Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Analysis and identification of complex stochastic systems admitting a flocking structure

Giulio Bottegal * **Giorgio Picci** **

* *ACCESS Linnaeus Centre, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden (e-mail: bottegal@kth.se)*
** *Department of Information Engineering, University of Padova, Padova, Italy (e-mail: picci@dei.unipd.it)*

Abstract: We discuss a new modeling paradigm for large dimensional aggregates of stochastic systems by Generalized Factor Analysis (GFA) models. These models describe the data as the sum of a *flocking* plus an uncorrelated *idiosyncratic* component. The flocking component describes a sort of collective orderly motion which admits a much simpler mathematical description than the whole ensemble while the idiosyncratic component describes weakly correlated noise. The extraction of the dynamic flocking component is discussed for time-stationary systems.

1. INTRODUCTION

In this paper we elaborate on a new paradigm on stochastic modeling of complex systems proposed in [Bottegal and Picci, 2013a] based on the theory of *Generalized Factor Analysis (GFA)* [Chamberlain and Rothschild, 1983, Forni and Lippi, 2001, Deistler and Zinner, 2007, Anderson and Deistler, 2008, Deistler et al., 2010b,a]. The underlying idea is to split the overall motion of a large ensemble of interacting random units into a *stochastic flocking* plus a noise of a special character which is called the *idiosyncratic* component. The first component describes the average random motion of the system by a rather simple statistical model while the second aims at describing the stochastic dynamics which pertains exclusively to individual fluctuations about the average.

The word *Flocking* is used to describe a commonly observed behavior in gregarious animals by which many equal individuals tend to group and follow, at least approximately, a common path in space. The phenomenon has been studied very actively in recent years; see e.g. [Reynolds, 1987, Vicsek et al., 1995, Veerman et al., 2005, Brockett, 2010] and the literature on this subject is now huge, consisting of hundreds of papers which would be impossible to discuss here. The mechanism of *formation* of flocks has also been intensely studied in the literature. There is now a quite articulated theory addressing the convergence to a flocking structure under a variety of assumptions on the communication strategy among agents, specific nonlinearities of the dynamics, the kind of permissible local control actions etc. see e.g. [Jadbabaie et al., 2003, Fagnani and Zampieri, 2008, Olfati-Saber et al., 2007, Tahbaz-Salehi and Jadbabaie, 2010, Cucker and Smale, 2007, Olfati-Saber, 2006, Shen, 2007, Tanner et al., 2007] and references therein.

Here we want to address a different and perhaps more basic problem: given observations of the motion of a large set of interacting agents and assuming statistical steady state, find out whether there is a flocking component in the collective motion and estimate its characteristics. The rationale for this search is that the very concept of flocking implies an *orderly motion* which must then admit a much simpler mathematical description than that of the whole ensemble. Once a flocking component (if present) has been discovered, the motion of the ensemble can naturally be split into flocking plus a random term (the idiosyncratic component) which describes local random disagreements of the individual agents or the effect of external disturbances. Hence extracting a flocking structure is essentially a parsimonious modeling problem. Prediction of the future behavior and control of a complex ensemble of random agents could then reasonably be restricted to the flocking component and be based on the simple model thereof.

Static GFA models describe a zero-mean stochastic sequence $\mathbf{y} := \{\mathbf{y}(k),\, k \in \mathbb{Z}_+\}$ (represented as a random column vector with an infinite number of components) by a linear model of the form

$$\mathbf{y} = \sum_{i=1}^{q} f_i \mathbf{x}_i + \tilde{\mathbf{y}} \tag{1}$$

where, in analogy to finite-dimensional Factor Analysis models, the random variables $\mathbf{x}_i$, $i = 1, \dots, q$ are called the *common factors* and the deterministic vectors $f_i \in \mathbb{R}^\infty$ are the *factor loadings*. The $\mathbf{x}_i$ can be taken, without loss of generality, to be orthonormal so as to form a q-dimensional random vector $\mathbf{x}$ with $\mathbb{E}\, \mathbf{x}\mathbf{x}^\top = I_q$. The random vector $\tilde{\mathbf{y}}$, uncorrelated with (orthogonal to) $\mathbf{x}$ is the idiosyncratic component. We shall denote the linear combination $\hat{\mathbf{y}} := \sum f_i \mathbf{x}_i$ by $\hat{\mathbf{y}}$ so that (1) can be written $\mathbf{y} = \hat{\mathbf{y}} + \tilde{\mathbf{y}}$ for short.

The idiosyncratic term is no longer required to have uncorrelated components as in the classical Factor Analysis

⋆ The first author was partially supported by the Swedish Research Council under contract 621-2009-4017.

models, but to satisfy instead a *zero-average* condition. This condition implies that the covariance of any two variables $\tilde{\mathbf{y}}(k)$ and $\tilde{\mathbf{y}}(j)$, say $\tilde{\sigma}(k,j)$ tends to zero when $|k-j| \to \infty$.

It has been shown that with this new definition the inherent non-uniqueness of classical finite-dimensional Factor Analysis models does not occur. Moreover in this generalized context the dimension q of the latent factors vector can be characterized as the number of "infinite eigenvalues" of the covariance matrix of $\mathbf{y}$.

The overall covariance of the observed process $\mathbf{y}$ can then be decomposed in the sum of two contributions.

- A *long range* correlation structure which describes the component of $\mathbf{y}$ driven by the latent vector. The *long range* property means that the covariance of two variables $\hat{\mathbf{y}}(k)$ and $\hat{\mathbf{y}}(j)$, say $\hat{\sigma}(k,j)$ does not go to zero when $|k-j| \to \infty$.
- A *short range* correlation structure which corresponds to the idiosyncratic component $\tilde{\mathbf{y}}$. The *short range* property means that the covarinace of two variables $\mathbf{y}(k)$ and $\mathbf{y}(j)$, say $\tilde{\sigma}(k,j) \to 0$ when $|k-j| \to \infty$.

We shall discuss this decomposition for dynamic systems restricting to the case of processes which are stationary with respect to the time variable which is a natural assumption to make in view of statistical inference.

2. DYNAMIC GFA MODELS

Consider an infinite aggregate of random "agents" indexed by a discrete variable $k \in \mathbb{Z}_+$ each described by a scalar output variable [1] $\mathbf{y}(k,t)$, which evolves randomly in (discrete) time. The overall evolution of the ensemble is then described by an infinite dimensional random process $\mathbf{y} := \{\mathbf{y}(t)\,;\, t \in \mathbb{Z}\}$ with components $\mathbf{y}(k,t)$, an infinite column vector of zero mean random variables of finite variance. We shall assume that the infinite covariance matrix,

$$\Sigma(\tau) := \mathbb{E}\,\mathbf{y}(t+\tau)\mathbf{y}(t)^\top$$

is well-defined, independent of t and of positive type. We shall call $\mathbf{y}$ a *time-stationary random field.* Let $F \in \mathbb{R}^{\infty \times q}$; we shall say that the q columns of F are *strongly linearly independent* if the $n \times n, (n \geq q)$ upper left corner of $FF^\top$ has q nonzero eigenvalues which tend to infinity as $n \to \infty$. This concept is introduced in [Bottegal and Picci, 2013a] and cannot be discussed further here for reasons of space.

Definition 1. A time-stationary random field has a dynamic GFA representation *of rank q if it can be written as the sum of two uncorrelated components,*

$$\mathbf{y}(t) = F\mathbf{x}(t) + \tilde{\mathbf{y}}(t) \qquad (2)$$

where the q columns of F are strongly linearly independent, the q dimensional process $\mathbf{x}(t)$, with $\mathbb{E}\,\mathbf{x}(t)\mathbf{x}(t)^\top = I_q$, is jointly (weakly) stationary with $\tilde{\mathbf{y}}(t)$ and the covariance matrix $\tilde{\Sigma}(\tau) := \mathbb{E}\,\tilde{\mathbf{y}}(t+\tau)\tilde{\mathbf{y}}(t)^\top$ is, for all τ, a bounded linear operator in ℓ^2.

Let $\ell^2(\Sigma)$ denote the Hilbert space of infinite sequences $a := \{a(k),\, k \in \mathbb{Z}_+\}$ such that $\|a\|_\Sigma^2 := a^\top \Sigma a < \infty$. When $\Sigma = I$ we use the standard symbol ℓ^2, denoting the corresponding norm by $\|\cdot\|_2$. The following definition was introduced in [Forni and Lippi, 2001]:

A sequence of elements $\{a_n\}_{n \in \mathbb{Z}_+} \subset \ell^2 \cap \ell^2(\Sigma)$ is an averaging sequence *(AS) for* $\mathbf{y}$, *if* $\lim_{n\to\infty} \|a_n\|_2 = 0$.
We say that a sequence of random variables $\mathbf{y}$ *is* idiosyncratic *if* $\lim_{n\to\infty} a_n^\top \mathbf{y} = 0$ *for any averaging sequence* $a_n \in \ell^2 \cap \ell^2(\Sigma)$.

Whenever an infinite covariance matrix Σ defines a bounded operator on ℓ^2, one has $\ell^2(\Sigma) \subset \ell^2$; in this case an AS can be seen just as a sequence of linear functionals in ℓ^2 converging strongly to zero. For example the sequence of elements in ℓ^2

$$a_n = \frac{1}{n}[\underbrace{1 \ldots 1}_{n}\, 0 \ldots]^\top \qquad (3)$$

is an averaging sequence for any Σ. On the other hand, let P_n denote the compression of the n-th power of the left shift operator to the space ℓ^2; i.e. $[P_n a](k) = a(k-n)$ for $k \geq n$ and zero otherwise. Then $\lim_{n\to\infty} P_n a = 0$ for all $a \in \ell^2$ [Halmos, 1961] so that $\{P_n a\}_{n\in\mathbb{Z}_+}$ is an AS for any $a \in \ell^2$.

Example 2. Let $\mathbb{1}$ be an infinite column vector of 1's and let $\mathbf{x}(t)$ be a zero-mean scalar process uncorrelated with $\tilde{\mathbf{y}}(t)$, a zero-mean process such that for each fixed t the random sequence $\{\tilde{\mathbf{y}}(k,t)\,;\, k = 0,1,\ldots\}$ is ergodic [2]. Consider the process

$$\mathbf{y}(t) = \mathbb{1}\mathbf{x}(t) + \tilde{\mathbf{y}}(t)$$

and the averaging sequence (3). Since

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \tilde{\mathbf{y}}(k,t) = \mathbb{E}\,\tilde{\mathbf{y}}(0,t) = 0$$

(limit in L^2) we have

$$\lim_{n\to\infty} a_n^\top \mathbf{y}(t) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{y}(k,t) = \mathbf{x}(t)\,;$$

hence we can recover the latent factor by averaging. More generally, if $\tilde{\mathbf{y}}$ is idiosyncratic $\lim_{n\to\infty} a_n^\top \tilde{\mathbf{y}}(t) = 0$ for any averaging sequence and for all t so one could recover $\mathbf{x}$ from AS's such that $\lim_{n\to\infty} a_n^\top \mathbb{1}$ exists and is non zero. □

The following proposition shows that for a stationary random field $\mathbf{y} := \{\mathbf{y}(t)\,;\, t \in \mathbb{Z}\}$, constructing dynamic GFA representations is in a sense equivalent to constructing static GFA representations for the vector $\mathbf{y}(0)$, or which is the same by stationarity, a static GFA representation for any of the the vectors $\mathbf{y}(t)\,;\, t \in \mathbb{Z}$. Hence, at least in principle, the dynamic problem can be reduced to the static one.

Proposition 3. The stationary random field $\mathbf{y} := \{\mathbf{y}(t)\,;\, t \in \mathbb{Z}\}$ *has a dynamic GFA representation* (2) *if and only if* $\mathbf{y}(0)$ *has a static GFA represenattion with the same factor loading matrix* F, $\mathbf{x} \equiv \mathbf{x}(0)$ *and* $\tilde{\mathbf{y}} \equiv \tilde{\mathbf{y}}(0)$.

A proof can be found in [Bottegal and Picci, 2013b].

The following criterion, originally stated for the static case by [Chamberlain and Rothschild, 1983], can in principle be

[1] This assumption is done for ease of notation; finite dimensional output variables can be treated in the same way.

[2] And hence has a short range correlation structure, in the sense described above.

used to check the existence of a flocking component in a time-stationary random field:

Theorem 4. For a time-stationary random field, a flocking structure exists with q factors if and only if q eigenvalues of the steady state covariance matrix Σ_n *of the n-dimensional random subvector* $\mathbf{y}^n(t)$ *of* $\mathbf{y}(t)$*, tend to infinity with n while the others remain bounded.*

Below is a sharpened version of Theorem 4 which will be used later.

Corollary 5. A stationary random field $\mathbf{y} := \{\mathbf{y}(t)\,;\, t \in \mathbb{Z}\}$ *has a flocking structure, if and only if its steady state covariance matrix* Σ *has a decomposition*

$$\Sigma = \hat{\Sigma} + \tilde{\Sigma}\,; \qquad \hat{\Sigma} = FF^\top \tag{4}$$

where $F \in \mathbb{R}^{\infty \times q}$ *has strongly linearly independent columns and* $\tilde{\Sigma}$ *is a bounded operator in* ℓ^2*. In other words,* Σ *must admit a decomposition as the sum of a bounded plus an unbounded finite rank perturbation of rank q. In particular* $\|\Sigma\|_2 = +\infty$.

3. STATISTICAL ESTIMATION

In this section, we focus on the problem of detecting and estimating a flocking component in a stationary random field.

Assume that we have sample estimates of the covariance of subvectors $\mathbf{y}^n(t) = [\,\mathbf{y}(0,t)\,\mathbf{y}(1,t)\,\ldots\,\mathbf{y}(n,t)\,]^\top$ of the process $\mathbf{y}$, computed using a large enough time window of observations $\{y(t)\,;\, t = 1, \ldots, N\}$. Since the process is stationary, the limit

$$\hat{\Sigma}_n(N) := \frac{1}{N}\sum_{t=1}^{N} y^n(t)[y^n(t)]^\top$$

converges to the true covariance $\Sigma_n = \mathbb{E}\,\mathbf{y}^n(t)[\mathbf{y}^n(t)]^\top$. Following [Chamberlain and Rothschild, 1983, Forni and Lippi, 2001] the idea is to do PCA on the covariance estimates for increasing n. If the data admit a GFA structure, there will be q eigenvalues of Σ_n which tend to grow without bound as $n \to \infty$ while the others stay bounded. The q corresponding eigenvectors will tend as $n \to \infty$ to the q factor loadings $f_1, \ldots, f_q$ and therefore provide asymptotically the GFA decomposition of the $\Sigma(0)$ matrix

$$\Sigma(0) = FF^\top + \tilde{\Sigma}(0) \tag{5}$$

where $\tilde{\Sigma}(0)$ is the part of $\Sigma(0)$ corresponding to the bounded eigenvalues which can in principle be isolated by the PCA procedure. After F and $\tilde{\Sigma}(0)$ are estimated, the stochastic realization procedure described in [Bottegal and Picci, 2013b] permits to construct the factor vector $\mathbf{x}$ and the idiosyncratic component $\tilde{\mathbf{y}}$ of the GFA representation of $\mathbf{y}$. The identification of the time varying factor variables $\mathbf{x}_i(t)$ of $\mathbf{y}$ from the observations $\mathbf{y}(k,t)$ can be done by averaging on the space variable. Since there are only q components to be estimated one should select q independent averaging sequences to construct samples of the $\mathbf{x}_i(t)$ at different time instants. From these samples one can then apply standard time-series identification techniques.

4. LINEAR DYNAMIC SYSTEMS AND GFA

We would like to gain some understanding of the structure of linear dynamical systems which admit a flocking component. A simple class of systems which is in principle amenable to analysis is that of random fields described by linear evolution equations of the general form

$$\mathbf{y}(t+1) = \mathcal{A}\mathbf{y}(t) + \mathbf{w}(t) \tag{6}$$

where $\mathbf{w}$ is a string of uncorrelated stationary white noise processes and $\mathcal{A}$ is a linear operator acting on infinite sequences. We assume that the evolution is asymptotically stable and is stationary in time so that the variance matrix of $\mathbf{y}(t)$ is a constant positive definite matrix, which should then satisfy an infinite dimensional Lyapunov equation

$$\Sigma = A\Sigma A^\top + Q \tag{7}$$

where A is a matrix representation of the operator $\mathcal{A}$ and Q is the variance matrix of the white noise which we assume an infinite diagonal matrix with uniformly bounded positive entries. In this case, a GFA model of $\mathbf{y}$ (if any exists) will also be stationary and the structure of the model can be inferred by analyzing the covariance matrix Σ. When the matrix of the operator A has a nested lower triangular structure, that is when the evolution of the first n agents is not influenced by that of the agents of index $k > n$, the solution of the Lyapunov equations (7) can sometimes be obtained explicitly. The n-dimensional random process $\mathbf{y}^n(t)$, obeys an equation of the form

$$\mathbf{y}^n(t+1) = A_n\mathbf{y}^n(t) + \mathbf{w}^n(t)\,, \quad n = 1, 2, \ldots \tag{8}$$

where the A_n's, the upper left $n \times n$ submatrices of A, are lower triangular with a nested structure of the type

$$A_{n+1} = \begin{bmatrix} A_n & 0 \\ b_n^\top & a_{n+1} \end{bmatrix}, \tag{9}$$

where $|a_{n+1}| < 1$ so that the asymptotic stability of A_n is preserved. The input process $\mathbf{w}^n(t)$ is an n-dimensional white noise with variance $\mathbb{E}\mathbf{w}^n(t)\mathbf{w}^n(s)^\top = Q_n\delta_{t,s}$. We are interested in the asymptotic covariance matrix of $\mathbf{y}^n(t)$. Questions regarding the existence of flocking components can be answered by analyzing the structure of the solution to the family of Lyapunov equations

$$\Sigma_n = A_n\Sigma_n A_n^\top + Q_n \qquad n = 1, 2, \ldots \tag{10}$$

when $n \to \infty$. Some types of families of matrices $\{A_n\}_{n\in\mathbb{N}}$ are considered below.

Autonomous agents In this scenario, the behavior of each agent is independent of the others, being just an autoregressive motion of the type

$$\mathbf{y}_k(t+1) = a_k\mathbf{y}_k(t) + \mathbf{w}_k(t) \quad , \quad \sup_{k\in\mathbb{N}} |a_k| < 1\,. \tag{11}$$

In this case, $A_n = \mathrm{diag}\{a_1, \ldots, a_n\}$. Assuming also Q diagonal (with uniformly bounded elements), the family of Lyapunov equations (10) admits diagonal (nested) solutions with uniformly bounded elements. Hence, in this case the resulting sequence is idiosyncratic, with uncorrelated components. Hence, there is no flocking structure.

Flocking by following a leader This is the case for some hierarchical leadership models as discussed in [Shen, 2007]. A very simple instance is the following model where each agent evolving with the same scalar random dynamics wants to follow a "leader" signal $\mathbf{y}_0(t)$ by applying the same proportional control law based on the measurement of its position with respect to $\mathbf{y}_0(t)$:

$$\mathbf{y}_0(t+1) = a\mathbf{y}_0(t) + \mathbf{w}_0(t)\,, \qquad |a| < 1$$
$$\mathbf{y}_k(t+1) = \mathbf{y}_0(t) + a[\mathbf{y}_k(t) - \mathbf{y}_0(t)] + \mathbf{w}_k(t)\,, \quad k = 1, 2, \ldots$$

The question is if following a leader should, under appropriate circumstances, produce a random flock. Rewriting the model in matrix form

$$\begin{bmatrix} \mathbf{y}_0(t+1) \\ \mathbf{y}_1(t+1) \\ \dots \\ \mathbf{y}_n(t+1) \end{bmatrix} = \begin{bmatrix} a & 0 & \dots & 0 \\ 1-a & a & & \vdots \\ \vdots & 0 & \ddots & 0 \\ 1-a & \dots & & a \end{bmatrix} \begin{bmatrix} \mathbf{y}_0(t) \\ \mathbf{y}_1(t) \\ \dots \\ \mathbf{y}_n(t) \end{bmatrix} + \begin{bmatrix} \mathbf{w}_0(t) \\ \mathbf{w}_1(t) \\ \dots \\ \mathbf{w}_n(t) \end{bmatrix}$$

and computing the covariance matrices of $\mathbf{y}^n(t)$ by solving the Lyapunov equation (10), provides the following answer.

Theorem 6. Assume for simplicity that $Q_n = I_n$. The solution of the Lyapunov equation (10) *tends for $n \to \infty$ to*

$$\Sigma = ff^\top + \tilde{\Sigma}$$

where $f \in \mathbb{R}^\infty$ has components

$$f_k = \begin{cases} a/(1-a^4)^{\frac{1}{2}} & , \ k=1 \\ (1+a^2)^{\frac{1}{2}}/[(1+a)(1-a^2)^{\frac{1}{2}}] & , \ k>1, \end{cases}$$

and $\tilde{\Sigma}$ is a bounded operator in ℓ^2. Hence

$$\mathbf{y}(t) = f\mathbf{x}(t) + \tilde{\mathbf{y}}(t), \quad \mathbf{x}(t) = (1-a^4)^{\frac{1}{2}}\mathbf{y}_1(t-1), \quad \operatorname{Var}\tilde{\mathbf{y}}(t) = \tilde{\Sigma}.$$

Following the previous agent Let the leader be described by the same first order dynamics as in the previous example. Assume instead that each agent has no measurements of $\mathbf{y}_0$ and tries just to follow the previous agent by using the same kind of control law namely

$$\mathbf{y}_k(t+1) = \mathbf{y}_{k-1}(t) + a(\mathbf{y}_k(t) - \mathbf{y}_{k-1}(t)) + \mathbf{w}_k(t),$$

where $k = 1, 2, \dots$ and $|a| < 1$. Does this field have a flocking component? The control gain a may depend on k and, say, increase exponentially with the distance as

$$a(k) = 1 - \lambda^k, \qquad k > 1$$

where $0 < \lambda < 1$ so that the spectrum of the system (6) has an accumulation at $z = 1$. In this case the solution of the Lyapunov equation (7) is unbounded see [Przyluski, 1980].

Infinite dimensional distributed average consensus We may model this adjustment in discrete time by a symmetric linear relation

$$\mathbf{y}_k(t+1) = a_k\mathbf{y}_k(t) + \sum_{j\in N_k} a_{k,j}(\mathbf{y}_j(t) - \mathbf{y}_k(t)) + \mathbf{w}_k(t), \quad (12)$$

where $k = 1, 2, \dots$ and the sum is over the set of neighbors N_k of each state k, which we assume to be a finite set. The overall motion can be described as

$$\mathbf{y}(t+1) = A\mathbf{y}(t) + \mathbf{w}(t) \quad (13)$$

starting at some initial state $\mathbf{y}(0)$. Here A is a matrix with positive elements such that

$$A = A^\top \qquad A\mathbb{1} = \mathbb{1}$$

an infinite doubly stochastic matrix. The state of (13) is not stationary since has a random walk component. We want to see if for some averaging sequence $\{a_n\}$ the limit

$$\lim_{n\to\infty} a_n^\top \mathbf{x}(t)$$

is non-zero. This would imply the existence of a flocking component. Problems of this kind have been studied in the finite-dimensional setting in [Xiao et al., 2007]. Here we study a slightly different model, obtained by modifying (12) so as to deal with an infinite number of agents:

(1) for each $n \geq n_0$, where n_0 is a fixed initial integer, consider a symmetric doubly stochastic matrix A_n, which achieves consensus on the first n agents;
(2) define $\bar{A}_n := (1 - \frac{1}{n})A_n$, a sequence of matrices such that consensus is reached as $n \to \infty$.

Denoting by $\bar{A}$ the limit of the sequence $\{\bar{A}_n, n \in \mathbb{N}\}$, the following result holds.

Theorem 7. The model

$$\mathbf{y}(t+1) = \bar{A}\mathbf{y}(t) + \mathbf{w}(t) \quad , Q = I \quad (14)$$

admits a flocking structure. The relative GFA decomposition has one ($q = 1$) latent factor.

4.1 Flocking and the structure of A

By Corollary 5, if the model (6) has a flocking structure, the spectral norm of the solution of the related Lyapunov equation must be unbounded. Such a property can be linked to the structure of the operator matrix A. Define the *radius of stability* of A as [Ackermann et al., 1993]

$$r(A) = \inf_{0\leq\theta\leq 2\pi} \|(\lambda I - A)^{-1}\|_2 . \quad (15)$$

Theorem 8. Assume $\mathbf{y}(t)$ satisfies (6)*, with $Q = I$ in the related Lyapunov equation* (7)*. Then a necessary condition for $\mathbf{y}(t)$ to have a flocking structure is that $r(A) \to 0$.*

Proof : The result follows from the inequalities

$$\frac{1}{2r(A) + r^2(A)} \leq \|\Sigma\|_2 \leq \frac{1}{r^2(A)} \quad (16)$$

of [Gahinet et al., 1990] and [Tippett and Marchesin, 1999] respectively. □

Remark 9. The above theorem applies in particular to the model of (14). In this case the matrix is symmetric and the radius of stability is just the distance of the largest eigenvalue from the unit circle, that is n^{-1}. However, when the matrix A is "highly" non-normal, as in the leader follower case, the behaviour of the stability radius is quite unpredictable and depends on the pseudospectrum of A. Problems of this kind are widely discussed in [Trefethen and Embree, 2005].

Quite unfortunately, the unboundedness (in the 2-norm sense) of the solution to the Lyapunov equation (7) does not generally imply the existence of a flocking structure. See the example below [Tippett and Marchesin, 1999].

Example 10. We consider the dynamics of a discretized one-dimensional advection equation

$$\frac{\partial y}{\partial t} + \frac{\partial y}{\partial x} = 0 . \quad (17)$$

We discretize the space variable x, define $\mathbf{y}_i(t) = y(x = i, t)$, $i \in \mathbb{N}$ and assume that independent random excitations with flat spectrum (white noise) are applied at every space location. Then the model (8) applies also here, with

$$A_n = \begin{bmatrix} 0 & & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \\ & \dots & 1 & 0 \end{bmatrix} . \quad (18)$$

a shift matrix. Again, the nesting property of the $\{A_n\}_{n\in\mathbb{N}}$ is satisfied. The associated family of Lyapunov equations (10) admits the solutions

$$\Sigma_n = \operatorname{diag}\{1, 2, \dots, n\}, \quad (19)$$

and $\|\Sigma\|_2 \to \infty$ as n grows to infinity. See the infinite dimensional unilateral shift example in [Przyluski, 1980]. Since the off-diagonal elements of Σ are all equal to 0, which means that there is no cross-correlation, no flocking structure can exist. This situation is also described in Example 2 in [Bottegal and Picci, 2011]. This represents a limit case, where Σ has unbounded elements and has no unique GFA representation (it can be viewed as a ∞-factor sequence).

5. CONCLUSIONS

We have discussed a new modeling paradigm for large dimensional aggregates of random systems based on the theory of Generalized Factor Analysis. The analysis of interesting classes of random fields, such as the linear evolution equation in (6), by using the decomposition of the steady state covariance has just been touched upon shortly. Their statistical identification can in principle be done by a limiting PCA procedure.

REFERENCES

J. Ackermann, A. Bartlett, D. Kaesbauer, W. Sienel, and R. Steinhauser. *Robust control.* Springer, 1993.

B.D.O. Anderson and M. Deistler. Generalized linear dynamic factor models - a structure theory. In *Proc. of the IEEE Decision and Control conference*, 2008.

G. Bottegal and G. Picci. A note on generalized factor analysis models. In *Proc. 50th Decision and Control Conference (CDC)*, pages 1485–1490, Orlando FLA, USA, 2011.

G. Bottegal and G. Picci. Modeling random flocks through generalized factor analysis. In *Proc. European Control Control Conference (ECC13)*, Zürich, Switzerland, 2013a.

G. Bottegal and G. Picci. Modeling complex systems by generalized factor analysis. Technical report, Department of Information Engineering, University of Padova, 2013b. URL `http://arxiv.org/abs/1301.2498`.

R. W. Brockett. On the control of a flock by a leader. In *Differential equations and topology. I*, volume 268 of *Tr. Mat. Inst. Steklova*, pages 56–63. MAIK Nauka/Interperiodica, Moscow, 2010.

G. Chamberlain and M. Rothschild. Arbitrage, factor structure and meanvariance analysis on large asset markets. *Econometrica*, 51(5):1281–1304, 1983.

F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Transactions on Automatic Control*, pages 852–862, 2007.

M. Deistler and C. Zinner. Modelling high-dimensional time series by generalized dynamic factor models: an introductory survey. *Communications on Information and Systems*, 7(2):153–166, 2007.

M. Deistler, B. D. O. Anderson, W. Chen, and A. Filler. Modelling high dimensional time series by generalized factor models, (semi-plenary lecture delivered at the MTNS 2010 Symposium. In *Proceedings of the Mathematical Theory of Networks and Systems Symposium (MTNS 2010)*, pages 323–329, Budapest, 2010a.

M. Deistler, B. D. O. Anderson, A. Filler, C. Zinner, and W. Chen. Generalized linear dynamic factor models: An approach via singular autoregressions. *European Journal of Control*, 3:211–224, 2010b.

F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale networks. *IEEE J. Selected Areas Commun.*, 26:634649, 2008.

M. Forni and M. Lippi. The generalized dynamic factor model: representation theory. *Econometric Theory*, 17: 1113–1141, 2001.

P. M. Gahinet, A. J. Laub, C. S. Kenney, and G. A. Hewer. Sensitivity of the stable discrete-time lyapunov equation. *Automatic Control, IEEE Transactions on*, 35 (11):1209–1217, 1990.

P. R. Halmos. Shifts on Hilbert spaces. *Journal für Reine und Angewandte Mathematik*, 208:102–112, 1961.

A Jadbabaie, J. Lin, and S. Morse. Coordination of groups of mobile agents using nearest neighbor rule. *IEEE Transactions on Automatic Control*, AC-48:988–1001, 2003.

R. Olfati-Saber. Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Transactions on Automatic Control*, 51:401420, 2006.

R. Olfati-Saber, J. A. Fax, and R. Murray. Consensus and cooperation in networked multi-agent systems. *Proc. of the IEEE*, 95:215–233, 2007.

K.M. Przyluski. The Lyapunov equation and the problem of stability for linear bounded discrete-time systems in Hilbert space. *Applied Mathematics and Optimization*, 6(1):97–112, 1980.

C.W. Reynolds. Flocks, herds and schools: a distributed behavioral model. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques SIGGRAPH 87*, page 2534. ACM Press: New York, 1987.

J. Shen. Cucker-smale flocking under hierarchical leadership. *SIAM Journ. Appl. Math*, 68:694–719, 2007.

A. Tahbaz-Salehi and A. Jadbabaie. Consensus over ergodic stationary graph processes. *IEEE Transactions on Automatic Control*, 55:225–230, 2010.

H. G. Tanner, A. Jadbabaie, and G. J. Pappas. Flocking in fixed and switching networks. *IEEE Trans. Automat. Contr.*, 52:863–868, 2007.

M. K. Tippett and D. Marchesin. Upper bounds for the solution of the discrete algebraic lyapunov equation. *Automatica*, 35(8):1485–1489, 1999.

L.N. Trefethen and M. Embree. *Spectra and pseudospectra: the behavior of nonnormal matrices and operators.* Princeton University Press, 2005.

J. J. P. Veerman, John S. Caughman, G. Lafferriere, and A. Williams. Flocks and formations. *J. Stat.Phys.*, 121, Vol 5-6:901–936, 2005.

T. Vicsek, A. Czirók, E. Ben-Jacoba, I. Cohen, and O. Schochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75:12261229, 1995.

L. Xiao, S. Boyd, and S.-J. Kim. Distributed average consensus with least-mean-square deviation. *Journal of Parallel and Distributed Computing*, 67(1):33–46, 2007.

APPENDIX

Proof of Theorem 6

Consider first the case $n = 3$ and write the solution to the related Lyapunov equation as

$$\Sigma_3 = \begin{bmatrix} p_1 & p_2 & p_3 \\ p_2 & p_4 & p_5 \\ p_3 & p_5 & p_6 \end{bmatrix} . \tag{20}$$

Then, simple calculations show that

$$p_1 = \frac{1}{1-a^2} \quad , \quad p_2 = p_3 = \frac{a}{(1+a)(1-a^2)} \quad ,$$

$$p_4 = p_6 = \frac{1}{1-a^2} + \frac{1}{(1+a)^2} + 2\frac{a^2}{(1+a)^2(1-a^2)}$$

$$p_5 = \frac{1}{(1+a)^2} + 2\frac{a^2}{(1+a)^2(1-a^2)} \tag{21}$$

Now assume that, for a given $n \geq 3$, the solution to the equation $X_n - A_n X_n A_n^\top = I_n$ has the form

$$\Sigma_n = \begin{bmatrix} p_1 & p_3 & p_3 & p_3 & \dots & p_3 \\ p_3 & p_4 & p_5 & p_5 & \dots & p_5 \\ p_3 & p_5 & p_4 & p_5 & \dots & p_5 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ p_3 & p_5 & \dots & p_5 & p_4 & p_5 \\ p_3 & p_5 & \dots & p_5 & p_5 & p_4 \end{bmatrix} ; \tag{22}$$

our goal is to show that Σ_{n+1} has an analogous structure, that is

$$\Sigma_{n+1} = \begin{bmatrix} \Sigma_n & p \\ p^\top & p_4 \end{bmatrix} , \tag{23}$$

where $p = [p_3 \; p_5 \; \dots \; p_5]^\top$. To this end, express the variable X_{n+1} as

$$X_{n+1} = \begin{bmatrix} X_n & z \\ z^\top & u \end{bmatrix}$$

and the matrix A_{n+1} as

$$A_{n+1} = \begin{bmatrix} A_n & 0 \\ b^\top & a \end{bmatrix} ,$$

where $b = [1-a \; 0 \; \dots \; 0]^\top$. Then the related Lyapunov equation has the form

$$\begin{bmatrix} X_n & z \\ z^\top & u \end{bmatrix} - \begin{bmatrix} A_n & 0 \\ b^\top & a \end{bmatrix} \begin{bmatrix} X_n & z \\ z^\top & u \end{bmatrix} \begin{bmatrix} A_n^\top & b \\ 0 & a \end{bmatrix} = I_{n+1} , \tag{24}$$

which can be rewritten as

$$\begin{bmatrix} X_n - A_n X_n A_n^\top & (I_n - aA_n)z - A_n X_n b \\ z^\top (I_n - aA_n^\top) - b^\top X_n A_n^\top & (1-a^2)u - b^\top X_n b - 2ab^\top z \end{bmatrix} =$$

$$= \begin{bmatrix} I_n & 0 \\ 0 & 1 \end{bmatrix} . \tag{25}$$

The top-left block of (25) admits the solution given by (22). Then, by inserting this into the top-right block, one then gets $z = p$. Finally, by exploiting the former findings, from the bottom-right block one has $u = p_4$, and hence the solution is exactly (23). Hence, one can easily observe that the matrix $\bar{\Sigma}_n$, obtained by discarding the first row and column from Σ_n, has the structure

$$\begin{bmatrix} p_5 & p_5 & \dots \\ p_5 & p_5 & \\ \vdots & & \ddots \end{bmatrix} + \mathrm{diag}\{p_4 - p_5, \dots, p_4 - p_5\} \tag{26}$$

that is, it admits a rank-one plus diagonal decomposition, where the vector generating the rank-one matrix is $\bar{f} = [\sqrt{p_5} \; \sqrt{p_5} \; \dots]$, with $\sqrt{p_5} = (1+a^2)^{\frac{1}{2}}/((1+a)(1-a^2)^{\frac{1}{2}})$, while the elements of the diagonal matrix are $p_4 - p_5 = 1/(1-a^2)$. Now, to complete the proof we need to show that also the matrix Σ_n admits a similar decomposition, i.e.

$$\Sigma_n = \begin{bmatrix} f_1 \\ \bar{f} \end{bmatrix} \begin{bmatrix} f_1 & \bar{f}^T \end{bmatrix} + \mathrm{diag}\{\sigma_1^2, 1/(1-a^2), \dots, 1/(1-a^2)\} .$$

This can be done be observing that, for any integer $k > 0$, it has to be $p_3 = f_1 \bar{f}(k)$, and so $f_1 = a/(1-a^4)^{\frac{1}{2}}$. Moreover, σ_1^2 is easily found by computing $\sigma_1^2 = p_1 - f_1^2 = 1$. Finally, since by comparing the leader dynamics

$$\mathbf{y}_1(t) = a\mathbf{y}_1(t-1) + \mathbf{w}_1(t-1)$$

with its GFA decomposition $\mathbf{y}_1(t) = f_1 \mathbf{x}(t) + \tilde{\mathbf{y}}_1(t)$, where both $\tilde{\mathbf{y}}_1(t)$ and $\mathbf{w}_1(t-1)$ are white noise with the same variance, it has to be $\mathbf{x}(t) = (1-a^4)^{\frac{1}{2}} \mathbf{y}_1(t-1)$.

Proof of Theorem 7

For $n \geq n_0$, consider the Lyapunov equation

$$\Sigma_n = \bar{A}_n \Sigma_n \bar{A}_n^\top + I_n ,$$

whose solution can be written

$$\Sigma_n = \sum_{j=0}^{\infty} \bar{A}_n^j (\bar{A}_n^j)^\top . \tag{27}$$

Since $\bar{A}_n$ is symmetric, for every j the decomposition

$$\bar{A}_n^j (\bar{A}_n^j)^T = U_n S_n^{2j} U_n^\top$$

holds, with S_n being the matrix of the singular values of A and U_n a unitary matrix whose columns are the (normalized) eigenvectors of $\bar{A}_n$. Note that one of such singular values is $\left(1 - \frac{1}{n}\right)^2$ and the relative eigenvector is $\frac{1}{\sqrt{n}} \mathbb{1}_n$, i.e. the normalized vector of all 1's in $\in \mathbb{R}^n$. The other eigenvalues are strictly stable. Then we can express Σ_n as

$$\begin{aligned} \Sigma_n &= U_n \left(\sum_{j=0}^{\infty} S_n^{2j} \right) U_n^\top \\ &= \frac{\mathbb{1}}{\sqrt{n}} \left(\sum_{j=0}^{\infty} \left(1 - \frac{1}{n}\right)^{2j} \right) \frac{\mathbb{1}}{\sqrt{n}}^\top + \tilde{U}_n \left(\sum_{j=0}^{\infty} \tilde{S}_n^{2j} \right) \tilde{U}_n^\top \\ &= \mathbb{1} \frac{n}{2n+1} \mathbb{1}^\top + \tilde{U}_n \left(\sum_{j=0}^{\infty} \tilde{S}_n^{2j} \right) \tilde{U}_n^\top , \end{aligned} \tag{28}$$

where $\tilde{U}_n$ and $\tilde{S}_n$ are obtained from U_n and S_n by removing the parts related to the eigenvalue $\left(1 - \frac{1}{n}\right)^2$. Now, take the averaging sequence (3)

$$a_n = \frac{1}{n} \left[\mathbb{1}_n^\top \, 0 \, \dots \right] \quad , \mathbb{1}_n \in \mathbb{R}^n \tag{29}$$

and apply it to Σ_n, that is, compute $\frac{1}{n} \mathbb{1}_n^\top \Sigma_n \mathbb{1}_n \frac{1}{n}$. Then, letting $n \to \infty$, the second term on the right hand side of (28) vanishes, while the first term gives

$$\frac{\mathbb{1}_n^\top \mathbb{1}_n \mathbb{1}_n^\top \mathbb{1}_n}{n(2n+1)} = \frac{n}{2n+1} , \tag{30}$$

which converges asymptotically to a finite value. One can easily verify that the averaging sequence (29) is the only sequence converging to nonzero values.