Sparse Grids

The sparse grid method is a general **numerical discretization** technique for multivariate problems. This approach, first introduced by the Russian mathematician Smolyak in 1963 [27], constructs a multidimensional multilevel basis by a special truncation of the tensor product expansion of a one-dimensional multilevel basis (see Figure 1 for an example of a sparse grid).

Discretizations on sparse grids involve only $O(N(\log N)^{d-1})$ degrees of freedom, where d is the problem dimension and N denotes the number of degrees of freedom in one coordinate direction. The accuracy obtained this way is comparable to one using a full tensor product basis involving $O(N^d)$ degrees of freedom, if the underlying problem is smooth enough, i.e., if the solution has bounded mixed derivatives.

This way, the **curse of dimension**, i.e., the exponential dependence of conventional approaches on the dimension d, can be overcome to a certain extent. This makes the sparse grid approach particularly attractive for the numerical solution of moderate- and higher-dimensional problems. Still, the classical sparse grid method is not completely independent of the dimension due the above logarithmic term in the complexity.

Sparse grid methods are known under various names, such as hyperbolic cross points, discrete blending, Boolean interpolation or splitting extrapolation. For a comprehensive introduction to sparse grids, see [5].

In computational finance, sparse grid methods have been employed for the valuation of **multi-asset options** such as basket [25] or outperformance options [13], various types of **path-dependent derivatives** and **likelihood estimation** [20] due to the high dimension of the arising partial differential equations or integration problems.

One-Dimensional Multilevel Basis

The first ingredient of a sparse grid method is a one-dimensional multilevel basis. In the classical sparse grid approach, a hierarchical basis based on standard hat functions,

$$\phi(x) \;:=\; \begin{cases} 1-|x| & \text{if } x \in [-1,1]\,, \\ 0 & \text{otherwise} \end{cases}\;, \qquad (1)$$

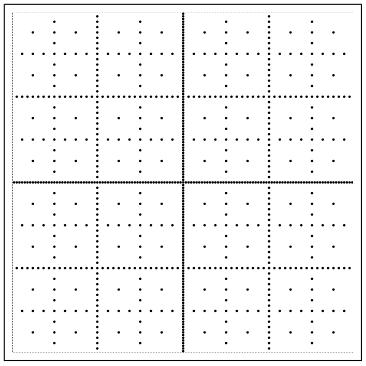

Figure 1: A regular two-dimensional sparse grid of level 7.

is used. Then, a set of equidistant grids Ω_l of level l on the unit interval $\bar{\Omega} = [0,1]$ and mesh width 2^{-l} is considered. The grid points $x_{l,i}$ are given by

$$x_{l,i} := i \cdot h_l,\; 0 \le i \le 2^l\,. \qquad (2)$$

The standard hat function is then taken to generate a family of basis functions $\phi_{l,i}(x)$ having support $[x_{l,i} - h_l, x_{l,i} + h_l]$ by dilation and translation, i.e.,

$$\phi_{l,i}(x) \;:=\; \phi\left(\frac{x - i\cdot h_l}{h_l}\right). \qquad (3)$$

Thereby, the index i indicates the location of a basis function or a grid point. This basis is usually termed *nodal basis* or *Lagrange basis* (see Figure 2, bottom). These basis functions are then used to define function spaces V_l consisting of piecewise linear functions[1]

$$V_l \;:=\; \text{span}\left\{\phi_{l,i} :\; 1 \le i \le 2^l - 1\right\}. \qquad (4)$$

With these function spaces, the hierarchical increment spaces W_l,

$$W_l := \text{span}\,\{\phi_{l,i} : i \in I_l\} \qquad (5)$$

using the index set

$$I_l = \{i \in \mathbb{N} : 1 \le i \le 2^l - 1, i \text{ odd}\} \qquad (6)$$

are defined. These increment spaces satisfy the relation

$$V_l \;=\; \bigoplus_{k \le l} W_k\,. \qquad (7)$$

The basis corresponding to W_l is called *hierarchical*

[1] In order to simplify this exposition, we assume that the functions in V_l are zero on the boundary of $\bar{\Omega}$. This restriction can be overcome by adding appropriate boundary basis functions.

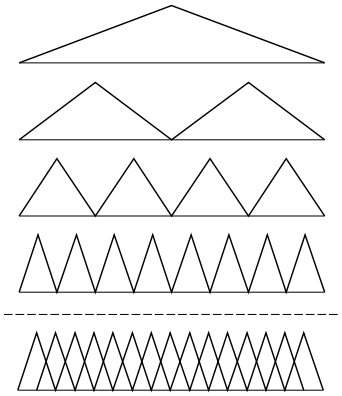

Figure 2: Piecewise linear hierarchical basis (top) vs. nodal basis (bottom) of level 4.

basis (see Figure 2, top) and any function $u \in V_l$ can be uniquely represented as

$$u(x) = \sum_{k=1}^{l} \sum_{i \in I_k} v_{k,i} \cdot \phi_{k,i}(x) \; . \tag{8}$$

with coefficient values $v_{k,i} \in \mathbb{R}$. Note that the supports of all basis functions $\phi_{k,i}$ spanning W_k are mutually disjoint.

Tensor Product Construction

From this one-dimensional hierarchical basis, a multi-dimensional basis on the d-dimensional unit cube $\bar{\Omega} := [0,1]^d$ is obtained by a tensor product construction. With the multi-index $\mathbf{l} = (l_1, \ldots, l_d) \in \mathbb{N}^d$, which indicates the level in a multivariate sense, the set of d-dimensional standard rectangular grids $\Omega_\mathbf{l}$ on $\bar{\Omega}$ with mesh size $\mathbf{h_l} := (h_{l_1}, \ldots, h_{l_d}) := 2^{-\mathbf{l}}$ are considered. Each grid $\Omega_\mathbf{l}$ is equidistant with respect to each individual coordinate direction, but, in general, may have varying mesh sizes in the different directions. The grid points $\mathbf{x_{l,i}}$ of the grid $\Omega_\mathbf{l}$ are the points

$$\mathbf{x_{l,i}} := (x_{l_1,i_1}, \ldots, x_{l_d,i_d}), \; \mathbf{1} \leq \mathbf{i} \leq 2^\mathbf{l} - \mathbf{1} \, , \tag{9}$$

where for the above multi-indices, all arithmetic operations are to be understood component-wise.

Then, for each grid point $\mathbf{x_{l,i}}$, an associated piecewise d-linear basis function $\phi_\mathbf{l,i}(\mathbf{x})$ (see Figure 3) is defined as the product of the one-dimensional basis

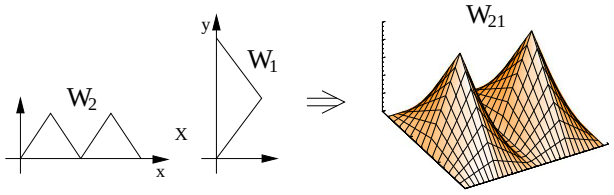

Figure 3: Tensor product approach to generate the piecewise bilinear basis functions $\phi_{(2,1),(1,1)}$ and $\phi_{(2,1),(1,1)}$ from the one-dimensional basis functions $\phi_{2,1}, \phi_{2,2}$ and $\phi_{1,1}$.

functions

$$\phi_\mathbf{l,i}(\mathbf{x}) \; := \; \prod_{j=1}^{d} \phi_{l_j,i_j}(x_j) \, . \tag{10}$$

Each of the multidimensional (nodal) basis functions $\phi_\mathbf{l,i}$ has a support of size $2 \cdot \mathbf{h_l}$. These basis functions are again used to define function spaces $V_\mathbf{l}$ consisting of piecewise d-linear functions which are zero on the boundary of $\bar{\Omega}$,

$$V_\mathbf{l} \; := \; \text{span} \left\{ \phi_\mathbf{l,i} : \; \mathbf{1} \leq \mathbf{i} \leq 2^\mathbf{l} - \mathbf{1} \right\} \, . \tag{11}$$

Similar to the one-dimensional case, the hierarchical increments $W_\mathbf{l}$ are defined by

$$W_\mathbf{l} \; := \; \text{span} \left\{ \phi_\mathbf{l,i} : \mathbf{i} \in \mathbf{I_l} \right\} \tag{12}$$

with the index set

$$\begin{aligned} \mathbf{I_l} \; := \; & \Big\{ \mathbf{i} \in \mathbb{N}^d : \mathbf{1} \leq \mathbf{i} \leq 2^\mathbf{l} - \mathbf{1}, \\ & \; i_j \text{ odd for all } 1 \leq j \leq d \; \Big\} \, . \end{aligned} \tag{13}$$

This way, the hierarchical increment spaces $W_\mathbf{l}$ are related to the nodal spaces $V_\mathbf{l}$ by

$$V_\mathbf{l} \; = \; \bigoplus_{\mathbf{k} \leq \mathbf{l}} W_\mathbf{k} \, . \tag{14}$$

Again, the supports of all multidimensional hierarchical basis functions $\phi_\mathbf{l,i}$ spanning $W_\mathbf{l}$ are mutually disjoint. Also, again each function $u \in V_\mathbf{l}$ can uniquely be represented by

$$u_\mathbf{l}(\mathbf{x}) = \sum_{\mathbf{k}=\mathbf{1}}^{l} \sum_{\mathbf{i} \in \mathbf{I_k}} v_\mathbf{k,i} \cdot \phi_\mathbf{k,i}(\mathbf{x}) \tag{15}$$

with hierarchical coefficients $v_\mathbf{k,i} \in \mathbb{R}$.

Classical Sparse Grids

The classical sparse grid construction arises from a cost to benefit analysis in function approximation. Thereby, functions $u : \Omega \to \mathbb{R}$ which have bounded mixed second derivatives

$$D^{\boldsymbol{\alpha}} u \;:=\; \frac{\partial^{|\boldsymbol{\alpha}|_1} u}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \tag{16}$$

for $|\alpha|_\infty \leq 2$ are considered. These functions belong to the Sobolev space $H_2^{\text{mix}}(\bar{\Omega})$ with

$$\begin{aligned} H_2^{\text{mix}}(\bar{\Omega}) \;&:=\; \{u : \bar{\Omega} \to \mathbb{R} : D^\alpha u \in L_2(\Omega), \\ &\qquad |\alpha|_\infty \leq 2, u|_{\partial\Omega} = 0\} \,. \end{aligned} \tag{17}$$

Here, the two norms $|\alpha|_1$ and $|\alpha|_\infty$ for multi-indices are defined by

$$|\alpha|_1 := \sum_{j=1}^{d} \alpha_j \quad \text{and} \quad |\alpha|_\infty := \max_{1 \leq j \leq d} \alpha_j \,. \tag{18}$$

For functions $u \in H_2^{\text{mix}}(\bar{\Omega})$, the hierarchical coefficients $v_{\mathbf{l},\mathbf{i}}$ decay as

$$|v_{\mathbf{l},\mathbf{i}}| = O(2^{-2|\mathbf{l}|_1}) \,. \tag{19}$$

On the other hand, the size (i.e., the number of degrees of freedom) of the subspaces $W_\mathbf{l}$ is given by

$$|W_\mathbf{l}| = O(2^{|\mathbf{l}|_1}) \,. \tag{20}$$

An optimization with respect to the number of degrees of freedom and the resulting approximation accuracy directly leads to *sparse grid* spaces $\hat{V}_n$ of level n defined by

$$\hat{V}_n := \bigoplus_{|\mathbf{l}|_1 \leq n+d-1} W_\mathbf{l} \,. \tag{21}$$

In comparison to the standard *full grid* space

$$V_n := V_{(n,\dots,n)} = \bigoplus_{|\mathbf{l}|_\infty \leq n} W_\mathbf{l} \,, \tag{22}$$

which corresponds to cubic sectors of subspaces, sparse grids use triangular or simplicial sectors, see Figure 4. The dimension of the space $\hat{V}_n$, i.e., the number of degrees of freedom or grid points, is given by

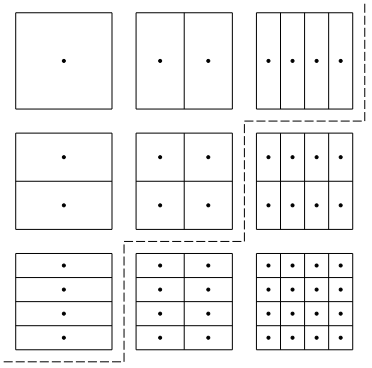

Figure 4: All subspaces $W_\mathbf{l}$ for levels $|\mathbf{l}|_\infty \leq 3$ which together form the full grid space V_3. The corresponding sparse grid space $\hat{V}_3$ consists of all subspaces above the dashed line ($|\mathbf{l}|_1 \leq 4$).

$$\begin{aligned} \left|\hat{V}_n\right| \;&=\; \sum_{i=0}^{n-1} 2^i \cdot \binom{d-1+i}{d-1} \\ &=\; O(h_n^{-1} \cdot |\log_2 h_n|^{d-1}) \,. \end{aligned} \tag{23}$$

This shows the order $O(2^n n^{d-1})$, which is a significant reduction of the number of degrees of freedom and, thus, of the computational and storage requirement compared to the order $O(2^{nd})$ of the dimension of the full grid space $|V_n|$.

On the other hand, the approximation accuracy of the sparse grid spaces for functions $u \in H_2^{\text{mix}}(\bar{\Omega})$ is in the L_p-norms given by

$$||u - \hat{u}_n||_p = O(h_n^2 \cdot n^{d-1}) \,. \tag{24}$$

For the corresponding full grid spaces, the accuracy is

$$||u - u_n||_p = O(h_n^2) \,. \tag{25}$$

This shows the crucial advantage of the sparse grid space $\hat{V}_n$ in comparison with the full grid space V_n: the number of degrees of freedom is significantly reduced, whereas the accuracy is only slightly deteriorated. This way, the curse of dimensionality can be overcome, at least to some extent. The dimension still enters through logarithmic terms both in the computational cost and the accuracy estimate as well as in the constants hidden in the order notation.

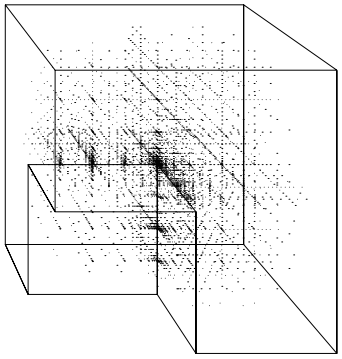

Figure 5: An at a corner singularity adaptively refined three-dimensional sparse grid.

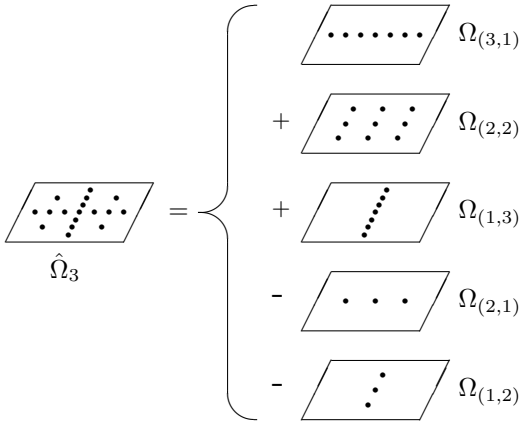

Figure 6: The combination technique in two dimensions for level $n = 3$: combine coarse full grids $\Omega_{\mathbf{l}}$, $|\mathbf{l}|_1 \in \{3, 4\}$, with mesh widths 2^{-l_1} and 2^{-l_2} to get a sparse grid $\hat{\Omega}_n$ corresponding to $\hat{V}_n$.

Extensions and Applications

The classical sparse grid concept has been generalized in various ways. First, there are special sparse grids which are optimized with respect to the energy seminorm [4]. These energy-based sparse grids are further sparsified and possess a cost complexity of $O(h_n^{-1})$ for an accuracy of $O(h_n)$. Thus, the dependence on the dimension d in the order is completely removed (but is still present in the hidden constants [9]). A generalization to sparse grids which are optimal with respect to other Sobolev norms can be found in [14]. In case the underlying space is not known a priori, dimension-adaptive methods [12] can be applied to find optimized sparse grids.

The sparse grid approach based on piecewise linear interpolation can be extended to non-smooth problems by locally adaptive methods [2], see Figure 5 for an adaptively refined sparse grid. Furthermore, it can be generalized to higher order polynomial [5] or wavelet discretizations (e.g., interpolets or prewavelets) [16, 26], which allows to utilize additional properties (such as higher polynomial exactness or vanishing moments) of the basis.

Sparse grids have been applied for the solution of different kinds of low- and moderate-dimensional partial differential equations, such as elliptic [5, 28], parabolic [1, 15] and hyperbolic [18] problems. In this context, finite element methods [2], finite difference methods [8] and finite volume methods [21] have been used in the discretization process.

For the solution of partial differential equations, often the so-called *combination technique* [10] is employed. Here, a sparse grid solution is obtained by a combination of anisotropic full grid solutions according to the combination formula

$$\hat{u}_n(\mathbf{x}) = \sum_{n \le |\mathbf{l}|_1 \le n+d-1} (-1)^{n+d-|\mathbf{l}|_1-1} \binom{d-1}{|\mathbf{l}|_1 - n} u_{\mathbf{l}}(\mathbf{x}) \quad (26)$$

where $u_{\mathbf{l}}(\mathbf{x})$ is a full grid solution on an anisotropic grid with mesh width $2^{-\mathbf{l}}$, see Figure 6 for a two-dimensional example. The combination technique can be further optimized with respect to the underlying differential operator [19].

The sparse grid approach can also be used for numerical integration, e.g., for the computation of expectations [11, 24]. Thereby, the classical sparse grid construction starts with a sequence of one-dimensional quadrature formulas $Q_l f$ using n_l points to integrate a function f on the unit interval $[0, 1]$,

$$Q_l f := \sum_{i=1}^{n_l} w_{li} \cdot f(x_{li}) \,. \quad (27)$$

Using the difference quadrature formulas

$$\Delta_k f := (Q_k - Q_{k-1}) f \;\text{ with }\; Q_0 f := 0 \,, \quad (28)$$

the sparse grid quadrature formula $\hat{Q}_n f$ of level n for a d–dimensional function f on the cube $[0, 1]^d$ is then defined by

$$\hat{Q}_n f := \sum_{|\mathbf{l}|_1 \le n+d-1} (\Delta_{l_1} \otimes \ldots \otimes \Delta_{l_d}) f \,. \quad (29)$$

Again, this construction can be improved by using spatially adaptive or dimension-adaptive refinement [3, 12].

The sparse grid methodology has also been successfully applied to the solution of integral equations [17], interpolation and approximation [22] and data analysis [7, 23].

References

[1] Balder, R. & Zenger, C. (1996) The solution of the multidimensional real Helmholtz equation on sparse grids, *SIAM J. Sci. Comp.* **17**, 631-646.

[2] Bungartz, H. (1992) An adaptive Poisson solver using hierarchical bases and sparse grids, in *Iterative Methods in Linear Algebra*, R. Beauwens (ed.), North-Holland, 293-310.

[3] Bungartz, H. & Dirnstorfer, S. (2003) Multivariate quadrature on adaptive sparse grids, *Computing* **71**, 89-114.

[4] Bungartz, H. & Griebel, M. (1999) A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives, *J. Complexity* **15**, 1-121.

[5] Bungartz, H. & Griebel, M. (2004) Sparse grids, *Acta Numerica* **13**, 147-269.

[6] Delvos, F. & Schempp, W. (1990) *Boolean methods in interpolation and approximation*, Longman.

[7] Garcke, J., Griebel, M. & Thess, M. (2001) Data mining with sparse grids, *Computing* **67**, 225-253.

[8] Griebel, M. (1998) Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences, *Computing*, **61**, 151-179.

[9] Griebel, M. (2006) Sparse grids and related approximation schemes for higher dimensional problems, in Proc. FoCM05, L. Pardo, A. Pinkus, E. Suli & M. Todd (eds.), Cambridge Univ. Press.

[10] Griebel, M., Schneider, M. & Zenger, C. (1992) A combination technique for the solution of sparse grid problems, in *Iterative Methods in Linear Algebra*, P. de Groen & R. Beauwens (eds.), Elsevier, pp. 263-281.

[11] Gerstner, T. & Griebel, M. (1998) Numerical integration using sparse grids, *Numer. Alg.* **18**, 209-232.

[12] Gerstner, T. & Griebel, M. (2003) Dimension-adaptive tensor-product quadrature, *Computing* **71**, 65-87.

[13] Gerstner, T. & Holtz, M. (2008) Valuation of performance-dependent options, *Appl. Math. Fin.* **15**, 1-20.

[14] Griebel, M. & Knapek, S. (2000) Optimized tensor-product approximation spaces, *Const. Approx.* **16**, 525-540.

[15] Griebel, M. & Oeltz, D. (2007) A sparse grid space-time discretization scheme for parabolic problems, *Computing* **81**, 1-34.

[16] Griebel, M. & Oswald, P. (1995) Tensor product type subspace splitting and multilevel iterative methods for anisotropic problems, *Adv. Comp. Math.* **4**, 171-206.

[17] Griebel, M., Oswald, P. & Schiekofer, T. (1999) Sparse grids for boundary integral equations, *Numer. Math.*, **83**, 279-312.

[18] Griebel, M. & Zumbusch, G. (1999) Adaptive sparse grids for hyperbolic conservation laws, in *Hyperbolic Problems: Theory, Numerics, Applications*, M. Fey & R. Jeltsch (eds.), Birkhäuser, pp. 411-422.

[19] Hegland, M., Garcke, J. & Challis, V. (2007) The combination technique and some generalisations, *Lin. Alg. Appl.* **420**, 249-275.

[20] Heiss, F. & Winschel, V. (2008) Likelihood approximation by numerical integration on sparse grids, *J. Econometrics*, to appear.

[21] Hemker, P. (1995) Sparse-grid finite-volume multigrid for 3D-problems, *Adv. Comp. Math.*, **4**, 83-110.

[22] Klimke, A. & Wohlmuth, B. (2005) Algorithm 847: Spinterp: piecewise multilinear hierarchical sparse grid interpolation in MATLAB, *ACM Trans. Math. Soft.* **31**, 561-579.

[23] Laffan, S., Nielsen, O., Silcock, H. & Hegland, M. (2005) Sparse grids: a new predictive modelling method for the analysis of geographical data, *Int. J. Geo. Inform. Sci.* **19**, 267-292.

[24] Novak, E. & Ritter, K. (1996) High dimensional integration of smooth functions over cubes, *Numer. Math.* **75**, 79-97.

[25] Reisinger, C. & Wittum, G. (2007) Efficient hierarchical approximation of high-dimensional option pricing problems, *SIAM J. Sci. Comp.* **29**, 440-458.

[26] Schwab, C. & Todor, R. (2003) Sparse finite elements for stochastic elliptic problems: higher order moments, *Computing* **71**, 43-63.

[27] Smolyak, S. (1963) Interpolation and quadrature formulas for the classes W_s^a and E_s^a, *Soviet Math. Dokl.* **4**, 240-243.

[28] Zenger, C. (1991) Sparse grids, in *Parallel Algorithms for Partial Differential Equations*, W. Hackbusch (ed.), Vieweg, pp. 241-251.

(*see also* **Derivative Pricing, Numerical Methods**)

THOMAS GERSTNER, MICHAEL GRIEBEL