A Convex Approach to Frisch-Kalman Problem

Di Zhao, Anders Rantzer, and Li Qiu

***Abstract*— This paper proposes a convex approach to the Frisch-Kalman problem that identifies the linear relations among variables from noisy observations. The problem was proposed by Ragnar Frisch in 1930s, and was promoted and further developed by Rudolf Kalman later in 1980s. It is essentially a rank minimization problem with convex constraints. Regarding this problem, analytical results and heuristic methods have been pursued over a half century. The proposed convex method in this paper is demonstrated to outperform several commonly adopted heuristics when the noise components are relatively small compared with the underlying data.**

I. INTRODUCTION

The identification from noisy data has become an important problem of statistics and, via applications, of econometrics, biometrics, psychometrics and so on. Among various problems with different models on the data and noise, the Frisch-Kalman problem (scheme) [1]–[3], which is rooted in the work of Charles Spearman [4] in 1904, has attracted much attention and been investigated since 1930s [1]–[3], [5]–[9].

Given a finite family of n (random) variables $\{\omega_1, \omega_2, \ldots, \omega_n\}$ that are linearly dependent, we call them the true or underlying data, and in general, we have no direct access to their exact values. Instead, we can measure or observe their values in a noisy environment. The observed data $\{x_1, x_2, \ldots, x_n\}$ are corrupted by noise variables $\{\delta_1, \delta_2, \ldots, \delta_n\}$ additively, i.e.,

$$x_i = \omega_i + \delta_i,\ i = 1, 2, \ldots, n.$$

One may ask naturally: can we identify the linear relations among the true data from the observed (noisy) data samples? For this purpose, what else do we need to know about the data and noise? A well-established answer to the problems is given by the Frisch-Kalman scheme.

Denote by Σ the covariance matrix of the observed data $\{x_i\}$, which may be obtained from repeated experiments and measurements. Denote by Ω and Δ the covariance matrices of the true data $\{\omega_i\}$ and noise $\{\delta_i\}$, respectively. The key assumption in the Frisch-Kalman scheme is that the noise components are mutually uncorrelated and independent from the true data, in which case the following decomposition holds:

$$\Sigma = \Omega + \Delta,$$

and Δ is nonnegative and diagonal. Such a decomposition is called the factor analytic decomposition [6], [10] as is used in a statistical method — the factor analysis. The Frisch-Kalman scheme suggests one way to identify the linear relations via the minimization of the rank of $\Sigma - \Delta$ over all possible noise covariance matrices Δ. Regarding this scheme, a particularly important problem, which aims at finding the exact class of the observed covariance matrices Σ such that the maximum corank of $\Sigma - \Delta$ over all Δ is one, has been investigated since 1940s [2], [3], [6], [7], [9], [11].

The Frisch-Kalman problem is essentially a rank minimization problem with convex constraints. It is closely related to the low-rank matrix completion problem [12]–[14], where one wishes to complete a partially known matrix so that its rank is as small as possible. The nuclear norm minimization [13] has been pursued as a suitable heuristic for general rank minimization problems. In terms of the Frisch-Kalman problem, the nuclear norm heuristic reduces to the well-studied minimum trace factor analysis [5], [6], [9], [15]. As generalizations to the nuclear norm, a family of low-rank inducing norms, called the $r*$-norms [16], [17] or spectral r-support norms [18], have been recently proposed, which improve the performance of the nuclear norm heuristic for rank minimization problems. In addition to the low-rank inducing norms, other surrogates have been studied for the rank function, for example, the logarithm of the determinant (log-det) [19].

In this paper, we propose a convex approach to the Frisch-Kalman problem by first reformulating the problem into a norm minimization problem with a rank constraint, then relaxing it into a convex problem that is essentially a semi-definite programming (SDP) [20]. The reformulated Frisch-Kalman problem additionally penalizes the variances of noise components, which is motivated by the application scenarios when the noise are well-bounded with respect to the underlying data. For example, population census and mapping in developed countries [21], channel estimations in slow fading channels [22], long-term global surface temperature measure [23], and so on. Comparisons with the existing heuristic methods, including the nuclear norm minimization [13], the $r*$-norm minimization [16] and the log-det heuristic [19], show that the proposed method has high success rates and strictly outperforms the others when the noise components are well bounded with respect to the underlying data.

*This work was supported in part by the Research Grants Council of Hong Kong Special Administrative Region, China, under Theme-based Research Scheme T23-701/14N, and by the Swedish Research Council, Sweden, under contract 2016-04764.

D. Zhao and L. Qiu are with the Department of Electronic & Computer Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China. Emails: dzhaoaa@ust.hk, eeqiu@ust.hk

A. Rantzer is with the Department of Automatic Control and the Excellence Center ELLIIT, Lund University, Lund, Sweden. Email: rantzer@control.lth.se

The rest of the paper is organized as follows. In Section II, basic notation and preliminary results are introduced. In Section III, the main algorithm is developed. In Section IV, comparisons with the existing heuristic methods are shown via simulations. Finally, in Section V, the study is concluded and future research directions are introduced.

II. Preliminaries

A. Notation

Let $\mathbb{R}$ be the real field, and $\mathbb{R}^n$ be the linear space of n-dimensional vectors over $\mathbb{R}$. For $x \in \mathbb{R}^n$, its Euclidean norm is denoted by $\|x\|$.

For matrix $A \in \mathbb{R}^{m\times n}$, its element in the ith row and jth column is denoted by $[A]_{ij}$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, its transpose is by A^T, its range is by

$$\mathcal{R}(A) := \{y \in \mathbb{R}^m |\ y = Ax \text{ for some } x \in \mathbb{R}^n\},$$

its kernel is by

$$\mathcal{K}(A) := \{x \in \mathbb{R}^n |\ Ax = 0\},$$

and its kth singular value is by $\sigma_k(A)$, $k = 1, 2, \ldots, l$, in a nonincreasing order, where $l = \min\{m, n\}$. The largest and smallest singular values are specially denoted by $\bar{\sigma}(A) := \sigma_1(A)$ and $\underline{\sigma}(A) := \sigma_l(A)$, respectively. The operator norm (spectral norm) and the Frobenius norm of A are respectively denoted by

$$\|A\| := \bar{\sigma}(A) \text{ and } \|A\|_F := \sqrt{\sum_{k=1}^{l} \sigma_k^2(A)}.$$

The r-norm [16] of A, $r = 1, 2, \ldots, l$, is defined via

$$\|A\|_r := \sqrt{\sum_{i=1}^{r} \sigma_i^2(A)}.$$

Clearly, $\|A\|_F = \|A\|_l$. Denote its singular value decomposition (SVD) as

$$A = USV^T = \sum_{i=1}^{l} \sigma_i(A) u_i v_i^T,$$

where U, V are unitary. For $A, B \in \mathbb{R}^{m\times n}$, their inner product is defined via

$$\langle A, B\rangle := \mathrm{tr}(A^T B).$$

For $X \in \mathbb{R}^{n\times n}$, the diagonal matrix that keeps the diagonal terms of X is denoted by $\mathrm{diag}(X)$. For $x \in \mathbb{R}^n$, the diagonal matrix with its ith diagonal term given by x_i is denoted by $\mathrm{diag}^*(x)$.

Some frequently used special sets of matrices are as follows.

- Denote by $\mathbb{S}^n$ the set of all symmetric matrices in $\mathbb{R}^{n\times n}$.
- Denote by $\mathbb{S}^n_0$ the set of symmetric matrices in $\mathbb{R}^{n\times n}$ with zero diagonals.
- Denote by $\mathbb{S}^n_+$ ($\mathbb{S}^n_{++}$, resp.) the set of all positive semi-definite (definite, resp.) matrices in $\mathbb{R}^{n\times n}$.
- Denote by $\mathbb{D}^n$ the set of all diagonal matrices in $\mathbb{R}^{n\times n}$.
- Denote by $\mathbb{D}^n_+$ the set of all nonnegative diagonal matrices in $\mathbb{R}^{n\times n}$.

B. Standard Low-Rank Approximation

Let $A \in \mathbb{R}^{m\times n}$ and $l = \min\{m, n\}$. Consider the following standard rank approximation problem:

$$\min_B \{\|A - B\|_F \mid \mathrm{rank}(B) \le r\}, \tag{1}$$

for $r = 1, 2, \ldots, l$. Based on the Schmidt-Mirsky theorem [24, Chapter IV], all solutions to problem (1) is given by

$$\mathrm{svd}_r(A) := \left\{ \sum_{i=1}^{r} \sigma_i(A) u_i v_i^T \;\middle|\; A = \sum_{i=1}^{l} \sigma_i(A) u_i v_i^T \text{ is SVD} \right\},$$

which is called the set of all standard rth order SVD-approximation to A. Clearly, for every $B \in \mathrm{svd}_r(A)$, it holds that

$$\mathrm{rank}(B) = r. \tag{2}$$

When $\sigma_r(A) > \sigma_{r+1}(A)$, $\mathrm{svd}_r(A)$ is a singleton and its only element is denoted by

$$[A]_r := \sum_{i=1}^{r} \sigma_i(A) u_i v_i^T.$$

The optimal value to the problem is given by

$$\min_B \{\|A - B\|_F \mid \mathrm{rank}(B) \le r\} = \left\| \begin{bmatrix} \sigma_{r+1}(A) & \cdots & \sigma_l(A) \end{bmatrix} \right\| = \sqrt{\|A\|_F^2 - \|A\|_r^2}.$$

C. Frisch-Kalman Problem

The Frisch-Kalman problem is defined via the following optimization [1]–[3], [9].

Definition 1. *Given $\Sigma \in \mathbb{S}^n_{++}$, determine*

$$\begin{aligned} \mathrm{mr}(\Sigma) := \min_{\Omega,\Delta} \{ & \mathrm{rank}(\Omega) |\ \Sigma = \Omega + \Delta, \\ & \Omega \in \mathbb{S}^n_+, \Delta \in \mathbb{D}^n_+\}. \end{aligned} \tag{3}$$

A matrix Δ is said to be feasible to problem (3), if Δ is diagonal and $\Sigma \ge \Delta \ge 0$. A trivial upper bound to the problem is given by $\mathrm{mr}(\Sigma) \le n - 1$, which can be obtained by selecting a feasible $\Delta = \underline{\sigma}(\Sigma) I$. The Frisch-Kalman problem is, in general, non-convex, and many heuristic convex approaches have been proposed and investigated [6], [9], [13], [14].

III. Proposed Convex Approach

In this section, we develop a convex approach to solving the Frisch-Kalman problem.

A. Reformulation and Relaxation

Consider the factor analytic decomposition $\Sigma = \Omega + \Delta$. In the context of Frisch-Kalman scheme, $\Omega \in \mathbb{S}^n_+$ is the unknown covariance matrix of some linearly dependent true data variables, and hence it is expected to have a low rank. The matrix Δ is the covariance matrix of an uncorrelated noise vector, and hence it must be nonnegative diagonal. Finally, $\Sigma \in \mathbb{S}^n_{++}$, as the sum of Ω and Δ, is the covariance matrix of the noisy data under the assumption that the data and noise are independent. In many practical situations, the variances of the noise may be much smaller than those of the true data. To take advantage of the additional preknowledge, we may penalize the "size of noise" as in the following reformulation of Frisch-Kalman problem.

Given an integer $r \in [1, n]$, we reformulate the Frisch-Kalman problem into the following norm minimization problem with a rank constraint:

$$\min_{\Omega}\{\|\Sigma - \Omega\|_F^2 |\ \mathrm{rank}(\Omega) \le r,\ \Sigma \ge \Omega \ge 0,\ \Sigma - \Omega \in \mathbb{D}^n\}. \tag{4}$$

Here, the object function is simply the sum of squares of all the entries in the diagonal matrix $\Sigma - \Omega$. The rank function is moved from the object function in (3) to the constraints in (4). If this problem is feasible, then we obtain immediately $\mathrm{mr}(\Sigma) \le r$. In other words, we can search for $\mathrm{mr}(\Sigma)$ via solving a sequence of feasibility problems of (4) with different levels of $r \in [1, n]$. However, the reformulated problem (4) is still non-convex. To proceed, we develop some further relaxations in the following.

We introduce a symmetric matrix with zero diagonals, namely, $\Lambda \in \mathbb{S}_0^n$, as the dual variable. Based on the reformulated Frisch-Kalman problem (4), we have the following series of equalities and inequalities:

$$\begin{aligned}
&\min_{\Omega}\{\|\Sigma - \Omega\|_F^2 |\ \mathrm{rank}(\Omega) \le r, \Sigma \ge \Omega \ge 0, \Sigma - \Omega \in \mathbb{D}^n\} \\
&= \min_{\Omega}\max_{\Lambda}\{\|\Sigma - \Omega\|_F^2 + 2\langle \Lambda, \Sigma - \Omega\rangle | \\
&\qquad\qquad \mathrm{rank}(\Omega) \le r,\ \Sigma \ge \Omega \ge 0, \Lambda \in \mathbb{S}_0^n\} \\
&\ge \max_{\Lambda}\min_{\Omega}\{\|\Sigma - \Omega\|_F^2 + 2\langle \Lambda, \Sigma - \Omega\rangle | \\
&\qquad\qquad \mathrm{rank}(\Omega) \le r,\ \Sigma \ge \Omega \ge 0, \Lambda \in \mathbb{S}_0^n\} \\
&\ge \max_{\Lambda}\min_{\Omega}\{\|\Sigma + \Lambda - \Omega\|_F^2 - \|\Sigma + \Lambda\|_F^2 + 2\langle \Lambda, \Sigma\rangle \\
&\qquad\qquad + \|\Sigma\|_F^2 |\ \mathrm{rank}(\Omega) \le r, \Lambda \in \mathbb{S}_0^n\} \\
&= \max_{\Lambda}\{-\|\Sigma + \Lambda\|_r^2 + 2\langle \Lambda, \Sigma\rangle + \|\Sigma\|_F^2 |\ \Lambda \in \mathbb{S}_0^n\},
\end{aligned} \tag{5}$$

where the first equality is due to that the maximization over $\Lambda \in \mathbb{S}_0^n$ forces $\Sigma - \Omega$ to be diagonal, the first inequality follows from the max-min inequality, the last inequality is due to that the constraint of $\Sigma \ge \Omega \ge 0$ is removed, and the last equality follows from the standard SVD-approximation to $\Sigma + \Lambda$ in Section II-B.

Here, the problem (5) is a maximization of a concave function with convex constraints, hence it is a convex problem as in (6), which is our targeted convex relaxation to the original non-convex problem. Using similar tricks in [16], we can equivalently transform (5) into the following SDP:

$$\begin{aligned}
\max_{T,\Lambda,\gamma}\ & -\mathrm{tr}(T) - \gamma(n - r) + 2\langle \Lambda, \Sigma\rangle + \|\Sigma\|_F^2 \\
\text{s.t. }\ & \Lambda \in \mathbb{S}_0^n,\ T - \gamma I \in \mathbb{S}_+^n, \\
& \begin{bmatrix} T & \Sigma + \Lambda \\ \Sigma + \Lambda & I \end{bmatrix} \in \mathbb{S}_+^{2n}.
\end{aligned} \tag{6}$$

B. Proposed Algorithm

Suppose we have solved the SDP in (6) and obtained an optimal dual variable $\Lambda^\star$. What is the most appropriate value for the primal variable Ω based on the dual optimum? The following theorem shows how we obtain the optimal primal variable $\Omega^\star$ when the duality gap is zero, i.e.,

$$\begin{aligned}
&\min_{\Omega}\{\|\Sigma - \Omega\|_F^2 |\ \mathrm{rank}(\Omega) \le r,\ \Sigma \ge \Omega \ge 0,\ \Sigma - \Omega \in \mathbb{D}^n\} \\
&= \max_{\Lambda}\{-\|\Sigma + \Lambda\|_r^2 + 2\langle \Lambda, \Sigma\rangle + \|\Sigma\|_F^2 |\ \Lambda \in \mathbb{S}_0^n\}.
\end{aligned} \tag{7}$$

Theorem 1. *Let $\Sigma \in \mathbb{S}_{++}^n$ and equality (7) be true. Then a solution to (4) satisfies*

$$\Omega^\star \in \mathrm{svd}_r(\Sigma + \Lambda^\star), \tag{8}$$

where $\Lambda^\star$ solves (6).

Proof. Since equality (7) is true, all the inequalities above (5) are actually equalities. Hence a solution to (4) necessarily solves the following problem:

$$\begin{aligned}
\min_{\Omega}\{&\|\Sigma + \Lambda^\star - \Omega\|_F^2 - \|\Sigma + \Lambda^\star\|_F^2 + 2\langle \Lambda^\star, \Sigma\rangle \\
&+ \|\Sigma\|_F^2 |\ \mathrm{rank}(\Omega) \le r\}.
\end{aligned} \tag{9}$$

By the standard SVD-approximation shown in (1), the solution to (4) satisfies

$$\Omega^\star \in \mathrm{svd}_r(\Sigma + \Lambda^\star),$$

which completes the proof. □

As we see from the theorem, $\Omega^\star$ may be selected as an appropriate candidate to test the feasibility of (4). In this case, we may first obtain an $\Omega^\star \in \mathrm{svd}_r(\Sigma + \Lambda^\star)$, then check whether $\Omega^\star$ is feasible to (4). It is clear that $\mathrm{rank}(\Omega^\star) = r$ due to (2), hence it suffices to check whether $\Sigma \ge \Omega^\star \ge 0$ and $\Sigma - \Omega^\star \in \mathbb{D}^n$.

Based on the above developments, we propose the following algorithm involving only convex optimizations to solve the Frisch-Kalman problem.

Algorithm 1 Proposed Method to Frisch-Kalman Problem

Step 1 Given $\Sigma \in \mathbb{S}_{++}^n$. Set the initial searching rank as $r \in [1, n - 1]$.
Step 2 Compute $\Lambda^\star$ via the SDP in (6).
Step 3 Compute an $\Omega^\star \in \mathrm{svd}_r(\Sigma + \Lambda^\star)$. Check whether $\Sigma \ge \Omega^\star \ge 0$ and $\Sigma - \Omega^\star \in \mathbb{D}^n$. If not, let $r := r + 1$ and go to **Step 2**.
Step 4 An upper bound of the Frisch-Kalman problem (3) is obtained as $r^\star := \mathrm{rank}(\Omega^\star) \ge \mathrm{mr}(\Sigma)$.

As the Frisch-Kalman problem is relaxed in Algorithm 1, some conditions under which the relaxation is tight have been investigated, which are omitted here due to the space limitation. The detailed developments and demonstrations can be found in [25].

C. Application to Shapiro Problem

Consider the following variant of the Frisch-Kalman problem, called the Shapiro problem [6], where the constraint that Δ is nonnegative is relaxed. Investigation into such a relaxed problem brings about more direct understanding on how the off-diagonal entries of Σ affect the minimization of its rank.

Definition 2 (Shapiro Problem)**.** *Given $\Sigma \in \mathbb{S}^n_{++}$, determine*

$$\mathrm{mr}_s(\Sigma) := \min_{\Omega,\Delta}\{\mathrm{rank}(\Omega)|\ \Sigma = \Omega + \Delta, \\ \Omega \in \mathbb{S}^n_+, \Delta \in \mathbb{D}^n\}. \tag{10}$$

Actually, Shapiro and Frisch-Kalman problems share many similar properties. Naturally, we can apply the above algorithm to Shapiro problem with slight modifications, i.e., replacing **Step 3** with

Step 3* Compute $\Omega^\star \in \mathrm{svd}_r(\Sigma + \Lambda^\star)$. Check whether $\Omega^\star \in \mathbb{S}^n_+$ and $\Sigma - \Omega^\star \in \mathbb{D}^n$. If not, let $r := r + 1$ and go to **Step 2**.

The obtained rank $r^\star$ satisfies that $\mathrm{mr}_s(\Sigma) \leq r^\star$.

D. Extension to the Complex-Valued Case

Denote by $\mathbb{H}_+$ ($\mathbb{H}_{++}$, resp.) the set of all positive semi-definite (definite, resp.) matrices in $\mathbb{C}^{n\times n}$. The Frisch-Kalman problem can be extended to the case with complex-valued matrices as follows.

Definition 3 (Complex-Valued Frisch-Kalman Problem)**.** *Given $\Sigma \in \mathbb{H}^n_{++}$, determine*

$$\mathrm{mr}(\Sigma) := \min_{\Omega,\Delta}\{\mathrm{rank}(\Omega)|\ \Sigma = \Omega + \Delta, \\ \Omega \in \mathbb{H}^n_+, \Delta \in \mathbb{D}^n_+\}. \tag{11}$$

In this case, we may directly apply Algorithm 1 to the above problem by suitably replacing all the involved symmetric matrices with the Hermitian ones.

IV. Comparison with Existing Methods

Various heuristic methods have been investigated for solving rank minimization problems. In this section, we compare our proposed method with several mostly adopted existing methods on solving the Frisch-Kalman problem.

A. Nuclear Norm Minimization

In the context of factor analysis, nuclear norm (trace) minimization has been pursued as a suitable heuristic; see, for instance, [6], [9], [13]. The nuclear norm of a matrix is defined as the sum of all its singular values. With this heuristic, the Frisch-Kalman problem is relaxed into

$$\min_{\Delta}\{\mathrm{tr}(\Sigma - \Delta)|\ \Delta \in \mathbb{D}^n_+, \Sigma \geq \Delta \geq 0\}. \tag{12}$$

One way to analyze the corresponding conditions on tight relaxation, i.e., when the solutions to (12) solve the Frisch-Kalman problem, is via investigating the restricted isometry property (RIP) [13] of an associated linear operator. However, it can be shown with some simple calculation that the RIP conditions are not applicable to the Frisch-Kalman problem.

B. Low-Rank Inducing $r*$-norm

A series of matrix norms, called the $r*$-norms (or spectral r-support norms) [16]–[18], are defined by

$$\|M\|_{l_\infty, r*} := \max_{\|X\|_{l_1,r}\leq 1} \langle X, M\rangle, \tag{13}$$

where $X, M \in \mathbb{R}^{m\times n}$, $r = 1, 2, \dots, \min\{m, n\}$, and

$$\|X\|_{l_1,r} := \sum_{k=1}^{r} \sigma_k(X)$$

is the Ky Fan r-norm. When $r = 1$, $\|X\|_{l_1,1}$ reduces to the spectral norm and its dual norm $\|M\|_{l_\infty,1*}$ reduces to the nuclear norm. Therefore, the $r*$-norms include the well-known nuclear norm as a special case. With these low-rank inducing norms, the Frisch-Kalman problem may be relaxed into

$$\min_{\Delta}\{\|\Sigma - \Delta\|_{l_\infty,r*}|\ \Delta \in \mathbb{D}^n, \Sigma \geq \Delta \geq 0\}. \tag{14}$$

Via similar developments in [17], we can transform (14) into the following SDP:

$$\begin{aligned} &\min_{W,\Delta,\gamma}\ \gamma \\ &\text{s.t. } W \in \mathbb{S}^n_+,\ \Delta \in \mathbb{D}^n_+,\ \Sigma - \Delta \in \mathbb{S}^n_+, \\ &\begin{bmatrix} \gamma I - W & \Sigma - \Delta \\ \Sigma - \Delta & I \end{bmatrix} \in \mathbb{S}^{2n}_+, \\ &\mathrm{tr}(W) = \gamma(n - r). \end{aligned} \tag{15}$$

When applying it to the Frisch-Kalman problem, we search for the lowest-rank solution by sequentially solving (15) with $r = 1, 2, \dots, n$.

C. Log-Det Heuristic

The logarithm of the determinant has been used as a smooth approximation for the rank function; see, for instance, [19]. For $X \in \mathbb{S}^n_+$, the function $\log\det(X + \delta I)$, where $\delta > 0$, is used as a smooth surrogate for $\mathrm{rank}(X)$. Since $\log\det(X + \delta I)$ is actually non-convex in X, local minimization methods are proposed in [19] by solving trace minimization problems iteratively. In this case, the Frisch-Kalman problem is approximately solved via the following iterations:

$$\Delta_0 = 0,\ \delta > 0, W_k = (\Sigma - \Delta_k + \delta I)^{-1},$$
$$\Delta_{k+1} = \arg\min_{\Delta}\{\mathrm{tr}(W_k(\Sigma - \Delta))|\ \Delta \in \mathbb{D}^n, \Sigma \geq \Delta \geq 0\}.$$

D. Simulation Result

We compare the proposed algorithm with the existing methods, including the nuclear norm, $r*$-norms and log-det heuristics, based on randomly generated data for both $\hat{\Omega}$ and Δ. The detailed randomization is given by the following steps.

1) Generate matrix $X \in \mathbb{R}^{r\times n}$ with $[X]_{ij}$ being standard i.i.d. Gaussian random variables, i.e., $[X]_{ij} \sim \mathbb{N}(0, 1)$. Compute $\hat{\Omega} = X^T X$.
2) Generate $d \in \mathbb{R}^n$ with d_i, $i = 1, 2, \dots, n$, being i.i.d. and uniformly distributed on $[0, 1]$. Generate $\tilde{\Delta}$ with prescribed norm $\|\tilde{\Delta}\|_F = \|\Delta\|_F$ according to
$$\tilde{\Delta} = \frac{\|\Delta\|_F}{\|d\|}\mathrm{diag}^*(d), \tag{16}$$
such that $\Sigma = \hat{\Omega} + \tilde{\Delta} > 0$.

Given randomly generated matrices $\Sigma \in \mathbb{S}^n_{++}$, we check whether the heuristic methods will solve the Frisch-Kalman

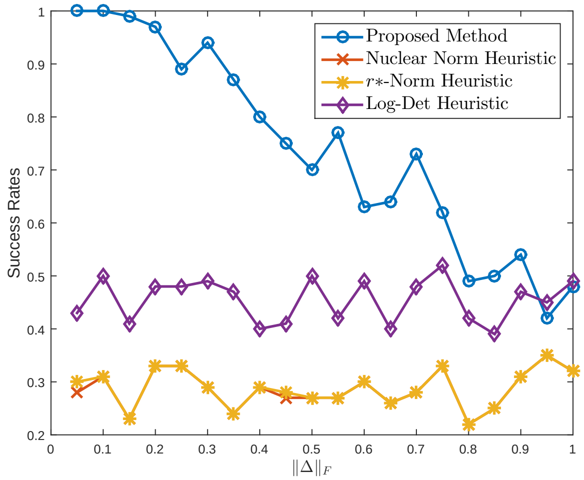

Fig. 1: Success rates for the proposed, the nuclear norm, $r*$-norms and log-det heuristics to solve the Frisch-Kalman problem in (3), respectively, where Σ is randomly generated as described in the context with parameters $n = 10$, $r = 5$, and $T = 100$.

problem. For each level of $\|\Delta\|_F$, we repeat the the above procedures for T times, and count for the success rates. Here, the "success rate" refers to the percentage of experiments in which the recovered rank $r^\star$ satisfies that $r^\star \leq r$.

It can be seen from Fig. 1 that the success rates for the proposed method highly depend on the "size of noise", namely, the value $\|\Delta\|_F$, while those of the other heuristics do not. When $\|\Delta\|_F$ is close to zero, the success rate of Algorithm 1 approaches one. Practically, we may consider a suitable combination of all these heuristics.

V. Conclusion and Future Work

A heuristic convex method is proposed for the century-old Frisch-Kalman problem. Simulation results show that the method is accurate under the condition that the noise components are relatively small compared with the underlying data.

For future research, the proposed method may be improved via, for example, the combination with other heuristics, pre-processing on the observed data to remove possible outliers and so on. Another direction is to apply the method or the underlying ideas to solve more general rank minimization problems, such as the low-rank matrix completion, the data compression and so on.

References

[1] R. Frisch, *Statistical Confluence Analysis by Means of Complete Regression Systems*. Universitetets Økonomiske Instituut, 1934, vol. 5.

[4] C. Spearman, "General intelligence, objectively determined and measured," *American Journal of Psychology*, vol. 15, pp. 201–293, 1904.

[2] R. E. Kalman, "System identification from noisy data," in *Dynamical Systems II*, A. Bednarek and L. Cesari, Eds. NY: Academic Press, 1982, vol. 210, pp. 135–164.

[3] ——, "Identification of noisy systems," *Russian Mathematical Surveys*, vol. 40, no. 4, p. 25, 1985.

[5] W. Ledermann, "On a problem concerning matrices with variable diagonal elements," *Proceedings of the Royal Society of Edinburgh*, vol. 60, no. 1, pp. 1–17, 1940.

[6] A. Shapiro, "Rank-reducibility of a symmetric matrix and sampling theory of minimum trace factor analysis," *Psychometrika*, vol. 47, no. 2, pp. 187–199, 1982.

[7] B. D. Anderson and M. Deistler, "Dynamic errors-in-variables systems with three variables," *Automatica*, vol. 23, no. 5, pp. 611–616, 1987.

[8] T. Söderström, "Errors-in-variables methods in system identification," *Automatica*, vol. 43, no. 6, pp. 939–958, 2007.

[9] L. Ning, T. T. Georgiou, A. Tannenbaum, and S. Boyd, "Linear models based on noisy data and the Frisch scheme," *SIAM Review*, vol. 57, no. 2, pp. 167–197, 2015.

[10] A. L. Comrey and H. B. Lee, *A First Course in Factor Analysis*. NY: Psychology Press, 1992.

[11] O. Reiersøl, "Confluence analysis by means of lag moments and other methods of confluence analysis," *Econometrica*, vol. 9, no. 1, pp. 1–24, 1941.

[12] E. J. Candes and Y. Plan, "Matrix completion with noise," *Proc. of the IEEE*, vol. 98, no. 6, pp. 925–936, June 2010.

[13] B. Recht, M. Fazel, and P. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.

[14] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proc. of the Forty-fifth Annual ACM Symposium on Theory of Computing*, 2013, pp. 665–674.

[15] J. Saunderson, V. Chandrasekaran, P. A. Parrilo, and A. S. Willsky, "Diagonal and low-rank matrix decompositions, correlation matrices, and ellipsoid fitting," *SIAM J. on Matrix Analysis and Appl.*, vol. 33, no. 4, pp. 1395–1416, 2012.

[16] C. Grussler, A. Rantzer, and P. Giselsson, "Low-rank optimization with convex constraints," *IEEE Trans. Automat. Contr.*, vol. 63, no. 11, pp. 4000–4007, 2018.

[17] C. Grussler and P. Giselsson, "Low-rank inducing norms with optimality interpretations," *SIAM J. on Optimization*, vol. 28, no. 4, pp. 3057–3078, 2018.

[18] A. M. McDonald, M. Pontil, and D. Stamos, "New perspectives on k-support and cluster norms," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 5376–5413, 2016.

[19] M. Fazel, H. Hindi, and S. Boyd, "Log-det heuristic for matrix rank minimization with applications to hankel and euclidean distance matrices," in *Proc. of the 2003 American Contr. Conference (ACC)*, vol. 3, 2003, pp. 2156–2162.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. London: Cambridge University Press, 2004.

[21] A. Tatem and C. Linard, "Population mapping of poor countries," *Nature*, vol. 474, pp. 36–36, 2011.

[22] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. London: Cambridge University Press, 2005.

[23] J. Hansen, M. Sato, R. Ruedy, K. Lo, D. W. Lea, and M. Medina-Elizade, "Global temperature change," *Proc. of the National Academy of Sciences*, vol. 103, no. 39, pp. 14 288–14 293, 2006.

[24] G. W. Stewart, *Matrix Perturbation Theory*. San Diego, CA: Academic Press, 1990.

[25] D. Zhao, A. Rantzer, and L. Qiu, "A convex approach to Frisch-Kalman problem," *arXiv:1909.03932 [math.OC]*, 2019.