TECHNICAL RESEARCH REPORT

Globally Convergent Algorithms for Robust Pole Assignment by State Feedback

by A.L. Tits and Y. Yang

T.R. 95-3

INSTITUTE FOR SYSTEMS RESEARCH

Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

Globally Convergent Algorithms for Robust Pole Assignment by State Feedback

André L. Tits and Yaguang Yang
Department of Electrical Engineering
and Institute for Systems Research
University of Maryland at College Park
College Park, MD 20742 USA

January 11, 1995

Abstract

It is observed that an algorithm proposed in 1985 by Kautsky, Nichols and Van Dooren (KNV) amounts to maximize, at each iteration, the determinant of the candidate closed-loop eigenvector matrix X with respect to one of its columns (with unit length constraint), subject to the constraint that it remains a valid closed-loop eigenvector matrix. This interpretation is used to prove convergence of the KNV algorithm. It is then shown that a more efficient algorithm is obtained if $\det(X)$ is concurrently maximized with respect to two columns of X, and that such a scheme is easily extended to the case where the eigenvalues to be assigned include complex conjugate pairs. Variations exploiting the availability of multiple processors are suggested. Convergence properties of the proposed algorithms are established. Their superiority is demonstrated numerically.

1 Introduction

Given $A \in \mathbf{R}^{n\times n}$ and $B \in \mathbf{R}^{n\times m}$, the state feedback pole assignment problem is that of determining a "feedback" matrix $F \in \mathbf{R}^{m\times n}$ such that the eigenvalues of $A+BF$ ("closed-loop" eigenvalues) lie at prescribed locations. In the case of more than one input ($m > 1$), the solution to this problem is generally non-unique. Many authors have investigated using the available degrees of freedom to achieve low sensitivity of the closed-loop eigenvalues to perturbations in A, B, and F [1–8]. This is often referred to as the robust pole assignment problem. An early contribution to robust pole assignment was that of Gourishanker and Ramer [1,2], who proposed heuristics to minimize the sensitivity of the closed-loop eigenvalues with respect to specified perturbations in the entries of A. Later, Cavin and Bhattacharyya [3] suggested to minimize the trace of $(I-X^\top X)^2$, where X is the closed-loop eigenvector matrix, Dickman [4] proposed to use the sum of the squares of the off-diagonal entries in the Schur form of X as robustness measure to be minimized, and Sun [7] adopted as cost functions two measures of robustness suggested by Kautsky *et al.* [9] (which in turn were inspired from Wilkinson [10]). Among these authors, Sun is the only one who provided a convergence analysis: he showed that one of his algorithms generates a sequence whose limit points are all stationary in a certain sense.

To date, the most computationally attractive methods are perhaps those proposed by Kautsky *et al.* [9]. Unlike most other methods, these algorithms require very few arithmetic operations at each iteration. The present work is directly inspired from Kautsky *et al.*'s 1985 paper [9]. In [9], four algorithms are proposed for the robust pole assignment problem. These algorithms are based on the fact that sensitivity of the eigenvalues of a non-defective matrix to perturbations in its entries is directly related to the condition number of the associated eigenvector matrix. Kautsky *et al.* first show that X is an eigenvector matrix of $A + BF$ for some F that assigns the prescribed eigenvalues if and only if the columns x_i of X belong to certain subspaces $\mathcal{S}_i$. Accordingly, all four of their algorithms successively modify the columns of X, repeatedly sweeping through the set of columns, in an attempt to make these columns "closer to orthogonal" to each other, while forcing them to remain in the respective $\mathcal{S}_i$'s. In their first algorithm ("Method 0"), conceptually the simplest among the four, each column x_i of X is successively made closest to the orthogonal complement of the subspace spanned by the other $n-1$ columns, while being forced to remain in $\mathcal{S}_i$. The operation count per iteration is low, and numerical experiments discussed in [9] indicate that Method 0 performs as well as or better than the other proposed algorithms ("Methods 1 to 3") on difficult problems. However, it is also observed in [9] that, on some problems, Method 0 ostensibly fails to converge (but good results are obtained after a few iterations).

The contribution of the present paper is as follows. First we observe that Method 0 of [9] amounts to maximize, at each iteration, the determinant of the candidate closed-loop eigenvector matrix X with respect to one of its columns (with unit length constraint), subject to the constraint that it remain a valid closed-loop eigenvector matrix. Using this interpretation we prove convergence of Method 0 of [9] (the "lack of convergence" observed in [9] turns out to be an instance of slow convergence). It is then shown that a more efficient algorithm is obtained if $\det(X)$ is concurrently maximized with respect to *two* columns of X, and that such a scheme is easily extended to the case where the eigenvalues to be assigned include complex conjugate pairs. Variations exploiting the availability of multiple processors are suggested. Convergence properties of the new algorithms are established. The superiority of the new algorithms is demonstrated on a large sample of numerical examples.

The remainder of this paper is organized as follows. In Section 2 we review Kautsky *et al.*'s Method 0. Section 3 is devoted to a convergence analysis for this method. In Section 4, we present the new algorithms for the case of real pole assignment and analyze their convergence. Section 5 discusses the extension to complex pole assignment. Section 6 addresses implementation issues. Section 7 presents some numerical results. Finally Section 8 is devoted to concluding remarks.

2 Preliminaries

Given a positive integer q, we make use of the function $i_q(\cdot) : \{1, 2, \ldots\} \to \{1, 2, \ldots, q\}$ defined by the equivalence

$$i_q(p) \equiv p \pmod{q}.$$

Given two complex vectors u and v and a complex matrix M, we denote by σ_i the ith largest singular value of M, by $\Re e(M)$ (resp. $\Im m(M)$) the real matrix whose entries are the real (resp. imaginary) parts of the corresponding entries of M, by $u^\top$ and u^* the transpose and conjugate transpose of u and by $M^\top$ and M^* those of M, by $\langle u, v\rangle$ the complex inner product u^*v (u and v are said to be orthogonal if $\langle u, v\rangle = 0$), by $\|u\|$ the Euclian norm $\sqrt{\langle u, u\rangle}$ of u and by $\|M\|$ the Frobenius norm $\sqrt{\mathrm{Tr}(M^*M)}$ of M, where Tr denotes the trace. Next, given a normed vector space $\mathcal{V}$, we denote by $\mathbf{B}\mathcal{V}$ its unit sphere, i.e.,

$$\mathbf{B}\mathcal{V} \equiv \{\mathbf{x} : \mathbf{x} \in \mathcal{V}, \|\mathbf{x}\| = \mathbf{1}\}.$$

Finally, we denote by j (non-italics) the square root of -1.

Let $A \in \mathbf{R}^{n\times n}$, $B \in \mathbf{R}^{n\times m}$ and let $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$ be the desired closed-loop eigenvalue locations. A key role will be played by subspaces $\mathcal{S}_i \subset \mathbf{C}^n$, $i = 1, \ldots, n$, defined by

$$\mathcal{S}_i = \{x : (A - \lambda_i I)x \in \mathcal{R}_{\mathbf{C}}(B)\}, \tag{1}$$

where $\mathcal{R}_{\mathbf{C}}(B) \subset \mathbf{C}^n$ is the "complex range" of B, i.e.,

$$\mathcal{R}_{\mathbf{C}}(B) = \{By : y \in \mathbf{C}^m\},$$

and by their real parts

$$\mathcal{S}_i^{\mathbf{R}} = \{x \in \mathbf{R}^n : x \in \mathcal{S}_i\}. \tag{2}$$

In this and the next two sections, we focus on the case where all λ_i's are real. The following result is used in [9].

Theorem 2.1 *Let $X = [x_1, \ldots, x_n]$, nonsingular, and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ be two real $n \times n$ matrices. Then*

$$(A + BF)X = X\Lambda \tag{3}$$

for some real matrix F if and only if $x_i \in \mathcal{S}_i^{\mathbf{R}}$, $i = 1, \ldots, n$. □

Given a suitable X, the corresponding feedback matrix F is readily obtained via QR factorization of B (see [9]).

Thus there exists a real F such that $A + BF$ is similar to Λ (by similarity transformation) if and only if there exists $X = [x_1, \ldots, x_n]$ nonsingular such that $x_i \in \mathcal{S}_i, i = 1, \ldots, n$. These conditions may fail to hold only if either (A, B) is not controllable or every $A + BF$ with the required eigenvalues is defective. In [9], necessary conditions are given for the existence of a nondefective $A + BF$ with the required eigenvalues. We simply note here that a suitable F does exist whenever (A, B) is controllable and the λ_i's are distinct. From now on we assume that a nonsingular matrix X satisfying (3) can indeed be obtained.

Clearly, it is enough to consider matrices X whose columns have unit length and whose determinant is strictly positive. Thus, let

$$\mathbf{X}_{\mathbf{R}} = \{X = [x_1, \ldots, x_n] \in \mathbf{R}^{n\times n} : x_i \in \mathbf{B}\mathcal{S}_i^{\mathbf{R}},\ \ i = 1, \ldots, n,\ \ \det(X) > 0\}.$$

To robustly assign the desired poles Method 0 in [9] makes use of iteration maps $f_i : \mathbf{X}_{\mathbf{R}} \to \mathbf{X}_{\mathbf{R}}$, $i = 1, 2, \ldots, n$, defined essentially[1] as follows. Given $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n] \in \mathbf{X}_{\mathbf{R}}$,

$$f_i(\hat{X}) = [\hat{x}_1, \ldots, \hat{x}_{i-1}, \tilde{\xi}, \hat{x}_{i+1}, \ldots, \hat{x}_n],$$

where $\tilde{\xi} \in \mathbf{R}^n$ solves

$$\text{maximize } \ \langle \xi, u_i(\hat{X})\rangle \ \text{ s.t. } \ \xi \in \mathbf{B}\mathcal{S}_i^{\mathbf{R}}, \tag{4}$$

where $u_i(\hat{X}) \in \mathbf{R}^n$ is the unit length vector orthogonal to $\{\hat{x}_1, \ldots, \hat{x}_{i-1}, \hat{x}_{i+1}, \ldots, \hat{x}_n\}$ such that $\langle \hat{x}_i, u_i(\hat{X})\rangle > 0$. Thus the ith column of $f_i(\hat{X})$ is a unit vector along the projection of $u_i(\hat{X})$ on $\mathcal{S}_i$ and thus, in some sense, is as close as possible to being orthogonal to the other columns of $\hat{X}$. Note that since $\hat{X}$ is nonsingular, the direction and the orientation of $u_i(\hat{X})$ are uniquely determined, the solution $\tilde{\xi}$ to (4) is unique, and, since $\langle \tilde{\xi}, u_i(\hat{X})\rangle \geq \langle \hat{x}_i, u_i(\hat{X})\rangle > 0$, $f_i(\hat{X})$ is nonsingular. The iteration proposed in [9] is essentially as follows.

Algorithm 2.1 (KNV Algorithm):

[1] In [9] the sign of $\det(X^0)$ and the orientation of the new column of X^k are left unspecified. Our choice forces $\det(X^k)$ to have constant sign, which allows a notationally simpler convergence analysis.

Step 1. Pick $X^0 \in \mathbf{X_R}$. Set $k := 1$.

Step 2. Set $X^k := f_{i_n(k)}(X^{k-1})$.

Step 3. Set $k := k+1$ and go to Step 2.

□

Computation of $f_i(\hat{X})$ given $\hat{X}$ is inexpensive. The main task, computation of $u_i(\hat{X})$, can be effected by means of a QR factorization of $[\hat{x}_1, \ldots, \hat{x}_{i-1}, \hat{x}_{i+1}, \ldots, \hat{x}_n]$. Kautsky *et al.* [9] note that the QR factorization to be carried out at iteration $k > 0$ can be obtained by a rank-one update of that computed at iteration $k-1$, requiring only $O(n^2)$ operations; and that the subsequent projection on $\mathcal{S}_i$ that yields the ith column of $f_i(\hat{X})$ requires $O(nm)$ operations (if all $\mathcal{S}_i$'s have dimension m), for a total of $O(n^2) + O(mn)$ operations per iteration.

In the next section, we show that, for the sequence $\{X^k\}$ constructed by Algorithm 2.1, $\{\det(X^k)\}$ is nondecreasing and every limit point $\check{X}$ of $\{X^k\}$ is a stationary point for the problem

$$\text{maximize } \det(X) \quad \text{s.t. } X \in \mathbf{X_R}, \tag{$\mathcal{P}$}$$

in the sense that

$$\langle\langle \frac{\partial \det(\check{X})}{\partial X}, H \rangle\rangle = 0, \quad \forall H \in T(\check{X})$$

where $T(\check{X})$ is the tangent plane to $\mathbf{X_R}$ at $\check{X}$, and $\langle\langle \frac{\partial \det(\check{X})}{\partial X}, H \rangle\rangle$ is the directional derivative of $\det(\cdot)$ at $\check{X}$ in direction H, i.e.,

$$\langle\langle \frac{\partial \det(\check{X})}{\partial X}, H \rangle\rangle = \operatorname{Tr}\left(\frac{\partial \det(\check{X})^\top}{\partial X} H \right).$$

Use of $\det(X)$ as a robustness indicator was suggested in [11]. Relationships between determinant and condition number are explored in [12] and [13].

3 Convergence of the KNV Algorithm

First, we show that, given $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n] \in \mathbf{X_R}$, and given $i \in \{1, \ldots, n\}$, $f_i(\hat{X})$ is the unique maximizer for

$$\text{maximize } \det(X) \;\; \text{s.t. } X = [\hat{x}_1, \ldots, \hat{x}_{i-1}, \xi, \hat{x}_{i+1}, \ldots, \hat{x}_n],\; \xi \in \mathbf{B}\mathcal{S}_i^{\mathbf{R}}. \tag{$\mathcal{P}^i(\hat{X})$}$$

The key is the following lemma.

Lemma 3.1 *Let* $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n] \in \mathbf{R}^{n \times n}$ *be such that* $\det(\hat{X}) > 0$, *let* $\xi \in \mathbf{R}^n$, *let* $i \in \{1, 2, \ldots, n\}$, *and let*

$$X = [\hat{x}_1, \cdots, \hat{x}_{i-1}, \xi, \hat{x}_{i+1}, \cdots, \hat{x}_n],$$

and

$$X_- = [\hat{x}_1, \cdots, \hat{x}_{i-1}, \hat{x}_{i+1}, \cdots, \hat{x}_n].$$

Then $X_-^\top X_-$ *is nonsingular and*

$$\det(X) = \langle \xi, u_i(X) \rangle \sqrt{\det(X_-^\top X_-)}. \tag{5}$$

Proof: Let P be the permutation matrix such that

$$XP = [\xi, X_-], \tag{6}$$

and let $Q \in \mathbf{R}^{n\times(n-1)}$ and $R \in \mathbf{R}^{(n-1)\times(n-1)}$ be any two matrices such that $Q^\top Q = I$ and

$$X_- = QR. \tag{7}$$

Using (6) and (7), we obtain

$$\begin{aligned}
(\det(X))^2 &= \det(X^\top X) \\
&= \det(P^\top X^\top XP) = \det([\xi, X_-]^\top[\xi, X_-]) \\
&= \det\left(\begin{bmatrix} \xi^\top\xi & \xi^\top X_- \\ X_-^\top\xi & X_-^\top X_- \end{bmatrix}\right) \\
&= (\xi^\top\xi - \xi^\top X_-(X_-^\top X_-)^{-1}X_-^\top\xi)\det(X_-^\top X_-) \\
&= (\xi^\top\xi - \xi^\top QR(R^\top R)^{-1}R^\top Q^\top\xi)\det(X_-^\top X_-) \\
&= (\xi^\top\xi - \xi^\top QQ^\top\xi)\det(X_-^\top X_-).
\end{aligned} \tag{8}$$

Now note that $[Q, u_i(X)]$ is an orthogonal matrix so that

$$QQ^\top + u_i(X)u_i(X)^\top = I.$$

Thus

$$\begin{aligned}
(\det(X))^2 &= \xi^\top u_i(X)u_i(X)^\top\xi\det(X_-^\top X_-) \\
&= \langle\xi, u_i(X)\rangle^2\det(X_-^\top X_-).
\end{aligned} \tag{9}$$

Finally note that $\langle\xi, u_i(X)\rangle$ and $\det(X)$ have the same sign since (i) they are both linear in ξ, (ii) in view of (9) they vanish simultaneously and (iii) they are both positive at $\xi = x_i$. The claim follows. □

Our next result relates problem $(\mathcal{P}^i(\hat{X}))$ to the KNV Algorithm.

Proposition 3.1 *For any nonsingular $\hat{X}$ and any $i \in \{1, \ldots, n\}$, $\mathcal{P}^i(\hat{X})$ has a unique maximizer X^+ given by $X^+ = f_i(\hat{X})$.*

Proof: The claim directly follows from Lemma 3.1 and the definitions of $f_i(\hat{X})$ and $\mathcal{P}^i(\hat{X})$. □

Proposition 3.2 *For $k = 1, 2, \ldots$, X^k constructed by Algorithm 2.1 is the unique maximizer of $\mathcal{P}^{i_n(k)}(X^{k-1})$,*

$$0 < \det(X^0) \leq \det(X^1) \leq \cdots$$

and $\{\det(X^k)\}$ converges to a finite limit. Moreover, if there exists k_0 such that

$$\det(X^{k_0+n}) = \det(X^{k_0}), \tag{10}$$

then

$$X^k = X^{k_0} \qquad \forall k \geq k_0.$$

Proof: The first claim follows immediately from Proposition 3.1. The second claim follows from Proposition 3.1 and feasibility of X^{k-1} for $(\mathcal{P}^{i_n(k)}(X^{k-1}))$. The third claim follows from the second claim and the fact that, since $\|x_i^k\| = 1$, for $i = 1, \ldots, n$, $\det(X^k) \leq 1$ for all k. The last claim follows from the uniqueness of the global maximizer of $(\mathcal{P}^i(X^{k_0}))$ for any i. □

Lemma 3.2 *For $k = 1, 2, \ldots$, let $\{X^k\}$ be the sequence constructed by Algorithm 2.1. Then the sequence $\{X^k - X^{k-1}\}$ goes to zero as k goes to infinity.*

Proof: For $k = 1, 2, \ldots$, write X^k as $[x_1^k, \ldots, x_n^k]$, and let $X_-^k \in \mathbf{R}^{n\times(n-1)}$ be given by

$$X_-^k = [x_1^{k-1}, x_2^{k-1}, \cdots, x_{i_n(k)-1}^{k-1}, x_{i_n(k)+1}^{k-1}, \cdots, x_n^{k-1}], \tag{11}$$

and $u^k \in \mathbf{R}^n$ be the unique (in view of Proposition 3.2) unit vector satisfying

$$(X_-^k)^T u^k = 0, \quad \langle x_{i_n(k)}^{k-1}, u^k \rangle > 0, \tag{12}$$

i.e., $u^k = u_{i_n(k)}(X^{k-1})$. Also, for $i = 1, \ldots, n$, let P_i be the orthogonal projection matrix onto $\mathcal{S}_i$ and, for $k = 1, 2, \ldots$, let $\alpha_k = \|P_{i_n(k)} u^k\|^{-1}$. Thus

$$x_{i_n(k)}^k = \alpha^k P_{i_n(k)} u^k. \tag{13}$$

First, we claim that the sequence $\{\sqrt{\det(X_-^{k^T} X_-^k)}\}$ is bounded away from zero and the sequence $\{\alpha^k\}$ is bounded. Indeed, since for $k = 1, 2, \ldots$, u^k and all columns of X_-^k have unit length, $\sqrt{\det(X_-^{k^T} X_-^k)} \leq 1$ and $|\langle x_{i_n(k)}^k, u^k \rangle| \leq 1$ for all k. Since, in view of Proposition 3.2, $\det(X^k)$ is bounded away from zero, it follows from Lemma 3.1 that $\sqrt{\det(X_-^{k^T} X_-^k)}$ and $\langle x_{i_n(k)}^k, u^k \rangle$ are both bounded away from zero. Since $\|P_{i_n(k)} u^k\| = \langle x_{i_n(k)}^k, u^k \rangle$, our first claim follows. Since, in view of Lemma 3.1 and Proposition 3.2,

$$\lim_{k\to\infty} \left(\langle x_{i_n(k)}^k, u^k \rangle - \langle x_{i_n(k)}^{k-1}, u^k \rangle \right) \sqrt{\det(X_-^{k^T} X_-^k)} = \lim_{k\to\infty} (\det(X^k) - \det(X^{k-1})) = 0,$$

it follows that

$$\lim_{k\to\infty} \left(\langle x_{i_n(k)}^k, u^k \rangle - \langle x_{i_n(k)}^{k-1}, u^k \rangle \right) = 0,$$

i.e., since $x_{i_n(k)}^{k-1} = x_{i_n(k)}^{k-n}$ and since $i_n(k) = i_n(k-n)$,

$$\lim_{k\to\infty} \left(\langle \alpha^k P_{i_n(k)} u^k, u^k \rangle - \langle \alpha^{k-n} P_{i_n(k)} u^{k-n}, u^k \rangle \right) = 0.$$

Since α^k is bounded and $P_{i_n(k)}$ is a projection matrix, it follows that

$$\lim_{k\to\infty} \left(\langle (\alpha^k)^2 P_{i_n(k)}^2 u^k, u^k \rangle - \langle \alpha^{k-n} \alpha^k P_{i_n(k)}^2 u^{k-n}, u^k \rangle \right) = 0,$$

i.e.,

$$\lim_{k\to\infty} \langle x_{i_n(k)}^k - x_{i_n(k)}^{k-1}, x_{i_n(k)}^k \rangle = 0.$$

Since

$$\begin{aligned} 1 &= \|x_{i_n(k)}^{k-1}\|^2 = \|x_{i_n(k)}^k + (x_{i_n(k)}^{k-1} - x_{i_n(k)}^k)\|^2 \\ &= 1 + 2\langle x_{i_n(k)}^k, x_{i_n(k)}^{k-1} - x_{i_n(k)}^k \rangle + \|x_{i_n(k)}^{k-1} - x_{i_n(k)}^k\|^2, \end{aligned} \tag{14}$$

it follows that

$$\lim_{k\to\infty}\{x^k_{i_n(k)} - x^{k-1}_{i_n(k)}\} = 0.$$

The claims follows. □

This yields a strong convergence result concerning problems $(\mathcal{P}^i(\cdot))$.

Proposition 3.3 *Let $\check{X}$ be a limit point of the sequence $\{X^k\}$ generated by Algorithm 2.1. Then for $i = 1,\ldots,n$, $\check{X}$ solves $(\mathcal{P}^i(\check{X}))$.*

Proof: Let $\mathcal{K}$ be an infinite index set such that

$$X^k \to \check{X} \text{ as } k \to \infty,\ k \in \mathcal{K}.$$

In view of Lemma 3.2, for any fixed integer ℓ,

$$X^{k+\ell} \to \check{X} \text{ as } k \to \infty,\ k \in \mathcal{K}. \tag{15}$$

Now let $i \in \{1,\ldots,n\}$ and for every $k \in \mathcal{K}$ let $j(k) \in \{k, k+1, \ldots, k+n-1\}$ be such that $i_n(j(k)) = i$. Since $\{X^k\}$ is constructed by Algorithm 2.1,

$$X^{j(k)} = f_{i_n(j(k))}(X^{j(k)-1}) = f_i(X^{j(k)-1}), \quad \forall k \in \mathcal{K}. \tag{16}$$

Clearly the iteration map f_i is continuous over $\mathbf{X_R}$. Since, in view of Proposition 3.2, $\check{X}$ is nonsingular (thus $\check{X} \in \mathbf{X_R}$), and since $0 \le j(k) - k \le n$ for all k, it follows from (15) and (16), taking limits as $k \to \infty$, $k \in \mathcal{K}$, that

$$\check{X} = f_i(\check{X}).$$

The claim then follows from Proposition 3.1. □

This leads to the main result of this section. Recall that, given $s \in \mathbf{W}$, $\mathbf{W}$ a subset of a normed vector space $\mathcal{V}$, the tangent cone to $\mathbf{W}$ at s is defined by

$$TC(\mathbf{W}, s) = \{h : \exists o : \mathbf{R} \to \mathcal{V} \text{ s.t. } s + th + o(t) \in \mathbf{W}\ \ \forall t \ge 0,\ \frac{o(t)}{t} \to 0, \text{ as } t \to 0,\ t \ne 0\}$$

and that if a continuouly Fréchet-differentiable function $f : \mathcal{V} \to \mathbf{R}$ achieves its minimum over $\mathbf{W}$ at $\check{s} \in \mathbf{W}$, then

$$f'(\check{s}, h) \ge 0, \quad \forall h \in TC(\mathbf{W}, \check{s}),$$

where $f'(\check{s}, h)$ is the directional derivative of f at $\check{s}$ in direction h. Given $\check{X} \in \mathbf{X_R}$, let $\mathbf{X}^i_\mathbf{R}(\check{X})$ denote the feasible set for $(\mathcal{P}^i(\check{X}))$, i.e.,

$$\mathbf{X}^i_\mathbf{R}(\check{X}) = \{X = [\check{x}_1, \ldots, \check{x}_{i-1}, x, \check{x}_{i+1}, \ldots, \check{x}_n] \in \mathbf{R}^{n\times n} : x \in \mathbf{B}\mathcal{S}^\mathbf{R}_i\}$$

It is readily verified that the tangent cones to $\mathbf{X}^i_\mathbf{R}(\check{X})$ and to $\mathbf{X_R}$ at $\check{X}$ are subspaces ($\check{X}$ is "regular"). Let $T_i(\check{X})$ and $T(\check{X})$ be these subspaces (tangent planes).

Theorem 3.1 *Every limit point of $\{X^k\}$ is a stationary point for $(\mathcal{P})$ in the sense that*

$$\langle\langle \frac{\partial \det(\check{X})}{\partial X}, H \rangle\rangle = 0, \quad \forall H \in T(\check{X}).$$

Proof: Since $\check{X}$ solves $(\mathcal{P}^i(\check{X}))$,

$$\langle\langle \frac{\partial \det(\check{X})}{\partial X}, H \rangle\rangle = 0, \ \ \forall H \in T_i(\check{X}).$$

It is readily checked that any $H \in T_i(\check{X})$ has all its entries equal to zero except for the ith column and that a matrix H is in $T(\check{X})$ if and only if, for all $i = 1, \ldots, n$, its ith column is the ith column of a matrix in $T_i(\check{X})$. Thus

$$T(\check{X}) = T_1(\check{X}) \oplus \cdots \oplus T_n(\check{X})$$

and the claim follows. □

Remark 3.1 An open problem is: Under what circumstances does the sequence $\{X^k\}$ actually converge? First, in view of Lemma 3.2, $\{X^k\}$ converges whenever it has an isolated limit point. Next, for $k = 1, 2, \ldots$, let d^k be the $i_n(k)$th column of $(X^{k-1})^{-\top}$. Since $\{X^k\}$ is bounded away from singularity, it is clear that $\{X^k\}$ converges if and only if $\{d^{nk+i}\}_{k=0}^{\infty}$ converges for all $i \in \{1, \ldots, n\}$. Let $\check{X} = [\check{x}_1, \ldots, \check{x}_n]$ be a limit point of $\{X^k\}$; also, for $i = 1, \ldots, n$, let $\check{d}_i$ be the corresponding limit point of $\{d^{nk+i}\}_{k=0}^{\infty}$ and let $\check{u}_i = \check{d}_i / \sqrt{\|P_i \check{d}_i\|}$, where P_i is the orthogonal projection matrix onto $\mathcal{S}_i$, ($\{\|P_i d^{nk+i}\|\}$ is bounded away from zero). Then for $i = 1, \ldots, n$,

$$\langle \check{u}_i, P_i \check{u}_i \rangle = \frac{1}{\|P_i \check{d}_i\|} \langle \check{d}_i, P_i \check{d}_i \rangle = \langle \check{d}_i, \check{x}_i \rangle = 1$$

On the other hand, since $\{X^k - X^{k-1}\}$ goes to 0, $\langle \check{d}_i, \check{x}_j \rangle = 0$ for $i, j = 1, \ldots, n, \ i \neq j$. It follows that

$$\langle \check{u}_i, P_j \check{u}_j \rangle = \delta_{ij} \ \ i, j = 1, \ldots, n, \qquad (17)$$

where δ_{ij} is the Kronecker delta. Thus a sufficient condition for $\{X^k\}$ to converge is that, for one of its limit points, the corresponding $\check{u}_1, \ldots, \check{u}_n$ be isolated solution of (17). If two or more $\mathcal{S}_i$'s are identical (in particular if $m = n$), then (17) has no isolated solutions. If this is not the case, however, we conjecture that, at least generically, the Jacobian of the mapping $\mathbf{R}^{n^2} \to \mathbf{R}^{n^2}$ defined by (17) is nonsingular at $\check{u}_1, \ldots, \check{u}_n$ (this is always true for $m = 1$ when the λ_i's are all distinct), thus, that solutions to (17) are isolated, implying that $\{X^k\}$ converges.

4 Updating Two Columns at a Time

The convergence analysis of Section 3 was based on the observation that the KNV algorithm amounts to attempting to solve $(\mathcal{P})$ by performing global maximizations successively with respect to each column of X, while fixing all other columns, and repeating this process until convergence is observed. Such partial global maximizations can be performed efficiently due to the linearity of $\det(X)$ with respect to the entries of any given column when all other columns are kept constant.

Suppose now that global maximization with respect to $q > 1$ columns could be performed efficiently as well, specifically at a cost not much greater than that of successively performing q global maximizations with respect to each column taken individually. Likely, the speed of convergence of the overall process would be increased (in the extreme case when $q = n$ the convergence would be achieved in a single iteration). It turns out that this is indeed possible, with $q = 2$, as demonstrated next. The key is Proposition 4.1 below.

For simplicity of exposition, suppose that n is even, say $n = 2p$. Given $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n] \in \mathbf{X_R}$, let $U_i(\hat{X}) \in \mathbf{R}^{n\times n}$, $i = 1, \ldots, p$, be defined by

$$U_i(\hat{X}) = (uv^\top - vu^\top)$$

where u, $v \in \mathbf{R}^n$ form an orthonormal basis for the orthogonal complement of the set

$$\{\hat{x}_1, \ldots, \hat{x}_{2(i-1)}, \hat{x}_{2i+1}, \ldots, \hat{x}_n\},$$

and satisfy the inequality

$$\langle \hat{x}_{2i-1}, u\rangle\langle \hat{x}_{2i}, v\rangle \geq \langle \hat{x}_{2i-1}, v\rangle\langle \hat{x}_{2i}, u\rangle \tag{18}$$

(note that the latter can be achieved by proper choice of the orientation of u and v). It is readily checked that $U_i(\hat{X})$ is thus uniquely determined (although u and v are not) and is continuous as a function of $\hat{X}$.

Proposition 4.1 *Let* $\hat{X} = [\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{2p}] \in \mathbf{R}^{n\times n}$ *be such that* $\det(\hat{X}) > 0$, *let* ξ, $\eta \in \mathbf{R}^n$, *let* $i \in \{1, 2, \ldots, p\}$, *and let*

$$X = [\hat{x}_1, \cdots, \hat{x}_{2i-2}, \xi, \eta, \hat{x}_{2i+1}, \cdots, \hat{x}_{2p}]$$

and

$$X_= = [\hat{x}_1, \cdots, \hat{x}_{2i-2}, \hat{x}_{2i+1}, \cdots, \hat{x}_{2p}].$$

Then $X_=^\top X_=$ *is nonsingular, and*

$$\det(X) = \langle \xi, U_i(\hat{X})\eta\rangle\sqrt{\det(X_=^\top X_=)}. \tag{19}$$

This result (proved in the appendix) suggests the iteration maps $g_i : \mathbf{X_R} \to \mathbf{X_R}$, $i = 1, 2, \ldots, p$, given by

$$g_i(\hat{X}) = [\hat{x}_1, \ldots, \hat{x}_{2(i-1)}, \tilde{\xi}, \tilde{\eta}, \hat{x}_{2i+1}, \ldots, \hat{x}_n]$$

where $(\tilde{\xi}, \tilde{\eta})$ is closest to $(\hat{x}_{2i-1}, \hat{x}_{2i})$, in the Euclidean sense, among the solutions of

$$\text{maximize } \langle \xi, U_i(\hat{X})\eta\rangle \ \text{ s.t. } \xi \in \mathbf{B}\mathcal{S}^{\mathbf{R}}_{2i-1}, \ \eta \in \mathbf{B}\mathcal{S}^{\mathbf{R}}_{2i}, \tag{20}$$

i.e., among the left-right singular vector pairs corresponding to the largest singular value of the projection of $U_i(\hat{X})$ on $\mathcal{S}^{\mathbf{R}}_{2i-1}$ on the left and on $\mathcal{S}^{\mathbf{R}}_{2i}$ on the right (more on this in Section 6 below). In view of Proposition 4.1, for $i = 1, \ldots, p$, $g_i(\hat{X})$ is closest to $\hat{X}$ among the global maximizers for the problem

$$\begin{aligned} &\text{maximize } \det(X) \\ &\text{s.t. } X = [\hat{x}_1, \ldots, \hat{x}_{2(i-1)}, \xi, \eta, \hat{x}_{2i+1}, \ldots, \hat{x}_n], \ \xi \in \mathbf{B}\mathcal{S}^{\mathbf{R}}_{2i-1}, \ \eta \in \mathbf{B}\mathcal{S}^{\mathbf{R}}_{2i}. \end{aligned} \tag{$\mathcal{P}_2^i(\hat{X})$}$$

Following Algorithm 2.1, Algorithm 4.1 below cyclically applies g_1 through g_p.

Algorithm 4.1 *:*

Step 1. Pick $X^0 \in \mathbf{X_R}$. *Set* $k := 1$.

Step 2. Set $X^k := g_{i_p(k)}(X^{k-1})$.

Step 3. Set $k := k+1$ *and go to Step 2.* □

Remark 4.1: $U_i(\hat{X})$ always has rank two. Problem (20) has exactly two solutions ($(\tilde{\xi}, \tilde{\eta})$ and $(-\tilde{\xi}, -\tilde{\eta})$) in case of nonrepeated largest singular value, and a continuum of solutions otherwise. $(\tilde{\xi}, \tilde{\eta})$ is selected closest to $(\hat{x}_{2i-1}, \hat{x}_{2i})$ in an attempt to force convergence of the sequence $\{X^k\}$.

The following result, which parallels Proposition 3.2, is a direct consequence of the above.

Proposition 4.2 *For* $k = 1, 2, \ldots$, X^k *constructed by Algorithm* 4.1 *is the unique maximizer of* $\mathcal{P}_2^{i_p(k)}(X^{k-1})$,

$$0 < \det(X^0) \leq \det(X^1) \leq \cdots$$

and $\{\det(X^k)\}$ *converges to a finite limit. Moreover, if there exists* k_0 *such that*

$$\det(X^{k_0+n}) = \det(X^{k_0}), \tag{21}$$

then

$$X^k = X^{k_0} \qquad \forall k \geq k_0.$$

□

In Section 3 we showed that, for the sequence $\{X^k\}$ generated by the KNV Algorithm, the distance $\|X^k - X^{k-1}\|$ between successive iterates converges to zero as $k \to \infty$. It appears that this is not guaranteed in the case of Algorithm 4.1. This is due to the discontinuity of the map g_i at points $\hat{X}$ where the largest singular value of the projection of $U_i(\hat{X})$ on $\mathcal{S}_{2i-1}$ on the left and on $\mathcal{S}_{2i}$ on the right is repeated. Still, in all our numerical experiments, we observed that $\|X^k - X^{k-1}\|$ did converge to zero. We next show that, whenever this occurs, all limit points of $\{X^k\}$ satisfy certain optimality conditions.

Proposition 4.3 *Let* $\check{X}$ *be a limit point of the sequence* $\{X^k\}$ *constructed by Algorithm 4.1, and suppose that* $\{X^k - X^{k-1}\} \to 0$ *as* $k \to \infty$. *Then, for* $i = 1, \ldots, p$, $\check{X}$ *solves* $(\mathcal{P}_2^i(\check{X}))$.

Proof: Let $\mathcal{K}$ be an infinite index set such that

$$X^k \to \check{X}, \quad \text{as} \quad k \to \infty, \quad k \in \mathcal{K}.$$

Since $i_p(k) \in \{1, \ldots, p\}$, a finite set, for all k, there must exist $\hat{\imath} \in \{1, \ldots, p\}$ such that $i_p(k) = \hat{\imath}$ for infinitely many $k \in \mathcal{K}$. Thus there exists an infinite index set $\mathcal{L}$ such that

$$X^{\ell n + \hat{\imath}} \to \check{X}, \quad \text{as} \quad \ell \to \infty, \quad \ell \in \mathcal{L}.$$

Since $\{X^k - X^{k-1}\} \to 0$ as $k \to \infty$, it follows that, for all $i \in \{1, \ldots, p\}$,

$$X^{\ell n + i} \to \check{X}, \quad \text{as} \quad \ell \to \infty, \quad \ell \in \mathcal{L}. \tag{22}$$

Now, by construction, we know that, for $i = 1, \ldots, p$, $\ell = 1, 2, \ldots$

$$\det(X^{\ell n+i}) = \max_{\xi \in \mathbf{B}\mathcal{S}_{2i-1}^{\mathbf{R}},\ \eta \in \mathbf{B}\mathcal{S}_{2i}^{\mathbf{R}}} \det[x_1^{\ell n+i}, \ldots, x_{2i-2}^{\ell n+i}, \xi, \eta, x_{2i+1}^{\ell n+i}, \ldots, x_n^{\ell n+i}]. \tag{23}$$

Since maximization over a compact set preserves continuity, letting $\ell \to \infty$, $\ell \in \mathcal{L}$ in both sides of (23) and taking (22) into account, we get

$$\det(\check{X}) = \max_{\xi \in \mathbf{B}\mathcal{S}^{\mathbf{R}}_{2i-1},\ \eta \in \mathbf{B}\mathcal{S}^{\mathbf{R}}_{2i}} \det[\check{x}_1, \ldots, \check{x}_{2i-2}, \xi, \eta, \check{x}_{2i+1}, \ldots, \check{x}_n],$$

which proves the claim. □

A direct consequence is that $\check{X}$ also solves $\left(\mathcal{P}^j(\check{X})\right)$ for $j = 1, \ldots, 2p$. The following result can then be proved using an argument identical to that used in the proof of Theorem 3.1.

Theorem 4.1 : *Suppose that $\{X^k - X^{k-1}\} \to 0$ and let $\check{X}$ be a limit point of the sequence $\{X^k\}$ constructed by Algorithm 4.1. Then $\check{X}$ is stationary for $(\mathcal{P})$ in the sense that*

$$\langle\langle \frac{\partial \det(\check{X})}{\partial X}, H \rangle\rangle = 0, \quad \forall H \in T(\check{X}).$$

□

As just pointed out, we cannot in general guarantee that $\|X^k - X^{k-1}\|$ goes to zero. A sufficient condition for this to hold is given next, suggesting that it may hold in some generic sense.

Proposition 4.4 : *Suppose that, for every $i \in \{1, \ldots, p\}$ for which $\lambda_{2i} \neq \lambda_{2i-1}$, and every accumulation points $\check{X}$ of $\{X^k\}$, it holds that*

$$\sigma_1(S_{2i-1}^\top U_i(\check{X}) S_{2i}) \neq \sigma_2(S_{2i-1}^\top U_i(\check{X}) S_{2i}),$$

where S_{2i-1} and S_{2i} are matrices whose columns form orthogonal bases for $\mathcal{S}_{2i-1}$ and $\mathcal{S}_{2i}$. Then $\{X^k - X^{k-1}\} \to 0$ as $k \to \infty$. □

Corollary 4.1 : *If, for $i = 1, \ldots, p$, $\lambda_{2i} = \lambda_{2i-1}$, then $\{X^k - X^{k-1}\} \to 0$ as $k \to \infty$.* □

Algorithm 4.1 can be modified in such a way that statements in Proposition 4.3 and Theorem 4.1 hold unconditionally. The key is, instead of operating cyclically over all pairs of columns of X, to always select the pair yielding the largest increase in the determinant value. This is done in Algorithm 4.2 below. Note that, on a sequential machine, the CPU cost of each iteration of this algorithm is roughly p times that of each iteration of Algorithm 4.1. If p processors are used, however, the CPU times are identical.

Algorithm 4.2 :

Step 1. Pick $X^0 \in \mathbf{X}_{\mathbf{R}}$. Set $k := 1$.

Step 2. Set $X^k := g_{i_k}(X^{k-1})$ for some $i_k \in \{1, \ldots, p\}$ satisfying

$$\det(g_{i_k}(X^{k-1})) \geq \det(g_i(X^{k-1})), \quad i = 1, \ldots, p.$$

Step 3. Set $k := k + 1$ and go to Step 2. □

This idea can be pushed further by selecting, at each iteration, the "best" among all $n(n-1)/2$ pairs of columns of X. Again, this is especially advantageous if a large number of processors are available. It is done in Algorithm 4.3 below, where iteration maps $\tilde{g}_{ij} : \mathbf{X_R} \to \mathbf{X_R}$, $i, j = 1, 2, \ldots, n$, $i < j$, are defined as follows. Given $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n] \in \mathbf{X_R}$,

$$\tilde{g}_{ij}(\hat{X}) = [\hat{x}_1, \ldots, \hat{x}_{i-1}, \tilde{\xi}, \hat{x}_{i+1}, \ldots, \hat{x}_{j-1}, \tilde{\eta}, \hat{x}_{j+1}, \ldots, \hat{x}_n]$$

where $(\tilde{\xi}, \tilde{\eta})$ is the closest to $(\hat{x}_i, \hat{x}_j)$, in the Euclidean sense, among the solutions of

$$\text{maximize } \langle \xi, U_i(\hat{X})\eta \rangle \;\text{ s.t. }\; \xi \in \mathbf{B}\mathcal{S}_i^{\mathbf{R}}, \;\; \eta \in \mathbf{B}\mathcal{S}_j^{\mathbf{R}}.$$

Algorithm 4.3 *:*

Step 1. Pick $X^0 \in \mathbf{X_R}$. *Set* $k := 1$.

Step 2. Set $X^k := \tilde{g}_{i_k j_k}(X^{k-1})$ *for some* $i_k, j_k \in \{1, \ldots, n\}$, $i_k < j_k$ *satisfying*

$$\det(\tilde{g}_{i_k j_k}(X^{k-1})) \geq \det(\tilde{g}_{ij}(X^{k-1})), \;\; \forall i, j \in \{1, \ldots, n\}, \;\; i < j.$$

Step 3. Set $k := k + 1$ *and go to Step 2.*

□

As for the previous algorithm, the following holds.

Proposition 4.5 *For* $k = 1, 2, \ldots$, X^k *constructed by Algorithm* 4.2 *and* 4.3 *satisfies*

$$0 < \det(X^0) \leq \det(X^1) \leq \cdots$$

and $\{\det(X^k)\}$ *converges to a finite limit. Moreover, if there exists* k_0 *such that*

$$\det(X^{k_0+n}) = \det(X^{k_0}), \tag{24}$$

then

$$X^k = X^{k_0} \qquad \forall k \geq k_0$$

□

The following stronger convergence result holds.

Proposition 4.6 *: Let* $\check{X}$ *be a limit point of the sequence* $\{X^k\}$ *constructed by either Algorithm 4.2 or Algorithm 4.3. Then, for* $i = 1, \ldots, p$, $\check{X}$ *solves* $(\mathcal{P}_2^i(\check{X}))$.

Proof. Suppose Algorithm 4.2 is used. Let $\mathcal{K}$ be an infinite index set such that $X^{k-1} \to \check{X}$, as $k \to \infty$, $k \in \mathcal{K}$. First note that

$$\det(X^{k-1}) \leq \det(g_i(X^{k-1})) \leq \det(g_{i_k}(X^{k-1})) = \det(X^k), \;\; \forall i, \;\; \forall k,$$

where the first relation follows the definition of g_i, and the second and third relations follow from the construction of $\{X^k\}$. In view of Proposition 4.5, $\det(X^{k-1})$ and $\det(X^k)$ converge to the same limit, $\det(\check{X})$, as $k \to \infty$. Thus, for $i = 1, \ldots, p$,

$$\det(g_i(X^{k-1})) \to \det(\check{X}) \;\text{ as } k \to \infty.$$

In view of the definition of g_i, for $i = 1, \dots, p$, for any $X = [x_1, \dots, x_n] \in \mathbf{X_R}$,

$$\det(g_i(X)) = \max_{\xi \in \mathbf{B}\mathcal{S}^{\mathbf{R}}_{2i-1},\ \eta \in \mathbf{B}\mathcal{S}^{\mathbf{R}}_{2i}} \det([x_1, \dots, x_{2(i-1)}, \xi, \eta, x_{2i+1}, \dots, x_n]).$$

Since the maximization over a compact set preserves continuity, $\det(g_i(X))$ is continuous in X. Therefore for $i = 1, \dots, p$,

$$\det(g_i(X^{k-1})) \to \det(g_i(\check{X})) \text{ as } k \to \infty,\ k \in \mathcal{K}.$$

It follows that for $i = 1, \dots, p$,

$$\det(g_i(\check{X})) = \det(\check{X}).$$

Thus the claim holds. The case of Algorithm 4.3 is handled similarly. □

The following result can then be proved using an argument identical to the one used in the proof of Theorem 3.1.

Theorem 4.2 : *Let $\check{X}$ be a limit point of the sequence $\{X^k\}$ constructed by either Algorithm 4.2 or Algorithm 4.3. Then $\check{X}$ is stationary for $(\mathcal{P})$ in the sense that*

$$\langle\langle \frac{\partial \det(\check{X})}{\partial X}, H \rangle\rangle = 0, \quad \forall H \in T(\check{X}).$$

□

Remark 3.2: Note that, similarly, the KNV Algorithm could be modified by selecting the "best" column at each iteration. It is readily verified that, with such an algorithm, Theorem 3.1 would still hold.

5 Possibly Complex Prescribed Eigenvalues

Consider now the case where the set of desired poles includes a number of complex conjugate pairs. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues to be assigned. For the sake of simplicity of exposition, again assume that n is even. Moreover assume that $\{\lambda_1, \dots, \lambda_n\} \cap \mathbf{R} = \{\lambda_1, \dots, \lambda_{2p}\}$, i.e., $\lambda_1, \dots, \lambda_{2p}$ are real and $\lambda_{2p+1}, \dots, \lambda_n$ are complex; let c be the number of complex pairs, i.e., $c = n/2 - p$, and assume that $\lambda_{2i} = \overline{\lambda}_{2i-1}$, $i = p+1, \dots, p+c$. Clearly, candidate sets of eigenvectors of the closed loop matrix $A + BF$ must include c complex conjugate pairs. Moreover, as in the real case, they must satisfy additional conditions. The next theorem extends Theorem 2.1.

Theorem 5.1 *Let $X = [x_1, \dots, x_n]$, nonsingular, and $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ be two complex $n \times n$ matrices such that (i) for $j = 1, \dots, 2p$, $\lambda_j \in \mathbf{R}$ and $x_j \in \mathbf{R}^n$, and (ii) for $i = p+1, \dots, p+c$, $\lambda_{2i} = \overline{\lambda}_{2i-1}$ and $x_{2i} = \overline{x}_{2i-1}$. Then (i) $X \Lambda X^{-1}$ is real, and (ii) $(A + BF)X = X\Lambda$ for some real matrix F if and only if $x_j \in \mathcal{S}_j$, $j = 1, \dots, n$.*

Proof: Let $P \in \mathbf{R}^{n \times n}$ be the permutation matrix that exchanges columns $2i-1$ and $2i$, $i = p+1, \dots, p+c$, of the matrix it postmultiplies. Thus, denoting complex conjugate by a bar, $X = \overline{X}P$, $\overline{\Lambda} = P\Lambda P$, and $P^{-1} = P$. Then

$$\overline{X \Lambda X^{-1}} = \overline{X} P \Lambda P \overline{X}^{-1} = \overline{X} P \Lambda (\overline{X} P)^{-1} = X \Lambda X^{-1},$$

proving the first claim. Now suppose that for some real matrix F, $(A+BF)X = X\Lambda$. Then

$$AX - X\Lambda = -BFX$$

implying that

$$(A - \lambda_j I)x_j \in \mathbf{R_C}(B), \;\; j = 1, \ldots, n, \tag{25}$$

i.e., $x_j \in \mathcal{S}_j$ for $j = 1, \ldots, n$. Finally, suppose that (25) holds, i.e., for some $y_j \in \mathbf{C}^m$, $j = 1, \ldots, n$,

$$(A - \lambda_j I)x_j = By_j.$$

Thus

$$AX - X\Lambda = BY,$$

with $Y := [y_1, \ldots, y_n] \in \mathbf{C}^{m \times n}$, i.e.,

$$A - X\Lambda X^{-1} = BYX^{-1}.$$

The left hand side is a real matrix. Thus

$$BYX^{-1} = \Re e(BYX^{-1}) = B\Re e(YX^{-1}).$$

The last claim then follows by setting $F = \Re e(YX^{-1})$. □

In view of this result, we will focus on a modification of problem $(\mathcal{P})$ given by

$$\text{maximize } |\det(X)| \text{ s.t. } X \in \mathbf{X}, \tag{$\tilde{\mathcal{P}}$}$$

where $\mathbf{X}$ (a generalization of $\mathbf{X_R}$) is defined by

$$\begin{aligned} \mathbf{X} \;=\; & \{X = [x_1, \ldots, x_n] \in \mathbf{C}^{n \times n} : x_i \in \mathbf{B}\mathcal{S}_i, i = 1, \ldots, n;\;\; x_{2i-1} = \overline{x}_{2i}, i = p+1, \ldots, p+c; \\ & x_i \in \mathbf{R}^n, i = 1, \ldots, 2p;\;\; \det(X) \neq 0\}. \end{aligned} \tag{26}$$

(The reason for the absolute value is discussed below.)

Again consider constructing a sequence $\{X^k\}$ of candidate closed loop eigenvector matrices, i.e., of matrices $X^k \in \mathbf{X}$. Since self-conjugacy of certain pairs of columns of X^k must be preserved, only algorithms that modify such pairs concurrently can be considered. This rules out extension of the KNV algorithm. Thus let $\hat{X} \in \mathbf{X}$ and consider instead problem $(\mathcal{P}_2^i(\hat{X}))$, which forms the basis of Algorithm 4.1. For use in the present context, for $i = p+1, \ldots, p+c$, two modifications are in order. First, obviously, maximization must be under the additional constraints that $\eta = \bar{\xi}$. Second, maximizing $\det(X)$ may not be appropriate, as its value may not be real (depending on the parity of c, it is either real or pure imaginary). While it was pointed out in Section 4 that, for $i = 1, \ldots, p$, $g_i(\hat{X})$ is the closest to $\hat{X}$ among the global maximizations for $(\mathcal{P}_2^i(\hat{X}))$, it is worth noting that it also is the closest to $\hat{X}$ among the global maximizers of $|\det(X)|$ subject to the same constraints as in $(\mathcal{P}_2^i(\hat{X}))$ (and that, when $|\det(X)|$ is substituted for $\det(X)$, $\mathbf{B}\mathcal{S}_{2i-1}^{\mathbf{R}}$ and $\mathbf{B}\mathcal{S}_{2i}^{\mathbf{R}}$ can be replaced by $\mathbf{B}\mathcal{S}_{2i-1}$ and $\mathbf{B}\mathcal{S}_{2i}$ without affecting the solution). The latter point of view will be used here. Thus, we now substitute for $(\mathcal{P}_2^i(\hat{X}))$ the problems $(\tilde{\mathcal{P}}_2^i(\hat{X}))$ (for $i = 1, \ldots, c+p$) given by

$$\text{maximize } |\det(X)| \;\; \text{s.t. } X = [\hat{x}_1, \ldots, \hat{x}_{2(i-1)}, \xi, \eta, \hat{x}_{2i+1}, \ldots, \hat{x}_n] \in \mathbf{X}. \tag{$\tilde{\mathcal{P}}_2^i(\hat{X})$}$$

It turns out that, as was the case for $(\mathcal{P}_2^i(\hat{X}))$, $(\tilde{\mathcal{P}}_2^i(\hat{X}))$ can be solved efficiently. The key is Proposition 5.1 below. Given $\hat{X} = [\hat{x}_1, \cdots, \hat{x}_n] \in \mathbf{X}$, for $i = 1, \ldots, p$, let $\tilde{U}_i(\hat{X}) = uv^\top - vu^\top \in$

$\mathbf{R}^{n\times n}$, where u, $v \in \mathbf{R}^n$ form an orthonormal basis for the orthogonal complement of the span of

$$\{\hat{x}_1, \ldots, \hat{x}_{2i-2}, \hat{x}_{2i+1}, \ldots, \hat{x}_{2p}\} \cup \{\Re e(\hat{x}_{2j}), \Im m(\hat{x}_{2j}), j = p+1, \ldots, p+c\},$$

and, for $i = p+1, \ldots, p+c$, let $\tilde{U}_i(\hat{X}) = u\overline{u}^\top - \overline{u}u^\top$, where $u = u_R + \mathrm{j}u_I$ is such that $\sqrt{2}u_R$, $\sqrt{2}u_I \in \mathbf{R}^n$ form an orthonormal basis for the orthogonal complement of the set

$$\{\hat{x}_1, \ldots, \hat{x}_{2p}\} \cup \{\Re e(\hat{x}_{2j}), \Im m(\hat{x}_{2j}), j = p+1, \ldots, p+c, j \neq i\},$$

and satisfy the inequality

$$\langle \Re e(\hat{x}_{2i}), u_R\rangle \langle \Im m(\hat{x}_{2i}), u_I\rangle \geq \langle \Re e(\hat{x}_{2i}), u_I\rangle \langle \Im m(\hat{x}_{2i}), u_R\rangle. \tag{27}$$

Again, it is readily checked that $\tilde{U}_i(\hat{X})$ is thus uniquely determined and is continuous as a function of $\hat{X}$.

Lemma 5.1 *: Let $i \in \{p+1, \ldots, p+c\}$, and let $u = u_R + \mathrm{j}u_I$ be such that $\sqrt{2}u_R$, $\sqrt{2}u_I \in \mathbf{R}^n$ form an orthonormal basis for the orthogonal complement of the set*

$$\{\hat{x}_1, \ldots, \hat{x}_{2p}\} \cup \{\Re e(\hat{x}_{2j}), \Im m(\hat{x}_{2j}), j = p+1, \ldots, p+c, j \neq i\}.$$

Then $\{u, \overline{u}\}$ is an orthonormal basis for the null space of $[\hat{x}_1, \cdots, \hat{x}_{2i-2}, \hat{x}_{2i+1}, \cdots, \hat{x}_n]^*$. □

Proposition 5.1 *Let $\hat{X} = [\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n] \in \mathbf{C}^{n\times n}$ with $\hat{x}_1, \ldots, \hat{x}_{2p} \in \mathbf{R}^n$, and $\hat{x}_{2i-1} = \overline{\hat{x}}_{2i}$, $i = p+1, \ldots, p+c$ be such that $\det(\hat{X}) \neq 0$, let $\eta \in \mathbf{C}^n$, let $i \in \{p+1, \ldots, p+c\}$, and let*

$$X = [\hat{x}_1, \cdots, \hat{x}_{2i-2}, \overline{\eta}, \eta, \hat{x}_{2i+1}, \cdots, \hat{x}_n]$$

and

$$X_= = [\hat{x}_1, \cdots, \hat{x}_{2i-2}, \hat{x}_{2i+1}, \cdots, \hat{x}_n].$$

Then $X_=^* X_=$ is nonsingular, and

$$|\det(X)| = |\langle \eta, \tilde{U}_i(\hat{X})\eta\rangle| \sqrt{\det(X_=^* X_=)}. \tag{28}$$

This result suggests the iteration maps $h_i : \mathbf{X} \to \mathbf{X}$, $i = 1 \ldots, p+c$, defined as follows. For $i = 1, \ldots, p$,

$$h_i(\hat{X}) = [\hat{x}_1, \ldots, \hat{x}_{2(i-1)}, \tilde{\xi}, \tilde{\eta}, \hat{x}_{2i+1}, \ldots, \hat{x}_n]$$

where $(\tilde{\xi}, \tilde{\eta})$ is the closest to $(\hat{x}_{2i-1}, \hat{x}_{2i})$, in the Euclidean sense, among the solutions of

$$\text{maximize } |\langle \xi, \tilde{U}_i(\hat{X})\eta\rangle| \text{ s.t. } \xi \in \mathbf{B}\mathcal{S}_{2i-1}, \ \eta \in \mathbf{B}\mathcal{S}_{2i}, \tag{29}$$

i.e., among the left-right singular vector pairs corresponding to the largest singular value of the projection of $U_i(\hat{X})$ on $\mathcal{S}_{2i-1}$ on the left and on $\mathcal{S}_{2i}$ on the right (more on this in Section 6 below). For $i = p+1, \ldots, p+c$,

$$h_i(\hat{X}) = [\hat{x}_1, \ldots, \hat{x}_{2(i-1)}, \tilde{\eta}, \overline{\tilde{\eta}}, \hat{x}_{2i+1}, \ldots, \hat{x}_n]$$

where $\tilde{\eta}$ is closest to $\hat{x}_{2i-1}$, in the Euclidian sense, among the solutions of

$$\text{maximize } \left|\langle \eta, \tilde{U}_i(\hat{X})\eta\rangle\right| \qquad \text{s.t. } \eta \in \mathbf{B}\mathcal{S}_{2i}. \tag{30}$$

i.e., among the unit length eigenvectors corresponding the largest eigenvalue of the projection of the Hermitian matrix $\tilde{U}_i(\hat{X})$ on $\overline{\mathcal{S}}_{2i}$ on the left and on $\mathcal{S}_{2i}$ on the right (more on this in Section 6). In view of Proposition 5.1, and of an obvious extension of Proposition 4.1 to the mixed case, for $i = 1, \ldots, p+c$, $h_i(\hat{X})$ is the closest to $\hat{X}$ among all global maximizers for $(\tilde{\mathcal{P}}_2^i(\hat{X}))$.

Algorithm 5.1 below cyclically applies h_1 through h_{p+c}.

Algorithm 5.1 :

Step 1. Pick $X^0 \in \mathbf{X}$. *Set* $k := 1$.

Step 2. Set $X^k := h_{i_{p+c}(k)}(X^{k-1})$.

Step 3. Set $k := k+1$ *and go to Step 2.* □

The following result, which parallels Proposition 3.2, is an immediate consequence of the above.

Proposition 5.2 *For* $k = 1, 2, \ldots$, X^k *constructed by Algorithm* 5.1 *is the unique maximizer of* $\mathcal{P}_2^{i_p(k)}(X^{k-1})$,

$$0 < |\det(X^0)| \le |\det(X^1)| \le \cdots$$

and $\{|\det(X^k)|\}$ *converges to a finite limit. Moreover, if there exists* k_0 *such that*

$$|\det(X^{k_0+n})| = |\det(X^{k_0})|, \tag{31}$$

then

$$X^k = X^{k_0} \qquad \forall k \ge k_0.$$

□

Proposition 5.3 *Let* $\check{X}$ *be a limit point of the sequence* $\{X^k\}$ *constructed by Algorithm 5.1, and suppose that* $\{X^k - X^{k-1}\} \to 0$ *as* $k \to \infty$. *Then, for* $i = 1, \ldots, p+c$, $\check{X}$ *solves* $(\tilde{\mathcal{P}}_2^i(\check{X}))$.
□

With a suitable interpretation of the notation, a result analogous to Theorem 4.1 follows.

Theorem 5.2 *Suppose that* $\{X^k - X^{k-1}\} \to 0$ *and let* $\check{X}$ *be a limit point of the sequence* $\{X^k\}$ *constructed by Algorithm 5.1. Then* $\check{X}$ *is stationary for* $(\tilde{\mathcal{P}})$ *in the sense that*

$$\langle\langle \frac{\partial |\det(\check{X})|}{\partial X}, H \rangle\rangle = 0, \quad \forall H \in T(\check{X}).$$

□

The notation used in this statement is to be understood as follows. First note that since $|\det(\check{X})| > 0$ at $\check{X}$, $|\det(X)|$ is a continuously differentiable function of $x_1, \ldots, x_{2p}$, and of the real and imaginary parts x_{2i}^R, and x_{2i}^I of x_{2i}, $i = p+1, \ldots, p+c$ (where $x_1, \ldots, x_n$ are the columns of X). By $\frac{\partial|\det(\check{X})|}{\partial X}$, we mean the value at $\check{X}$ of the Fréchet derivative of $|\det(X)|$ with respect to $(x_1, \ldots, x_{2p}; x_{2i}^R, x_{2i}^I, i = p+1, \ldots, p+c)$. Similarly, in the definition of $T(\check{X})$, the constraints defining the feasible set of $(\tilde{\mathcal{P}})$ are considered as functions of these real variables.

Again, there appears to be no guarantee that $\|X^k - X^{k-1}\| \to 0$ as $k \to \infty$. However, a result similar to Proposition 4.4 holds, again suggesting that it may hold in some generic sense.

Proposition 5.4 *: Suppose that, for every $i \in \{p+1, \ldots, p+c\}$ and every accumulation points $\hat{X}$ of $\{X^k\}$, it holds that*

$$\theta_1(S_{2i}^* \tilde{U}_i(\hat{X}) S_{2i}) \neq \theta_2(S_{2i}^* \tilde{U}_i(\hat{X}) S_{2i}),$$

where θ_ℓ, $\ell = 1, 2$ are the nonzero eigenvalues of $S_{2i}^* \tilde{U}_i(\hat{X}) S_{2i}$, and S_{2i} are matrices whose columns form orthonormal bases for $\mathcal{S}_{2i}$. Then $\{X^k - X^{k-1}\} \to 0$ as $k \to \infty$. □

Algorithm 5.1 can be modified in such a way that statements in Proposition 5.3 and Theorem 5.2 hold unconditionally. We therefore propose a modification of Algorithm 5.1 similar to Algorithm 4.2.

Algorithm 5.2 *:*

Step 1. Pick $X^0 \in \mathbf{X}$. Set $k := 1$.

Step 2. Set $X^k := h_{i_k}(X^{k-1})$ for some $i_k \in \{1, \ldots, p+c\}$ satisfying

$$|\det(h_{i_k}(X^{k-1}))| \geq |\det(h_i(X^{k-1}))|, \quad i = 1, \ldots, p+c.$$

Step 3. Set $k := k+1$ and go to Step 2.

□

Our last algorithm is the obvious extension of Algorithm 4.3. Let $\tilde{h}_{ij} : \mathbf{X} \to \mathbf{X}$ be defined as follows for every pair $(i, j) \in J$, where

$$J = \{(i,j) : i, j \in \{1, \ldots, 2p\}, i < j\} \cup \{(2\ell - 1, 2\ell) : \ell = p+1, \ldots, p+c\}.$$

Given $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n] \in \mathbf{X}$, (i) if $i, j \in \{1, \ldots, 2p\}$, then

$$\tilde{h}_{ij}(\hat{X}) = [\hat{x}_1, \ldots, \hat{x}_{i-1}, \tilde{\xi}, \hat{x}_{i+1}, \ldots, \hat{x}_{j-1}, \tilde{\eta}, \hat{x}_{j+1}, \ldots, \hat{x}_n]$$

where $(\tilde{\xi}, \tilde{\eta})$ is the closest to $(\hat{x}_i, \hat{x}_j)$, in the Euclidean sense, among the solutions of

$$\text{maximize } \langle \xi, \tilde{U}_i(\hat{X}) \eta \rangle \text{ s.t. } \xi \in \mathbf{B}\mathcal{S}_i, \quad \eta \in \mathbf{B}\mathcal{S}_j;$$

(ii) if $i = 2\ell - 1$, $j = 2\ell$ for $\ell \in \{p+1, \ldots, p+c\}$, then $\tilde{h}_{ij}$ is equal to h_ℓ. The idea is to select, at each iteration, the "best" among all possible $p(2p-1) + c$ pairs of columns of X.

Algorithm 5.3 *:*

Step 1. Pick $X^0 \in \mathbf{X}$. Set $k := 1$.

Step 2. Set $X^k := \tilde{h}_{i_k j_k}(X^{k-1})$ for some $(i_k, j_k) \in J$ satisfying

$$|\det(\tilde{h}_{i_k j_k}(X^{k-1})| \geq |\det(\tilde{h}_{ij}(X^{k-1})|, \quad \forall (i,j) \in J.$$

Step 3. Set $k := k + 1$ and go to Step 2. □

The following are obvious extensions of Proposition 4.6 and Theorem 4.2.

Proposition 5.5 *Let $\check{X}$ be a limit point of the sequence constructed by either Algorithm 5.2 or Algorithm 5.3. Then, for $i = 1, \ldots, p + c$, $\check{X}$ solves $(\tilde{\mathcal{P}}_2^i(\check{X}))$.*

Theorem 5.3 *Let $\check{X}$ be a limit point of the sequence constructed by either Algorithm 5.2 or Algorithm 5.3. Then, $\check{X}$ is stationary for $(\tilde{\mathcal{P}})$ in the sense that*

$$\langle\langle \frac{\partial |\det(\check{X})|}{\partial X}, H \rangle\rangle = 0, \quad \forall H \in T(\check{X}).$$

□

6 Implementation Issues

For simplicity, when counting floating point operations, we will assume that all subspaces $\mathcal{S}_i$, $i = 1, \ldots, n$, have dimension of m (this is the case when B is full column rank and none of the prescribed closed-loop eigenvalues λ_i is an eigenvalue of A). Qualitatively, our results do not require this assumption.

For $j = 1, \ldots, n$, let $S_j \in \mathbf{C}^{n \times n}$ be a matrix whose columns form an orthonormal basis for the subspace $\mathcal{S}_j$. Let $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n] \in \mathbf{X}$, and for $i = p+1, \ldots, p+c$, let $\hat{x}_{2i}^R = \Re e(\hat{x}_{2i})$, and $\hat{x}_{2i}^I = \Im m(\hat{x}_{2i})$. Then, for $i = 1, \ldots, p$, $h_i(\hat{X})$ is obtained by replacing the $(2i-1)$th and $(2i)$th columns of $\hat{X}$ by $\tilde{\xi}$ and $\tilde{\eta}$, where $[\tilde{\xi}^\top, \tilde{\eta}^\top]$ is the closest, in the Euclidian sense, to $[\hat{x}_{2i-1}^\top, \hat{x}_{2i}^\top]$ among all the vectors of the form $[(S_{2i-1}\tilde{\mu})^\top, (S_{2i}\tilde{\nu})^\top]$, with $[\tilde{\mu}^\top, \tilde{\nu}^\top]^\top$ a solution of

$$\text{maximize } \mu^\top S_{2i-1}^\top U_i(\hat{X}) S_{2i} \nu \;\text{ s.t. }\; \mu, \nu \in \mathbf{BR}^m. \tag{32}$$

For $i = p+1, \ldots, p+c$, $h_i(\hat{X})$ is obtained by replacing the $(2i-1)$th and $(2i)$th columns of $\hat{X}$ by $\tilde{\eta}$ and $\overline{\tilde{\eta}}$, where $\tilde{\eta}$ is closest to $\hat{x}_{2i-1}$, in the Euclidian sense, among all the vectors of the form $S_{2i}\tilde{\mu}$, with $\tilde{\mu}$ a solution of

$$\text{maximize } \left|\langle \mu, S_{2i}^* \tilde{U}_i(\hat{X}) S_{2i} \mu \rangle\right| \qquad \text{s.t. } \; \mu \in \mathbf{BC}^m. \tag{33}$$

Remark 6.1: When the current iterate X^k is such that, for some i, the two nonzero singular values of $S_{2i-1}^\top \tilde{U}_i(X^k) S_{2i}$ (for $i \in \{1, \ldots, p\}$) or the two nonzero eigenvalues of $S_{2i-1}^* \tilde{U}_i(X^k) S_{2i}$ (for $i \in \{p+1, \ldots, p+c\}$) are different but close to each other, computation of a singular vector (resp. eigenvector) associated to the largest singular value (resp. eigenvalue) is ill-conditioned. Still, typically, the computed vector will be close to the two-dimensional subspace spanned by the singular vectors (resp. eigenvectors) associated with the two nonzero singular values (resp. eigenvalues). As a result, the next value of $\det(X)$ will be close to the "correct" next value, and the overall behavior of the algorithm will likely be satisfactory. This intuition can be formalized. Indeed, it is readily shown that the convergence properties established in Section 5 still hold if, in the definition of the map h_i (and similarly for h_{ij}), (i) if $i \in \{1, \ldots, p\}$, $\tilde{\xi}$ and $\tilde{\eta}$ are replaced with any vectors $\xi \in \mathbf{B}\mathcal{S}_{2i-1}$ and $\eta \in \mathbf{B}\mathcal{S}_{2i}$ such that

$$\langle \xi, \tilde{U}_i(X^k)\eta \rangle \geq \langle x_{2i-1}^k, \tilde{U}_i(X^k) x_{2i}^k \rangle + \theta(\langle \tilde{\xi}, \tilde{U}_i(X^k)\tilde{\eta} \rangle - \langle x_{2i-1}^k, \tilde{U}_i(X^k) x_{2i}^k \rangle);$$

and (ii) if $i \in \{p+1, \ldots, p+c\}$, $\tilde{\eta}$ is replaced with any vector $\eta \in \mathbf{B}\mathcal{S}_{2i}$ such that

$$|\langle \eta, \tilde{U}_i(X^k)\eta\rangle| \geq |\langle x_{2i}^k, \tilde{U}_i(X^k)x_{2i}^k\rangle| + \theta(|\langle \tilde{\eta}, \tilde{U}_i(X^k)\tilde{\eta}\rangle| - |\langle x_{2i}^k, \tilde{U}_i(X^k)x_{2i}^k\rangle|).$$

Here $\theta \in (0,1)$ is a fixed parameter. □

It remains to specify, for given $\hat{X}$, an efficient way to evalute $h_i(\hat{X})$, $i = 1, \ldots, p+c$.

6.1 Updating real columns

The following results are easily established.

Proposition 6.1 *: Let $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n] \in \mathbf{X}$, let $i \in \{1, \ldots, p\}$, let $\sigma_1 \geq \sigma_2$ be the top two singular values of $S_{2i-1}\tilde{U}_i(\hat{X})S_{2i}$, and for $j = 1, 2$, let μ_j, $\nu_j \in \mathbf{B}\mathbf{R}^m$ form a left-right singular vector pair associated with σ_j with the property that $\langle \mu_1, \mu_2\rangle = 0$ and $\langle \nu_1, \nu_2\rangle = 0$. Then for $i = 1, \ldots, p$, $h_i(\hat{X}) = [\hat{x}_1, \ldots, \hat{x}_{2i-2}, \tilde{\xi}, \tilde{\eta}, \hat{x}_{2i+1}, \ldots, \hat{x}_n]$, where $\begin{bmatrix} \tilde{\xi} \\ \tilde{\eta} \end{bmatrix} = \sqrt{2}\zeta/\|\zeta\|$ and*

(i) if $\sigma_1 > \sigma_2$, ζ is the orthogonal projection of $\begin{bmatrix} \hat{x}_{2i-1} \\ \hat{x}_{2i} \end{bmatrix}$ on the span of $\left\{\begin{bmatrix} S_{2i-1}\mu_1 \\ S_{2i}\nu_1 \end{bmatrix}\right\}$, i.e.,

$$\zeta = \begin{bmatrix} S_{2i-1}\mu_1 \\ S_{2i}\nu_1 \end{bmatrix} \begin{bmatrix} S_{2i-1}\mu_1 \\ S_{2i}\nu_1 \end{bmatrix}^\top \begin{bmatrix} \hat{x}_{2i-1} \\ \hat{x}_{2i} \end{bmatrix} = (\langle \hat{x}_{2i-1},\ S_{2i-1}\mu_1\rangle + \langle \hat{x}_{2i},\ S_{2i}\nu_1\rangle) \begin{bmatrix} S_{2i-1}\mu_1 \\ S_{2i}\nu_1 \end{bmatrix} \tag{34}$$

(ii) if $\sigma_1 = \sigma_2$, ζ is the orthogonal projection of $\begin{bmatrix} \hat{x}_{2i-1} \\ \hat{x}_{2i} \end{bmatrix}$ on the span of $\left\{\begin{bmatrix} S_{2i-1}\mu_1 \\ S_{2i}\nu_1 \end{bmatrix}, \begin{bmatrix} S_{2i-1}\mu_2 \\ S_{2i}\nu_2 \end{bmatrix}\right\}$, i.e.,

$$\zeta = \begin{bmatrix} S_{2i-1} & 0 \\ 0 & S_{2i} \end{bmatrix} \begin{bmatrix} \mu_1 & \mu_2 \\ \nu_1 & \nu_2 \end{bmatrix} \begin{bmatrix} \mu_1 & \mu_2 \\ \nu_1 & \nu_2 \end{bmatrix}^\top \begin{bmatrix} S_{2i-1} & 0 \\ 0 & S_{2i} \end{bmatrix}^\top \begin{bmatrix} \hat{x}_{2i-1} \\ \hat{x}_{2i} \end{bmatrix}. \tag{35}$$

For the prescribed real eigenvalues, the computation of $h_i(\hat{X})$, $i = 1, \ldots, p$, for a given $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n]$ is therefore carried out as follows.

(i) Compute an orthonormal basis $\{u, v\} \subset \mathbf{R}^n$ for the orthogonal complement of

$$\{\hat{x}_1, \ldots, \hat{x}_{2i-2}, \hat{x}_{2i+1} \ldots, \hat{x}_n\}.$$

(ii) Evaluate $M = S_{2i-1}^\top(uv^\top - vu^\top)S_{2i}$.

(iii) Compute associated orthonormal bases for the (1- or 2- dimensional) left and right singular spaces of M corresponding to the largest singular value of M.

(iv) Compute $(\tilde{\xi}, \tilde{\eta})$ as per Proposition 6.1 to obtain $h_i(\hat{X})$.

In Step (i), an appropriate choice for $\{u, v\}$ is the last two columns of the Q matrix in any QR factorization of

$$[\hat{x}_1, \ldots, \hat{x}_{2i-2}, \hat{x}_{2i+1}, \ldots, \hat{x}_{2p}, \hat{x}_{2p+2}^R, \hat{x}_{2p+2}^I, \ldots, \hat{x}_n^R, \hat{x}_n^I].$$

This factorization can be obtained at low cost by means of two successive column deletion operations followed by two successive column insertion operations, starting from the decomposition of

$$[\hat{x}_1, \ldots, \hat{x}_{2i-4}, \hat{x}_{2i-1}, \ldots, \hat{x}_{2p}, \hat{x}_{2p+2}^R, \hat{x}_{2p+2}^I, \ldots, \hat{x}_n^R, \hat{x}_n^I],$$

(where for notational simplicity, we have assumed that $i \geq 2$); see e.g., [14, Section 12.6]. (This is similar to the suggestion in [9] that one rank-one update of a related QR factorization be performed at each iteration of the QR algorithm.) The operation count is almost exactly twice that of the computation of $u_i(\hat{X})$ in the KNV iteration (in which a single column is updated). When $m << n$, the operation count of Steps (ii) to (iv) is small compared to that of Step (i).

6.2 Updating complex conjugate columns

Proposition 6.2 *: Let $\hat{X} = [\hat{x}_1, \ldots, \hat{x}_n] \in \mathbf{X}$, let $i \in \{p+1, \ldots, p+c\}$, let θ_1 and θ_2, with $|\theta_1| \geq |\theta_2|$, be the two nonzero eigenvalues of $S_{2i}^* \tilde{U}_i(\hat{X}) S_{2i}$, and let μ_ℓ, $\ell = 1, 2$, denote unit length eigenvectors associated with θ_ℓ with the property that $\langle \mu_1,\ \mu_2 \rangle = 0$. Then for $i = p+1, \ldots, p+c$, $h_i(\hat{X}) = [\hat{x}_1, \ldots, \hat{x}_{2i-2}, \overline{\tilde{\eta}}, \tilde{\eta}, \hat{x}_{2i+1}, \ldots, \hat{x}_n]$, where $\tilde{\eta} = \zeta / \|\zeta\|$ and*

(i) if $|\theta_1| > |\theta_2|$, ζ is the orthogonal projection of $\hat{x}_{2i}$ on the span $\{\mathcal{R}_{\mathbf{C}}(S_{2i}\mu_1)\}$, i.e.,

$$\zeta = S_{2i} \mu_1 \mu_1^* S_{2i}^* \hat{x}_{2i}, \tag{36}$$

(ii) if $|\theta_1| = |\theta_2|$, ζ is the orthogonal projection of $\hat{x}_{2i}$ on the span $\{\mathcal{R}_{\mathbf{C}}(S_{2i}\mu_1), \mathcal{R}_{\mathbf{C}}(S_{2i}\mu_2)\}$, i.e.,

$$\zeta = S_{2i} [\mu_1,\ \mu_2][\mu_1,\ \mu_2]^* S_{2i}^* \hat{x}_{2i}. \tag{37}$$

□

Therefore, given $i \in \{p+1, \ldots, p+c\}$, $h_i(\hat{X})$ can be computed as follows.

(i) Compute an orthonormal basis $\{\sqrt{2}u_R,\ \sqrt{2}u_I\} \subset \mathbf{R}^n$ for the orthogonal complement of

$$\{\hat{x}_1, \ldots, \hat{x}_{2p}\} \cup \{\hat{x}_{2j}^R, \hat{x}_{2j}^I, j = p+1, \ldots, p+c, j \neq i\},$$

(ii) Evaluate $M = S_{2i}^* (u\overline{u}^\top - \overline{u}u^\top) S_{2i}$ where $u = u_R + \mathrm{j} u_I$.

(iii) Compute orthonormal bases for the (1- or 2-dimensional) eigenspaces of M corresponding to the eigenvalue of M with the largest magnitude.

(iv) Compute $\tilde{\eta}$ as per Proposition 6.2 to obtain $h_i(\hat{X})$.

As was the case for the updating of real columns, the computation of (i) can be performed by means of two successive column deletion and then two successive column insertion operations starting from the QR factorization of

$$[\hat{x}_1, \ldots, \hat{x}_{2p}, \hat{x}_{2p+1}^R, \hat{x}_{2p+1}^I, \ldots, \hat{x}_{2i-2}^R, \hat{x}_{2i-2}^I, \hat{x}_{2i+2}^R, \hat{x}_{2i+2}^I, \ldots, \hat{x}_n^R, \hat{x}_n^I],$$

using rank one QR factorization, where for notational simplicity, we have assumed that $i < p+c$. The operation count is close to that for the update of two real columns.

7 Numerical Tests

Algorithms 5.1 to 5.3 were tried on thousands of randomly generated problems, with n ranging up to 32, and m up to 16. For every case, the sequence $\{X^k\}$ seemed to converge (recall that we were not able to prove that this is always the case). Furthermore while convergence of $\{X^k\}$ was sometimes slow, convergence of $\{\det(X^k)\}$ (and of sequences of other robustness indicators, see below) was always fast.

Systematic testing on 2000 examples was carried out to assess the relative merits of the KNV algorithm and of Algorithms 5.1 to 5.3. Note that the KNV algorithm, including a heuristic extension to the case where the set of prescribed eigenvalues includes complex conjugate pairs, is implemented as `place` in the MATLAB Control Toolbox. We also tested the "best column" variation on KNV suggested in Remark 3.2 (below we refer to this algorithm as "bestKNV"). In all cases, n was 10, m was 4, A and B were generated according to the normal distribution $\mathcal{N}(0,1)$, and computation was terminated when $|(\det(X^k) - \det(X^{k-1}))/\det(X^k)| < 0.001$. The sequences $\{X^k\}$ generated by the various algorithms often converged to different limit points $\check{X}$. To compare the relative robustness of the corresponding closed loops systems, we computed, in addition to our robustness indicator $\det(\check{X})$, the real and complex stability radii $\rho_{\mathbf{R}}(\check{X})$ and $\rho_{\mathbf{C}}(\check{X})$ of the corresponding closed loop matrix (see e.g., [15], [16]), and four robust indicators used in [9], namely the condition number $\kappa_2(\check{X})$ of $\check{X}$, the Frobenius norm $\|F(\check{X})\|$ of the corresponding feedback matrix, and the Euclidean and max norms $\|c(\check{X})\|$ and $\|c(\check{X})\|_\infty$ of a vector whose components are the individual condition numbers associated to the closed loop eigenvalues (low values are desired for the latter four indicators, while high values of $\det(\check{X})$, $\rho_{\mathbf{R}}(\check{X})$ and $\rho_{\mathbf{C}}(\check{X})$ are sought). (The stability radii were computed using M-files due to V. Balakrishnan and L. Qiu, based on the algorithms in [16–18].)

First, since the KNV algorithm was designed for assignment of real eigenvalues, we ran all five algorithms on 1000 problems with purely real prescribed eigenvalues. These eigenvalues were the negative of the absolute value of a random variable distributed according to $\mathcal{N}(0,1)$. For all algorithms, the initial state X^0 was the one generated by `place` as follows. First, orthonormal bases for the subspaces $\mathcal{S}_i$ are generated by means of two QR factorizations for each $\mathcal{S}_i$, using the MATLAB `qr` command (specifically, the basis for $\mathcal{S}_i$ consists of the last rank(B) columns of the Q matrix in the QR factorization of $(A - \lambda_i I)^* Q_1$, where the columns of Q_1 are the last $n - \text{rank}(B)$ columns of the Q matrix in the QR factorization B). Given these bases, for $i = 1, \ldots, p$, the ith column of X^0 is taken to be the first basis vector of $\mathcal{S}_i$ such that the first i columns of X^0 are numerically linearly independent as determined by the `rank` command in MATLAB.

On Figures 1 to 7 the performances of the five algorithms are compared, as measured by the seven robustness indicators. For each curve, the ordinate indicates the number of problems for which the ratio (resp. its inverse, in the case of "det", $\rho_{\mathbf{R}}$, and $\rho_{\mathbf{C}}$, for which large values are best) of the indicator value achieved by the corresponding algorithm to the best indicator value among all algorithms is less than (resp., more than) the abscissa value (a log scale is used for the abscissa). Figure 1 shows that, with regards to indicator "det", as expected, Algorithm 5.3 performs best (in particular, it achieves the highest value among all algorithms in almost 90% of the cases, and it never achieved less than about 70% of the highest value $(\det(\check{X}_{BEST})/\det(\check{X}) \leq 1.3)$), while bestKNV performs worst (somewhat surprising). With regard to the other indicators (Figures 2 to 7), bestKNV tends to always perform badly while `place`, Algorithm 5.1 and 5.3 are roughly equivalent and Algorithm 5.2 trails (also somewhat surprising to us). We also compared the number of iterations to convergence of the various algorithms, in two different modes. In "sequential" mode we counted as one "CPU time unit" 2 iterations of `place`, 1 iteration of Algorithm 5.1, 1/4 iteration of Algorithm 5.2, 2/9 iteration of bestKNV, and 1/44 iteration of Algorithm 5.3 (as these take about the same amount of CPU time on a sequential machine). In "parallel" mode we counted as one CPU time unit two iterations of `place` or bestKNV, and 1 iteration of Algorithm 5.1 to Algorithm 5.3 (as these take about the same amount of CPU time, when appropriately implemented on a parallel machine with a large enough number of processors). The results are shown on Figures 8 and 9 where, for each algorithm, the ordinate indicates the number of problems for which the ratio of number of CPU time units to convergence for the corresponding algorithm to the smallest number of CPU time units to convergence among all algorithms is less that the abscissa value (again, a log scale is used for the abscissa). It is clear that Algorithm 5.1 is best in sequential mode and that Algorithm 5.3 is best in parallel mode.

Next, we ran all algorithms except bestKNV (it does not extend to the complex case), on 1000 problems with $p = 3$ and $c = 2$, i.e., six prescribed real eigenvalues and two prescribed complex pairs of eigenvalues. The real eigenvalues and the real parts of the complex eigenvalues were generated as above, the imaginary parts were generated according to $\mathcal{N}(0,1)$. Again, for all four algorithms, the initial iterate X^0 was the one generated by `place`. In this general case, it is obtained as follows. First the orthonormal bases for the subspace $\mathcal{S}_i$ are obtained as in the purely real case. Then the first $2p$ columns X^0 are selected as before and, for $i = p+1, \ldots, p+c$, columns $2i-1$ and $2i$ of X^0 are taken to be the first basis vector of $\mathcal{S}_i$, and its complex conjugate, such that the first $2i$ columns of X^0 are numerically linearly independent as determined by the `rank` command in MATLAB.

The performance of the four algorithms as measured by the seven robustness indicators is portrayed in Figures 10 to 16, where the abscissa and ordinate are as in Figures 1 to 7. Here the difference between the quality of the solutions achieved by the various algorithms is much more striking than in the case of purely real eigenvalue assignment. Algorithm 5.2 and 5.3 are generally best, and `place` is always worst. Finally, as for the first 1000 test problems, we compared the number of CPU time units in sequential mode and in parallel mode. Here, in sequential mode, one CPU time unit corresponds to 1/5 of a "sweep" (2 iterations) of `place` (update all columns of X), 1 iteration of Algorithm 5.1, 1/4th of an iteration of Algorithm of 5.2, and 1/16th of an iteration of Algorithm 5.3. As for the case of purely real eigenvalues, in the parallel mode, one unit corresponds to 1/5th of a sweep of place, and one iteration of Algorithms 5.1 to 5.3. The results are presented in Figures 17 and 18 (same format as Figures 8 and 9). In both modes, `place` is fastest (but it converges to poor solutions!). Among the other three algorithms, as expected, Algorithm 5,1 is best in sequential mode, and Algorithm 5.3 is best in parallel mode.

8 Conclusion

New algorithms have been proposed for the problem of robust pole assignment by state feedback. Convergence properties have been established. The proposed algorithms are amenable to efficient implementation. Specifically, $h_i(X)$ (which updates two columns) can be computed in roughly twice the time it takes to compute $f_i(X)$, used in the KNV algorithm (which updates a single column). Extensive numerical tests suggest that the proposed algorithms are superior to their competitors.

A MATLAB M-file implementation of Algorithms 5.1 to 5.3 is available from the authors.

9 Acknowledgment

The second author wishes to thank Yung-Shan Chou for helpful discussions.

A Appendix

Proof of Proposition 4.1: Let P be a permutation matrix such that

$$XP = [\xi, \eta, X_{=}], \tag{38}$$

and let $Q \in \mathbf{R}^{n\times(n-2)}$ and $R \in \mathbf{R}^{(n-2)\times(n-2)}$ be any two matrices such that $Q^\top Q = I$ and

$$X_= = QR. \tag{39}$$

Using (38) and (39) we have

$$\begin{aligned}
(\det(X))^2 &= \det(X^\top X)\\
&= \det(P^\top X^\top XP) = \det([\xi,\eta,X_=]^\top[\xi,\eta,X_=])\\
&= \det\begin{pmatrix} \begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}[\xi,\eta] & \begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}X_= \\ X_=^\top[\xi,\eta] & X_=^\top X_= \end{pmatrix}\\
&= \det\left(\begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}[\xi,\eta] - \begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}X_=(X_=^\top X_=)^{-1}X_=^\top[\xi,\eta]\right)\det(X_=^\top X_=)\\
&= \det\left(\begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}[\xi,\eta] - \begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}QR(R^\top R)^{-1}R^\top Q^\top[\xi,\eta]\right)\det(X_=^\top X_=)\\
&= \det\left(\begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}[\xi,\eta] - \begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}QQ^\top[\xi,\eta]\right)\det(X_=^\top X_=)
\end{aligned}\tag{40}$$

Now, $U_i(\hat{X}) = (uv^\top - vu^\top)$ with $\{u,v\}$ an orthonormal basis for the nullspace of $X_=^\top$ satisfying (18). Thus $[Q,u,v]$ is orthogonal, therefore $QQ^\top + \begin{bmatrix} u & v\end{bmatrix}\begin{bmatrix}u^\top\\ v^\top\end{bmatrix} = I$ and

$$\begin{aligned}
(\det(X))^2 &= \det\left(\begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}[u,v]\begin{bmatrix}u^\top\\ v^\top\end{bmatrix}[\xi,\eta]\right)\det(X_=^\top X_=)\\
&= \left(\det\left(\begin{bmatrix}\xi^\top\\ \eta^\top\end{bmatrix}[u,v]\right)\right)^2\det(X_=^\top X_=)\\
&= (\langle\xi,u\rangle\langle\eta,v\rangle - \langle\xi,v\rangle\langle\eta,u\rangle)^2\det(X_=^\top X_=).
\end{aligned}\tag{41}$$

Next,

$$\operatorname{sgn}(\det(X)) = \operatorname{sgn}(\langle\xi,u\rangle\langle\eta,v\rangle - \langle\xi,v\rangle\langle\eta,u\rangle)$$

since (i) the arguments in both sides are quadratic in ξ,η, (ii) in view of (41) they vanish simultaneously and (iii) in view of (18) and since $\det(\hat{X}) > 0$ they are both positive at $\xi = \hat{x}_{2i-1}$, and $\eta = \hat{x}_{2i}$. Thus

$$\begin{aligned}
\det(X) &= (\langle\xi,u\rangle\langle\eta,v\rangle - \langle\xi,v\rangle\langle\eta,u\rangle)\sqrt{\det(X_=^\top X_=)}\\
&= \langle\xi,U_i(\hat{X})\eta\rangle\sqrt{\det(X_=^\top X_=)}.
\end{aligned}\tag{42}$$

The claim follows. □

Proof of Proposition 4.4: From Propositions 4.1 and 4.2,

$$\begin{aligned}
&\lim_{k\to\infty}(\langle x^k_{2i_p(k)-1}, U_{i_p(k)}(X^{k-1})x^k_{2i_p(k)}\rangle - \langle x^{k-1}_{2i_p(k)-1}, U_{i_p(k)}(X^{k-1})x^{k-1}_{2i_p(k)}\rangle)\sqrt{\det(X_=^{k-1\top}X_=^{k-1})}\\
=\ &\lim_{k\to\infty}(\det(X^k) - \det(X^{k-1})) = 0
\end{aligned}$$

Since, from Propositions 4.1 and 4.2, $\det(X_{\underline{=}}^{k-1^\top} X_{\underline{=}}^{k-1})$ is bounded away from zero, it follows that

$$(\langle x^k_{2i_p(k)-1}, U_{i_p(k)}(X^{k-1})x^k_{2i_p(k)}\rangle - \langle x^{k-1}_{2i_p(k)-1}, U_{i_p(k)}(X^{k-1})x^{k-1}_{2i_p(k)}\rangle) \to 0, \quad \text{as } k \to \infty. \tag{43}$$

Next, by construction,

$$\begin{bmatrix} x^k_{2i_p(k)-1} \\ x^k_{2i_p(k)} \end{bmatrix} = \begin{bmatrix} S_{2i_p(k)-1} & 0 \\ 0 & S_{2i_p(k)} \end{bmatrix} \begin{bmatrix} \mu^k \\ \nu^k \end{bmatrix}, \tag{44}$$

for some left-right unit length singular vector pair $\{\mu^k, \nu^k\}$ corresponding to the largest singular value σ_1^k of $S^\top_{2i_p(k)-1}U_{i_p(k)}(X^{k-p-1})S_{2i_p(k)}$, i.e., for some unit length μ^k, ν^k satisfying

$$\mu^{k^\top} S^\top_{2i_p(k)-1}U_{i_p(k)}(X^{k-1})S_{2i_p(k)}\nu^k = \sigma_1^k.$$

Letting $(\mu^k_-,\ \nu^k_-)$ be the unique vectors satisfying

$$x^{k-1}_{2i_p(k)-1} = S_{2i_p(k)-1}\mu^k_-, \quad x^{k-1}_{2i_p(k)} = S_{2i_p(k)}\nu^k_-, \tag{45}$$

i.e.,

$$\mu^k_- = S^\top_{2i_p(k)-1}x^{k-1}_{2i_p(k)-1}, \quad \nu^k_- = S^\top_{2i_p(k)}x^{k-1}_{2i_p(k)},$$

we can thus write (43) as

$$\sigma_1^k - \mu_-^{k^\top} S^\top_{2i_p(k)-1}U_{i_p(k)}(X^{k-1})S_{2i_p(k)}\nu^k_- \to 0 \ \text{ as } k \to \infty. \tag{46}$$

Now for $j = 1, 2$, let $(\mu^k_j,\ \nu^k_j)$ be a left-right unit length singular pair corresponding to the jth singular value σ^k_j of $S^\top_{2i_p(k)-1}U_{i_p(k)}(X^{k-1})S_{2i_p(k)}$ such that $\langle \mu^k_1, \mu^k_2\rangle = \langle \nu^k_1, \nu^k_2\rangle = 0$. Since $S^\top_{2i_p(k)-1}U_{i_p(k)}(X^{k-1})S_{2i_p(k)}$ has rank 2, we have

$$S^\top_{2i_p(k)-1}U_{i_p(k)}(X^{k-1})S_{2i_p(k)} = \sum_{j=1}^{2} \mu^k_j \nu^{k^\top}_j \sigma^k_j,$$

and (46) becomes

$$\sigma_1^k - \sum_{j=1}^{2} \langle \mu^k_-, \mu^k_j\rangle\langle \nu^k_j, \nu^k_-\rangle \sigma^k_j \to 0 \ \text{ as } k \to \infty. \tag{47}$$

Let $\mathcal{K}_1 = \{k : \lambda_{2i_p(k)} \neq \lambda_{2i_p(k)-1}\}$ and $\mathcal{K}_2 = \{k : \lambda_{2i_p(k)} = \lambda_{2i_p(k)-1}\}$. First, for all accumulation points $\hat{X}$ of the sequence $\{X^k\}$, it follows from our premise that

$$\sigma_1(S^\top_{2i_p(k)-1}U_{i_p(k)}(\hat{X})S_{2i_p(k)}) \neq \sigma_2(S^\top_{2i_p(k)-1}U_{i_p(k)}(\hat{X})S_{2i_p(k)}) \quad \forall k \in \mathcal{K}_1. \tag{48}$$

Therefore, since $U_{i_p(k)}(\cdot)$ is continuous, for k large enough, $k \in \mathcal{K}_1$, we have $\sigma^k_1 \neq \sigma^k_2$, hence $\mu^k_1 = \mu^k$, $\nu^k_1 = \nu^k$, and since μ^k, ν^k, μ^k_- and ν^k_- all have unit length, (47) implies

$$\langle \mu^k_-, \mu^k\rangle\langle \nu^k_-, \nu^k\rangle \to 1, \quad k \to \infty, \quad k \in \mathcal{K}_1,$$

which in turn implies

$$|\langle \mu^k_-, \mu^k\rangle| \to 1 \ \text{ as } k \to \infty, \quad k \in \mathcal{K}_1. \tag{49}$$

It follows that

$$\|\mu^k - \mu_-^k\|^2 = \|\mu^k\|^2 + \|\mu_-^k\|^2 - 2\langle \mu^k, \mu_-^k\rangle \to 0, \text{ as } k \to \infty, \quad k \in \mathcal{K}_1.$$

Similarly $\|\nu^k - \nu_-^k\|^2 \to 0$, as $k \to \infty$, $k \in \mathcal{K}_1$. Therefore, in view of (45) and (44),

$$\|X^k - X^{k-1}\| \to 0, \text{ as } k \to \infty, \quad k \in \mathcal{K}_1. \tag{50}$$

Next, for all $k \in \mathcal{K}_2$, let $\zeta_-^k = [\mu_-^{k^\top}, \nu_-^{k^\top}]^\top$, and for $j = 1, 2$, let $\zeta_j^k = [\mu_j^{k^\top}, \nu_j^{k^\top}]^\top$. Also for $k = 1, 2, \ldots$, let $\tilde{\zeta}^k$ be the length $\sqrt{2}$ vector along the orthogonal projection of ζ_-^k on the span of $\{\zeta_1^k, \zeta_2^k\}$, i.e.,

$$\tilde{\zeta}^k = \frac{\langle \zeta_-^k, \zeta_1^k\rangle}{2}\zeta_1^k + \frac{\langle \zeta_-^k, \zeta_2^k\rangle}{2}\zeta_2^k. \tag{51}$$

Since for $k \in \mathcal{K}_2$, $\lambda_{2i_p(k)} = \lambda_{2i_p(k)-1}$, then $\mathcal{S}_{2i_p(k)} = \mathcal{S}_{2i_p(k)-1}$, and their bases $S_{2i_p(k)}$ and $S_{2i_p(k)-1}$ can be chosen identical. Denote this matrix by S. Then

$$S_{2i_p(k)-1}^\top U_{i_p(k)}(X^{k-1}) S_{2i_p(k)} = S^\top(uv^\top - vu^\top)S$$

for some u, $v \in \mathbf{BR}^m$, and the nonzero singular values of this matrix are equal to the positive square roots of the eigenvalues of the 2×2 matrix

$$\begin{bmatrix} (S^\top u)^\top \\ (S^\top v)^\top \end{bmatrix} \begin{bmatrix} S^\top u, & S^\top v \end{bmatrix} \begin{bmatrix} (S^\top v)^\top \\ -(S^\top u)^\top \end{bmatrix} \begin{bmatrix} S^\top v, & -S^\top u \end{bmatrix} = (\|S^\top u\|^2 \|S^\top v\|^2 - (u^\top S S^\top v)^2) I. \tag{52}$$

Therefore $\sigma_1^k = \sigma_2^k$, and (47) yields

$$\langle \mu_-^k, \mu_1^k\rangle\langle \nu_-^k, \nu_1^k\rangle + \langle \mu_-^k, \mu_2^k\rangle\langle \nu_-^k, \nu_2^k\rangle \to 1, \text{ as } k \to \infty, \quad k \in \mathcal{K}_2. \tag{53}$$

Now, since μ_j^k, ν_j^k, $j = 1, 2$, μ_-^k and ν_-^k all have unit length, it follows from the Cauchy-Schwartz inequality that, for all $k \in \mathcal{K}_2$,

$$|\langle \mu_-^k, \mu_1^k\rangle\langle \nu_-^k, \nu_1^k\rangle + \langle \mu_-^k, \mu_2^k\rangle\langle \nu_-^k, \nu_2^k\rangle| \le \sqrt{\langle \mu_-^k, \mu_1^k\rangle^2 + \langle \mu_-^k, \mu_2^k\rangle^2}\sqrt{\langle \nu_-^k, \nu_1^k\rangle^2 + \langle \nu_-^k, \nu_2^k\rangle^2} \le 1. \tag{54}$$

This, in conjunction with (53), implies that

$$\langle \mu_-^k, \mu_1^k\rangle^2 + \langle \mu_-^k, \mu_2^k\rangle^2 \to 1, \text{ as } k \to \infty, \quad k \in \mathcal{K}_2, \tag{55}$$

and

$$\langle \nu_-^k, \nu_1^k\rangle^2 + \langle \nu_-^k, \nu_2^k\rangle^2 \to 1, \text{ as } k \to \infty, \quad k \in \mathcal{K}_2. \tag{56}$$

It follows from (53), (55), and (56) that

$$\langle \zeta_-^k, \zeta_1^k\rangle^2 + \langle \zeta_-^k, \zeta_2^k\rangle^2 \to 4 \text{ as } k \to \infty, \quad k \in \mathcal{K}_2.$$

Thus, since $\|\zeta_-^k\|^2 = 2$,

$$\langle \zeta_-^k, \zeta_-^k - \frac{\langle \zeta_-^k, \zeta_1^k\rangle}{2}\zeta_1^k - \frac{\langle \zeta_-^k, \zeta_2^k\rangle}{2}\zeta_2^k\rangle \to 0 \text{ as } k \to \infty, \quad k \in \mathcal{K}_2.$$

and, using (51) and the fact that $\frac{1}{2}(\langle \zeta_-^k, \zeta_1^k\rangle\zeta_1^k + \langle \zeta_-^k, \zeta_2^k\rangle\zeta_2^k)$ is the orthogonal projection of ζ_-^k on the span of $\{\zeta_1^k, \zeta_2^k\}$,

$$\|\zeta_-^k - \tilde{\zeta}^k\|^2 = \|\zeta_-^k - \frac{\langle \zeta_-^k, \zeta_1^k\rangle}{2}\zeta_1^k - \frac{\langle \zeta_-^k, \zeta_2^k\rangle}{2}\zeta_2^k\|^2 \to 0 \text{ as } k \to \infty, \quad k \in \mathcal{K}_2. \tag{57}$$

Now, since for $k \in \mathcal{K}_2$, $\sigma_1^k = \sigma_2^k$, it follows from the definition of the map $g_{i_p(k)}$ that

$$\begin{bmatrix} x^k_{2i_p(k)-1} \\ x^k_{2i_p(k)} \end{bmatrix} = \begin{bmatrix} S_{2i_p(k)-1} & 0 \\ 0 & S_{2i_p(k)} \end{bmatrix} \tilde{\zeta}^k.$$

This, together with (45) and (57) implies that

$$\|X^k - X^{k-1}\| \to 0, \quad k \to \infty, \quad k \in \mathcal{K}_2. \tag{58}$$

Since $\mathcal{K}$ is the union of $\mathcal{K}_1$ and $\mathcal{K}_2$, the claim then follows from (50) and (58). □

Proof of Proposition 5.1: Let P be the permutation matrix such that

$$XP = (\bar{\eta}, \eta, X_=), \tag{59}$$

and let $Q \in \mathbf{C}^{n\times(n-2)}$ and $R \in \mathbf{C}^{(n-2)\times(n-2)}$ be any two matrices such that $Q^*Q = I$ and

$$X_= = QR. \tag{60}$$

Using (59) and (60) we have

$$\begin{aligned}
|\det(X)|^2 &= \det(X^*X) \\
&= \det(P^\top X^*XP) = \det((\bar{\eta},\eta,X_=)^*(\bar{\eta},\eta,X_=)) \\
&= \det\begin{pmatrix} \begin{bmatrix} \bar{\eta}^* \\ \eta^* \end{bmatrix}[\bar{\eta},\eta] & \begin{bmatrix} \bar{\eta}^* \\ \eta^* \end{bmatrix} X_= \\ X_=^*[\bar{\eta},\eta] & X_=^*X_= \end{pmatrix} \\
&= \det\left(\begin{bmatrix} \bar{\eta}^* \\ \eta^* \end{bmatrix}[\bar{\eta},\eta] - \begin{bmatrix} \bar{\eta}^* \\ \eta^* \end{bmatrix} X_=(X_=^*X_=)^{-1}X_=^*[\bar{\eta},\eta]\right)\det(X_=^*X_=) \\
&= \det\left(\begin{bmatrix} \bar{\eta}^* \\ \eta^* \end{bmatrix}[\bar{\eta},\eta] - \begin{bmatrix} \bar{\eta}^* \\ \eta^* \end{bmatrix} QR(R^*R)^{-1}R^*Q^*[\bar{\eta},\eta]\right)\det(X_=^*X_=) \\
&= \det\left(\begin{bmatrix} \bar{\eta}^* \\ \eta^* \end{bmatrix}[\bar{\eta},\eta] - \begin{bmatrix} \bar{\eta}^* \\ \eta^* \end{bmatrix} QQ^*[\bar{\eta},\eta]\right)\det(X_=^*X_=)
\end{aligned} \tag{61}$$

Since by Lemma 5.1, $\{u, \bar{u}\}$ is an orthonormal basis for the nullspace of $X_=^*$, thus $[Q, u, \bar{u}]$ is unitary, therefore $QQ^* + \begin{bmatrix} u, \bar{u} \end{bmatrix}\begin{bmatrix} u^* \\ \bar{u}^* \end{bmatrix} = I$ and

$$\begin{aligned}
|\det(X)|^2 &= \det\left(\begin{bmatrix} \bar{\eta}^* \\ \eta^* \end{bmatrix}[u,\bar{u}]\begin{bmatrix} u^* \\ \bar{u}^* \end{bmatrix}[\bar{\eta},\eta]\right)\det(X_=^*X_=) \\
&= \left|\det\left(\begin{bmatrix} u^* \\ \bar{u}^* \end{bmatrix}[\bar{\eta},\eta]\right)\right|^2 \det(X_=^*X_=) \\
&= \left|\det\left[\begin{bmatrix} \bar{u}^\top \\ u^\top \end{bmatrix}[\bar{\eta},\eta]\right]\right|^2 \det(X_=^*X_=) \\
&= \left|\left((u^\top\eta)(\bar{u}^\top\bar{\eta}) - (\bar{u}^\top\eta)(u^\top\bar{\eta})\right)\right|^2 \det(X_=^*X_=) \\
&= \left|\langle \eta, (\bar{u}u^\top - u\bar{u}^\top)\eta\rangle\right|^2 \det(X_=^*X_=)
\end{aligned}$$

The claim follows by taking the square root on the both sides. □

References

[1] V. Gourishanker & K. Ramer, "Pole assignment with minimum eigenvalue sensitivity to plant parameter variations," *Internat. J. Control* 23 (1976), 493–504.

[2] K. Ramer & V. Gourishanker, "Utilization for the design freedom of pole assignment feedback controllers of unrestricted rank," *Int. J. Control* 24 (1976), 424–430.

[3] K. R. Cavin & S. P. Bhattacharyya, "Robust and well-conditioned eigenstructure assignment via Sylvester's equation," *American Control Conference* (1982).

[4] A. Dickman, "On the robustness of multivariable linear feedback system in state space representation," *IEEE Trans. Automat. Control* 32 (1987), 407–410.

[5] L.H. Keel, S.P. Bhattacharyya & J.W. Howze, "Robust control with structured perturbations," *IEEE Trans. Automat. Control* 33 (1988), 68–78.

[6] L. H. Keel & S. P. Bhattacharyya, "State space design of low-order stabilizers," *IEEE Trans. Automat. Control* 35 (1990), 82–86.

[7] J. G. Sun, "On numerical methods for robust pole assignment in control system design," *J. Computational Math.* 5 (1987), 119–134.

[8] R. Byers & S.G. Nash, "Approaches to Robust Pole Assignment," *International Journal of Control* 45 (1989), 97–107.

[9] J. Kautsky, N. K. Nichols & P. Van Dooren, "Robust pole assignment in linear state feedback," *Internat. J. Control* 41 (1985), 1129–1155.

[10] J. H. Wilkinson, *The algebraic eigenvalue problem*, Charendon Press, Oxford, 1965.

[11] Y. Yang, "A new condition number of eigenvalue and its application in control theory," *J Computational Math.* 7 (1989), 15–21.

[12] T. Kato, "Estimation of Iterated Matrices, with Application to the Von Neumann Condition," *Numerische Mathematik* 2 (1960), 22–29.

[13] M. Marcus, "An Inequality Connecting the p-Condition Number and the Determinant," *Numerische Mathematik* 4 (1962), 350–353.

[14] G. H. Golub & C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University press, Baltimore, 1989.

[15] D. Hinrichsen & A.J. Pritchard, "Stability Radius for Structured Perturbations and the Algebraic Riccati Equation," *Systems Control Lett.* 8 (1986), 105–113.

[16] L. Qiu, B. Bernhardsson, A. Rantzer, E.J. Davison, P.M. Young & J.C. Doyle, "On the Real Structured Stability Radius," *Proceedings of the 12th IFAC World Congress* 8 (1993), 71–78.

[17] S. Boyd & V. Balakrishnan, "A Regularity Result for the Singular Values of a Transfer Matrix and a Quadratically Convergent Algorithm for Computing its L_∞-norm," *Systems Control Lett.* 15 (1990), 1–7.

[18] L. Qiu, A. L. Tits & Y. Yang, "On the Computation of the Real Hurwitz-Stability Radius," in *Proceedings of the 32'th IEEE Conference on Decision and Control*, Orlando Florida, December 1994, 1905–1906.

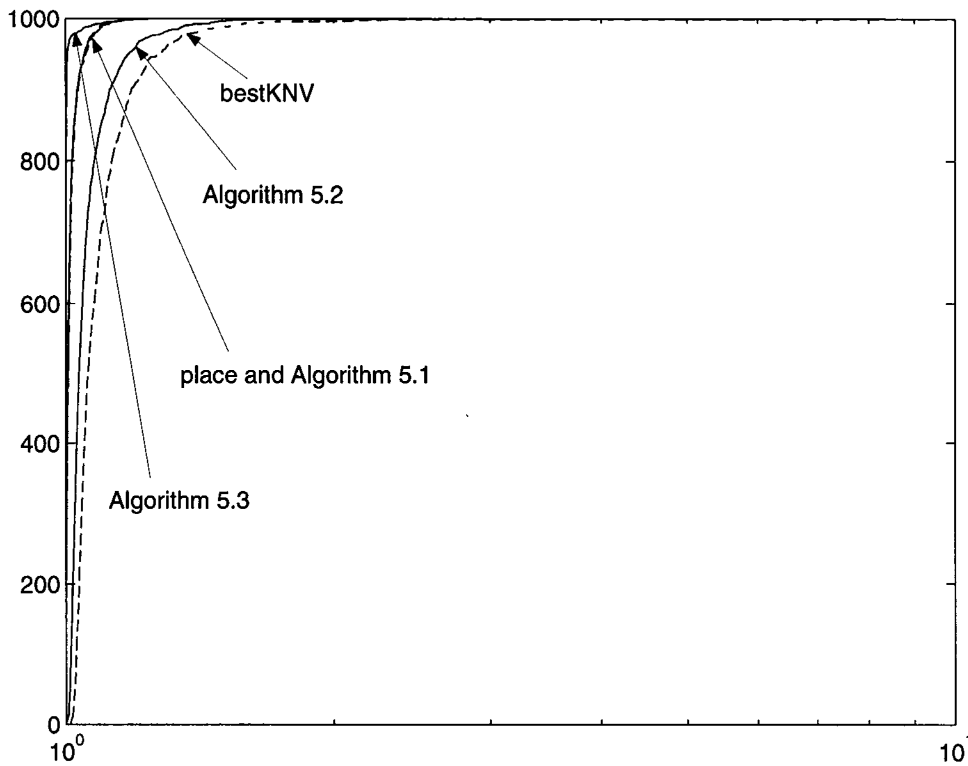

Figure 1: Real assignment: Cumulative distribution of $\det(\check{X}_{BEST})/\det(\check{X})$.

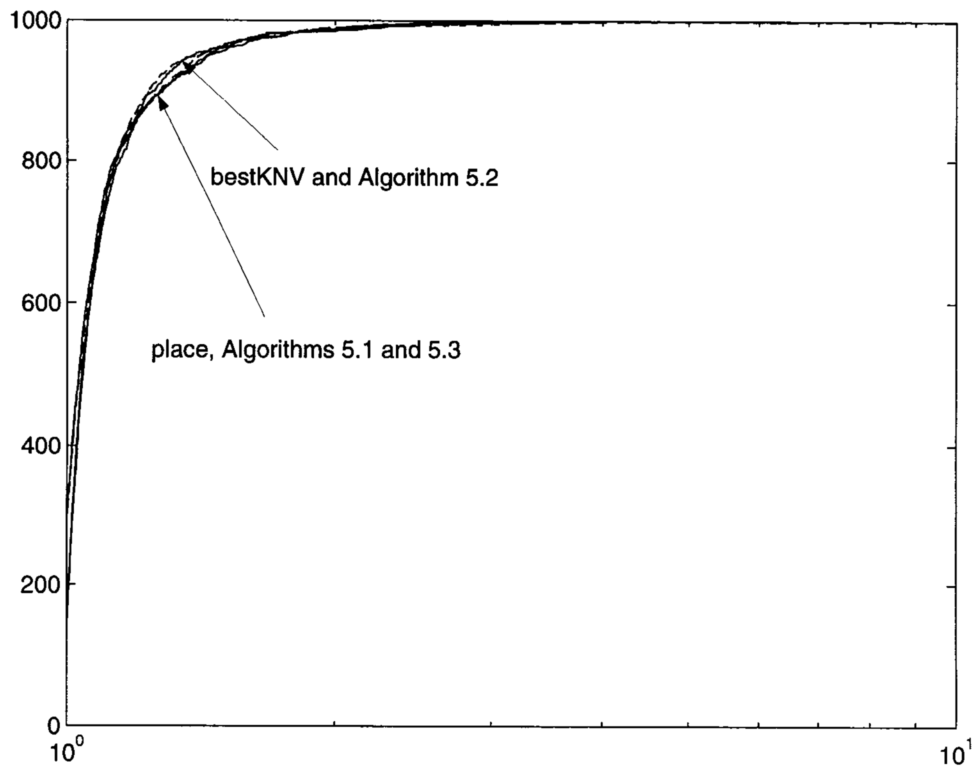

Figure 2: Real assignment: Cumulative distribution of $\rho_{\mathbf{R}}(\check{X}_{BEST})/\rho_{\mathbf{R}}(\check{X})$.

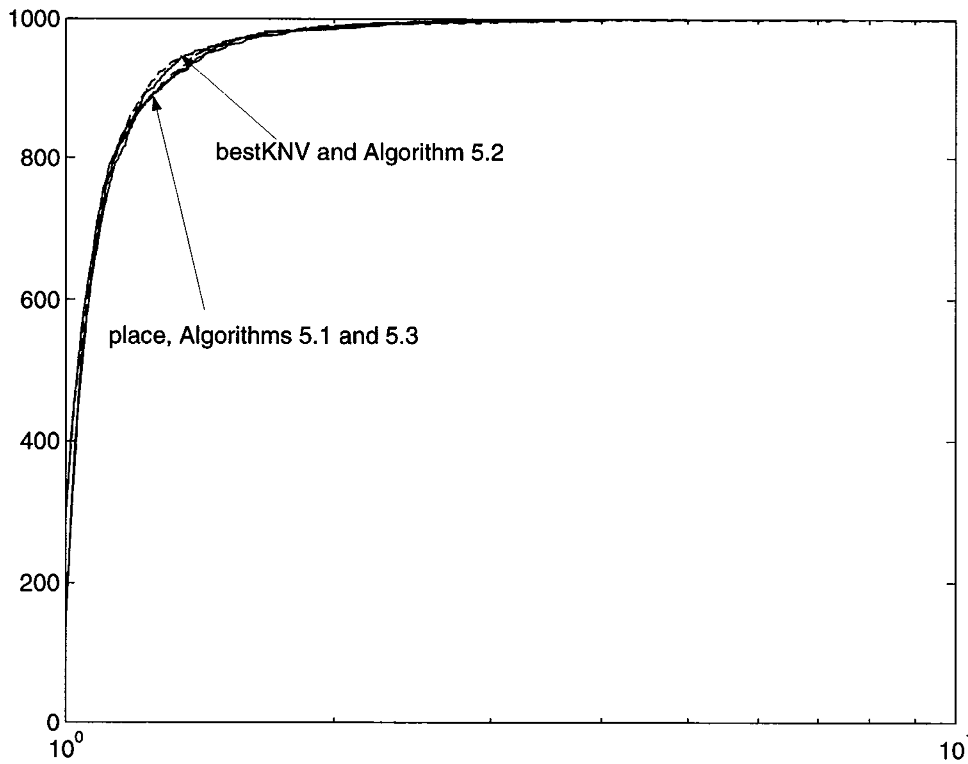

Figure 3: Real assignment: Cumulative distribution of $\rho_{\mathbf{C}}(\check{X}_{BEST})/\rho_{\mathbf{C}}(\check{X})$.

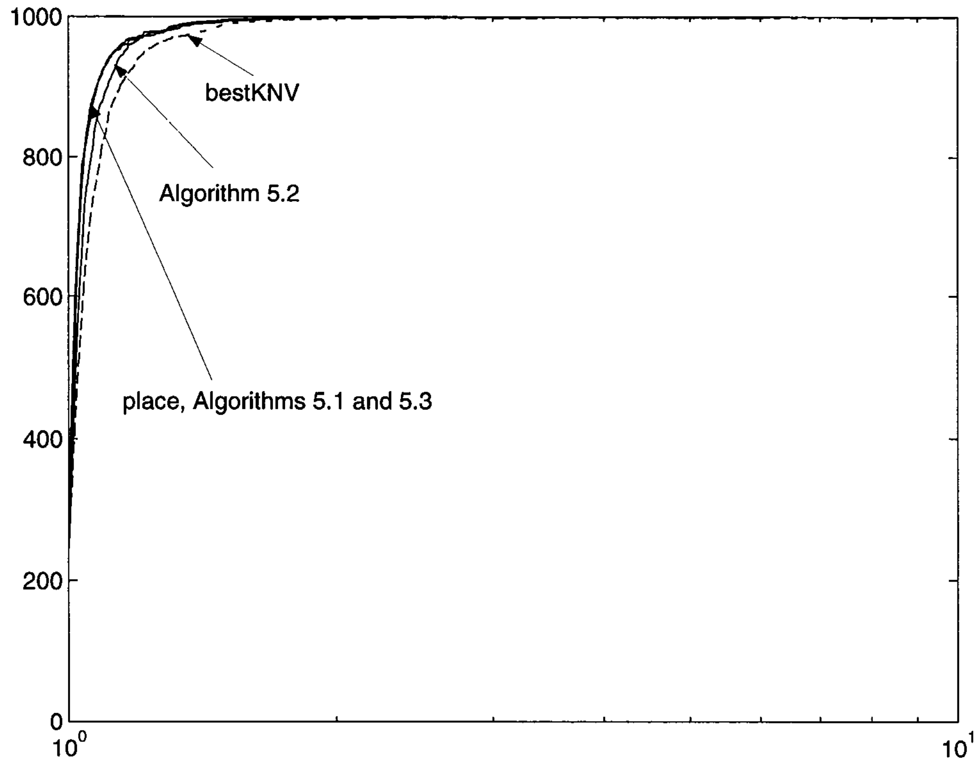

Figure 4: Real assignment: Cumulative distribution of $\kappa_2(\check{X})/\kappa_2(\check{X}_{BEST})$.

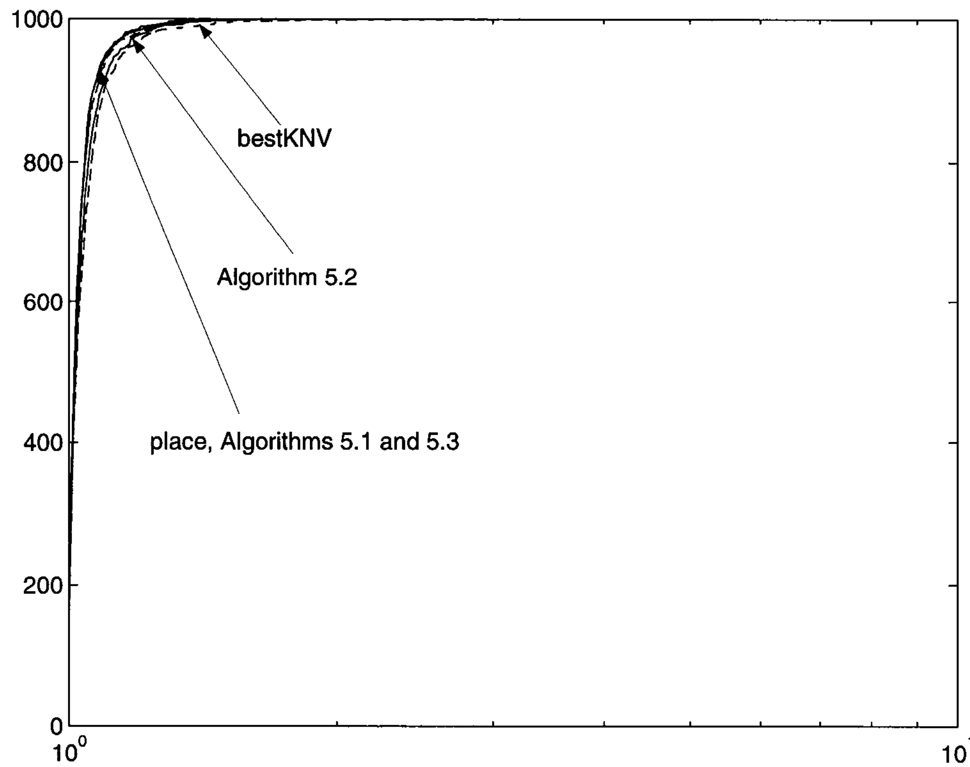

Figure 5: Real assignment: Cumulative distribution of $\|F(\check{X})\|/\|F(\check{X}_{BEST})\|$.

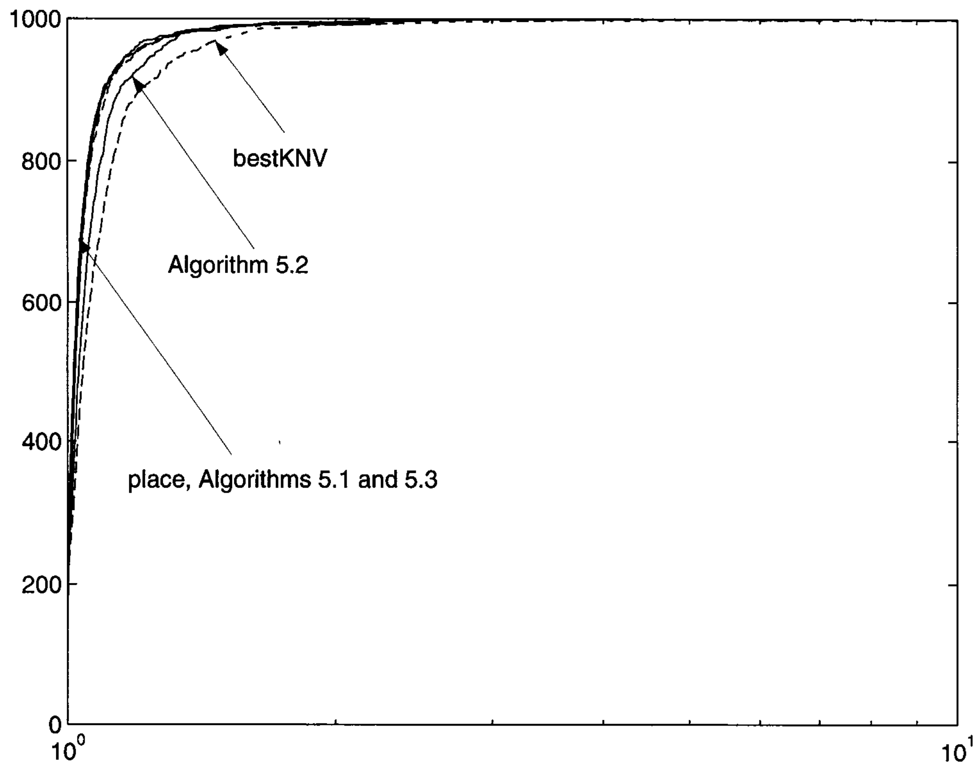

Figure 6: Real assignment: Cumulative distribution of $\|c(\check{X})\|/\|c(\check{X}_{BEST})\|$.

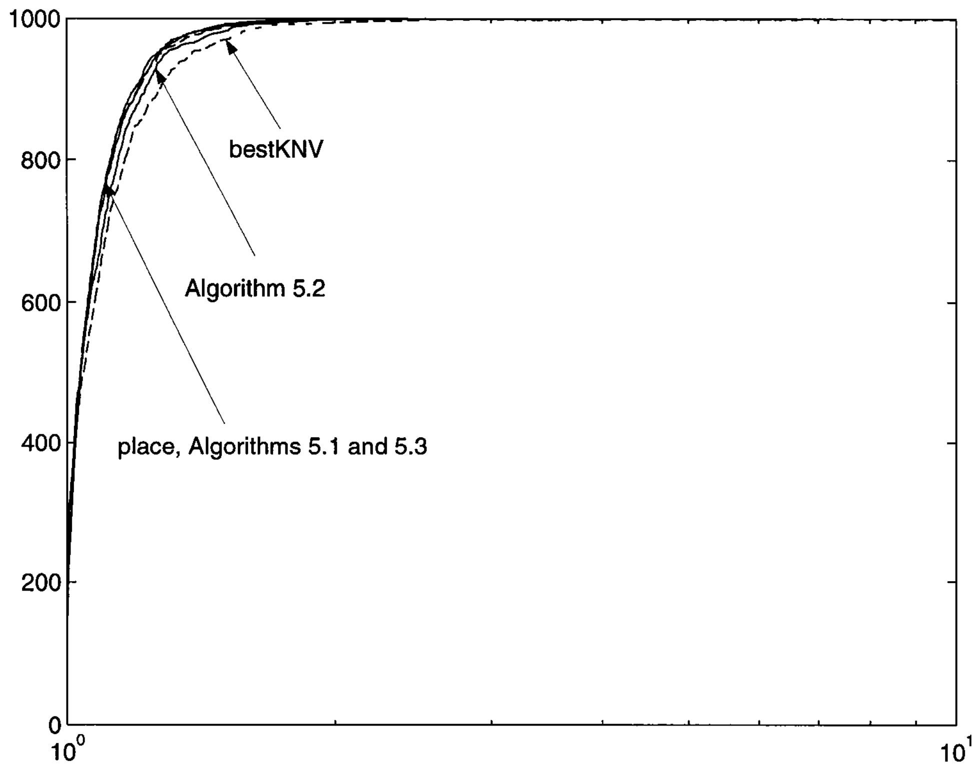

Figure 7: Real assignment: Cumulative distribution of $\|c(\check{X})\|_\infty / \|c(\check{X}_{BEST})\|_\infty$.

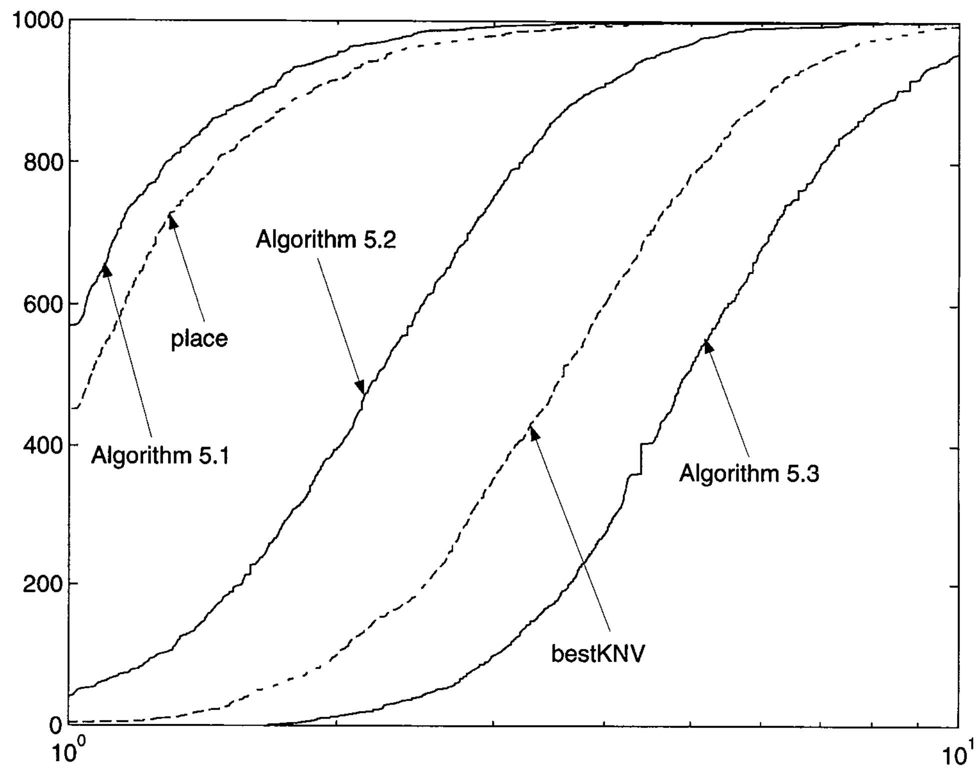

Figure 8: Real assignment: Cumulative distribution of (CPU time/best CPU time) counted according to sequential mode.

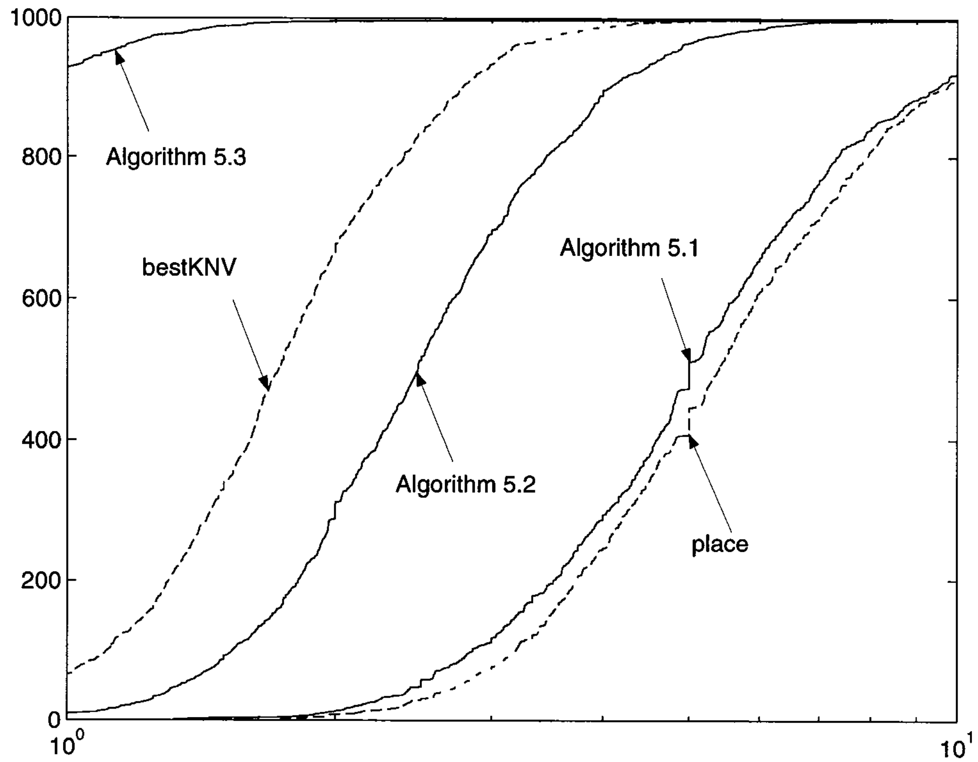

Figure 9: Real assignment: Cumulative distribution of (CPU time/best CPU time) counted according to parallel mode.

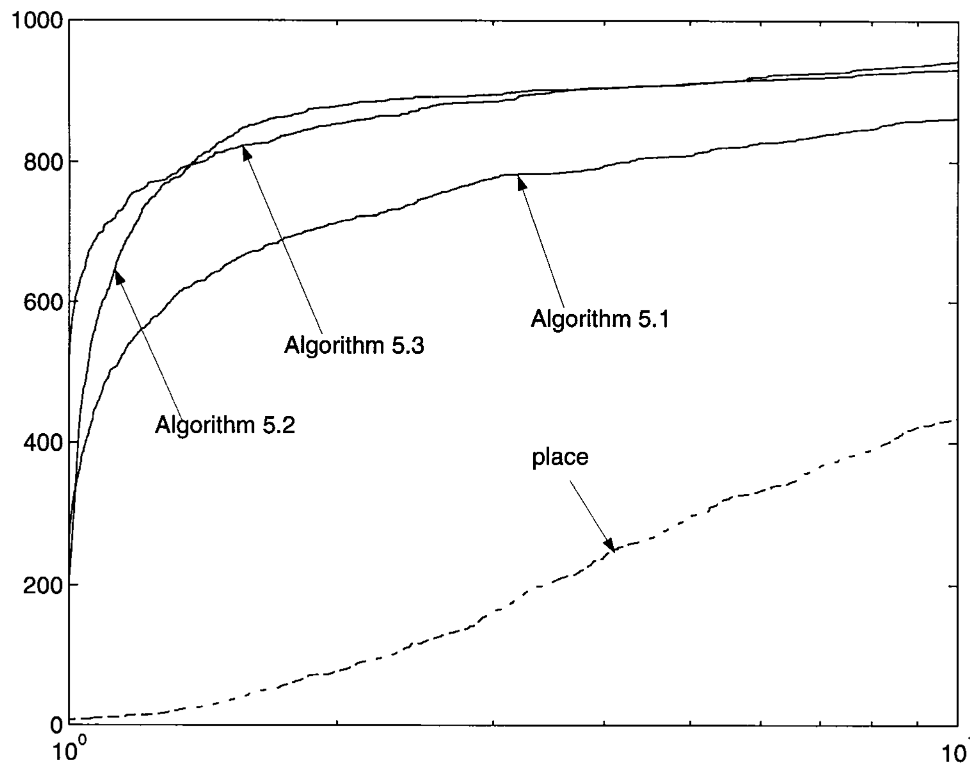

Figure 10: Mixed assignment: Cumulative distribution of $\det(\check{X}_{BEST})/\det(\check{X})$.

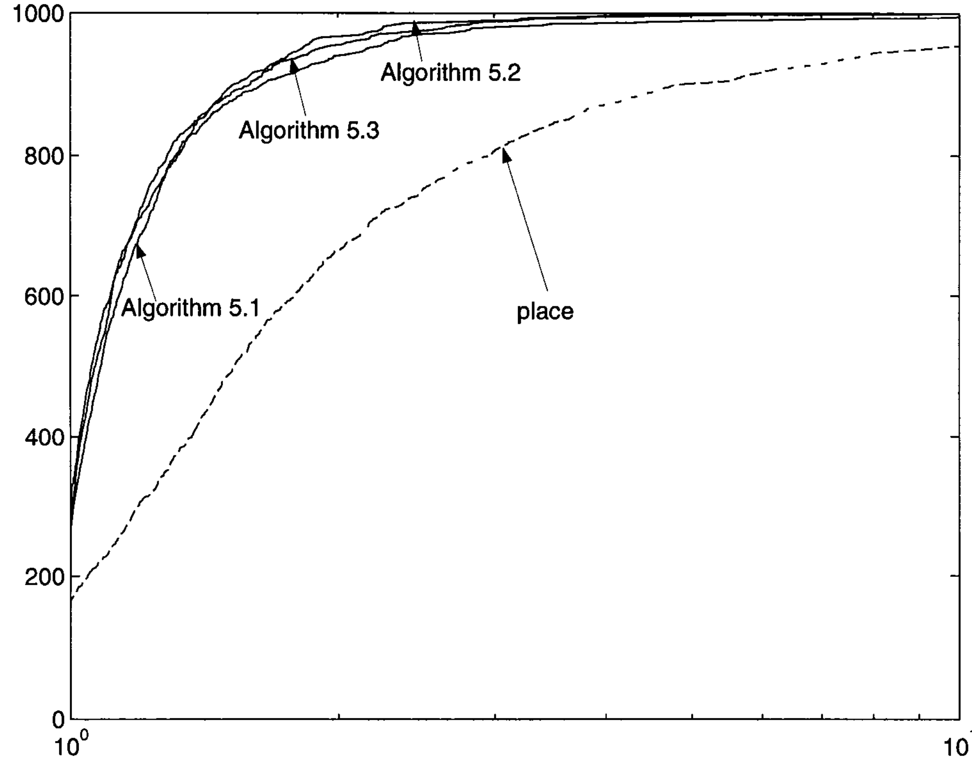

Figure 11: Mixed assignment: Cumulative distribution of $\rho_{\mathbf{R}}(\check{X}_{BEST})/\rho_{\mathbf{R}}(\check{X})$.

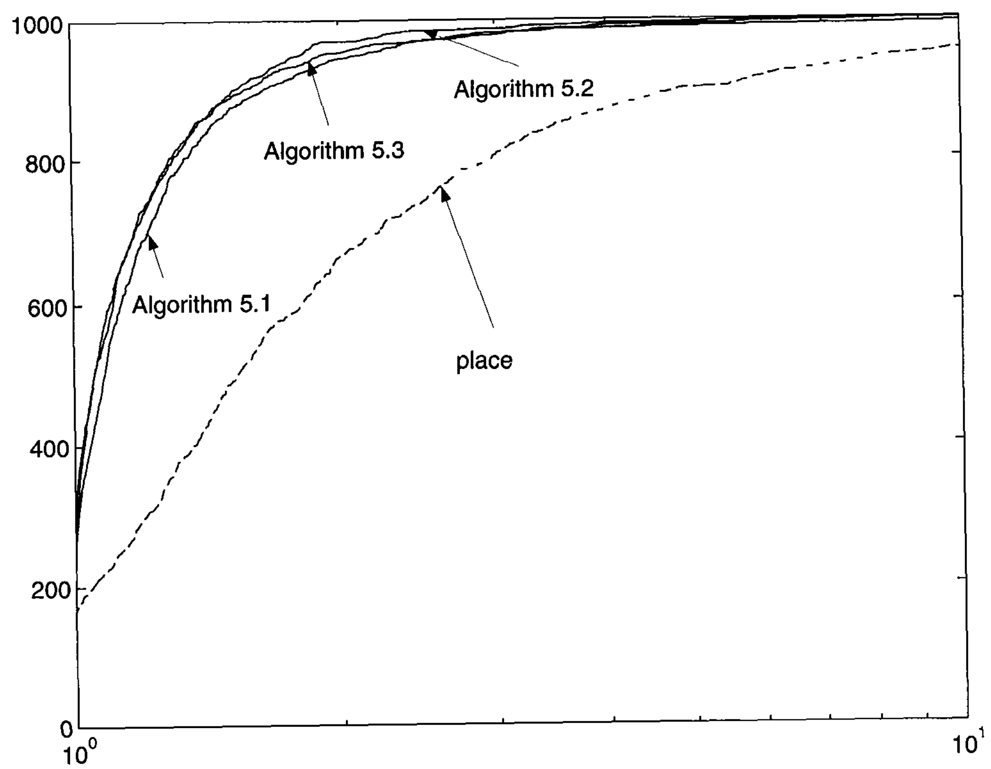

Figure 12: Mixed assignment: Cumulative distribution of $\rho_{\mathbf{C}}(\check{X}_{BEST})/\rho_{\mathbf{C}}(\check{X})$.

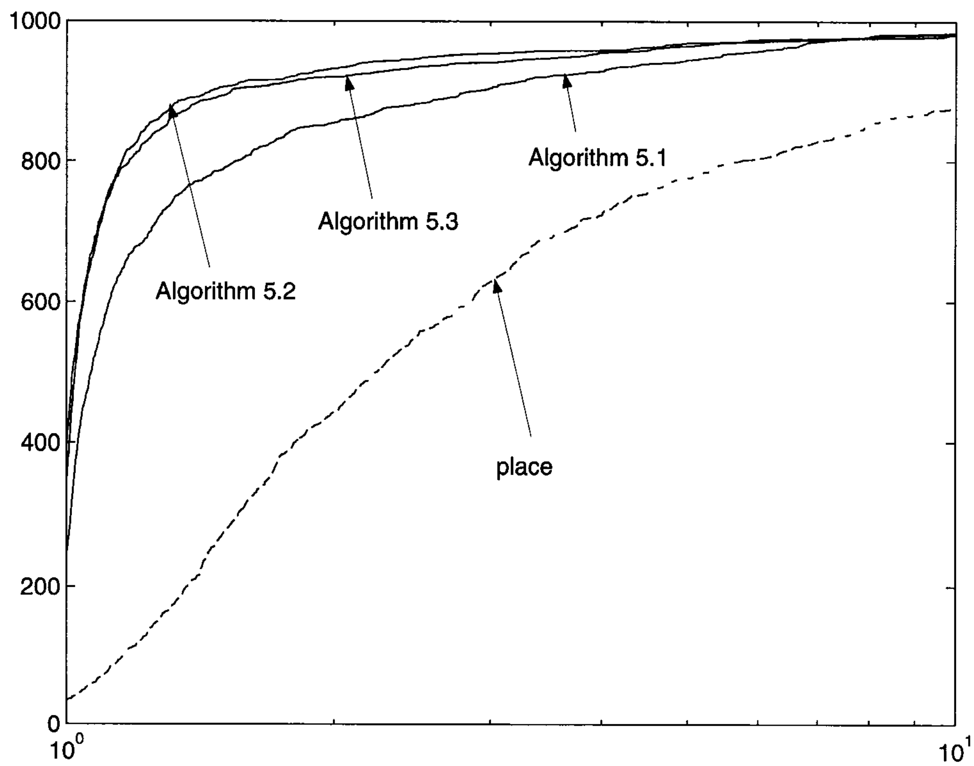

Figure 13: Mixed assignment: Cumulative distribution of $\kappa_2(\check{X})/\kappa_2(\check{X}_{BEST})$.

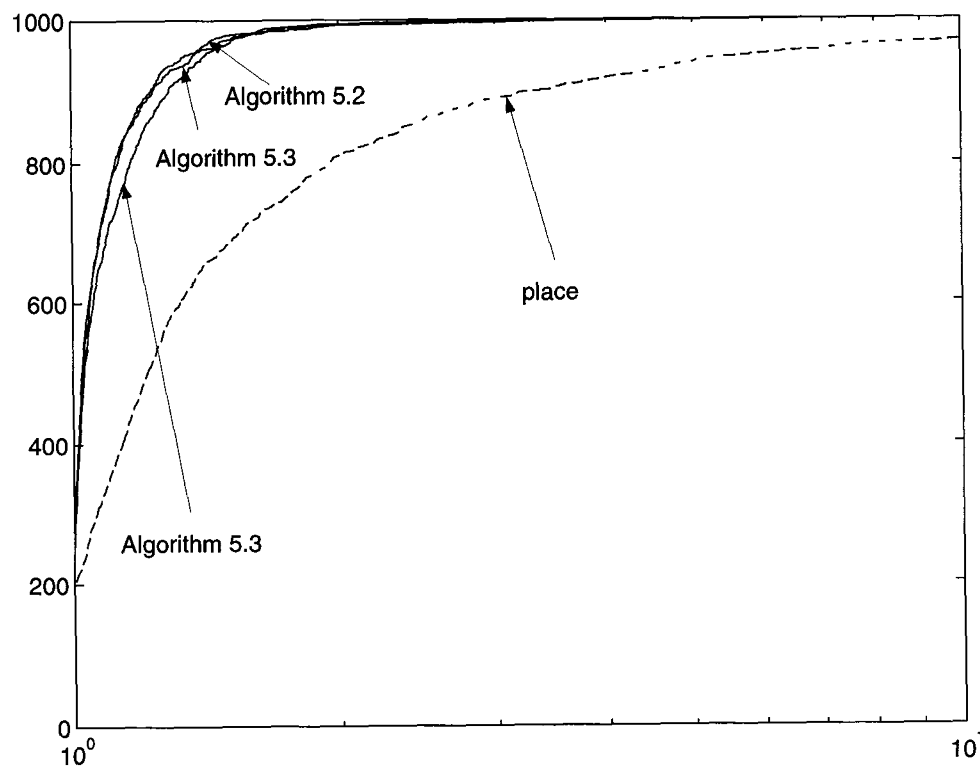

Figure 14: Mixed assignment: Cumulative distribution of $\|F(\check{X})\|/\|F(\check{X}_{BEST})\|$.

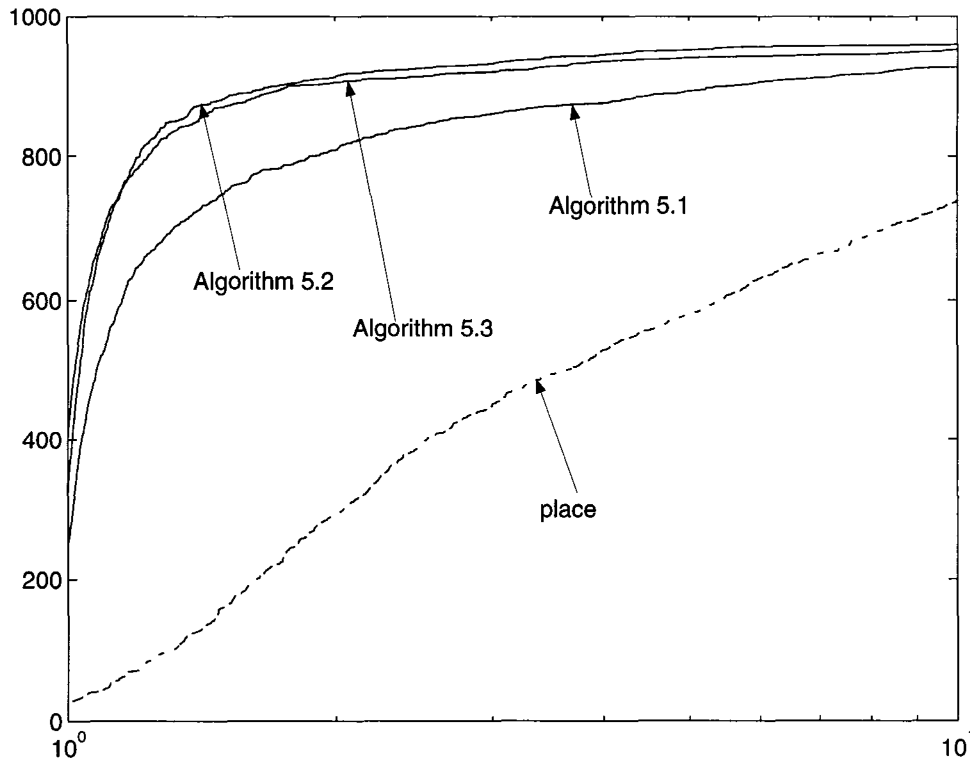

Figure 15: Mixed assignment: Cumulative distribution of $\|c(\check{X})\|/\|c(\check{X}_{BEST})\|$.

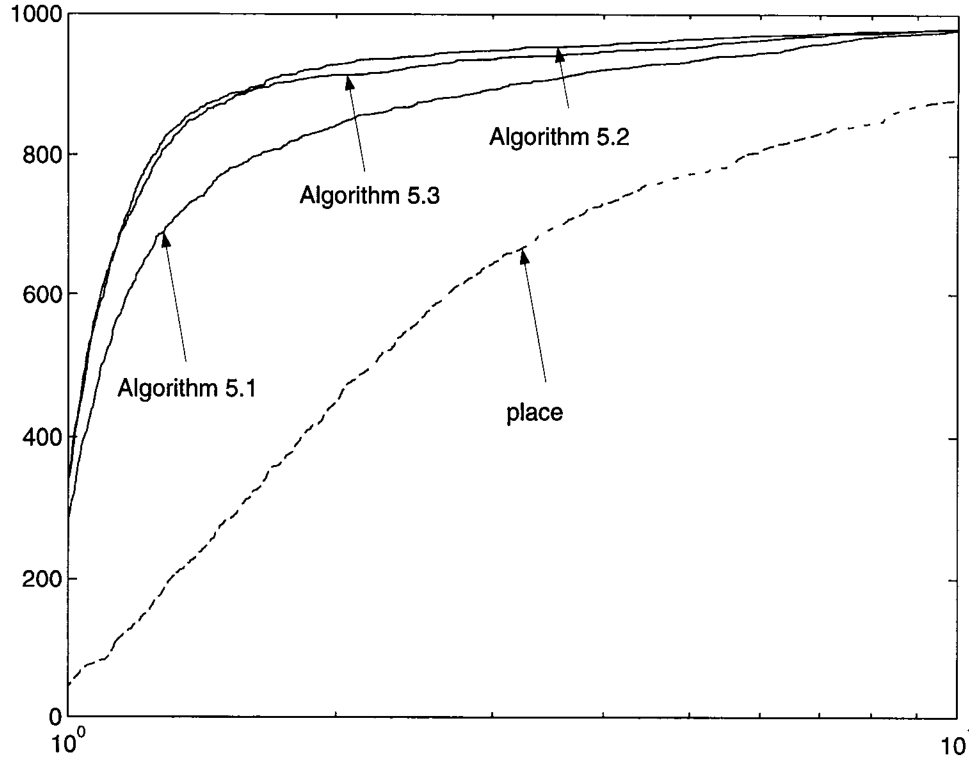

Figure 16: Mixed assignment: Cumulative distribution of $\|c(\check{X})\|_\infty / \|c(\check{X}_{BEST})\|_\infty$.

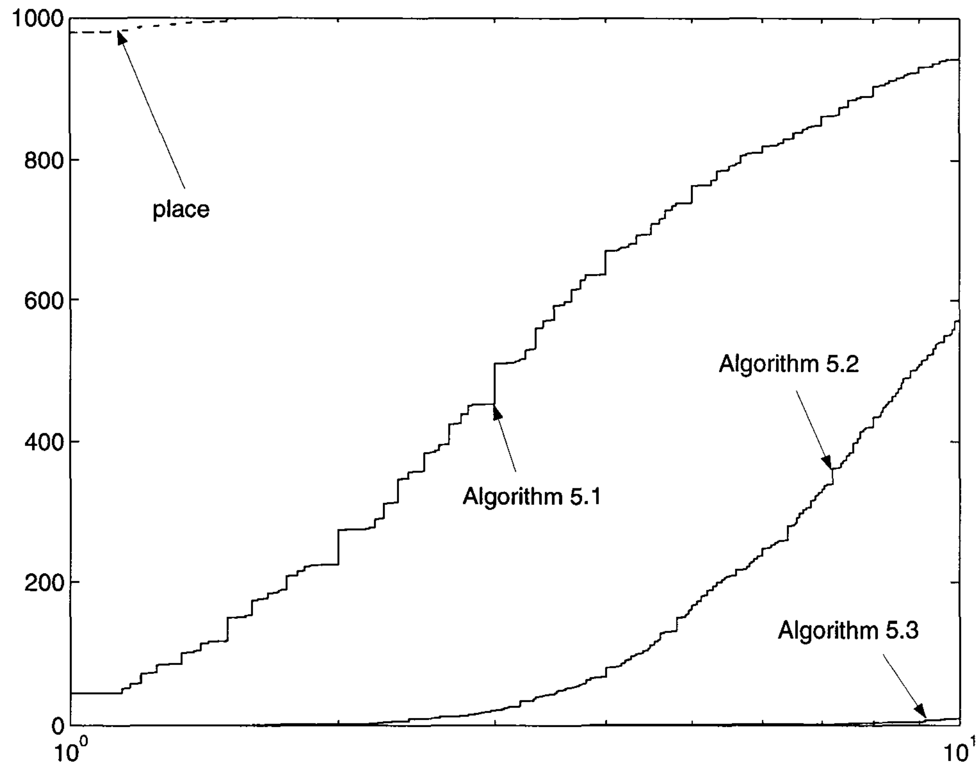

Figure 17: Mixed assignment: Cumulative distribution of (CPU time/best CPU time) counted according to sequential mode.

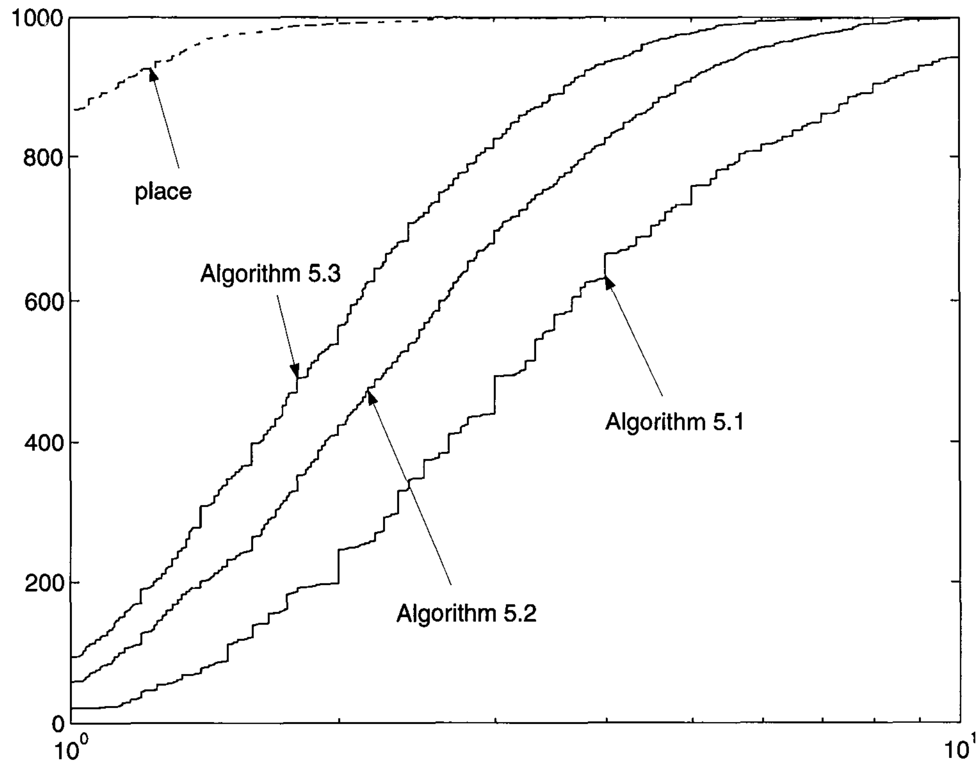

Figure 18: Mixed assignment: Cumulative distribution of (CPU time/best CPU time) counted according to parallel mode.