Semi–Monotonic Augmented Lagrangians for Optimal Control and Parameter Identification

Dalibor Lukáš and Zdeněk Dostál

Abstract Optimization and inverse problems governed with partial differential equations are often formulated as constrained nonlinear programming problems via the Lagrange formalism. The nonlinearity is treated using the sequential quadratic programming. A numerical solution then hinges on an efficient iterative method for the resulting saddle–point systems. In this paper we apply a semi–monotonic augmented Lagrangians method, recently proposed and analyzed by the second author, for equality and simple–bound constrained quadratic programming subproblems arising from optimal control and parameter identification. Provided multigrid preconditioning of primal and dual space inner products and of the Hessian the algorithm converges at $O(1)$ matrix–vector multiplications. Numerical results are given for applications in image segmentation and 2–dimensional magnetostatics discretized using lowest–order Lagrange finite elements.

1 Introduction

Many engineering problems involves a solution to partial differential equations (PDE), which describe a physical field under consideration, and a design of some parameters that influence data of the PDE so that the solution poses required properties. A typical example is optimal control, where the design parameters control forcing terms in the PDE. Another example is parameter identification, where the design parameters are material coefficients of the PDE operator such that the corresponing solution fits best a given (measured) field. The latter problem is rather close

Dalibor Lukáš
VŠB–Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava–Poruba, Czech Republic,
e-mail: dalibor.lukas@vsb.cz

Zdeněk Dostál
VŠB–Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava–Poruba, Czech Republic,
e-mail: zdenek.dostal@vsb.cz

to an optimal topology design, in which we additionaly require the coefficients to be discrete–valued.

In either case we consider the following optimization problem:

$$\min_{u \in \mathscr{U}_{\text{ad}}} \mathscr{I}(u,y) \tag{1}$$

where $\mathscr{U}_{\text{ad}} := \{u \in \mathscr{U} \,:\, \underline{u} \le u \le \overline{u}\}$ for $\underline{u}, \overline{u} \in \mathscr{U}$, $\underline{u} < \overline{u}$, where $\mathscr{U}$ denotes a Hilbert space of design parameters, $y \in \mathscr{Y}$ with $\mathscr{Y}$ being another Hilbert space of state PDE solutions, and where $\mathscr{I} : \mathscr{U} \times \mathscr{Y} \to \mathbb{R}$ denotes a twice continuously differentiable objective functional. Let us denote by $\mathscr{U}'$ and $\mathscr{Y}'$ the related dual spaces. As we mentioned, the problem (1) is additionaly subjected to a PDE state equation. This reads in case of optimal control as follows:

$$\mathscr{B}(y) = \mathscr{G}(u) \text{ on } \mathscr{Y}', \tag{2}$$

where $\mathscr{B} : \mathscr{Y} \to \mathscr{Y}'$ is an elliptic linear and continuous PDE operator and $\mathscr{G} : \mathscr{U} \to \mathscr{Y}'$ is a linear and continuous PDE right–hand side controlled by the design. In case of parameter identification the state equation takes the following form:

$$\mathscr{B}(u,y) = \mathscr{G} \text{ on } \mathscr{Y}', \tag{3}$$

where $\mathscr{B} : \mathscr{U} \times \mathscr{Y} \to \mathscr{Y}'$ is a bilinear continuous PDE operator, which is for a fixed $u \in \mathscr{U}_{\text{ad}}$ elliptic in y, and $\mathscr{G} \in \mathscr{Y}'$.

In terms of sequential quadratic programming (SQP), the problems (1) s.t. (2) and (1) s.t. (3) can be sequentially approximated by equality and simple–bound constrained quadratic programming (qp) subproblems. Let us now adopt the following notation: $V := \mathscr{U} \times \mathscr{Y}$, $V_{\text{ad}} := \mathscr{U}_{\text{ad}} \times \mathscr{Y}$, $A^i := \mathscr{I}''(u^i,y^i)$, $f^i := -\mathscr{I}'(u^i,y^i)$. Then, the constrained qp–subproblems reads as follows:

$$\min_{v \in V_{\text{ad}}} h^i(v) \quad \text{s.t.} \quad B^i(v) = g^i, \quad i = 0,1,2,\ldots \tag{4}$$

where $v := (u,y)$, $h^i(u,y) := (1/2)\langle A^i(u,y),(u,y)\rangle_V - \langle f^i,(u,y)\rangle_V$ is the quadratic approximation (up to a constant) of $\mathscr{I}(u,y)$ at the last iteration point $(u^i,y^i) \in V_{\text{ad}}$, where $\langle .,.\rangle_V$ denotes the duality pairing, and where $B^i(u,y) = g^i$ is the linearization of the state PDE equation at (u^i,y^i), i.e. $B^i := (\mathscr{G},\mathscr{B})$ and $g^i := \mathscr{G}(u^i) - \mathscr{B}(y^i)$ in case of (2), while $B^i := (\mathscr{B}'_u(u^i,y^i), \mathscr{B}'_y(u^i,y^i))$ and $g^i := \mathscr{G} - \mathscr{B}(u^i,y^i)$ in case of (3). For a detail presentation and analysis of some SQP schemes we refer to [1, 2].

The concern of our paper is an efficient solution to problems (4). We base our exposition on a Semi–Monotonic Augmented Lagrangian method for Bound and Equality constrained qp–problems (SMALBE), which has been recently proposed and analyzed by the second author in [3]. Note that we have recently applied a similar method to the Stokes problem, see [5]. The method relies on uniformly bounded spectra of Hessians A^i, which we can often assure via a geometric multigrid preconditioning. Then, the algorithm is proven to converge at $O(1)$ matrix–vector multiplications provided we have an optimal convergence method for the inner simple–

bound constrained quadratic programming subproblems. Such a method was proposed and analyzed in [4], it is based on conjugate gradients and we call it Modified Proportioning with Reduced Gradient Projections (MPRGP). The rest of the paper is organized as follows: In Section 2 we describe the algorithms SMALBE and MPRGP preconditioned with multigrid and we refer to the main theoretical result on the optimal convergence. Finally, In Section 3 we present two benchmarks, namely an optimal control for image segmentation and a parameter identification for 2–dimensional magnetostatics, and give numerical results in terms of convergence.

2 The Algorithm

Let us consider the problem (4) and from now on skip the index i. Denote by $Q := \mathscr{Y}$ the space of Lagrange multipliers. Let I_V and I_Q denote the inner product (Riesz isomorphism) operators on the Hilbert spaces V and Q, respectively, let $g \in \mathrm{Range}(B)$, $\mathrm{Range}(B)$ be closed, and let $V_{\mathrm{BE}} := \{v \in V_{\mathrm{ad}} \;:\; Bv = g\}$ be nonempty. Denote by $B^T : Q \to V'$ the adjoint operator to B. We assume that there exists $\underline{\rho} > 0$ such that the operator $A + \underline{\rho} B^T I_Q^{-1} B$ is elliptic with the constant $\underline{\lambda} > 0$. For arbitrary $v \in V_{\mathrm{ad}}$, $q \in Q$ and $\rho \geq \underline{\rho}$ the augmented Lagrange functional related to (4) reads as follows: $L(v,q,\rho) := h(v) + \langle Bv - g, q\rangle_Q + (\rho/2)\|Bv - g\|^2_{Q'}$, the related Fréchet derivative is $F(v,q,\rho) := L'_v(v,q,\rho) = Av - f + B^T q + \rho B^T I_Q^{-1}(Bv - g)$. Note that evaluations of dual norms are due to the Riesz theorem, e.g. $\|\varphi\|_{V'} = \sqrt{\langle \varphi, I_V^{-1}\varphi\rangle_V}$. Then, the problem (4) is equivalent to the following saddle–point problem:

$$\min_{v \in V_{\mathrm{ad}}} \max_{q \in Q} L(v,q,\rho)$$

and it poses a unique and stable solution $v^* \in V_{\mathrm{ad}}$, while a related Lagrange multiplier $q^* \in Q$ need not be unique.

2.1 Outer Iterations: SMALBE

For numerical solution we will make use of the classical augmented Lagrangian algorithm, where in outer iterations we maximize over Q and increase the penalty parameter ρ while in inner iterations we solve the following simple–bound constrained qp–subproblems:

$$\min_{v \in V_{\mathrm{ad}}} L(v,q,\rho),$$

where $q \in Q$ and $\rho \geq \underline{\rho}$ are fixed. Note that V_{ad} is convex nonempty and closed and that the latter problem is equivalent to the following variational inequality:

$$\text{Find } v^* \in V_{\mathrm{ad}} \quad \text{such that} \quad \langle F(v^*,q,\rho), v - v^*\rangle_V \geq 0 \quad \forall v \in V_{\mathrm{ad}}. \tag{5}$$

Next, for $v \in V_{\mathrm{ad}}$ by $M(v) := \{w \in V \mid \exists \bar{t} > 0\ \forall t \in [-\bar{t}, \bar{t}]\ :\ v + tw \in V_{\mathrm{ad}}\}$ denote a vector subspace of feasible full–directions, by $N(v) := M(v)^{\perp}$ its orthogonal complement, by $N^+(v) := \{w \in N(v) \mid \exists \bar{t} > 0\ \forall t \in (0, \bar{t}]\ :\ v + tw \in V_{\mathrm{ad}}\}$ a half–space of feasible half–directions, and by $N^-(v) := -N^+(v)$ the other half–space of nonfeasible half–directions. Then, we can equivalently translate the variational inequality (5) into the following nonsmooth equality:

$$[F^P(v^*, q, \rho)](w) = 0 \quad \forall w \in V \qquad \text{or} \qquad \|F^P(v^*, q, \rho)\|_{V(v)^*} = 0, \tag{6}$$

where for any $v \in V_{\mathrm{ad}}$, $w \in V$, decomposed into $w = w_M + w_N^+ + w_N^-$ with $w_M \in M(v)$, $w_N^+ \in N^+(v)$, and $w_N^- \in N^-(v)$, we define the following projection of $F(v, q, \rho)$:

$$[F^P(v, q, \rho)](w) = \langle F(v, q, \rho), w_M \rangle_V + \min\left\{\langle F(v, q, \rho), w_N^+ \rangle_V, 0\right\},$$

where the additive terms are applications of the so–called free and chopped gradient, respectively, and where

$$\|F^P(v, q, \rho)\|^2_{V(v)^*} := \|F(v, z, \rho)\|^2_{M(v)'} + \|F^P(v, z, \rho)\|^2_{N(v)^*},$$
$$\|F^P(v, z, \rho)\|_{N(v)^*} := \sup_{w_N^+ \in N^+(v)} \left|[F^P(v, q, \rho)](w_N^+)\right| / \|w_N^+\|_V.$$

Now we can present Algorithm 1, which is a modification of the classical augmented Lagrangian algorithm such that we additionaly employ an adaptive precision control for solution to the simple–bound constrained subproblems (6), and a special update rule for ρ assuring a monotonic increase of L. For details we refer to [3].

Algorithm 1 Semi–monotonic augmented Lagrangians with adaptive prec. control

Given $\eta > 0$, $\beta > 1$, $\nu > 0$, $\rho^{(0)} \geq \underline{\rho}$, $q^{(0)} \in Q$, precision $\varepsilon > 0$, feasibility precision $\varepsilon_{\mathrm{feas}} > 0$
for $k := 0, 1, 2, \ldots$ **do**
  Find $v^{(k+1)} \in V_{\mathrm{ad}} : \|F^P(v^{(k+1)}, q^{(k)}, \rho^{(k)})\|_{V(v^{(k+1)})^*} \leq \min\left\{\nu \|Bv^{(k+1)} - g\|_{Q'}, \eta\right\}$
  if $\|F^P(v^{(k+1)}, q^{(k)}, \rho^{(k)})\|_{V(v^{(k+1)})^*} \leq \varepsilon$ and $\|Bv^{(k+1)} - g\|_{Q'} \leq \varepsilon_{\mathrm{feas}}$ **then**
    break
  end if
  $q^{(k+1)} := q^{(k)} + \rho^{(k)} I_Q^{-1} B v^{(k+1)}$
  if $k > 0$ and $L(v^{(k+1)}, q^{(k+1)}, \rho^{(k)}) < L(v^{(k)}, q^{(k)}, \rho^{(k-1)}) + \frac{\rho^{(k)}}{2} \|Bv^{(k+1)} - g\|^2_{Q'}$ **then**
    $\rho^{(k+1)} := \beta \rho^{(k)}$
  else
    $\rho^{(k+1)} := \rho^{(k)}$
  end if
end for
$v^{(k+1)}$, $q^{(k+1)}$ is the solution.

2.2 Inner Iterations: MPRGP

In the inner iterations, i.e. the first line in the for–cycle of Algorithm 1, we shall approximately solve the auxiliary subproblems (6), for which we recommend to use Algorithm 2. It is based on the conjugate gradient method with proportioning and reduced gradient projections. We denote by $\|A\|_V := \sup_{\|w\|_V=1} \|Aw\|_{V'}$ the norm of the linear continuous mapping. For more details we refer to [3, 4].

Algorithm 2 Modified proportioning with reduced gradient projections

Given $\Gamma > 0$, $\overline{\alpha} \in (0, 2\|A\|_V^{-1}]$, $\eta > 0$, $\nu > 0$, $v^{(0)} \in V_{\text{ad}}$, $q \in Q$, $\rho \geq \underline{\rho}$, prec. $\varepsilon > 0$, $\varepsilon_{\text{feas}} > 0$
for $l := 0, 1, 2, \ldots$ **do**
 if $\|F^P(v^{(l)}, q, \rho)\|_{V(v^{(l)})^*} \leq \min\left\{\nu\|Bv^{(l)} - g\|_{Q'}, \eta\right\}$ or $\left(\|F^P(v^{(l)}, q, \rho)\|_{V(v^{(l)})^*} \leq \varepsilon\right.$ and $\left.\|Bv^{(l)} - g\|_{Q'} \leq \varepsilon_{\text{feas}}\right)$ **then**
  break
 end if
 if $\|F^P(v^{(l)}, q, \rho)\|_{N(v)^*} < \Gamma\|F(v^{(l)}, q, \rho)\|_{M(v)'}$ **then**
  Generate $v^{(l+1)}$ by the conjugate gradient step
  if $v^{(l+1)} \notin V_{\text{ad}}$ **then**
   Generate $v^{(l+1/2)}$ by the maximal feasible conjugate gradient step
   Generate $v^{(l+1)}$ as a feasible $(\overline{\alpha})$ addition of the free gradient and project onto V_{ad}
   Restart conjugate gradients with the free gradient
  end if
 else
  Restart conjugate gradients with the chopped gradient
  Generate $v^{(l+1)}$ by the conjugate gradient step
 end if
end for
$v^{(l+1)}$ is the solution.

2.3 Analysis: Linear Complexity, Multigrid Preconditioning

Let us present the main theoretical result, see [3, Theorem 5]. Assume a class $\mathscr{T}$ of the following finite–dimensional optimization problems: for $t \in \mathscr{T}$

$$\min_{v_t \in V_{\text{BE}}^t} h_t(v) \tag{7}$$

with $V_{\text{BE}}^t := \{v \in \mathrm{I\!R}^{n_t} \mid B_t v = 0 \text{ and } v \geq \underline{v}_t\}$, $h_t(v) := (1/2)v^T A_t v - f_t^T v$, $A_t \in \mathrm{I\!R}^{n_t \times n_t}$ symmetric positive definite, $B_t \in \mathrm{I\!R}^{m_t \times n_t}$, and $f_t, \underline{v}_t \in \mathrm{I\!R}^{n_t}$. Assume also $0 \in V_{\text{BE}}^t$. The following theorem gives us optimality of Algorithm 1 in terms of matrix–vector multiplications provided the inner iterations are implemented by Algorithm 2. By $\|.\|$ we now denote the Euclidean norm.

Theorem 1. *Let* $0 < a_{\min} < a_{\max}$ *and* $0 < b_{\max}$ *be given constants and let the class of problems (7) satisfies* $a_{\min} \leq \lambda_{\min}(A_t) \leq \lambda_{\max}(A_t) \leq a_{\max}$ *and* $\|B_t\| \leq c_{\max}$. *Let* $v_t^{(k)}$, $q_t^{(k)}$ *and* $\rho_t^{(k)}$ *be generated by Algorithm 1 for (7) with* $\|f_t\| \geq \eta_t > 0$, $\beta > 1$, $\nu > 0$, $\rho_0 := \rho_t^{(0)} > 0$, $q_t^{(0)} = 0$. *Let* $s \geq 0$ *denote the smallest integer such that* $\beta^s \rho_0 \geq \nu^2 / a_{\min}$ *and let the inner iterations be implemented by Algorithm 2 with the parameters* $\Gamma > 0$ *and* $\overline{\alpha} \in (0, (a_{\max} + \beta^s \rho_0 c_{\max}^2)^{-1}]$ *to generate the iterates* $v_t^{(k,0)}, v_t^{(k,1)}, \ldots, v_t^{(k,l)} =: v_t^{(k)}$ *for the solution of (7) starting from* $v_t^{(k,0)} := v_t^{(k-1)}$ *with* $v_t^{-1} := 0$, *where* $l = l_{t,k}$ *is the first index satisfying* $\|F^P(v_t^{(k,l)}, q_t^{(k)}, \rho_t^{(k)})\| \leq \nu \|B_t v_t^{(k,l)}\|$ *or* $\|F^P(v_t^{(k,l)}, q_t^{(k)}, \rho_t^{(k)})\| \leq \varepsilon \|f_t\| \min\{1, \nu^{-1}\}$. *Then Algorithm 1 generates an approximate solution* $v_t^{(k_t)}$ *of any problem (7) which satisfy*

$$\nu^{-1} \|v_t^{(k,l)}, q_t^{(k)}, \rho_t^{(k)})\| \leq \|B_t v_t^{(k,l)}\| \leq \varepsilon \|f_t\|$$

at $O(1)$ *matrix–vector multiplications by the Hessian of the augmented Lagrangian* L_t *for (7).*

Theorem 1 indicates that Algorithm 1 works optimally provided a uniformly bounded spectra of all the Hessians A_t, where t typically denotes a discretization parameter. Thus, we can construct a multigrid preconditioner for A denoted by $\widehat{A}$, and substitute each occurence of v by $\widehat{A}^{-1/2}\widehat{v}$, which will guarantee boundeness of spectra of A_t as well as linear complexity of number of the inner CG–iterations. However, under this substitution we would change the simple–bound constraint to a linear inequality constraint, which might be more tricky to handle. Therefore, we recommend to use a diagonal preconditioner for A only. Note also that Theorem 1 is much less general than our exposition in the previous sections, but we believe that a generalization of the result will be as straightforward as it has recently turned up in case of equality constrained quadratic programming, see [5].

Additionaly, for proper measurements of dual norms we need applications of inverses of I_V and I_Q to be of the linear complexity too. Thus, we can replace applications of I_V^{-1} and I_Q^{-1} by approximate inverse applications $\widehat{I_V}^{-1}$ and $\widehat{I_Q}^{-1}$ using multigrid again.

3 Numerical Results

We present numerical results for two benchmark problems. First, we consider an optimal control problem for image segmentation. Given $\Omega := (0,1)^2$, a noisy colour (red, green, blue components) image data $y^{\mathrm{d}} \in \left[L^2(\Omega)\right]^3$ and a regularization parameter $\mu > 0$, we look for sources $u \in \left[L^2(\Omega)\right]^3$ that produced homogeneous colour segments in the image. This leads to the following optimal control problem:

$$\min_{(u,y) \in \mathscr{U}_{\mathrm{ad}} \times \mathscr{Y}} \left\{ \frac{1}{2} \|y(x) - y^{\mathrm{d}}(x)\|^2_{[L^2(\Omega)]^3} + \frac{\mu}{2} \|u(x)\|^2_{[L^2(\Omega)]^3} \right\} \text{ s.t. } -\triangle y = u \text{ on } \mathscr{Y}',$$

where $\mathscr{U}_{\mathrm{ad}} := \{u \in [L^2(\Omega)]^3 \mid 0 \le u(x) \le 1 \text{ a.e. in } \Omega\}$ and $\mathscr{Y} := \left[H_0^1(\Omega)\right]^3$.

By numerical experiments we realized that it is enough to proceed with an SQP method, where we neglect the simple–bound constraint and solve the following unconstrained saddle–point system and then project the resulting u^i onto $\mathscr{U}_{\mathrm{ad}}$:

$$\begin{pmatrix} I_{L^2} & 0 & \text{sym.} \\ 0 & \mu I_{L^2} & \text{sym.} \\ -\triangle & -I_{L^2\to H^1} & 0 \end{pmatrix} \begin{pmatrix} y_j^i \\ u_j^i \\ q_j^i \end{pmatrix} = \begin{pmatrix} I_{L^2} y_j^{\mathrm{d}} \\ 0 \\ 0 \end{pmatrix} \quad \text{for} \quad j = 1,2,3,$$

where I_{L^2} stands for the identity (inner product) operator on $L^2(\Omega)$ and $I_{L^2\to H^1}$ stands for the orthogonal projection from $L^2(\Omega)$ to $H^1(\Omega)$. We use a finite element method and employ linear nodal Lagrange elements for both $\mathscr{U}$ and $Q = \mathscr{Y}$. We construct geometric multigrid preconditioners for $I_Q = -\triangle$ and $I_V = I_{[L^2(\Omega)]^3}$ so that a point diagonal smoother with 3 symmetric smoothing iterations is applied for the latter. For $I_Q = -\triangle$ we test a point additive smoother as well as a block Gauss–Seidel smoother with 3 symmetric smoothing iterations again. Since we have neglected the bound constraint, we use the preconditioned conjugate gradients (PCG) method instead of MPRGP. Numerical results for the first SQP iteration and for the red component with $\mu := 10^{-4}$ and relative precisions $\varepsilon = \varepsilon_{\mathrm{feas}} = 10^{-3}$ are depicted in Fig. 1 and Table 1. The results are similar for the other colour components. Note that the number of iterations holds about a constant. Yet we have to improve our implementation in Matlab in order to get large–scale simulations in a reasonable time.

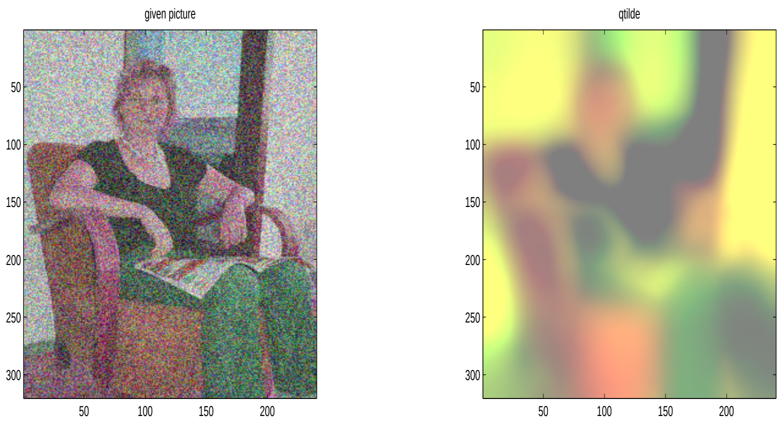

Fig. 1 Image segmentation: original noisy image y^{d} and the reconstructed smooth segments u^1

Second, we consider a parameter identification problem for 2–dimensional magnetostatics. Given a rectangular domain $\Omega \subset \mathbb{R}^2$, a measured magnetic field distribution $B^{\mathrm{g}} = \mathrm{curl}(y^{\mathrm{g}})$, where $y^{\mathrm{g}} \in H_0^1(\Omega)$, a forcing electric current term $g \in L^2(\Omega)$

Table 1 Numerical experiments for optimal control in image segmentation

level l	$\dim(Q_l)$	point additive smoother: **SMALBE**/PCG **iterations**	point additive smoother: **total PCG iterations**	block Gauss–Seidel smoother: **SMALE**/PCG **iterations**	block Gauss–Seidel smoother: **total PCG iterations**
1	336	**5**/7,6,9,7,9	**38**	**5**/7,6,9,7,9	**38**
2	1271	**5**/9,10,10,11,13	**53**	**5**/7,6,9,9,9	**40**
3	4941	**5**/9,11,11,12,12	**55**	**5**/7,9,8,6,9	**39**
4	19481	**5**/8,11,12,13,13	**57**	out of time	
5	77361	**5**/9,11,11,12,13	**56**	out of time	

such that $\operatorname{div}(g) = 0$, reluctivity of the air $\nu_0 := 4\pi 10^{-7}$, a minimal reluctivity of ferromagnetic components $\nu_1 = \nu_0/5000$ and a regularization parameter $\mu > 0$, we search for a distribution $u \in L^2(\Omega)$ of the magnetic reluctivity that has caused the measured magnetic field B^{g}. This leads to the following problem:

$$\min_{(u,y)\in\mathscr{U}_{\mathrm{ad}}\times\mathscr{Y}} \left\{ \frac{1}{2}\|\nabla y(x) - \nabla y^{\mathrm{g}}(x)\|^2_{L^2(\Omega)} + \frac{\mu}{2}\|u(x)\|^2_{L^2(\Omega)} \right\}$$

$$\text{s.t.} \quad -\operatorname{div}\left((\nu_0 + (\nu_1 - \nu_0)u)\,\nabla y\right) = g \text{ on } \mathscr{Y}',$$

where $\mathscr{U}_{\mathrm{ad}} := \{u \in L^2(\Omega) \mid 0 \le u(x) \le 1 \text{ a.e. in } \Omega\}$ and $\mathscr{Y} := H^1_0(\Omega)^3$.

SQP approximations now lead to simple–bound and equality constrained qp–subproblems with the following Hessian:

$$\begin{pmatrix} -\triangle & 0 & \text{sym.} \\ 0 & \mu I_{L^2} & \text{sym.} \\ -\operatorname{div}(q^i\nabla\cdot) & -\operatorname{div}(\cdot\nabla y^i) & 0 \end{pmatrix}.$$

However, it turned out by numerical experiments that at some SQP iterations the design search set V^i_{BE} is empty. As a remedy we relax the upper bound constraint such that $0 \le u(x) \le \gamma^i$, where $\gamma^i \to 1_+$. For approximation we use linear Lagrange finite elements (fe) for $\mathscr{Y} = Q$ and elementwise constant basis for $\mathscr{U}$. We build a geometric multigrid preconditioner for $-\triangle$ with 3 symmetric Gauss–Seidel smoothing steps. The inner product on $\mathscr{U}$ can be inverted directly, since the fe–approximations of $I_{\mathscr{U}}$ are diagonal matrices. Therefore, we can also use the tensor–product preconditioner for the Hessians without changing the simple–bound constraint into a linear constraint, and we can make use of MPRGP. Numerical results for the first and second SQP iteration with $\mu := 10^{-4}$ and relative precisions $\varepsilon = \varepsilon_{\mathrm{feas}} = 10^{-3}$ are depicted in Fig. 2 and Table 2. While the number of SMALBE iterations seems to be about constant, yet we have to improve preconditioning of $A + \rho B^T I_Q^{-1} B$, see the increasing numbers of CG–iterations as well as expansion steps in Table 2.

Acknowledgements This work has been supported by the Czech Grant Agency under the grant GAČR 201/05/P008, by the Czech Ministry of Education under the project MSM6198910027, and by the Czech Academy of Science under the project AVČR 1ET400300415.

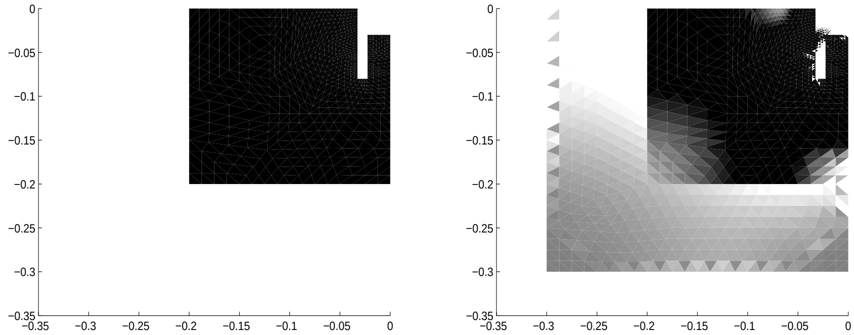

Fig. 2 Parameter identification in 2–dimensional magnetostatics: original and reconstructed ferromagnetics distribution u

Table 2 Numerical experiments for parameter identification in 2–dimensional magnetostatics

level primal / dual DOFs (y's + u's) / (q's)	1st SQP iteration **SMALBE / total inner steps** PCG + exp. + prop. steps	2nd SQP iteration **SMALBE / total inner steps** PCG + expansion + proportioning steps
1 91+187 / 91	**7 / 15+0+0** 1, 3, 2, 2, 3, 2, 2	**5 / 78+16+1** 15+13+0, 26+2+1, 12+1+0, 16+0+0, 9+0+0
2 373+784 / 373	**7 / 17+0+0** 2, 3, 2, 3, 3, 2, 2	**5 / 152+18+3** 20+11+0, 27+6+1, 35+1+1, 37+0+0, 33+0+1
3 1574+2992 / 1574	**8 / 22+0+0** 2, 3, 3, 3, 3, 3, 3, 2	**4 / 185+52+3** 19+30+0, 73+19+2, 46+3+0, 47+0+1
4 6292+11968 / 6292	**8 / 30+0+0** 3, 4, 4, 4, 4, 4, 4, 3	**5 / 377+176+5** 26+80+0, 160+87+4, 76+7+1, 73+2+0, 42+0+0

References

1. Burger, M., Mühlhuber, W.: Iterative regularization of parameter identification problems by sequential quadratic programming methods. Inverse Probl. **18**(4), 943–969 (2002)
2. Burger, M., Mühlhuber, W.: Numerical approximation of an SQP-type method for parameter identification. SIAM J. Numer. Anal. **40**(5), 1775–1797 (2002)
3. Dostál, Z.: An optimal algorithm for bound and equality constrained quadratic programming problems with bounded spectrum. Computing **78**(4), 311–328 (2006)
4. Dostál, Z., Schöberl, J.: Minimizing quadratic functions over non–negative cone with the rate of convergence and finite termination. Comput. Optim. Appl. **30**(1), 23–44 (2005)
5. Lukáš, D., Dostál, Z.: Optimal multigrid preconditioned semi–monotonic augmented Lagrangians applied to the Stokes problem. Numer. Linear Algebra Appl. **14**(9), 741–750 (2007)