A Semi-Lagrangian Scheme using adaptive Sparse Grids for front propagation

Irene Klompmaker

Institut für Mathematik
Technische Universität Berlin
joint work with O. Bokanowski, J. Garcke, M. Griebel

17.02.2011

Contents

- Reinforcement Learning: so-called *agent* must learn from experience good/optimal behaviour
- agent interacts with a dynamic environment
- RL-problems closely related to OC-problems: at least, they lead to an discounted infinite time horizon problem, i.e. one has to solve the HJB

$$\lambda V(x) = \min_{a \in A} \left[f(x,a) DV(x) + r(x,a) \right],$$

  but in contrast state dynamics f and reinforcement r are (at least partially) unknown
- RL-methods
  - *model based* RL (simulate the environment by approximating f and r)
  - *model free* RL (based on observation, no model of the environment)
- all of them use Dynamic Programming methods

- R. Munos ([2]): problems in continuous state space
- adaption of monotone and consistent schemes (finite differences, finite elements/SL) to the RL-case
- but: curse of dimensionality
- idea: use in a similar way sparse grids in order to make progress for higher dimensional RL-problems
- sparse grids: discretization technique that allows to some extent to cope with the curse of dimensionality

- Time dependent finite time horizon problems:
-

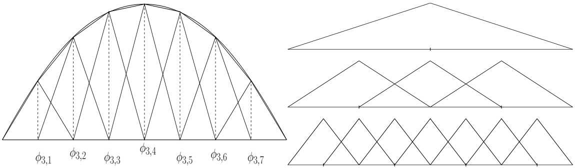
$$u_t + \max_{a \in \mathcal{A}} (f(x,\alpha) \cdot \nabla u) = 0, \quad t \geq 0, \ \ x \in \Omega \quad (1a)$$

$$u(0,x) = \varphi^0(x), \quad x \in \Omega, \quad (1b)$$

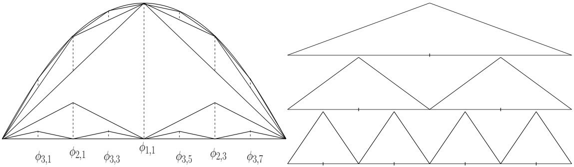
- $\mathcal{A} = \{a : [0,\infty) \to A, \ a(\cdot) \text{ measurable}\}$, A is a compact subset of $\mathbb{R}^m$,
- $\dot{y}(s) = f(y(s), \alpha(s))$, $y(0) = x$, f Lipschitz
- focus on $\Gamma_0 = \{x | u(t,x) = 0\}$, front propagation
- SL-scheme:
  - Initialize grid $\tilde{\Omega}_0$ with v_0, an approximation of φ^0.
  - Iterate for $n = 0, \dots, N-1$,

$$v_{n+1} = \min_{a \in A} v_n^{SG}(y_x^a(-\tau)).$$

Interpolation with Hierarchical Basis



nodal basis for $V_1 \subset V_2 \subset V_3$



hierarchical basis for $V_3 = W_3 \bigoplus W_2 \bigoplus V_1$

- Hierarchical Subspaces:

$$W_{\underline{l}} \ := \ V_{\underline{l}} \setminus \bigoplus_{t=1}^{d} V_{\underline{l}-\underline{e}_t}, \tag{2}$$

- Approximation Space

$$V_n \ := \bigoplus_{|\underline{l}|_1 \leq n} W_{\underline{l}} \tag{3}$$

- Each function $f \in V_n$ can be represented as

$$f(\underline{x}) \ = \sum_{|\underline{l}|_\infty \leq n} \sum_{\underline{j} \in \mathrm{B}_{\underline{l}}} \alpha_{\underline{l},\underline{j}} \cdot \phi_{\underline{l},\underline{j}}(\underline{x}), \tag{4}$$

where $\alpha_{\underline{l},\underline{j}} \in \mathbb{R}$ are called *hierarchical surplus*. They specify what has to be added to the hierarchical representation from level $l-1$ to obtain the one of level l.

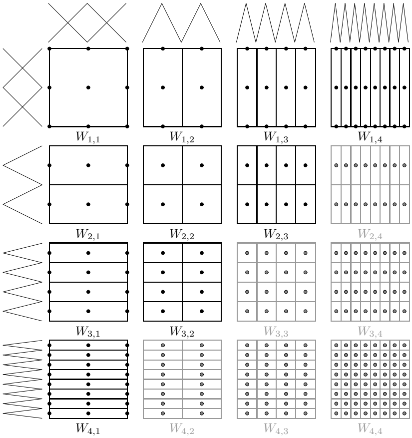


Figure: Supports of the basisfunctions of the hierarchical subspaces $W_{\underline{l}}$ of the space V_4

- number of basis functions used for $f \in V_n$ is $(2^n + 1)^d$: curse of dimensionality
- for $f \in H^2_{mix}(\bar{\Omega})$ it can be shown that for its hierarchical components $f_{\underline{l}} := \sum_{\underline{j} \in \mathrm{B}_{\underline{l}}} \alpha_{\underline{l},\underline{j}} \cdot \phi_{\underline{l},\underline{j}}(\underline{x}) \in W_{\underline{l}}$ it holds

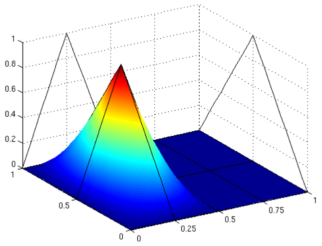
$$\|f_{\underline{l}}\|_2 \leq C(d) \cdot 2^{-2 \cdot |\underline{l}|_1} \cdot |f|_{H^2_{mix}},$$

  i.e. "importance" of basis functin depends on size of support
- Griebel, Zenger ([1, 5]): Sparse Grids
- Sparse Grid Approximation Space:

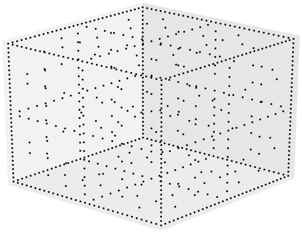
$$V_n^s := \bigoplus_{|\underline{l}|_1 \leq n} W_{\underline{l}}. \tag{5}$$

- For each $f \in V_n^s$

$$f_n^s(\underline{x}) = \sum_{|\underline{l}|_1 \leq n} \sum_{\underline{j} \in \mathrm{B}_{\underline{l}}} \alpha_{\underline{l},\underline{j}} \phi_{\underline{l},\underline{j}}(\underline{x}). \tag{6}$$



(a) Basis function $\phi_{2,1}$ on 2-dim grid



(b) Three-dimensional sparse grid of level $n = 5$

Figure: Example for employed basis function and sparse grid.

Approximation Properties

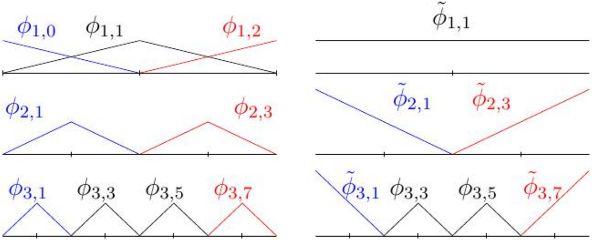
- $h_n := 2^{-n}$, f (sufficiently smooth) defined over a d-dimensional domain
- sparse grid: approximation order $\mathcal{O}(h_n^2 \cdot \log(h_n^{-1})^{d-1})$ with $\mathcal{O}(h_n^{-1} \cdot \log(h_n^{-1})^{d-1})$ points
- full grid: approximation order $\mathcal{O}(h_n^2)$ with $\mathcal{O}(h_n^{-d})$ points

- spatial adaptivity: representation of functions $f \notin H^2_{mix}$, more efficient representation of functions that show significantly differing characteristics (e.g. very steep regions beyond flat ones)
- error indicator based on hierarchical basis: refine if

$$\|\alpha_{\underline{l},\underline{j}}\phi_{\underline{l},\underline{j}}\| = |\alpha_{\underline{l},\underline{j}}| \cdot \gamma > \varepsilon,$$

  where γ depends on the norm we take into account for the refinement (f.e. $\gamma = 1$ for $\|\cdot\|_\infty$)
- in the same way coarsening (against over-refinement): coarsen if $|\alpha_{\underline{l},\underline{j}}| \cdot \gamma < \eta$
- both refinement and coarsening have to keep the grid consistent

Modified basis functions for boundary treatment ([3])



(a) Hierarchical basis functions $\phi_{i,j}$ up to level 3 (b) Modified basis functions $\phi_{i,j}$ up to level 3

Consider an adaptive sparse grid Ω_k and its corresponding sparse grid function $v_k \in V_n^s$ at some time $t_k = k\tau$, where $\tau := T/N$ is the time step.

1. Initialize $\tilde{\Omega}_0$ with initial grid function v_0 (by interpolating initial function φ^0 on an adaptive SG, using spatial adaptivity with refinement constant ε, coarsen with coarsening constant η)
2. Iterate in time for $k = 0, \ldots, N-1$,
   (a) Initialize $\tilde{\Omega}_{k+1} = \tilde{\Omega}_k$.
   (b) Compute $v_{k+1}(x) = \min_{a \in \mathcal{A}} v_k(y_x^\alpha(-\tau))$ for all $x \in \tilde{\Omega}_{k+1}$
   (c) While refinement is needed (using constant ε):
      (i) Refine $\tilde{\Omega}_{k+1}$
      (ii) Compute $v_{k+1}(x) = \min_{\alpha \in \mathcal{A}} v_k(y_x^\alpha(-\tau))$ on new points $x \in \tilde{\Omega}_{k+1}$ from (i)
   (d) Coarsen $\tilde{\Omega}_{k+1}$ according to constant η

- Consider

$$v_t + f(x) \cdot \nabla v = 0, \quad t \geq 0, \quad x \in \Omega \tag{7a}$$

$$v(0,x) = \varphi(x), \quad x \in \Omega, \tag{7b}$$

- where $f = (-1, \ldots, -1)$ and

$$\varphi(x) := -\frac{r_0}{2} + \frac{1}{2r_0}\|x - a\|_2^2, \quad \text{with } r_0 = 0.5$$

- zero level set $\{x, \varphi(x) = 0\}$ represents the sphere of radius r_0 centered at a

$t = 0$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	379	1.46_{-2}			1.77_{-2}		
2.00_{-3}	763	3.65_{-3}	1.00	−1.98	4.42_{-3}	1.00	−1.98
5.00_{-4}	1,531	9.14_{-4}	1.00	−1.99	1.10_{-3}	1.00	−2.00
1.25_{-4}	3,067	2.28_{-4}	1.00	−2.00	2.76_{-4}	1.00	−1.99

$t = 1$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	166	1.55_{-1}			2.64_{-1}		
2.00_{-3}	304	2.70_{-2}	1.26	−2.89	4.54_{-2}	1.27	−2.91
5.00_{-4}	595	9.70_{-3}	0.74	−1.52	1.65_{-2}	0.73	−1.51
1.25_{-4}	1,168	1.69_{-3}	1.26	−2.59	2.84_{-3}	1.27	−2.61

Table: Example 1: convergence table for $d = 3$, initial data ($t = 0$) and terminal data ($t = 1$) after $N = 10$ time steps.

$\underline{t = 0}$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	625	1.95_{-2}			2.30_{-2}		
2.00_{-3}	1,265	4.87_{-3}	1.00	−1.97	5.77_{-3}	1.00	−1.96
5.00_{-4}	2,545	1.21_{-3}	1.00	−1.99	1.44_{-3}	1.00	−1.99
1.25_{-4}	5,105	3.04_{-4}	1.00	−1.98	3.59_{-4}	1.00	−2.00

$\underline{t = 1}$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	229	2.07_{-1}			3.47_{-1}		
2.00_{-3}	413	3.61_{-2}	1.26	−2.96	5.98_{-2}	1.27	−2.98
5.00_{-4}	801	1.29_{-2}	0.74	−1.55	2.17_{-2}	0.73	−1.53
1.25_{-4}	1,565	2.25_{-3}	1.26	−2.61	3.73_{-3}	1.27	−2.63

Table: Example 1: convergence table for $d = 4$, initial data ($t = 0$) and terminal data ($t = 1$) after $N = 10$ time steps.

				$t = 1$			
d	n	$e_{L^\infty_{loc}}$	$\alpha_{d,e}$	$e_{L^2_{loc}}$	$\alpha_{d,e}$	$\alpha_{d,n}$	$\frac{n(d)}{n(d-1)}$
2	393	6.46_{-3}		1.07_{-2}			
3	595	9.70_{-3}	1.00	1.65_{-2}	1.07	1.02	1.51
4	801	1.29_{-2}	0.99	2.17_{-2}	0.95	1.03	1.35
5	1,011	1.61_{-2}	0.99	2.63_{-2}	0.86	1.04	1.26
6	1,255	1.94_{-2}	1.02	3.05_{-2}	0.81	1.19	1.24

Table: Example 1, scaling analysis ($2 \leq d \leq 6$) for $t = 1$, $N = 10$ time steps, using $\varepsilon = 5\ 10^{-4}$ and $\eta = \varepsilon/10$

		$t = 1$	
d	n	$e_{L^\infty_{loc}}$	$e_{L^2_{loc}}$
2	100^2	1.32_{-1}	9.14_{-2}
3	100^3	1.26_{-1}	8.49_{-2}
4	100^4	1.03_{-1}	7.21_{-2}
5	100^5	-	-

Table: Example 1, **finite difference scheme**, uniform grid with 100^d mesh points.

- We consider the following transport equation

$$v_t + f(x) \cdot \nabla v = 0, \quad t \geq 0, \quad x \in \Omega \tag{8a}$$
$$v(0,x) = \varphi(x), \quad x \in \Omega, \tag{8b}$$

- $\Omega = [-2,2]^d$ and $t = 0.5$
- assume $-f$ to be an inward pointing flow, the backward characteristics of the flow stay inside the domain Ω
-

$$f(x) = \left(\max \left(1 - \frac{\|x\|}{r_{\max}} \right)_+ \right)^3 f_1(x), \quad r_{\max} = 1 \text{ or } 1.5$$

where $f_1(x)$ is a rotation in $\mathbb{R}^d$
- $\varphi(x) = x_2$

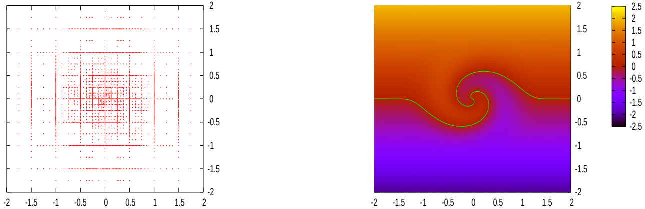


Figure: Example 2, left: resulting grid, right: graph of resulting function (with zero levelset) in 2D, for $\varepsilon = 2\ 10^{-3}$ and $\eta = \varepsilon/5$

$\underline{t=1}$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	845	1.92_{-1}			1.22_{-1}		
4.00_{-3}	1,423	9.21_{-2}	1.06	-1.41	5.63_{-2}	1.12	-1.48
2.00_{-3}	2,189	4.33_{-2}	1.09	-1.75	2.75_{-2}	1.03	-1.66
1.00_{-3}	3,405	2.08_{-2}	1.06	-1.66	1.57_{-2}	0.81	-1.27
5.00_{-4}	5,135	1.08_{-2}	0.95	-1.60	8.68_{-3}	0.85	-1.44
2.50_{-4}	7,863	6.03_{-3}	0.84	-1.37	4.82_{-3}	0.85	-1.38

Table: Example 2: convergence table for $d = 2$, terminal data ($t = 1$) after $N = 10$ time steps.

$\underline{t = 1}$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	4,281	4.57_{-1}			1.17_{-1}		
4.00_{-3}	8,439	2.61_{-1}	0.81	-0.83	6.08_{-2}	0.94	-0.96
2.00_{-3}	15,593	1.03_{-1}	1.34	-1.51	2.99_{-2}	1.02	-1.16
1.00_{-3}	26,131	4.81_{-2}	1.10	-1.47	1.71_{-2}	0.81	-1.08
5.00_{-4}	44,497	2.32_{-2}	1.05	-1.37	9.77_{-3}	0.81	-1.05
2.50_{-4}	74,155	1.36_{-2}	0.77	-1.05	5.64_{-3}	0.79	-1.08

Table: Example 2: convergence table for $d = 3$, terminal data $(t = 1)$ after $N = 10$ time steps.

	$t = 1$	
d	n	time (sec)
2	2189	2.54
3	15593	112.52
4	90973	598.80
5	462327	6134.13

Table: Example 1, $N = 10$ time steps, computing times

- We consider the eikonal equation:

$$v_t + \|\nabla v\|_2 = 0, \quad t \geq 0, \quad x \in \Omega \tag{9a}$$
$$v(0, x) = \varphi(x), \quad x \in \Omega, \tag{9b}$$

- $\Omega = (-2, 2)^d$
- $\varphi(x) := q(\|x\|)$ with

$$q(x) := -\frac{r_0}{2} + \frac{x^2}{2r_0}$$

  and $r_0 = 0.5$ (q is chosen such that $q(x) = 0$ for $x = r_0$ and $q'(r_0) = 1$)
- zero level set $\{x, \varphi(x) = 0\}$ represents the sphere of radius r_0
- $f(x, \alpha) = c(x) \cdot \alpha$, $a \in A = B(0, 1)^d$, $c(x) = 1$
- simplification: assume the optimal control $\alpha = \frac{x}{||x||}$

$t = 0$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	189	9.75_{-3}			1.24_{-2}		
2.00_{-3}	381	2.43_{-3}	1.00	-1.98	2.91_{-3}	1.05	-2.07
5.00_{-4}	765	5.85_{-4}	1.03	-2.04	7.24_{-4}	1.00	-2.00
1.25_{-4}	1,533	1.52_{-4}	0.97	-1.94	1.83_{-4}	0.99	-1.98

$t = 1$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	261	1.26_{-1}			1.29_{-1}		
2.00_{-3}	705	3.40_{-2}	0.94	-1.32	3.23_{-2}	1.00	-1.39
5.00_{-4}	1,889	1.12_{-2}	0.80	-1.13	1.01_{-2}	0.84	-1.18
1.25_{-4}	4,745	2.92_{-3}	0.97	-1.46	2.99_{-3}	0.88	-1.32

Table: Example 3: convergence table for $d = 2$, initial data ($t = 0$) and terminal data ($t = 1$) after $N = 10$ time steps.

$t = 0$

d	n	$e_{L^\infty_{loc}}$	$\alpha_{d,e}$	$e_{L^2_{loc}}$	$\alpha_{d,e}$	$\alpha_{d,n}$	$\frac{n(d)}{n(d-1)}$
2	189	9.75_{-3}		1.24_{-2}			
3	373	1.46_{-2}	1.00	1.87_{-2}	1.01	1.68	1.97
4	617	1.94_{-2}	0.99	2.41_{-2}	0.88	1.75	1.65
5	921	2.40_{-2}	0.95	2.85_{-2}	0.75	1.80	1.49

$t = 1$

d	n	$e_{L^\infty_{loc}}$	$\alpha_{d,e}$	$e_{L^2_{loc}}$	$\alpha_{d,e}$	$\alpha_{d,n}$	$\frac{n(d)}{n(d-1)}$
2	261	1.26_{-1}		1.29_{-1}			
3	1,085	1.54_{-1}	0.49	1.67_{-1}	0.64	3.51	4.16
4	5,401	1.72_{-1}	0.38	1.96_{-1}	0.56	5.58	4.98
5	47,153	2.12_{-1}	0.94	1.63_{-1}	-0.83	9.71	8.73

Table: Example 3, scaling analysis ($2 \leq d \leq 5$) for initial data ($t = 0$) and terminal data ($t = 1$) after $N = 10$ time steps, using $\varepsilon = 8\ 10^{-3}$ and $\eta = \varepsilon/5$ for the adaptive procedure.

- $$v_t + \sum_{i=1,\dots,d} |\partial_{x_i} v| = 0, \quad t \geq 0, \quad x \in \Omega \tag{10a}$$
  $$v(0,x) = \varphi(x), \quad x \in \Omega, \tag{10b}$$
- $\varphi(x) := q(\|x\|)$ as above
- (10a) is equivalent to

  $$v_t + \max_{u \in \mathcal{U}} u \cdot \nabla v = 0, \quad t \geq 0, \quad x \in \Omega$$

  where $\mathcal{U} = \{u = (u_1, \dots, u_d),\ u_i = \pm 1\}$ ($\mathcal{U}$ is a set of 2^d controls).
- for the scheme: use explicit paths $y_x^u(-h) = x - hu$ for all $u \in \mathcal{U}$.

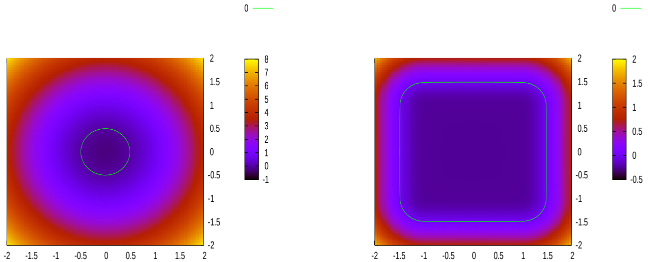


Figure: Example 4, top left: left: graph of initial function, right: resulting function for $\varepsilon = 2\ 10^{-3}$ and $\eta = \varepsilon/5$

$t = 0$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	273	9.75_{-3}			1.24_{-2}		
2.00_{-3}	529	2.43_{-3}	1.00	−2.10	2.91_{-3}	1.05	−2.19
5.00_{-4}	1,041	5.85_{-4}	1.03	−2.10	7.24_{-4}	1.00	−2.06
1.25_{-4}	2,065	1.52_{-4}	0.97	−1.97	1.83_{-4}	0.99	−2.01

$t = 1$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	185	3.80_{-1}			4.46_{-1}		
2.00_{-3}	313	6.49_{-2}	1.27	−3.36	8.03_{-2}	1.24	−3.26
5.00_{-4}	585	2.40_{-2}	0.72	−1.59	2.94_{-2}	0.72	−1.61
1.25_{-4}	1,081	4.05_{-3}	1.28	−2.90	5.51_{-3}	1.21	−2.73

Table: Example 4 using common basis functions with boundary points, convergence table for $d = 2$, initial data ($t = 0$) and terminal data ($t = 1$) after $N = 40$ time steps.

$t = 0$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	189	9.75_{-3}			1.24_{-2}		
2.00_{-3}	381	2.43_{-3}	1.00	-1.98	2.91_{-3}	1.05	-2.07
5.00_{-4}	765	5.85_{-4}	1.03	-2.04	7.24_{-4}	1.00	-2.00
1.25_{-4}	1,533	1.52_{-4}	0.97	-1.94	1.83_{-4}	0.99	-1.98

$t = 1$

ε	n	$e_{L^\infty_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$	$e_{L^2_{loc}}$	$\alpha_{\varepsilon,e}$	$\alpha_{n,e}$
8.00_{-3}	70	3.80_{-1}			4.51_{-1}		
2.00_{-3}	149	6.49_{-2}	1.27	-2.34	8.15_{-2}	1.23	-2.26
5.00_{-4}	285	2.40_{-2}	0.72	-1.53	2.97_{-2}	0.73	-1.56
1.25_{-4}	533	4.05_{-3}	1.28	-2.84	5.57_{-3}	1.21	-2.67

Table: Example 4 using modified basis functions, convergence table for $d = 2$, initial data ($t = 0$) and terminal data ($t = 1$) after $N = 40$ time steps.

$t = 0$

d	n	$e_{L^\infty_{loc}}$	$\alpha_{d,e}$	$e_{L^2_{loc}}$	$\alpha_{d,e}$	$\alpha_{d,n}$	$\frac{n(d)}{n(d-1)}$
2	381	2.43_{-3}		2.91_{-3}			
3	757	3.65_{-3}	1.00	4.45_{-3}	1.05	1.69	1.99
4	1,257	4.87_{-3}	1.00	5.80_{-3}	0.92	1.76	1.66
5	1,881	6.09_{-3}	1.00	6.98_{-3}	0.83	1.81	1.50

$t = 1$

d	n	$e_{L^\infty_{loc}}$	$\alpha_{d,e}$	$e_{L^2_{loc}}$	$\alpha_{d,e}$	$\alpha_{d,n}$	$\frac{n(d)}{n(d-1)}$
2	149	6.49_{-2}		8.15_{-2}			
3	229	9.73_{-2}	1.00	1.10_{-1}	0.74	1.06	1.54
4	313	1.29_{-1}	0.98	1.39_{-1}	0.81	1.09	1.37
5	401	1.62_{-1}	1.02	1.67_{-1}	0.82	1.11	1.28

Table: Example 4, scaling analysis ($2 \leq d \leq 5$) for initial data ($t = 0$) and terminal data ($t = 1$) after $N = 40$ time steps, using modified basis functions and $\varepsilon = 2\ 10^{-3}$ and $\eta = \varepsilon/5$ for the adaptive procedure.

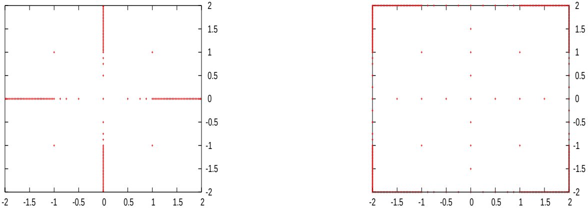


Figure: Example 4, left: resulting grid using modified basis functions, right: resulting grid using common hat basis functions with boundary, $\varepsilon = 2\ 10^{-3}$ and $\eta = \varepsilon/5$

- We consider the eikonal equation:

$$v_t + \|\nabla v\|_2 = 0, \quad t \geq 0, \quad x \in \Omega \tag{11a}$$

$$v(0,x) = \varphi(x), \quad x \in \Omega, \tag{11b}$$

- initial function: $\varphi(x) := min(q(||x-a||), q(||x-b||), \; a \neq b$
- $f(x,\alpha) = c(x) \cdot \alpha$, $a \in A = B(0,1)^2$, $c(x) = 1$
- discretization of A

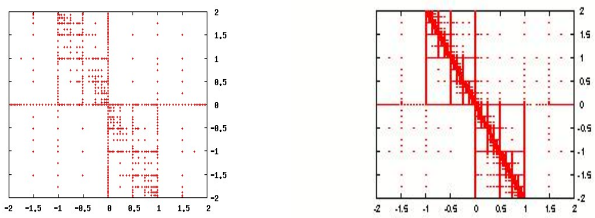


Figure: Example 5, left: L_2-based refinement, $e_{L^\infty_{loc}} = 3.33e-01$; right: L_∞-based refinement, $e_{L^\infty_{loc}} = 2.90e-01$

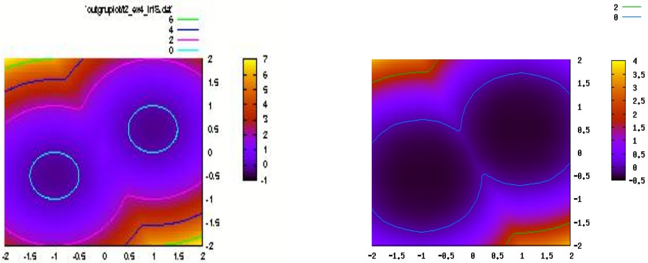


Figure: Example 5, left: initial function, right: function for $t = 0.75$ after $N = 100$ iterations, using L_2-based refinement

Conclusions

- Adaptive SGSL scheme works
- Discontinuous solutions, irregular solutions: refinement strategies
- Convergence? No monotonicity!

M. Griebel.
A parallelizable and vectorizable multi-level algorithm on sparse grids.
In W. Hackbusch, editor, *Parallel Algorithms for partial differential equations, Notes on Numerical Fluid Mechanics*, volume 31, pages 94–100. Vieweg Verlag, Braunschweig, 1991.

R. Munos.
A study of reinforcement leaning in the continuous case by the means of viscosity solutions.
Machine Learning, 40:265–299, 2000.

D. Pflüger.
Spatially Adaptive Sparse Grids for High Dimensional Problems.
PhD thesis, TU München, 2010.

H. Yserentant.

On the multi-level splitting of finite element spaces.
Numerische Mathematik, 49:379–412, 1986.

C. Zenger.
Sparse grids.
In W. Hackbusch, editor, *Parallel Algorithms for Partial Differential Equations, Proceedings of the Sixth GAMM-Seminar, Kiel, 1990*, volume 31 of *Notes on Num. Fluid Mech.*, pages 241–251. Vieweg-Verlag, 1991.