49th IEEE Conference on Decision and Control
December 15-17, 2010
Hilton Atlanta Hotel, Atlanta, GA, USA

Singular Autoregressions for Generalized Dynamic Factor Models

Manfred Deistler, *Fellow, IEEE,* Alexander Filler, Brian D.O. Anderson, *Life Fellow, IEEE,*
Weitian Chen, *Member, IEEE* and Elisabeth Felsenstein

Abstract— **We consider Generalized Linear Dynamic Factor Models in a stationary context, where the latent variables and thus the static and dynamic factors are the sum of a linearly regular and a linearly singular stationary process and the noise process is linearly regular. The linearly singular component may be useful for modeling e.g. business cycles or seasonal fluctuations in the observed variables. We present a structure theory for this case. The emphasis is laid on the autoregressive case. In general the stationary solutions of the autoregressive models considered here consist of a linearly regular and a linearly singular part. The linearly singular part corresponds to the homogeneous solution of a system having stable roots as well as roots of modulus one. We discuss the solutions of the Yule Walker equations for this case.**

Index Terms— **Generalized Dynamic Factor Models, Linearly Regular and Linearly Singular Stationary Processes, Yule Walker Equations, Identification, High Dimensional Time Series**

I. Introduction

Modeling high dimensional time series is an important issue in many applications. As is well known traditional approaches for modeling high dimensional time series are plagued by the so called "curse of dimensionality". For instance if unrestricted autoregressions with maximum lag p are used and N is the cross-sectional dimension, the dimension of the parameter space is equal to $N^2 p$ whereas the amount of the data is NT, where T is the sample size. There are several approaches to overcome this curse of dimensionality in modeling of high dimensional time series, such as grey box models, panel time series models or factor models. In this contribution we deal with Generalized Linear Dynamic Factor Models (GDFMs) which have been introduced in [7], [8] and, in a slightly different form, in [15], [16]. The idea is to generalize and combine linear dynamic factor models with strictly idiosyncratic[1] noise as analyzed in [13] and [14] and generalized linear static factor models, introduced in [3] and [4].

The basic idea of GDFMs is that the N-dimensional observation at time t, y_t^N say, can be represented as

$$y_t^N = \hat{y}_t^N + u_t^N \tag{1}$$

where $(\hat{y}_t^N)$ is the process of latent variables, which only depend on a few factors and thus are co-moving and a noise part (u_t^N) which now, in contrast to strictly idiosyncratic dynamic factor models, is weakly dependent in the cross-sectional dimension and therefore is called weakly idiosyncratic. The generalization from strict to weak idiosyncratic noise is crucial in defining the difference between GDFMs and "classical" dynamic factor models.

The focus of this contribution is on structure theory. In structure theory, we treat identification in an idealized setting, where we commence from population rather than from sample second moments. Although this is an idealized setting, certain problems occurring in actual estimation from finite data can already be solved at the level of structure theory ([8],[5]). Up to now GDFMs have been used for applications in macro economics and finance, both for forecasting, but recently also for analysis of economic relations. We think, that these models are of great potential use for engineering too, in particular in case of "over-sensoring", e.g. in situations where a sensor network is used to collect many data to cooperatively monitor physical or environmental conditions [11].

The novelty of the analysis given in this paper is that we allow for a linearly singular component (harmonic process) to be present in the latent variables.

In section II we discuss the model class in detail and in setion III we are concerned with finding a model for the latent variables $\hat{y}_t$. This is decomposed in two steps, finding a linear static transformation between the latent variables and the static factors and then finding a linear dynamic model for

Support by the FWF (Austrian Science Fund under contracts P17378 and P20833/N18), the ARC (Australian Research Council under Discovery Project Grant DP1092571) and National ICT Australia Ltd. is gratefully acknowledged. Ltd. is funded through the Australian Governments Backing Australias Ability initiative, and in part through the Australian Research Council.

M. Deistler is with the Department of Mathematical Methods in Economics, Technical University of Vienna, 8/119 Argentinierstrasse, A 1040 Vienna, Austria. e-mail: deistler@tuwien.ac.at

A. Filler is with the Department of Mathematical Methods in Economics, Technical University of Vienna, 8/119 Argentinierstrasse, A 1040 Vienna, Austria. e-mail: alexander.filler@tuwien.ac.at

B.D.O. Anderson is with the Research School of Information Sciences and Engineering, Australian National University, Canberra, ACT 0200, Australia and Canberra Research Laboratory, National ICT Australia Ltd., PO Box 8001, Canberra, ACT 2601, Australia. e-mail: brian.anderson@anu.edu.au

W. Chen is with the Research School of Information Sciences and Engineering, Australian National University, Canberra, ACT 0200, Australia. e-mail: weitian.chen@anu.edu.au

E. Felsenstein is a student at the Department of Mathematical Methods in Economics, Technical University of Vienna, 8/119 Argentinierstrasse, A 1040 Vienna, Austria. e-mail: elisabeth.felsenstein@gmail.co

[1] "Idiosyncratic" means specific for a particular univariate time series, as opposed to "co-moving". We use the term "strictly idiosyncratic" noise if the noise components are mutually uncorrelated in cross-section, i.e. if the noise spectrum is diagonal.

the static factors from the Yule Walker equations. Moreover the second moments of the linearly regular part and of the linearly singular part of the static factors are described.

II. THE MODEL CLASS

We impose the following assumptions:

Assumption 1. $\mathbb{E}\hat{y}_t^N = \mathbb{E}u_t^N = 0 \quad \forall t$

Assumption 2. $\mathbb{E}[\hat{y}_t^N u_s^{N\prime}] = 0 \quad \forall s,t$

Assumption 3. *$(\hat{y}_t^N)$ and (u_t^N) are wide sense stationary, where (u_t^N) is (linearly) regular with absolutely summable covariances and $(\hat{y}_t^N)$ is a sum of a (linearly) regular part $(\hat{y}_t^{rN})$ and a (linearly) singular part $(\hat{y}_t^{sN})$ according to Wold decomposition (see [12],[9]). In addition we assume that $(\hat{y}_t^{rN})$ has absolutely summable covariances and while $(\hat{y}_t^{sN})$ is a harmonic process (see [6]) of the form*

$$\hat{y}_t^{sN} = \sum_{j=1}^{h} e^{i\lambda_j t} C_j^N z_j \tag{2}$$

where $C_j^N \in \mathbb{C}^{N\times 1}$ and the complex valued one dimensional random variables z_j satisfy

- $|z_j| = 1, j = 1,...,h$
- $\mathbb{E}z_j = 0, j = 1,...,h$
- $\mathbb{E}z_j\bar{z}_l = 0, \forall j \neq l$, *where $\bar{}$ denotes complex conjugation*
- $\lambda_{j+1} = -\lambda_{h-j}$, $z_{j+1} = \bar{z}_{h-j}$ *and* $C_{j+1}^N = \bar{C}_{h-j}^N$ *for* $j = 0,1,\cdots,h/2-1$

and the frequencies are ordered in decreasing size, i.e. $\lambda_j > \lambda_{j+1}$.

As in general a harmonic process is non ergodic and as we only consider a single trajectory, the randomness may be confined to the phases of the z_j. Thus the normalization condition $|z_j| = 1$ can be justified.

By these assumptions, the spectral densities of the respective linearly regular parts exist and in an obvious notation we can write

$$f_y^N(\lambda) = f_{\hat{y}}^N(\lambda) + f_u^N(\lambda). \tag{3}$$

Clearly the linearly singular parts of (y_t^N) and $(\hat{y}_t^N)$ are identical with spectral distribution function $F^{sN}(\lambda) = \sum_{j:\lambda_j\leq\lambda} C_j^N \mathbb{E}z_j\bar{z}_j C_j^{N*} = \sum_{j:\lambda_j\leq\lambda} C_j^N C_j^{N*}$, where * means conjugation and transposition, and thus the spectral distribution function F_y^N of (y_t^N) is of the form $F_y^N(\lambda) = \int_{-\pi}^{\lambda} f_y^N(\omega)d\omega + F^{sN}(\lambda)$.

By inclusion of the linearly singular component we allow for exact oscillations in the observed data. Note that we consider the case where the frequencies of the oscillations are not necessarily known a priori. Of course, the task of parameter identification is greatly simplified by knowing such frequencies.

In the applications we have in mind not only the sample size T but also the cross-sectional dimension N is large. For instance in typical economic situations N is 100 or 120. This is a justification for developing an asymptotic theory where both the sample size T and the cross-sectional dimension N are tending to infinity. Accordingly we consider a sequence of Generalized Dynamic Factor Models, indexed by N. While the assumptions 1 - 3 are rather unspecific, the additional assumptions 4 - 10 below are central for the definition of GDFMs.

Assumption 4. *The double-indexed sequence $(y_{it}\,|\, i \in \mathbb{N}, t \in \mathbb{Z})$ corresponds to a nested sequence of models, in the sense that $\hat{y}_{it}$ and u_{it} do not depend on N for $i \leq N$, where e.g. y_{it} denotes the i-th component of y_t^N.*

Assumption 5. *There is an N_0 such that for all $N \geq N_0, f_{\hat{y}}^N$ is a rational spectral density with constant rank $q < N$ on $[-\pi,\pi]$, and with q independent of N.*

Assumption 6. *The number of oscillations h in* (2) *is independent of N ($N \geq$ some N_0).*

Assumption 7 (Strong dependence for the linearly regular part of the latent variables)**.** *The first q (i.e. the q largest) eigenvalues of $f_{\hat{y}}^N$ diverge to infinity for all frequencies, as $N \to \infty$.*

Assumption 8 (Strong dependence for the linearly singular part of the latent variables)**.** *The nonzero eigenvalues of $C_{j+1}^N C_{j+1}^{N*} + C_{h-j}^N C_{h-j}^{N*}$ diverge for $N \to \infty$.*

Assumption 9 (Weak dependence for the noise)**.** *The largest eigenvalue of the spectral density of u_t^N is uniformly bounded for all frequencies and N.*

In addition in our analysis we assume

Assumption 10. *The dimension, n say, of a minimal state space realization of a stable and mini-phase spectral factor of $f_{\hat{y}}^N$ is independent of N ($N \geq$ some N_0).*

Let us discuss some of the assumptions now. As can be shown, weak dependence means that the influence of the noise can be filtered out by letting N going to infinity. On the other hand strong dependence means that the space spanned by the latent variable is not affected by such a filtering operation (see [8]). A practical example would be a technical system endowed with a large number of sensors. Here the latent variables would represent the unobserved true variables and the idiosyncratic noise would be the vector of measurement errors. The linearly singular component describes sinusoidal variations in the latent variables. The intuitive meaning of strong dependence is that there is sufficient comovement on a sufficient number of latent variables, both for the linearly regular and the linearly singular component. A simple mathematical example for this would be as follows

$$y_t^N = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \varepsilon_t + e^{i\lambda_1 t} \begin{bmatrix} 1+i \\ 1-i \\ 1+i \\ 1-i \\ \vdots \end{bmatrix} z_1 + e^{-i\lambda_1 t} \begin{bmatrix} 1-i \\ 1+i \\ 1-i \\ 1+i \\ \vdots \end{bmatrix} \bar{z}_1 + \begin{bmatrix} u_t^1 \\ \vdots \\ u_t^N \end{bmatrix} \tag{4}$$

Here ε_t as well as $u_t^i, i = 1,...,N$ are white noise processes which are mutually uncorrelated and both of these processes are uncorrelated with the complex variables $z_1, \bar{z}_1$. $\mathbb{E}\varepsilon_t^2 = \mathbb{E}|u_t^i|^2 = 1$. Therefore

$$f_{\hat{y}} = 1/(2\pi) \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \\ 1 & & 1 \end{bmatrix}$$

and thus the largest eigenvalue of $f_{\hat{y}}$ is $N/(2\pi)$; thus the linearly regular part of the latent variables is strongly dependent. An analogous statement holds for the linearly singular part. Clearly the noise in this example is weakly dependent. By premultiplying (4) with $1/N \begin{bmatrix} 1 & 1 & 1 & 1 & \dots \\ 1 & -1 & 1 & -1 & \dots \end{bmatrix}$ we obtain the static factor (as defined in Section III) $z_t = \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix} + e^{i\lambda_1 t} \begin{bmatrix} z_1 \\ iz_1 \end{bmatrix} + e^{-i\lambda_1 t} \begin{bmatrix} \bar{z}_1 \\ -i\bar{z}_1 \end{bmatrix}$.

From now on, unless the contrary is stated explicitly, for the sake of simplicity of notation, we will omit the superscript N.

Let us discuss the linearly singular components in more detail. Note that $s_t = \sum_{j=1}^{h} e^{i\lambda_j t} z_j$ is a one dimensional (minimal) dynamic factor[2] for the linearly singular component $\hat{y}_t^s$. Because of assumption 3 we can write s_t as

$$s_t = \sum_{j=1}^{h/2} 2(Re(z_j)\cos\lambda_j t - Im(z_j)\sin\lambda_j t) \tag{5}$$

where Re and Im denote the real and the imaginary part of a complex number respectively. Note that the consideration of vectors z_j (of dimension larger than one) doesn't make sense in our context, since the covariance matrix $\mathbb{E}z_j z_j^*$ could not be estimated consistently. Even if there are two independent vectors affecting different outputs, with the same frequency, for a single sample function this would be indistinguishable from the rank 1 situation argued here. For this reason, z_j is assumed to be one dimensional.

The linearly singular part of the latent variables $\hat{y}_t^s$ is obtained by a linear dynamic transformation of s_t with a transfer function $w^s(\lambda)$ described by $w^s(\lambda) = \begin{cases} C_j & \text{for } \lambda_j, j = 1,..,h \\ 0 & \lambda \neq \lambda_j \end{cases}$. Then clearly $\mathbb{E}C_j|z_j|^2 C_j^*$ is of rank one.

III. Modeling the Static Factors, the Yule Walker Equations

As has been said already, by our assumptions, asymptotically, the noise can be filtered out from the observations (see [8]). For this reason, in our structural analysis, we restrict ourselves to the problem of finding a model for the latent variables. In addition, we commence from the population rather than from the sample second moments of the latent variables.

[2]A (minimal) dynamic factor is a process (of minimal dimension) which gives the latent variables (or its linearly regular and singular components respectively) as a linear dynamic transformation.

A static factor z_t is a random variable such that

$$\hat{y}_t = Hz_t, \quad H \in \mathbb{R}^{N\times r} \tag{6}$$

holds. We are primarily interested in minimal static factors, i.e. in static factors of minimal dimension. As is easy to see, minimal static factors can be obtained from

$$\mathbb{E}\hat{y}_t\hat{y}_t' = H\mathbb{E}z_t z_t' H' \tag{7}$$

where $r = \text{rk}\,\mathbb{E}\hat{y}_t\hat{y}_t'$, and they are unique up to premultiplication by constant nonsingular (real) matrices.

In general, static factors consist of a linearly regular and a linearly singular part. By assumptions 6 and 10, the dimension r of a minimal static factor is independent of N, and clearly $r < N$ holds, from a certain N_0 onwards. For this reason, it is convenient to decompose the modeling procedure for $(\hat{y}_t)$ into two steps:

1) modeling of the linear static relation between latent variables and minimal static factors. As is seen from (7) this is straight forward.
2) modeling the dynamics of the static factors, which is the core problem considered here.

Consider the decomposition of the static factors, corresponding to Wold decomposition

$$z_t = z_t^r + z_t^s \tag{8}$$

where (z_t^r) is linearly regular and (z_t^s) is linearly singular. The analysis of the linearly regular part is described in detail in [5]. The linearly regular part has a Wold representation

$$z_t^r = \sum_{j=0}^{\infty} k_j \varepsilon_{t-j}, k_j \in \mathbb{R}^{r\times q}, \mathbb{E}\varepsilon_t\varepsilon_t' = I \tag{9}$$

The coefficients k_j in (9) are unique up to postmultiplication by an orthogonal matrix which independent on j. By (6) and assumption 5 the spectral density f_z of (z_t^r) is rational and has the same rank q as $f_{\hat{y}}$. Note that $r \geq q$ holds. Clearly, ε_t is a dynamic factor for the linearly regular part $\hat{y}_t^r$ of $\hat{y}_t$. Also note, that the dimension, r^r say, of a static factor for the linearly regular part $\hat{y}_t^r$ may be smaller than r, because the rank of $\mathbb{E}\hat{y}_t^r\hat{y}_t^{r'}$ may be smaller than the rank of $\mathbb{E}\hat{y}_t\hat{y}_t'$. As has been shown in [1], for $r^r > q$, generically the linearly regular part $\hat{y}_t^r$ can be represented by an autoregression.

Given the dynamic factor in (5) a static factor for the linearly singular part can be obtained as

$$z_t^s = 2\begin{bmatrix} (Rez_1 - Imz_1)\cos\lambda_1 t \\ (-Imz_1 - Rez_1)\sin\lambda_1 t \\ \vdots \\ (-Imz_{h/2} - Rez_{h/2})\sin\lambda_{h/2} t \end{bmatrix} \tag{10}$$

Note that this static factor is not necessarily a minimal one. The relation between the static and the dynamic linearly singular factor is the following

$$z_t^s = \begin{bmatrix} 1 & 0 & \dots & & & \dots & 0 & 1 \\ i & 0 & \dots & & & \dots & 0 & -i \\ 0 & 1 & 0 & \dots & \dots & 0 & 1 & 0 \\ 0 & i & 0 & & & 0 & -i & 0 \\ & & \ddots & & & \iddots & & \\ & & & 1 & 1 & & & \\ & & & i & -i & & & \end{bmatrix} \begin{bmatrix} k_1^s(\lambda) \\ \vdots \\ k_h^s(\lambda) \end{bmatrix} s_t \qquad (11)$$

where $k_j^s(\lambda) = \begin{cases} 1 & \lambda = \lambda_j \\ 0 & \text{else} \end{cases}$ is the transfer function for z_j, i.e. $z_j = k_j^s(\lambda) s_t$. Since z_t^s is a harmonic process, it can be obtained as the homogeneous solution of an autoregression with innovation equal to zero, where all zeros of the autoregression polynomial are of modulus one. Now it is easy to show, that the best linear least squares forecast of z_t given $z_{t-1}, z_{t-2}, \dots$ only depends on the finite past $z_{t-1}, \dots, z_{t-p}$ and the forecasting error is of the form $k_0 \varepsilon_t$; In other words, z_t is the stationary solution of an autoregressive system of the form

$$z_t = a_1 z_{t-1} + \dots + a_p z_{t-p} + b\varepsilon_t \qquad (12)$$

where $a_j \in \mathbb{R}^{r\times r}$ and $b = k_0 \in \mathbb{R}^{r\times q}$, with

$$\det(I - a_1 z - \dots - a_p z^p) \neq 0, |z| < 1 \qquad (13)$$

where z is both the backward shift and a complex variable, and, for $r > q$, the variance matrix of the one step ahead forecast error is singular. For the case of (linearly) regular processes, such systems have been discussed in detail in [5].

The Yule Walker equations of the process z_t are of the form

$$(a_1, \dots, a_p)\Gamma_p = (\gamma_1, \dots, \gamma_p) \qquad (14)$$

$$\Sigma = \gamma_0 - (a_1, \dots, a_p)(\gamma_1, \dots, \gamma_p)' \qquad (15)$$

where $\gamma_j = \mathbb{E} z_t z_{t-j}'$, $\Gamma_p = \begin{pmatrix} \gamma_0 & \cdots & \cdots & \gamma_{p-1} \\ \vdots & \gamma_0 & & \vdots \\ \vdots & & \ddots & \vdots \\ \gamma_{p-1}' & \cdots & \cdots & \gamma_0 \end{pmatrix}$ and $\Sigma = \mathbb{E} b\varepsilon_t (b\varepsilon_t)' = bb'$.

Let $a(z) = I - a_1 z - \dots - a_p z^p$ where $(a_1, \dots, a_p)$ is a solution of the Yule-Walker equations (14). Since (z_t) is stationary its linearly regular part (z_t^r) is stationary too with innovations (ε_t) corresponding to Wold decomposition (9). The fact that $(b\varepsilon_t)$ is linearly regular implies that $a(z) z_t^s = 0$ holds and thus $a(z) z_t^r = b\varepsilon_t$ must hold. Let $k(z) = \sum_{j=0}^{\infty} k_j z^j$ and $z_t^r = k(z)\varepsilon_t$, then $a(z)k(z)\varepsilon_t = b\varepsilon_t$. Since (ε_t) is persistently exciting, the transfer function $a^{-1}(z)b$ has poles only outside the unit circle. Now we define $\tilde{a}(z)$ by

$$a(z) = c(z)\tilde{a}(z), b = c(z)b \qquad (16)$$

where $c(z)$ is a greatest common left divisor of $(a(z), b)$ with $c(0) = I$. The existence of such a left coprime pair $(\tilde{a}(z), b)$ has been shown in [2]. Since $(\tilde{a}(z), b)$ is left coprime, $\tilde{a}(z)$ defined by (16) is stable, i.e. $\det(\tilde{a}(z)) \neq 0, |z| \geq 1$, as otherwise $\tilde{a}^{-1}(z)b = a^{-1}(z)b$ would have a pole inside or on the unit circle. Note that $\tilde{a}(z)$ isn't a solution of the Yule Walker equations if the linearly singular part z_t^s is nonzero. Further note, that the transfer function $a^{-1}(z)b$, which corresponds to the Wold decomposition (9), is unique up to postmultiplication by constant orthogonal matrices. Thus we can summarize the results in the following theorem.

Theorem 1. *Every solution of the Yule Walker equations* (14) *together with every matrix b of full column rank q satisfying* $bb' = \Sigma$, *with* Σ *defined in* (15), *uniquely defines the transfer function* $a^{-1}(z)b$ *corresponding to the linearly regular part* z_t^r *and the spectral density* f_z *of* z_t^r *given by*

$$f_z(\lambda) = 1/(2\pi) a^{-1}(e^{-i\lambda}) bb' a^{-1*}(e^{-i\lambda}) \qquad (17)$$

An analogous result has been obtained from a slightly different point of view in [10].
From (17)

$$\Gamma_p^r = \mathbb{E} \begin{bmatrix} z_t^r \\ \vdots \\ z_{t-p+1}^r \end{bmatrix} \left[z_t^{r\prime}, \dots, z_{t-p+1}^{r\prime} \right]$$

can be obtained and then

$$\Gamma_p^s = \Gamma_p - \Gamma_p^r \qquad (18)$$

is the corresponding block Toeplitz matrix of the linearly singular process (z_t^s). Analogously to (2), the linearly singular process z_t^s is given by

$$z_t^s = \sum_{j=1}^{h} e^{i\lambda_j t} D_j z_j, D_j \in \mathbb{C}^{r\times 1} \qquad (19)$$

and its spectral distribution function F_z^s is given by

$$F_z^s(\lambda) = \sum_{j:\lambda_j \leq \lambda} D_j D_j^* \qquad (20)$$

As Γ_p^s uniquely determines all covariances of (z_t^s) and as these covariances are in a one-to-one relation with the spectral distribution function (20), $D_j D_j^*$ is uniquely determined from Γ_p^s.

IV. Conclusions

In this contribution we have presented a Generalized Linear Dynamic Factor Model where the latent variables are the sum of a linearly regular and a linearly singular component whereas the noise is a linearly regular process only. By inclusion of the linearly singular component we allow for exact oscillations in the observed data. We analyze the Yule Walker equations for the static factors for this case and describe how to obtain the second moments of the linearly regular and the linearly singular part of the static factors.

References

[1] B. D. O. Anderson and M. Deistler. Properties of zero-free transfer function matrices. *SICE*, 1:284–292, 2008.

[2] B. D. O. Anderson, M. Deistler, W. Chen, and A. Filler. Ar models of singular spectral matrices. In *Conference on Decision and Control*, 2009.

[3] G. Chamberlain. Funds, factors and diversification in arbitrage pricing models. *Econometrica*, 51(5):1305–1324, 1983.
[4] G. Chamberlain and M. Rothschild. Arbitrage, factor structure and meanvariance analysis on large asset markets. *Econometrica*, 51(5):1281–1304, 1983.
[5] M. Deistler, B. D. O. Anderson, A. Filler, Ch. Zinner, and W. Chen. Generalized dynamic factor models-an approach via singular autoregressions. *European Journal of Control*, 16(3):211–224, 2010.
[6] M. Deistler and W. Scherrer. The prague lectures. *Manuscript, TU Vienna*, 1994.
[7] M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic factor model: identification and estimation. *The Review of Economics and Statistics*, 65:453–473, 2000.
[8] M. Forni and M. Lippi. The generalized dynamic factor model: representation theory. *Econometric Theory*, 17:1113–1141, 2001.
[9] E. J. Hannan. *Multiple Time Series*. Wiley, 1970.
[10] Y. Inouye. Modeling of multichannel time series and extrapolation of matrix-valued autocorrelation sequences. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(1):45–55, 1983.
[11] K. Roemer and F. Mattern. The design space of wireless sensor networks. *IEEE Wireless Communications*, 11(6):54–61, 2004.
[12] Yu. A. Rozanov. *Stationary Random Processes*. Holden-Day, San Francisco, 1967.
[13] T. J. Sargent and C. A. Sims. *Business cycle modelling without pretending to have too much a priori economic theory*. Minneapolis: Federal Reserve Bank of Minneapolis, 1977.
[14] W. Scherrer and M. Deistler. A structure theory for linear dynamic errors-in-variables models. *SIAM, J. on Control Optim.*, 36(6):2148–2175, 1998.
[15] J. H. Stock and M. W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the Americal Statistical Association*, 97:1167–79, 2002.
[16] J. H. Stock and M. W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20:147–62, 2002.