Parallelized Dimensional Decomposition for Dynamic Stochastic Economic Models

Università della Svizzera italiana | Faculty of Informatics | Institute of Computational Science ICS

Aryan Eftekhari[a], Simon Scheidegger[b,c], Olaf Schenk[a]

[a] Institute of Computational Science, Faculty of Informatics, Universitá della Svizzera italiana,
[b] Department of Banking and Finance, University of Zürich, [c] Hoover Institution, Stanford University

Dynamic Stochastic Economic Models

Dynamic Stochastic Models

These models can used to analyze, for example, business cycles and effects of policies. The specific model being looked at is the International Real Business Cycle (IRBC), which describes the dynamics of N countries characterized by a $d = 2N$ dimensional state-space $X \subset [0,1]^d$, comprising of productivity and capital stock.

Stochastic Transitions & Equilibrium Conditions: For $x_t \in X$, the transition $t \to t+1$ can then be represented by the distribution of next period's state x_{t+1}, conditional on the current state x_t and policy p. In addition, the period-to-period equilibrium conditions $E(\cdot)$ must be satisfied by the policy p.

[E1] $x_{t+1} \sim \mathcal{D}(\cdot|x_t, p(x_t))$

$$\mathbb{E}\left[E(x_t, x_{t+1}, p(x_t), p(x_{t+1}))|x_t, p(x_t)\right] = 0, \quad \forall t$$

Time Iteration Algorithm: We use the term "time iteration" for an algorithm that iteratively updates the policy function by solving the period-to-period first-order equilibrium conditions [E1].

Motivation

Interpolation is required on a $2N$-dimensional state-space. With $N \gg 2$, we encounter the so called *Curse-of-Dimensionality*.

■ **Full Grid:** Discretization into m $1D$ segments (Cartesian Grid) leads to *exponential increase* of computational resources, $\mathcal{O}(\mathbf{m^d})$.

■ **Sparse Girds:** Sufficient reduction in computational resources for medium sized problems $\mathcal{O}(\mathbf{m}\log(\mathbf{m})^{\mathbf{d-1}})$, but in higher dimensions attainting required *grid resolution leads to massive increases* in grid-points.

■ **Dimensional Decomposition:** Provides an avenue to both *reduce total grid points* and achieve *gradual increase in grid points with refinement*. For a first-order decomposition, grid points increase linearly with dimension, $\mathcal{O}(\mathbf{md})$.

Proposition: Utilize dimensional decomposition in conjunction with adaptive sparse grids (referred to as **DDSG**) to generate a highly parrallelized interpolation routine for the iterative solution method.

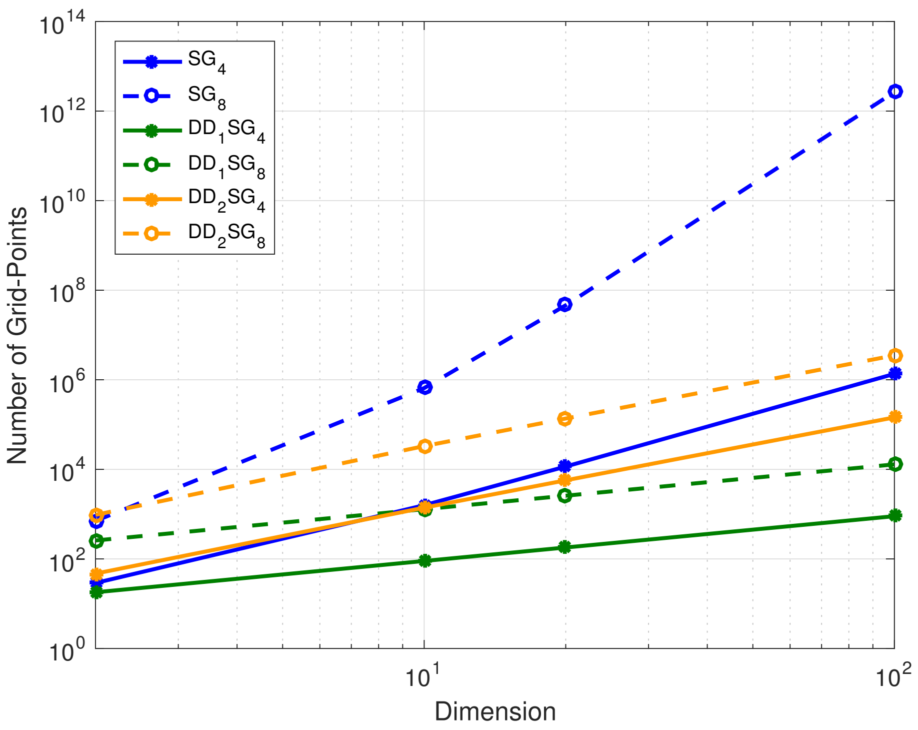

Fig 1: Number of interpolation grid-points for SG and DDSG order 1 and 2 at refinement level 4 and 8. DDSG exhibits both a reduction of grid points but also a gradual increase in point with changing refinement levels.

Dimensional Decomposition

General Dimensional Decomposition: The solution to the minimization problem of [E3] is the unique and optimal component function. Specifying the weight function w dictates the properties of the method.

[E3]
$$\underset{f_{\mathbf{u}^*}(\mathbf{x}_{\mathbf{u}^*})}{\text{minimize}} \ \frac{1}{2}\left\|\sum_{\mathbf{u}\subseteq\mathcal{S}} f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) - f(\mathbf{x})\right\|_w^2$$
$$\text{subject to} \int f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}})w(\mathbf{x}_{\mathbf{u}})dx_i = 0$$
$$\forall i \in \mathbf{u} \subseteq \mathcal{S}, i \neq \emptyset$$

Cut-HDMR: A variate of the general formulation of [E3], but using the Dirac Delta (product measure) as its weight function.

[E4]
$$w(\mathbf{x})d\mathbf{x} = \prod_{i=1}^{d} \delta(x_i - \bar{x}_i)dx_i,$$
$$f(\mathbf{x}) = \sum_{\mathbf{u}\subseteq\mathcal{S}} f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}})$$
$$f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}}) = f(\mathbf{x})|_{\mathbf{x}=\bar{\mathbf{x}}\setminus\mathbf{x}_{\mathbf{u}}} - \sum_{\mathbf{v}\subset\mathcal{S}} f_{\mathbf{v}}(\mathbf{x}_{\mathbf{v}})$$

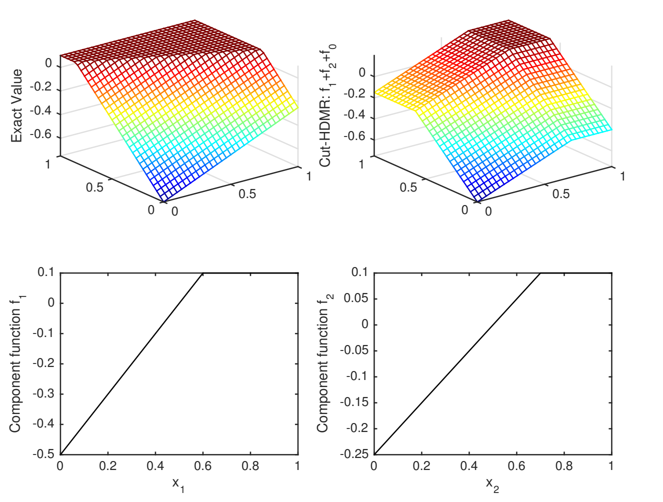

Fig 2: 2D function decomposed into a set of 1D functions. No smoothing behavior, as the component functions only retain point-wise information.

Significance Measure: Early truncation of equation [E4] is required for efficient approximation.

[E5]
$$\eta_{\mathbf{u}} = \frac{\left\|\int f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}})d\mathbf{x}\right\|_2}{\left\|\sum_{\mathbf{v}\subset\mathcal{S},|\mathbf{v}|\leq|\mathbf{u}|-1}\int f_{\mathbf{u}}(\mathbf{x}_{\mathbf{u}})d\mathbf{x}\right\|_2}$$

Sparse Grids

Function Approximation: The sparse grid interpolation $\mathcal{SG}$, is the hierarchical summation of subspaces $W_{\mathbf{k}}$ such that $\|\mathbf{k}\|_1 \leq \ell + N - 1$.

[E6]
$$\mathcal{SG}f(\mathbf{x}) = \sum_{\|\mathbf{k}\|_1\leq\ell+N-1} \mathcal{I}_{\mathbf{k}}f(\mathbf{x})$$
$$\mathcal{I}_{\mathbf{k}}f(\mathbf{x}) = \sum_{\mathbf{i}\in\mathbf{I}_{\mathbf{k}}} \alpha_{\mathbf{i},\mathbf{k}}\Phi_{\mathbf{i},\mathbf{k}}(\mathbf{x})$$

Basis Function: One-dimensional piecewise linear basis function are generalized to higher dimensions using tensor product construction.

[E7]
$$\Phi_{\mathbf{k},\mathbf{i}}(\mathbf{x}) := \bigotimes_{j=1}^{|\mathbf{x}|} \phi_{k,i}(x_j)$$
$$\phi_{k,i}(x) = \phi\left(\frac{x - ih_k}{h_k}\right)$$

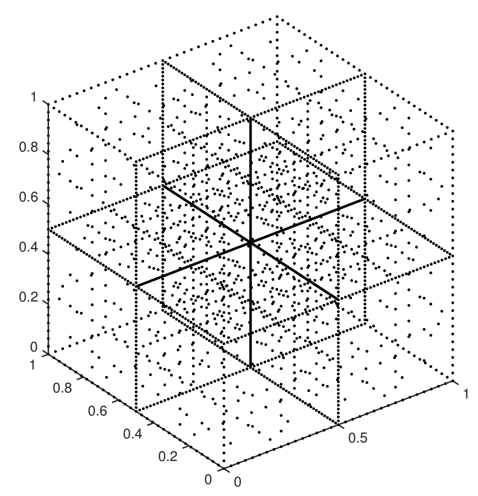

Fig 3: 3D sparse grid at refinement level 7.

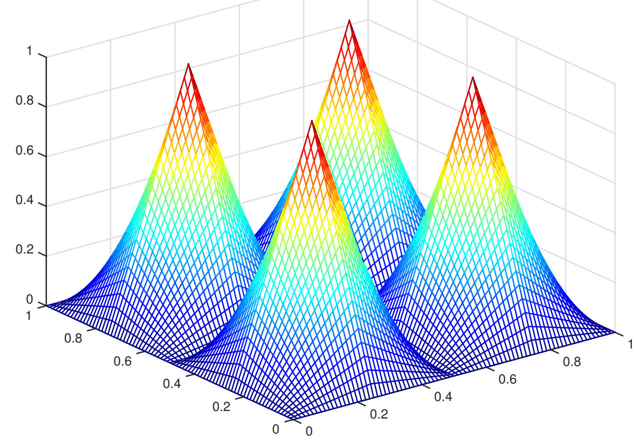

Fig 4: Example of a 2D basis function of the hierarchical increment space $W_{(2,2)}$.

Parallelized Computation Scheme & Results

DDSG Time Iteration Algorithm

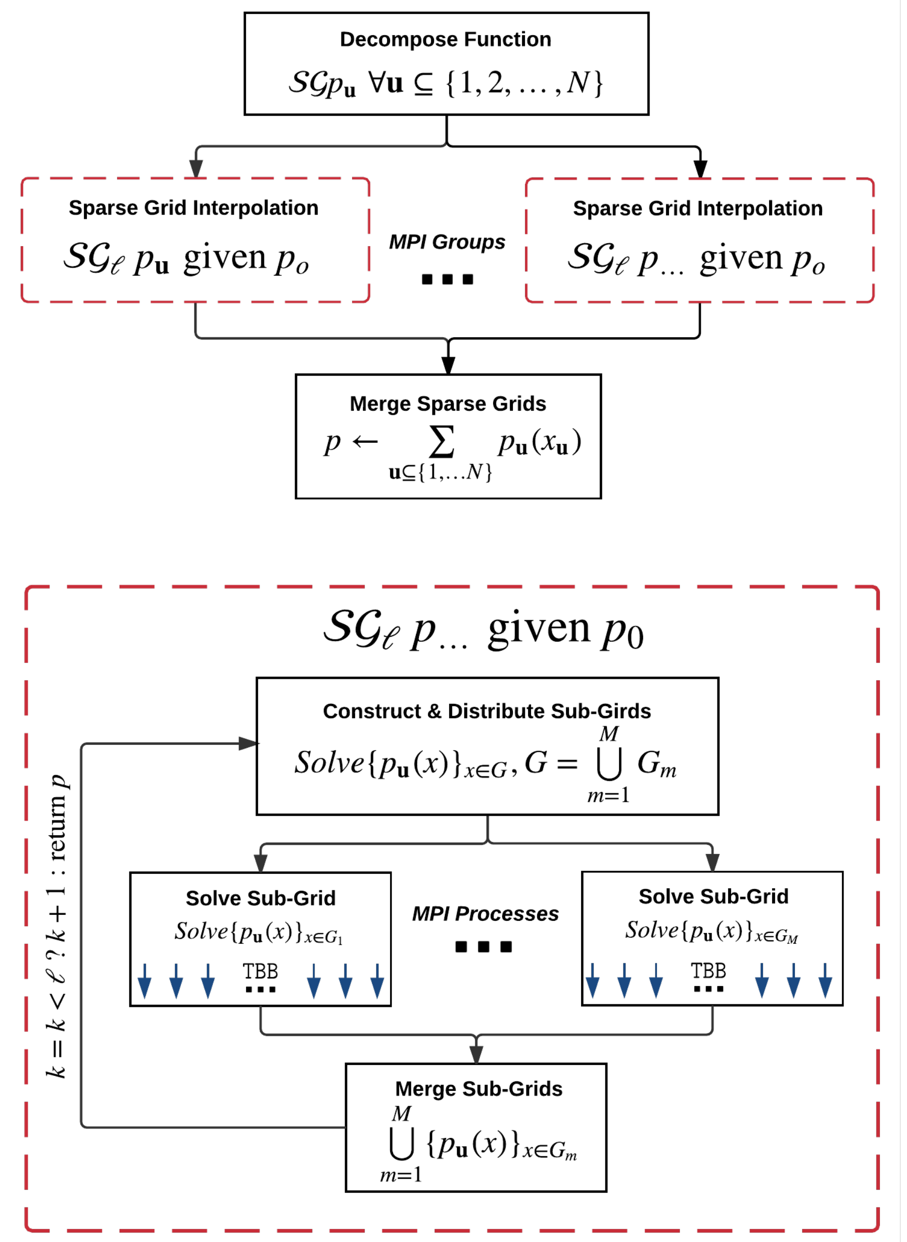

Fig 5: The first level of parallelism from cut-HDMR (top) uses `MPI_Groups` and provides completely separable computation (dashed boxes). The sparse grid routine offers the second level of parallelism (bottom), utilizing `MPI` and Intel(R) `TBB` with synchronization of resources at every refinement level.

Convergence: For dimensions > 8, DDSG takes less execution time for similar convergence in comparison to SG.

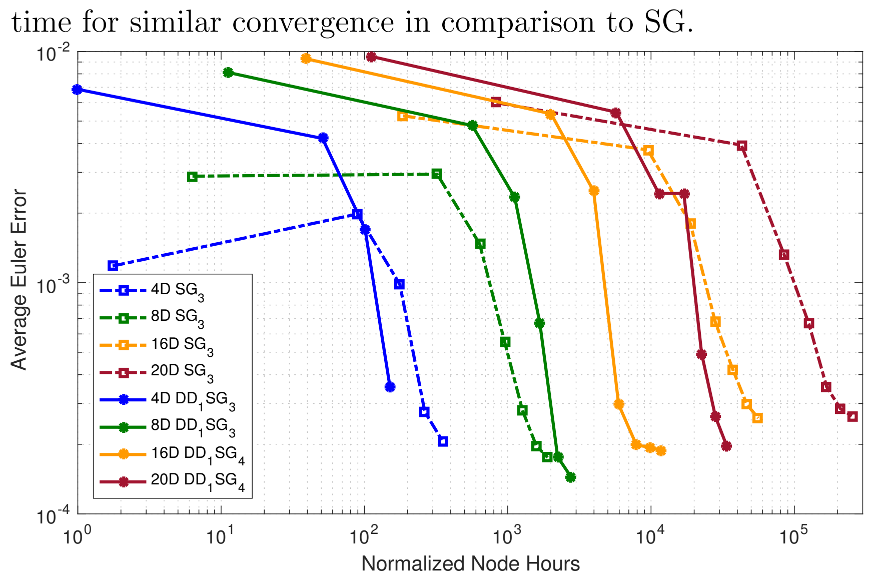

Fig 6: Convergence plot for different sized models using SG refinement level 3 and first order DDSG at refinement level 3 for 4D and 8D and level 4 for 16D and 20D. Normalization is referenced to first iteration time of 4D DDSG.

Scalability: Strong scaling for 50D model with 36k and 80k grid points, using DDSG order 2 at refinement level 4 and 5.

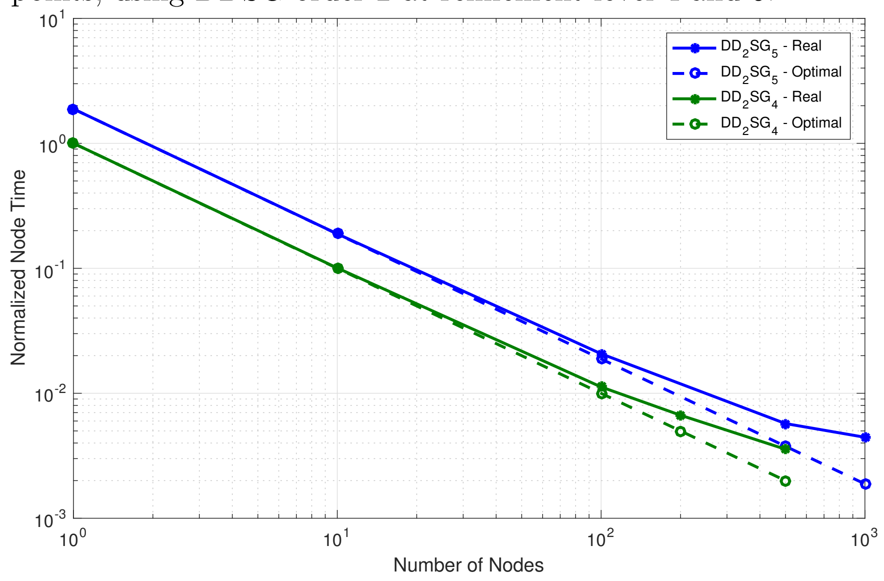

Fig 7: Normalized node hours vs number of nodes, for actual and theoretical parrallelized execution time. Normalization is referenced to single node execution time.

Global Solution: The global solutions for IRBC models up to 200 dimensions, computed using DDSG.

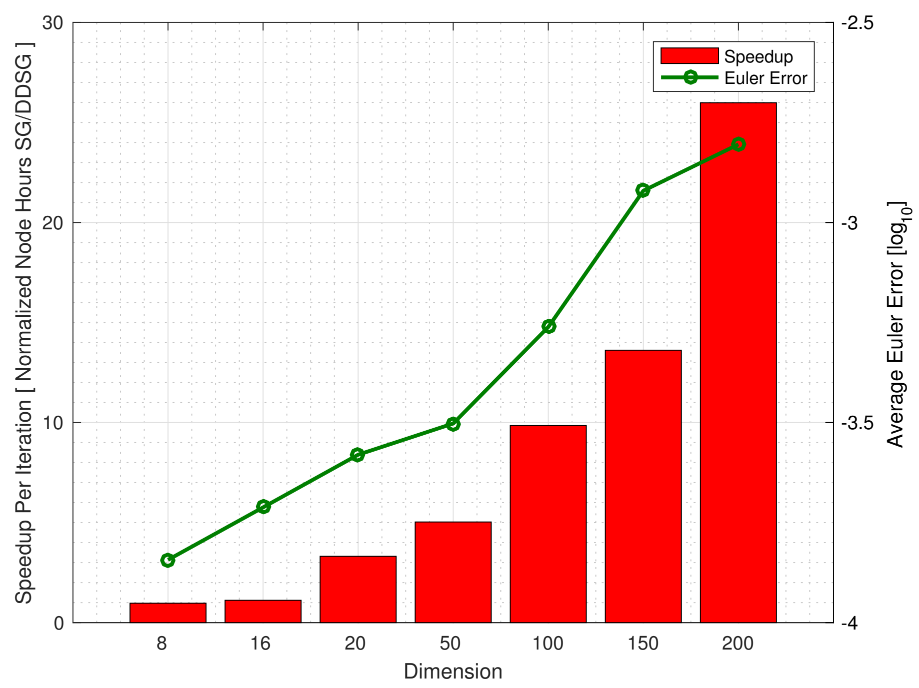

Fig 8: The red bars show the speedup of DDSG order 1 over sparse grids, both at refinement level 2. The green line shows the average Euler error measured after sufficient iterations.

Conclusion: The proposed methodology introduces the following advancements:

■ **Efficient Time-Iteration:** Dimensional decomposition with sparse grids (**DDSG**) allows for significant reduction in the number grid-points. Furthermore, we are able to increase the refinement levels without significant increases in grid points.

■ **Increased Parallelism:** The framework developed introduced significant parallelism with strong scaling that is well suited for large-scale HPC systems.

■ **Global Solution:** We compute the global optimal policy for a 200D dynamic stochastic economic model, which is far beyond what is possible with modern solution methods such Adaptive Sparse Grids.

References: [1] Johannes Brumm, Dmitry Mikushin, Simon Scheidegger, and Olaf Schenk. Scalable high-dimensional dynamic stochastic economic modeling. Journal of Computational Science, 11:12 25, 2015. [2] Johannes Brumm and Simon Scheidegger. Using adaptive sparse grids to solve high-dimensional dynamic models. Available at SSRN 2349281, 2013. [3] Hans-Joachim Bungartz and Michael Griebel. Sparse grids. Acta Numerica, 13:1123, 2004. [4] J. Garcke and M. Griebel. Sparse Grids and Applications. Lecture Notes in Computational Science and Engineering Series. Springer-Verlag GmbH, 2012. [5] Herschel Rabitz Genyuan Li. General formulation of hdmr component functions with independent and correlated variables. 2012. [6] Sheng-Wei Genyuan Li and Herschel Rabitz. High dimensional model representation (hdmr): Concepts and applications. 1999. [7] Markus Holtz. Sparse Grid Quadrature in High Dimensions with Applications in Finance and Insurance. Lecture Notes in Computational Science and Engineering. Springer, Dordrecht, 2011. [8] Giles Hooker. Generalized functional anova diagnostics for high dimensional functions of dependent variables. 2007. [9] Xiang Ma and Nicholas Zabaras. An adaptive high-dimensional stochastic model representation technique for the solution of stochastic partial differential equations. J. Comput. Phys., 2010. [10] Dirk Pfluger. Spatially Adaptive Sparse Grids for High-Dimensional Problems. PhD thesis, Munchen, August 2010. [11] Sharif Rahman. A generalized anova dimensional decomposition for de- pendent probability measures.2007.