SIAM J. MATRIX ANAL. APPL.
Vol. 31, No. 4, pp. 1835–1847

APPROXIMATE NULLSPACE ITERATIONS FOR KKT SYSTEMS*

KAZUFUMI ITO†, KARL KUNISCH‡, VOLKER SCHULZ§, AND ILIA GHERMAN§

Abstract. We investigate a linear iteration scheme for solving Karush–Kuhn–Tucker systems arising from optimization problems with linear equality constraints. The iterations are motivated by the simplicity of the proposed combination of iterations for the forward and adjoint systems that need to be solved and for which efficient solvers may already be available. Convergence results are derived, and their practical relevance is investigated by means of a numerical example.

Key words. KKT systems, iterative solvers, optimization

AMS subject classifications. 65F10, 65K05, 90C20, 93C20

DOI. 10.1137/080724952

1. Introduction.

We consider the quadratic optimization problem with equality constraints

$$\text{(QP)} \qquad \begin{aligned} &\min_{x,p} \tfrac{1}{2}\left(x^\top H_x\, x + x^\top H_{xp}\, p + p^\top H_{px}\, x + p^\top H_p\, p\right) + f_x^\top x + f_p^\top p \\ &\text{subject to (s.t.)} \quad C_x\, x + C_p\, p + c = 0, \end{aligned}$$

where $x \in \mathbb{R}^{n_x}$, $p \in \mathbb{R}^{n_p}$ are the variable vectors of the optimization problem, the vectors $c, f_x \in \mathbb{R}^{n_x}$, $f_p \in \mathbb{R}^{n_p}$, and the matrices arising are of consistent dimensions:

$$\begin{aligned} &H_x \in \mathbb{R}^{n_x \times n_x}, \quad H_p \in \mathbb{R}^{n_p \times n_p}, \quad H_{xp} = H_{px}^\top \in \mathbb{R}^{n_x \times n_p}, \\ &C_x \in \mathbb{R}^{n_x \times n_x}, \quad C_p \in \mathbb{R}^{n_x \times n_p}. \end{aligned}$$

We assume as an important structural property that C_x is nonsingular and utilize a nullspace basis Z such that $[C_x\ C_p]Z = 0$:

$$Z = \begin{bmatrix} -C_x^{-1} C_p \\ I \end{bmatrix}.$$

Considering the Hessian of (QP)

$$H := \begin{bmatrix} H_x & H_{xp} \\ H_{px} & H_p \end{bmatrix},$$

we assume that the respective reduced Hessian

$$S = Z^\top H Z = H_p - H_{px} C_x^{-1} C_p - C_p^\top C_x^{-\top} H_{xp} + C_p^\top C_x^{-\top} H_x C_x^{-1} C_p \tag{1.1}$$

is positive definite, which guarantees the unique solvability of (QP). These linear-quadratic problems typically arise as subproblems within large-scale model-based optimization tasks as, e.g., in [HSBG05, HSB08], where the constraint consists of a

*Received by the editors May 22, 2008; accepted for publication (in revised form) by A. S. Lewis January 25, 2010; published electronically March 31, 2010.
http://www.siam.org/journals/simax/31-4/72495.html

†Department of Mathematics, North Carolina State University, Raleigh, NC 27695 (kito@math.ncsu.edu).

‡Institut für Mathematik, Karl-Franzens-Universität Graz, Heinrichstr. 36, A-8010 Graz, Austria (karl.kunisch@uni-graz.at).

§Department of Mathematics, University of Trier, Universitätsring 15, 54286 Trier, Germany (volker.schulz@uni-trier.de, ilia.gherman@uni-trier.de).

discretized partial differential equation (PDE). However, within this paper, we focus on the linear-quadratic problem itself.

Problem (QP) is known to satisfy a KKT system of the form

$$\begin{bmatrix} H_x & H_{xp} & C_x^\top \\ H_{px} & H_p & C_p^\top \\ C_x & C_p & 0 \end{bmatrix} \begin{pmatrix} x \\ p \\ \lambda \end{pmatrix} = - \begin{pmatrix} f_x \\ f_p \\ c \end{pmatrix}, \tag{1.2}$$

and we should note that the reduced Hessian S can be interpreted as the Schur complement of the KKT matrix in (1.2) with respect to the variables (x, λ). The block structure in the system matrix has motivated research in block-structured preconditioners for Krylov subspace methods. In [BS01] three different preconditioners for the solution of (QP) with minimal residual method (MINRES) and symmetric LQ method (SYMMLQ) [PS75] are analyzed and tested. In [BG05], a preconditioner very similar to our methods below is used and applied to accelerate GMRES [SS86] iterations for the linear system. In [HA01] again a block-structured preconditioner is used to accelerate QMR [FN94] for model-based output least-squares problems. The article [DW06] again studies approximate block factorizations in the context of Krylov iterations.

However, the application of Krylov-subspace methods might not be practical in situations where a very high degree of modularity is required. This is particularly the case when one aims to utilize previously existing iterative solvers for the equality constraints in (QP). Furthermore, a generalization to nonlinear iterations for related nonlinear problems is not obvious in a Krylov subspace context. Besides that, the preconditioners used in the publications mentioned have so far not been analyzed as iterative solvers. This is an interesting subject in itself and is the purpose of this paper. The aim is to perform investigations for model-based quadratic programming similar to those in [BWY90] for variational saddle-point problems like the Stokes equation. In addition to modularity, large-scale applications may impede the use of a direct sparse solver. Instead an iterative refinement is used.

In particular, we investigate the following linear iteration for solving (QP):

$$(x^{k+1}, p^{k+1}, \lambda^{k+1}) = (x^k, p^k, \lambda^k) + (\Delta x^k, \Delta p^k, \Delta \lambda^k), \tag{1.3a}$$

where

$$\begin{bmatrix} 0 & 0 & A_a \\ 0 & B & C_p^\top \\ A_f & C_p & 0 \end{bmatrix} \begin{pmatrix} \Delta x^k \\ \Delta p^k \\ \Delta \lambda^k \end{pmatrix} = - \begin{bmatrix} H_x & H_{xp} & C_x^\top \\ H_{px} & H_p & C_p^\top \\ C_x & C_p & 0 \end{bmatrix} \begin{pmatrix} x^k \\ p^k \\ \lambda^k \end{pmatrix} - \begin{pmatrix} f_x \\ f_p \\ c \end{pmatrix} \tag{1.3b}$$

involves the approximations

$$A_f \approx C_x, \quad A_a \approx C_x^\top, \quad B \approx Z^\top H Z.$$

The matrices $A_f, A_a \in \mathbb{R}^{n_x \times n_x}$ are supposed to be square, nonsingular, and cheaply invertible, and $B \in \mathbb{R}^{n_p \times n_p}$ is symmetric positive definite. We will see below that in the case $A_f = C_x, A_a = C_x^\top, B = S$, iteration (1.3) stops after three iterations at the solution of (QP). This may motivate choosing $B \approx S$ always. However, it is the main point of the paper to observe that another choice for B in the case $A_f \neq C_x, A_a \neq C_x^\top$ should be preferred.

In fact, in practical examples (see, e.g., [GS05]), the following fact has been observed that seems surprising at first glance: The method iteration (1.3) works

better if B is an approximation to the reduced Hessian consistent with the choice of A_f, A_a as approximations to $C_x, C_x^\top$ in the form

$$S_A = H_p - H_{px}A_f^{-1}C_p - C_p^\top A_a^{-1}H_{xp} + C_p^\top A_a^{-1}H_x A_f^{-1}C_p \tag{1.4}$$

rather than the exact reduced Hessian S from (1.1). We will give an explanation for this observation. One should note that this is in line with similar studies for variational saddle-point problems as in [BWY90]. However, the convergence theory there cannot be carried over to iteration (1.3).

Since the iteration concept is based on an approximate nullspace representation (C_x is replaced by A_f, and $C_x^\top$ is replaced by A_a), we call it an approximate nullspace iterative technique. The iterations considered in this paper are related to the so-called piggyback iterations in [GF02, Gri06]. This is described in more detail in section 4. In contrast to [BS01, HA01, BG05, DW06], we do not consider here preconditioners to be used within some Krylov method but rather focus on complete linear iterations.

The matrix

$$R := \begin{bmatrix} 0 & 0 & A_a \\ 0 & B & C_p^\top \\ A_f & C_p & 0 \end{bmatrix}$$

arising in (1.3) can also be considered as a preconditioner, and related matrices are used in [BS01, HA01, BG05] as preconditioners within a Krylov iteration setting. In [BS01], an eigenvalue analysis is performed for the case $A_f = C_x, A_a = C_x^\top, B = S$, and approximations are numerically tested within a Krylov framework. In [HA01] the output-least-squares structure is exploited for the generation of B. In [BG05], quasi-Newton update techniques based on an outer iteration are recommended for the generation of B.

In section 2 we give the main theoretical results of this paper and their proofs. Section 3 is devoted to the application of these results in the context of a generic optimal control problem often found in the literature. Section 4 presents an algorithmic reformulation of the linear iteration as a preparation for nonlinear variants and for comparison with other similar iterative strategies. Numerical experiments supporting the theory of section 2 are given in section 5.

2. Convergence results. In this section, we show that iteration (1.3) is convergent under certain circumstances, and we also give criteria for the convergence. The convergence theory of this section is based on a perturbation analysis. First we show finite-step convergence for a reduced-type exact solver, where B in R is chosen as S_A.

LEMMA 2.1. *If A_f, A_a and the Schur complement S_A are invertible, then iteration* (1.3) *with $B = S_A$ converges after three steps to the exact solution of the (perturbed) problem*

$$\begin{bmatrix} H_x & H_{xp} & A_a \\ H_{px} & H_p & C_p^\top \\ A_f & C_p & 0 \end{bmatrix} \begin{pmatrix} x \\ p \\ \lambda \end{pmatrix} + \begin{pmatrix} f_x \\ f_p \\ c \end{pmatrix} = 0.$$

Proof. We consider the iteration matrix M and show that it is nilpotent. First we give the exact inverse in block form:

$$\begin{bmatrix} 0 & 0 & A_a \\ 0 & S_A & C_p^\top \\ A_f & C_p & 0 \end{bmatrix}^{-1} = \begin{bmatrix} A_f^{-1}C_pS_A^{-1}C_p^\top A_a^{-1} & -A_f^{-1}C_pS_A^{-1} & A_f^{-1} \\ -S_A^{-1}C_p^\top A_a^{-1} & S_A^{-1} & 0 \\ A_a^{-1} & 0 & 0 \end{bmatrix}.$$

Now we compute explicitly the iteration matrix

$$M = I - \begin{bmatrix} 0 & 0 & A_a \\ 0 & S_A & C_p^\top \\ A_f & C_p & 0 \end{bmatrix}^{-1} \begin{bmatrix} H_x & H_{xp} & A_a \\ H_{px} & H_p & C_p^\top \\ A_f & C_p & 0 \end{bmatrix}$$
$$= \begin{bmatrix} M_{11} & M_{12} & 0 \\ (S_A^{-1} C_p^\top A_a^{-1} H_x - S_A^{-1} H_{px}) & (S_A^{-1} C_p^\top A_a^{-1} H_{xp} + S_A^{-1}(S_A - H_p)) & 0 \\ -A_a^{-1} H_x & -A_a^{-1} H_{xp} & 0 \end{bmatrix},$$

where

$$M_{11} = -A_f^{-1} C_p S_A^{-1} C_p^\top A_a^{-1} H_x + A_f^{-1} C_p S_A^{-1} H_{px},$$
$$M_{12} = -A_f^{-1} C_p S_A^{-1} C_p^\top A_a^{-1} H_{xp} - A_f^{-1} C_p S_A^{-1}(S_A - H_p).$$

We study M^2 and keep in mind definition (1.4) for S_A. Separately we investigate each block of the 3×3-block matrix M^2 that is not obviously zero. In order to simplify the notation, we define the formal expression

$$\mathcal{Z} := C_p^\top A_a^{-1} H_x A_f^{-1} C_p - H_{px} A_f^{-1} C_p - C_p^\top A_a^{-1} H_{xp} - S_A + H_p.$$

Of course, $\mathcal{Z} = 0$, but we have to keep in mind this formal expression in order to be able to understand the subsequent arguments. We compute the blocks

$$(M^2)_{(1,1)} = A_f^{-1} C_p S_A^{-1} \mathcal{Z} S_A^{-1} C_p^\top A_a^{-1} H_x - A_f^{-1} C_p S_A^{-1} \mathcal{Z} S_A^{-1} H_{px} = 0,$$
$$(M^2)_{(1,2)} = A_f^{-1} C_p S_A^{-1} \mathcal{Z} S_A^{-1} C_p^\top A_a^{-1} H_{xp} + A_f^{-1} C_p S_A^{-1} \mathcal{Z} S_A^{-1}(S_A - H_p) = 0,$$
$$(M^2)_{(2,1)} = S_A^{-1} \mathcal{Z} S_A^{-1} H_{xp} - S_A^{-1} \mathcal{Z} S_A^{-1} C_p A_a^{-1} H_x = 0,$$
$$(M^2)_{(2,2)} = -S_A^{-1} \mathcal{Z} S_A^{-1} C_p A_a^{-1} H_{xp} - S_A^{-1} \mathcal{Z} S_A^{-1}(S_A - H_p) = 0.$$

Therefore, M^2 is of the form

$$M^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ * & * & 0 \end{bmatrix},$$

and obviously $M^3 = MM^2 = 0$. ∎

Henceforth we use the following abbreviations:

$$K := \begin{bmatrix} H_x & H_{xp} & C_x^\top \\ H_{px} & H_p & C_p^\top \\ C_x & C_p & 0 \end{bmatrix}, \quad \tilde{K} := \begin{bmatrix} H_x & H_{xp} & A_a \\ H_{px} & \tilde{H}_p & C_p^\top \\ A_f & C_p & 0 \end{bmatrix}, \qquad \tilde{H}_p := H_p - S_A + B.$$

Note that the matrix B is the exact Schur complement for $\tilde{K}$ that arises by eliminating the first and third variables. Therefore, we can conclude from Lemma 2.1 that

$$(I - R^{-1}\tilde{K})^3 = 0.$$

The iteration matrix for iteration (1.3) is given by

$$I - R^{-1}K = I - R^{-1}\tilde{K} + R^{-1}(\tilde{K} - K) =: M + N, \tag{2.1}$$

where M is a nilpotent matrix of nilpotency degree 3 and N can be considered as a perturbation of M. For any matrix norm $\|.\|$ subordinate to a vector-space norm, this yields

$$\|(M+N)^3\| \leq \left[\|M^2+MN+NM+N^2\| + \|M^2+NM\| + \|M\|^2\right]\|N\|.$$

The following lemma is based on a perturbation argument and estimates the influence of the perturbation induced by $\tilde{K}$.

LEMMA 2.2. *Define*

$$\theta := \|M^2+MN+NM+N^2\| + \|M^2+NM\| + \|M\|^2, \quad r := \|N\| = \|R^{-1}(\tilde{K}-K)\|.$$

If $\theta r < 1$*, then iteration* (1.3) *converges with the convergence rate bounded above by*

$$\kappa := \sqrt[3]{\theta r} < 1.$$

Proof. Since $\rho(B) = \lim_{n\to\infty} \|B^n\|^{\frac{1}{n}}$ for any square matrix B, we obtain

$$\begin{aligned}\rho(I-R^{-1}K) &= \sqrt[3]{\rho((M+N)^3)} = \sqrt[3]{\lim_{n\to\infty} \|(M+N)^{3n}\|^{1/n}} \\ &= \sqrt[3]{\lim_{n\to\infty} \|(M+N)^{3n}\|^{1/n}} \\ &\leq \sqrt[3]{\theta r} < 1. \qquad \square\end{aligned}$$

Remark 1. We would like to point out that the problem of perturbation of the spectrum of a matrix with a single eigenvalue, which is 0 in our case (for the matrix M), has received a considerable amount of attention in the literature (see, e.g., [CI97] and the references given there). For our work, we prefer to utilize the estimate resulting from Lemma 2.2 rather than other perturbation results.

Henceforth we need to specify a norm, which we fix as the ℓ_2-norm.

THEOREM 2.3. *We define the numerical spectral norms*

$$r_A^f := \|I - A_f^{-1}C_x\|_2, \qquad r_A^a := \|I - A_a^{-1}C_x^\top\|_2, \qquad r_S := \|I - B^{-1}S_A\|_2.$$

If

$$\max\{r_A^f, r_A^a, r_S\} < 1/\tilde{\theta}, \tag{2.2}$$

with

$$\tilde{\theta} := \theta\varphi \quad \text{and } \varphi := \left\| \begin{bmatrix} A_f^{-1}C_pB^{-1}C_p^\top & -A_f^{-1}C_p & I \\ -B^{-1}C_p^\top & I & 0 \\ I & 0 & 0 \end{bmatrix} \right\|_2,$$

then iteration (1.3) *converges with the convergence rate bounded above by*

$$\kappa = \sqrt[3]{\theta r} < 1.$$

Proof. We observe that

$$
\begin{aligned}
&R^{-1}(\tilde{K}-K)\\
&\quad=\begin{bmatrix} A_f^{-1}C_pB^{-1}C_p^\top A_a^{-1} & -A_f^{-1}C_pB^{-1} & A_f^{-1}\\ -B^{-1}C_p^\top A_a^{-1} & B^{-1} & 0\\ A_a^{-1} & 0 & 0\end{bmatrix}\begin{bmatrix} 0 & 0 & A_a-C_x^\top\\ 0 & B-S_A & 0\\ A_f-C_x & 0 & 0\end{bmatrix}\\
&\quad=\begin{bmatrix} A_f^{-1}C_pB^{-1}C_p^\top & -A_f^{-1}C_p & I\\ -B^{-1}C_p^\top & I & 0\\ I & 0 & 0\end{bmatrix}\begin{bmatrix} 0 & 0 & I-A_a^{-1}C_x^\top\\ 0 & I-B^{-1}S_A & 0\\ I-A_f^{-1}C_x & 0 & 0\end{bmatrix},
\end{aligned}
$$

where

$$
\left\|\begin{bmatrix} 0 & 0 & I-A_a^{-1}C_x^\top\\ 0 & I-B^{-1}S_A & 0\\ I-A_f^{-1}C_x & 0 & 0\end{bmatrix}\right\|_2=\max\{r_A^f,r_A^a,r_S\}.
$$

The application of Lemma 2.2 gives

$$
\begin{aligned}
\rho(I-R^{-1}K)\le\sqrt[3]{\|(I-R^{-1}K)^3\|_2}&=\underbrace{\sqrt[3]{\theta\|(R^{-1}(\tilde{K}-K))^3\|_2}}_{\kappa}\\
&\le\sqrt[3]{\theta\varphi\max\{r_A^f,r_A^a,r_S\}}<1. \qquad \square
\end{aligned}
$$

Remark 2. Since M and N as defined in (2.1) depend on A_f and A_a and therefore so does $\tilde{\theta}$, we need to verify that condition (2.2) can be satisfied for some choice of A_f and A_a. In particular, we observe that $\tilde{\theta}$ stays bounded from above when $A_f\to C_x$ and $A_a\to C_x^\top$ because then

$$
\begin{aligned}
&N\to 0,\\
&M\to I-\begin{bmatrix} 0 & 0 & C_x^\top\\ 0 & S & C_p^\top\\ C_x & C_p & 0\end{bmatrix}^{-1}\begin{bmatrix} H_x & 0 & C_x^\top\\ 0 & H_p & C_p^\top\\ C_x & C_p & 0\end{bmatrix}=:M_{\lim},\\
&\Rightarrow\theta\to 2\|M_{\lim}^2\|+\|M_{\lim}\|^2=:\theta_{\lim},\\
&\varphi\to\left\|\begin{bmatrix} C_x^{-1}C_pS^{-1}C_p^\top & -C_x^{-1}C_p & I\\ -S^{-1}C_p^\top & I & 0\\ I & 0 & 0\end{bmatrix}\right\|_2=:\varphi_{\lim},\\
&\Rightarrow\tilde{\theta}\to\theta_{\lim}\varphi_{\lim}<\infty.
\end{aligned}
$$

Theorem 2.3 shows that the convergence behavior of the approximate nullspace iteration is limited both by the approximation quality in the forward and adjoint systems and by the approximation quality of the consistent Schur complement in a worst-case fashion. Often, one might choose $A_a=A_f^\top=:A$. In this situation, we can give a refined version of Theorem 2.3, if additionally $C_x^\top=C_x$ and one chooses the spectral norm

$$
\left\|\begin{pmatrix} x\\ p\\ \lambda\end{pmatrix}\right\|_R:=\left\|\begin{pmatrix} A^{1/2}x\\ B^{1/2}p\\ A^{1/2}\lambda\end{pmatrix}\right\|_2.
$$

COROLLARY 2.4. *Choose* $A_a = A_f^\top = A$, *where* A *is a symmetric matrix and* $\|.\| := \|.\|_R$, *and define*

$$\rho_A := \rho(I - A^{-1}C_x), \qquad \rho_S := \rho(I - B^{-1}S_A).$$

Then if $\max\{\rho_A^f, \rho_S\} < 1/\bar{\theta}$ *with*

$$\bar{\theta} := \theta\bar{\varphi} \quad \text{and } \bar{\varphi} := \rho\left(\begin{bmatrix} A^{-1/2}C_pB^{-1}C_p^\top A^{-1/2} & -A^{-1/2}C_pB^{-1/2} & I \\ -B^{-1/2}C_p^\top A^{-1/2} & I & 0 \\ I & 0 & 0 \end{bmatrix}\right),$$

then iteration (1.3) *converges with the convergence rate bounded above by*

$$\kappa = \sqrt[3]{\theta r} < 1.$$

Proof. We give more refined representations of the factors

$$\begin{aligned} &\left\| \begin{bmatrix} A^{-1}C_pB^{-1}C_p^\top & -A^{-1}C_p & I \\ -B^{-1}C_p^\top & I & 0 \\ I & 0 & 0 \end{bmatrix} \right\|_R \\ &\quad = \left\| \begin{bmatrix} A^{-1/2}C_pB^{-1}C_p^\top A^{-1/2} & -A^{-1/2}C_pB^{-1/2} & I \\ -B^{-1/2}C_p^\top A^{-1/2} & I & 0 \\ I & 0 & 0 \end{bmatrix} \right\|_2 \\ &\quad = \rho\left(\begin{bmatrix} A^{-1/2}C_pB^{-1}C_p^\top A^{-1/2} & -A^{-1/2}C_pB^{-1/2} & I \\ -B^{-1/2}C_p^\top A^{-1/2} & I & 0 \\ I & 0 & 0 \end{bmatrix}\right) \end{aligned}$$

and

$$\begin{aligned} &\left\| \begin{bmatrix} 0 & 0 & I - A^{-1}C_x^\top \\ 0 & I - B^{-1}S_A & 0 \\ I - A^{-1}C_x & 0 & 0 \end{bmatrix} \right\|_R \\ &\quad = \left\| \begin{bmatrix} 0 & 0 & I - A^{-1/2}C_xA^{-1/2} \\ 0 & I - B^{-1/2}S_AB^{-1/2} & 0 \\ I - A^{-1/2}C_xA^{-1/2} & 0 & 0 \end{bmatrix} \right\|_2 \\ &\quad = \rho\left(\begin{bmatrix} 0 & 0 & I - A^{-1/2}C_xA^{-1/2} \\ 0 & I - B^{-1/2}S_AB^{-1/2} & 0 \\ I - A^{-1/2}C_xA^{-1/2} & 0 & 0 \end{bmatrix}\right) \\ &\quad = \max\{\rho_A^f, \rho_S\}. \end{aligned}$$

We conclude analogously to the proof of Theorem 2.3 that

$$\begin{aligned} \rho(I - R^{-1}K) \le \sqrt[3]{\|(I - R^{-1}K)^3\|_R} &\le \underbrace{\sqrt[3]{\theta\|(R^{-1}(\tilde{K} - K))^3\|_R}}_{\kappa} \\ &\le \sqrt[3]{\theta\bar{\varphi}\max\{\rho_A, \rho_S\}} < 1. \end{aligned}$$

∎

In many cases a good choice for A approximating C_x will be available. However, the only part of S_A that is easily accessible is H_p. Therefore, a natural question arises about the usefuleness of just using H_p as an approximation to S_A. For the analysis of this effect, one has to take into account more refined problem characteristics. This is performed in the next section.

3. Application to optimal control. The iterative methods discussed above are of particular importance for the solution of optimal control problems. A generic version of them is the problem

$$\begin{aligned} \min_{x,p} \frac{1}{2}(x-\hat{x}, x-\hat{x})_2 + \frac{\mu}{2}(p,p)_2 \\ \text{s.t.} \qquad L\,x + \Pi p = 0\,, \end{aligned}$$

where $\mu > 0$, $L : H^2(\Omega) \cap H_0^1(\Omega) \to L_2(\Omega)$ is a linear mapping for functions defined on an open region Ω and $(\cdot,\cdot)_2$ is the scalar product in L_2. The operator Π is assumed to be bounded. Discretization, e.g., by finite differences with mesh size h, gives

$$H_x = h^d M_1, \quad H_p = h^d M_1, \quad C_x = L_h = h^{-2} M_2, \quad C_p = \Pi_h,$$

where d is the space dimension of Ω and

$$\|M_1\| \le m_1 < \infty, \qquad \|M_2\| \le m_2 < \infty, \qquad \|\Pi_h\| \le \pi < \infty.$$

For iteration purposes, we may assume that the approximation A to L_h is of the form

$$A = h^{-2} W \quad \text{with} \quad \frac{1}{w} I \le W \le wI$$

for some $w \in \mathbb{R}, w \ge 0$. Then, the "wrong" Schur complement takes the form

$$\begin{aligned} S_A &= \mu h^d M_1 + \Pi_h^\top h^2 W^{-\top} h^d M_1 h^2 W^{-1} \Pi_h \\ &= \mu h^d M_1 + h^{d+4} \Pi_h^\top W^{-\top} M_1 W^{-1} \Pi_h. \end{aligned}$$

The H_p part of the Schur complement is easily available. Therefore, we consider the choice of $B = H_p$ in the Schur complement part of the iterations. The resulting iteration matrix takes the form

$$I - B^{-1} S_A = I - H_p^{-1} S_A = -\frac{h^4}{\mu} \Pi_h^\top W^{-\top} M_1 W^{-1} \Pi_h.$$

Now we can make the following observations in Corollary 3.1.

COROLLARY 3.1. *For discretized optimal control problems with the characteristics described above, the convergence rate ρ_S for the Schur complement can become arbitrarily small if B is chosen as $B = H_p$ and μ is large enough or the discretization h is fine enough.*

Proof. For the norm $\|.\| = (.,.)^{1/2}$ we see that

$$\begin{aligned} \rho_S = \rho(I - H_p^{-1} S_A) &\le \left\| \frac{h^4}{\mu} \Pi_h^\top W^{-\top} M_1 W^{-1} \Pi_h \right\| \\ &\le \frac{h^4}{\mu} \pi^2 w^2 m_1. \end{aligned}$$

Therefore, $\rho_S < 1$ if μ is large enough or h is small enough. ∎

Now, we can easily achieve $\rho_S < 1$. The forward and adjoint system can also be assumed to be solvable with $\rho_A < 1$. If we take a close look at Theorem 2.3 and Corollary 2.4, we see that these properties for ρ_S, ρ_A are not enough to guarantee overall convergence. At least in Corollary 2.4, we observe that $\bar{\varphi}$ is close to 1 if h is small enough or μ is large enough. However, the parameter θ increases noticeably for decreasing h, as observed in numerical experiments below. Therefore, for small h the conditions for ρ_S, ρ_A become very restrictive in order to be able to apply the convergence Theorems 2.3 and 2.4. But the numerical results below show that convergence is also achieved in cases where Theorem 2.3 and Corollary 2.4 are not applicable.

4. Algorithmic reformulation. For comparison with other iterations and for generalization to nonlinear problems, it is useful to write iteration (1.3) using the notation (QP). We define the Lagrangian as

$$L(x,p,\lambda) := \frac{1}{2}x^\top H_x\, x + x^\top H_{xp}\, p + p^\top H_{px}\, x + p^\top H_p\, p + f_x^\top x + f_p^\top p + \lambda^\top \left(C_x x + C_p p + c\right).$$

Then the necessary optimality conditions are

$$\begin{aligned}
\nabla_x L(x,p,\lambda) &= H_x x + H_{xp} p + C_x^\top \lambda + f_x = 0,\\
\nabla_p L(x,p,\lambda) &= H_{px} x + H_p p + C_p^\top \lambda + f_p = 0,\\
\nabla_\lambda L(x,p,\lambda) &= C_x x + C_p p + c = 0.
\end{aligned}$$

Iteration (1.3) can now be rewritten in a more compact form as

$$\lambda^{k+1} = \lambda^k - A_a^{-1}\nabla_x L(x^k, p^k, \lambda^k), \tag{4.1}$$

$$p^{k+1} = p^k - B^{-1}\nabla_p L(x^k, p^k, \lambda^{k+1}), \tag{4.2}$$

$$x^{k+1} = x^k - A_f^{-1}\nabla_\lambda L(x^k, p^{k+1}, \lambda^{k+1}). \tag{4.3}$$

In this way, the iteration can be directly applied to nonlinear optimization problems that use a nonlinear iteration for the forward problem, i.e., solve the equality constraint in (QP) for x, in this case more appropriately written in the form $c(x,p) = 0$. This approach has been applied successfully in [HS04, HSBG05, HSB08] for aerodynamic shape optimization.

Note in iteration (4.1)–(4.3) the usage of novel information as soon as it is available. This is the principal difference between iteration (4.1)–(4.3) and the so-called piggyback iterations as proposed in [GF02, Gri06]. Because piggyback iterations are derived from automatic differentiation, where functions and derivatives are evaluated simultaneously, the piggyback iteration for problem (QP) is written in this notation as

$$\begin{aligned}
\lambda^{k+1} &= \lambda^k - A_a^{-1}\nabla_x L(x^k, p^k, \lambda^k),\\
p^{k+1} &= p^k - B^{-1}\nabla_p L(x^k, p^k, \lambda^k),\\
x^{k+1} &= x^k - A_f^{-1}\nabla_\lambda L(x^k, p^k, \lambda^k),
\end{aligned}$$

or in matrix notation as

$$\begin{bmatrix} 0 & 0 & A_a \\ 0 & B & 0 \\ A_f & 0 & 0 \end{bmatrix} \begin{pmatrix} \Delta x^k \\ \Delta p^k \\ \Delta \lambda^k \end{pmatrix} = - \begin{bmatrix} H_x & H_{xp} & C_x^\top \\ H_{px} & H_p & C_p^\top \\ C_x & C_p & 0 \end{bmatrix} \begin{pmatrix} x^k \\ p^k \\ \lambda^k \end{pmatrix} - \begin{pmatrix} f_x \\ f_p \\ c \end{pmatrix}, \tag{4.4}$$

where the appropriate definition of the design preconditioner B is discussed in [GF02, Gri06].

5. Numerical experiments. The purpose of this section is to illustrate the theoretical results from the previous section by application to a common model problem, which serves as a standard problem in PDE-constrained optimization. We do not claim that this problem is in any way challenging. The sole purpose of this study is to highlight certain features of the iterative method discussed above and its theoretical properties.

For a given open computational region Ω, we investigate the problem

$$\begin{aligned} \min_{x,p} \frac{1}{2}\int_\Omega (x(\xi)-\bar{x}(\xi))^2 d\xi + \frac{\mu}{2}\int_\Omega p(\xi)^2 d\xi \\ \text{s.t.} \qquad -\triangle x(\xi) = p(\xi), \quad \forall \xi \in \Omega, \\ x(\xi) = 0 \quad \forall \xi \in \partial\Omega. \end{aligned}$$

The variables x and p are functions defined on the domain Ω, and $\triangle$ denotes the Laplacian operator. The aim of the problem is to track a given function $\bar{x}$ with the solution of a differential equation. Since we aim only to illustrate certain numerical effects and do not pretend to attack any real-life problem in this paper (this has been done and will be done again by the authors in different publications), we choose $\Omega = [0,1]$. Then this model problem simplifies to

$$\begin{aligned} \min_{x,p} \frac{1}{2}\int_0^1 (x(\xi)-\bar{x}(\xi))^2 d\xi + \frac{\mu}{2}\int_0^1 p(\xi)^2 d\xi \\ \text{s.t.} \qquad -x''(\xi) = p(\xi), \quad \forall \xi \in [0,1], \\ x(0) = 0 = x(1). \end{aligned}$$

The function $\bar{x}$ is chosen as (cf. Figure 5.1 later in this article)

$$\bar{x}(\xi) = \begin{cases} 0.8 - \xi, & 0 \le \xi \le 0.4, \\ -2.6 + 2\xi, & 0.4 < \xi \le 1. \end{cases}$$

This problem is discretized by finite differences on a regular mesh with mesh size $h = 1/(N-1)$, where $N = n_x$:

$$\begin{aligned} x_\ell &:= x(\ell h), \quad \ell = 0, \dots, N, \\ p_\ell &:= p(\ell h), \quad \ell = 0, \dots, N, \\ -x''(\ell h) &\approx \frac{1}{h^2}(-x_{\ell-1} + 2x_\ell - x_{\ell+1}), \quad \ell = 1, \dots, N-1, \\ \int_0^1 (x(\xi)-\bar{x}(\xi))^2 d\xi &\approx h\sum_{\ell=1}^{N-1} (x_\ell - \bar{x}(\ell h))^2, \\ \int_0^1 p(\xi)^2 d\xi &\approx h\sum_{\ell=1}^{N-1} p(\xi)^2. \end{aligned}$$

For the sake of simplicity, we omit values at 0 and 1 so that our vectors of unknowns are

$$x = (x_1, \dots, x_{N-1})^\top, \qquad p = (p_1, \dots, p_{N-1})^\top.$$

The discretized problem is now of the form (QP) with

$$H_x = hI, \quad H_{xp} = 0 = H_{px}^\top, \quad H_p = \mu h I, \qquad C_p = I, \qquad f_x = -\bar{x}, \quad f_p = 0, \quad c = 0,$$

where I is the identity in $\mathbb{R}^{N-1}$ and

$$C_x = \frac{1}{h^2}\begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix}.$$

In order to perform numerical convergence tests with varying approximations A to C_x, we construct these approximations by Jacobi steps, i.e., for $D := \operatorname{diag}(C_x)$ we define

$$A_0^{-1} := D^{-1}, \tag{5.1}$$

$$A_1^{-1} := D^{-1}(I + (I - C_x D^{-1})), \tag{5.2}$$

$$A_2^{-1} := D^{-1}(I + (I + (I - C_x D^{-1}))(I - C_x D^{-1})), \tag{5.3}$$

$$\vdots \quad \vdots$$

and therefore

$$A_i^{-1} = A_0^{-1}((A_0 - C_x)A_{i-1}^{-1} + I).$$

Thus

$$\begin{aligned} 1 > \rho(I - A_1^{-1}C_x) > \rho(I - A_2^{-1}C_x) = \rho(I - A_1^{-1}C_x)^2 > \rho(I - A_3^{-1}C_x) \\ = \rho(I - A_1^{-1}C_x)^3 > \cdots . \end{aligned}$$

In the case of A_0, we can give the Schur complement analytically:

$$S_{A_0} = H_p + C_p^\top A_0^{-\top} H_x A_0^{-1} C_p = \left(\mu h + \frac{h^5}{4}\right) I. \tag{5.4}$$

In all other cases, the formulas become more complicated. We treat B analogously: we choose $B_0 := H_p$ and B_j^{-1} as the approximation to S_A^{-1} after j Richardson iterations with H_p. That means

$$B_0 := H_p = \mu h I, \tag{5.5}$$

$$B_j^{-1} := B_0^{-1}(I - C_p^\top A_i^{-\top} H_x A_i^{-1} C_p B_{j-1}^{-1}). \tag{5.6}$$

Additionally, we investigate the cases $B = S_A$ and $B = S$ from (1.1).

The norms to be used later are chosen as

$$\left\| \begin{pmatrix} x \\ p \\ \lambda \end{pmatrix} \right\| := \left(\|x\|^2 + \|p\|^2 + \|\lambda\|^2 \right)^{1/2} ,$$

where $\|x\| := (h \sum_{\ell=1}^{N-1} x_\ell^2)^{1/2}$ with analogous definitions for p and λ. This norm is the discrete approximation to the continuous $L^2(\Omega)$-norm.

We investigate for the setting $N = 101$ ($h = \frac{1}{100}$) and $\mu = 0.001$ the convergence rates of Corollary 2.4. We perform the approximate nullspace iterations until

$$\left\| \begin{bmatrix} H_x & H_{xp} & C_x^\top \\ H_{px} & H_p & C_p^\top \\ C_x & C_p & 0 \end{bmatrix} \begin{pmatrix} x^k \\ p^k \\ \lambda^k \end{pmatrix} + \begin{pmatrix} f_x \\ f_p \\ c \end{pmatrix} \right\| \Big/ \left\| \begin{pmatrix} f_x \\ f_p \\ c \end{pmatrix} \right\| \leq 10^{-3} .$$

The problem solution is plotted in Figure 5.1.

Table 5.1 summarizes the results. In column i, we denote how many Jacobi iterations are performed for obtaining A_i^{-1} in (5.1)–(5.3). Column j gives analogous information for B_j^{-1}. The notation $j = S_A$ means that we choose $B = S_A$, and $j = S$ means that we choose $B = S$. Furthermore, ρ_A denotes the spectral radius of the matrix $I - A_i^{-1}C_x$. Analogously, ρ_S denotes the spectral radius of the iteration matrix of the Schur complement (cf. (5.6)), and ρ_{It} denotes the spectral radius of the overall

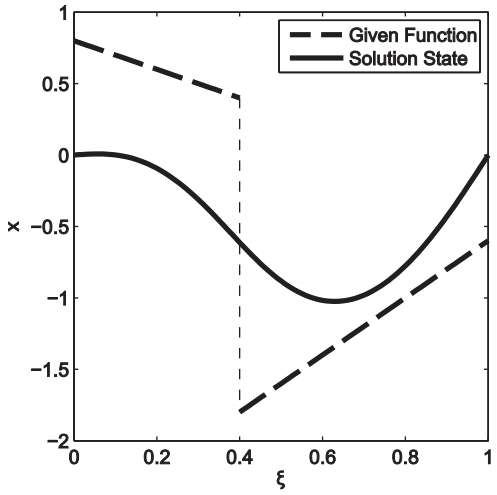

FIG. 5.1. *Solution (state, solid line) and function $\bar{x}$ (dashed line) to be tracked.*

TABLE 5.1
Convergence results for $N = 101$ and $\mu = 0.001$.

i	j	ρ_A	ρ_S	ρ_{It}	# It.
0	0	0.9995	0.0000	1.0011	—
0	1	0.9995	0.0000	1.0011	—
0	3	0.9995	0.0000	1.0011	—
0	S_A	0.9995	0.0000	1.0011	—
0	S	0.9995	0.9113	1.0011	—
3	0	0.9980	0.0000	0.9980	2483
3	1	0.9980	0.0000	0.9980	2483
3	3	0.9980	0.0000	0.9980	2483
3	S_A	0.9980	0.0000	0.9980	2483
3	S	0.9980	0.9112	0.9982	3461
5	0	0.9970	0.0001	0.9970	2317
5	1	0.9970	0.0000	0.9970	2317
5	3	0.9970	0.0000	0.9970	2317
5	S_A	0.9970	0.0000	0.9970	2317
5	S	0.9970	0.9112	0.9975	2963

iteration matrix $(I - R^{-1}K)$. In the last column "# It." refers to the number of iterations (1.3).

We observe that the KKT iteration does not converge for the choice $A = D$ $(i = 0)$, although the iterations in the components (forward, adjoint, design) are convergent. This is in line with Theorem 2.3 and Corollary 2.4, which state that each spectral radius of the components has to be below a certain limit, perhaps significantly below 1, in order to guarantee convergence of the overall KKT iteration.

Furthermore, we observe the effect of choosing $B = S$: the convergence deteriorates, i.e., the choice $B = S_A$ (and $B \approx S_A$) leads always to better convergence properties than $B \approx S$, resulting also in smaller iteration numbers.

6. Conclusions. The aim of this paper is to investigate defect-correcting iterations of the type (1.3) for the solution of linear-quadratic optimization problems. Theoretical foundations for iterations, well established in practice, including Jacobi

and Gauss–Seidel iterations [HA93] in the context of potentially large-scale problems, are given, and the following facts are established:

- Iteration (1.3) is convergent if (A_f, A_a, B) are close enough to $(C_x, C_x^\top, S_A)$.
- The matrix B, the preconditioner for the Schur complement, should be chosen close to S_A rather than S.

Acknowledgments. The third author wishes to thank the University of Graz for providing support for his stay, during which most of the ideas in this paper were developed. Furthermore, the authors thank Andreas Griewank for stimulating discussions on several issues of this paper, in particular the comparison with the piggy-back approach. Also, the authors are grateful to Omar Ghattas for feedback on the paper in manuscript version, which helped to improve the presentation of the material. Finally, the authors are grateful for many detailed and helpful hints from the referees.

REFERENCES

[BWY90] R. Bank, B. Welfert, and H. Yserentant, *A class of iterative methods for solving saddle point problems*, Numer. Math., 56 (1990), pp. 645–666.

[BS01] A. Battermann and E. W. Sachs, *Block preconditioners for KKT systems in PDE-governed optimal control problems*, in Fast Solution of Discretized Optimization Problems, Internat. Ser. Numer. Math. 138, K.-H. Hoffmann, R. H. W. Hoppe, and V. Schulz, eds., Birkhäuser, Basel, Switzerland, 2001, pp. 1–18.

[BG05] G. Biros and O. Ghattas, *Parallel Lagrange–Newton–Krylov–Schur methods for PDE-constrained optimization. Part* I: *The Krylov–Schur solver*, SIAM J. Sci. Comput., 27 (2005), pp. 687–713.

[CI97] G. Cho and I. Ipsen, *If a Matrix Has Only a Single Eigenvalue, How Sensitive Is This Eigenvalue?*, Technical report CRSC-TR97-20, Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, Raleigh, NC, 1997.

[DW06] H. S. Dollar and A. J. Wathen, *Approximate factorization constraint preconditioners for saddle-point matrices*, SIAM J. Sci. Comput., 27 (2006), pp. 1555–1572.

[FN94] R. W. Freund and N. M. Nachtigal, *QMR: A quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.

[Gri06] A. Griewank, *Projected Hessians for preconditioning in one-step one-shot design optimization*, in Large-Scale Nonlinear Optimization, G. DiPilb and M. Roma, eds., Springer, New York, 2006, pp. 151–172.

[GF02] A. Griewank and C. Faure, *Reduced functions, gradients and Hessians from fixed point iterations for state equations*, Numer. Algorithms, 30 (2002), pp. 113–139.

[GS05] I. Gherman and V. Schulz, *Preconditioning of one-shot pseudo-timestepping methods for shape optimization*, Proc. Appl. Math. Mechanics, 5 (2005), pp. 741–742.

[HA01] E. Haber and U. Ascher, *Preconditioned all-at-once methods for large, sparse parameter estimation problems*, Inverse Problems, 17 (2001), pp. 1847–1864.

[HA93] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Appl. Math. Sci. 95, Springer, Berlin, 1993.

[HS04] S. B. Hazra and V. Schulz, *Simultaneous pseudo-timestepping for PDE-model based optimization problems*, BIT, 44 (2004), pp. 457–472.

[HSB08] S. B. Hazra, V. Schulz, and J. Brezillon, *Simultaneous pseudo-time stepping for* 3*D aerodynamic shape optimization*, J. Numer. Math., 16 (2008), pp. 139–161.

[HSBG05] S. B. Hazra, V. Schulz, J. Brezillon, and N. Gauger, *Aerodynamic shape optimization using simultaneous pseudo-timestepping*, J. Comput. Phys., 204 (2005), pp. 46–64.

[PS75] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[SS86] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.