Regularization of parameter estimation by adaptive discretization using refinement and coarsening indicators†

Hend Ben Ameur‡ and Barbara Kaltenbacher⋆

‡ FSB and ENIT-Lamsin, université de Tunis, BP 37, Le Belvédère, Tunis, Tunisia,
E-mail : Hend.Benameur@enit.rnu.tn

⋆ SFB 013 *Numerical and Symbolic Scientific Computing, Johannes Kepler Universität Linz, Freistädterstr. 313, Linz, Austria, E-mail: barbara@sfb013.uni-linz.ac.at*

Abstract. This paper is concerned with the ill-posed problem of identifying a hydraulic transmissivity in an isotropic and confined aquifer in two space dimensions. To define a regularization by adaptive discretization of the parameter, we use refinement and coarsening indicators, which give the first order effect on the optimal data misfit of adding or removing degrees of freedom to a current set of parameters. The direct problem is discretized by a mixed hybrid finite element method. A combination of the direct problem discretization with the adaptive discretrization for the transmissivity allows us to prove convergence of the algorithm as the mesh size goes to zero, stability of the finite-dimensional approximation, and convergence for noisy data with an appropriate stopping rule.

† Supported by the Austrian Science Foundation 'Fonds zur Förderung der wissenschaftlichen Forschung' in the Special Research Program SFB F013 and under grant T7-TEC

INTRODUCTION

We estimate hydraulic transmissivities, which are space dependent coefficients in a parabolic partial differential equations, by minimizing an objective function defined as a least-squares misfit between observations and the corresponding quantities computed with a "chosen" descritization of the transmissivities. In order to find a discretization of the transmissivity which reduces the number of unknowns, we use refinement and coarsening technique [Benameur-Chavent-Jaffré, 2001]. This method, compared to the multiscale one [Liu, 1993, Chavent-Liu, 1989], has the advantage of avoiding overparameterization. At a given refinement level the parameters are estimated by minimizing the misfit to the data. Then we compute refinement and coarsening indicators which indicate the effect on the optimal data misfit of adding or removing some degrees of freedom and so lead us to chose the "best" parameter discretization in the next refinement level.
Using the idea of regularization by discretization [Engl-Hanke-Neubauer, 1996, Kaltenbacher-Schöberl, 2001], we prove a convergence result of this method of parameter estimation. This result is composed by three steps:

• Convergence of the algorithm to the exact parameter for exact data, as the mesh size goes to zero,

• Stability of the finite dimentional problem (regularization by discretization), and

• Convergence for noisy data with appropriate stopping rule.

We apply the refinement and coarsening indicator technique to the estimation of the distributed transmissivity parameter T in the partial differential equation

$$S\frac{\partial \Phi}{\partial t} + \mathrm{div}\ (-T \ \mathrm{grad}\ \Phi) = Q \text{ in } \Omega, \tag{1}$$

where Ω is a domain in $\mathbb{R}^2$. This equation governs a two-dimensional groundwater flow in an isotropic and confined aquifer, subject to boundary and initial conditions:

$$\begin{aligned} &\Phi = \Phi_d \text{ on } \Gamma_D, \quad (-T \ \mathrm{grad}\ \Phi)\cdot n = q_d \text{ on } \Gamma_N, \\ &\Phi(0) = \Phi_0 \text{ in } \Omega, \end{aligned} \tag{2}$$

where Γ_D and Γ_N is a partition of the boundary of Ω supporting respectively Dirichlet and Neumann conditions and where Φ is the piezometric head, $S(x,y)$ is the storage coefficient, $T(x,y)$ is the hydraulic transmissivity, $Q(x,y,t)$ is the distributed source term, Φ_d, q_d are given boundary piezometric head and source terms, Φ_0 is a given initial piezometric head and $\vec{n}$ is the unit normal vector to Ω.
The domain Ω is discretized with a mesh $\mathcal{T}_h$ for the state variable Φ (the size of each cell in $\mathcal{T}_h$ is h^2), and equations (1),(2) are approximated by a mixed-hybrid finite element method, which is suitable to the case where the parameter T have discontinuities [Chavent-Roberts, 1997]. Time discretization is done by finite differences with a time step Δt.

The transmissivity T is a piecewise constant space function. Both the zonation (the partition of Ω into zones where T is constant) and the value of T on each zone are to be determined.

We define our objective function by:

$$J(T) = \frac{1}{2} \sum_{i,j} \Delta t \; h^2 |\Phi(x_j, t_i) - \Phi_{i,j}^{obs}|^2. \tag{3}$$

where $\Phi_{i,j}^{obs}$ is the piezometric head measured in the point x_j at the time t_i and $\Phi(x_j, t_i)$ is the model output for current transmissivity values.

In order to have a computational cost independent of the number of transmissivity values to estimate, we use the Quasi-Newton gradient algorithm described in [Bonnans et al., 1997], which is known to be efficient in the case of a large number of unknown parameters. The gradient of J is computed by the adjoint state method [Sun, 1994].

1. ADAPTIVE DISCRETISATION OF THE HYDRAULIC TRANSMISSIVITY

We consider a partition of Ω into zones where the transmissivity T is constant. We denote by $\mathcal{Z}$ this zonation. The degrees of freedom of the transmissivity to be estimated are the values of the transmissivity in each zone of $\mathcal{Z}$. We want this number to be significantly lower than the number of elements in the computing mesh $\mathcal{T}_h$. For this purpose we use the refinement and coarsening indicators technique introduced in [Benameur-Chavent-Jaffré, 2001]. Given a current parameterization, these indicators tell us where to insert new discontinuities of the transmissivity and which ones can be removed.

1.1. Refinement indicators

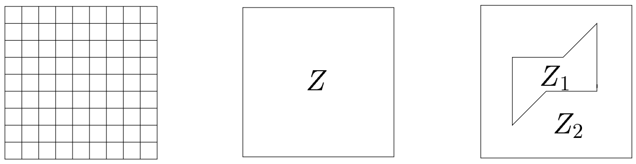
Let $(\mathcal{P}_1)$ be an initial problem where the hydraulic transmissivity is constant in all the domain (figure 1(b)). So we have only one value of the transmissivity to estimate, which is done by minimizing the objective function J with respect to this single variable. We note $(\mathcal{P}_2)$ the parameter estimation problem corresponding to a tentative partition of Ω into two zones Z_1 and Z_2 having different transmissivities (figure 1(c)). We note $T^* = (T_1^*, T_2^*)$ the solution of $(\mathcal{P}_2)$ obtained by minimizing the objective function with respect to the corresponding two variables. If $B = T_1^* - T_2^*$ were known, then the solution of minimizing J under the constraint $T_1 - T_2 = B$ is necessarily $T^* = (T_1^*, T_2^*)$, the solution of $(\mathcal{P}_2)$. But for $B = 0$ the $T = (T_0^*, T_0^*)$ solution of the constrained problem is obviously the solution of $(\mathcal{P}_1)$.

To the problem of minimizing the objective function under these constraints:

$$J(T^*) = \min_{AT=B} J(T),$$

we associate the Lagrangian function defined by

$$\mathcal{L}_B(T, \lambda) = J(T) - < \lambda, AT - B > \tag{4}$$



(a) Computing mesh $\mathcal{T}$

(b) Partition $\mathcal{Z}^1$ associated to prob. $(\mathcal{P}_1)$

(c) Partition $\mathcal{Z}^2$ associated to prob. $(\mathcal{P}_2)$

Figure 1. An example of parameterization refinement

where λ is the Lagrange multiplier associated to the constraint $AT = B$. Then the Lagrange condition ensures that T^* is obtained by solving

$$\frac{\partial \mathcal{L}_B}{\partial T}(T^*, \lambda^*) = \nabla J(T^*) - A^T \lambda^* = 0 \tag{5}$$

$$\frac{\partial \mathcal{L}_B}{\partial \lambda}(T^*, \lambda^*) = AT^* - B = 0. \tag{6}$$

If we denote by $J^*_B = J(T^*) = \mathcal{L}_B(T^*, \lambda^*)$ the optimal misfit associated to the right hand side B of the constraint, we deduce from (4), (5), and (6) that

$$\frac{\partial J^*_B}{\partial B}|_{B=0} = \frac{\partial \mathcal{L}}{\partial B}(T^*, \lambda^*)|_{B=0} = \lambda^*. \tag{7}$$

Such that we have the first order development:

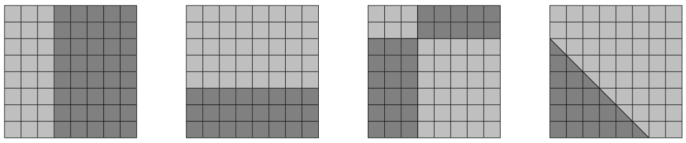
$$J^*_B = J^*_0 + \lambda^*(T^*_1 - T^*_2) + ...$$

Therefore the Lagrange multiplier, or refinement indicator gives us the sensitivity of the optimal data misfit J^*_B to the perturbation B. Without solving $(\mathcal{P}_2)$, this indicator, which can be easily deduced from equation (5), indicates through its absolute value, if the suggested refinement is likely to induce an important variation of the optimal objective function. In the case of our example, the refinement indicator associated to splitting the single zone of figure 1(b)) into two zones of figure 1(c) is :

$$\lambda^* = \frac{\partial J}{\partial T_1}(T^*) = -\frac{\partial J}{\partial T_2}(T^*), \text{ where } T = (T^*_0, T^*_0). \tag{8}$$

Such a refinement is defined by introducing a curve called “cut” (the boundary of Z_2 in fig. 1(c)). It divides the domain into only two zones of different transmissivities. In practice, the families of elementary cuts shown in fig. 2 are used.

In all cases, refinement indicators can be computed easily, from the partial derivatives $\frac{\partial J}{\partial T_j}(T^*)$ of the objective function with respect to transmissivities [Benameur-Chavent-Jaffré, 2001]. Hence indicators associated to a large number of tentative cuts can be evaluated, and only the few ones with the largest absolute values are retained.



(a) (b) (c) (d)

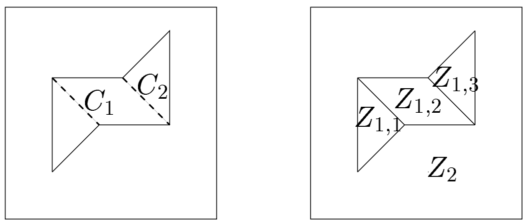
Figure 2. Four elementary families of cuts: (a) vertical, (b) horizontal, (c) checkerboard, (d) oblique

The objective function is then minimized with a new discretization of the transmissivity obtained by implementing the cut corresponding to one of the selected indicators in the current zonation. Finally, the refinement $\mathcal{Z}'$ of $\mathcal{Z}$ which corresponds to the cut providing the largest decrease of J is implemented; $\mathcal{Z}'$ has only one more degree of freedom than $\mathcal{Z}$, but produces a significant decrease of the objective function.

1.2. Coarsening indicators

Consider the example shown in fig. 1, and suppose that in a first step the cut represented in fig. 1(c) is selected: the zonation $\mathcal{Z}$ contains two zones, an interior zone Z_1 and an exterior zone Z_2, with the corresponding optimal values of transmissivity T_1^* and T_2^*. Suppose that refinement indicators corresponding to various positions and forms of cuts have been computed and that two cuts C_1 and C_2 in the interior zone (fig. 3 left), which both correspond to an indicator λ with a strong modulus, have been selected.

Before choosing definitively this new zonation, one wants to make sure that the four degrees of freedom $Z_{1,1}$, $Z_{1,2}$, $Z_{1,3}$ and Z_2 are not over-abundant (fig. 3 right). To do that, one checks whether the current discontinuity $T_1^* - T_2^*$ between Z_1 and Z_2 could be set to zero on some part of $\partial Z_1 \cap \partial Z_2$ without increasing the optimal misfit $J^*(T_1^*, T_2^*)$. We notice that



Two selected cuts New parametrization to study

Figure 3. A refinement for figure 1(c)

minimizing the two variables objective function $J(T_1, T_2)$ over the current parametrization of fig. 1(c), is equivalent to minimizing $J(T_{1,1}, T_{1,2}, T_{1,3}, T_2)$ a function of four variables, and aggregating for example the two zones $Z_{1,1}$ and Z_2 is equivalent to set to 0 the discontinuity $T_{1,1} - T_2$. As detailed in subsection 1.1,the effect of this aggregation on the optimal data misfit,at first order, is measured by the corresponding Lagrange multiplier

$$\lambda_1 = \frac{\partial J}{\partial T_{1,1}}(T_1^*, T_2^*). \tag{9}$$

Similarly the effects of aggregating $Z_{1,2}$ with Z_2 and $Z_{1,3}$ with Z_2 are measured respectively by

$$\begin{aligned} \lambda_2 &= \frac{\partial J}{\partial T_{1,2}}(T_1^*, T_2^*), \\ \lambda_3 &= \frac{\partial J}{\partial T_{1,3}}(T_1^*, T_2^*). \end{aligned} \tag{10}$$

Notice that one has

$$\lambda_1 + \lambda_2 + \lambda_3 = \frac{\partial J}{\partial T_1}(T_1^*, T_2^*) = -\frac{\partial J}{\partial T_2}(T_1^*, T_2^*) = 0. \tag{11}$$

Multiplying (11) by $T_1^* - T_2^*$ we see that

$$\Delta J_1^* + \Delta J_2^* + \Delta J_3^* = 0 \tag{12}$$

Hence one at least of the ΔJ_j^* is negative, so that aggregating the degrees of freedom $Z_{1,j}$ and Z_2 for which $\Delta J_j^* < 0$ will lower the number of degrees of freedom and enhance, at least at first order, the fit to the data. After this coarsening indicator step, the new zonation $\mathcal{Z}'$ will have only one more degree of freedom than $\mathcal{Z}$ (if only one of the ΔJ_j^* is negative), or even the same number of degrees of freedom (if two of the ΔJ_j^* are negative).

1.3. Algorithm

We give now the algorithm used to define adaptive discretization of transmissivity using refinement and coarsening indicators:

1. Choose an initial zonation $\mathcal{Z}$.

2. **Do** until data are satisfactorily fitted:

    3. Estimate the transmissivity with the current zonation $\mathcal{Z}$ by minimizing J.

    4. **For** every zone Z_i of $\mathcal{Z}$ **do**

        Compute all the refinement indicators I corresponding to the chosen family of cuts.

        Enddo

    5. Compute I_{max} the largest absolute value of all computed refinement indicators in all parts Z_i. Select all cuts corresponding to refinement indicators which are larger than 80% of I_{max} (this percentage can be adjusted)

    6. **If** some of these cuts generate subdomains with more than one connected component **then**

Compute the refinement indicators corresponding to the subcuts associated to each connected component (this will be the case each time checker board cut is selected!).

Update the set of selected cuts according to the 80% rule.

Endif

7. **Else if** Two or more selected cuts divide the same part Z_i **then**

   Compute the corresponding coarsening indicators.

   Aggregate the subdomains where coarsening indicators allow it.

   Endif

8. **If** there is at most one selected cut in each zone **then**

   Minimize J successively with all parameterizations associated to all selected cuts.

   Keep only the cut which induces the largest decrease of J

   Endif

9. Update the current zonation according to the selected refinement or coarsening.

10. **Enddo**

2. CONVERGENCE ANALYSIS

To prove that the proposed algorithm indeed yields a good approximation to the parameter estimation problem of identifying the transmissivity T in the parabolic differential equation (1) we have to consider mainly three aspects:

- The computed approximation should, in the case of exact data, converge to the exact solution of the infinite-dimensional identification problem as the mesh size of the underlying computing grid goes to zero.
- For any fixed mesh size, the discrete problem should be stable with respect to data perturbations. This is important in view of the fact that the original infinite-dimensional problem is unstable in the sense that even small noise in the data can lead to large deviations in the searched for parameter. So we want to show that here regularization is done by discretization.
- Due to the mentioned instability of the infinite-dimensional identification problem, noise in the data will, in the worst case, be amplified by a factor that grows with progressive refinement. Therefore, the refinement level has to be chosen in an optimal way according to a trade-off between approximation error and propagated data noise, such that convergence is obtained as the noise level goes to zero.

For this purpose we will introduce an auxiliary finite-dimensional problem, namely the one that would result from a discretization of the transmissivity using one unknown for each grid cell of the underlying computing mesh $\mathcal{T}_h$. We show, on one hand, a stability estimate (see Lemma 1), that later on will be useful to prove convergence and analyze stability. On the other hand (see Lemmas 2 and 3) we prove that the adaptive discretization algorithm described in Section 2, working with a possibly small number of zones instead of a computationally expensive fine discretization based on the underlying computing mesh $\mathcal{T}_h$, indeed arrives at or sufficiently close to the solution of the auxiliary finite-dimensional problem. Altogether (see Theorems 1, 2) we can conclude convergence of our algorithm in the exact data case as the mesh size of the computing grid $\mathcal{T}_h$ goes to zero. In the presence of data noise $\delta > 0$ we can give an a priori discretization level choice leading to convergence as $\delta \to 0$.

To get more concrete, we first of all write down the discrete forward problem we are working with in our computation. Time discretization is done by a backwards Euler scheme, we denote by a superscript n the respective time step, $n = 1\ldots, \frac{t}{\Delta t}$, where t is the final time. A mixed hybrid finite element space discretization of (1) (cf. e.g. [Benameur, 1999], [Chavent-Roberts, 1997]) can be derived by introducing a variable Ψ for $T\mathrm{grad}\phi$, and using the Raviart-Thomas finite elements of lowest order as ansatz and test functions. This results in the system

$$\begin{aligned}
\forall q \in M_h \ &: \ \sum_{K\in\mathcal{T}_h}\left(\int_K S_K \frac{{\Phi_h}^{(n)} - {\Phi_h}^{(n-1)}}{\Delta t}\, q + \int_K \mathrm{div}\Psi_h^{(n)} q\right) && = \sum_{K\in\mathcal{T}_h}\int_K Q_h^{(n)} q,\\
\forall v \in X_h \ &: \ \sum_{K\in\mathcal{T}_h}\left(\int_K \frac{1}{T_K}\Psi_h^{(n)} v - \int_K \Phi_h^{(n)}\mathrm{div} v + \sum_{E\in\partial K}\int_E \sigma_h^{(n)}\ v_{K,E}\ \nu_{K,E}\right) && = 0,\\
\forall \mu \in N_h \ &: \ \sum_{E\in\mathring{\mathcal{A}}}\int_E (\Psi_{K_1}^{(n)}.\nu_{K_1,E} + \Psi_{K_2}^{(n)}.\nu_{K_2,E})\mu && = 0,\\
\forall E \subset \Gamma_D \ &: \ \sigma_h^{(n)}|_E && = \Phi_{d,E}\ ,\\
\forall E \subset \Gamma_N \ &: \ \int_E \Psi_h^{(n)}.\nu && = q_{D,E},
\end{aligned} \tag{13}$$

for all $n \in \{1, \ldots, \frac{t}{\Delta t}\}$,

$$ {\Phi_h}^{(0)} = \Phi_{0,h} \tag{14}$$

Here $\mathcal{T}_h$ is the computing mesh consisting of rectangular cells K, $E \in \mathcal{A}(\mathring{\mathcal{A}})$ are the cell's edges (in Ω), $\nu_{K,E}$ the unit outer normal to K on E, Φ_h is supposed to be an element of M_h, the space of piecewise (in each cell K) constant functions, σ_h is the restriction of Φ to the edges and contained in N_h, the space of piecewise constant functions on the edges, and Ψ_h is in $X_h = \bigcup_{K\in\mathcal{T}_h} \mathrm{P}_1 \times \mathrm{P}_1$, which means that on each cell the fist component of Ψ_h is linear with respect to the first variable and the second component is linear with respect to the second variable. T_h, the discrete version of the transmissivity, that is used here, is assumed to be piecewise constant with respect to $\mathcal{T}_h$ (T_K denotes its value on the cell K) and bounded away from zero, so that the direct problem of computing $(\Phi_h^{(n)}, \Psi_h^{(n)}, \sigma_h^{(n)})$ from (13) is well-posed in each time step. Conversely one can consider the problem of reconstructing the cell wise constant function T_h in M_h from measurements of $\Phi_h^{(n)}$ in each cell and at each time step. This will be just what we consider our auxiliary identification problem, for which the following lemma gives a stability estimate:

Lemma 1 *Assume that $\Phi_h^{(n)}$ bounded away from zero is given for each cell K at each time step n and that $\|T_h\|_{L^\infty(\Omega)} \leq C$ with some constant C independent of h. Then $T_h \in M_h$ is uniquely determined by (13) and the assumption that it is known on the Dirichlet boundary edges. Moreover, for any two combinations $(T_{i,h}, \Phi_{i,h}^{(n)})$, $i = 1, 2$, satisfying (13) with $(T_h, \Phi_h^{(n)}, \sigma_h^{(n)}, \Psi_h^{(n)})$ replaced by $(T_{i,h}, \Phi_{i,h}^{(n)}, \sigma_{i,h}^{(n)}, \Psi_{i,h}^{(n)})$, and $\Phi_{i,h}^{(n)}$ bounded away from zero by mesh size independent constants, we have the stability estimate*

$$\|T_{1,h} - T_{2,h}\|_{L^2(\Omega)} \leq C \frac{1}{\Delta t} \|\Phi_{1,h} - \Phi_{2,h}\|_{L^2((0,t)\times\Omega)} \tag{15}$$

for some constant C independent of h and Δt.

Note that for arbitrary $\Phi_h^{(n)} \in M_h$ the solution T_h of (13) might depend on time, but here we are only interested in time independent solutions T_h of (13).
We mention that we do not let the time step Δt tend to zero, so that $C\frac{1}{\Delta t}$ on the right hand side of (15) can be considered as a constant. Nevertheless we keep track of Δt in our estimates, since it is likely to be small, so that it might be of interest with which exponent it appears.

Proof. We reformulate (13) equivalently as

$$\begin{aligned}
\forall q \in M_h \ &: \ \sum_{K\in\mathcal{T}_h} \int_K \operatorname{div}\Psi_h^{(n)} q &&= \sum_{K\in\mathcal{T}_h} \int_K \left(Q_h^{(n)} - S_K \frac{{\Phi_h}^{(n)} - {\Phi_h}^{(n-1)}}{\Delta t} \right) q, \\
\forall v \in X_h \ &: \ \sum_{K\in\mathcal{T}_h} \left(\int_K \Psi_h^{(n)} v - \int_K \tilde{T}_h^{(n)} \operatorname{div} v + \sum_{E\in\partial K} \int_E \tilde{\sigma}_h^{(n)} \, v_{K,E} \, \nu_{K,E} \right) &&= 0, \\
\forall \mu \in N_h \ &: \ \sum_{E\in\mathring{\mathcal{A}}} \int_E (\Psi_{K_1}^{(n)}.\nu_{K_1,E} + \Psi_{K_2}^{(n)}.\nu_{K_2,E})\mu &&= 0, \\
E \subset \Gamma_D \ &: \ \tilde{\sigma}_h^{(n)}|_E &&= T_E \cdot \Phi_{d,E},
\end{aligned} \tag{16}$$

with, in addition

$$E \subset \Gamma_N \ : \ \int_E \Psi^{(n)}.\nu = q_{D,E},$$

for all $n \in \{1, \ldots, \frac{t}{\Delta t}\}$, and (14), where we have introduced new variables $\tilde{T}_h^{(n)} := T_h \cdot \Phi_h^{(n)}$, $\tilde{\sigma}_h^{(n)} := T_h \cdot \sigma_h^{(n)}$. Assuming that T_h is known on the boundary of Ω, we can see (16) as an equation satisfied by $(\tilde{T}_h^{(n)}, \tilde{\sigma}_h^{(n)}, \Psi_h^{(n)})$ at each time step n, with $\Phi_h^{(n)}$ as data. This system (16) together with

$$T_h = \frac{\tilde{T}_h^{(n)}}{\Phi_h^{(n)}} \ ,$$

is equivalent to our auxiliary identification problem. Now taking a closer look at the left hand sides of (16) we see that on each time step, $(\tilde{T}_h^{(n)}, \tilde{\sigma}_h^{(n)}, \Psi^{(n)})$ just solves the mixed hybrid discretization of the Laplace equation with right hand side determined by the data. The uniqueness and stability assertions in Lemma 1 are therefore immediate consequences of well-known results on well-posedness (with mesh-independent constants) of the discretization of the Laplace equation by Raviart-Thomas mixed finite elements (cf. e.g., Subsections IV.1 and V.1 in [Brezzi-Fortin, 1991]). Namely they imply

$$\|\tilde{T}_{1,h}^{(n)} - \tilde{T}_{2,h}^{(n)}\|_{L^2(\Omega)} \leq C\|\frac{\Phi_{1,h}^{(n)} - \Phi_{1,h}^{(n-1)}}{\Delta t} - \frac{\Phi_{2,h}^{(n)} - \Phi_{2,h}^{(n-1)}}{\Delta t}\|_{L^2(\Omega)} ,$$

(where the subscripts $_1$ and $_2$ have to be understood as introduced in the lemma and we keep the notation C for a general constant with possible changing value), which together with the simple estimate

$$\begin{aligned} \|T_{1,h} - T_{2,h}\|_{L^2(\Omega)} &= \|\frac{1}{\Phi_{1,h}^{(n)}}(\tilde{T}_{1,h}^{(n)} - \tilde{T}_{2,h}^{(n)}) + T_{2,h}(\frac{\Phi_{2,h}^{(n)}}{\Phi_{1,h}^{(n)}} - 1)\|_{L^2(\Omega)} \\ &\leq C(\|\tilde{T}_{1,h}^{(n)} - \tilde{T}_{2,h}^{(n)}\|_{L^2(\Omega)} + \|T_{2,h}\|_{L^\infty(\Omega)}\|\Phi_{1,h}^{(n)} - \Phi_{2,h}^{(n)}\|_{L^2(\Omega)}) \end{aligned}$$

yields

$$\begin{aligned} \|T_{1,h} - T_{2,h}\|_{L^2(\Omega)} \leq C \Bigg(&\|\frac{\Phi_{1,h}^{(n)} - \Phi_{1,h}^{(n-1)}}{\Delta t} - \frac{\Phi_{2,h}^{(n)} - \Phi_{2,h}^{(n-1)}}{\Delta t}\|_{L^2(\Omega)} \\ &+ \|T_{2,h}\|_{L^\infty(\Omega)}\|\Phi_{1,h}^{(n)} - \Phi_{2,h}^{(n)}\|_{L^2(\Omega)}\Bigg) , \end{aligned} \tag{17}$$

for all $n \in \{1, \ldots, \frac{t}{\Delta t}\}$, from which we can conclude, by taking the sum over all time steps n on both sides side in (17), and by Cauchy-Schwarz inequality

$$\frac{t}{\Delta t}\|T_{1,h} - T_{2,h}\|_{L^2(\Omega)} \leq \frac{C}{\Delta t}\sum_{n=1}^{t/\Delta t}\|\Phi_{1,h}^{(n)} - \Phi_{2,h}^{(n)}\|_{L^2(\Omega)} \leq C\frac{t^{\frac{1}{2}}}{(\Delta t)^2}\|\Phi_{1,h} - \Phi_{2,h}\|_{L^2((0,t)\times\Omega)} .$$

□

As we will show in more detail in the proof of our main results Theorems 1, 2, Lemma 1 will be crucial for showing stability of the auxiliary identification problem as well as convergence of its solution towards a solution of the original infinite dimensional parameter estimation problem. Before doing so we have to answer the question, whether, for a fixed finite computing mesh, the approximation produced by the proposed adaptive discretization algorithm has something to do with this auxiliary finite dimensional problem. This is done in the next two lemmas.

First of all, note that the algorithm described in Subsection 1.3 has to stop after a finite number of refinement steps, since we deal with a finite number of overall degrees of freedom, bounded by the number of cells in the computing mesh $\mathcal{T}_h$, and the objective function is supposed to strictly decrease when going from one step of the adaptive discretization algorithm to the next. By the definition of the refinement indicators, stopping refinement is characterized by vanishing refinement indicators on one hand, which in turn, by equation (5) implies that the derivatives of the objective function J with respect to each possible cut of each zone in the current zonation is equal to zero. On the other hand, since we are at a minimizer with respect to the current zonation, also the derivative of J with respect to each zone vanishes. Now Lemma 2 says that this allows us to conclude that we are at a critical point. Here in order to deal with a computationally practicable situation we replace being "equal to zero" by being "small".

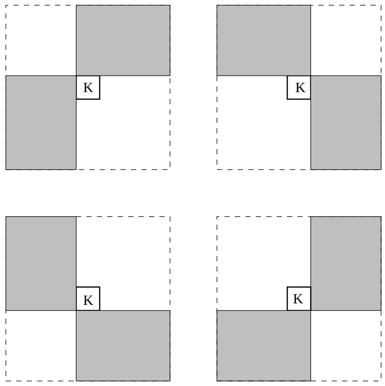


Figure 4. Four checkerboard cuts whose superposition gives $2 \times (Z_i \setminus K)$

Lemma 2 *Let T_h^* denote the discrete solution produced by the algorithm described in Subsection 1.3 with the stopping criterion*

$$|\frac{\partial J}{\partial T_i}(T_h^*)| \leq \epsilon_{1,h} \tag{18}$$

for each zone Z_i in the current zonation and

$$|\lambda| \leq \epsilon_{2,h} \tag{19}$$

for the refinement indicators λ corresponding to each cut. Then we have

$$\forall K \in \mathcal{T}_h \ : \quad |\frac{\partial J}{\partial T_K}(T_h^*)| \leq \epsilon_h \tag{20}$$

where $\epsilon_h \leq C(\epsilon_{1,h} + \epsilon_{2,h})$ with some constant C independent of h and Δt.

Proof. The proof is based on the fact that each cell K of $\mathcal{T}_h$ in a fixed zone Z_i can be represented by a linear combination of cuts. More precisely, choosing the four checkerboard cuts corresponding to the four corners of K and not containing K (see Fig. 4) and considering their superposition, we get two times the whole zone Z_i minus the element K. So we only have to weight this superposition of checkerboard cuts with $\frac{1}{2}$ and subtract it from the whole zone to obtain K. Now (19), (5) and the fact that the norm of the matrix A^T is bounded by a constant C independent of h, implie

$$|\frac{\partial J}{\partial T_{i,cb}}(T_h^*)| \leq C\epsilon_{2,h}$$

for each $T_{i,cb}$ resulting from some checkerboard cut of T_i. Therfore and by (18), in sum we get

$$|\frac{\partial J}{\partial T_K}(T_h^*)| \leq \epsilon_{1,h} + \frac{C}{2}(\epsilon_{2,h} + \epsilon_{2,h} + \epsilon_{2,h} + \epsilon_{2,h}) =: \epsilon_h \ .$$

□

Note that, as opposed to Lemmas 1 and 3, Lemma 2 is a very general assertion on the adaptive discretization strategy of Section 2, that retains its validity and its own interest also in applications to completely different parameter identification problems.

Finally, in order to conclude from (20) that we are at or close to a solution of the auxiliary estimation problem formulated with one unknown per grid cell, we make use of the following Lemma 3, which states that a critical point of the objective function will even be a global minimizer. Moreover, the fact that the data appearing in the objective function is close to the (discrete) output $\Phi_h^\dagger$ of the projection onto M_h of some exact transmissivity $T_h^\dagger =: Proj_{M_h} T^\dagger$ comes into play. Namely, we assume that

$$\|\Phi_h^{obs} - \Phi_h^\dagger\|_{L^2((0,t)\times\Omega)} \leq \delta \tag{21}$$

where δ is the noise level.

Lemma 3 *Let T_h^* satisfy (20) and denote by Φ_h^* the output corresponding to T_h^* (via (13)) with data satisfying (21). Assume, moreover, that T_h^* is close to $T_h^\dagger$:*

$$\|T_h^* - T_h^\dagger\|_{L^\infty(\Omega)} \leq c \tag{22}$$

for some constant c sufficiently small but independent of h and Δt. Then we have, with some constant C independent of h and Δt,

$$\|\Phi_h^* - \Phi_h^\dagger\|_{L^2((0,t)\times\Omega)} \leq C(\delta + \frac{\epsilon_h}{\Delta t})$$

Proof. Denote by F_h the discrete "parameter to solution map", that maps a cell-wise constant transmissivity to the Φ_h part of a solution to (13), (14). Then we can rewrite the value of the objective function at T_h as

$$J(T_h) = \frac{1}{2}\|F_h(T_h^*) - \Phi_h^{obs}\|^2_{L^2((0,t)\times\Omega)}$$

and its derivative at T_h^* into some direction $dT \in M_h$ as

$$J'(T_h^*)[dT] = \langle F_h(T_h^*) - \Phi_h^{obs}, F_h'(T_h^*)[dT]\rangle_{L^2((0,t)\times\Omega)}$$

The difference between $F_h(T_h^*)$ and $F_h(T_h^\dagger)$ can be expressed in terms of a solution to a mixed hybrid discrete system, namely $F_h(T_h^*) - F_h(T_h^\dagger) = d\Phi_1$ where $(d\Phi_1, d\sigma_1, d\Psi_1) \in (M_h \times N_h \times X_h)^{t/\Delta t}$ solves

$$\begin{aligned}
\forall q \in M_h \ : \quad & \sum_{K\in\mathcal{T}_h}\left(\int_K S_K \frac{d\Phi^{(n)} - d\Phi^{(n-1)}}{\Delta t}\, q + \int_K \mathrm{div} d\Psi^{(n)}\right) q && = 0,\\
\forall v \in X_h \ : \quad & \sum_{K\in\mathcal{T}_h}\left(\int_K \frac{1}{T_K^*} d\Psi^{(n)} v - \int_K d\Phi^{(n)}\mathrm{div} v + \sum_{E\in\partial K}\int_E d\sigma^{(n)}\ v_{K,E}\ \nu_{K,E}\right) &&\\
& \qquad = \sum_{K\in\mathcal{T}_h}\left(\int_K f^{(n)}\mathrm{div} v + \sum_{E\in\partial K}\int_E g^{(n)}\ v_{K,E}\ \nu_{K,E}\right), &&\\
\forall \mu \in N_h \ : \quad & \sum_{E\in\overset{\circ}{\mathcal{A}}}\int_E (d\Psi^{(n)}_{K_1}.\nu_{K_1,E} + d\Psi^{(n)}_{K_2}.\nu_{K_2,E})\mu && = 0,\\
\forall E \subset \Gamma_D \ : \quad & d\sigma^{(n)}|_E && = 0,\\
\forall E \subset \Gamma_N \ : \quad & \int_E d\Psi^{(n)}.\nu && = 0,
\end{aligned} \tag{23}$$

for all $n \in \{1, \ldots \frac{t}{\Delta t}\}$ and

$$d\Phi^{(0)} = 0, \tag{24}$$

with

$$f^{(n)} := \frac{T_h^\dagger - T_h^*}{T_h^*} \, \Phi_h^{(n)}(T_h^\dagger) \ , \quad g^{(n)} := -\frac{T_h^\dagger - T_h^*}{T_h^*} \, \sigma_h^{(n)}(T_h^\dagger) \ .$$

Here $\Phi_h(T_h)$, $\sigma_h(T_h)$ denote the Φ- and σ- components of the solution $(\Phi_h, \sigma_h, \Psi_h) \in (M_h \times N_h \times X_h)^{t/\Delta t}$ of (13), (14) with some discrete transmissivity T_h.
The formal derivative $F_h'(T_h^*)$ of the parameter to solution map into direction dT is the primal part $d\Phi_2$ of a solution $(d\Phi_2, d\sigma_2, d\Psi_2)$ to (23), (24) with

$$f^{(n)} := \frac{dT}{T_h^*} \, \Phi_h^{(n)}(T_h^*) \ , \quad g^{(n)} := -\frac{dT}{T_h^*} \, \sigma_h^{(n)}(T_h^*)$$

Indeed, the first order Taylor remainder $F_h(T_h^* + dT) - F_h(T_h^*) - F'(T_h^*)[dT]$ (along with apropriate dual solutions) satisfies (23), (24) with $f^{(n)} := \frac{dT}{T_h^*} \, (\Phi_h^{(n)}(T_h^* + dT) - \Phi_h^{(n)}(T_h^*))$, $g^{(n)} := \frac{dT}{T_h^*} \, (\sigma_h^{(n)}(T_h^* + dT) - \sigma_h^{(n)}(T_h^*))$ so that using stability estimates for solutions of (23), (24), as well as the fact that on the finite dimensional spaces under consideration all norms are equivalent (with constants possibly depending on h), it can be easily checked that this is really the Fréchet derivative. More precisely

$$\begin{aligned} &\|F_h(T_h^* + dT) - F_h(T_h^*) - F_h'(T_h^*)[dT]\|_{L^2((0,t)\times\Omega)} \\ &\quad \leq \bar{C}\|dT\|_{L^\infty(\Omega)}\|F_h(T_h^* + dT) - F_h(T_h^*)\|_{L^2((0,t)\times\Omega)} \end{aligned} \tag{25}$$

with some constant $\bar{C}$ independent of h.
Therefore, by (22) and (21) we get, with the special choice $dT := T_h^\dagger - T_h^*$ and $\bar{C}$ as in (25)

$$\begin{aligned} &|J'(T_h^*)[dT]| \\ &\geq \langle F_h(T_h^*) - \Phi_h^{obs}, F_h(T_h^*) - F_h(T_h^\dagger)\rangle_{L^2((0,t)\times\Omega)} \\ &\quad -\bar{C}\|T_h^\dagger - T_h^*\|_{L^\infty}\|F_h(T_h^*) - F_h(T_h^\dagger)\|_{L^2((0,t)\times\Omega)}\|F_h(T_h^*) - \Phi_h^{obs}\|_{L^2((0,t)\times\Omega)} \\ &\geq \|F_h(T_h^*) - F_h(T_h^\dagger)\|_{L^2((0,t)\times\Omega)} \Big((1 - \bar{C}c)\|F_h(T_h^*) - F_h(T_h^\dagger)\|_{L^2((0,t)\times\Omega)} - (1 + \bar{C}c)\delta \Big) \ . \end{aligned} \tag{26}$$

On the other hand (20) implies

$$|J'(T_h^*)[dT]| = |\sum_{K\in\mathcal{T}_h} h^2 \frac{\partial J}{\partial T_K}(T_h^*) dT_K| \leq \epsilon_h \|dT\|_{L^1(\Omega)} \leq C\epsilon_h\|dT\|_{L^2(\Omega)} \tag{27}$$

and from Lemma 1 we can conclude for the choice of dT we made:

$$\|dT\|_{L^2(\Omega)} = \|T_h^* - T_h^\dagger\|_{L^2(\Omega)} \leq C\frac{1}{\Delta t}\|F_h(T_h^*) - F_h(T_h^\dagger)\|_{L^2((0,t)\times\Omega)} \tag{28}$$

Collecting estimates (26), (27), (28) yields the assertion of Lemma 3. □

Remark 1 *Note that the assertion of Lemma 3 remains valid if instead of assuming (22), i.e., closeness in L^∞, we consider only measurements at one time instance $t = \Delta t$, i.e., the double sum in the definition of J is reduced to a simple sum (over the cells). In this situation we can choose $dT := (T_h^\dagger - T_h^*)\frac{\Phi_h^1(T_h^\dagger)}{\Phi_h^1(T_h^*)}$ to even obtain $F_h'(T_h^*)[dT] = F_h(T_h^*) - F_h(T_h^\dagger)$ and hence a similar estimate to (26).*

By (22) and (25), the first order Taylor remainder can be estimated by differences of function values multiplied by a small constant independent of h. This corresponds to a so-called tangential cone condition for the infinite-dimensional problem, i.e., a condition on the nonlinearity of the forward operator F. Such a condition is often used in convergence analysis of regularization methods for general nonlinear ill-posed problems (see e.g. [Hanke-Neubauer-Scherzer, 1995], [Kaltenbacher-Neubauer-Scherzer]). An alternative assumption on the nonlinear forward operator F that is used in convergence proofs, is a condition on the invariance of the range of F'. This is here satisfied in the situation described in Remark 1 (or in the steady state case $S \equiv 0$), without any restrictions on closeness of T_h^* to $T_h^\dagger$.

Remark 2 *An assumption, we have more or less explicitely made in the proofs above, was uniform boundedness of T_h^*. This is not only necessary for the estimate in Lemma 1 but also for the well-posedness of the system (23), (24). This L^∞- boundedness can be guaranteed by introducing box constraints, i.e. considering instead of the unconstrained optimization problem $J = \min\limits_{T_h}!$, the minimization of the objective function J subject to a constraint*

$$0 < \underline{T} \leq T_h \leq \overline{T} \ , \tag{29}$$

where the upper and lower bounds $\overline{T}, \underline{T}$ may vary over the domain and are chosen such that the exact solution $T_h^\dagger$ satisfies these constraints with strict inequalities; moreover, using a priori information on the exact solution, we might be able to choose these bounds such that all transmissivities satisfying (29) automatically fulfill condition (22) of Lemma 3. Incorporating box constraints into the algorithm described in the foregoing section and in the proofs of this section is quite straightforward, nevertheless we shortly mention, for the sake of exactness, the points where modifications appear:

- *For the re-definition of the refinement indicators we go out from a discrete transmissivity T_h^* that, along with the Lagrange multipliers $\overline{\Lambda}^*, \underline{\Lambda}^* \in \mathbb{R}^{n_{zon}}$ corresponding to the upper and lower box constraint, respectively, solves $J = \min\limits_{T_h}!$, subject to (29), for some fixed zonation consisting of n_{zon} zones. The refinement indicators λ^* have to be defined replacing (5) by*

$$\nabla J(T_h^*) + \overline{\Lambda}^* - \underline{\Lambda}^* - A^T\lambda^* = 0$$

- *Similarly, the criterion (18) for stopping the minimization process of finding T_h^* has to be replaced by*

$$|\frac{\partial J}{\partial T_i}(T_h^*) + \overline{\Lambda}^* - \underline{\Lambda}^*| \leq \epsilon_{1,h}$$

- *Like in the unconstrained case (Lemma 2) we can then conclude*

$$\forall K \in \mathcal{T}_h \ : \quad |\frac{\partial J}{\partial T_K}(T_h^*) + \overline{\Lambda}^* - \underline{\Lambda}^*| \leq \epsilon_h \ .$$

Consequently, T_h^* together with the natural cell wise extension of the zone wise given Lagrange multipliers $\overline{\Lambda}^*, \underline{\Lambda}^*$, satisfies (approximately) the first order Kuhn-Tucker *conditions and is therefore (close to) a critical point of the constrained uniformly discretized minimization problem. This means that for all those increments dT, that point into the interior of the feasible set at cells where one of the box constraints is active, but are arbitrary elsewhere, we have (cf. (27))*

$$-J'(T_h^*)[dT] \leq -\epsilon_h \|dT\|_{L^1(\Omega)}$$

The special choice of $dT = T_h^\dagger - T_h^*$ (which is an admissible increment in the constrained setting, since $T_h^\dagger$ lies in the interior of the feasible set) gives, just as described in Lemma 3, $-J'(T_h^*)[dT] \geq c\|F_h(T_h^*) - F_h(T_h^\dagger)\|^2_{L^2(\Omega)} - C\delta$. The rest is obvious.

To obtain our main convergence results, we need, in addition to Lemmas 1–3, an approximation property of the space M_h. These spaces are known (cf., e.g., [Ciarlet, 1978], Theorem 3.1.4) to satisfy

$$\|u - Proj_{M_h} u\|_{L^2(\Omega)} \leq Ch \text{ and } \|u - Proj_{M_h} u\|_{L^\infty(\Omega)} \stackrel{h\to 0}{\longrightarrow} 0 \quad \forall u \in L^\infty(\Omega) \cap H^1(\Omega)\ . \tag{30}$$

Since we want to admit also the practically relevant situation of transmissivities with jumps, we make use of the following version

$$\begin{aligned} \|T - Proj_{M_h} T\|_{L^2(\Omega)} \leq C\sqrt{h} \text{ and } \|T - Proj_{M_h} T\|_{L^\infty(\Omega)} \stackrel{h\to 0}{\longrightarrow} 0 \\ \forall T \in L^\infty(\Omega), T \text{ piecewise } H^1\ . \end{aligned} \tag{31}$$

This can be shown analogously to (30) when taking into account the fact that for only piecewise H^1-functions T, the set Σ_h of elements K containing jumps of T has cardinality $O(h^{-1})$, so that

$$\|T - Proj_{M_h} T\|^2_{L^2(\Omega)} = \underbrace{\sum_{K \in \mathcal{T}_h \setminus \Sigma_h} \|T - Proj_{M_h} T\|^2_{L^2(K)}}_{\leq Ch^2 \text{ by } (30)} + \underbrace{\sum_{K \in \Sigma_h} \|T - Proj_{M_h} T\|^2_{L^2(K)}}_{\leq Ch^{-1}\|T\|_{L^\infty(\Omega)} h^2}\ .$$

Moreover, the inverse inequality

$$\|u_h\|_{L^\infty(\Omega)} \leq Ch^{-1} \|u_h\|_{L^2(\Omega)} \quad \forall u_h \in M_h \tag{32}$$

(cf., e.g., [Ciarlet, 1978], Theorem 3.2.6) will be used.

First we treat the noise free case

$$\Phi^{obs} = \Phi^\dagger\ . \tag{33}$$

Theorem 1 *Assume that $T^\dagger \in L^\infty(\Omega)$ piecewise $H^1(\Omega)$ and that $\Phi^\dagger$ is uniformly bounded away from zero in $(0,t) \times \Omega$. Let T_h^*, the output of the algorithm described in Subsection 1.3, be uniformly bounded (cf. Remark 2) and satisfy the closeness assumption (22) of Lemma 3.*
In the noise free case (33) with the choice

$$\epsilon_{1,h} = o(h)\ , \quad \epsilon_{2,h} = o(h) \tag{34}$$

in (18) and (19), one obtains

$$\|T_h^* - T^\dagger\|_{L^2(\Omega)} = O(\sqrt{h}) \tag{35}$$

and

$$\|T_h^* - T^\dagger\|_{L^\infty(\Omega)} \to 0 \text{ as } h \to 0\ . \tag{36}$$

Next we consider the situation of noisy data with

$$\|\Phi_h^{obs} - \Phi_h^\dagger\|_{L^2((0,t)\times\Omega)} \leq \delta \quad \forall h > 0 \tag{37}$$

and an appropriately chosen mesh size.

Theorem 2 *Let the assumptions of Theorem 1 be satisfied.*
If the data are contaminated with noise of noise level δ in (37), then with the choice

$$\epsilon_{1,h} \sim \Delta t \ \delta \ , \quad \epsilon_{2,h} \sim \Delta t \ \delta \tag{38}$$

in (18) and (19), one gets with

$$h \sim \delta^2 \tag{39}$$

that

$$\|T_h^* - T^\dagger\|_{L^2(\Omega)} = O(\delta) \tag{40}$$

and with

$$h \to 0 \quad \frac{\delta}{h} \to 0 \ \text{as} \ \delta \to 0$$

that

$$\|T_h^* - T^\dagger\|_{L^\infty(\Omega)} \to 0 \ \text{as} \ \delta \to 0 \ . \tag{41}$$

Proof. (Theorem 2)
If (37) holds, then by Lemmas 2 and 3

$$\|\Phi_h^* - \Phi_h^\dagger\|_{L^2((0,t)\times\Omega)} \leq C(\delta + \frac{\epsilon_h}{\Delta t})$$

so that by Lemma 1 and estimate (31) we get

$$\|T_h^* - T^\dagger\|_{L^2(\Omega)} \leq \|T_h^* - T_h^\dagger\|_{L^2(\Omega)} + \|T_h^\dagger - T^\dagger\|_{L^2(\Omega)} \leq C\left(\frac{1}{\Delta t}(\delta + \frac{\epsilon_h}{\Delta t}) + \sqrt{h}\right)$$

and with (32)

$$\|T_h^* - T^\dagger\|_{L^\infty(\Omega)} \leq Ch^{-1}\|T_h^* - T_h^\dagger\|_{L^2(\Omega)} + \|T_h^\dagger - T^\dagger\|_{L^\infty(\Omega)} \leq C\left(\frac{1}{\Delta t}h^{-1}(\delta + \frac{\epsilon_h}{\Delta t}) + o_h(1)\right)$$

which by the choice (38), (39) gives (40), (41). □

Proof. (Theorem 1)
The proof follows from the estimates in the proof of Theorem 2 by just setting $\delta := 0$. □

By Remark 1, the assertion of Theorems 1 and 2 remain valid if instead of the closeness condition (22) we assume that measurements at only one time step are taken into account.

Note that from Theorems 1, 2 one can immediately deduce respective convergence estimates for intermediate norms $L^p(\Omega)$, $2 < p < \infty$ of the error by using interpolation, i.e., just the Hölder inequality.

Remark 3 *The estimate (40) of the L^2-error in terms of the noise level is more typical for well-posed than for ill-posed problems, where one usually has a slower rate than $O(\delta)$ or even no convergence rate. Indeed, it is a simple consequence of Lemma 1 (with $h \to 0$) that the transmissivity estimation problem is well-posed with respect to the $L^2(\Omega)$ topology in parameter space, and the $L^2((0,t) \times \Omega)$ topology in data space.*

3. NUMERICAL EXPERIMENTS

We consider now a case of three zones of transmissivity as in fig. 5, and we consider that $\mathcal{T}_h$ is composed of 36 cells. We start with a constant transmissivity (fig. 5) and we follow the steps of the algorithm described in section 1.3. We show in figures 6, 7, 8, 9, 10 the zonations obtained at each iteration and the transmissivity calculated on these zonations. In this case one notice, observing fig. 10, that only a close approximation of the exact zonation and transmissivity is recovered. The value of the objective function corresponding to transmissivity computed on this zonation is small (10^{-6}). We continue according to the algorithm and we compte refinement indicators corresponding to this zonation, their values are of the order of 10^{-2}. Using the first order approximation: $\lambda = \dfrac{\Delta J}{\Delta T}$, we conclude that with such refinement the variation of the tranmissivity will be of 10^{-4}. Knowing the order of values of transmissivity in Ω, such variation can be neglected, this is then a stopping rule for our refinement process.

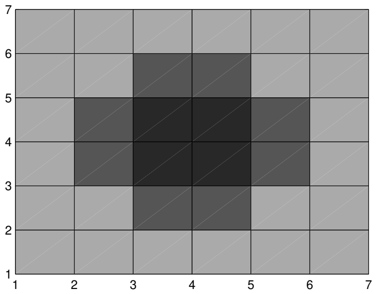


Figure 5. Exact transmissivity: unknown of the inverse problem

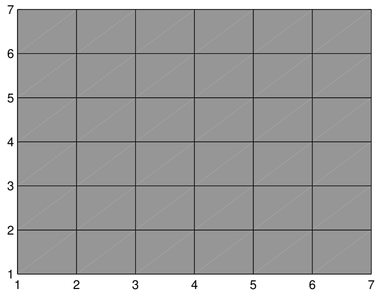


Figure 6. Transmissivity computed as a constant

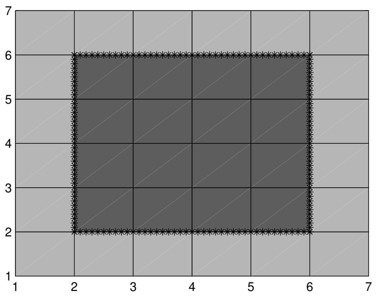


Figure 7. Zonation obtained at the first iteration and the corresponding computed transmissivity

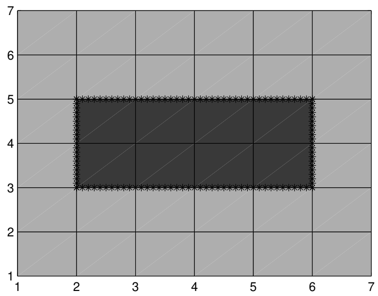


Figure 8. Zonation obtained at the second iteration and the corresponding computed transmissivity

We consider now measured data contaminated with noise of noise level 10% and 20%, using the same $\mathcal{T}_h$ as above the same results are obtained. We refine our computing mesh, now $\mathcal{T}_h$ is composed of 144 cells, and again we consider data contaminated with the same noise

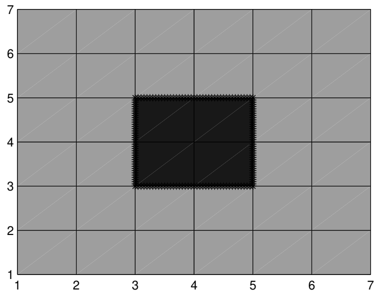


Figure 9. Zonation obtained at the third iteration and the corresponding computed transmissivity

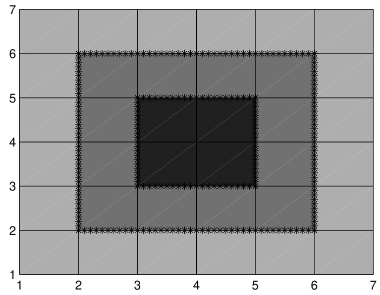


Figure 10. Zonation obtained at the fourth iteration and the corresponding computed transmissivity

levels . As in the 36 cells case, we start with a constant transmissivity and we follow the steps of the algorithm described in section 1.3. At the last iteration, a parameterisation similar to that in fig. 10 is recovered. The number of iterations is the same as in the above case, but at each iteration the set of selected cuts according to the rule "80% of I_{max}" is larger then that obtained in the 36 cells case, and so the number of minimisations to be done at each iteration is also larger, which means that computing cost is higher.
Once again, we refine $\mathcal{T}_h$ into 24×24 grid, and we fellow the steps of algorithm 1.3 to find the parameterisation and the values of the transmissivity in Ω. At each iteration we have to minimize the objective function with a "large" number of cuts corresponding to refinement indicators larger than 80% of I_{max}. In spite of that, for both noise levels, after seven iterations we stopped our adaptive discretization because values of refinement indicators indicate that decreasing the objective function by refining discretization according to algorithm 1.3 is hopeless. At this step, the same parameterization is obtained with the two noise levels, values of transmissivities are not much different, and as we see in fig. 12, we are far away from the exact parameterization.

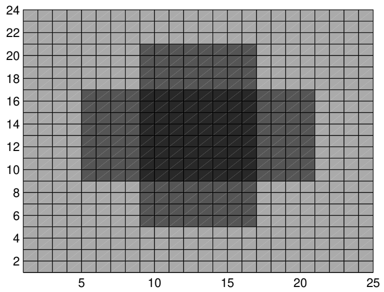


Figure 11. Exact transmissivity: unknown of the inverse problem

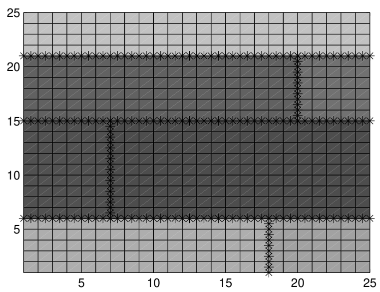


Figure 12. Zonation obtained at the seventh iteration and the corresponding computed transmissivity with noise level=20%

Table 1

	$\delta = 10\%$		$\delta = 20\%$	
	iterations in alg. Subsec. 1.3	minimizations of J	iterations in alg. Subsec. 1.3	minimizations of J
6×6 grid	4	4	4	5
12×12 grid	4	10	4	12
24×24 grid	7	55	7	59

The number of iterations in the adaptive disretization algorithm described in Subsection 1.3 as well as the total number of objective function minimizations for noised data and different grid sizes is displayed in Table 1. We note that the coarser the discretization, the higher the admissible noise level. In the case of the 6×6 grid we still obtained reasonable results with even 50% noise. This illustrates the regularizing effect of coarse discretization.

Here we do not emphasize the numerical verification of convergence as $h \to 0$. (Note that in the test example the jumps in the exact transmissivity are aligned to the grid). This task is thoroughly treaten in the paper [Benameur-Hackl-Mühlhuber, 2001].

4. CONCLUSION

In this paper we do a convergence analysis of the method of adaptive discretization using refinement and coarsening indicators, as proposed in [Benameur-Chavent-Jaffré, 2001]. We consider application of this method to the problem of hydraulic transmissivity identification and prove that it is a regularization in the sense given by [Engl-Hanke-Neubauer, 1996]. In case of exact data it converges to the exact solution of the infinite dimensional identification problem, as the mesh size goes to zero. For noisy data, the mesh size acts as a regularization parameter. This is also verified in numerical experiments.
In a current work [Benameur-Hackl-Mühlhuber, 2001], we are showing that the technique of refinement and coarsening indicators can be used also for other parameter identification problems and for other ways of discretizing the parameter and the state.

[Benameur, 1999] H. Benameur. *Problèmes inverses géométriques en élasticité et pour les écoulements en milieu poureux.* PhD thesis, École Nationale d'Ingénieurs de Tunis and INRIA Rocquencourt, 1999.

[Benameur-Chavent-Jaffré, 2001] H. Benameur, G. Chavent, and J. Jaffré. Refinement and coarsening indicators for adaptive parametrization: application to the estimation of hydraulic transmissivities. submitted.

[Benameur-Hackl-Mühlhuber, 2001] H. Benameur, B. Hackl, W. Mühlhuber. An adaptive refinement strategy for parameter identification problems. in preparation.

[Brezzi-Fortin, 1991] F. Brezzi, M. Fortin. *Mixed and Hybrid Finite Element Methods.* Springer, New York - Berlin - Heidelberg, 1991.

[Bonnans et al., 1997] J.F. Bonnans, J.-Ch. Gilbert, C. Lemaréchal, and C. Sagastizabal. *Optimisation Numérique Aspects théoriques et pratiques.* Springer-Verlag, 1997.

[Chavent-Bissel, 1998] G. Chavent and R. Bissell. Indicator for the refinement of parametrization. In *Proceedings of the International Symposium on Inverse Problems in Engineering Mechanics, Nagano, Japan*, pages 185–190, March 1998.

[Chavent-Liu, 1989] G. Chavent and J. Liu. Multiscale parametrization for the estimation of a diffusion coefficient in elliptic and parabolic problems. In A. El Jai and M. Amouroux, editors, *5th IFAC Symposium on Control of Distributed Parameter Systems, Perpignan, June 26-29, 1989*, pages 315–324. Université de Perpignan, France.

[Chavent-Roberts, 1997] G. Chavent and J. Roberts. A unified physical presentation of mixed, mixed hybrid finite elements and standard finite differences approximations of the determination of velocities in waterflow problems. *Advances in Water Ressources*, 14:326–348, 1997.

[Chardaire-Rivière et al., 1990] C. Chardaire-Rivière, G. Chavent, Jérôme Jaffré, and J. Liu. Multiscale representation for simultaneous estimation of relative permeabilities and capillary pressure, paper SPE 20501. In *65th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, New Orleans, Louisiana, September 23-26, 1990*, pages 303–312. Society of Petroleum Engineers, Richardson, Texas.

[Chardighy et al., 1996] E. Chardigny, P. Siegel, R. Mosé, and Ph. Ackerer. Parameter identification for complex groundwater systems. In A. Aldama and al., editors, *Computational Methods in Water Resources XI, Cancun, Mexico, July 22-26, 1996*, pages 305–312. Computational Mechanics Publications.

[Ciarlet, 1978] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North Holland, Amsterdam - New York - Oxford, 1978.

[Engl-Hanke-Neubauer, 1996] H.W. Engl, M. Hanke, and A.Neubauer. *Regularization of Inverse Problems.* Kluwer, Dortdrecht, 1996.

[Eppstein-Dougherty, 1996] M.J. Eppstein and D.E. Dougherty. Simultaneous estimation of transmissivity values and zonation. *Water Resources Research*, 32:3321–3336, 1996.

[Hanke-Neubauer-Scherzer, 1995] M. Hanke, A. Neubauer, and O. Scherzer. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. *Numerische Mathematik* 72:21–37, 1995.

[Jegou, 1997] S. Jegou. Using Maple for symbolic differentiation to solve inverse problems. *Maple Tech*, 4:32–40, 1997.

[Kaltenbacher-Neubauer-Scherzer] B. Blaschke(-Kaltenbacher), A. Neubauer, and O. Scherzer. On convergence rates for the iteratively regularized Gauß–Newton method. *IMA Journal of Numerical Analysis*, 17:421–436, 1997.

[Kaltenbacher-Schöberl, 2001] B. Kaltenbacher and J. Schöberl. A Saddle Point Variational Formulation for Projection–Regularized Parameter Identification. to appear in Numerische Mathematik, 2001.

[Liu, 1993] J. Liu. A multiresolution method for distributed parameter estimation. *SIAM J. Sci. Comput*, 14:389–405, 1993.

[Sun, 1994] N.Z. Sun. *Inverse problems in groundwater modeling.* Kluwer Academic Publishers, 1994.