Concepts for Higher Order Finite Elements on Sparse Grids

H.-J. Bungartz *

Abstract

On the way to an efficient implementation of finite element algorithms related to the p- and h-p-versions on sparse grids, we present a general concept for the construction of hierarchical bases of higher order suitable for sparse grid methods. For the solution of partial differential equations, this approach allows us to profit both from the efficiency of sparse grid discretizations and from the advantages of higher order basis functions with regard to their approximation accuracy.

We discuss the general relations of sparse grids and higher order techniques, and we report the results of some first numerical experiments concerning piecewise biquadratic hierarchical basis functions.

Key words: finite element method, hierarchical bases, higher order techniques, partial differential equations, p- and h-p-version, sparse grids.

AMS subject classifications: 65N22, 65N30, 65N50.

1 Introduction

Since their presentation in 1990 [18], sparse grids have turned out to be a very interesting approach to use for the efficient solution of partial differential equations and for a lot of other topics in numerical analysis like numerical integration [5] or FFT [14]. In comparison to the standard full grid approach, the number of grid points can be reduced significantly from $O(N^d)$ to $O(N(\log_2(N))^{d-1})$ or even $O(N)$ in the d-dimensional case; whereas the accuracy of the sparse grid interpolant and of the approximation to the solution of the given boundary value problem, resp., is only slightly deteriorated. For piecewise d-linear basis functions, an accuracy of the order $O(N^{-2}(\log_2(N))^{d-1})$ with respect to the L_2- or the maximum norm and of the order $O(N^{-1})$ with respect to the energy norm has been shown [7]. Furthermore, regular sparse grids can be extended in a very simple and natural manner to adaptive ones, which makes the hierarchical sparse grid concept applicable to problems that require adaptive grid refinement, too.

For the two-dimensional case, the results mentioned above show that, apart from the logarithmic factor and with respect to the L_2-norm, sparse grid techniques with piecewise bilinear (biquadratic, ...) hierarchical basis functions correspond to full grid methods of fourth (sixth, ...) order. In the three-dimensional case, the gain in order is even more impressive. Therefore, sparse grid methods themselves can be considered as an approach of higher order. Additionally, together with polynomials of higher degree as basis functions, sparse grids are well-suited for the efficient realization of higher order finite element methods. Finally, implementing p- or h-p-version-type algorithms on sparse grids seems to be a very promising approach that allows us to profit not only from the sparse grid efficiency, but from the advantages of usual h-adaptivity, and the improved approximation quality of higher order basis functions.

In this paper, first, a short introduction to sparse grid methods recalls their most important properties. Furthermore, an overview of existing high order concepts for sparse grids is provided. Then, we present a new approach for generating higher order hierarchical bases on sparse grids, followed by some first numerical results for the case of piecewise biquadratic basis functions. Finally, some concluding remarks and an outlook on further work to be done will close the discussion.

2 Sparse grids

The use of hierarchical bases for finite element discretizations as proposed by Yserentant [17] and Bank, Dupont,

*Institut für Informatik der Technischen Universität München, D-80290 München, Germany.

ICOSAHOM'95: Proceedings of the Third International Conference on Spectral and High Order Methods.

and Yserentant [4] instead of standard nodal bases stood at the beginning of the sparse grid idea, together with a tensor-product-type approach for the generalization from the one-dimensional to the d-dimensional case. For the corresponding subspace splitting of a full grid discretization space in two dimensions with piecewise bilinear hierarchical basis functions as in figure 1, it can be seen that the dimension (i.e., the number of grid points) of all subspaces with $i_1 + i_2 = c$ is 2^{c-2}. Furthermore, it has been shown in [7] that the contribution of all those subspaces with $i_1 + i_2 = c$ to the interpolant of a function u is of the same order $O(2^{-2c})$ with respect to the L_2- or maximum norm and $O(2^{-c})$ with regard to the energy norm, if u fulfills the smoothness requirement $\frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} \in \mathcal{C}^0(\bar{\Omega})$ for the two-dimensional and $\frac{\partial^{2d} u}{\partial x_1^2 \ldots \partial x_d^2} \in \mathcal{C}^0(\bar{\Omega})$ for the general d-dimensional case, respectively. Here, Ω denotes the underlying domain. Therefore, it turns out to be more reasonable to deal with a triangular subspace scheme as given in figure 2 instead of using the quadratic scheme of figure 1. This leads us to the so-called *sparse grids*. For a formal definition of sparse grids, see [6], [7], or [18], e.g.

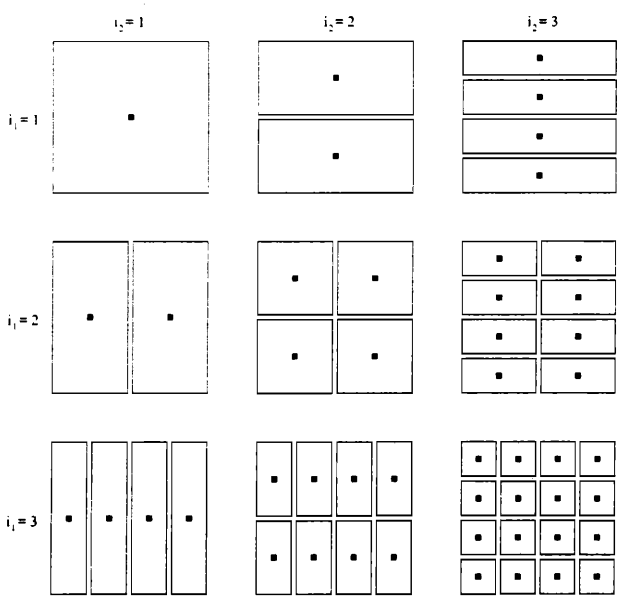

Figure 1: Subspace splitting of a full grid space.

Besides the regular sparse grids that result from skipping certain subspaces according to figure 2, adaptive grid refinement can be realized in the sparse grid context in a very straightforward way. Since we use recursive dynamic data structures like binary trees for the implementation, and since the value of a hierarchical basis function, the hierarchical surplus, can be used itself to indicate the smoothness

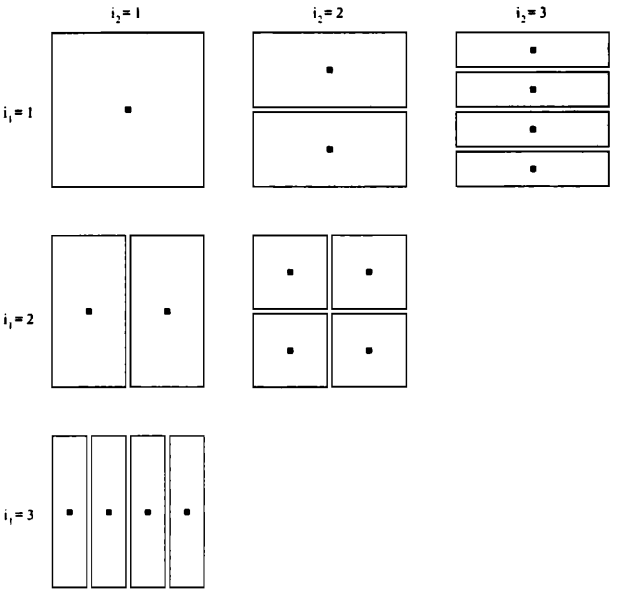

Figure 2: Subspace splitting of a sparse grid space.

of u at the corresponding grid point and, consequently, the necessity to refine the grid here, no additional work has to be done to implement adaptive refinement. Figure 3 shows a two-dimensional regular sparse grid, and figure 4 shows a three-dimensional adaptive one with singularities at the re-entrant corner and along the edges.

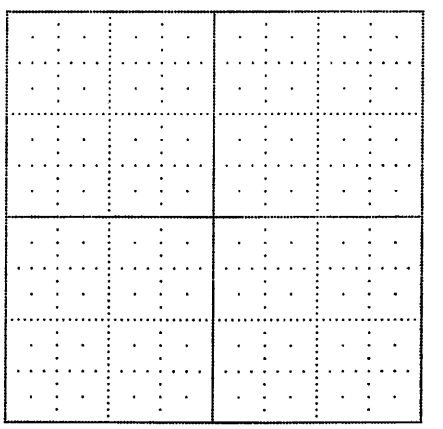

Figure 3: Regular sparse grid.

Speaking about the most important properties of sparse grids, we at least have to look at the number of grid points involved and at the approximation accuracy of piecewise d-linear hierarchical basis functions on sparse grids. For a

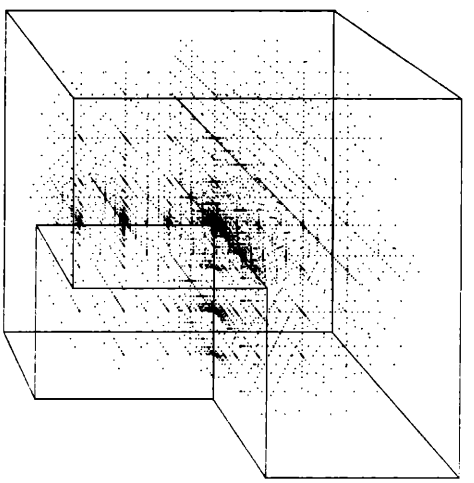

Figure 4: Adaptive sparse grid.

detailed analysis, we once again refer to [7] and [18]. For a d-dimensional problem, the approach described above and illustrated in figure 2 leads to regular sparse grids with $O(N(\log_2(N))^{d-1})$ grid points, if N denotes the number of grid points in one dimension (i. e., $\frac{1}{N}$ is the smallest mesh width occurring). A variant also discussed in [7] even leads to regular sparse grids with $O(N)$ grid points. These results have to be compared with the $O(N^d)$ points of regular full grids. Concerning the approximation quality, the accuracy of the sparse grid interpolant is only slightly deteriorated from $O(N^{-2})$ to $O(N^{-2}(\log_2(N))^{d-1})$ with respect to the L_2- or maximum norm. With regard to the energy norm, both the sparse grid interpolant and the finite element approximation to the solution of the given boundary value problem stay of the order $O(N^{-1})$.

Thus, sparse grids enable us to gain a factor of 2 in accuracy for arbitrary number d of dimensions by just doubling the number of grid points. Since the smoothness requirements can be overcome by adaptive grid refinement, sparse grids are a very efficient approach for the solution of partial differential equations.

Recently, the class of problems that can be treated with sparse grid methods has been significantly extended. First experiments with time-dependent problems have been reported by Balder et al. in [3]; Pflaum [15] generalized the algorithm for the solution of the Poisson equation to the case of general elliptic differential operators of second order in two dimensions, and Dornseifer developed a mapping technique to deal with curvilinear domains. Furthermore, systems of equations like the Stokes equations are the focus of present sparse grid interest.

3 A bicubic approach on sparse grids

A first step towards higher order techniques on sparse grids has been done by Störtkuhl [16]. There, the main emphasis is put on the solution of the two-dimensional Stokes equations. Using the stream function ψ and the vorticity ω as variables, this system of two partial differential equations of second order can be reduced to the fourth order biharmonic equation $\Delta^2\psi = 0$. The corresponding bilinear form a is symmetric, H_0^2-elliptic, and given by

$$a(u,v) = \int_\Omega \left(\frac{\partial^2 u}{\partial x^2} \cdot \frac{\partial^2 v}{\partial x^2} + 2 \cdot \frac{\partial^2 u}{\partial x \partial y} \cdot \frac{\partial^2 v}{\partial x \partial y} + \right.$$
$$\left. + \frac{\partial^2 u}{\partial y^2} \cdot \frac{\partial^2 v}{\partial y^2} \right) d(x,y).$$

Since this approach requires the use of C^1-elements, a piecewise cubic hierarchical Hermite basis is defined for the one-dimensional case, first. Here, we get two basis functions (i. e., two degrees of freedom to fix the value of the function and its first derivative) per grid point. For $d = 2$, the usual sparse grid tensor product approach leads to a piecewise bicubic hierarchical basis with four degrees of freedom per grid point. The resulting four different types of basis functions are shown in figure 5.

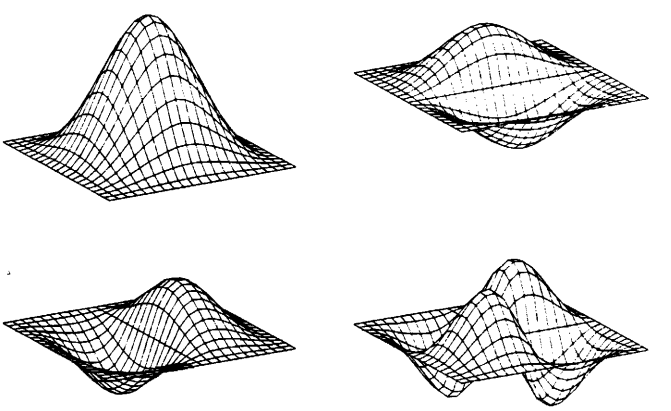

Figure 5: The four types of bicubic basis functions.

In the following, we present an alternative approach (cf. [8]) based on C^0-elements with still one degree of freedom per grid point.

4 Another concept for generating higher order hierarchical bases on sparse grids

4.1 A quadratic hierarchical basis

For reasons of clarity, let us study the one-dimensional case of a regular grid with N grid points, $N = 2^n + 1$, $n \in \mathbb{N}$, and N values to be interpolated, first. For the construction of a piecewise quadratic interpolant, one has to fix three degrees of freedom in each interval between two neighboring grid points. This leads to a total of $3N - 3$ degrees of freedom for the whole problem. It is well-known that quadratic splines are perhaps the most common way to construct a suitable interpolant. With splines, we need N degrees of freedom to get an interpolant and twice $N - 2$ degrees of freedom to make the interpolant both continuous and differentiable at the inner grid points. With one more condition fixed (some kind of boundary condition, e. g.), the interpolant is definitely determined. Thus, the higher order of the polynomials used leads to more smoothness of the interpolant. This effect is especially attractive, if smooth functions are to be interpolated, or if partial differential equations of higher order (like the biharmonic equation, e. g., see [16]) have to be solved. However, in a lot of other situations (like the numerical treatment of singularities, e. g.), it seems to be neither necessary nor desirable.

Therefore, we suggest a construction that leads to an interpolant (N degrees of freedom) which is only continuous ($N - 2$ degrees of freedom). The remaining $N - 1$ degrees of freedom are fixed by interpolation conditions outside the respective interval. For instance, the parabolic interpolant between two neighboring grid points i and $i + 1$, $1 \leq i \leq N - 1$, could be determined by either the values at the grid points $i - 1$, i, and $i + 1$ (if $i > 1$), or the values at the nodes i, $i + 1$, and $i + 2$ (if $i < N - 1$), or even the values at the grid points i, $i + 1$ and an arbitrary third point. Since we want to define hierarchical bases, it turns out to be the best choice to determine the third grid point for interpolation by means of an hierarchical criterion: If i is a grid point on the finest level only, i. e., if i is even, then $i - 1$, i, and $i + 1$ are taken into account. If, on the other hand, i is a coarse grid point (i. e. odd) and if, thus, $i + 1$ does appear on the finest grid only, then i, $i + 1$, and $i + 2$ are the points chosen for interpolation. The result of this approach is shown in figure 6. On the intervals $[2k + 1, 2k + 3]$, $0 \leq k \leq (N - 3)/2$, the resulting overall interpolant is quadratic, but at the (coarse) grid points $2k + 1$, it may not be differentiable.

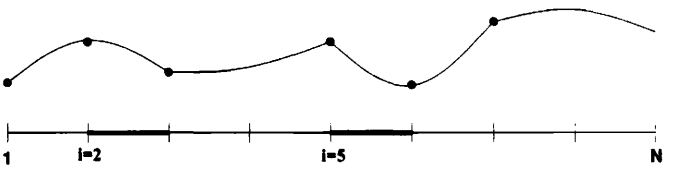

Figure 6: Piecewise quadratic C^0-interpolant.

Starting from these considerations, we now introduce a piecewise quadratic hierarchical basis. To explain the principles, we first look at the well-known piecewise linear case in one dimension. If we add appropriate basis functions at the coarse grid points to the hierarchical basis functions of each level, we get nested spaces of piecewise linear functions on the different levels (see figure 7). Here, a coarse grid function can be constructed by summing up three neighboring fine grid functions with the weights $\frac{1}{2}$, 1, and $\frac{1}{2}$. This is important for a simple switch from one level to another, and it is necessary for the efficient implementation of sparse grid algorithms.

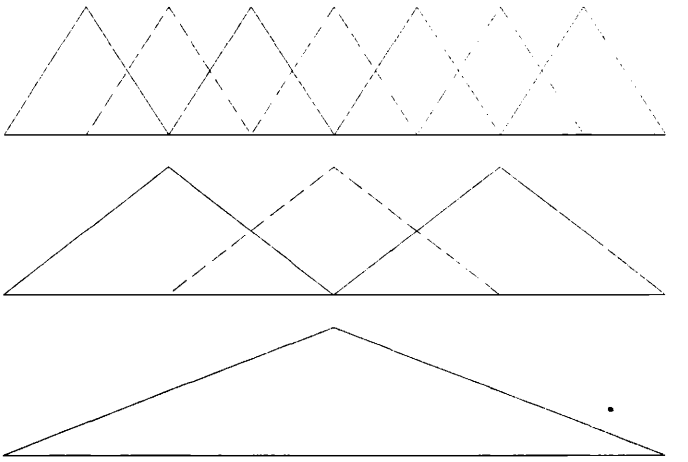

Figure 7: Linear hierarchical basis and nodal point bases on each level.

The quadratic case turns out to be a little bit more complicated, because it is not possible to get a quadratic basis function on the coarse grid as a weighted sum of three neighboring quadratic basis functions on the fine grid. However, if we sum up two quadratic fine grid functions with the weight $\frac{1}{4}$ and one standard piecewise linear coarse grid function with the weight 1 as indicated in figure 8, we get the desired quadratic function on the coarse grid.

Now, figure 9 shows our piecewise quadratic hierarchical basis (solid lines), together with the extension to a nodal

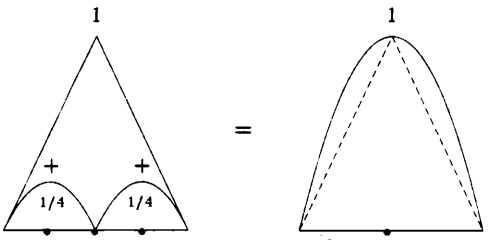

Figure 8: Switching from fine to coarse level with quadratic hierarchical basis functions.

point basis on each level (dashed lines). Note that each of these nodal bases consists of basis functions whose supports vary in size.

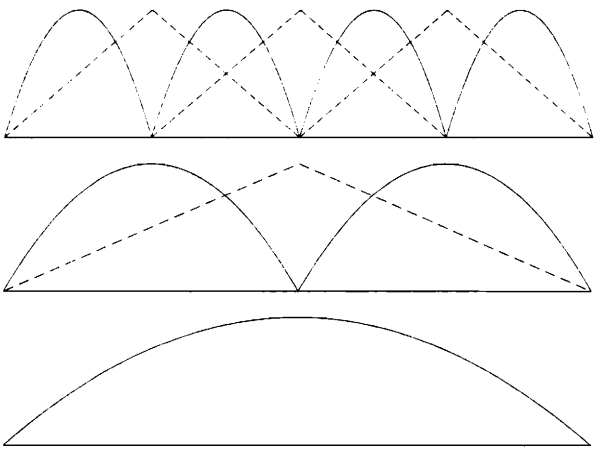

Figure 9: Quadratic hierarchical basis and nodal point bases on each level.

As in the linear case, the generalization to a d-dimensional piecewise d-quadratic hierarchical basis with $d > 1$ is done by the tensor product approach that is typical for the sparse grid context.

Another important problem we have to deal with is the question of how to calculate the (quadratic) hierarchical surplus. Again, we first look at the one-dimensional case. The linear hierarchical surplus $v_m^{(l)}$ in a grid point m with hierarchical neighbors $e(m)$ and $w(m)$ is given by

$$(1) \qquad v_m^{(l)} = u_m - \frac{1}{2} \cdot \left(u_{e(m)} + u_{w(m)}\right),$$

where u_m, $u_{e(m)}$, and $u_{w(m)}$ denote the values of the underlying function u at the respective grid points. Remember that the hierarchical neighbors of a grid point m are just the two ends of the support of the hierarchical basis function located in m. The corresponding formula for the quadratic hierarchical surplus $v_m^{(q)}$ depends on the hierarchical relations of the involved grid points. Figure 10 illustrates the situation if $e(m)$ is the father of m (with respect to the underlying binary tree) and if $e(e(m))$ is the father of $e(m)$.

A short calculation leads to

$$\begin{aligned} v_m^{(q)} &= -\frac{3}{8} \cdot u_{w(m)} + u_m - \frac{3}{4} \cdot u_{e(m)} + \frac{1}{8} \cdot u_{e(e(m))} \\ &= \left(-\frac{1}{2} \cdot u_{w(m)} + u_m - \frac{1}{2} \cdot u_{e(m)}\right) - \\ &\quad -\frac{1}{4}\left(-\frac{1}{2} \cdot u_{w(m)} + u_{e(m)} - \frac{1}{2} \cdot u_{e(e(m))}\right), \end{aligned}$$

i.e., the quadratic hierarchical surplus at a grid point m can be easily calculated with the help of the linear hierarchical surplus at m and the linear surplus at the father of m:

$$(2) \qquad v_m^{(q)} = v_m^{(l)} - \frac{1}{4} \cdot v_{e(m)}^{(l)}.$$

Thus, as in the linear case, nothing else has to be stored other than the linear hierarchical surplus. Again, the tensor product approach leads to a generalization of this result to the d-dimensional case with $d > 1$. For $d = 2$, e.g., we immediately get

$$(3) \quad v_m^{(q)} = v_m^{(l)} - \frac{1}{4} \cdot v_{e(m)}^{(l)} - \frac{1}{4} \cdot v_{n(m)}^{(l)} + \frac{1}{16} \cdot v_{ne(m)}^{(l)},$$

where $e(m)$ denotes the father of m in x-direction, $n(m)$ the father of m in y-direction, and $ne(m)$ the father of $n(m)$ in x-direction (see figure 11). For arbitrary d, the quadratic hierarchical surplus is given by

$$(4) \qquad v_m^{(q.d)} = \left[1, -\frac{1}{4}\right]^d \cdot v_\bullet^{(l.d)}.$$

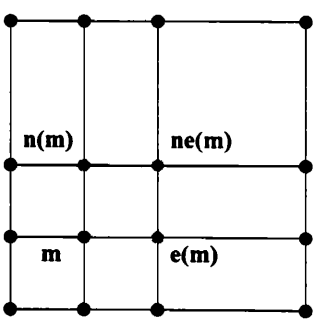

Figure 11: Calculation of the quadratic hierarchical surplus for $d = 2$.

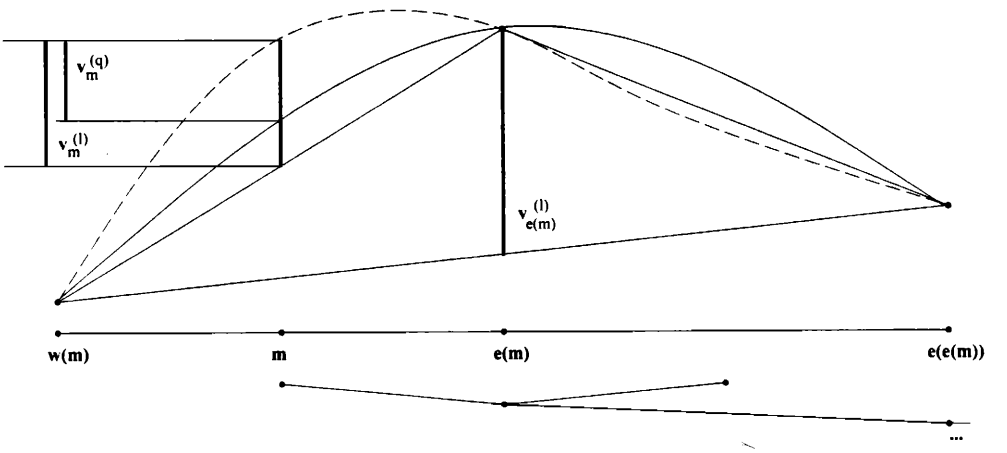

Figure 10: Linear and quadratic hierarchical surplus.

4.2 Theoretical results

Now, we turn to the approximation properties of sparse grids with the quadratic hierarchical basis introduced above. To this end, we study the behavior of the interpolation error with respect to the L_2-, the maximum, and the energy norm. According to finite element theory, the latter gives insight into the error of the finite element solution, too.

In the main, the notation and the argumentation follow the linear case from [7]. Because of (2) and (4), we look at

$$J_m^{(q,1)} := J_m^{(l,1)} - \frac{1}{4} \cdot J_{e(m)}^{(l,1)} \tag{5}$$

for the one-dimensional case or

$$J_m^{(q,d)} := \left[1, -\frac{1}{4}\right]^d \cdot J_\bullet^{(l,d)} \tag{6}$$

for the general d-dimensional case. Here, for some sufficiently smooth function $u^{(d)}$ of d independent variables, $J_m^{(l,d)}$ is the integral well-known from linear sparse grid theory,

$$J_m^{(l,d)} := \int_{-h_1}^{+h_1} \dots \int_{-h_d}^{+h_d} \left(\prod_{j=1}^{d} \frac{-h_j}{2} \cdot w_j(x_j)\right) \cdot \frac{\partial^{2d} u^{(d)}(x_1, \dots, x_d)}{\partial x_1^2 \dots \partial x_d^2} \, dx_d \dots dx_1, \tag{7}$$

at a grid point m (here normed to 0) with assigned piecewise linear hierarchical basis function $\prod_{j=1}^{d} w_j(x_j)$,

$$w_j(x_j) := \begin{cases} \frac{h_j + x_j}{h_j}, & \text{if } -h_j \le x_j \le 0, \\ \frac{h_j - x_j}{h_j}, & \text{if } 0 \le x_j \le h_j, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Since we know from [7] that $J_m^{(l,d)}$ is just the linear hierarchical surplus $v_m^{(l,d)}$ at point m, with (4) and (6) we get

$$\begin{aligned} J_m^{(q,d)} &= \left[1, -\frac{1}{4}\right]^d \cdot J_\bullet^{(l,d)} \\ &= \left[1, -\frac{1}{4}\right]^d \cdot v_\bullet^{(l,d)} \\ &= v_m^{(q,d)}. \end{aligned} \tag{9}$$

In the following, we study the situation for $d = 1$, first. Together with (7) and (8) for $d = 1$, (5) leads to

$$\begin{aligned} J_m^{(q,1)} &= J_m^{(l,1)} - \frac{1}{4} \cdot J_{e(m)}^{(l,1)} \\ &= \int_{-h_1}^{3h_1} t_1(x_1) \cdot \frac{\partial^2 u^{(1)}(x_1)}{\partial x_1^2} \, dx_1, \end{aligned}$$

where $e(m)$ again denotes the hierarchical father of m with assigned support $[-h_1, 3h_1]$ and

$$t_1(x_1) := \frac{1}{8} \cdot \begin{cases} -3x_1 - 3h_1, & -h_1 \le x_1 \le 0, \\ 5x_1 - 3h_1, & 0 \le x_1 \le h_1, \\ -x_1 + 3h_1, & h_1 \le x_1 \le 3h_1. \end{cases} \tag{10}$$

By partial integration for each sub-interval $[-h_1, 0]$, $[0, h_1]$, and $[h_1, 3h_1]$, and by elimination of the resulting h_1^2-terms (which is possible here in contrast to the linear case), we get

$$J_m^{(q,1)} = -\int_{-h_1}^{3h_1} T_1(x_1) \cdot \frac{\partial^3 u^{(1)}(x_1)}{\partial x_1^3}\, dx_1, \tag{11}$$

$$T_1(x_1) := \frac{1}{16} \cdot \cdot \begin{cases} -3x_1^2 - 6h_1x_1 - 3h_1^2, & -h_1 \le x_1 \le 0, \\ 5x_1^2 - 6h_1x_1 - 3h_1^2, & 0 \le x_1 \le h_1, \\ -x_1^2 + 6h_1x_1 - 9h_1^2, & h_1 \le x_1 \le 3h_1. \end{cases} \tag{12}$$

Together with (6) and (7), this result can be used to derive the generalization for the d-dimensional case. After a short calculation, we get

$$J_m^{(q,d)} = (-1)^d \cdot \int_{-h_1}^{3h_1} \cdots \int_{-h_d}^{3h_d} \left(\prod_{j=1}^{d} T_j(x_j) \right) \cdot \cdot \frac{\partial^{3d} u^{(d)}(x_1, ..., x_d)}{\partial x_1^3 \ldots \partial x_d^3}\, dx_d \ldots dx_1, \tag{13}$$

where $T_j(x_j)$ is defined in an analogous way to (12).

With (9) and (13), we are able to give two bounds for the quadratic hierarchical surplus $v_m^{(q,d)}$:

$$\left| v_m^{(q,d)} \right| = \left| J_m^{(q,d)} \right| \le \left\| \frac{\partial^{3d} u^{(d)}}{\partial x_1^3 \ldots \partial x_d^3} \right\|_\infty \cdot \frac{1}{2^d} \cdot h_1^3 \cdot \ldots \cdot h_d^3 \tag{14}$$

and

$$\left| v_m^{(q,d)} \right| \le \left\| \frac{\partial^{3d} \left(u^{(d)} \cdot \varphi_m^{(d)} \right)}{\partial x_1^3 \ldots \partial x_d^3} \right\|_2 \cdot \cdot \left(\frac{17}{160} \right)^{d/2} \cdot h_1^{5/2} \cdot \ldots \cdot h_d^{5/2}, \tag{15}$$

where $\varphi_m^{(d)}(x_1, \ldots x_d)$ denotes the characteristic function of the support of the basis function located at point $(h_1, ..., h_d)$, if m is normed to the origin. Note that (14) and (15) are correct only if we really have a quadratic surplus in each direction. On the coarsest level in some coordinate direction j (i.e., $i_j = 1$ or $x_j = \frac{1}{2}$ for the unit square), however, figure 9 shows that, in spite of using a quadratic basis function, we have to use the linear hierarchical surplus with respect to the boundary values. Therefore, we get for the general case of a point m with k indices $i_1, ..., i_k$ equaling 1 and the others being greater than 1 ($0 \le k \le d$)

$$\left| v_m^{(q,d)} \right| \le \left\| \frac{\partial^{3d-k} u^{(d)}}{\partial x_1^2 \ldots \partial x_k^2 \partial x_{k+1}^3 \ldots \partial x_d^3} \right\|_\infty \cdot \cdot \frac{1}{2^d} \cdot h_1^2 \cdot \ldots \cdot h_k^2 \cdot h_{k+1}^3 \cdot \ldots \cdot h_d^3 \tag{16}$$

and

$$\left| v_m^{(q,d)} \right| \le \left\| \frac{\partial^{3d-k} \left(u^{(d)} \cdot \varphi_m^{(d)} \right)}{\partial x_1^2 \ldots \partial x_k^2 \partial x_{k+1}^3 \ldots \partial x_d^3} \right\|_2 \cdot \cdot \left(\frac{1}{6} \right)^{k/2} \cdot \left(\frac{17}{160} \right)^{(d-k)/2} \cdot \cdot h_1^{3/2} \cdot \ldots \cdot h_k^{3/2} \cdot h_{k+1}^{5/2} \cdot \ldots \cdot h_d^{5/2}. \tag{17}$$

Thus, with the following definition

$$\left| u^{(d)} \right|_\infty := \sup_{\alpha_i \in \{2,3\}} \left\{ \left\| \frac{\partial^{\alpha_1 + \ldots + \alpha_d} u^{(d)}}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \right\|_\infty \right\}, \tag{18}$$

$$\left| u^{(d)} \right|_2 := \sup_{\alpha_i \in \{2,3\}} \left\{ \left\| \frac{\partial^{\alpha_1 + \ldots + \alpha_d} u^{(d)}}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \right\|_2 \right\},$$

we get

$$\left| v_m^{(q,d)} \right| \le \left| u^{(d)} \right|_\infty \cdot \frac{1}{2^d} \cdot \cdot h_1^2 \cdot \ldots \cdot h_k^2 \cdot h_{k+1}^3 \cdot \ldots \cdot h_d^3 \tag{19}$$

and

$$\left| v_m^{(q,d)} \right| \le \left| u^{(d)} \cdot \varphi_m^{(d)} \right|_2 \cdot \left(\frac{1}{6} \right)^{k/2} \cdot \left(\frac{17}{160} \right)^{(d-k)/2} \cdot \cdot h_1^{3/2} \cdot \ldots \cdot h_k^{3/2} \cdot h_{k+1}^{5/2} \cdot \ldots \cdot h_d^{5/2}. \tag{20}$$

Finally, we have to calculate the L_2- and maximum norm of the d-quadratic hierarchical basis function $\prod_{j=1}^d g_j(x_j)$,

$$g_j(x_j) = \frac{h_j^2 - x_j^2}{h_j^2}, \qquad -h_j \le x_j \le h_j, \tag{21}$$

which is now used instead of the piecewise d-linear $\prod_{j=1}^d w_j(x_j)$ defined in (8). Obviously, the maximum norm of $\prod_{j=1}^d g_j(x_j)$ is 1; and for the L_2-norm, we get

$$\left\| \prod_{j=1}^{d} g_j(x_j) \right\|_2 = \left(\frac{16}{15} \right)^{d/2} \cdot h_1^{1/2} \cdot \ldots \cdot h_d^{1/2}. \tag{22}$$

Now, we are ready to apply standard sparse grid approximation theory to the situation of piecewise d-quadratic

hierarchical basis functions. We are first interested in the difference between a sufficiently smooth function $u^{(d)}$ and its piecewise d-quadratic sparse grid interpolant $\tilde{u}_{n,I}^{(d)}$ of level n with a smallest occurring mesh width of 2^{-n}. Analogously to the linear case, (19), (20), and (22) lead to

$$\begin{aligned} \left\| u^{(d)} - \tilde{u}_{n,I}^{(d)} \right\|_\infty &\leq \frac{|u^{(d)}|_\infty}{14} \cdot \left(\frac{15}{112}\right)^{d-1} \cdot B_n^{(d)}, \\ \left\| u^{(d)} - \tilde{u}_{n,I}^{(d)} \right\|_2 &\leq \frac{|u^{(d)}|_2}{14} \cdot \left(\frac{15}{112}\right)^{d-1} \cdot B_n^{(d)}, \end{aligned} \tag{23}$$

where

$$B_n^{(d)} := \binom{n+d-1}{d-1} \cdot 8^{-n}. \tag{24}$$

Consequently, we get for the sparse grid interpolation error $u^{(d)} - \tilde{u}_{n,I}^{(d)}$

$$\begin{aligned} \left\| u^{(d)} - \tilde{u}_{n,I}^{(d)} \right\|_\infty &= O\left(N^{-3}(\log_2(N))^{d-1}\right), \\ \left\| u^{(d)} - \tilde{u}_{n,I}^{(d)} \right\|_2 &= O\left(N^{-3}(\log_2(N))^{d-1}\right), \end{aligned} \tag{25}$$

where $N = 2^n + 1$ denotes the maximum number of grid points in one direction. Thus, in comparison to the standard regular full grid, the accuracy of the interpolation is only slightly deteriorated by the logarithmic factor $(\log_2(N))^{d-1}$. Note that, according to the above argumentation and analogously to the piecewise linear case, $u^{(d)}$ has to fulfill the following smoothness requirement, e. g.:

$$\frac{\partial^{3d} u^{(d)}}{\partial x_1^3 \dots \partial x_d^3} \in C^0(\bar{\Omega}). \tag{26}$$

With respect to the energy norm, we again have to look at our d-quadratic hierarchical basis function $\prod_{j=1}^d g_j(x_j)$, first:

$$\begin{aligned} &\left\| \prod_{j=1}^d g_j(x_j) \right\|_E^2 \\ &= \int_{-h_1}^{+h_1} \dots \int_{-h_d}^{+h_d} \sum_{l=1}^d \left(\frac{\partial \prod_{j=1}^d g_j(x_j)}{\partial x_l} \right)^2 dx_d \dots dx_1 \\ &= \int_{-h_1}^{+h_1} \dots \int_{-h_d}^{+h_d} \sum_{l=1}^d \left(\prod_{j \neq l} \frac{(h_j^2 - x_j^2)^2}{h_j^4} \cdot \frac{4x_l^2}{h_l^4} \right) dx_d \dots dx_1 \\ &= \sum_{l=1}^d \left(\left(\frac{16}{15}\right)^{d-1} \cdot \left(\prod_{j \neq l} h_j \right) \cdot \frac{8}{3} \cdot \frac{1}{h_l} \right) \\ &= \frac{8}{3} \cdot \left(\frac{16}{15}\right)^{d-1} \cdot \sum_{l=1}^d \frac{h_1 \cdot \dots \cdot h_d}{h_l^2}. \end{aligned}$$

As above, this result concerning $\prod_{j=1}^d g_j(x_j)$ and (19) are the starting point for standard sparse grid analysis, which finally results in

$$\left\| u^{(d)} - \tilde{u}_{n,I}^{(d)} \right\|_E = O\left(4^{-n}\right) = O\left(N^{-2}\right), \tag{27}$$

the desired bound for the sparse grid interpolation error $u^{(d)} - \tilde{u}_{n,I}^{(d)}$ with regard to the energy norm. Thus, as in the linear case, the order of the energy error does not deteriorate when we switch from full grids to sparse. Since it is a well-known fact from finite element analysis that the finite element solution $\tilde{u}_n^{(d)}$ of a given boundary value problem is a best approximation to the solution $u^{(d)}$ on the underlying grid, we also get the following result concerning the error $u^{(d)} - \tilde{u}_n^{(d)}$ of the finite element approximation:

$$\left\| u^{(d)} - \tilde{u}_n^{(d)} \right\|_E = O\left(4^{-n}\right) = O\left(N^{-2}\right). \tag{28}$$

5 First numerical experiments

In this section, we report the results of some first numerical experiments concerning the piecewise quadratic hierarchical basis described above. For that, we study the Laplace equation on the unit square with Dirichlet boundary conditions as a simple model problem:

$$\begin{aligned} \Delta u(x,y) &= 0 \quad \text{on } \bar{\Omega} = [0,1]^2, \\ u(x,y) &= \sin(\pi y) \cdot \frac{\sinh(\pi(1-x))}{\sinh(\pi)}. \end{aligned}$$

Figure 12 shows the approximation to the solution calculated on the regular sparse grid of level 10 and its error.

For the solution of the linear system that results from the finite element discretization on the sparse grid, a simple Gauss-Seidel-iteration was used. The numerical results for this model problem are given in table 1. There, n denotes the level of the regular sparse grids considered (i. e., 2^{-n} is the smallest mesh width occurring). $\|e\|_\infty$ indicates the maximum norm of the sparse grid error $u^{(d)} - \tilde{u}_n^{(d)}$, and $\|e\|_E$ denotes its energy norm. Finally, ρ_∞ and ρ_E indicate the rates of reduction from level n to level $n+1$ of the respective error, and dof_n denotes the number of degrees of freedom, i. e. the number of grid points of the respective sparse grid. In table 1 and in figure 13, one can clearly see the $O(4^{-n}) = O(N^{-2})$-behaviour of the energy norm, and the convergence with respect to the maximum norm turns out to be just slightly worse than $O(8^{-n})$, as was to be expected due to the logarithmic factor in (25).

Furthermore, in figure 14, the results for the piecewise biquadratic case are compared to the piecewise bilinear

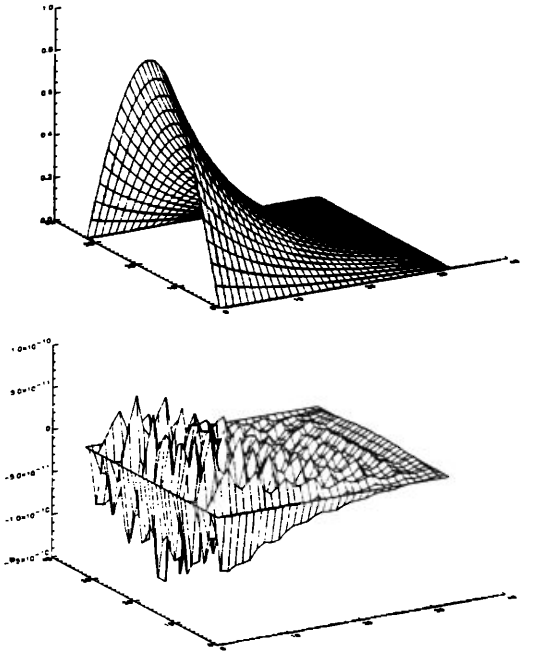

Figure 12: Sparse grid solution (above) and error (below) of the model problem.

n	$\|\|e\|\|_\infty$	ρ_∞	$\|\|e\|\|_E$	ρ_E	dof_n
1	$2.53\ 10^{-3}$		$2.43\ 10^{-1}$		1
		1.02		3.02	
2	$2.48\ 10^{-3}$		$8.05\ 10^{-2}$		5
		2.67		3.41	
3	$9.28\ 10^{-4}$		$2.36\ 10^{-2}$		17
		3.14		3.77	
4	$2.96\ 10^{-4}$		$6.26\ 10^{-3}$		49
		4.09		3.91	
5	$7.24\ 10^{-5}$		$1.60\ 10^{-3}$		129
		5.36		3.98	
6	$1.35\ 10^{-5}$		$4.02\ 10^{-4}$		321
		6.19		3.98	
7	$2.18\ 10^{-6}$		$1.01\ 10^{-4}$		769
		6.83		4.02	
8	$3.19\ 10^{-7}$		$2.51\ 10^{-5}$		1793
		7.28		3.99	
9	$4.38\ 10^{-8}$		$6.29\ 10^{-6}$		4097
		7.60		4.01	
10	$5.76\ 10^{-9}$		$1.57\ 10^{-6}$		9217
		7.76		3.99	
11	$7.42\ 10^{-10}$		$3.93\ 10^{-7}$		20481
		7.89		4.00	
12	$9.41\ 10^{-11}$		$9.82\ 10^{-8}$		45057
		7.94		3.99	
13	$1.19\ 10^{-11}$		$2.46\ 10^{-8}$		98305
		7.97		4.01	
14	$1.49\ 10^{-12}$		$6.14\ 10^{-9}$		212993

Table 1: Error on the regular sparse grid of level n.

situation. Here, both times, adaptive sparse grids were used. Again, the advantages of the quadratic approach can be seen clearly.

Now, let us turn to a second example:

$$\begin{aligned} \Delta u(x,y) &= 0 \quad \text{on } \bar{\Omega} = [0,1]^2, \\ u(x,y) &= \cos(4\pi(x-y)) \cdot \frac{\sinh(4\pi(x+y))}{\sinh(8\pi)}. \end{aligned}$$

Here, again, u fulfills the smoothness requirement (26). However, in contrast to our first example, adaptive grid refinement is very helpful (see figure 15).

Figure 16 illustrates the results for piecewise bilinear and piecewise biquadratic basis functions. In both cases, adaptive grid refinement has been used. In comparison with figure 14, the gain that can be achieved with the biquadratic approach is smaller. This had to be expected, since adaptive mesh refinement is very efficient in such situations, and it indicates that a combined process of adaptive grid refinement and adaptive choice of the polynomial degree p of the basis functions might be the appropriate strategy for sparse grids, too.

6 Concluding remarks

In this paper, some first steps towards an efficient implementation of higher order techniques on sparse grids have been discussed. The approach of section 4 leads to hierarchical bases of polynomials of higher degree $p > 1$, but still results in C^0-(sparse grid)-interpolants. However, the number of degrees of freedom per grid point does not increase with growing p. Obviously, the concepts presented for the quadratic case can be generalized to the situation with cubic polynomials, and so on, which will be in the centre of future work. Finally, h-p-version-type algorithms [1, 12, 13] are to be developed for sparse grids, too.

The following tables 2 and 3 show why higher order techniques on sparse grids seem to be a very promising approach to the efficient numerical treatment of partial differential equations. Each row in both tables corresponds to a fixed number d of dimensions of the underlying problem, and each column stands for a certain polynomial degree p of the basis functions used. If M denotes the overall number of unknowns (i. e., $M = N^d$ for a regular full grid and $M = O(N(\log_2(N))^{d-1})$ or $M = O(N)$, respectively, for regular sparse grids), then, we can indicate the order of approximation with respect to the energy norm by $M^{-\alpha}$.

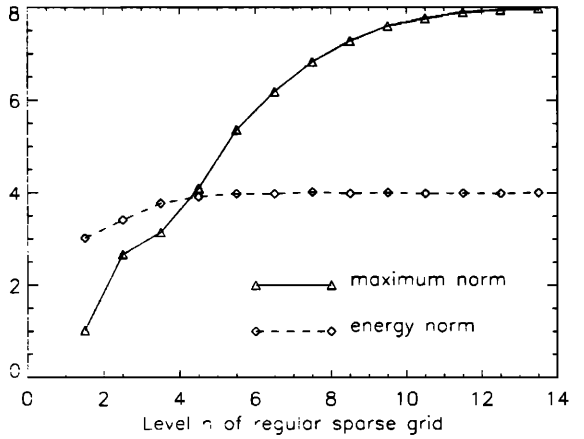

Figure 13: Rates ρ_∞ and ρ_E of error reduction.

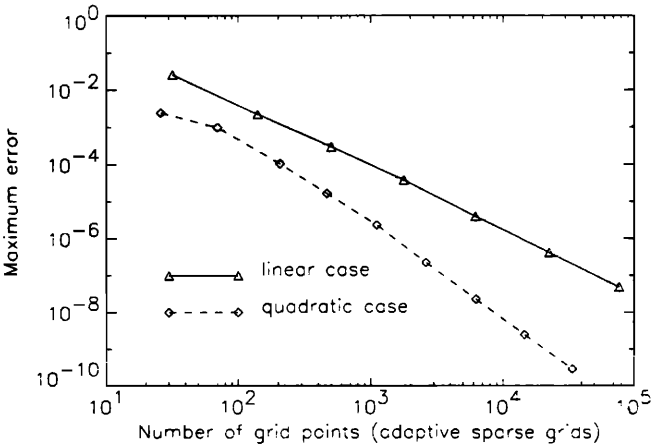

Figure 14: Maximum error vs. number of grid points (linear and quadratic case).

The entries in both tables now show the respective values of α. For example, if we want to achieve second order with respect to the number of unknowns on full grids, i.e. $\alpha = 2$, we have to use quadratic polynomials in the one-dimensional case, quartic ones for $d = 2$; and, for three-dimensional problems, even polynomials of degree $p = 6$ have to be used. With sparse grids, in contrast to that, p does not depend on d. For $\alpha = 2$, quadratic polynomials are sufficient for arbitrary d.

$d\backslash p$	1	2	3	4	5	6
1	1	**2**	3	4	5	6
2	1/2	1	3/2	**2**	5/2	3
3	1/3	2/3	1	4/3	5/3	**2**

Table 2: Approximation order $M^{-\alpha}$ for various d and p on full grids.

At this point, we have to go into the smoothness requirements of sparse grid techniques. For the quadratic case, they are given in (26). At first glance, these seem to be quite restrictive, especially for larger p. However, as in the linear case, the inherent h-adaptivity of sparse grid techniques should be able to deal with non-smooth situations, too. Furthermore, we can learn from tables 2 and 3 that, with respect to the overall number of unknowns, sparse grids can manage with smaller values of p than full grids. Therefore, especially for achieving high approximation quality for three-dimensional problems, sparse grids even turn out to be advantageous regarding smoothness requirements.

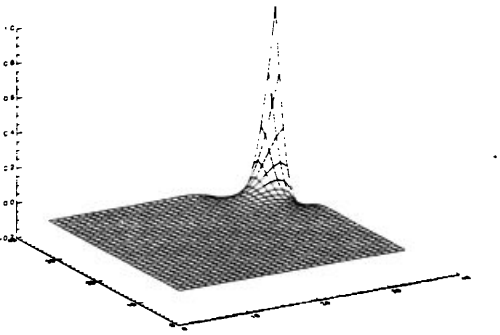

Figure 15: Sparse grid solution of the second example.

7 Acknowledgments

This work is supported by the Bayerische Forschungsstiftung via FORTWIHR — The Bavarian Consortium for High Performance Scientific Computing. I am indebted to Prof. Christoph Zenger for many fruitful discussions and suggestions.

References

[1] I. Babuška and M. Suri. The p- and h-p-versions of

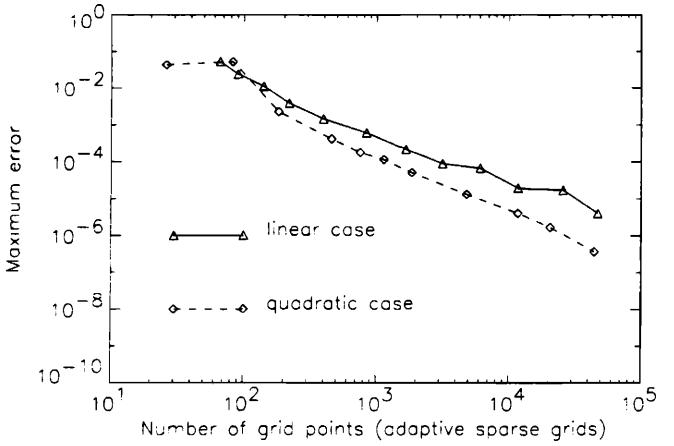

Figure 16: Maximum error vs. number of grid points (linear and quadratic case).

$d\backslash p$	1	2	3	4	5	6
1	1	**2**	3	4	5	6
2	1	**2**	3	4	5	6
3	1	**2**	3	4	5	6

Table 3: Approximation order $M^{-\alpha}$ for various d and p on sparse grids.

the finite element method: An overview. *Comput. Methods Appl. Mech. Engrg.*, 80:5–26, 1990.

[2] I. Babuška, B. A. Szabó, and I. N. Katz. The p-version of the finite element method. *SIAM J. Numer. Anal.*, 18(3):515–545, 1981.

[3] R. Balder, U. Rüde, S. Schneider, and C. Zenger. Sparse grid and extrapolation methods for parabolic problems. In A. Peters, G. Wittum, B. Herrling, and U. Meissner, editors, *Proceedings of the 10th International Conference on Computational Methods in Water Resources, Heidelberg, July 1994*. Kluwer academic publishers, 1994.

[4] R. E. Bank, T. Dupont, and H. Yserentant. The hierarchical basis multigrid method. *Numerische Mathematik*, 52:427–458, 1988.

[5] T. Bonk. A new algorithm for multi-dimensional adaptive numerical quadrature. In W. Hackbusch, editor, *Proceedings of the 9th GAMM-Seminar, Kiel, January 1993*. Vieweg, Braunschweig, 1994.

[6] H.-J. Bungartz. An adaptive Poisson solver using hierarchical bases and sparse grids. In P. de Groen and R. Beauwens, editors, *Iterative Methods in Linear Algebra: Proceedings of the IMACS International Symposium, Brussels, 2.-4. 4. 1991*, pages 293–310. Elsevier, Amsterdam, 1992.

[7] H.-J. Bungartz. *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung.* Dissertation, Institut für Informatik, TU München, 1992.

[8] H.-J. Bungartz. *Higher Order Finite Elements on Sparse Grids.* SFB Report 342/01/95 A, Institut für Informatik, TU München, 1995.

[9] H.-J. Bungartz, M. Griebel, and U. Rüde. Extrapolation, combination, and sparse grid techniques for elliptic boundary value problems. *Comput. Methods Appl. Mech. Engrg.*, 116:243–252, 1994.

[10] M. Griebel. Parallel multigrid methods on sparse grids. In W. Hackbusch and U. Trottenberg, editors, *Multigrid Methods III: Proceedings of the 3rd European Conference on Multigrid Methods. Bonn, October 1990*, pages 211–221. Int. Ser. Num. Math. 98, Birkhäuser, Basel, 1991.

[11] M. Griebel. A parallelizable and vectorizable multilevel algorithm on sparse grids. In W. Hackbusch, editor, *Parallel Algorithms for Partial Differential Equations: Proceedings of the 6th GAMM-Seminar, Kiel, January 1990, Notes on Numerical Fluid Mechanics 31*, pages 94–100. Vieweg, Braunschweig, 1991.

[12] B. Guo and I. Babuška. The h-p-version of the finite element method (Part 1: The basic approximation results). *Computational Mechanics*, 1:21–41, 1986.

[13] B. Guo and I. Babuška. The h-p-version of the finite element method (Part 2: General results and applications). *Computational Mechanics*, 1:203–220, 1986.

[14] K. Hallatschek. Fouriertransformation auf dünnen Gittern mit hierarchischen Basen. *Numer. Math.*, 63(1):83–97, 1992.

[15] C. Pflaum. *A multi-level-algorithm for the finite-element-solution of general second order elliptic differential equations on adaptive sparse grids.* SFB Report 342/12/94 A, Institut für Informatik, TU München, 1994.

[16] T. Störtkuhl. *Ein numerisches, adaptives Verfahren zur Lösung der biharmonischen Gleichung auf dünnen Gittern.* Dissertation, Institut für Informatik, TU München, 1994.

[17] H. Yserentant. On the multilevel splitting of finite element spaces. *Numer. Math.*, 49:379–412, 1986.

[18] C. Zenger. Sparse grids. In W. Hackbusch, editor, *Parallel Algorithms for Partial Differential Equations: Proceedings of the 6th GAMM-Seminar, Kiel, January 1990, Notes on Numerical Fluid Mechanics 31.* Vieweg, Braunschweig, 1991.