [27]. Using $\vec{P} = (\mathbf{x}_p, \mathbf{y}_p)$ and $\vec{Q} = (\mathbf{x}_q, \mathbf{y}_q)$ denote two vectors $\mathbf{y}_p$ and $\mathbf{y}_q$ at positions $\mathbf{x}_p$ and $\mathbf{x}_q$ with $|\mathbf{y}_q| = 1$, the one-sided distance used in [27] is given by:

$$d_p(\vec{P}, \vec{Q}) = \sqrt{|\mathbf{y}_p|^2 - |\mathbf{y}_p| \, \mathbf{y}_p \cdot \mathbf{y}_q + w|\mathbf{y}_p|^2 |\mathbf{x}_p - \mathbf{x}_q|^2} \ .$$

Here, w is a constant scaling factor that serves to balance the emphasis on the variation of the vector fields with respect to the spatial distributions. For a given non-uniformly distributed vector fields with a density distribution $\rho(\mathbf{x}_p)$, a centroid $\vec{m} = (\mathbf{x}_m, \mathbf{y}_m)$ of a spatial region C is given by the minimizer of the mean distance square:

$$E(\vec{m}, C) = \int_C \rho(\mathbf{x}_p) d^2_p(\vec{P}, \vec{m}) \ d\mathbf{x}_p \ .$$

The Voronoi regions and the CVTs may thus be defined accordingly [27]. An application of the CVT based vector field clustering and simplification is illustrated in figure 2.

These variants are useful in the application of CVTs to surface and anisotropic mesh generations, vector field clustering and simplifications. Naturally, one may also consider CVTs for additive and multiplicative Voronoi tessellations.

4 Algorithms for CVTs

In the vector quantization literature, CVTs give rise to the optimal vector quantizers. In the seminal work of Lloyd on the least square quantization [35], one of the algorithms proposed for computing optimal vector quantizers is an iterative algorithm consisting of the following simple steps: starting from an initial quantization (a Voronoi tessellation corresponding to an old set of generators), a new set of generators is defined by the mass centers of the Voronoi regions. This process is continued until certain stopping criterion is met. It is easy to see that the Lloyd algorithm is an energy descent iteration of the energy functional, which gives strong indications to its practical convergence, we refer to [12] for some discussion on the recent development of a rigorous convergence theory. In particular, the global convergence of the Lloyd iteration has been established for any density functions in the one-dimensional case. Results on convergence of subsequences have also been given there.

Lloyd's algorithms and their variants have been proposed and studied in many contexts for different applications [29, 33]. For modern applications of the CVT concept in large scale scientific and engineering problems such as data communication and mesh generation, efficient algorithms for computing the CVTs play crucial roles. In recent works, we examined a number of different approaches to speed up the convergence of Lloyd iteration.

In [10], the direct application of the Newton method has been studied and a hybrid Lloyd-Newton scheme has also been proposed. The fast local quadratic convergence of the Newton iteration acts to accelerate the convergence of the original Lloyd's method. The robustness and the energy descent property of the Lloyd's algorithm, meanwhile, serve to provide good starting point for the Newton' method. Both analytical and numerical results on the hybrid scheme have been given in [10].

The ideas of multilevel algorithms have been explored in [10] and [11] recently. Two main strategies have been proposed for applying them in the nonlinear optimization context with one being a full nonlinear multilevel construction and another one combining a linearization outer iteration and an inner linear multilevel iteration. The multilevel scheme is implemented through a hierarchical space decomposition and the minimization of the error functional with subspace corrections. The main characteristics of such schemes, such as the dynamic nonlinear preconditioning and uniform convergence with respect to the problem size, have been discussed and some numerical results demonstrating its superiority over traditional methods have been provided [10, 11].

For probabilistic type of methods, the classical *MacQueen's method* [37], is a very elegant random sequential sampling method for the computation of CVTs. A generalization of MacQueen's method given in [33] for the probabilistic construction of CVTs provides motivations to the algorithms studied in [26]. The new algorithms in [26] may also be viewed as stochastic implementation of the Lloyd's algorithm with an automated estimation of the metric tensor in the context of the EM methods.

5 Applications of CVTs: some examples

CVTs have diverse application as outlined earlier. As an illustration, we present some applications of optimal CVTs, with an emphasis on ideal point distributions, best mode selections and optimal spatial tessellations and triangulations.

In figure 2, some examples of using CVTs for vector field simplification and representation are presented.

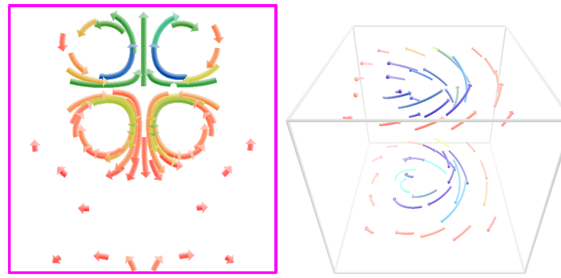

Figure 2: Simplified vector field representation based on the CVTs: 2d and 3d examples.

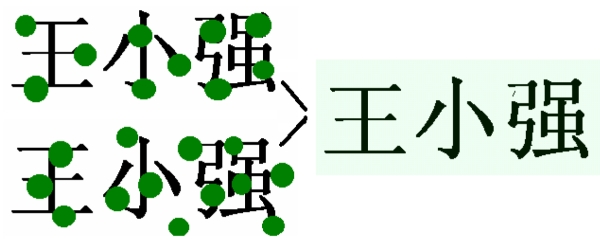

Figure 3: CVT based image analysis: processed image (right) based on images from a two-channel input.

The CVTs are characterized by the one sided distance and the centroid concept introduced in the section 3. They are produced for the analysis of complex flow fields [26].

CVTs have also been applied to image segmentation, thresholding, color compression. For example, the color compression problem can be formulated as to find the optimal distribution of the representative colors (among all the colors in the image) and to find the optimal tessellation of the set of colors which are used to define color replacement rules [13]. In figure 3, we present an example of CVT based multi-channel image segmentation and restoration [18]. Here, two stained copies of the same text image (the Chinese name of one of the authors) are given as a two-channel input, CVTs are then defined for vectors which represent the colors in both of these two images. The optimal clustering gives to an image whose colors try to be the common colors in both images and thus allows us to remove the stains and recognizes the original Chinese characters.

In figure 4, CVTs constrained to the sphere as de-

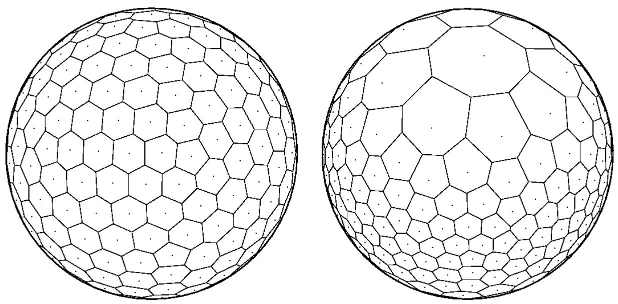

Figure 4: Constrained spherical CVTs with uniform and nonuniform densities [16].

fined in [16] are presented. The optimal CVTs corresponding to a constant density gave a nearly uniform tessellation of the sphere. Moreover, optimal CVTs in general have been used in mesh generation and optimization [14, 22, 23]. The Delaunay triangulations are the duals of Voronoi tessellations, thus, optimal Delaunay triangulations may result from optimal CVTs. On the sphere, it is known that no perfectly uniform triangulations exist on the sphere in general, but the dual of the optimal CVTs on sphere gives a nearly uniform triangulation. Let us also give another example for a two dimensional square. In figure 5, a comparison of a Delaunay triangulation based on a uniformly sampled vertices with that based on a CVT is compared. In [20], it has been shown that the finite volume scheme for the convection diffusion equation based on the CVTs and their dual Delaunay triangulations has higher order accuracy than that based on a generic pair of Voronoi tessellation and Delaunay triangulation.

The proper orthogonal decompositions (POD) have been used to systematically extract the most energetic modes in the study of complex dynamic systems. POD is closely related to the statistical method known as Karhunen-Loève analysis or the method of empirical orthogonal eigenfunctions, and it is intimately associated with the more well-known concept of singular value decomposition. There have been many studies devoted to the use of POD for obtaining low-dimensional dynamical system approximations; see, for example

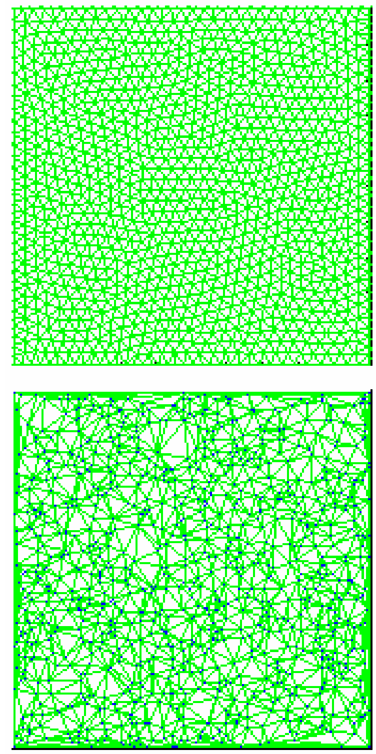

Figure 5: Delaunay triangulations with uniformly sampled vertices and CVT generated vertices [14].

[2, 3, 9, 28, 31, 32, 36, 40, 41, 42, 44]. While having no assurance to success, the popularity of POD analysis is mostly due to its simplicity. In [15], a new clustering approach was introduced by combining the POD with CVTs into a hybrid method for model reduction. The optimality of such an approach and various practical implementation strategies have also been discussed.

In the proper orthogonal decompositions (POD) technique, dominant features from experimental or numerical data are extracted through a set of orthogonal basis functions which are related to the eigenfunctions of the correlation matrix of the data. More specifically, for n vectors (which are commonly called snapshots) $\widetilde{\mathbf{x}}_j \in \mathbb{R}^N$, $j = 1, \ldots, n$, let $\{\mathbf{x}_j = \widetilde{\mathbf{x}}_j - \widetilde{\boldsymbol{\mu}}\}_{j=1}^n$ be a set of modified snapshots where $\widetilde{\boldsymbol{\mu}} = (\sum_{j=1}^n \widetilde{\mathbf{x}}_j)/n$. Let $d \le n$, the POD basis $\{\boldsymbol{\phi}_i\}_{i=1}^d$ of cardinality d is then found by successively solving

$$\lambda_i = \max_{\|\boldsymbol{\phi}_i\|=1} \frac{1}{n} \sum_{j=1}^{n} |\boldsymbol{\phi}_i^T \mathbf{x}_j|^2 \qquad \text{and} \qquad \boldsymbol{\phi}_i^T \boldsymbol{\phi}_\ell = 0$$

for $\ell \leq i-1$ and $i = 1, \ldots, d$. Here, we will only consider the case $n < N$.

The POD basis is optimal in the following sense [31]: let $\{\boldsymbol{\psi}_i\}_{i=1}^n$ denote an arbitrary orthonormal basis for the span of the modified snapshot set $\{\mathbf{x}_j\}_{j=1}^n$. Let $P_{\boldsymbol{\psi},d}\mathbf{x}_j$ be the projection of $\mathbf{x}_j$ in the subspace spanned by $\{\boldsymbol{\psi}_i\}_{i=1}^d$ and let the *error* be defined by $\mathcal{E} = \sum_{j=1}^n \|\mathbf{x}_j - P_{\boldsymbol{\psi},d}\mathbf{x}_j\|^2$. Then, the minimum error is obtained when $\boldsymbol{\psi}_i = \boldsymbol{\phi}_i$ for $i = 1, \ldots, d$, i.e., when the $\boldsymbol{\psi}_i$'s are the POD basis vectors.

To combine the CVT concept with POD, we again essentially need to specify the two ingredients in the definition of CVTs: distances and centroids. The generators themselves are selected modes or basis functions.

First, the square of the distance from a one dimensional subspace spanned by a vector $\mathbf{x}$ to a d-dimensional subspace $\mathcal{Z}$ is defined in conventional term:

$$\delta^2(\mathbf{x}, \mathcal{Z}) = 1 - \frac{1}{\|\mathbf{x}\|^2} \sum_{i=1}^{d} |\mathbf{x}^T \boldsymbol{\theta}_i|^2 \,,$$

where $\{\boldsymbol{\theta}_i\}_{i=1}^d$ forms an orthonormal basis for $\mathcal{Z}$. Then, given a set of vectors (e.g., modified snapshots) $W = \{\mathbf{x}_j\}_{j=1}^n$ and a set of d-dimensional subspaces $\{\mathcal{Z}_i\}_{i=1}^k$ (which are the generators), we define the generalized Voronoi tessellation of W by

$$\mathcal{V}_i = \{\mathbf{x}_j \in W \ \mid \ \delta^2(\mathbf{x}_j, \mathcal{Z}_i) \leq \delta^2(\mathbf{x}_j, \mathcal{Z}_\ell) \quad \forall\, \ell \neq i \ \}$$

for $i = 1, \ldots, k$.

Second, given a set of vectors $\mathcal{V} = \{\mathbf{x}_j\}$ that span an m-dimensional subspace of $\mathbb{R}^N$, the generalized centroid of $\mathcal{V}$ may be defined by an orthonormal basis $\{\boldsymbol{\phi}_i\}_{i=1}^d$ which minimizes

$$\mathcal{D} = \sum_{\mathbf{x}_j \in \mathcal{V}} \|\mathbf{x}_j - P\mathbf{x}_j\|^2 \,,$$

where P denotes the projection operator into the d-dimensional subspace spanned by $\{\boldsymbol{\phi}_i\}_{i=1}^d$. The optimal basis $\{\boldsymbol{\phi}_i\}_{i=1}^d$ is in fact the d-dimensional POD basis for the set $\mathcal{V}$. Moreover, the generators $\{\mathcal{Z}_j\}$ need not have the same cardinality for different subspaces. Thus, we use $\mathbf{d} = \{d_i\}_{i=1}^k$ to denote a multi-index.

If a set of finite subspaces $\{\mathcal{Z}_j\}_{i=1}^k$ with dimensions $\mathbf{d} = \{d_i\}_{i=1}^k$, respectively, along with the corresponding generalized Voronoi tessellation $\{\mathcal{V}_j\}_{i=1}^k$ is a *CVT* (i.e., the $\mathcal{Z}_i$'s are themselves the centroids of the $\mathcal{V}_i$'s), then the union of basis vectors corresponding to the CVT is called a *centroidal Voronoi orthogonal decomposition (CVOD)* [15]. CVOD can be viewed as an optimal basis or mode selection procedure. It is a generalization of CVT for which the set W of modified snapshots is divided into k clusters or generalized Voronoi regions $\{\mathcal{V}_i\}_{i=1}^k$ and for which the generators are d_i-dimensional spaces and each of which is spanned by the d_i-dimensional POD basis inside the same cluster. CVOD is also a generalization of POD in the sense that the set of modified snapshots is first divided into k clusters and a POD basis is separately determined for each cluster. If $d_i = 1$ for $i = 1, \ldots, k$, CVOD reduces to the standard CVT, while if $k = 1$, CVOD educes to the standard POD.

Since a nonuniform density function can be used in the standard CVT construction, we may also define the weighted CVOD with a prescribed *discrete density* or a set of weights, i.e., for a density function ρ with $\{\rho(\mathbf{x}_j) = \rho_j\}_{j=1}^n$, the more general CVOD minimizes the functional

$$\mathcal{G}(\{\mathcal{Z}_i, \mathcal{V}_i\}) = \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in \mathcal{V}_i} \rho_j \delta^2(\mathbf{x}_j, \mathcal{Z}_i) = \sum_{i=1}^{k} |\mathcal{V}_i| \sum_{j=d_i+1}^{|\mathcal{V}_i|} \lambda_{i_j}$$

where $|\mathcal{V}_i|$ denotes the cardinality of the Voronoi set or cluster $\mathcal{V}_i$ and the λ_{i_j}'s are the eigenvalues (in decreasing order) of the (weighted) local correlation matrix of the snapshots in the cluster.

Once a reduced set of modes or a reduced basis is obtained, either by POD or CVT or CVOD, it can be used to define a low-order model for a complex system in a more automated fashion. For instance, let

$F(t, X, u(X, t)) = 0$ be a system of partial differential equations with suitable boundary and/or initial conditions for the unknown function u and some system parameter t. The CVOD based model reduction is performed as follows.

Algorithm: CVOD based model reduction

1 Construct a set of modified snapshots $\{u_j\}_1^n$ by solving $F(t, X, u(X, t)) = 0$ for different t.

2 Calculate the CVOD for the set $\{u_j\}_1^n$ for some integer k and multi-index $\{d_j\}_{j=1}^k$ to obtain a set of CVOD basis vectors $\{\phi_m\}_{m=1}^{|\mathbf{d}|}$.

3 For $1 \leq m \leq |\mathbf{d}|$, solve the reduced system $\left\langle \phi_m, F(t, X, \sum_{l=1}^{|\mathbf{d}|} \beta_l \phi_l) \right\rangle = 0$ with a suitable inner product.

The above procedures has been applied to fluid control problems [5] and problems in homogenization. Let us consider a hypothetical setting where a two dimensional Possion equation with a highly oscillatory diffusion coefficient and homogeneous Dirichlet boundary condition appears as a state equation that needs to be numerically solved with many different right hand side functions. Due to the oscillation in the coefficient, the solutions of the PDEs in general exhibit very complex behavior thus demand detailed resolution with a fine grid. However, it is likely that for a large class of smooth right hand side functions, the corresponding solutions are highly clustered and stay in a relatively low dimensional space. One may take some solutions already computed as snapshots and perform the CVOD model reduction so that new solutions can be computed with the reduced basis when the equations are to be solved again. The computed solutions can be very good approximations to those computed directly with a fine grid but the latter comes at much higher computational cost. An example of this is given in figure 6 where the computed numerical solution using a reduced CVOD basis with four generators is shown along with the error between

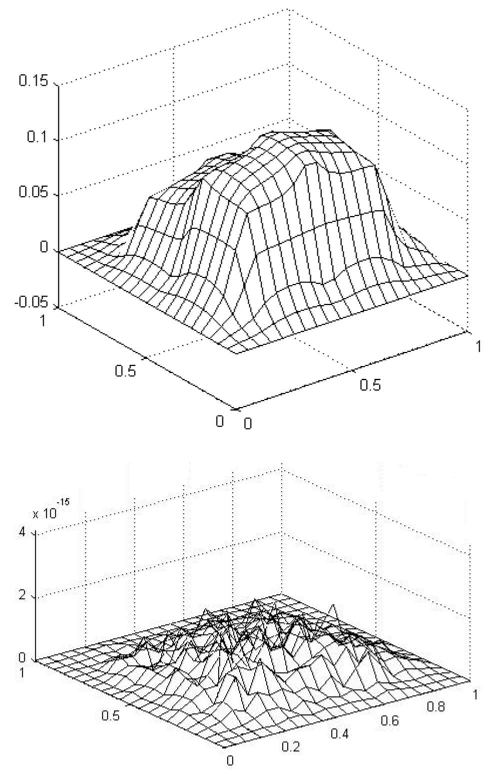

Figure 6: A solution for a two dimensional homogenization problem computed with 4 CVOD basis and the error between it and the fine grid solution.

it and the solution computed on a much finer grid. Applications of CVOD in MD and ab-initio simulations are now also under investigation.

The advantage of CVOD over POD is that POD mostly is a linear decomposition technique, CVOD is on the other hand nonlinear in nature and it introduces the concept of clustering into the decomposition. CVOD also reduces the amount of work relative to the full POD analysis. POD involves the solution of an $n \times n$ eigenproblem, where n is the number of snapshots; CVOD instead requires the solution of several smaller eigenproblems. Yet, one of the most important features of CVOD lies also in its universality and simplicity, it can be developed as an on-line toolbox that can constantly probe the computation results and seek for clustering patterns and identifying dominant modes. This serves the goal of manifold learning, model reduction and knowledge discovery. For complex dynamic systems, how to efficiently compute the CVOD dynamically through some adaptive and hierarchical processes becomes important. In this regard, the recent works on multilevel methods for computing the CVTs [10] hold great promises.

6 Conclusion

Centroidal Voronoi Tessellations are useful mathematical concepts and practical tools. They are simple, as the definitions are easy to describe; they are universal, as they appear in various context; they are progressive, as iterative algorithms provide a gradual construction of CVTs and result in improvement of the tessellation and point distribution; and best of all, they are optimal, as they minimizes error functionals, variances and distortion measures.

In this paper, the basic concepts of CVTs are described and some recent algorithmic advances of CVTs are discussed. Drawing from our recent works, several examples of applications are presented. When the number of generators becomes large, it has been conjectured [13] that CVT's enjoys the equi-partition of error property, following the celebrated Gersho's conjecture [25]; it is natural to extend such a conjecture to CVOD. Such an error equi-partition property may be explored in future works to find adaptive strategies for the construction of CVTs. Finally, we are surely going to see more and more new and exciting applications of CVTs in the future.

Acknowledgement. We would like to thank Professors Max Gunzburger (Florida State Univ), Lili Ju (Univ South Carolina) and Desheng Wang (Nanyang Univ) for their collaboration on projects related to the works presented here.

References

[1] P. Alliez, T. deVerdiere, O. Devillers and M. Isenburg, Centroidal Voronol diagrams for isotropic surface remeshing, *Graphical Models*, **67** (2005), 204-231.

[2] G. Berkooz, P. Holmes, and J. Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, *Ann. Rev. Fluid. Mech.*, **25** (1993), 539–575.

[3] G. Berkooz and E. Titi, Galerkin projections and the proper orthogonal decomposition for equivariant equations, *Physics Letters A*, **174** (1993), 94–102.

[4] I. Boier-Martin, H Rushmeier and J. Jin, Parameterization of Triangle Meshes over Quadrilateral Domains, Proceedings 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing, ACM Intern. Conf. Proc. Series, **71** (2004), 193-203.

[5] J. Burkardt, M Gunzburger and H.-C. Lee, Centroidal Voronoi Tessellation-Based Reduced-Order Modeling of Complex Systems, to appear.

[6] M. Cappellari and Y. Copin, Adaptive spatial binning of integral-field spectroscopic data using Voronoi tessellations, *Monthly Notices Royal Astronomical Soc.*, **342** (2003), 345-354.

[7] D. Cohen-Steiner, P. Alliez, and M. Desbrun, Variational shape approximation, *ACM Transactions on Graphics*, **23** (2004), 905-914.

[8] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, Coverage control for mobile sensing networks *IEEE Tran. Robotics and Automation*, **20** (2004), 243-255.

[9] A. Deane, I. Kevrekidis, G. Karniadakis, and S. Orszag, Low-dimensional models for complex geometry flows: applications to grooved channels and circular cylinders, *Phys. Fluids A*, **3** (1991), 2337–2354.

[10] Q. Du and M. Emelianenko, Acceleration schemes for computing centroidal Voronoi tessellations, preprint, 2005

[11] Q. Du and M. Emelianenko, Uniform Convergence of a Multilevel Energy-based Quantization Scheme, preprint, 2005

[12] Q. Du, M. Emelianenko and L. Ju, Convergence Properties of the Lloyd Algorithm for Computing the Centroidal Voronoi Tessellations, *SIAM J. Num. Anal.*, in press, 2005

[13] Q. Du, V. Faber, and M. Gunzburger, Centroidal Voronoi tessellations: applications and algorithms, *SIAM Review*, **41** (1999), 637–676.

[14] Q. Du and M. Gunzburger, Grid Generation and Optimization Based on Centroidal Voronoi Tessellations, *Appl. Math. and Comput.*, **133** (2002), 591–607.

[15] Q. Du and M. Gunzburger, Centroidal Voronoi Tessellation Based Proper Orthogonal Decomposition Analysis, *International series of Numerical Mathematics*, Birkhauser, **143** (2002), 137-150.

[16] Q. Du, M. Gunzburger, and L. Ju, Constrained centroidal Voronoi tessellations on general surfaces, *SIAM J. Sci. Comput.*, **24** (2003), 1488–1506.

[17] Q. Du, M. Gunzburger and L. Ju, Meshfree, probabilistic determination of point sets and support regions for meshless computing, *Comput. Meths. Appl. Mech. Engrg.*, **191** (2002), 1349–1366.

[18] Q. Du, M. Gunzburger, L. Ju and X. Wang, Centroidal Voronoi tessellation algorithms for image processing, *J. of Mathematical Imaging and Vision*, in press, 2005.

[19] Q. Du, Z. Huang and D. Wang, Mesh and Solver Coadaptation in Finite Element Methods for Anisotropic Problems, *Numerical Methods for Differential Equations*, 21, pp.859-874, 2005

[20] Q. Du and L. Ju, Finite volume methods and spherical centroidal Voronoi tessellations, *SIAM J. Numer. Anal.*, in press, 2005

[21] Q. Du and D. Wang, Tetrahedral mesh generation and optimization based on centroidal Voronoi tessellations, *Int. J. Numer. Meth. Eng.*, **56** (2002), 1355–1373.

[22] Q. Du and D. Wang, Anisotropic centroidal Voronoi tessellations and their applications, *SIAM J. Sci. Comput.*, **26** (2005), 737-761.

[23] Q. Du and D. Wang, New progress in robust and quality Delaunay mesh generation, *J. Computational and Applied Mathematics*, in press, 2005

[24] Q. Du and D. Wang, Mesh Optimization Based on the Centroidal Voronoi Tessellation, to appear in *International Journal of Numerical Analysis and Modeling*, 2005

[25] Q. Du and D. Wang, On the Optimal Centroidal Voronoi Tessellations and the Gersho's Conjecture in the Three Dimensional Space, *Computers and Mathematics with Applications*, **49** (2005), 1355-1373.

[26] Q. Du and X. Wang, Centroidal Voronoi tessellation based algorithms for vector fields visualization and segmentation, *IEEE Proceedings of Visualization 2004 (VIS2004)*, IEEE, pp43-50, Austin, TX, 2004.

[27] Q. Du and X. Wang, Tessellation and Clustering by Mixture Models and Their Parallel Implementations, *Proceeding of the fourth SIAM international conference on Data Mining*, Lake Buena Vista, FL, 2004, SIAM, pp257-268.

[28] M. Graham and I. Kevrekidis, Pattern analysis and model reduction: some alternative approaches to the Karhunen-Lóeve decomposition, *Comp. Chem. Engng.*, **20** (1996), 495–506.

[29] R. Gray and D. Neuhoff, Quantization, *IEEE Trans. Inform. Theory*, **44** (1998), 2325–2383.

[30] S. Hiller, H. Hellwig and O. Deussen, Beyond stippling - Methods for distributing objects on the plane, *Computer Graphics Forum*, **22** (2003), 515-522.

[31] P. Holmes, J. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, (1996).

[32] P. Holmes, J. Lumley, G. Berkooz, J. Mattingly, and R. Wittenberg, Low-dimensional models of coherent structures in turbulence, *Phys. Rep.*, **287** (1997), 337–384.

[33] L. Ju, Q. Du, and M. Gunzburger, Probablistic methods for centroidal Voronoi tessellations and their parallel implementations, *Parallel Computing*, **28** (2002), 1477–1500.

[34] T. Kanungo, D. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Wu, An efficient k-means clustering algorithm: Analysis and implementation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **24** (2002), 881-892.

[35] S. Lloyd; Least squares quantization in PCM, *IEEE Trans. Infor. Theory*, **28** (1982), 129–137.

[36] J. Lumley, *Stochastic Tools in Turbulence*, Academic, New York, (1971).

[37] J. MacQueen; Some methods for classification and analysis of multivariate observations, in *Proc. 5th Berkeley Symposium on Math. Stat. and Prob. I,* 1967, pp. 281–297.

[38] A. Mendes and I. Themido, Multi-outlet retail site location assessment. *Inter. Trans. in Operational Research*, **11** (2004), 1-18.

[39] A. Okabe, B. Boots, K. Sugihara, and S. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, 2nd edition, Wiley, Chichester, (2000).

[40] M. Rathinam and L. Petzold, A new look at proper orthogonal decomposition, *SIAM J. Numer. Anal.*, 41 (2003), 1893-1925.

[41] S. Ravindran, Proper orthogonal decomposition in optimal control of fluids, *Int. J. Numer. Meth. Fluids*, **34** (2000), 425–448.

[42] N. Smaoui and D. Armbruster, Symmetry and the Karhunen-Loève analysis, *SIAM J. Sci. Comput.*, **18** (1997), 1526–1532.

[43] S. Valette and J. Chassery, Approximated Centroidal Voronoi Diagrams for Uniform Polygonal Mesh Coarsening, *Computer Graphics Forum*, 2004, **23**, pp.381–390.

[44] S. Volkwein, Optimal control of a phase field model using the proper orthogonal decomposition, *ZAMM*, **81** (2001), 83–97.

[45] C. Wager, B. Coulland N. Lange, Modelling spatial intensity for replicated inhomogeneous point patterns in brain imaging, *J. Royal Stat. Soc. B*, **66** 2004, 429–446.

[46] G. Yagawa, Node-by-node parallel finite elements: a virtually meshless method, *International Journal for Numerical Methods in Engineering*, **60** (2004), 69–102.