UNIFORM CONVERGENCE OF A NONLINEAR ENERGY-BASED MULTILEVEL QUANTIZATION SCHEME*

QIANG DU† AND MARIA EMELIANENKO‡

Abstract. A popular vector quantization scheme can be constructed by the Centroidal Voronoi tessellations (CVTs) which also have many other applications in diverse areas of science and engineering. The development of efficient algorithms for their construction is a key to the successful applications of CVTs in practice. This paper studies the details of a new optimization based multilevel algorithm for the numerical computation of the CVTs. The rigorous proof of its uniform convergence in one space dimension and the results of computational simulations are provided. They substantiate recent claims on the significant speedup demonstrated by the new scheme in comparison with traditional methods.

Key words. Optimal quantization, Centroidal Voronoi Tessellations, Lloyd's method, multilevel method, uniform convergence

1. Introduction. Optimal vector quantization is used in many applications such as image and data compression, pattern recognition, image rendering and so on [20]. A vector quantizer maps N-dimensional vectors in the domain $\Omega \subset \mathbb{R}^N$ into a finite set of vectors $\{\mathbf{z}_i\}_{i=1}^k$. Each vector $\mathbf{z}_i$ is called a code vector or a *codeword*, and the set of all the codewords is called a codebook. A special quantization scheme is given by the Voronoi tessellation which associates with each codeword $\mathbf{z}_i$, also called a *generator*, a nearest neighbor region that is called a Voronoi region $\{V_i\}_{i=1}^k$. That is, for each i, V_i consists of all points in the domain Ω that are closer to $\mathbf{z}_i$ than to all the other generating points, and a Voronoi tessellation refers to the tessellation of a given domain by the Voronoi regions $\{V_i\}_{i=1}^k$ associated with a set of given generating points $\{\mathbf{z}_i\}_{i=1}^k \subset \Omega$ [1, 32].

With a suitably defined distortion measure, an optimal quantizer can be described as a Centroidal Voronoi tessellation. 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

$$\mathbf{z}_i^* = \Big(\int_{V_i} \mathbf{y}\rho(\mathbf{y})\,d\mathbf{y}\Big)\Big(\int_{V_i} \rho(\mathbf{y})\,d\mathbf{y}\Big)^{-1}. \tag{1.1}$$

Then, an *optimal quantization* may be constructed through a *centroidal Voronoi tessellation* (CVT) for which the generators of the Voronoi tessellation are the centroids of their respective Voronoi regions, in other words, $\mathbf{z}_i = \mathbf{z}_i^*$ for all i. Such a connection between the CVTs and optimal quantization schemes has been explored extensively in the literature [12].

Given a set of points $\{\mathbf{z}_i\}_{i=1}^k$ and a tessellation $\{V_i\}_{i=1}^k$ of the domain, we may define the *energy functional* or the *distortion value* for the pair $(\{\mathbf{z}_i\}_{i=1}^k, \{V_i\}_{i=1}^k)$ by:

$$\mathcal{G}\Big(\{\mathbf{z}_i\}_{i=1}^k, \{V_i\}_{i=1}^k\Big) = \sum_{i=1}^k \int_{V_i} \rho(\mathbf{y})|\mathbf{y}-\mathbf{z}_i|^2\,d\mathbf{y}\,. \tag{1.2}$$

*This research was supported in part by grants from US NSF DMS-0409297 and NSF ITR-0205232.

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

‡Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213. (masha@cmu.edu).

The minimizer of $\mathcal{G}$, that is, the optimal quantizer, necessarily forms a CVT which illustrates the optimization property of the CVT [12]. The terms optimal quantizer and CVT are thus to be used interchangeably in the sequel. For more studies on optimal quantization schemes, we refer to [20, 22, 36]. We note that, besides providing an optimal least square vector quantizer design in electrical engineering applications, the CVT concept also has applications in diverse areas such as astronomy, biology, image and data analysis, resource optimization, sensor networks, geometric design, and numerical partial differential equations [4, 7, 12, 13, 14, 15, 16, 17, 25, 26, 31, 37, 38]. We refer to [12] for a more comprehensive review of the mathematical theory and diverse applications of CVTs.

In the seminal work of Lloyd on the least square quantization [30], one of the algorithms proposed for computing optimal quantizers is an fixed point type 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. For algorithms on the computation of Voronoi tessellations, we refer to [1, 18, 19]. It is easy to see that the Lloyd algorithm is an energy descent iteration of the energy functional (1.2), which gives strong indications to its practical convergence. We refer to [10] for some discussions on the recent development of a rigorous convergence theory.

Lloyd's algorithms and their variants have been proposed and studied in many contexts for different applications [14, 20, 29, 34]. A particular extension using parallel and probabilistic sampling was given in [26] which allows efficient and mesh free implementation of the Lloyd's algorithm. Still, Lloyd algorithm is at best linearly convergent. Moreover, it slows down as the number of generators gets large.

For modern applications of the CVT concept in large scale scientific and engineering problems such as data communication, vector quantization and mesh generation, efficient algorithms for computing the CVTs play crucial roles.

Recently, a new multilevel approach to the optimal quantization problem has been developed [8, 11]. The new multilevel scheme offers considerable speed up over the traditional methods such as the celebrated Lloyd iteration. It can be combined with other techniques together to accelerate the computation of CVTs and the optimal vector quantizers. In this paper, we present a rigorous mathematical theory for the new algorithm. We focus on the main characteristics of this multilevel scheme, namely, the uniform convergence with respect to the problem size. Though the result is shown in one space dimension only, such a result is the first of its type in the vector quantization field and on the computation of CVTs. Proofs for higher dimensional cases are more involved, and they are now being worked out in our ongoing study [11].

The rest of the paper is organized as follows. In section 2, the optimization-based nonlinear multilevel algorithm is introduced. In section 3, the uniform convergence theory is established. Numerical results demonstrating its superiority over traditional methods are given in section 4. Final conclusions are made in section 5.

2. Optimization-based nonlinear multilevel algorithm. Since the original concept of centroidal Voronoi tessellations is related to the solution of a nonlinear optimization problem, and the monotone energy descent property is preserved by the Lloyd's iteration [12], we may thus investigate whether monotone energy reduction can be achieved in a multilevel procedure which would also improve the performance of the simple-minded iteration.

The problem of constructing a CVT is nonlinear in nature, hence standard linear multigrid theory cannot be directly applied. There are still several ways one could implement a nonlinear multilevel scheme in this context (see [8, 9, 27, 28]). A Newton type acceleration method, studied earlier in [9], is based on some global linearization as the outer loop, coupled with other fast solvers in the inner loop. Alternatively, we now study an approach that overcomes the difficulties of the nonlinearity by essentially relying on the direct energy minimization without any type of global linearization.

Let us define the energy functional

$$\mathcal{H}\Big(\{\mathbf{z}_i\}_{i=1}^k\Big) = \mathcal{G}\Big(\{\mathbf{z}_i\}_{i=1}^k, \{V_i\}_{i=1}^k\Big) \tag{2.1}$$

where $\{V_i\}_{i=1}^k$ forms the Voronoi tessellation with generators $\{\mathbf{z}_i\}_{i=1}^k$. The CVTs and optimal quantizers are closely related to the problem of minimization of the functional $\mathcal{H}$. In fact, we may note that the vector of generators of a CVT forms a critical point of $\mathcal{H}$ and vice versa [12]. That is, at a CVT (or optimal quantizer), we have $\nabla\mathcal{H} = 0$. This is one of the important characterizations of the CVTs which will be used in the later discussion.

In our study here, we follow the ideas presented in the literature on the extension of multigrid ideas to nonlinear optimization problems (see [5, 21, 23, 24, 35] and the references cited therein). However, we observe that a direct application of multilevel schemes to the minimization of $\mathcal{H}$ is not the best strategy. We illustrate later in this section that an alternative minimization formulation can be introduced to make the problem of computing CVTs more amenable for the multilevel framework.

2.1. Space decomposition. Since the energy functional is in general non-convex, it turns out to be very effective to relate our problem to an equivalent optimization problem through a technique that mimics the role of a dynamic nonlinear preconditioner. More precisely, denote $R = \mathtt{diag}(R_i^{-1}, R_2^{-1}, ..., R_k^{-1})$ a diagonal matrix whose diagonal entries $\{R_i = \int_{V_i} \rho(\mathbf{y})\, d\mathbf{y}\}$ are the masses of the corresponding Voronoi cells. It is easy to deduce that $R\nabla\mathcal{H} = 0$ at a CVT. Hence we arrive at an equivalent formulation of the minimization problem: $\min ||R\nabla\mathcal{H}||^2$, with respect to the standard Euclidean norm.

A key observation is that as R varies with respect to the generators, the above transformation or *dynamic preconditioning* keeps the modified objective functional convex in a suitably large neighborhood of the minimizer and therefore makes the new formulation more amenable to analysis than the original problem. Now, if we define the set of iteration points $\mathbf{W}$ by

$$\mathbf{W} = \{(w_i)|_{i=0}^{k+1}| \quad 0 = w_0 \leq w_i \leq w_{i+1} \leq w_{k+1} = 1,\ \forall\, 0 \leq i \leq k\}\ ,$$

we can design a new multilevel algorithm based on the following nonlinear optimization problem

$$\min_{\mathbf{Z}\in\mathbf{W}} \tilde{\mathcal{H}}(\mathbf{Z}), \text{ where } \tilde{\mathcal{H}}(\mathbf{Z} = \{\mathbf{z}_i\}_{i=0}^{k+1}) = ||R\nabla\mathcal{H}(\{\mathbf{z}_i\}_{i=1}^k)||^2\ . \tag{2.2}$$

The functional $\tilde{\mathcal{H}}$ may be regarded as a dynamically preconditioned energy. Viewing $\mathbf{W}$ as a set of grid points in the unit interval, we may denote by $\mathcal{T} = \mathcal{T}_J$ a finite element mesh corresponding to $\mathbf{W}$, and consider a sequence of nested quasi-uniform finite element meshes $\mathcal{T}_1 \subset \mathcal{T}_2 \subset \ldots \mathcal{T}_J$, where $\mathcal{T}_i$ consists of all finite element meshes $\{\tau_j^i\}_{j=1}^{n_i}$ with mesh parameter h_i, such that $\cup_{j=1}^{n_i}\tau_j^i = \Omega$. Corresponding to each finite

element partition $\mathcal{T}_i$ $(i = 1, ..., J)$, there is a finite element space $\mathbf{W}_i$ defined by

$$\mathbf{W}_i = \{v \in H^1(\Omega) \ | \quad v|_\tau \in \mathcal{P}_1(\tau), \forall \tau \in \mathcal{T}_i\} \ .$$

For each $\mathbf{W}_i$, there corresponds a nodal basis $\{\psi_j^i\}_{j=1}^{n_i}$, such that $\psi_j^i(x_k^i) = \delta_{jk}$, where δ_{jk} is the usual Kronecker Delta function and $\{x_k^i\}_{k=1}^{n_i}$ is the set of all nodes of the elements of $\mathcal{T}_i$ with $x_1^J = 0, x_{n_J}^J = 1$. Define the corresponding one-dimensional subspaces $\mathbf{W}_{i,j} = \text{span}\{\psi_j^i\}$. Then the decomposition can be regarded as

$$\mathbf{W}_J = \sum_{i=1}^{J} \sum_{j=1}^{n_i} \mathbf{W}_{i,j} = \bigoplus_{i=1}^{J} \bar{\mathbf{W}}_i$$

where $\bar{\mathbf{W}}_i = \mathbf{W}_i / \mathbf{W}_{i-1}$ for $i > 1$ and $\bar{\mathbf{W}}_1 = \mathbf{W}_1$. Now clearly for each $\psi_j^i \in \mathbf{W}_i$, we can find a vector $\bar{\psi}_j^i = \{\bar{\psi}_{jm}^i\} \in \mathbb{R}^{n_J}$, such that $\psi_j^i(x) = \sum_{m=1}^{n_J} \bar{\psi}_{jm}^i \psi_m^J(x)$, for $x \in \Omega$.

The above setup naturally applies to higher dimensional cases, where $\mathbf{W}$ consists of grids points in a higher dimensional domain and is discussed in more details in [11]. Here we will restrict our attention to the 1-dimensional case. It may be then noted that the set of basis functions

$$Q_i = [\bar{\psi}_1^i, \ldots, \bar{\psi}_{n_i}^i]^T \in \mathbb{R}^{n_i \times k}$$

used at each iteration can be pre-generated using the recursive procedure: $Q_J = I_{k \times k}$ and $Q_{J-s} = (\Pi_{i=1}^{s} P_{J-i}) Q_J$ where P_i is the basis transformation from space $\mathbf{W}_{i+1}$ to $\mathbf{W}_i$ which plays a role of a restriction operator.

2.2. Description of the algorithm. Using the above notations, we design the Algorithm 2.1 which is a multilevel successive subspace correction algorithm [5, 39]. Each step of the procedure outlined below involves solving a system of nonlinear equations which plays the role of relaxation. We can use the Newton iteration to solve this nonlinear system. Solution at current iterate is updated after each nonlinear solve by the Gauss-Seidel type procedure, hence the resulting scheme is successive in nature. The "slash" cycle can be defined as follows.

ALGORITHM 2.1. **Successive correction $V(\nu_1, 0)$ scheme**
Input:
k, number of generators; $u_1 = \{z_i\}_{i=0}^{k+1}$, the ends plus the initial set of generators.
Output after n cycles:
$u_n = \{z_i\}_{i=0}^{k+1}$, the ends plus the set of generators for CVT $\{V_i\}_{i=1}^{k}$.
Method: For $n > 1$, given u_n, do
1. For $i = 1 : J$
$\bar{u}_{n+\frac{i-1}{J}} = u_{n+\frac{i-1}{J}}$
For $l = 1 : \nu_1$
$\bar{u}_{n+\frac{i-1}{J}} = \bar{u}_{n+\frac{i-1}{J}} + \alpha_{j\,l}^0 \bar{\psi}_j^i \in \bar{\mathbf{W}}_i$ *sequentially for $1 \le j \le n_i$,*
such that $\tilde{\mathcal{H}}(\bar{u}_{n+\frac{i-1}{J}} + \alpha_{j\,l}^0 \bar{\psi}_j^i) = \min_{\alpha_{jl}} \tilde{\mathcal{H}}(\bar{u}_{n+\frac{i-1}{J}} + \alpha_{j\,l} \bar{\psi}_j^i)$,
endfor.
$u_{n+\frac{i}{J}} = \bar{u}_{n+\frac{i-1}{J}} = u_{n+\frac{i-1}{J}} + e_n^i$, *where* $e_n^i = \sum_{l=1}^{\nu_1} \sum_{j=1}^{n_i} \alpha_{j\,l}^0 \bar{\psi}_j^i$
endfor.
2. On the coarsest level, $u_{n+1} \leftarrow$CoarseGridSolve(u_{n+1}).
3. n = n+1
4. Repeat the procedure 1 to 3 until some stopping criterion is met.

Here ν_1, ν_2 denote the number of Gauss-Seidel iterations (smoothings) used at each level. Although it is enough to have $\nu_{1,2} = 1$ in theory, larger values need to be used in practice due to the numerical error in solving the nonlinear system. The values $\nu_{1,2} \leq 3$ usually suffice for the optimization to reach saturation. In the above description, no post-smoothings are used to make the analysis more transparent. A complete $V(\nu_1, \nu_2)$ cycle with ν_2 post-smoothings can be defined and anlayzed in a similar fashion. The algorithm uses a procedure $CoarseGridSolve(\mathbf{Z})$, which, as the name indicates, refers to finding the solution at the coarsest level. In our implementation, this procedure consists of applying Lloyd method for a few steps or until saturation. In general, other efficient optimization methods, as well as Newton's method, can be used in order to quickly damp the error, since the number of unknowns on the coarsest grid remains relatively small. The algorithm essentially only depends on the proper space decompositions and its relation to the set of generators, thus it is applicable in any dimension. The more general forms, including detailed description of the grid coarsening procedure, will be discussed in our subsequent works.

3. The uniform convergence theory. The uniform convergence of the new multilevel scheme can be rigorously proved, at least, for a large class of density functions in one dimensional space. This is the first step towards a more comprehensive theoretical analysis of this type of multilevel schemes in general. Here, let us first establish some important properties of the energy functional defined in 2.2. Then, we introduce our main convergence results. Without further notice, all the analysis in this section is restricted to the one dimensional case only and we further without loss of generaity assume that the domain is simply the unit interval $\Omega = (0, 1)$, which can always be transformed into any other intervals.

3.1. Technical lemmas. First, we can supply each $y = u - v$, where $u, v \in \mathbf{W}$ and $y_0 = y_{k+1} = 0$, with the following norm:

$$||y||^2_{\mathbf{W}} = \frac{1}{k} \sum_{i=1}^{k+1} (y_i - y_{i-1})^2 \ .$$

In the discussion that follows we say that a functional F satisfies the *convexity* and *continuity* properties in W, if there exist constants $K > 0, L > 0$, satisfying

$$(F'(w) - F'(v), w - v) \geq K||w - v||^2_W, \forall w, v \in W \ , \tag{3.1}$$

$$(F'(w) - F'(v), w - v) \leq L||w - v||^2_W, \forall w, v \in W \ . \tag{3.2}$$

To simplify the presentation, let us introduce the following notations: for $i = 1, \ldots, k$, $u_i^- = \frac{u_i + u_{i-1}}{2}$, $u_{i-1}^+ = \frac{u_{i-1} + u_i}{2}$, $a_i = u_i - u_{i-1}$, $b_i = w_i - w_{i-1}$, and $x_i = u_i - w_i$, with $u_0 = w_0 = 0$, $u_{k+1} = w_{k+1} = 1$ being the fixed ends of the interval.

Let us now first turn our attention to the case of a constant density. In this simple case, with the above notation, we get the following result for the preconditioned energy functional,

PROPOSITION 3.1. *Let $\rho(x) = 1$ be the density function on $[0, 1]$. Then the following relation holds for all $u, w \in W$:*

$$(\tilde{\mathcal{H}}'(u) - \tilde{\mathcal{H}}'(w), u - w) = \frac{1}{2} \sum_{i=1}^{k+1} (a_i - b_i)^2 \ .$$

Proof.

$$\frac{\partial\tilde{\mathcal{H}}}{\partial u_i} = 2(u_i - T_i) = 2\Big(u_i - \frac{u_i^+ + u_i^-}{2}\Big) = \frac{1}{2}(a_i - a_{i+1})$$

$$(\tilde{\mathcal{H}}'(u) - \tilde{\mathcal{H}}'(w), u - w) = \frac{1}{2}\sum_{i=1}^{k}(u_i - w_i)(a_i - a_{i+1} - b_i + b_{i+1}) =$$

$$= \frac{1}{2}\sum_{i=1}^{k}(u_i - w_i)(a_i - b_i) - \frac{1}{2}\sum_{i=1}^{k}(u_{i-1} - w_{i-1})(a_i - b_i) = \frac{1}{2}\sum_{i=1}^{k+1}(a_i - b_i)^2 \ .$$

□

COROLLARY 3.1. *For a constant density, the functional $F = k^{-1}\tilde{\mathcal{H}}$ satisfies the continuity and convexity conditions with $K = L = 1/2$ for all points in $\mathbf{W}$.*

Note that this is a simple consequence of the fact $\sum_{i=1}^{k+1}(a_i - b_i)^2 = k||u - w||^2_{\mathbf{W}}$.

One can extend the above result to a broader class of density functions. First, the following auxiliary lemma can be verified.

LEMMA 3.2. *Let $m \le g(x) \le M$ on $[0,1]$, $M' = \sup_{x\in[0,1]} g'(x)$ and define*

$$Q_i(u) = \frac{\int_{u_i^-}^{u_i^+}(2u - (u_i^+ + u_i^-))g(u)du}{(u_i^+ - u_i^-) + \epsilon\int_{u_i^-}^{u_i^+} g(u)du}.$$

Then we get

$$|Q_i(u) - Q_i(w)| \le (M + M' + \epsilon M^2 + 3\epsilon MM')\alpha_i \ ,$$

where $\alpha_i = |u_i^+ - w_i^+| + |u_i^- - w_i^-|$.

Proof. Denote

$$Q_i(u) = \frac{\int_{u_i^-}^{u_i^+}(2u - (u_i^+ + u_i^-))g(u)du}{(u_i^+ - u_i^-) + \epsilon\int_{u_i^-}^{u_i^+} g(u)du} = \frac{N_i}{D_i} \ .$$

In order to represent the above expression in a more convenient form, we employ the following change of variables argument:

$$\int_\alpha^\beta f(x)\,dx = (\beta - \alpha)\int_{-\frac{1}{2}}^{\frac{1}{2}} f\Big(\frac{\alpha + \beta}{2} + z(\beta - \alpha)\Big)\,dz \ .$$

Introducing $-\frac{1}{2} \le z \le \frac{1}{2}$, we can rewrite $u(z) = \dfrac{u_i^+ + u_i^-}{2} + z(u_i^+ - u_i^-)$ and $w(z) = \dfrac{w_i^+ + w_i^-}{2} + z(w_i^+ - w_i^-)$. Then for the numerator we have:

$$\begin{aligned} N_i(u) &= \textstyle\int_{u_i^-}^{u_i^+}(2u - (u_i^+ + u_i^-))g(u)du \\ &= 2(u_i^+ - u_i^-)\textstyle\int_{-\frac{1}{2}}^{\frac{1}{2}} z(u_i^+ - u_i^-)g(u(z))\,dz \\ &= 2(u_i^+ - u_i^-)^2\textstyle\int_{-\frac{1}{2}}^{\frac{1}{2}} g(u(z))z\,dz \end{aligned}$$

while the denominator is equal to

$$D_i(u) = (u_i^+ - u_i^-)\Big(1 + \epsilon\int_{-\frac{1}{2}}^{\frac{1}{2}} g(u(z))\,dz\Big) \ .$$

For the ratio $Q_i = \dfrac{N_i}{D_i}$ we have

$$\frac{N_i(u)}{D_i(u)} = \frac{2(u_i^+ - u_i^-)\int_{-\frac{1}{2}}^{\frac{1}{2}} g(u(z))z\,dz}{1+\epsilon\int_{-\frac{1}{2}}^{\frac{1}{2}} g(u(z))\,dz}\ .$$

For simplicity let us redefine the modified numerator as

$$\tilde{N}_i(u) = 2(u_i^+ - u_i^-)\int_{-\frac{1}{2}}^{\frac{1}{2}} g(u(z))z\,dz$$

and denominator as

$$\tilde{D}_i(u) = 1+\epsilon\int_{-\frac{1}{2}}^{\frac{1}{2}} g(u(z))\,dz\ .$$

In the new notations,

$$\begin{aligned}|Q_i(u)-Q_i(w)| &\le \tfrac{1}{\tilde{D}_i(u)\tilde{D}_i(w)}\Big|\tilde{N}_i(u)\tilde{D}_i(w)-\tilde{N}_i(w)\tilde{D}_i(u)\Big| \\ &\le \tfrac{1}{2}\Big|\tilde{N}_i(u)-\tilde{N}_i(w)\Big|\Big(\tilde{D}_i(u)+\tilde{D}_i(w)\Big) \\ &\quad +\tfrac{1}{2}\Big(\tilde{N}_i(u)+\tilde{N}_i(w)\Big)\Big|\tilde{D}_i(u)-\tilde{D}_i(w)\Big|\ .\end{aligned}$$

Notice further that $|u(z)-w(z)| \le |u_i^+ - w_i^+| + |u_i^- - w_i^-| = \alpha_i$. It follows that

$$\begin{aligned}&\Big|\tilde{D}_i(u)-\tilde{D}_i(w)\Big| = \epsilon\Big|\int_{-\frac{1}{2}}^{\frac{1}{2}}(g(u(z))-g(w(z)))\,dz\Big| \le \epsilon M'\alpha_i\ , \\ &\Big|\tilde{N}_i(u)-\tilde{N}_i(w)\Big| = 2\Big|\int_{-\frac{1}{2}}^{\frac{1}{2}}((u_i^+-u_i^-)g(u(z))-(w_i^+-w_i^-)g(w(z)))z\,dz\Big| \\ &\qquad \le (M+M')\alpha_i\ .\end{aligned}$$

Finally, since $1 \le \tilde{D}_i \le 1+\epsilon M$, and $\tilde{N}_i \le 2M$, we have

$$|Q_i(u)-Q_i(w)| \le \frac{1}{2}\Big(2(M+M')(1+\epsilon M)+4\epsilon MM'\Big)\alpha_i$$

so that

$$|Q_i(u)-Q_i(w)| \le (M+M'+\epsilon M^2+3\epsilon MM')\alpha_i\ . \tag{3.3}$$

This gives the result of the lemma. □

With the help of Lemma 3.2 we can derive the following

PROPOSITION 3.2. *For any $\rho(x) = 1+\epsilon g(x)$ with ϵ suitably small, there exist constants C_l and C_u such that for any $u, w \in \mathbf{W}$ and with the notation defined earlier,*

$$C_l\sum_{i=1}^{k+1}(a_i-b_i)^2 \le (\tilde{\mathcal{H}}'(u)-\tilde{\mathcal{H}}'(w), u-w) \le C_u\sum_{i=1}^{k+1}(a_i-b_i)^2\ .$$

Proof. Direct calculation gives

$$\frac{\partial\tilde{\mathcal{H}}}{\partial u_i} = \frac{2}{M_i(u)}\Big(\int(u_i-u)du+\epsilon\int(u_i-u)g(u)du\Big)$$

$$= \frac{(u_i^+ - u_i^-)(2u_i - u_i^+ - u_i^-) + 2\epsilon \int (u_i - u) g(u) du}{(u_i^+ - u_i^-) + \epsilon \int g(u) du}$$

$$= 2u_i - u_i^+ - u_i^- + \frac{2\epsilon \int (u_i - u) g(u) du - \epsilon (u_i^+ - u_i^-) \int g(u) du}{(u_i^+ - u_i^-) + \epsilon \int g(u) du} =$$

$$= \frac{1}{2}\Big(a_i - a_{i+1}\Big) - \epsilon \Big(\frac{\int (2u - (u_i^+ + u_i^-)) g(u) du}{(u_i^+ - u_i^-) + \epsilon \int g(u) du} \Big) = \frac{1}{2}\Big(a_i - a_{i+1}\Big) - \epsilon Q_i$$

where the integrals all refer to the integral over $[u_i^-, u_i^+]$. Then,

$$(\tilde{\mathcal{H}}'(u) - \tilde{\mathcal{H}}'(w), u - w) = \frac{1}{2} \sum_{i=1}^{k} (a_i - b_i)^2 - \epsilon (Q(u) - Q(w), u - w) \; . \tag{3.4}$$

The first term in (3.4) comes from the constant part of the density and hence complies with the results of the previous theorem. It remains to get a similar estimation for the second term. From the Cauchy inequality,

$$|(Q(u) - Q(w), u - w)| \leq \sum_{i=1}^{k} |Q_i(u) - Q_i(w)| \cdot |u_i - w_i| \; . \tag{3.5}$$

Combining (3.3), (3.4) and (3.5), we get

$$\begin{aligned}
(\tilde{\mathcal{H}}'(u) - \tilde{\mathcal{H}}'(w), u - w) &\leq \sum_{i=1}^{k} \frac{(a_i - b_i)^2}{2} \\
&\quad + 2\epsilon (M + M' + \epsilon M^2 + 3\epsilon M M') \sum_{i=1}^{k} (|u_i^+ - w_i^+| + |u_i^- - w_i^-|) |u_i - w_i| \\
&\leq \sum_{i=1}^{k} \frac{(a_i - b_i)^2}{2} + 4\epsilon (M + M' + \epsilon M^2 + 3\epsilon M M') \sum_{i=1}^{k} |u_i - w_i|^2 \; .
\end{aligned}$$

Since $x_0 = 0$, we have the inequality $\sum\limits_{i=1}^{k} x_i^2 \leq 2k \sum\limits_{i=1}^{k} (x_i - x_{i-1})^2$. So,

$$\begin{aligned}
&(\tilde{\mathcal{H}}'(u) - \tilde{\mathcal{H}}'(w), u - w) \\
&\quad \leq \Big(\tfrac{1}{2} + 8k\epsilon (M + M' + \epsilon M^2 + 3\epsilon M M') \Big) \sum_{i=1}^{k} (a_i - b_i)^2 \leq k C_u ||u - w||_W^2 \; ,
\end{aligned}$$

for some constant C_u with $k\epsilon$ being suitably and uniformly small. Same arguments applied to the lower bound yield

$$\begin{aligned}
&(\tilde{\mathcal{H}}'(u) - \tilde{\mathcal{H}}'(w), u - w) \\
&\quad \geq \Big(\tfrac{1}{2} - 8k\epsilon (M + M' + \epsilon M^2 + 3\epsilon M M') \Big) \sum_{i=1}^{k} (a_i - b_i)^2 \geq k C_l ||u - w||_W^2 \; .
\end{aligned}$$

for some constant C_l. The proposition is thus proved. □

Note that it follows from Proposition 3.2 that the functional $F = k^{-1}\tilde{\mathcal{H}}$ preserves the continuity and convexity for the perturbation of the constant density with $\epsilon = o(k^{-1})$.

In addition to showing that the energy functional possesses convexity and continuity properties, we also need the following conditions on the space decomposition to be satisfied:

Condition A.
$\forall v \in \mathbf{W}, \exists v_i \in \bar{\mathbf{W}}_i s.t. \sum_{i=1}^{J} v_i = v$, and

$$(\sum_{i=1}^{J} ||v_i||^2_{\bar{\mathbf{W}}_i})^{1/2} \leq C_1 ||v||_{\mathbf{W}} .$$

for some C_1 independent of k.

Condition B. 'Strengthened Cauchy-Schwartz'
$\forall w_{ij} \in \mathbf{W}, u_i \in \bar{\mathbf{W}}_i, v_j \in \bar{\mathbf{W}}_j \Rightarrow$

$$\sum_{i,j=1}^{J} (F'(w_{ij} + u_i) - F'(w_{ij}), v_j) \leq C_2 (\sum_{i=1}^{J} ||u_i||^2_{\bar{\mathbf{W}}_i})^{1/2} (\sum_{j=1}^{J} ||v_j||^2_{\bar{\mathbf{W}}_j})^{1/2} .$$

for some $C_2 \leq cL$, where c is independent of k.

These conditions are verified below. First, we have

THEOREM 3.3. *For the nested subspace decomposition with the choice of the 'hat' basis functions,* $(\sum_{i=1}^{J} ||v_i||^2_{\bar{\mathbf{W}}_i})^{1/2} = ||v||_{\mathbf{W}}$, *so that* $C_1 = 1$. *Moreover, for* $F = k^{-1}\tilde{\mathcal{H}}$, C_2 *can be estimated as* $C_2 = L \cdot \max_j (\sum_{l=1}^{J} 2^{-|j-l|}) \leq 2L$.

Proof. Notice that 'hat' functions form an orthogonal basis, so

$$(\sum_{i=1}^{J} ||v_i||^2_{\bar{\mathbf{W}}_i})^{1/2} = ||v||_{\mathbf{W}}$$

follows easily from calculation. As for the C_2, first notice that for any $w, u, v \in \mathbf{W}$, using the arguments similar to those in the proof of Propositions 3.1 and 3.2 and bounding the l_2-norm by the $\mathbf{W}$-norm, we can get

$$\begin{aligned}
(\tilde{\mathcal{H}}'(u+w) - \tilde{\mathcal{H}}'(w), v) &= \frac{1}{2} \sum_{i=1}^{k+1} (v_i - v_{i-1})(u_i - u_{i-1}) - \epsilon(Q(u+w) - Q(w), v) \\
&= \frac{1}{2} \sum_{supp(u) \cap \mathrm{supp}(v)} (v_i - v_{i-1})(u_i - u_{i-1}) - \epsilon(Q(u+w) - Q(w), v) \\
&\leq \frac{k}{2} ||u||_{\mathbf{W}, supp(u) \cap \mathrm{supp}(v)} \cdot ||v||_{\mathbf{W}, supp(u) \cap \mathrm{supp}(v)} + \epsilon \tilde{M} \sum_{supp(u) \cap \mathrm{supp}(v)} |u_i||v_i| \\
&\leq k(1/2 + 4\epsilon\tilde{M}) ||u||_{\mathbf{W}, supp(u) \cap \mathrm{supp}(v)} \cdot ||v||_{\mathbf{W}, supp(u) \cap \mathrm{supp}(v)} \\
&= kL ||u||_{\mathbf{W}, supp(u) \cap \mathrm{supp}(v)} \cdot ||v||_{\mathbf{W}, supp(u) \cap \mathrm{supp}(v)}
\end{aligned}$$

where $\tilde{M} = 2(M + M' + \epsilon M^2 + 3\epsilon M M')$ as defined in Proposition 3.2 and $L = 1/2 + 4\epsilon\tilde{M}$ is the continuity constant.

Now since $\forall u \in \mathbf{W}_j, v \in \mathbf{W}_l$, we have $supp(u) \cap supp(v) = 2^{-|j-l|} supp(v)$ in the case of one-dimensional *hat* basis. Moreover,

$$||v||_{\mathbf{W}, supp(v)\cap supp(u)} = 2^{-|j-l|}||v||_{\mathbf{W}}.$$

It follows that, since for any symmetric matrix $||Ax|| \le \rho(A)||x|| \le \max_i (\sum_{j=1}^{n} |a_{ij}|)||x||$,

$$\sum_{i,j=1}^{J} (\tilde{\mathcal{H}}'(w_{ij}+u_i) - \tilde{\mathcal{H}}'(w_{ij}), v_j) \le kL \sum_{i,j=1}^{J} 2^{-|i-j|}||u_i||_{\mathbf{W}}||v_j||_{\mathbf{W}}$$
$$\le L\Big(\max_j \sum_{i=1}^{J} 2^{-|i-j|}\Big)\Big(\sum_{i=1}^{J} ||u_i||^2_{\mathbf{W}_i}\Big)^{1/2}\Big(\sum_{j=1}^{J} ||v_j||^2_{\mathbf{W}_j}\Big)^{1/2}.$$

Henceforth, we have $C_2 = L \cdot \max_j \sum_{i=1}^{J} 2^{-|i-j|} \le 2L$.

Note that although the proof above is presented for the special case of the 'hat' basis that we used in our numerical implementation, similar arguments can be used to show that Conditions A and B hold for other suitable bases. Indeed, since the energy part of the argument does not depend on the decomposition, Conditions A and B will remain true for any basis as long as

$$(\sum_{i=1}^{J} ||v_i||^2_{\mathbf{W}_i})^{1/2} \le C_1||v||_{\mathbf{W}}$$

and

$$||v||_{\mathbf{W}, supp(v)\cap supp(u)} \le c2^{-|j-l|}||v||_{\mathbf{W}}$$

for some constants c, C_1 independent of k.

3.2. Uniform convergence theorem. Finally, putting together Conditions A and B and using convexity and continuity of $F = k^{-1}\tilde{\mathcal{H}}$ in $\mathbf{W}$, we are ready to prove the following uniform convergence result:

THEOREM 3.4. *Under Conditions A and B on space decomposition, Algorithm 3.1 converges uniformly in* $\mathbf{W}$ *for any density of the type* $\rho(x) = 1 + \epsilon g(x)$ *with sufficiently small* ϵ. *Moreover,* $d_n = \tilde{\mathcal{H}}(u_n) - \tilde{\mathcal{H}}(u)$ *satisfies*

$$d_n \le r d_{n-1}, \quad \rho \in (0,1)$$

where $r = \frac{C}{1+C}$ *and* $C = C_1^2 C_2^2 L/K^3$.

Before proceeding to its proof, let us first state a consequence:

COROLLARY 3.5. *In the case of a "'hat"' basis, the constants* C_1 *and* C_2 *can be estimated as* $C_1 = 1$ *and* $C_2 = 2L$, *so for example when* $\rho(x) = 1$, $C = 4$.

The proof of this result is similar to the one given in [35] and relies on the following technical lemma:

LEMMA 3.6. *Suppose the functional* F *satisfies the continuity and convexity conditions in* $\mathbf{W}$. *Then the following statements are true for all points* $v, w \in \mathbf{W}$:

$$F(u) - F(v) \ge (F'(v), u-v) + \tfrac{K}{2}||u-v||^2_{\mathbf{W}}$$
$$F(u) - F(v) \le (F'(v), u-v) + \tfrac{L}{2}||u-v||^2_{\mathbf{W}}.$$

Proof. Let $\phi(\lambda) = F(u + \lambda(v-u))$. Then $\phi'(\lambda) = (v-u, F'(u+\lambda(v-u)))$, $\phi(0) = F(u)$, $\phi(1) = F(v)$. First inequality can be verified using fundamental theorem of calculus and convexity assumption:

$$
\begin{aligned}
&F(u) - F(v) = \phi(0) - \phi(1) = -\int_0^1 \phi'(t)dt \\
&= -\int_0^1 (v-u, F'(u+t(v-u)))dt = \int_0^1 (u-v, F'(u+t(v-u)))dt \\
&= \int_0^1 (u-v, F'(u+t(v-u)) - F'(v))dt + \int_0^1 (u-v, F'(v))dt \\
&= (F'(v), u-v) + \int_0^1 (F'(u+t(v-u)) - F'(v), u+t(v-u)-v)\frac{dt}{1-t} \\
&\geq (F'(v), u-v) + K\int_0^1 ||(1-t)(u-v)||^2_{\mathbf{W}}\frac{dt}{1-t} = (F'(v), u-v) + \frac{K}{2}||u-v||^2_{\mathbf{W}} .
\end{aligned}
$$

The proof of the second inequality is analogous and follows from the continuity of functional F. □

3.3. Proof of Theorem 3.4.

Proof. Denote u to be the exact solution of (2.2). Consider u_n - the approximate solution after one J-level iteration of the Algorithm 2.1. At the i-th level, since the supports of the basis functions $\bar{\psi}^i_j$ are disjoint in $\bar{\mathbf{W}}_i$, we have

$$
u_{n+\frac{i}{J}} = u_{n+\frac{i-1}{J}} + e^i_n, \quad \tilde{\mathcal{H}}(u_{n+\frac{i-1}{J}} + e^i_n) \leq \tilde{\mathcal{H}}(u_{n+\frac{i-1}{J}} + v_i), \forall v_i \in \bar{\mathbf{W}}_i,
$$

where in the notations of the algorithm, $e^i_n = \sum_{l=1}^{\nu_1}\sum_{j=1}^{n_i} \alpha^0_{j\,l}\bar{\psi}^i_j \in \bar{\mathbf{W}}_i$.

First notice that since the minimizer $u_{n+\frac{i}{J}}$ satisfies $(\tilde{\mathcal{H}}'(u_{n+\frac{i}{J}}), v) = 0, \forall v \in \bar{\mathbf{W}}_i$, it follows from Lemma 3.6, that

$$
\begin{aligned}
&\tilde{\mathcal{H}}(u_n) - \tilde{\mathcal{H}}(u_{n+1}) = \sum_{i=1}^{J}(\tilde{\mathcal{H}}(u_{n+\frac{i-1}{J}}) - \tilde{\mathcal{H}}(u_{n+\frac{i}{J}})) \\
&\geq \sum_{i=1}^{J}\Big((\tilde{\mathcal{H}}'(u_{n+\frac{i}{J}}), u_{n+\frac{i-1}{J}} - u_{n+\frac{i}{J}}) + \frac{K}{2}||u_{n+\frac{i-1}{J}} - u_{n+\frac{i}{J}}||^2_{\bar{\mathbf{W}}_i}\Big) = \frac{K}{2}\sum_{i=1}^{J}||e^i_n||^2_{\bar{\mathbf{W}}_i} .
\end{aligned}
$$

Next, let us use Condition A to decompose $u_{n+1} - u = \sum_{i=1}^{J} v_i$. Then

$$
\begin{aligned}
&(\tilde{\mathcal{H}}'(u_{n+1}) - \tilde{\mathcal{H}}'(u), u_{n+1} - u) = (\tilde{\mathcal{H}}'(u_{n+1}), u_{n+1} - u) \\
&= \sum_{i=1}^{J}(\tilde{\mathcal{H}}'(u_{n+1}), v_i) = \sum_{i=1}^{J}(\tilde{\mathcal{H}}'(u_{n+1}) - \tilde{\mathcal{H}}'(u_{n+\frac{i}{J}}), v_i) \\
&= \sum_{i=1}^{J}\sum_{j\geq i+1}^{J}(\tilde{\mathcal{H}}'(u_{n+\frac{j}{J}}) - \tilde{\mathcal{H}}'(u_{n+\frac{j-1}{J}}), v_i) \\
&\leq C_2(\sum_{j=1}^{J}||e^j_n||^2_{\bar{\mathbf{W}}_j})^{1/2}(\sum_{i=1}^{J}||v_i||^2_{\bar{\mathbf{W}}_i})^{1/2} .
\end{aligned}
$$

Hence

$$(\tilde{\mathcal{H}}'(u_{n+1}) - \tilde{\mathcal{H}}'(u), u_{n+1} - u) \le C_1 C_2 (\sum_{j=1}^{J} ||e_n^j||^2_{\tilde{\mathbf{W}}_j})^{1/2} ||u_{n+1} - u||_{\mathbf{W}}$$
$$\le C_1 C_2 (\frac{2}{K}(\tilde{\mathcal{H}}(u_n) - \tilde{\mathcal{H}}(u_{n+1})))^{1/2} ||u_{n+1} - u||_{\mathbf{W}} .$$

Denote $r_n = \tilde{\mathcal{H}}(u_n) - \tilde{\mathcal{H}}(u)$, then $\tilde{\mathcal{H}}(u_n) - \tilde{\mathcal{H}}(u_{n+1}) = r_n - r_{n+1}$ and it follows from the inequality above that

$$\Big(\frac{2}{K}(r_n - r_{n+1})\Big)^{1/2} \ge \frac{(\tilde{\mathcal{H}}'(u_{n+1}) - \tilde{\mathcal{H}}'(u), u_{n+1} - u)}{C_1 C_2 ||u_{n+1} - u||_{\mathbf{W}}} .$$

Thus,

$$r_n - r_{n+1} \ge \frac{K}{2}\Big(\frac{(\tilde{\mathcal{H}}'(u_{n+1}) - \tilde{\mathcal{H}}'(u), u_{n+1} - u)}{C_1 C_2 ||u_{n+1} - u||_{\mathbf{W}}}\Big)^2$$
$$\ge \frac{K}{2}(C_1 C_2)^{-2} K^2 ||u_{n+1} - u||^2_{\mathbf{W}} \ge \frac{K^3}{C_1^2 C_2^2 L} r_{n+1} .$$

The last step of the argument uses the result of Lemma 3.6:

$$r_{n+1} = \tilde{\mathcal{H}}(u_{n+1}) - \tilde{\mathcal{H}}(u) \le \frac{L}{2} ||u_{n+1} - u||^2_{\mathbf{W}} .$$

As a consequence, we get

$$r_{n+1} \le \frac{C_1^2 C_2^2 L}{K^3}(r_n - r_{n+1}) \Rightarrow r_{n+1} \le \frac{C}{1+C} r_n, \text{where } C = \frac{C_1^2 C_2^2 L}{K^3} .$$

Since for any basis satisfying Conditions A and B, $C_1^2 C_2^2 L/K^3 = O(L^3/K^3)$, the convergence is uniform as long as the ratio L/K does not depend on k. By looking at the L and K estimates obtained in Proposition 3.2, it is easy to see that this condition is satisfied for any smooth perturbation of the constant density $1 + \epsilon g(x)$ with ϵ of the order of $O(1/k)$. This concludes the proof of the main theorem. □

It follows that, for a suitable choice of decomposition in 1D, the asymptotic convergence factor of our multilevel algorithm is independent of the size of the problem and the number of grid levels, which gives a significant speedup comparing to other methods, like the traditional Lloyd iteration. This claim is further substantiated by the following numerical examples. Although for the sake of simplicity we have presented the detailed theory only for the case of the $V(\nu_1, 0)$ multigrid cycle with no post smoothings, same conclusions can be drawn for the case of the full $V(\nu_1, \nu_2)$ cycle, and the results of the numerical experiments in both cases are outlined below.

4. Numerical examples. We now report some numerical results of using the new multilevel algorithm.

First we compare the results of our $V(1,1)$ multigrid implementation with the usual Lloyd iteration for the one dimensional problem. Then, we present some results for a two dimensional test problem in a parallelogram domain. The results are obtained with the Matlab 6.5 implementation of the new algorithm. The test runs are performed on a PC with a Pentium IV processor and 512MB RAM.

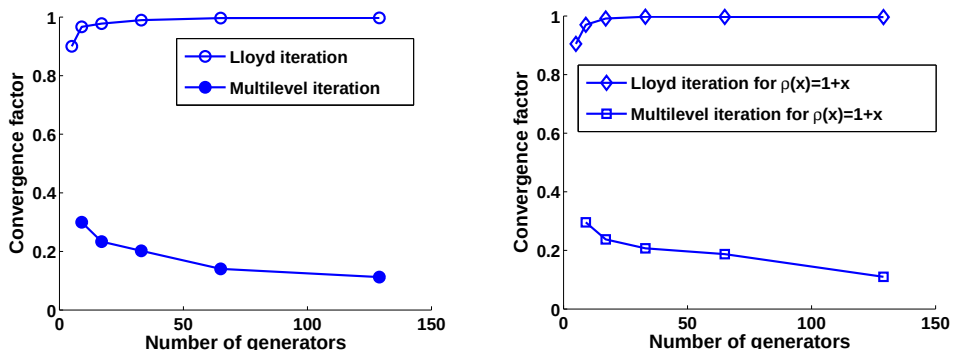

FIG. 4.1. *Plot of the convergence factor vs. the number of generators for the regular Lloyd (upper) and the multilevel (lower curves) iterations for* $\rho(x) = 1$ *(top) and* $\rho(x) = 1+0.1x$ *(bottom).*

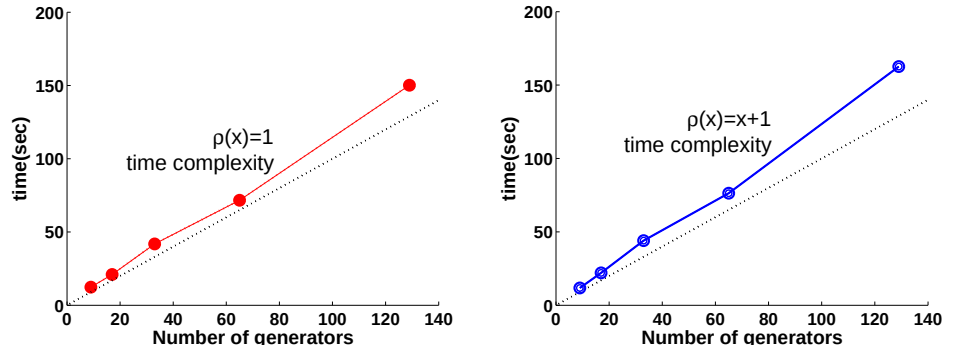

FIG. 4.2. *Computational time vs. problem size for the 1d implementation.*

The one dimensional implementation is very straightforward. Here, we take the unit interval and test a couple of different density functions, like $\rho(x) = 1$ and $\rho(x) = 1 + x$. We plot the convergence factor $\rho \approx |z_{n+1} - z_n|/|z_n - z_{n-1}|$ for each $V(1,1)$ cycle with respect to the total number of generators (grid points) involved.

Figure 4.1 substantiates the fact that the speed of convergence for the proposed scheme remains nearly constant as the number of generators increases. In figure 4.2, the computational time needed for the V(1,0) implementation of the multilevel method to reach 10^{-12} accuracy is given for $\rho(x) = 1$ and $\rho(x) = 1 + x$ respectively. The graph shows that in the 1-d case, the computational time *scales almost linearly* with the problem size.

The data in the table 4.1 shows the stabilization of the number of multigrid cycles $V(\nu_1, \nu_2)$ needed to reduce the error to $\epsilon = 10^{-12}$ in the constant density case. While there is a visible difference between the number of iterations required for $V(1,1)$ and $V(2,2)$ cycles respectively, saturation occurs if the values of the relaxation parameters are increased, which is why $\nu_{1,2} \leq 2$ in most of our calculations. The geometric rate of the energy and error reduction asserted by the Theorem 3.4 are confirmed by the experiments. Indeed, Figure 4.3 shows the convergence history of a $V(1,1)$-cycle against the total number of relaxations for the $k = 129$ case.

Next, the convergence factors for some two dimensional problems on a parallelogram domain are compared in Figure 4.4. The bottom graph shows the convergence

TABLE 4.1
The number of $V(\nu_1, \nu_2)$ cycles needed vs. the number of generators.

$k/V(\nu_1, \nu_2)$	V(1,0)	V(0,1)	V(1,1)	V(2,0)	V(0,2)	V(2,2)
3	7	8	6	6	7	4
5	11	11	8	8	8	6
9	13	14	9	9	9	7
17	18	18	12	12	12	8
33	21	20	13	12	13	8
65	21	22	12	12	12	8
129	21	21	12	12	12	8
257	20	23	12	12	13	7
513	20	22	12	11	13	7
1025	19	22	11	11	13	7

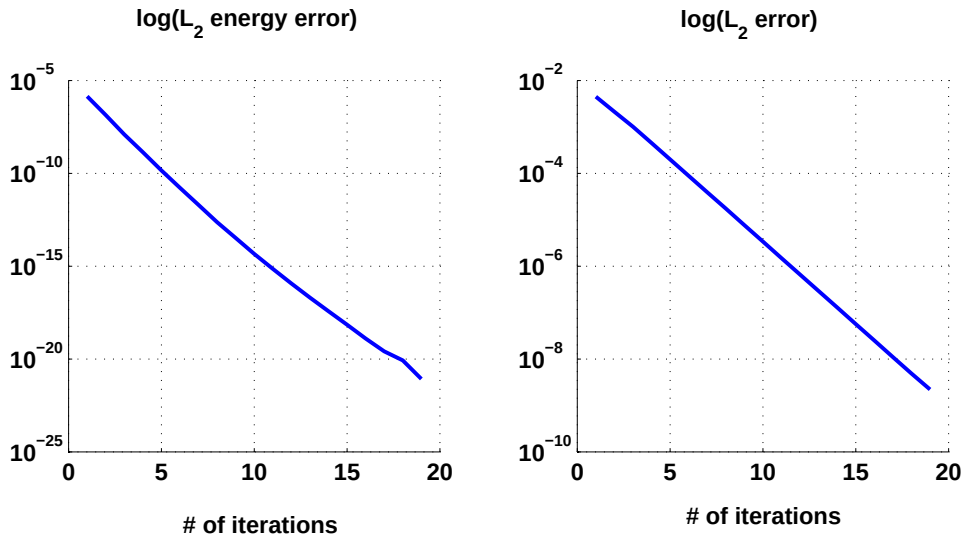

FIG. 4.3. *The energy reduction (left) and the convergence history (right) for* 129 *generators in the log-normal scale.*

factor for the compatible relaxation, that is, a relaxation performed on the grid with the exact solution given at the coarse nodes (see [2] for discusions on the compatible relaxation). This factor can serve as s lower bound on the convergence factor of the full multigrid cycle, and the quality of the coarse grid influences the distance between the two graphs. As we can see from Figure 4.4, the result of the compatible relaxation in this case comes in good agreement with the convergence factor of the whole V(1,0) cycle given on top, which is an indication of a good quality for the coarsening procedure.

The convergence history plots are given in Figure 4.5. The top curves in both graphs depict the error reductions given by the Lloyd iteration, while the graphs below correspond to the convergence of the multigrid scheme for various problem sizes. One can clearly see that the slopes of error reduction on the logarithmic scale do not depend on the number of generators. We note that even though our theoretical results are only proved in 1d here, it is clear that it remains valid in the higher dimensional implementations.

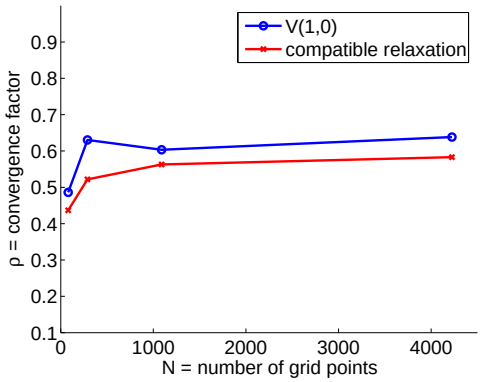

FIG. 4.4. Comparison of the convergence factors for the compatible relaxation and a V(1,0) cycle.

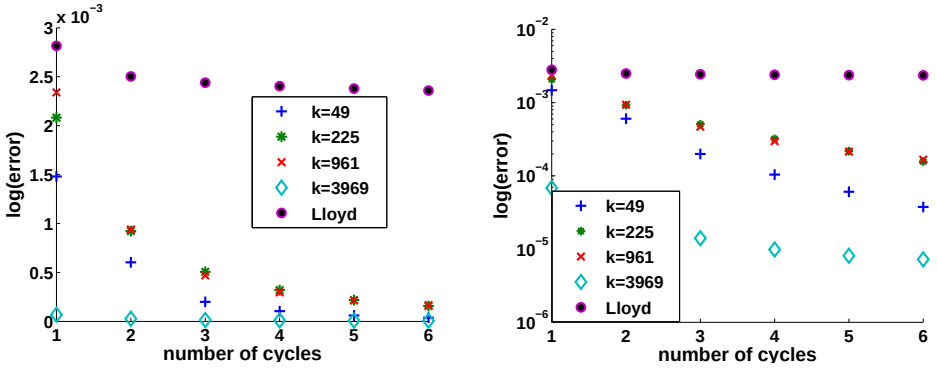

FIG. 4.5. (a) Convergence history for the multigrid scheme compared to the Lloyd scheme shown on the top curve in the normal scale (left) and log-normal scale (right).

5. Conclusion. Recently, several methods have been proposed for accelerating the convergence of the classical Lloyd iteration commonly used in the context of quantization and in the construction of centroidal Voronoi tessellations [8, 9]. These algorithmic advances are important for making the computation of optimal quantizer more efficient and for many other successful applications of CVTs. A few possible extensions that use multilevel techniques to accelerate the convergence of the CVTs have been suggested. One of such extensions uses some algebraic multigrid solvers in the spirit of [3, 6, 33] as preconditioners to accelerate the solution of the linear system at every Newton iteration [9], while the other adopts an energy based nonlinear multigrid approach with the use of a dynamic nonlinear preconditioning [8]. In this paper, we focus on the latter approach, and for the first time, a rigorous analysis of its convergence properties is presented for a class of one dimensional density functions, with the results of several numerical experiments given for both one and two dimensional cases. A more detailed analysis in the higher dimensional cases as well as more efficient and robust implementations of the energy based nonlinear multigrid approach are presently under investigation [11]. We conclude by commenting that there is obviously a great potential in using such multigrid methods to accelerate the optimal quantizer design and more generally the computation of centroidal Voronoi tessellations.

6. Acknowledgment. The authors would like to thank M. Gunzburger of FSU, L. Zikatanov and J. Xu of PSU for interesting discussions on the subject.

REFERENCES

[1] Aurenhammer F., Voronoi diagrams. A survey of a fundamental geometric data structure. *ACM Computing Surveys* 1990; **23**: 345-405.

[2] Brandt A., General highly accurate algebraic coarsening, *Electronic Transactions on Numerical Analysis* 2000; **10**: 1-20.

[3] Brezina M., Cleary A., Falgout R., Henson H., Jones J., Manteuffel T., McCormick S. and Ruge J., Algebraic multigrid based on element interpolation. *SIAM Journal on Scientific Computing* 2000; **22**: 1570-1592.

[4] Cappellari M. and Copin Y., Adaptive spatial binning of integral-field spectroscopic data using Voronoi tessellations. *Monthly Notices of the Royal Astronomical Society* 2003; **342**: 345–354.

[5] Chan T. and I. Sharapov, Subspace correction multilevel methods for elliptic eigenvalue problems, *Numer. Linear Algebra Appl.* 1993; **1**: 1–7.

[6] Chang Q. and Huang Z., Efficient algebraic multigrid algorithms and their convergence. *SIAM Journal on Scientific Computing* 2002; **24**: 597–618.

[7] Cortes J., Martinez S., Karata T. and Bullo F., Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation* 2004; **20**: 243–255.

[8] Du Q., Emelianenko M., Nonlinear optimization-based Multilevel Quantization Scheme, to appear in proceesings of DD16 conference, 2005.

[9] Du Q. and Emelianenko M., Acceleration of algorithms for the computation of centroidal Voronoi tessellations, to appear in *Numerical Linear Algebra and Applications*, 2005

[10] Du Q., Emelianenko M. and Ju L., Convergence Properties of the Lloyd Algorithm for Computing the Centroidal Voronoi Tessellations, to appear in *SIAM Journal on Numerical Analysis* 2005.

[11] Du Q., Emelianenko M. and Zikatanov L., An Energy-based Multilevel Quantization Scheme in Multidimension, in preparation, 2005.

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

[13] Du Q. and Gunzburger M., Grid Generation and Optimization Based on Centroidal Voronoi Tessellations, *Applied Mathematics and Computation*, 2002; **133**: 591-607.

[14] Du Q., Gunzburger M. and Ju L., Meshfree, Probabilistic Determination of Points, Support Spheres, and Connectivities for Meshless Computing. *Computational Methods in Applied Mechanics and Engineering* 2002; **191**: 1349-1366.

[15] Du Q., Gunzburger M. and Ju L., Constrained centroidal Voronoi tessellations on general surfaces, *SIAM Journal on Scientific Computing* 2003; **24**:1488-1506.

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

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

[18] Dwyer R., Higher-dimensional Voronoi diagrams in linear expected time, *Discrete and Computational Geometry* 1991; **6**: 343-367.

[19] Fortune S., A sweepline algorithm for Voronoi diagrams. *Algorithmica* 1987; **2**: 153-174.

[20] Gray R. and Neuhoff D., Quantization. *IEEE Transactions on Information Theory* 1998; **44**: 2325–2383.

[21] Gelman E. and J. Mandel., On multilevel iterative methods for optimization problems. *Math. Programming*, 1990; **48**:1–17,

[22] Gersho A., Asymptotically optimal block quantization. *IEEE Transactions on Information Theory* 1979; **25**: 373–380.

[23] W. Hackbusch and A. Reusken, Analysis of a damped nonlinear multilevel method. *Numer. Math.*, **55**:225

[24] R. Kornhuber, Adaptive monotone multigrid methods for nonlinear variational problems. Teubner, Stuttgart, 1997

[25] Hiller S., Hellwig H. and Deussen O. Beyond Stippling Methods for Distributing Objects on the Plane. *Computer Graphics Forum* 2003; **22**: 515–522.

[26] Ju L., Du Q. and Gunzburger M., Probabilistic methods for centroidal Voronoi tessellations and their parallel implementations. *Parallel Computing* 2002; **28**: 1477-1500.

[27] Koren Y., Yavneh I. and Spira A., A Multigrid Approach to the 1-D Quantization Problem,

preprint, 2003.
[28] Koren Y., Yavneh I., Adaptive Multiscale Redistribution for Vector Quantization. *SIAM Journal on Scientific Computing* 2005, to appear.
[29] Linde Y., Buzo A. and Gray R., An algorithms for vector quantizer design. *IEEE Transactions on Communications* 1980; **28**: 84–95.
[30] Lloyd S., Least square quantization in PCM. *IEEE Transactions on Information Theory* 1982; **28**: 129–137.
[31] Mendes A. and Themido I., Multi-outlet retail site location assessment. *International Transactions in Operational Research* 2004; **11**: 1–18.
[32] Okabe A., Boots B. and Sugihara K., *Spatial Tessellations; Concepts and Applications of Voronoi Diagrams,*. Wiley: Chichester, 1992.
[33] Ruge J. and Stuben K., Algebraic multigrid, in *Mutigrid methods. Fronties in applied mathematics. SIAM*; Philadelphia, 1987: 73-130.
[34] Sabin J. and Gray R., Global Convergence and Empirical Consistency of the Generalized Lloyd Algorithm, *IEEE Transactions on Information Theory* 1986; **Vol. IT-32**: no.2.
[35] Tai X. and Xu J., Global Convergence of Subspace Correction Methods for Some Convex Optimization Problems. *Mathematics of Computation* 2002; **71**: 105-124.
[36] Trushkin A., On the design of an optimal quantizer. *IEEE Transactions on Information Theory* 1993; **39**: 1180–1194.
[37] Wager C., Coull B. and Lange N., Modelling spatial intensity for replicated inhomogeneous point patterns in brain imaging. *Journal of Royal Statistical Society B* 2004; **66**:429–446.
[38] Valette S. and Chassery J., Approximated Centroidal Voronoi Diagrams for Uniform Polygonal Mesh Coarsening. *Computer Graphics Forum* 2004; **23**: 381–390.
[39] Xu J. and Zikatanov L., The method of alternating projections and the method of subspace corrections in Hilbert space. *Journal of Americal Mathematical Society*, 2002; **15**:573–597.