University of Granada

Department of Statistics and Operational Research

PhD Thesis

Doctoral program in Mathematics and Statistics

New advances in the estimation problem in systems with random failures

Irene García Garrido

Granada, 2015

Editor: Universidad de Granada. Tesis Doctorales
Autora: Irene García Garrido
ISBN: 978-84-9125-531-4
URI: http://hdl.handle.net/10481/42603

Universidad de Granada

Departamento de Estadística e Investigación Operativa

Tesis Doctoral

Programa de doctorado en Matemáticas y Estadística

Aportaciones al problema de estimación en sistemas con fallos aleatorios

Irene García Garrido

Granada, 2015

Tesis Doctoral

Aportaciones al problema de estimación en sistemas con fallos aleatorios

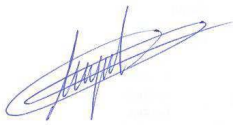
Memoria de tesis presentada por Dña. Irene García Garrido para optar al grado de Doctora por la Universidad de Granada.

Esta Tesis Doctoral ha sido dirigida por la Dra. Dña. Josefa Linares Pérez, Catedrática de Universidad del Departamento de Estadística e Investigación Operativa de la Universidad de Granada, y la Dra. Dña. Carmen Raquel Caballero Águila, Profesora Titular de Universidad del Departamento de Estadística e Investigación Operativa de la Universidad de Jaén.

Granada, 23 de junio de 2015

Directoras de la Tesis

Fdo.: Josefa Linares Pérez



Fdo.: Carmen Raquel Caballero Águila

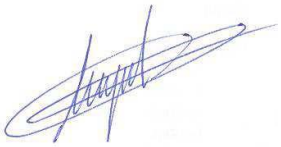
Doctoranda

Fdo.: Irene García Garrido

La doctoranda Irene García Garrido y las directoras de la tesis Josefa Linares Pérez y Carmen Raquel Caballero Águila garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por la doctoranda bajo la dirección de las directoras de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 23 de junio de 2015

Directoras de la Tesis



Fdo.: Josefa Linares Pérez

Fdo.: Carmen Raquel Caballero Águila

Doctoranda

Fdo.: Irene García Garrido

Esta tesis doctoral ha sido financiada por la beca del programa de Formación de Profesorado Universitario (FPU) del Ministerio de Educación, Cultura y Deporte, en su resolución del 20 de diciembre de 2011, con código de referencia AP2010-1553, así como por los proyectos No. MTM2011-24718 del Ministerio de Ciencia e Innovación, No. P07-FQM-02701 de la Junta de Andalucía y No. MTM2014-52291-P del Ministerio de Economía y Competitividad.

Dedicada a Antonio,
mis padres José Carlos y Eva M^a,
y mis hermanos Laura y Álvaro.

Agradecimientos

Comienzo agradeciendo a mis directoras de tesis, Josefa Linares Pérez y Raquel Caballero Águila, la dedicación y el apoyo que he recibido durante la elaboración de esta tesis doctoral. A Pepa porque desde mis últimos años en la Licenciatura de Matemáticas confió en mí y me estimuló para ampliar mis estudios en el área de la estadística, me introdujo en el campo de la investigación y me motiva e ilusiona de forma permanente. A Raquel por dedicarme su tiempo y paciencia en mi labor investigadora. A ambas les agradezco las reiteradas revisiones críticas para concretar ideas y a la vez perfeccionar la redacción y el estilo de los diversos trabajos desarrollados en esta tesis. I would also like to thank Zidong Wang for his hospitality, his guidance in research and the teaching of new knowledge during my stay at Brunel University in London.

A D. Ramón Gutiérrez Jáimez y D. Andrés González Carmona, anterior y actual directores del Departamento de Estadística e Investigación Operativa, por su amable acogida en este departamento.

A mi querido Antonio por estar siempre ahí, haciéndome sentir muy feliz cada día. Por su comprensión, paciencia, ánimo y por compartir conmigo ilusiones en lo personal y en lo profesional. También a su familia que siempre se preocupa por nuestro bienestar.

A mi familia por su constante ánimo y cariño. A mis padres, José Carlos y Eva Mª, porque sin pertenecer a este ámbito me han ayudado en la redacción de este trabajo de investigación. A mi hermana Laura porque me acompañó durante

los meses de estancia en Londres, por sus buenos consejos y porque me transmitió optimismo. A mi hermano Álvaro por hacerme el trabajo más ameno con sus pequeñas interrupciones durante mis tardes de estudio. A mis abuelas, Primy y "la Yaya" Julia, por estimularme y decirme constantemente que continúe en el "puesto de los valientes".

A mis amigos, en especial Lydia y Antonio, por su cariñosa acogida y compañía en Londres; y por estar siempre disponibles cuando los necesito. También nombrar a Rosa, Fran, Elvira, Elena, Marta y Cristina porque con ellos he podido salir de la rutina y relajarme en los momentos de más tensión.

Finalmente, a mis compañeros del Departamento de Estadística e Investigación Operativa porque hemos compartido agradables momentos durante estos años que perdurarán en mi memoria.

Contents

Abstract

The aim of this PhD thesis is to address least-squares estimation problems in discrete-time linear systems from noisy measurements derived from multiple sensors, affected by random parameters which model different situations of failure in the mechanism or the transmission of the measurements. According to the kind of systems considered, the main contributions of this PhD thesis are summarized below:

Sensor network systems with uncertain observations. These systems describe situations in which the mechanism of measurements may be randomly interrupted, in the sense that, at each instant of time, there is a positive probability that the corresponding observation is only noise, i.e., the observations may not contain information about the state. This kind of uncertainty is modeled by including in the observation equation not only an additive noise, but also a multiplicative noise component described by a sequence of Bernoulli random variables whose values, one or zero, indicate the presence or absence of the state in the corresponding measurement. In cases in which the Bernoulli variables are assumed to be correlated at instants that differ by m units of time, on the one hand, centralized and distributed fusion linear estimators are designed (**Chapter 1**) and, on the other, in order to improve the linear estimators, quadratic estimators are obtained using the centralized fusion method (**Chapter 2**).

Sensor network systems with failures in the measurements, in which the observations from the different sensors may contain only partial information about the

state. This kind of failure is more general than the previous one and it is described by a sequence of independent random variables with discrete probability distribution over the interval $[0, 1]$. For this class of systems, under the assumption that the system additive noises are autocorrelated and also cross-correlated, recursive linear filtering algorithms are derived using the centralized and distributed fusion methods (**Chapter 3**).

Sensor network systems with random parameter matrices. This kind of systems constitute a more general framework than the previous ones since the state and/or the observation equations may be affected by random parameter matrices, thus covering numerous real situations with random failures in the measurements. First, we consider independent random state transition matrices, and one-step correlated and cross-correlated random parameter matrices in the observation equation; it is also assumed that the system noises are autocorrelated and cross-correlated. Using the centralized fusion method, a recursive linear filtering algorithm is obtained and the results are applied to multi-sensor systems with failures in the measurements described by random variables with discrete distribution over the interval $[0, 1]$, and to multi-sensor systems with randomly delayed observations (**Chapter 4**). Second, the linear estimation problem in systems with independent random parameter matrices and correlated noises is addressed, using the distributed fusion method (**Chapter 5**). Finally, centralized quadratic estimators are obtained in systems with independent random parameter matrices and noises, and they are applied to systems with random failures in the measurements, described by different sequences of random variables with discrete probability distribution over the interval $[0, 1]$ (**Chapter 6**).

Resumen

Esta tesis doctoral tiene como objetivo el estudio de problemas de estimación mínimo cuadrática en sistemas lineales en tiempo discreto a partir de observaciones ruidosas procedentes de múltiples sensores, afectadas por parámetros aleatorios que modelizan diferentes tipos de fallo en el mecanismo de medidas o en la transmisión de las mismas. A continuación se resumen las principales aportaciones de la tesis en función del tipo de sistemas considerado:

Sistemas de redes de sensores con observaciones inciertas. Dichos sistemas modelizan situaciones en las que el mecanismo de medidas puede interrumpirse aleatoriamente de forma que, en cada instante de tiempo, existe una probabilidad positiva de que la observación correspondiente sea únicamente ruido, es decir, no contenga información sobre el estado del sistema. Para modelizar este tipo de incertidumbre, se supone que las observaciones disponibles para la estimación están afectadas, no solo por ruidos aditivos, sino también por una componente ruido multiplicativa descrita por una sucesión de variables aleatorias de Bernoulli, cuyos valores, uno o cero, indican la presencia o ausencia del estado en la observación correspondiente. Suponiendo que dichas variables están correladas en instantes que se diferencian m unidades de tiempo, por una parte, mediante los métodos de fusión centralizado y distribuido, se estudia el problema de estimación lineal (**Capítulo 1**) y, por otra, con objeto de mejorar a los estimadores lineales, se obtienen estimadores cuadráticos utilizando el método de fusión centralizado (**Capítulo 2**).

Sistemas de redes de sensores con fallos en las medidas, en los que las observa-

ciones de los distintos sensores pueden contener sólo información parcial del estado. Este tipo de fallo es más general que el anterior y está descrito por una sucesión de variables aleatorias escalares independientes con distribución de probabilidad discreta en el intervalo $[0, 1]$. Para esta clase de sistemas, bajo la hipótesis de que los ruidos aditivos del sistema son autocorrelados y también correlados entre sí, se establecen algoritmos recursivos de filtrado lineal mediante los métodos de fusión centralizado y distribuido (**Capítulo 3**).

Sistemas de redes de sensores con matrices aleatorias. Estos sistemas constituyen un marco más general que los anteriores, ya que las ecuaciones del estado y/o de las observaciones pueden estar afectadas por matrices aleatorias, lo cual permite modelizar una amplia variedad de fallos aleatorios en las medidas. En primer lugar, se considera que las matrices de transición del estado son independientes y las matrices de la ecuación de observación son autocorreladas y correladas entre sí en instantes consecutivos. Asimismo, los ruidos que intervienen en el sistema se suponen autocorrelados y correlados entre sí. Utilizando el método de fusión centralizado se obtiene un algoritmo recursivo de filtrado lineal y los resultados se aplican a sistemas con múltiples sensores y fallos en las medidas modelizados por variables con distribución discreta en el intervalo $[0, 1]$, y también a sistemas con observaciones retrasadas aleatoriamente procedentes de múltiples sensores (**Capítulo 4**). En segundo lugar, se aborda el problema de estimación lineal en sistemas con matrices aleatorias independientes y ruidos correlados, utilizando el método de fusión distribuido (**Capítulo 5**). Finalmente, se obtienen estimadores cuadráticos centralizados en sistemas con matrices aleatorias y ruidos independientes, y se aplican a sistemas con fallos aleatorios en las medidas, descritos mediante diferentes sucesiones de variables con distribución discreta en el intervalo $[0, 1]$ (**Capítulo 6**).

Introduction

In classical estimation theory for discrete-time linear stochastic systems, the problem of estimating the state from related observations is addressed assuming that the state is always present in the observations and that the disturbances within them are described only by additive noises. The solution of this problem is obtained from the conditional expectation of the state given the observations; that is, from the optimal least-squares (LS) estimator. The Kalman filter provides a recursive algorithm for the optimal LS estimator when the additive white noises and the initial state are Gaussian and mutually independent (or, equivalently, uncorrelated due to the assumption of Gaussianity). This type of algorithm has received considerable attention in the scientific community due to its wide applicability in many practical situations, such as video and laser tracking systems, satellite navigation and radar and meteorological applications [1]. Since the first publication of the Kalman filter [2] in 1960, numerous results and several solution methods have been reported about the state estimation problem from noisy observations. The estimation algorithms proposed depend on the system model that represents the possible relationships between the unknown state and the observable variables and are also influenced by the assumptions about the noise processes.

In non-Gaussian systems, the Kalman filter only provides the LS linear estimator and, in general, from a computational point of view, the optimal LS estimator is difficult to obtain. In consequence, it is necessary to search for suboptimal estimators that are better than the linear ones, while maintaining their recursivi-

ty and simplicity of calculation. Along these lines, quadratic and arbitrary-order polynomial estimation problems in linear systems with non-Gaussian noises have been analyzed by [3] and [4], respectively. Furthermore, polynomial estimators have been used to resolve various problems in the field of signal processing, such as prediction, detection and control, as well as problems of image restoration (see, for example, [5–7], among others) and it has been sufficiently proven that, in general, polynomial estimators significantly outperform linear ones.

Systems with multiplicative noises in the state and/or observation equations constitute another kind of non-Gaussian systems in which the Kalman filter does not provide the optimal LS estimator and for which it is necessary to look for sub-optimal estimators. Such systems have received considerable attention in recent years, as this kind of formulation arises in many applications, for example image processing problems and communication systems. Therefore, under different hypotheses and performance criteria, the LS linear estimation problem in systems with multiplicative noises is now an active research area (see [8], [9], and references therein).

Some classes of systems, where the influence of multiplicative noises affects only the measurements of the model, have important applications, for example, in cases in which there are intermittent failures in the observation mechanism, fading phenomena in propagation channels, accidental losses of measurements or data inaccessibility at certain times. This type of situation, termed *systems with missing measurements*, is modeled by including in the observation equation, in addition to the additive noise, a multiplicative noise component described by a sequence of binary random variables taking the values one and zero (Bernoulli random variables); the value one indicates that the state is present in the corresponding measurement at that time, whereas the value zero reflects the fact that the state is missing and, hence, the corresponding observation is only noise. Nahi [10] first analyzed the LS linear filtering problem in this kind of systems, assuming that the Bernoulli

variables modeling the uncertainty in the observations are independent and that the additive noises are uncorrelated, by obtaining an estimation algorithm with a recursive structure similar to that of the Kalman filter. Subsequently, Monzingo [11] completed these results by analyzing the LS smoothing problem. Jaffer and Gupta [12] studied the optimal estimation problem in this kind of systems, and concluded that computation of the optimal estimator requires exponential memory growth. This drawback motivates the development of a broad-based investigation of the estimation problem in the face of missing measurements, focused on obtaining suboptimal solutions under diverse assumptions about the state and the noise processes involved in the model.

The treatment of practical situations in which the uncertainty of the observations cannot be represented by independent random variables (this occurs, for example, in problems of signal transmission through multiple channels) was initially addressed by Hadidi and Schwartz [13] and by Monzingo [14], who established recursive algorithms for obtaining linear filtering and smoothing estimators, respectively. Jackson and Murthy [15] subsequently analyzed a different situation, assuming that the variables modeling the uncertainty are correlated at consecutive instants and that the proposed LS linear filtering algorithm can be applied, for example, in signal transmission models with stand-by sensors in which any failure in the transmission is detected immediately and the old sensor is then replaced, thus avoiding the possibility of the signal being missing in two successive observations. Afterwards, in [16], the polynomial filtering and smoothing problems were addressed under the same correlation assumption for the variables which model the phenomenon of missing measurements.

In all of the above-mentioned papers, the observations available for the estimation are assumed to come either from a single sensor or from multiple sensors with identical uncertainty characteristics. In recent years, this approach has been generalized to consider multi-sensor systems featuring different statistical proper-

ties at each sensor, motivated by the increasing development of sensor networks for data acquisition and signal processing, and by networked communication systems involving heterogeneous measurement devices (see e.g. [17] and [18], among others). Data measurements collected by sensors are usually imprecise. The uncertainty in sensor data can arise from various sources, including the inherent limitations to the accuracy with which the sensed data is acquired and external ones such as efficiency or battery life. This reality means that missing measurements, time delays and/or data packet dropouts often arise in multi-sensor systems. In addition, an issue of great importance in multi-sensor systems is the data fusion technique employed, that is, how the measured data from the different sensors are combined to address the estimation problem in order to achieve better estimations than could be achieved by the use of a single sensor. The most common methods to manage sensor data fusion in estimation problems are the centralized and distributed fusion methods.

In the first of these methods, all the sensor data are transmitted to a central processor (fusion center) where the data fusion is performed; specifically, the measured data from the different sensors are stacked as a single-sensor measurement, from which the optimal LS linear estimators are derived. The optimal centralized estimation problem, in the linear minimum variance sense, has been investigated for sensor network systems including some of the aforementioned sources of uncertainty in the observation mechanism. Specifically, centralized linear estimators are derived in [19], in which the phenomenon of missing measurements is described by sequences of independent Bernoulli variables; and in [20], where centralized linear and quadratic estimation problems are addressed under the assumption that the Bernoulli random variables modeling the uncertainty at each sensor are correlated at consecutive sampling times. The centralized estimation problem in multi-sensor systems with random delays is also addressed in [21], [22] under different delay rates for the different sensors; and in [23], [24] for systems with multiple sensors of

different packet dropout rates.

Nevertheless, the centralized fusion method has various drawbacks, including poor survivability and reliability, exacting communication requirements and high computational cost. Various distributed fusion algorithms have been proposed to overcome these disadvantages. In the distributed fusion method, each sensor estimates the state based on its own measurement data, and this estimate is then sent to the central processor (fusion center), where the fusion is performed in accordance with a specific information fusion criterion. For example, under the assumption of normal distribution, a distributed fusion estimator is proposed in [25], based on the maximum likelihood criterion, and [26] propose the distributed fusion criterion weighted by matrices in the linear minimum variance sense, which is equivalent to the maximum likelihood fusion criterion under the normality assumption. The distributed fusion estimation method in networked systems with unreliable network transmission, such as missing measurements, packet delays and dropouts, has subsequently been used by [27–31], among others.

In most of the above-mentioned papers, the presence of independent white noises is assumed; however, this assumption must be undermined in many real-world problems, in which noise correlation may be present. This problem arises, for example, when a target is taking an electronic countermeasure, e.g. noise jamming [32], or if the system noises are state-dependent, and then there may be cross-correlation, both between different sensor noises and between process noise and sensor noises. Moreover, if all the sensors are observed in the same noisy environment, the measurement noises of different sensors are usually correlated. In view of these considerations, the estimation problem in systems with correlated noises has become an important research topic. Whithin this context, the optimal Kalman filtering fusion problem in systems with cross-correlation between the process noise and the measurement noises at the same sampling time is addressed in [33]. In general, when the noise process and the measurement noises are correlated

and cross-correlated at different sampling times, it is difficult to obtain optimal estimators; this limitation has spurred broader research into suboptimal Kalman-type estimation problems. A Kalman-type recursive filter is presented in [34] for systems with finite-step correlated process noises, and the filtering problem with multi-step correlated process and measurement noises is investigated in [35]. The optimal robust non-fragile Kalman-type recursive filtering problem is studied in [36] for a class of uncertain systems with finite-step autocorrelated measurement noises and multiple packet dropouts. Finally, the problem of distributed weighted robust Kalman filter fusion is studied in [37] for linear systems with state-dependent multiplicative noise and autocorrelated and cross-correlated noises.

Clearly, systems with multiplicative noises in the state and/or observation equations are special cases of systems with *random parameter matrices*, which have important practical applications in many areas, such as the digital control of chemical processes, radar control, navigation systems or economic systems, and the estimation problem in this type of systems has acquired significant research interest (see, for instance [38] and references therein). The linear filtering problem in systems with independent random state transition and measurement matrices has been addressed in [39] and [40], who transformed the original system into one with deterministic parameter matrices and state-dependent process and measurement noises, to which the Kalman filter was applied. In [39], the Kalman filter is applied without providing any theoretical justification, while in [40], it is shown that in undemanding conditions, the transformed system satisfies the Kalman filter requirements and, hence, optimal linear estimators are derived for systems with independent random parameter matrices. Considering scalar measurements with random observation matrices, the quadratic LS filtering problem was recently addressed in [41], who applied the Kalman filter to a suitably augmented system with deterministic observation matrices.

However, there are many practical situations in which the random parameter

matrices are not independent; for example, when random sensor delays and/or multiple packet dropouts are converted into observation models with random measurement matrices [42], or when a nonlinear system is linearized around the random state estimate to apply the extended Kalman filter (for other realistic systems and backgrounds where the model parameter matrices are random and correlated, see [43] and [44]). In [45], systems with deterministic transition matrices and one-step correlated random measurement matrices are considered, and the optimal recursive state estimation is derived by converting the observation equation into one with deterministic measurement matrices and by applying the optimal Kalman filter to the case of one-step correlated measurement noises. For sensor network systems with random parameter matrices and autocorrelated and cross-correlated noises, the linear estimation problem has been addressed by [46] and by [47], among others.

The above comments constitute a brief overview of the background and the current state of the different problems addressed in the six chapters of this PhD thesis. A more specific analysis is presented in the introduction to each of these chapters. The general aim of the thesis is to study LS estimation problems in discrete-time linear systems from noisy measurements derived from multiple sensors, affected by random parameters and which model different situations of failure in the mechanism or the transmission of the measurements. Each of the chapters constitutes an original paper in which a specific topic related to this study is addressed. Therefore, each chapter is presented as an independent paper, with its own abstract, sections, appendices and references, in the same form as when it was originally published or submitted. To date, the papers in chapters 1-4 have been published and those in chapters 5 and 6 have been submitted and are under review.

This PhD thesis is organized as follows. **Chapters 1 and 2** consider linear discrete-time stochastic systems with missing measurements derived from multiple sensors when, at each sensor, the random variables modeling the missing mea-

surements are correlated at instants that differ by m units of time. Centralized and distributed fusion linear estimation problems are addressed in **Chapter 1**, after which the centralized fusion quadratic estimation problem is addressed in **Chapter 2**. Multi-sensor systems with missing measurements, with correlated and cross-correlated noises, when the missing measurements phenomenon is described by different sequences of scalar random variables with arbitrary discrete probability distribution over the interval $[0, 1]$, are considered in **Chapter 3**, in which the centralized and distributed fusion linear estimation problems are addressed. Next, in **Chapter 4**, systems with independent random transition matrices and correlated random parameter matrices in the observation equation, together with correlated system noises, are considered. The centralized fusion linear filter for this kind of systems is obtained and the results are applied to multi-sensor systems with missing measurements and random delays. In **Chapter 5**, the distributed fusion linear estimation problem in sensor network systems with independent random parameter matrices and correlated noises is investigated. Finally, the centralized quadratic filtering problem is addressed in **Chapter 6**, considering linear systems with independent random parameter matrices and with noises.

Introducción

En la teoría clásica de estimación en sistemas estocásticos lineales en tiempo discreto, se aborda el problema de estimación del estado a partir de observaciones relacionadas con él, suponiendo que el estado que se desea estimar está siempre presente en las observaciones y que las perturbaciones de éstas se deben, únicamente, a ruidos aditivos. La solución de este problema es la esperanza condicionada del estado dadas las observaciones; es decir, el estimador óptimo de menor error cuadrático medio. El célebre filtro de Kalman proporciona un algoritmo recursivo para la obtención del estimador óptimo de mínimos cuadrados en sistemas lineales bajo condiciones de gaussianidad e independencia mutua de los ruidos aditivos y el estado en el instante inicial. Dicho algoritmo ha recibido considerable atención en la comunidad científica debido a sus diversas aplicaciones tecnológicas, como por ejemplo, en sistemas de control modernos, sistemas de seguimiento láser y navegación por satélite, meteorología, etc. [1]. A partir de la publicación del filtro de Kalman [2] en 1960, numerosos resultados han contribuido a la existencia de una amplia literatura para tratar el problema de estimación del estado a partir de observaciones ruidosas. Los algoritmos de estimación propuestos, dependen de las relaciones establecidas entre el estado desconocido y las variables observables, y también de las hipótesis sobre los procesos ruido.

Es importante resaltar que en sistemas no gaussianos el filtro de Kalman sólo proporciona el estimador de menor error cuadrático medio en la subclase de estimadores lineales y, en general, el estimador óptimo es difícil de obtener desde

el punto de vista computacional. Esta situación motiva la necesidad de buscar estimadores subóptimos que mejoren al lineal, manteniendo sus propiedades de recursividad y simplicidad de cálculo. En esta línea, destacan los trabajos sobre estimación cuadrática y polinomial de grado arbitrario en sistemas lineales con ruidos no gaussianos establecidos en [3] y [4], respectivamente. Además, los estimadores polinomiales tienen diversas aplicaciones en problemas de procesamiento de señales, tanto en predicción, detección y control, como en problemas de restauración de imágenes (véanse, por ejemplo, [5–7], entre otros) y, en términos generales, ha quedado suficientemente probada la efectividad de los estimadores polinomiales frente a los lineales.

Sistemas con ruidos multiplicativos en las ecuaciones del estado y/o de la observación constituyen otro tipo de sistemas no gaussianos en los que el filtro de Kalman no proporciona el estimador óptimo y, por tanto, es necesario buscar estimadores subóptimos. Esta clase de sistemas ha recibido gran atención recientemente, debido a sus diversas aplicaciones prácticas como en problemas de procesamiento de imágenes y sistemas de comunicación. Por lo tanto, bajo diferentes hipótesis y tratamientos, el estudio del problema de estimación lineal en sistemas con ruidos multiplicativos se ha convertido en un área de investigación importante en los últimos años (véanse, por ejemplo [8], [9], y sus referencias).

Algunos tipos de sistemas en que la influencia del ruido multiplicativo afecta únicamente a las observaciones del modelo, son adecuados, por ejemplo, en el caso de fallos intermitentes en el mecanismo de observación, pérdida accidental de algunas medidas, o inaccesibilidad de datos durante ciertos instantes de tiempo. Este tipo de sistemas, denominado *sistemas con observaciones inciertas*, incluyen en la ecuación de observación, además de un ruido aditivo, una componente ruido multiplicativa, descrita por una sucesión de variables aleatorias de Bernoulli, cuyos valores uno o cero, indican la presencia o ausencia del estado en la observación correspondiente. Nahi [10] planteó por primera vez el problema de filtrado li-

neal de menor error cuadrático medio en este tipo de sistemas, suponiendo que las variables de Bernoulli que describen la incertidumbre en las observaciones son independientes y los ruidos aditivos del sistema están incorrelados. Para ello desarrolló un algoritmo recursivo, con estructura análoga a la del filtro de Kalman. Posteriormente, Monzingo [11] generalizó estos resultados para resolver el problema de suavizamiento de menor error cuadrático medio. Jaffer y Gupta [12] abordaron el problema de estimación óptima en sistemas con observaciones inciertas, concluyendo que el cálculo del estimador óptimo requiere un crecimiento exponencial de memoria. Este inconveniente motiva el desarrollo de una amplia investigación en el problema de estimación en dichos sistemas, centrada en la búsqueda de estimadores subóptimos bajo diversas hipótesis sobre el estado y los procesos ruido que intervienen en el modelo.

El tratamiento de situaciones prácticas en que la incertidumbre de las observaciones no puede representarse por variables aleatorias independientes (lo que ocurre, por ejemplo, en problemas de transmisión de señales a través de múltiples canales) fue abordado inicialmente por Hadidi y Schwartz [13] y Monzingo [14], quienes establecieron algoritmos recursivos para la obtención del filtro y suavizador lineal, respectivamente. Posteriormente, Jackson y Murthy [15] analizaron una situación diferente, suponiendo que las variables que modelizan la incertidumbre están correladas en instantes consecutivos y, el algoritmo de filtrado lineal de menor error cuadrático medio propuesto pudo aplicarse, por ejemplo, en modelos de transmisión de señales con sensores *stand-by* programados para ser automáticamente reemplazados al detectar la anomalía, evitando de esta forma la posibilidad de que el estado esté ausente en dos observaciones consecutivas. Más tarde, en [16] se abordaron los problemas de filtrado y suavizamiento polinomiales bajo los mismos supuestos de correlación sobre las variables que modelizan el fenómeno de incertidumbre en las observaciones.

En todos los trabajos a los que se ha hecho referencia anteriormente, se supone

que las observaciones disponibles para la estimación proceden de un único o de varios sensores con las mismas características de incertidumbre. En los últimos años, este estudio se ha generalizado a sistemas con múltiples sensores con diferentes propiedades estadísticas en cada sensor, motivado por el creciente desarrollo de redes de sensores para la adquisición de datos y procesamiento de señales, y por la utilización de sistemas de comunicación en red que incluyen dispositivos de medida heterogéneos (véanse, por ejemplo [17] y [18], entre otros). La incertidumbre en las medidas de los distintos sensores puede deberse a diversas fuentes, causadas por limitaciones intrínsecas en la precisión con la que se adquiere la información, y limitaciones externas tales como la eficiencia o la duración de las baterías. Por lo tanto, la posibilidad de observaciones inciertas, retrasos en la recepción de las medidas y/o pérdida de las medidas en ciertos instantes de tiempo se contempla frecuentemente en sistemas con múltiples sensores. Una cuestión fundamental para abordar el problema de estimación en dichos sistemas, es la técnica utilizada para fusionar las medidas procedentes de los distintos sensores, con el fin de lograr mejores estimaciones que las que se podrían alcanzar con la utilización de un único sensor. Para combinar la información procedente de los diferentes sensores, frecuentemente se utilizan los métodos de fusión centralizado y distribuido.

En el método de fusión centralizado, las observaciones procedentes de los sensores se transmiten a un procesador central (centro de fusión) en el que se realiza la fusión de la información. En concreto, en cada instante de tiempo, los datos obtenidos de los diferentes sensores se combinan formando un único único vector de medidas y, a partir de éstos, se obtienen estimadores lineales de mínimos cuadrados. El problema de estimación óptima mediante el método centralizado se ha investigado en sistemas de redes de sensores que incluyen en las observaciones algunas de las incertidumbres mencionadas anteriormente. Específicamente, en [19] se obtienen estimadores lineales centralizados suponiendo observaciones inciertas descritas mediante sucesiones de variables de Bernoulli independientes; y en [20],

se abordan los problemas de estimación lineal y cuadrática bajo el supuesto de que las variables de Bernoulli que modelizan la incertidumbre en cada sensor son correladas en instantes consecutivos. También se han obtenido estimadores centralizados en sistemas con múltiples sensores con retrasos aleatorios en [21], [22] bajo diferentes tasas de retraso en los diferentes sensores; y en [23], [24] para sistemas con múltiples sensores y pérdidas aleatorias de medidas, considerando también distintas tasas de pérdida.

El enfoque centralizado tiene algunos inconvenientes como, por ejemplo, el elevado coste computacional, que usualmente pueden reducirse con el método de fusión distribuido consistente en obtener, a partir de las observaciones procedentes de cada sensor, estimadores lineales locales de mínimos cuadrados, y posteriormente dichos estimadores se envían al procesador central (centro de fusión) donde se realiza la fusión de la información aplicando un determinado criterio. Por ejemplo, bajo el supuesto de distribución normal, en [25] se propone un estimador fusión distribuido basado en el criterio de máxima verosimilitud, mientras que en [26] se utiliza el método de fusión distribuido lineal ponderado por matrices en el sentido de mínima varianza, que es equivalente al anterior bajo hipótesis de normalidad. Más tarde, el método distribuido se ha utilizado en redes de sensores en las que los mecanismos de transmisión de las observaciones al procesador están sujetos a fallos, tales como observaciones inciertas, retrasos en la recepción y/o pérdidas de medidas en ciertos instantes de tiempo (véanse [27–31], entre otros).

En la mayoría de los trabajos mencionados anteriormente, se consideran ruidos blancos independientes; sin embargo, dicha hipótesis puede ser restrictiva en muchos problemas reales en que los ruidos pueden estar correlados. Esta situación surge, por ejemplo, cuando se utilizan contramedidas electrónicas para impedir que el enemigo identifique sus blancos, por ejemplo, interferencias [32], o cuando los ruidos del sistema dependen del estado, y por tanto, puede haber correlación cruzada entre los ruidos de las ecuaciones de observación y entre éstos con los de la ecuación del estado. También, si todas las observaciones procedentes de los

sensores se realizan en el mismo entorno ruidoso, los ruidos suelen ser correlados. Teniendo en cuenta estos comentarios, el problema de estimación en sistemas con ruidos correlados es un área de investigación importante, por ejemplo, en [33] se aborda el problema de filtrado fusión óptimo de Kalman en sistemas con ruidos con correlación cruzada en el mismo instante de tiempo. En general, cuando los ruidos del sistema están correlados y tienen correlación cruzada en diferentes instantes de tiempo, es difícil obtener estimadores óptimos; esta restricción ha fomentado una investigación más amplia relacionada con los problemas de estimación subóptima de tipo Kalman. En [34] se presenta un filtro recursivo de este tipo para sistemas en que los ruidos del estado están correlados en un número finito de instantes de tiempo, y en [35] se estudia el problema de filtrado en sistemas con ruidos del estado y la observación correlados en múltiples instantes de tiempo. En [36] se estudia el problema de filtrado recursivo de tipo Kalman robusto en sistemas con pérdidas aleatorias de medidas y ruidos autocorrelados en un número finito de instantes de tiempo. El problema de obtener el filtro fusión distribuido de tipo Kalman robusto y ponderado se estudia en [37] para sistemas lineales con ruido multiplicativo dependiente del estado e hipótesis de correlación.

Claramente, sistemas con ruidos multiplicativos en las ecuaciones del estado y/o de la observación, son casos particulares de *sistemas con matrices aleatorias* los cuales tienen aplicaciones prácticas importantes en áreas tales como control digital de procesos químicos, control de radares, sistemas de navegación o sistemas económicos. El problema de estimación en dichos sistemas ha sido ampliamente investigado como puede verse, por ejemplo, en [38] y sus referencias. El problema de filtrado lineal en sistemas con matrices aleatorias independientes, tanto en la ecuación del estado como en la de las observaciones, se ha abordado en [39] y [40], mediante la transformación del sistema original en uno con matrices determinísticas y ruidos dependientes del estado, al que se le ha aplicado el filtro de Kalman. En concreto, en [39] se aplicó el filtro de Kalman sin ninguna justificación teórica,

mientras que en [40] se probó que, bajo ciertas condiciones, el sistema transformado satisface las hipótesis necesarias para la aplicación del filtro de Kalman y, por lo tanto, se obtienen estimadores lineales óptimos. Recientemente, suponiendo medidas escalares y matrices aleatorias en la ecuación de observación, en [41] se aborda el problema de filtrado cuadrático, aplicándole el filtro de Kalman a un sistema aumentado apropiado con matrices determinísticas.

Sin embargo, en diversas situaciones prácticas las matrices aleatorias no son independientes. Esto ocurre, por ejemplo, cuando los sistemas con retrasos aleatorios y/o pérdidas se transforman en modelos de observación con matrices aleatorias [42], o cuando un sistema no lineal se linealiza para aplicar el filtro de Kalman extendido (otras situaciones en las que las matrices del sistema son aleatorias y correladas pueden verse en [43] y [44]). En [45] se consideran sistemas con matrices de transición determinísticas y matrices de observación aleatorias correladas en instantes consecutivos; mediante la transformación de la ecuación de observación en una ecuación con matrices determinísticas y aplicando el filtro de Kalman, se obtiene el estimador óptimo. El problema de estimación lineal en sistemas de redes de sensores con matrices aleatorias y ruidos autocorrelados y con correlación cruzada ha sido abordado en [46] y [47], entre otros.

Los comentarios anteriores constituyen un breve resumen sobre los antecedentes y el estado actual de los diferentes problemas abordados a lo largo de los seis capítulos en los que se ha estructurado esta tesis doctoral. Un análisis más detallado se establece en la introducción de cada uno de ellos. El objetivo general de la tesis es el estudio de problemas de estimación mínimo cuadrática en sistemas lineales en tiempo discreto con observaciones ruidosas procedentes de múltiples sensores, afectadas por parámetros aleatorios que modelizan diferentes tipos de fallo en el mecanismo o la transmisión de las medidas. Cada uno de los capítulos que la integran es un artículo original en el que se aborda un tema específico relacionado con este estudio. Por lo tanto, cada capítulo se presenta como un artículo

independiente, con sus propios abstract, secciones, apéndices y referencias, en la misma forma en la que fueron publicados o sometidos. En el momento de la presentación de esta tesis, los artículos de los capítulos 1-4 han sido publicados y los de los capítulos 5 y 6 han sido sometidos y están bajo revisión.

Esta tesis está organizada de la siguiente forma. En los **Capítulos 1 y 2** se consideran sistemas estocásticos lineales en tiempo discreto con observaciones inciertas procedentes de múltiples sensores cuando, en cada sensor, las variables aleatorias que modelizan el fenómeno de observaciones inciertas están correladas en instantes que se diferencian m unidades de tiempo. Para este tipo de sistemas, en el **Capítulo 1** se abordan los problemas de estimación mediante los métodos de fusión lineal centralizado y distribuido y, en el **Capítulo 2** se lleva a cabo el problema de estimación cuadrática mediante el método centralizado. En el **Capítulo 3** se consideran sistemas con múltiples sensores, ruidos correlados y con correlación cruzada, y fallos en las medidas descritos por sucesiones de variables aleatorias escalares con distribución discreta en el intervalo [0,1]. Bajo dichas condiciones, se establecen algoritmos recursivos de filtrado lineal mediante los métodos de fusión centralizado y distribuido. A continuación, en el **Capítulo 4**, se consideran sistemas con matrices aleatorias independientes en la ecuación del estado y correladas en la ecuación de observación, además de ruidos correlados. En este tipo de sistemas se obtiene el filtro lineal utilizando el método de estimación fusión centralizado y los resultados se aplican a sistemas con múltiples sensores y fallos en las medidas modelizados por variables con distribución discreta en el intervalo [0,1], y también a sistemas con observaciones retrasadas aleatoriamente procedentes de múltiples sensores. En el **Capítulo 5** se estudia el problema de estimación lineal fusión distribuido en sistemas de redes de sensores con matrices aleatorias independientes y ruidos correlados. Por último, el problema de filtrado cuadrático centralizado se aborda en el **Capítulo 6**, en el que se consideran sistemas lineales con matrices aleatorias y ruidos independientes.

Objectives

A basic assumption in classical estimation theory for linear stochastic systems is that the model parameter matrices are known; furthermore, the additive noises and the initial state are assumed to be Gaussian and mutually independent. As is well known, under these conditions, the systems are Gaussian and the Kalman filter provides the conditional expectation of the state given the observations, that is, the optimal LS estimator. However, there exist a considerable number of situations in which the joint distribution of the state and the observations is not Gaussian and the Kalman filter only provides the LS linear estimator. In these cases, the optimal LS estimator is not a linear function of the observations and, generally, cannot be readily obtained; this fact has motivated the search for suboptimal estimators which are computationally easier, such as linear estimators or, more generally, polynomial estimators.

In some real situations where sensor networks are used, the state estimation problem is addressed under the assumption that, at each sampling time, the available measurements always contain information about the current state. However, the unreliability of network characteristics often provokes problems such as the accidental loss of measurements, intermittent failures or random interruptions in the transmission mechanism. These situations, which are characterized by including random parameters in the observation equation, generally give rise to non-Gaussian systems (even if the additive noises are Gaussian). Hence, the Kalman filter does not provide the optimal LS estimator and suboptimal estimators must be sought.

In general terms, the aim of this PhD thesis is to address LS estimation problems in discrete-time linear systems from noisy measurements derived from multiple sensors, affected by random parameters which model different situations of failure in the mechanism or the transmission of the measurements. Specifically, the following systems are considered:

Sensor network systems with missing measurements. These systems are characterized by the fact that the observations contain only partial information about the state or, even, only noise. This kind of failure is modeled by including in the observation equation not only an additive noise, but also a multiplicative noise component. This multiplicative noise is usually described by a sequence of Bernoulli random variables whose values - one or zero - indicate the presence or absence of the state in the corresponding measurement. In cases in which the Bernoulli variables are assumed to be independent, the distribution of the multiplicative noise is determined by the probability of each particular observation containing the state. However, there exist many practical situations where this independence assumption is not realistic; for example, in signal transmission models with stand-by sensors in which any failure in the transmission is detected immediately and the old sensor is then replaced. The phenomenon of missing measurements in these cases has been modeled by Bernoulli variables correlated at consecutive instants, thus covering practical situations in which the state cannot be missing in two successive observations, such as the above-mentioned situation of transmission models with stand-by sensors. For such systems, the linear estimation problem has been studied extensively, and some results for the quadratic estimation problem have been obtained (see, for example [20] and references therein).

As a generalization of the previous situation, it is also possible that the failed sensor may not be replaced immediately but after m instants of time; in such situations, correlation among the Bernoulli random variables modeling the missing measurements at times k and $k+m$ must be considered. Since the estimation

problem had not been examined in this case, our attention is focused on investigating the LS linear and quadratic estimation problems in this kind of systems, considering, for each sensor, different sequences of Bernoulli variables correlated at instants that differ by m units of time.

A general framework for modeling situations in which measurements are missing consists of replacing the Bernoulli variables by random variables, with any discrete distribution on the interval $[0, 1]$, thus covering some practical applications where only partial information is missing. With this in mind, we examine the LS linear estimation problem in multi-sensor systems with missing measurements described in each sensor by this kind of variables. We also assume that the additive noises involved in the system are autocorrelated and cross-correlated, since the independent white noise assumption may be a limitation in many real-world problems in which noise correlation may be present, for example, when the process noise and the sensor measurement noises are dependent on the system state, or when all the sensors are observed in the same noisy environment.

Sensor network systems with random parameter matrices. Such systems constitute an interesting research topic due to the numerous realistic situations in which the state transition and/or the measurement matrices contain random parameters. Evidently, the systems with missing measurements described above are special cases of random measurement matrices. In addition, systems with independent random state transition matrices can be used, for example, to describe randomly variant dynamic systems with multiple models [40], or linear systems with state-dependent multiplicative noise [37]. Discrete-time systems with random state transition and measurement parameter matrices also arise in areas such as the digital control of chemical processes, systems with human operators, economic systems and stochastically-sampled digital control systems [39].

Moreover, as indicated, the fairly conservative assumption that the process and measurement noises are uncorrelated must be weakened in order to cover practical

situations where such noises are usually correlated. Also, in some situations when the estimation problem is studied, systems with random delays and packet dropouts are transformed into systems with correlated noises. Hence, the estimation problem with autocorrelated and cross-correlated noises is considered a challenging research topic (see, e.g. [46] and references therein). Accordingly, our aim is to address the LS linear estimation problem in systems that simultaneously include random parameter matrices and correlated noises. Furthermore, due to the importance of this kind of systems and the significant improvement that quadratic LS estimators may provide over linear estimators, we also study the quadratic LS estimation problem in systems with random parameter matrices in the state equation, as well as in the measurements equation. To date, this problem had only been addressed considering scalar measurements with random observation matrices (see [41]).

In the following, and taking into account the above considerations, the specific objectives of the successive chapters of this PhD thesis are described.

Linear discrete-time stochastic systems with missing measurements obtained from multiple sensors are considered in **Chapters 1 and 2**, assuming different sequences of m-step autocorrelated Bernoulli random variables to model the phenomenon of missing measurements in each of the different sensors. In this kind of systems, our aim is, on the one hand, to address the centralized and distributed fusion linear estimation problems (**Chapter 1**) and, on the other (in order to improve the linear estimators) to consider the centralized fusion quadratic estimation problem (**Chapter 2**).

In **Chapter 3**, we consider multi-sensor systems with missing measurements, using different sequences of scalar random variables with arbitrary discrete probability distribution over the interval $[0, 1]$ in order to model the missing phenomenon in each sensor. In addition, it is assumed that the process noise and all the sensor noises are one-step autocorrelated; that different sensor noises are one-step

cross-correlated; and that the process noise and each sensor noise are two-step cross-correlated. Under these assumptions, our goal is to study the centralized and distributed fusion linear estimation problems.

Next, in **Chapter 4**, we study the centralized fusion linear estimation problem in systems with independent random transition matrices, and one-step correlated and cross-correlated random parameter matrices in the observation equation; it is also assumed that the process and measurement noises are one-step autocorrelated and two-step cross-correlated. These assumptions are imposed in order to apply the results to two significant classes of systems with random failures in the measurements, namely:

- Multi-sensor systems with missing measurements and with correlated and cross-correlated noises, when the missing measurement phenomenon at each sensor is described by different sequences of correlated (at consecutive sampling times) scalar random variables with an arbitrary discrete probability distribution over the interval $[0, 1]$.

- Multi-sensor systems with randomly delayed measurements, with correlated and cross-correlated noises, when the delayed measurement phenomenon at each sensor is described by Bernoulli random variables correlated at consecutive sampling times.

In **Chapter 5**, we discuss the distributed fusion state estimation problem in sensor network systems with random parameter matrices and correlated noises. Unlike the situation described in the previous chapter, in order to obtain the distributed fusion estimators, we now assume independent random parameter matrices. This simplifies the process of obtaining the cross-covariance matrices between any two local estimators.

Finally, in **Chapter 6** we consider discrete-time stochastic systems with both independent random parameter matrices and noises. Our goal in this area is to

address the centralized quadratic filtering problem, which may represent a significant improvement on the linear one. The estimators obtained are applied to multi-sensor systems with state-dependent multiplicative noise and missing measurements described by independent sequences of scalar random variables with probability distribution over the interval $[0, 1]$.

Methodology

As mentioned in the previous section, the overall objective in this PhD thesis is to study the estimation problem in linear systems using noisy measurements from multiple sensors subject to random failures. Specifically, recursive algorithms for both linear and quadratic filtering and fixed-point smoothing problems are obtained under the LS optimality criterion. For this purpose, the orthogonal projection lemma and an innovation approach are used. Furthermore, to amalgamate the information from multiple sensors, both centralized and distributed fusion methods are used.

Accordingly, the methodology used throughout this PhD thesis is related to these issues: the derivation of linear and quadratic recursive estimation algorithms, and multi-sensor information fusion. Finally, the methodology used to illustrate the feasibility of the proposed algorithms is based on computer simulation results. Each of these issues is now described in more detail.

LS linear estimation methodology. To address the linear estimation problem, knowledge is required of the first and second order moments of the different processes involved in the system. The problem at hand is to derive recursive algorithms for the estimator of the system state, x_k, based on the measurements $y_1, \ldots, y_L$, with $L \geq k$, for which the orthogonal projection lemma and an innovation analysis approach are used (see, e.g. [48]).

– *Orthogonal Projection Lemma (OPL).* The LS linear estimator of the state,

x_k, based on the measurements, $\{y_1, \ldots, y_L\}$, $L \geq k$, denoted by $\widehat{x}_{k/L}$, is the only linear combination of $y_1, \ldots, y_L$ that satisfies the orthogonality property

$$E[(x_k - \widehat{x}_{k/L})y_s^T] = 0, \quad s \leq L.$$

– *Innovation approach.* Let $\mathcal{L}^2_{\mathbb{R}^n}(\Omega, \mathcal{A}, \mathcal{P})$ be the space of equivalence classes of n-dimensional random vectors defined on the probability space $(\Omega, \mathcal{A}, \mathcal{P})$, with finite second-order moments. $\mathcal{L}^2_{\mathbb{R}^n}(\Omega, \mathcal{A}, \mathcal{P})$ is a Hilbert space with the scalar product $< Y, W >= E[Y^T W]$.

Because the observations are generally nonorthogonal vectors, we use an innovation approach, which consists of transforming the observation process $\{y_k;\ k \geq 1\}$ into an equivalent process of orthogonal vectors $\{\nu_k;\ k \geq 1\}$; this is equivalent in the sense that each set $\{\nu_1, \ldots, \nu_L\}$ spans the same linear subspace of the space $\mathcal{L}^2_{\mathbb{R}^n}(\Omega, \mathcal{A}, \mathcal{P})$ as $\{y_1, \ldots, y_L\}$; that is,

$$\mathcal{L}(y_1, \ldots, y_L) = \mathcal{L}(\nu_1, \ldots, \nu_L) = \mathcal{L}_L.$$

The innovation process is constructed by the *Gram-Schmidt orthogonalization procedure*, using the following inductive reasoning. Let $\{\nu_1, \ldots, \nu_{k-1}\}$ be the set of orthogonal vectors satisfying $\mathcal{L}(\nu_1, \ldots, \nu_{k-1}) = \mathcal{L}(y_1, \ldots, y_{k-1})$. The next orthogonal vector, ν_k, corresponding to the new observation y_k, is then obtained by projecting y_k onto $\mathcal{L}_{k-1}$; specifically

$$\nu_k = y_k - Proj\{y_k \text{ onto } \mathcal{L}_{k-1}\},$$

and, because of the orthogonality of $\{\nu_1, \ldots, \nu_{k-1}\}$ the above projection can be found by projecting y_k along each of the previously found orthogonal vectors ν_i, for $i \leq k-1$,

$$Proj\{y_k \text{ onto } \mathcal{L}_{k-1}\} = \sum_{i=1}^{k-1} Proj\{y_k \text{ along } \nu_i\} = \sum_{i=1}^{k-1} E[y_k \nu_i^T] \left(E[\nu_i \nu_i^T]\right)^{-1} \nu_i.$$

Since the projection of y_k onto $\mathcal{L}_{k-1}$ is $\widehat{y}_{k/k-1}$, the one-stage LS linear predictor of y_k, we have

$$\widehat{y}_{k/k-1} = \sum_{i=1}^{k-1} E[y_k \nu_i^T] \left(E[\nu_i \nu_i^T]\right)^{-1} \nu_i, \quad k \geq 2.$$

Consequently, by starting with $\nu_1 = y_1 - E[y_1]$, for $k \geq 2$, the orthogonal vectors ν_k are given by $\nu_k = y_k - \widehat{y}_{k/k-1}$. Hence, ν_k can be considered as the "new information" or the "innovation" in y_k given $\{y_1, \ldots, y_{k-1}\}$.

In summary, the observation process $\{y_k;\ k \geq 1\}$ has been transformed into an equivalent white noise $\{\nu_k;\ k \geq 1\}$ known as the innovation process. Taking into account that both processes satisfy

$$\nu_i \in \mathcal{L}(y_1, \ldots, y_i) \quad \text{and} \quad y_i \in \mathcal{L}(\nu_1, \ldots, \nu_i), \quad \forall i \geq 1,$$

it is concluded that such processes are related to each other by a causal and causally invertible linear transformation, and therefore the innovation process is uniquely determined by the observations and reciprocally.

From this consideration it can be stated that the LS linear estimator, $\widehat{x}_{k/L}$, of the state, x_k, based on the observations until the instant L, is equal to the LS linear estimator based on the innovations $\nu_1, \ldots, \nu_L$:

$$\widehat{x}_{k/L} = \sum_{i=1}^{L} h_{k,i} \nu_i, \quad k \geq 1,$$

where the impulse-response function $h_{k,i}$, $i = 1, \ldots, L$, is calculated from the OPL,

$$E[(x_k - \widehat{x}_{k/L}) \nu_s^T] = 0, \quad s \leq L,$$

which leads to the Wiener-Hopf equation:

$$E[x_k \nu_s^T] = \sum_{i=1}^{L} h_{k,i} E[\nu_i \nu_s^T], \quad s \leq L.$$

Due to the whiteness of the innovation process, $E[\nu_i \nu_s^T] = 0$ for $i \neq s$ and the Wiener-Hopf equation is expressed as

$$E[x_k \nu_s^T] = h_{k,s} E[\nu_s \nu_s^T], \quad s \leq L,$$

consequently,

$$h_{k,s} = E[x_k \nu_s^T] \left(E[\nu_s \nu_s^T]\right)^{-1}, \quad s \leq L$$

and thus the following general expression for the LS linear estimator of the state is obtained

$$\widehat{x}_{k/L} = \sum_{i=1}^{L} E[x_k \nu_i^T] \left(E[\nu_i \nu_i^T]\right)^{-1} \nu_i, \quad k \geq 1.$$

From this general expression, the following recursive relation is immediately clear:

$$\widehat{x}_{k/L} = \widehat{x}_{k/L-1} + E[x_k \nu_L^T] \left(E[\nu_L \nu_L^T]\right)^{-1} \nu_L, \quad L \geq k,$$

and this recursive relation of the estimators provides the starting point to derive the recursive linear filtering ($L = k$) and fixed-point smoothing ($L > k$) algorithms.

LS quadratic estimation methodology. To address the quadratic estimation problem, the assumptions about the processes involved in the system must include the knowledge of the fourth-order moments of these processes. The problem is then to obtain the quadratic estimator of x_k based on $\{y_1, \ldots, y_L\}$, $L \geq k$; that is, the LS linear estimator of x_k based on the measurements $y_1, \ldots, y_L$ and their second-order Kronecker powers $y_1^{[2]}, \ldots, y_L^{[2]}$. This quadratic estimation problem is reformulated as a linear estimation problem defining augmented state and observation vectors by stacking the original vectors with their second-order Kronecker powers:

$$\mathcal{X}_k = \begin{pmatrix} x_k \\ x_k^{[2]} \end{pmatrix}, \qquad \mathcal{Y}_k = \begin{pmatrix} y_k \\ y_k^{[2]} \end{pmatrix}.$$

Since the space of linear transformations of $\mathcal{Y}_1, \ldots, \mathcal{Y}_L$, $L \geq k$, is equal to the space of linear transformations of $y_1, \ldots, y_L$ and $y_1^{[2]}, \ldots, y_L^{[2]}$, the LS quadratic estimator, $\widehat{x}_{k/L}^q$, is the LS linear estimator of x_k based on $\mathcal{Y}_1, \ldots, \mathcal{Y}_L$, which is obtained by extracting the first n entries of the LS linear estimator of $\mathcal{X}_k$ based on $\mathcal{Y}_1, \ldots, \mathcal{Y}_L$. Therefore, the quadratic estimation problem for the original state is reduced to the linear estimation problem for the augmented state.

In order to address the LS linear estimation problem of the augmented state based on the augmented measurements, the evolution of the vectors $\mathcal{X}_k$ and $\mathcal{Y}_k$ is analyzed taking into account the Kronecker product properties in the evolution of $x_k^{[2]}$ and $y_k^{[2]}$. For simplicity, since the additive noises of the new model are non-zero mean vectors, the augmented state and measurement equations are rewritten in terms of the centered augmented vectors, $X_k = \mathcal{X}_k - E[\mathcal{X}_k]$ and $Y_k = \mathcal{Y}_k - E[\mathcal{Y}_k]$. A new system, called the augmented system, is then obtained, and the second-order statistical properties of its initial state and the noise processes involved are established to obtain the LS linear estimator of X_k based on $Y_1, \ldots, Y_L$.

Finally, we find that this linear estimator provides the LS linear estimator of $\mathcal{X}_k$ based on $\mathcal{Y}_1, \ldots, \mathcal{Y}_L$, adding the mean vector $E[\mathcal{X}_k]$. Hence, the required quadratic filter $\widehat{x}_{k/L}^q$ is obtained by adding the mean $E[x_k]$ to the vector constituted of the first n entries of the LS linear estimator of X_k.

Multi-sensor information fusion methodology. As indicated previously, to address the state estimation problem, the observations are assumed to be transmitted by multiple sensors whose statistical properties are not necessarily the same for all the sensors. To merge the information coming from multiple sensors, centralized and distributed fusion methods are used.

– *Centralized fusion method.* This method is based on the fact that all the measured data are transmitted from the sensors to a fusion center to be processed. The observations obtained from the multiple sensors are then

merged as a single-sensor measurement and, from this, the optimal LS linear estimators, together with their estimation error covariances, are computed.

– *Distributed fusion method.* This method computes, at each sensor, a local optimal LS linear estimator using its own measurement data, and, subsequently, these estimators are sent to a fusion center where the LS matrix-weighted linear combination of the local estimators is provided. Hence, the distributed fusion estimation problem is addressed in two steps. First, local LS linear estimators of the state, together with their estimation error covariance matrices, are obtained by recursive algorithms. Second, the cross-covariance matrix of the estimation errors between any two local estimators is determined. These covariances, with the estimates and error covariance matrices of all the local subsystems, are then merged to determine the distributed matrix-weighted fusion estimators in the linear minimum variance sense.

Computer simulation methodology. To demonstrate the feasibility of the proposed recursive algorithms and to evaluate the performance of the estimators, different numerical simulation examples are presented. All of the algorithms designed are implemented by MATLAB programs which, at each iteration, simulate the state and the observed values and provide the estimates, together with the corresponding error covariance matrices, as a measure of the accuracy of the estimators.

Numerical and graphical comparisons between the estimation error variances confirm the theoretical results. On the one hand, the quadratic estimators are shown to be more accurate than the linear ones and, on the other, the fixed-point smoothing estimators are more effective than the filtering ones. We also analyze the performance of the centralized and distributed fusion estimators, finding that they both outperform the local ones. Moreover, the centralized and distributed fusion estimators have approximately the same error variances, with a slight inferiority

of the distributed estimator over the centralized one, which is compensated by its greater robustness and increased fault-tolerance abilities. Finally, comparisons with other estimators that have been reported reveal the superior performance of the proposed estimators for the system model considered in each specific situation.

References

[1] C. Chui and G. Chen, *Kalman Filtering: with Real-Time Applications*. Berlín: Springer, 2009.

[2] R. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME. Journal of Basic Engineering*, vol. 82 (Series D), pp. 35–45, 1960.

[3] A. D. Santis, A. Germani, and M. Raimondi, "Optimal quadratic filtering of linear discrete-time non-Gaussian systems," *IEEE Transactions on Automatic Control*, vol. 40, pp. 1274–1278, 1995.

[4] F. Carravetta, A. Germani, and M. Raimondi, "Polynomial filtering for linear discrete time non-Gaussian systems," *SIAM Journal on Control and Optimization*, vol. 34, pp. 1666–1690, 1996.

[5] S. V. Uppala and J. D. Sahr, "On the design of quadratic filters with applications to image processing," *IEEE Transactions on Image Processing*, vol. 6, no. 4, pp. 608–614, 1997.

[6] T. I. Laakso, A. Tarczynski, N. P. Murphy, and V. Välimäki, "Polynomial filtering approach to reconstruction and noise reduction of nonuniformly sampled signals," *Signal Processing*, vol. 80, pp. 567–575, 2000.

[7] M. D. Mora, A. Germani, and A. Nardecchia, "Restoration of images corrupted by additive non-Gaussian noise," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 48, pp. 859–875, 2001.

[8] O. Costa and G. Benites, "Linear minimum mean square filter for discrete-time linear systems with Markov jumps and multiplicative noises," *Automatica*, vol. 47, pp. 466–476, 2011.

[9] P. Fangfang and S. Shuli, "Distributed fusion estimation for multisensor multirate systems with stochastic observation multiplicative noises," *Mathematical Problems in Engineering*, vol. 2014, Article ID 373270, 8 pages, 2014.

[10] N. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Transactions on Information Theory*, vol. 15, no. 4, pp. 457–462, 1969.

[11] R. Monzingo, "Discrete optimal linear smoothing for systems with uncertain observations," *IEEE Transactions on Information Theory*, vol. 21, no. 3, pp. 271–275, 1975.

[12] A. G. Jaffer and S. C. Gupta, "Recursive bayesian estimation with uncertain observation," *IEEE Transactions on Information Theory*, vol. 17, pp. 614–616, 1971.

[13] M. T. Hadidi and S. C. Schwartz, "Linear recursive state estimators under uncertain observations," *IEEE Transactions on Automatic Control*, vol. 24, pp. 944–948, 1979.

[14] R. Monzingo, "Discrete linear recursive smoothing for systems with uncertain observations," *IEEE Transactions on Automatic Control*, vol. 26, no. 3, pp. 754–757, 1981.

[15] R. N. Jackson and D. N. P. Murthy, "Optimal linear estimation with uncertain observations," *IEEE Transactions on Information Theory*, vol. 22, no. 3, pp. 376–378, 1976.

[16] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo, J. Jiménez-López, and J. Linares-Pérez, "Signal polynomial smoothing from correlated interrupted observations based on covariances," *Mathematical Methods in the Applied Sciences*, vol. 30, pp. 1645–1665, 2007.

[17] V. Malyavej, I. Manchester, and A. Savkin, "Precision missile guidance using radar/multiple-video sensor fusion via communication channels with bit-rate constraints," *Automatica*, vol. 42, no. 5, pp. 763–769, 2006.

[18] R. Huang and G. Záruba, "Incorporating data from multiple sensors for localizing nodes in mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 9, pp. 1090–1104, 2007.

[19] F. Hounkpevi and E. Yaz, "Robust minimum variance linear state estimators for multiple sensors with different failure rates," *Automatica*, vol. 43, no. 7, pp. 1274–1280, 2007.

[20] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "Linear and quadratic estimation using uncertain observations from multiple sensors with correlated uncertainty," *Signal Processing*, vol. 91, pp. 330–337, 2011.

[21] F. Hounkpevi and E. Yaz, "Minimum variance generalized state estimators for multiple sensors with different delay rates," *Signal Processing*, vol. 87, pp. 602–613, 2007.

[22] J. Linares-Pérez, A. Hermoso-Carazo, R. Caballero-Águila, and J. Jiménez-López, "Least-squares linear filtering using observations coming from multiple sensors with one- or two-step random delay," *Signal Processing*, vol. 89, pp. 2045–2052, 2009.

[23] J. Ma and S. Sun, "Information fusion estimators for systems with multiple sensors of different packet dropout rates," *Information Fusion*, vol. 12, pp. 213–222, 2011.

[24] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "Least-squares linear estimators using measurements transmitted by different sensors with packet dropouts," *Digital Signal Processing*, vol. 22, pp. 1118–1125, 2012.

[25] K. Kim, "Development of track to track fusion algorithms," in *The American Control Conference, Baltimore, Maryland*, pp. 1037–1041, June 1994.

[26] S. Sun and Z. Deng, "Multi-sensor optimal information fusion Kalman filter," *Automatica*, vol. 40, no. 6, pp. 1017–1023, 2004.

[27] B. Shen, Z. Wang, and Y. Hung, "Distributed H_∞-consensus filtering in sensor networks with multiple missing measurements: The finite-horizon case," *Automatica*, vol. 46, pp. 1682–1688, 2010.

[28] B. Chen, L. Yu, W.-A. Zhang, and A. Liu, "Distributed robust fusion Kalman filtering for uncertain stochastic systems with random delays and missing measurements," in *8th Asian Control Conference (ASCC), Kaohsiung, Taiwan*, pp. 1288–1293, May 2011.

[29] M. Liu, D. Qi, S. Zhang, M. Qiu, and S. Zheng, "Optimal H_∞ fusion filters for a class of discrete-time intelligent systems with time delays and missing measurement," *Neurocomputing*, vol. 74, pp. 3741–3752, 2011.

[30] S. Sun and W. Xiao, "Distributed weighted fusion estimators with random delays and packet dropping," *Circuits Systems Signal Processing*, vol. 26, no. 4, pp. 591–605, 2007.

[31] J. Sun and C. Zhang, "Distributed optimal fusion estimators for multi-sensor systems with bounded random measurement delays and packet dropouts," *Journal of Computational Information Systems*, vol. 8, no. 10, pp. 4087–4094, 2012.

[32] E. Song, Y. Zhu, J. Zhou, and Z. You, "Optimal Kalman filtering fusion with cross-correlated sensor noises," *Automatica*, vol. 43, no. 8, pp. 1450–1456, 2007.

[33] X. Wang, S. Sun, K. Ding, and J. Xue, "Weighted measurement fusion white noise deconvolution filter with correlated noise for multisensor stochastic systems," *Mathematical Problems in Engineering*, vol. 2012, Article ID 257619, 16 pages, 2012.

[34] E. Song, Y. Zhu, and Z. You, "The Kalman type recursive state estimator with a finite-step correlated process noises," in *Proceedings of the IEEE International Conference on Automation and Logistics Qingdao, China*, pp. 196–200, September 2008.

[35] A. Fu, Y. Zhu, and E. Song, "The optimal Kalman type state estimator with multi-step correlated process and measurement noises," in *The 2008 International Conference on Embedded Software and Systems*, pp. 215–220, 2008.

[36] J. Feng, Z. Wang, and M. Zeng, "Optimal robust non-fragile Kalman-type recursive filtering with finite-step autocorrelated noises and multiple packet dropouts," *Aerospace Science and Technology*, vol. 15, no. 6, pp. 486–494, 2011.

[37] J. Feng, Z. Wang, and M. Zeng, "Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises," *Information Fusion*, vol. 14, pp. 78–86, 2013.

[38] J. Hu, Z. Wang, H. Dong, and H. Gao, "Recent advances on recursive filtering and sliding mode design for networked nonlinear stochastic systems: A survey," *Mathematical Problems in Engineering*, vol. 2013, Article ID 646059, 12 pages, 2013.

[39] W. L. D. Koning, "Optimal estimation of linear discrete-time systems with stochastic parameters," *Automatica*, vol. 20, pp. 113–115, 1984.

[40] Y. Luo, Y. Zhu, D. Luo, J. Zhou, E. Song, and D. Wang, "Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications," *Sensors*, vol. 8, no. 12, pp. 8086–8103, 2008.

[41] F. Cacace, A. Fasano, and A. Germani, "Quadratic filtering of non-Gaussian linear systems with random observation matrices," in *The International Federation of Automatic Control, Cape Town, South Africa*, pp. 2370–2375, August 2014.

[42] M. Sahebsara, T. Chen, and S. Shah, "Optimal H_2 filtering with random sensor delay, multiple packet dropout and uncertain observations," *International Journal of Control*, vol. 80, pp. 292–301, 2007.

[43] Y. Luo, Y. Zhu, X. Shen, and E. Song, "Novel data association algorithm based on integrated random coefficient matrices Kalman filtering," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 1, pp. 144–158, 2012.

[44] X. Shen, Y. Luo, Y. Zhu, and E. Song, "Globally optimal distributed Kalman filtering fusion," *Science China. Information Sciences*, vol. 55, no. 3, pp. 512–529, 2012.

[45] X. Shen, Y. Zhu, and Y. Luo, "Optimal state estimation of linear discrete-time systems with correlated random parameter matrices," in *Proceedings of the 30th Chinese Control Conference, Yantai, China*, pp. 1488–1493, July 2011.

[46] J. Hu, Z. Wang, and H. Gao, "Recursive filtering with random parameter matrices, multiple fading measurements and correlated noises," *Automatica*, vol. 49, pp. 3440–3448, 2013.

[47] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "Optimal state estimation for networked systems with random parameter matrices, cor-

related noises and delayed measurements," *International Journal of General Systems*, vol. 44, no. 2, pp. 142–154, 2015.

[48] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*. New Jersey: Prentice Hall, 2000.

Chapter 1

Information fusion algorithms for state estimation in multi-sensor systems with correlated missing measurements

Reference

Caballero-Águila, R., García-Garrido, I. and Linares-Pérez, J. (2014).
Information fusion algorithms for state estimation in multi-sensor systems with correlated missing measurements.
Applied Mathematics and Computation, 226: 548-563.
DOI:10.1016/j.amc.2013.10.068

Abstract

In this paper, centralized and distributed fusion estimation problems in linear discrete-time stochastic systems with missing observations coming from multiple sensors are addressed. At each sensor, the Bernoulli random variables describing the phenomenon of missing observations are assumed to be correlated at instants that differ m units of time. By using an innovation approach, recursive linear

filtering and fixed-point smoothing algorithms for the centralized fusion problem are derived in the least-squares sense. The distributed fusion estimation problem is addressed based on the distributed fusion criterion weighted by matrices in the linear minimum variance sense. For each sensor subsystem, local least-squares linear filtering and fixed-point smoothing estimators are given and the estimation error cross-covariance matrices between any two sensors are derived to obtain the distributed fusion estimators. The performance of the proposed estimators is illustrated by numerical simulation examples where scalar and two-dimensional signals are estimated from missing observations coming from two sensors, and the estimation accuracy is analyzed for different missing probabilities and different values of m.

Contents

1.1 Introduction

During the past decades, the estimation problem in multi-sensor systems has motivated a significant amount of research due to its increasing application in many

engineering fields (for example, in the fields of computer and communication) where sensor networks are used to obtain the whole available information on the system state and its estimation must be carried out from the observations provided by all the sensors (see for example [1] and references therein).

Although the use of sensor networks offers several advantages such as easier installation, simpler maintenance and reduced cost, since the measured data are sent to a processing center via a communication network, the unreliable network characteristics usually leads to other problems such as missing measurements (i.e. measured outputs containing noise only, also called uncertain observations), random communication packet losses and/or delays. These problems may occur in practical applications for many different reasons, such as random failures in the transmission mechanism, accidental loss of some measurements or data inaccessibility at certain times, etc.

The estimation problems in systems with only one or several of the aforementioned uncertainties has attracted considerable research attention, see e.g. [2]-[10] and references therein. To be more specific, the estimation problem in discrete-time nonlinear systems with uncertain observations has been studied in [2], [3]; the estimation problem from measurements subject to random delay which does not exceed one sampling period is addressed in [4], [5]; modifications of conventional linear estimation algorithms for systems with packet dropouts have been proposed in [6], [7]; the optimal linear estimation problem for systems with random delays and packet dropouts has been considered in [8], and for systems including the three sources of uncertainty in [9], [10].

All the above papers consider a single sensor or multiple sensors with the same uncertainty characteristics. However, this is not a realistic assumption in several application fields, for instance, in networked communication systems involving heterogeneous measurement devices, and hence multiple-sensor systems whose statistical properties are not necessarily the same for all the sensors require the

derivation of new estimation algorithms, as the conventional ones cannot be applied directly. A basic matter in systems with multiple sensors is how to fuse the measurement data from the different sensors to address the estimation problem. Mainly two methods are used to process the measured sensor data in estimation problems with sensor networks: centralized and distributed fusion methods.

In the centralized fusion method, all the measured data from sensors are communicated to the fusion center for being processed; specifically, the observations from multiple sensors are stacked as one sensor measurement (with greater dimension) and, hence, it does not require a particular fusion rule. In [11], [12] centralized linear minimum variance estimators are derived considering multiple sensors with different failure rates, and different delay rates are considered in [13], [14]. The optimal centralized problem, also in linear minimum variance sense, is investigated in [15], [16] for systems with multiple sensors of different packet dropout rates.

Nevertheless, as it is known, the centralized approach has several drawbacks due to augmentation, such as poor survivability, reliability, heavy communication and expensive computational cost, and various distributed fusion algorithms have been proposed to reduce these drawbacks. In the distributed fusion method, each sensor estimates the state based on its own measurement data, and then it sends such estimate to the fusion center for fusion according to a certain information fusion criterion. For example, under the assumption of normal distribution, a distributed fusion estimator is proposed in [17] based in maximum likelihood criterion, and the distributed fusion criterion weighted by matrices in the linear minimum variance sense is established in [18], which is equivalent to the maximum likelihood fusion criterion under normality assumption.

Recently, more attention has been paid to the distributed fusion estimation in networked systems with unreliable network transmission (see e.g. [19]-[23] and references therein). Distributed fusion estimators for multi-sensor systems with random delays were presented in [19], [20], and for systems with packet dropouts

in [15], [21]. Simultaneous packet delays and dropouts are considered in [22], [23].

Compared with the number of papers about multi-sensor systems with random communication packet delays and/or dropouts, to the best of the authors knowledge, the literature regarding distributed fusion estimation in multi-sensor systems with missing measurements is relatively scarcer, and most existing papers use independent Bernoulli variables to model the missing measurements [24]-[26].

In [12] centralized linear minimum variance estimators are obtained removing the assumption of independence of the Bernoulli variables describing the phenomenon of missing measurements. Specifically, different sequences of Bernoulli random variables correlated at consecutive sampling times are considered to model the uncertainty at each sensor. This form of correlation covers practical situations where the state cannot be missing in two successive observations and hence, transmission models with stand-by sensors, which are immediately substituted when a failure occurs, are appropriately managed with this model. However, the failed sensor may not be replaced immediately but after m instants of time; in such situations, correlation among the random variables modeling the missing measurements at times k and $k+m$ must be considered and new algorithms must be deduced.

In response to the above considerations, this paper deals with the centralized and distributed fusion estimation problems in multi-sensor systems with missing measurements when, at each sensor, the random variables modeling the missing measurements are correlated at instants that differ m units of time. The main contributions can be summarized as follows: i) centralized fusion filtering and fixed-point smoothing algorithms are proposed in multi-sensor systems with correlated missing measurements and the correlation form considered covers certain models in which the state cannot be missing in $m+1$ consecutive observations, thus generalizing the results in [12]; ii) the distributed fusion filtering and fixed-point smoothing problems are addressed in multi-sensor systems with missing measurements.

The paper is organized as follows: in Section 1.2 the problem formulation is described; more specifically, we introduce the linear state transition model perturbed by white noise, and the measurement model affected by additive white noise and multiplicative noise describing the phenomenon of missing measurements. Also, the pertinent assumptions to address the least-squares linear estimation problem are established. In Section 1.3, by using an innovation analysis approach and the orthogonal projection Lemma, recursive algorithms for the centralized fusion filter and fixed-point smoothers are presented (the derivation has been deferred to Appendixes A.1 and A.2). Next, in Section 1.4, the local least-squares linear estimators and the error cross-covariance matrices between any two local estimates are derived, then distributed fusion estimators are obtained based on the optimal fusion criterion weighted by matrices in the linear minimum variance sense. The performance of the proposed estimators is illustrated in Section 1.5 by two numerical simulation examples where local, distributed and centralized fusion estimators are compared. The paper ends with some concluding remarks in Section 1.6.

Notation: The notation used is standard. A^T represents the transpose of A, $\mathbb{R}^n$ denotes the n-dimensional Euclidean space, $\mathbb{R}^{m\times n}$ is the set of all real matrices of dimension $m \times n$, and I and 0 represent the identity matrix and zero matrix of appropriate dimension, respectively. The shorthand $Diag(M_1, \ldots, M_r)$ denotes a block diagonal matrix whose diagonal blocks are the matrices $M_1, \ldots, M_r$. If the dimension of a matrix is not explicitly stated, it is assumed to be compatible for algebraic operations. For time-varying matrices $F_k,\ k \geq 0$, the product $F_{k-1} \cdots F_i$ is denoted by $\mathbb{F}_{k,i}$. The Hadamard product of matrices C and D is denoted by $\circ$ ($[C \circ D]_{ij} = C_{ij}D_{ij}$).

Also, for arbitrary random vectors α and β, the following notation is used throughout the paper: $\mathrm{Cov}[\alpha, \beta] = E\left[(\alpha - E[\alpha])\,(\beta - E[\beta])^T\right]$ and $\mathrm{Cov}[\alpha] = \mathrm{Cov}[\alpha, \alpha]$, where $E[\cdot]$ stands for the mathematical expectation operator. $\widehat{\alpha}$ denotes the estimator of α and $\widetilde{\alpha} = \alpha - \widehat{\alpha}$ the estimation error.

1.2 Problem formulation

The problem at hand is to determine the least-squares (LS) linear filtering and fixed-point smoothing estimators of the state in linear discrete-time stochastic systems with missing measurements coming from multiple sensors. In this section, we present the system model and the hypotheses about the state and noise processes involved.

Consider a class of discrete-time linear stochastic systems with missing measurements coming from r sensors; the phenomenon of missing measurements (that is, observations containing only noise) occurs randomly and, for each sensor, a different sequence of Bernoulli random variables is used to model this phenomenon. Specifically, the following system is considered

$$x_k = F_{k-1}x_{k-1} + w_{k-1}, \quad k \geq 1, \tag{1.1}$$

$$y_k^i = \theta_k^i H_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2, \ldots, r \tag{1.2}$$

where $x_k \in \mathbb{R}^n$ is the state, $y_k^i \in \mathbb{R}$, $i = 1, 2, \ldots, r$, is the measurement collected by sensor i at sampling time k, $\{w_k;\, k \geq 0\}$ and $\{v_k^i;\, k \geq 1\}$, $i = 1, 2, \ldots, r$, are noise sequences, and $\{\theta_k^i;\ k \geq 1\}$, $i = 1, 2, \ldots, r$, are Bernoulli random variables whose values –one or zero– indicate whether the state is present or missing in the corresponding measure. F_k and H_k^i, $i = 1, 2, \ldots, r$, are known time-varying matrices with compatible dimensions, superscript i denotes the i-th sensor, and r is the number of sensors.

As is known, to address the LS linear estimation problem the state and the observations are required to have finite second-order moments; the following assumptions specify the first- and second-order moments required in the study of this problem, as well as the statistical properties assumed about the initial state and noise processes.

Assumption 1. The initial state x_0 is a random vector with $E[x_0] = \overline{x}_0$ and $Cov[x_0] = P_0$.

Assumption 2. The additive noises $\{w_k;\ k \geq 0\}$ and $\{v_k^i;\ k \geq 1\}$, $i = 1, 2, \ldots, r$, are zero-mean white sequences with covariances $Cov[w_k] = Q_k$ and $Cov[v_k^i] = R_k^i$, respectively.

Assumption 3. The multiplicative noises $\{\theta_k^i;\ k \geq 1\}$, $i = 1, 2, \ldots, r$, are sequences of Bernoulli random variables with $P[\theta_k^i = 1] = \overline{\theta}_k^i$. For $i = 1, 2, \ldots, r$, the variables θ_k^i and θ_s^i are independent for $|k - s| \neq 0, m$, and $Cov[\theta_k^i, \theta_s^i] = K_{k,s}^{\theta^i}$ are known for $|k - s| = 0, m$.

Assumption 4. The initial state x_0 and the noise processes $\{w_k;\ k \geq 0\}$, $\{v_k^i;\ k \geq 1\}$ and $\{\theta_k^i;\ k \geq 1\}$, for $i = 1, 2, \ldots, r$, are mutually independent.

Remark 1. Note that, when $\theta_k^i = 1$, which occurs with known probability $\overline{\theta}_k^i$, the state x_k is present in the measure y_k^i coming from the i-th sensor at time k, whereas if $\theta_k^i = 0$ the state is missing in the measured data at time k, which means that such observation only contains additive noise v_k^i with probability $1 - \overline{\theta}_k^i$. To model the phenomenon of missing measurements at each sensor, different sequences of Bernoulli random variables correlated at instants that differ m units of time are considered. This special form of correlation allows us to consider certain class of systems in which the state cannot be missing in $m + 1$ consecutive observations; specifically, sensor networks where sensor failures may happen and a failed sensor is substituted not immediately, but m sampling times after having failed. For instance, consider that, as in Section 1.5, $\theta_k^i = 1 - \gamma_{k+m}^i(1 - \gamma_k^i)$, with $\{\gamma_k^i;\ k \geq 1\}$ sequences of independent Bernoulli random variables. Hence, if $\theta_k^i = 0$, then $\gamma_{k+m}^i = 1$ and $\gamma_k^i = 0$, and consequently $\theta_{k+m}^i = 1$; this fact guarantees that, if the state is missing at time k, the output measurement at time $k + m$ necessarily contains the state. Therefore, there cannot be more than m consecutive measured data consisting of noise only.

Remark 2. From Assumption 3, θ_k^i and θ_s^i are independent for $|k - s| \neq 0, m$, and hence $K_{k,s}^{\theta^i} = 0$ for $|k - s| \neq 0, m$. Also, it is immediate that $K_{k,k}^{\theta^i} = \overline{\theta}_k^i(1 - \overline{\theta}_k^i)$ and

$K^{\theta^i}_{k,k-m} = E[\theta^i_k \theta^i_{k-m}] - \overline{\theta}^i_k \overline{\theta}^i_{k-m}$.

Remark 3. From Assumption 4, the Bernoulli sequences as well as the observation noises are independent from sensor to sensor. This condition is not necessary to deduce the centralized estimators and clearly it is not involved in the derivation of the local estimators. Such condition is just used to obtain the cross-covariance matrices of the local estimation errors, which are necessary to determine the matrix weights of the distributed fusion estimators.

Our aim is to solve the LS estimation problem of the state x_k based on the received measurements $\left\{y^i_1, y^i_2, \ldots, y^i_k, \ldots, y^i_{k+N}\right\}$, $N \geq 0$, $i = 1, 2, \ldots, r$, by using centralized and distributed fusion methods to process the measured sensor data. More specifically, our aim can be stated as follows:

(i) *Centralized fusion estimation problem.* Consider that all measurement data coming from r sensors are transmitted to a fusion center for being processed, and our aim is to obtain the LS linear filter, $\widehat{x}_{k/k}$, and fixed-point smoother, $\widehat{x}_{k/k+N}$, $N \geq 1$, by recursive algorithms.

(ii) *Distributed fusion estimation problem.* Firstly, recursive algorithms to obtain local LS linear filters, $\widehat{x}^i_{k/k}$, and fixed-point smoothers, $\widehat{x}^i_{k/k+N}$, $N \geq 1$, for $i = 1, 2, \ldots, r$, are derived. Secondly, distributed matrix-weighted fusion estimators $\widehat{x}^0_{k/k+N}$, $N \geq 0$, are established by applying the optimal information fusion criterion weighted by matrices in the linear minimum variance sense [18].

Remark 4. In both cases, recursive algorithms for the LS linear estimators will be established using an innovation approach and the orthogonal projection Lemma (OPL). Since the observations are generally nonorthogonal vectors, through the *Gram-Schmidt orthogonalization procedure*, the set of observations is transformed into an equivalent set of orthogonal vectors, *innovations*, defined as the differences

between each observation and the one-stage observation predictor. The fact that the innovation process is uniquely determined by the observations allows us to state that the LS linear estimator of the state based on the observations is equal to the LS linear estimator of the state based on the innovations. The advantage of this approach comes from the fact that the innovations constitute a white process, and the expression of the estimators as linear combination of the innovations provide the starting point to derive the recursive filtering and fixed-point smoothing algorithms.

1.3 Centralized fusion estimation

In this section, our aim is to obtain the optimal (under the LS criterion) linear estimator by the centralized fusion method, in which all the measurement data coming from r sensors are transmitted to a central site for being processed.

For this purpose, denoting $y_k = (y_k^1, \ldots, y_k^r)^T$, $v_k = (v_k^1, \ldots, v_k^r)^T$, $H_k = (H_k^{1T}, \ldots, H_k^{rT})^T$ and $\Theta_k = Diag(\theta_k^1, \ldots, \theta_k^r)$, equation (1.2) is equivalent to the following stacked measurement equation

$$y_k = \Theta_k H_k x_k + v_k, \quad k \geq 1. \tag{1.3}$$

Remark 5. The following properties of the noises in (1.3) are easily inferred from the model assumptions stated in the previous Section 1.2:

- The additive noise $\{v_k;\, k \geq 1\}$ is a zero-mean white process with covariance matrix $R_k = Diag(R_k^1, \ldots, R_k^r)$, $\forall k \geq 1$.

- The random matrices $\{\Theta_k;\, k \geq 1\}$ satisfy

$$\begin{aligned} &E[\Theta_k] = \overline{\Theta}_k = Diag(\overline{\theta}_k^1, \ldots, \overline{\theta}_k^r), \\ &E[(\Theta_k - \overline{\Theta}_k)^2] = \overline{\Theta}_k(I - \overline{\Theta}_k), \\ &E[(\Theta_k - \overline{\Theta}_k)(\Theta_{k-m} - \overline{\Theta}_{k-m})] = Diag(K_{k,k-m}^{\theta^1}, \ldots, K_{k,k-m}^{\theta^r}). \end{aligned}$$

- The initial state x_0 and the noise processes $\{w_k;\ k \geq 0\}$, $\{v_k;\ k \geq 1\}$ and $\{\Theta_k;\ k \geq 1\}$ are mutually independent.

Remark 6. By denoting $\theta_k = (\theta_k^1, \ldots, \theta_k^r)^T$, it is clear that $Cov[\theta_k] = K_{k,k}^{\theta} = \overline{\Theta}_k(I - \overline{\Theta}_k)$ and $Cov[\theta_k, \theta_{k-m}] = K_{k,k-m}^{\theta} = Diag(K_{k,k-m}^{\theta^1}, \ldots, K_{k,k-m}^{\theta^r})$. Moreover, for any random matrix G independent of $\{\Theta_k;\ k \geq 1\}$, using the Hadamard product, it is easily deduced [11] that $E[\Theta_k G \Theta_s] = E[\theta_k \theta_s^T] \circ E[G]$. Particularly, the next property (which will be needed later) is immediately clear

$$E[(\Theta_k - \overline{\Theta}_k)G(\Theta_{k-m} - \overline{\Theta}_{k-m})] = K_{k,k-m}^{\theta} \circ E[G]. \tag{1.4}$$

In the following theorems, using an innovation approach and the OPL, recursive algorithms for the linear filter, $\widehat{x}_{k/k}$, (Theorem 1.3.1) and the fixed-point smoother, $\widehat{x}_{k/k+N}$, for fixed k and $N \geq 1$, (Theorem 1.3.2) are derived.

Theorem 1.3.1 *For the system model* (1.1) *and measurement model* (1.3)*, under Assumptions 1-4, the LS linear filter* $\widehat{x}_{k/k}$ *is obtained as*

$$\widehat{x}_{k/k} = \widehat{x}_{k/k-1} + S_{k,k}\Pi_k^{-1}\nu_k, \quad k \geq 1; \quad \widehat{x}_{0/0} = \overline{x}_0, \tag{1.5}$$

where the state predictor, $\widehat{x}_{k/k-1}$*, is given by*

$$\widehat{x}_{k/k-1} = F_{k-1}\widehat{x}_{k-1/k-1}, \quad k \geq 1. \tag{1.6}$$

The innovation, ν_k*, satisfies*

$$\begin{aligned} &\nu_k = y_k - \overline{\Theta}_k H_k \widehat{x}_{k/k-1}, \quad k \leq m, \\ &\nu_k = y_k - \overline{\Theta}_k H_k \widehat{x}_{k/k-1} + \Psi_{k,k-m}\left(\nu_{k-m} - \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1}\nu_{k-i}\right), \quad k > m, \end{aligned} \tag{1.7}$$

where $\Psi_{k,k-m} = K_{k,k-m}^{\theta} \circ \left(H_k \mathbb{F}_{k,k-m} D_{k-m} H_{k-m}^T\right) \Pi_{k-m}^{-1}$*, with* $D_k = E[x_k x_k^T]$ *recursively calculated from*

$$D_k = F_{k-1}D_{k-1}F_{k-1}^T + Q_{k-1}, \quad k \geq 1; \qquad D_0 = P_0 + \overline{x}_0\overline{x}_0^T. \tag{1.8}$$

The matrices $T_{k,k-i}$ *are determined by*

$$\begin{aligned} &T_{k,k-i} = \overline{\Theta}_k H_k \mathbb{F}_{k,k-i} S_{k-i,k-i}, \quad 2 \leq k \leq m, \quad 1 \leq i \leq k-1, \\ &T_{k,k-i} = \overline{\Theta}_k H_k \mathbb{F}_{k,k-i} S_{k-i,k-i} - \Psi_{k,k-m} T_{k-i,k-m}^T, \quad k > m, \quad 1 \leq i \leq m-1. \end{aligned} \tag{1.9}$$

The matrix $S_{k,k}$ is calculated by

$$\begin{aligned}
&S_{k,k} = P_{k/k-1} H_k^T \overline{\Theta}_k, \quad k \leq m, \\
&S_{k,k} = P_{k/k-1} H_k^T \overline{\Theta}_k \\
&\qquad -\left(\mathbb{F}_{k,k-m} S_{k-m,k-m} - \sum_{i=1}^{m-1} \mathbb{F}_{k,k-i} S_{k-i,k-i} \Pi_{k-i}^{-1} T_{k-i,k-m} \right) \Psi_{k,k-m}^T, \; k > m,
\end{aligned} \tag{1.10}$$

where $P_{k/k-1}$, the prediction error covariance matrix, is obtained by

$$P_{k/k-1} = F_{k-1} P_{k-1/k-1} F_{k-1}^T + Q_{k-1}, \quad k \geq 1,$$

with $P_{k/k}$, the filtering error covariance matrix, given by

$$P_{k/k} = P_{k/k-1} - S_{k,k} \Pi_k^{-1} S_{k,k}^T, \quad k \geq 1; \qquad P_{0/0} = P_0.$$

The innovation covariance matrix, Π_k, satisfies

$$\begin{aligned}
&\Pi_k = K_{k,k}^{\theta} \circ \left(H_k D_k H_k^T \right) + R_k + \overline{\Theta}_k H_k S_{k,k}, \quad k \leq m, \\
&\Pi_k = K_{k,k}^{\theta} \circ \left(H_k D_k H_k^T \right) + R_k + \overline{\Theta}_k H_k S_{k,k} + S_{k,k}^T H_k^T \overline{\Theta}_k - \overline{\Theta}_k H_k P_{k/k-1} H_k^T \overline{\Theta}_k \\
&\qquad - \Psi_{k,k-m} \left(\Pi_{k-m} + \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} T_{k-i,k-m} \right) \Psi_{k,k-m}^T, \quad k > m.
\end{aligned} \tag{1.11}$$

Proof. See Appendix A.1 □

Theorem 1.3.2 *For the system model* (1.1) *and measurement model* (1.3)*, under Assumptions 1-4, the fixed-point smoothers, $\widehat{x}_{k/k+N}$, $N \geq 1$ are recursively obtained by*

$$\widehat{x}_{k/k+N} = \widehat{x}_{k/k+N-1} + S_{k,k+N} \Pi_{k+N}^{-1} \nu_{k+N}, \quad N \geq 1, \tag{1.12}$$

whose initial condition is the filter, $\widehat{x}_{k/k}$, given in Theorem 1.3.1.

The matrices $S_{k,k+N}$ are determined by

$$\begin{aligned} S_{k,k+N} &= \left(D_k \mathbb{F}_{k+N,k}^T - M_{k,k+N-1} F_{k+N-1}^T\right) H_{k+N}^T \overline{\Theta}_{k+N}, \quad k \leq m-N, \\ S_{k,k+N} &= \left(D_k \mathbb{F}_{k+N,k}^T - M_{k,k+N-1} F_{k+N-1}^T\right) H_{k+N}^T \overline{\Theta}_{k+N} \\ &\quad - \left(S_{k,k+N-m} - \sum_{i=1}^{m-1} S_{k,k+N-i} \Pi_{k+N-i}^{-1} T_{k+N-i,k+N-m}\right) \\ &\quad \times \Psi_{k+N,k+N-m}^T, \quad k > m-N. \end{aligned} \tag{1.13}$$

where the matrices $M_{k,k+N}$ are recursively obtained from

$$\begin{aligned} M_{k,k+N} &= M_{k,k+N-1} F_{k+N-1}^T + S_{k,k+N} \Pi_{k+N}^{-1} S_{k+N,k+N}^T, \\ M_{k,k} &= D_k - P_{k/k}. \end{aligned} \tag{1.14}$$

The fixed-point smoothing error covariance matrix, $P_{k/k+N}$, satisfies

$$P_{k/k+N} = P_{k/k+N-1} - S_{k,k+N} \Pi_{k+N}^{-1} S_{k,k+N}^T, \quad N \geq 1, \tag{1.15}$$

with initial condition $P_{k/k}$, the filtering error covariance matrix.
The innovations ν_{k+N}, their covariance matrices Π_{k+N}, the matrices $T_{k+N,k+N-i}$, $\Psi_{k+N,k+N-m}$, D_k and $P_{k/k}$ are given in Theorem 1.3.1.

Proof. See Appendix A.2. □

Remark 7. As indicated in Remark 3, the assumption that the Bernoulli sequences and the observation noises are independent from sensor to sensor is not required to obtain the centralized estimators. If this assumption is suppressed, one should take into account that, in Theorem 1.3.1, the covariance matrices $K_{k,k}^{\theta}$, $K_{k,k-m}^{\theta}$ and R_k would not be necessarily diagonal, and clearly $K_{k,k}^{\theta} \neq \overline{\Theta}_k(I - \overline{\Theta}_k)$.

1.4 Distributed fusion estimation

Our aim in this section is to find optimal distributed fusion estimators, in the linear minimum variance sense, based on the information provided by local LS linear estimators.

This estimation problem is tackled in two-stage fusion structure. In the first fusion stage, each sensor provides its local estimator based on its own measurement data along with their estimation error covariance matrices. In the second fusion stage, the cross-covariance matrix of the estimation errors between any two sensors from the first fusion stage are determined, and then, these covariances along with the estimates and error covariance matrices of all local subsystems are fused to determine the optimal matrix weights and the optimal fusion estimators in the linear minimum variance sense.

1.4.1 Local LS linear estimators

This section is concerned with the problem of obtaining, for each sensor subsystem of system (1.1) and (1.2), the local LS linear filter, $\widehat{x}^i_{k/k}$, and fixed-point smoothers, $\widehat{x}^i_{k/k+N}$, $N \geq 1$, along with their corresponding error covariance matrices from recursive algorithms. By using an innovation approach, these algorithms are established in the following theorems.

Theorem 1.4.1 *For the i-th sensor subsystem of system* (1.1) *and* (1.2) *under Assumptions 1-4, the local LS linear filter,* $\widehat{x}^i_{k/k}$*, is calculated by*

$$\widehat{x}^i_{k/k} = \widehat{x}^i_{k/k-1} + S^i_{k,k} \left(\Pi^{ii}_{k,k}\right)^{-1} \nu^i_k, \quad k \geq 1, \quad \widehat{x}^i_{0/0} = \overline{x}_0, \tag{1.16}$$

where the local state predictor, $\widehat{x}^i_{k/k-1}$*, satisfies*

$$\widehat{x}^i_{k/k-1} = F_{k-1} \widehat{x}^i_{k-1/k-1}, \quad k \geq 1. \tag{1.17}$$

The innovation, ν^i_k*, is given by*

$$\begin{aligned} \nu^i_k &= y^i_k - \overline{\theta}^i_k H^i_k \widehat{x}^i_{k/k-1}, \quad k \leq m, \\ \nu^i_k &= y^i_k - \overline{\theta}^i_k H^i_k \widehat{x}^i_{k/k-1} \\ &\quad - \Psi^i_{k,k-m} \left(\nu^i_{k-m} - \sum_{l=1}^{m-1} T^{iT}_{k-l,k-m} \left(\Pi^{ii}_{k-l,k-l}\right)^{-1} \nu^i_{k-l} \right), \ k > m, \end{aligned} \tag{1.18}$$

where $\Psi^i_{k,k-m} = K^{\theta^i}_{k,k-m} H^i_k \mathbb{F}_{k,k-m} D_{k-m} H^{iT}_{k-m} \left(\Pi^{ii}_{k-m,k-m}\right)^{-1}$, *with* D_{k-m} *given in Theorem 1.3.1.*

The matrices $T^i_{k,k-l}$ *are determined by*

$$T^i_{k,k-l} = \overline{\theta}^i_k H^i_k \mathbb{F}_{k,k-l} S^i_{k-l,k-l}, \quad 2 \leq k \leq m, \quad 1 \leq l \leq k-1,$$
$$T^i_{k,k-l} = \overline{\theta}^i_k H^i_k \mathbb{F}_{k,k-l} S^i_{k-l,k-l} - \Psi^i_{k,k-m} T^{iT}_{k-l,k-m}, \quad k > m, \quad 1 \leq l \leq m-1.$$

The innovation covariance matrix, $\Pi^{ii}_{k,k}$, *satisfies*

$$\Pi^{ii}_{k,k} = \overline{\theta}^i_k \left(1 - \overline{\theta}^i_k\right) H^i_k D_k H^{iT}_k + R^i_k + \overline{\theta}^i_k H^i_k S^i_{k,k}, \quad k \leq m,$$
$$\begin{aligned}\Pi^{ii}_{k,k} = {} & \overline{\theta}^i_k \left(1 - \overline{\theta}^i_k\right) H^i_k D_k H^{iT}_k + R^i_k + \overline{\theta}^i_k H^i_k S^i_{k,k} + \overline{\theta}^i_k S^{iT}_{k,k} H^{iT}_k \\ & - (\overline{\theta}^i_k)^2 H^i_k P^{ii}_{k/k-1} H^{iT}_k - \Psi^i_{k,k-m} \Bigg(\Pi^{ii}_{k-m,k-m} \\ & + \sum_{l=1}^{m-1} T^{iT}_{k-l,k-m} \left(\Pi^{ii}_{k-l,k-l}\right)^{-1} T^i_{k-l,k-m} \Bigg) \Psi^{iT}_{k,k-m}, \ k > m.\end{aligned}$$

The matrix $S^i_{k,k}$ *is derived by the following expression*

$$S^i_{k,k} = \overline{\theta}^i_k P^{ii}_{k/k-1} H^{iT}_k, \quad k \leq m,$$
$$\begin{aligned}S^i_{k,k} = {} & \overline{\theta}^i_k P^{ii}_{k/k-1} H^{iT}_k - \Bigg(\mathbb{F}_{k,k-m} S^i_{k-m,k-m} \\ & - \sum_{l=1}^{m-1} \mathbb{F}_{k,k-l} S^i_{k-l,k-l} \left(\Pi^{ii}_{k-l,k-l}\right)^{-1} T^i_{k-l,k-m} \Bigg) \Psi^{iT}_{k,k-m}, \quad k > m,\end{aligned}$$

where $P^{ii}_{k/k-1}$, *the prediction error covariance matrix, is obtained by*

$$P^{ii}_{k/k-1} = F_{k-1} P^{ii}_{k-1/k-1} F^T_{k-1} + Q_{k-1}, \quad k \geq 1,$$

with $P^{ii}_{k/k}$, *the filtering error covariance matrix, satisfying*

$$P^{ii}_{k/k} = P^{ii}_{k/k-1} - S^i_{k,k} \left(\Pi^{ii}_{k,k}\right)^{-1} S^{iT}_{k,k}, \quad k \geq 1, \quad P^{ii}_{0/0} = P_0.$$

Proof. This proof is analogous to that of Theorem 1.3.1 and hence it is omitted. □

Theorem 1.4.2 *For the i-th sensor subsystem of system* (1.1) *and* (1.2) *under Assumptions 1-4, the local LS linear fixed-point smoothers* $\widehat{x}^i_{k/k+N}$, $N \geq 1$, *are recursively calculated by*

$$\widehat{x}^i_{k/k+N} = \widehat{x}^i_{k/k+N-1} + S^i_{k,k+N} \left(\Pi^{ii}_{k+N,k+N}\right)^{-1} \nu^i_{k+N}, \quad N \geq 1, \tag{1.19}$$

whose initial condition is the local filter, $\widehat{x}^i_{k/k}$, *given in Theorem 1.4.1.*

The matrices $S^i_{k,k+N}$ *satisfy the following expressions*

$$\begin{aligned}
S^i_{k,k+N} &= \overline{\theta}^i_{k+N} \left(D_k \mathbb{F}^T_{k+N,k} - M^i_{k,k+N-1} F^T_{k+N-1}\right) H^{iT}_{k+N}, \quad k \leq m-N, \\
S^i_{k,k+N} &= \overline{\theta}^i_{k+N} \left(D_k \mathbb{F}^T_{k+N,k} - M^i_{k,k+N-1} F^T_{k+N-1}\right) H^{iT}_{k+N} \\
&\quad - \left(S^i_{k,k+N-m} - \sum_{l=1}^{m-1} S^i_{k,k+N-l} \left(\Pi^{ii}_{k+N-l,k+N-l}\right)^{-1} T^i_{k+N-l,k+N-m}\right) \\
&\quad \times \Psi^{iT}_{k+N,k+N-m}, \quad k > m-N,
\end{aligned}$$

where the matrices $M^i_{k,k+N}$ *are recursively obtained by*

$$\begin{aligned}
M^i_{k,k+N} &= M^i_{k,k+N-1} F^T_{k+N-1} + S^i_{k,k+N} \left(\Pi^{ii}_{k+N,k+N}\right)^{-1} S^{iT}_{k+N,k+N}, \\
M^i_{k,k} &= D_k - P^{ii}_{k/k}.
\end{aligned}$$

The fixed-point smoothing error covariance matrices, $P^{ii}_{k/k+N}$, *are given by*

$$P^{ii}_{k/k+N} = P^{ii}_{k/k+N-1} - S^i_{k,k+N} \left(\Pi^{ii}_{k+N,k+N}\right)^{-1} S^{iT}_{k,k+N}, \quad N \geq 1,$$

with initial condition the filtering error covariance matrix, $P^{ii}_{k/k}$.

The innovations ν^i_{k+N}, *their covariance matrices* $\Pi^{ii}_{k+N,k+N}$, *the matrices* $T^i_{k+N,k+N-l}$, $\Psi^i_{k+N,k+N-m}$, D_k *and* $P^{ii}_{k/k}$ *are given in Theorem 1.4.1.*

Proof. This proof is analogous to that of Theorem 1.3.2 and hence it is omitted. □

1.4.2 Distributed fusion estimators

Once the local LS linear filtering and fixed-point smoothing estimators given in Theorems 1.4.1 and 1.4.2 are available, we can easily obtain the distributed optimal weighted fusion estimators and their error covariance matrices, by applying the optimal information fusion criterion weighted by matrices in the linear minimum variance sense [18].

Theorem 1.4.3 *For the system model* (1.1) *and measurement model* (1.2)*, under Assumptions 1-4, the distributed optimal fusion filter,* $\widehat{x}^0_{k/k}$*, and fixed-point smoother,* $\widehat{x}^0_{k/k+N}$*,* $N \geq 1$*, are given by*

$$\widehat{x}^0_{k/k+N} = A^1_{k,k+N}\widehat{x}^1_{k/k+N} + \cdots + A^r_{k,k+N}\widehat{x}^r_{k/k+N}, \quad N \geq 0,$$

where $\widehat{x}^i_{k/k+N}$*,* $N \geq 0$ $(i = 1, 2 \ldots, r)$ *are calculated by the recursive algorithms established in Theorems 1.4.1 and 1.4.2.*

The optimal matrix weights $A^i_{k,k+N}$ $(i = 1, 2, \ldots, r)$ *are computed by*

$$A_{k,k+N} = \Sigma^{-1}_{k/k+N} e \left(e^T \Sigma^{-1}_{k/k+N} e \right)^{-1},$$

where the matrices $A_{k,k+N} = \left[A^1_{k,k+N}, \ldots, A^r_{k,k+N}\right]^T$ *and* $e = [I, \ldots, I]^T$ *are both* $nr \times n$ *matrices, and*

$$\Sigma_{k/k+N} = E\left[\left(\widetilde{x}^1_{k/k+N}, \ldots, \widetilde{x}^r_{k/k+N}\right) \left(\widetilde{x}^1_{k/k+N}, \ldots, \widetilde{x}^r_{k/k+N}\right)^T \right] = (P^{ij}_{k/k+N})$$

is a symmetric positive definite matrix of dimension $nr \times nr$*.*

The error covariance matrices of the distributed weighted fusion estimators are computed by

$$P^0_{k/k+N} = \left(e^T \Sigma^{-1}_{k/k+N} e \right)^{-1}, \quad N \geq 0,$$

and the following inequality holds: $P^0_{k/k+N} \leq P^{ii}_{k/k+N}$*,* $i = 1, 2, \ldots, r$*.*

Proof. The proof follows directly from the optimal information criterion weighted by matrices in the linear minimum variance sense [18] and therefore it is omitted. □

To apply the above Theorem 1.4.3, besides the local estimators, $\widehat{x}^i_{k/k+N},\ N \geq 0$ $(i = 1, 2 \ldots, r)$, and their error covariance matrices, $P^{ii}_{k/k+N}$, given in Theorems 1.4.1 and 1.4.2, we need to calculate the cross-covariance matrices, $P^{ij}_{k/k+N}$, between any two subsystems. Next, computation procedures for the cross-covariance matrices, $P^{ij}_{k/k+N},\ N \geq 0,\ i \neq j,\ i, j = 1, 2 \ldots, r$, will be presented, before which some useful lemmas will be given. The assumptions and notation in these lemmas are those of Theorems 1.4.1 and 1.4.2.

Lemma 1.4.1 *For $i \neq j,\ i, j = 1, 2 \ldots, r,\ L^{ij}_{k,k} = E\left[\widehat{x}^i_{k/k-1}\nu^{jT}_k\right]$, is calculated by*

$$\begin{aligned} L^{ij}_{k,k} &= \overline{\theta}^j_k \left(P^{jj}_{k/k-1} - P^{ij}_{k/k-1}\right) H^{jT}_k, \quad k \leq m, \\ L^{ij}_{k,k} &= \overline{\theta}^j_k \left(P^{jj}_{k/k-1} - P^{ij}_{k/k-1}\right) H^{jT}_k - \left(L^{ij}_{k,k-m} - \sum_{l=1}^{m-1} L^{ij}_{k,k-l} (\Pi^{jj}_{k-l,k-l})^{-1} T^j_{k-l,k-m}\right) \\ &\quad \times \Psi^{jT}_{k,k-m}, \quad k > m, \end{aligned} \tag{1.20}$$

where $L^{ij}_{k,s} = E\left[\widehat{x}^i_{k/k-1}\nu^{jT}_s\right],\ s < k$, is recursively obtained by

$$L^{ij}_{k,s} = F_{k-1} L^{ij}_{k-1,s} + F_{k-1} S^i_{k-1,k-1} (\Pi^{ii}_{k-1,k-1})^{-1} \Pi^{ij}_{k-1,s}, \quad s < k. \tag{1.21}$$

Proof. Taking into account expression (1.18) for the innovation ν^j_k, to obtain (1.20) for $L^{ij}_{k,k}$ it is enough to prove that

$$E\left[\widehat{x}^i_{k/k-1} y^{jT}_k\right] - \overline{\theta}^j_k E\left[\widehat{x}^i_{k/k-1} \widehat{x}^{jT}_{k/k-1}\right] H^{jT}_k = \overline{\theta}^j_k \left(P^{jj}_{k/k-1} - P^{ij}_{k/k-1}\right) H^{jT}_k. \tag{1.22}$$

Using (1.2) for y^j_k and the OPL, we have

$$E\left[\widehat{x}^i_{k/k-1} y^{jT}_k\right] = \overline{\theta}^j_k E\left[\widehat{x}^i_{k/k-1} \widehat{x}^{iT}_{k/k-1}\right] H^{jT}_k.$$

Now, since

$$E\left[\widehat{x}^i_{k/k-1}\widehat{x}^{jT}_{k/k-1}\right] = P^{ij}_{k/k-1} - D_k + E\left[\widehat{x}^i_{k/k-1}\widehat{x}^{iT}_{k/k-1}\right] + E\left[\widehat{x}^j_{k/k-1}\widehat{x}^{jT}_{k/k-1}\right],$$
$$E\left[\widehat{x}^j_{k/k-1}\widehat{x}^{jT}_{k/k-1}\right] = D_k - P^{jj}_{k/k-1},$$

we have that $E\left[\widehat{x}^i_{k/k-1}\widehat{x}^{iT}_{k/k-1}\right] - E\left[\widehat{x}^i_{k/k-1}\widehat{x}^{jT}_{k/k-1}\right] = P^{jj}_{k/k-1} - P^{ij}_{k/k-1}$ and (1.22) is easily derived.

Finally, from (1.17) for $\widehat{x}^i_{k/k-1}$ and (1.16) for $\widehat{x}^i_{k-1/k-1}$, expression (1.21) is immediately obtained. □

Lemma 1.4.2 *For $i \neq j$, $i, j = 1, 2 \ldots, r$, the innovation cross-covariance matrix $\Pi^{ij}_{k,s} = E\left[\nu^i_k \nu^{jT}_s\right]$ satisfies*

$$\begin{aligned}
\Pi^{ij}_{k,s} &= \overline{\theta}^i_k H^i_k \left(\mathbb{F}_{k,s} S^j_{s,s} - L^{ij}_{k,s}\right), \quad k \leq m, \quad 1 \leq s \leq k, \\
\Pi^{ij}_{k,s} &= \overline{\theta}^i_k H^i_k \left(\mathbb{F}_{k,s} S^j_{s,s} - L^{ij}_{k,s}\right) - \Psi^i_{k,k-m}\Bigg(\Pi^{ij}_{k-m,s} \\
&\quad - \sum_{l=1}^{m-1} T^{iT}_{k-l,k-m} (\Pi^{ii}_{k-l,k-l})^{-1} \Pi^{ij}_{k-l,s}\Bigg), \quad k > m, \quad k-m \leq s \leq k.
\end{aligned} \tag{1.23}$$

Proof. By using (1.18) for the innovation ν^i_k, and taking into account that, from (1.2),

$$E\left[y^i_k \nu^{jT}_s\right] = \overline{\theta}^i_k H^i_k E\left[x_k \nu^{jT}_s\right] = \overline{\theta}^i_k H^i_k \mathbb{F}_{k,s} S^j_{s,s},$$

expression (1.23) is obtained. □

Lemma 1.4.3 *For $i \neq j$, $i, j = 1, 2 \ldots, r$, $J^{ij}_{k/k+N-1,k+N} = E\left[\widehat{x}^i_{k/k+N-1}\nu^{jT}_{k+N}\right]$ satisfies*

$$\begin{aligned}
J^{ij}_{k/k+N-1,k+N} &= \overline{\theta}^j_{k+N}\left(E^{ii}_{k,k+N-1} - E^{ij}_{k,k+N-1}\right) F^T_{k+N-1} H^{jT}_{k+N}, \quad k \leq m-N, \\
J^{ij}_{k/k+N-1,k+N} &= \overline{\theta}^j_{k+N}\left(E^{ii}_{k,k+N-1} - E^{ij}_{k,k+N-1}\right) F^T_{k+N-1} H^{jT}_{k+N} \\
&\quad - \Bigg(J^{ij}_{k/k+N-1,k+N-m} - \sum_{l=1}^{m-1} J^{ij}_{k/k+N-1,k+N-l} (\Pi^{jj}_{k+N-l,k+N-l})^{-1} \\
&\quad \times T^j_{k+N-l,k+N-m}\Bigg) \Psi^{jT}_{k+N,k+N-m}, \quad k > m-N,
\end{aligned} \tag{1.24}$$

where, for $i,j=1,2\ldots,r$, $E^{ij}_{k,k+N}=E\left[\widehat{x}^i_{k/k+N}\widehat{x}^{jT}_{k+N/k+N}\right]$ *is recursively computed by*

$$\begin{aligned}E^{ij}_{k,k+N}&=E^{ij}_{k,k+N-1}F^T_{k+N-1}+J^{ij}_{k/k+N-1,k+N}(\Pi^{jj}_{k+N,k+N})^{-1}S^{jT}_{k+N,k+N}\\&\quad+S^i_{k,k+N}(\Pi^{ii}_{k+N,k+N})^{-1}(L^{ji}_{k+N,k+N})^T\\&\quad+S^i_{k,k+N}(\Pi^{ii}_{k+N,k+N})^{-1}\Pi^{ij}_{k+N,k+N}(\Pi^{jj}_{k+N,k+N})^{-1}S^{jT}_{k+N,k+N},\quad N\geq 1,\end{aligned}\tag{1.25}$$

with initial condition $E^{ij}_{k,k}=D_k+P^{ij}_{k/k}-P^{ii}_{k/k}-P^{jj}_{k/k}$.
For $l=1,2\ldots,m$, $J^{ij}_{k/k+N-1,k+N-l}=E\left[\widehat{x}^i_{k/k+N-1}\nu^{jT}_{k+N-l}\right]$ *satisfies*

$$J^{ij}_{k/k+N-1,k+N-l}=J^{ij}_{k/k+N-2,k+N-l}+S^i_{k,k+N-1}(\Pi^{ii}_{k+N-1,k+N-1})^{-1}\Pi^{ij}_{k+N-1,k+N-l}.\tag{1.26}$$

Proof. From (1.18) for the innovation ν^j_{k+N}, in order to obtain (1.24) we just need to prove that

$$\begin{aligned}E\left[\widehat{x}^i_{k+N/k+N-1}y^{jT}_{k+N}\right]-\overline{\theta}^j_{k+N}E\left[\widehat{x}^i_{k/k+N-1}\widehat{x}^{jT}_{k+N/k+N-1}\right]H^{jT}_{k+N}\\=\overline{\theta}^j_{k+N}\left(E^{ii}_{k,k+N-1}-E^{ij}_{k,k+N-1}\right)F^T_{k+N-1}H^{jT}_{k+N}.\end{aligned}\tag{1.27}$$

Using (1.2) for y^j_{k+N} and the OPL, we have

$$E\left[\widehat{x}^i_{k/k+N-1}y^{jT}_{k+N}\right]=\overline{\theta}^j_{k+N}E\left[\widehat{x}^i_{k/k+N-1}\widehat{x}^{iT}_{k+N/k+N-1}\right]H^{jT}_{k+N},$$

and since, from (1.17), $\widehat{x}^j_{k+N/k+N-1}=F_{k+N-1}\widehat{x}^j_{k+N-1/k+N-1}$, expression (1.27) is easily obtained.

On the other hand, by using (1.19) for $\widehat{x}^i_{k/k+N}$ and (1.16) for $\widehat{x}^j_{k+N/k+N}$, recursive expression (1.25) for $E^{ij}_{k,k+N}$ is immediately derived; its initial condition $E^{ij}_{k,k}$ is also easily obtained.

Finally, by using again (1.19) for $\widehat{x}^i_{k/k+N-1}$, recursive expression (1.26) is also immediately clear and the proof is completed. □

Remark 8: For $i=j$, since the innovation is a white process, it is clear that $L^{ii}_{k,k}=E\left[\widehat{x}^i_{k/k-1}\nu^{iT}_k\right]=0$ and $J^{ii}_{k/k+N-1,k+N}=E\left[\widehat{x}^i_{k/k+N-1}\nu^{iT}_{k+N}\right]=0$.

A set of recursive formulas to calculate the filtering and fixed-point smoothing error cross-covariance matrices $P^{ij}_{k/k+N}$, $i \neq j$, $i,j = 1,2\ldots,r$, $N \geq 0$, is now derived in the following theorem based on Lemmas 1.4.1-1.4.3.

Theorem 1.4.4 *The cross-covariance matrices, $P^{ij}_{k/k+N}$, $N \geq 1$, of the fixed-point smoothing errors between the i-th and the j-th sensor subsystems are recursively computed by*

$$\begin{aligned} P^{ij}_{k/k+N} &= P^{ij}_{k/k+N-1} + S^{i}_{k,k+N}(\Pi^{ii}_{k+N,k+N})^{-1}\Pi^{ij}_{k+N,k+N}(\Pi^{jj}_{k+N,k+N})^{-1}S^{jT}_{k,k+N} \\ &\quad - \left(S^{j}_{k,k+N} - J^{ij}_{k/k+N-1,k+N}\right)(\Pi^{jj}_{k+N,k+N})^{-1}S^{jT}_{k,k+N} \\ &\quad - S^{i}_{k,k+N}(\Pi^{ii}_{k+N,k+N})^{-1}\left(S^{i}_{k,k+N} - J^{ji}_{k/k+N-1,k+N}\right)^{T},\ N \geq 1. \end{aligned} \tag{1.28}$$

The initial condition, $P^{ij}_{k/k}$, the cross-covariance matrix of the filtering error between the i-th and the j-th sensor subsystems, satisfies

$$\begin{aligned} P^{ij}_{k/k} &= P^{ij}_{k/k-1} + S^{i}_{k,k}(\Pi^{ii}_{k,k})^{-1}\Pi^{ij}_{k,k}(\Pi^{jj}_{k,k})^{-1}S^{jT}_{k,k} \\ &\quad - \left(S^{j}_{k,k} - L^{ij}_{k,k}\right)(\Pi^{jj}_{k,k})^{-1}S^{jT}_{k,k} - S^{i}_{k,k}(\Pi^{ii}_{k,k})^{-1}\left(S^{i}_{k,k} - L^{ji}_{k,k}\right)^{T}, \quad k \geq 1, \\ P^{ij}_{k/k-1} &= F_{k-1}P^{ij}_{k-1/k-1}F^{T}_{k-1} + Q_{k-1}, \quad k \geq 1; \quad P^{ij}_{0/0} = P_0. \end{aligned} \tag{1.29}$$

Proof. By using (1.19) for $\widehat{x}^{i}_{k/k+N}$ and $\widehat{x}^{j}_{k/k+N}$, we have

$$\begin{aligned} P^{ij}_{k/k+N} &= P^{ij}_{k/k+N-1} - E\left[\left(x_k - \widehat{x}^{i}_{k/k+N-1}\right)\nu^{jT}_{k+N}\right](\Pi^{jj}_{k+N,k+N})^{-1}S^{jT}_{k,k+N} \\ &\quad - S^{i}_{k,k+N}(\Pi^{ii}_{k+N,k+N})^{-1}E\left[\nu^{i}_{k+N}\left(x_k - \widehat{x}^{j}_{k+N/k+N-1}\right)^{T}\right] \\ &\quad + S^{i}_{k,k+N}(\Pi^{ii}_{k+N,k+N})^{-1}\Pi^{ij}_{k+N,k+N}(\Pi^{jj}_{k+N,k+N})^{-1}S^{jT}_{k,k+N}. \end{aligned}$$

Taking into account that $E\left[x_k\nu^{jT}_{k+N}\right] = S^{j}_{k,k+N}$, $E\left[\widehat{x}^{i}_{k/k+N-1}\nu^{jT}_{k+N}\right] = J^{ij}_{k/k+N-1,k+N}$, $E\left[\nu^{i}_{k+N}x^{T}_{k}\right] = S^{iT}_{k,k+N}$ and $E\left[\nu^{i}_{k+N}\widehat{x}^{jT}_{k/k+N-1}\right] = J^{jiT}_{k/k+N-1,k+N}$, recursive expression (1.28), for the cross-covariance matrices of the local fixed-point smoothing errors, is obtained.

Finally, by using (1.16), and following an analogous reasoning, it is easy to get (1.29) for the cross-covariance matrices of the local filtering errors. □

1.5 Numerical simulation examples

In this section, two numerical examples are presented to show the effectiveness of the proposed estimation algorithms. To test and compare the performance of the proposed estimators, we ran a program in MATLAB, simulating at each iteration the state and the measured values and providing the centralized and distributed fusion filter and fixed-point smoothers, as well as the corresponding error covariance matrices.

1.5.1 Example 1

In this example, for the simulation, we consider that the system state is given by a scalar process, $\{x_k;\, k \geq 0\}$, generated by the following first-order autoregressive model,

$$x_k = 0.95x_{k-1} + w_{k-1}, \quad k \geq 1,$$

where the initial state is a zero-mean Gaussian variable with $Var[x_0] = 1$ and $\{w_k; k \geq 0\}$ is a zero-mean white Gaussian noise with $Var[w_k] = 0.1$, for all k.

Consider missing measurements coming from two sensors and perturbed by independent sequences of Bernoulli random variables $\{\theta_k^i;\, k \geq 1\}$, $i = 1, 2$, and by independent additive white noises, $\{v_k^i;\, k \geq 1\}$, $i = 1, 2$, with zero-mean and variances $Var[v_k^1] = 1$ and $Var[v_k^2] = 1.5$, for all k.

$$y_k^i = \theta_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2.$$

According to our theoretical model, it is assumed that, for each sensor, the uncertainty at any sampling time $k \geq 1$ depends only on the uncertainty at the previous time $k-m$. The variables θ_k^i modeling this type of uncertainty correlation in the output measurements are defined based on two independent sequences of independent Bernoulli random variables, $\{\gamma_k^i;\, k \geq 1\}$, $i = 1, 2$, with constant probabilities $P[\gamma_k^i = 1] = \gamma_i$. Specifically, the variables θ_k^i are defined as follows

$$\theta_k^i = 1 - \gamma_{k+m}^i(1 - \gamma_k^i), \quad i = 1, 2.$$

Thus, if $\theta_k^i = 0$, then $\gamma_{k+m}^i = 1$ and $\gamma_k^i = 0$, and hence, $\theta_{k+m}^i = 1$; this fact guarantees that, if the state is missing at time k, the output measurement at time $k+m$ necessarily contains the state. Therefore, there cannot be more than m consecutive measured data consisting of noise only.

Since the variables γ_k^i and γ_s^i are independent, θ_k^i and θ_s^i are also independent for $|k-s| \neq 0, m$. The mean of these variables is $\overline{\theta}^i = 1 - \gamma_i(1-\gamma_i)$ and its covariance function is given by

$$K_{k,s}^{\theta} = E[(\theta_k^i - \overline{\theta}^i)(\theta_s^i - \overline{\theta}^i)] = \begin{cases} 0, & |k-s| \neq 0, m, \\ -(1-\overline{\theta}^i)^2, & |k-s| = m, \\ \overline{\theta}^i(1-\overline{\theta}^i), & |k-s| = 0. \end{cases}$$

To compare the effectiveness of the proposed estimators, fifty iterations of the proposed algorithms have been performed and the results obtained for different values of the uncertainty probability and several values of m have been analyzed.

Let us observe that the means, $\overline{\theta}^i$, for $i = 1, 2$, of the variables θ_k^i, are the same if $1 - \gamma_i$ is used instead of γ_i; for this reason, only the case $\gamma_i \leq 0.5$ will be considered here.

Assuming that the Bernoulli variables θ_k^i, for $i = 1, 2$, of the measurement outputs are correlated at sampling times that differ three units of time ($m = 3$), the error variances of local, centralized and distributed fusion filters will be compared considering fixed values of the probabilities γ_1 and γ_2; specifically, $\gamma_1 = 0.1$, $\gamma_2 = 0.2$. In Figure 1.1, as mentioned in Theorem 1.4.3, we can see that the error variances of each local filter are higher than that of the distributed fusion filter. Although the distributed fusion filter has lower accuracy than the centralized one, this difference is slight. Besides, this is compensated by the fact that the distributed fusion structure is in general more robust, reduces the computational cost and improves the reliability due to its parallel structure.

Figure 1.2 displays the filtering and fixed-point smoothing error variances ($N = 2, 5$) for the centralized and distributed fusion methods. It can be seen that the

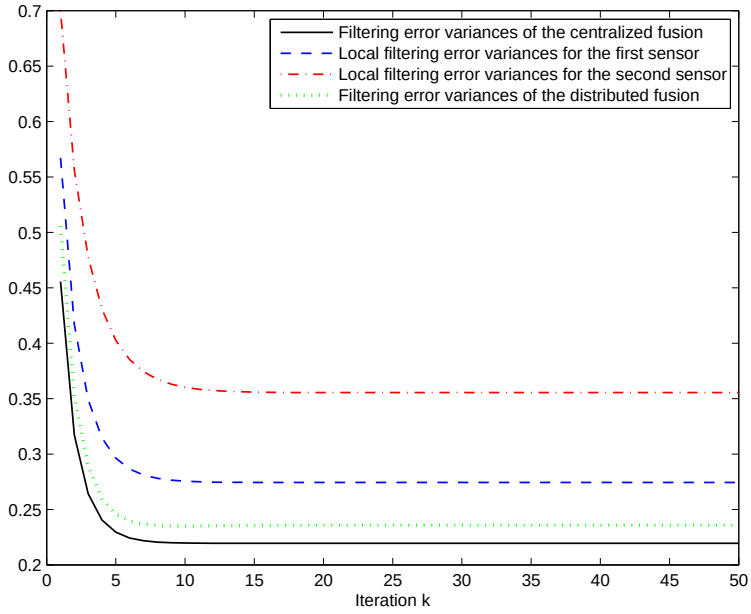


Figure 1.1: Filtering error variances for the centralized and distributed fusion methods for $\gamma_1 = 0.1$, $\gamma_2 = 0.2$, when $m = 3$.

error variances corresponding to the fixed-point smoothers are less than those of the filters and, consequently, the fixed-point smoothing estimates are more accurate. It is also verified that centralized and distributed fusion filter and smoothers have a similar accuracy. If we compare the smoothing error variances at each fixed-point k for $N = 2$ and $N = 5$, we observe that these estimators become more accurate as the number of available observations increases.

Finally, in order to show more precisely the dependence of the error variances on the values γ_1 and γ_2, Figure 1.3 displays the filtering error variances, at a fixed iteration (namely, $k = 50$) for $m = 3$, when both γ_1 and γ_2 are varied from 0.1 to 0.5, which provide different values of the probabilities $\overline{\theta}^1$ and $\overline{\theta}^2$. More specifically, we have considered the values $\gamma_i = 0.1, 0.2, 0.3, 0.4, 0.5$, which lead to the probabilities $\overline{\theta}^i = 0.91, 0.84, 0.78, 0.76, 0.75$, respectively.

In this figure, both graphs (corresponding to the centralized and distributed

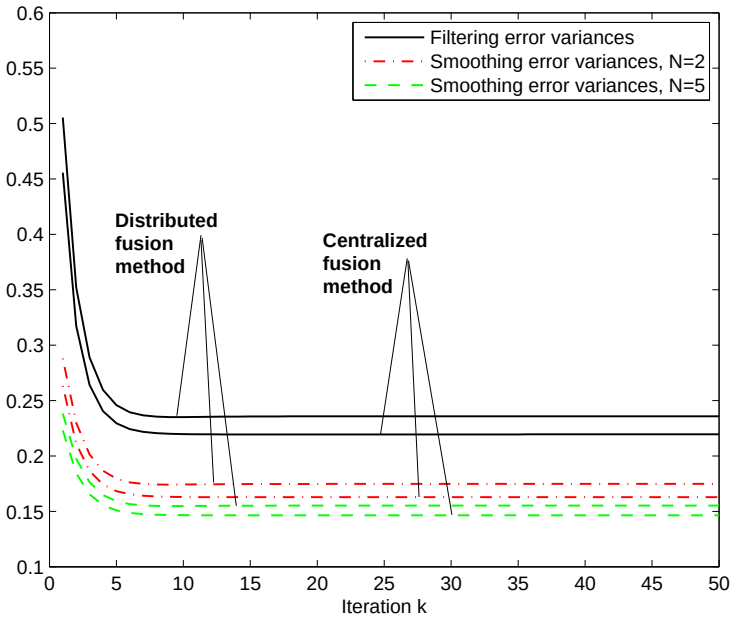


Figure 1.2: Filtering and smoothing error variances for the centralized and distributed fusion methods for $\gamma_1 = 0.1$, $\gamma_2 = 0.2$, when $m = 3$.

fusion filters, respectively) show that the performance of the filter diminishes as $\overline{\theta}^i$ becomes lower, due to the fact that the probability of observations containing the state decreases. Also, this figure confirms that both methods, centralized and distributed, have approximately the same accuracy, corroborating the previous results.

1.5.2 Example 2

In this example, the following discrete-time system with missing measurements has been considered

$$x_k = \left(1 + 0.2\sin\left(\frac{(k-1)\pi}{50}\right)\right)\begin{pmatrix}0.8 & 0\\ 0.9 & 0.2\end{pmatrix}x_{k-1} + w_{k-1}, \quad k \geq 1$$
$$y_k^i = \theta_k^i \begin{pmatrix}1 & 1\end{pmatrix} x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2$$

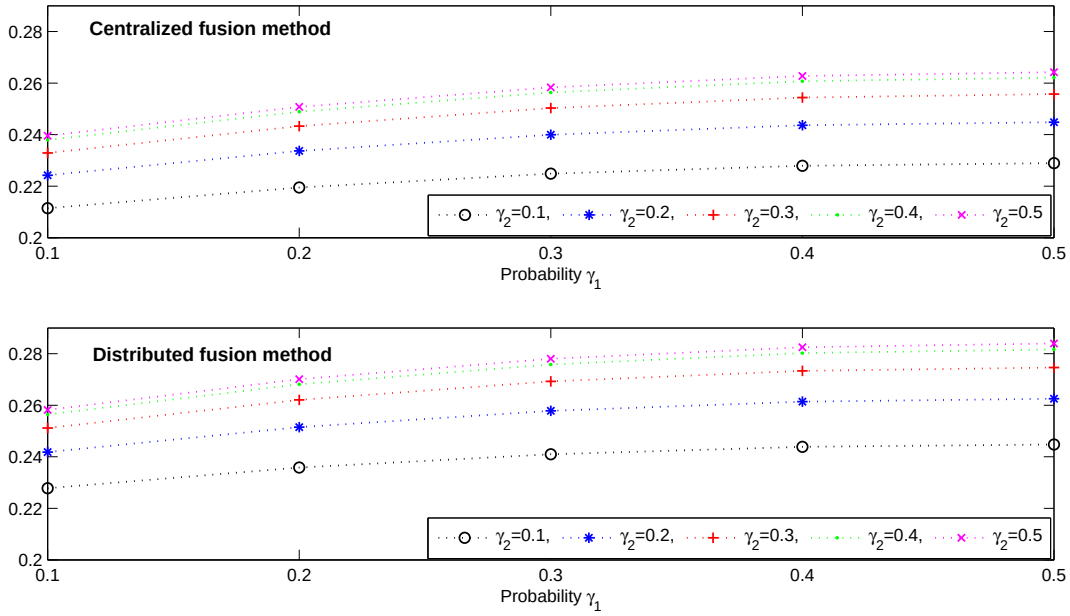


Figure 1.3: Filtering error variances for the centralized and distributed fusion methods at $k = 50$ versus γ_1, with γ_2 varying from 0.1 to 0.5 when $m = 3$.

where the initial state, x_0, is a zero-mean Gaussian vector with covariance matrix given by $Cov[x_0] = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$, the processes $\{w_k; k \geq 0\}$ and $\{v_k^i;\ k \geq 1\}$, $i = 1, 2$ are zero-mean white Gaussian noises with $Cov[w_k] = \begin{pmatrix} 0.36 & 0.3 \\ 0.3 & 0.25 \end{pmatrix}$, $Var[v_k^1] = 0.5$ and $Var[v_k^2] = 0.9$, $\forall k$, and the multiplicative noises $\{\theta_k^i;\ k \geq 1\}$, $i = 1, 2$ are defined as in Example 1.

Firstly, our aim is to check that the accuracy of the optimal distributed fusion filter is higher than that of any local filter, but lower than that of the centralized fusion filter. For this, two hundred iterations of the proposed algorithms have been carried out and the results corresponding to the first state component for $m = 3$ and probabilities $\gamma_1 = 0.1$ and $\gamma_2 = 0.2$ are shown graphically in Figure 1.4. As in Figure 1, the error variances of each local filter are higher than that of the distributed fusion filter and the centralized and distributed filters have a similar

accuracy. Analogous results for the second state component are obtained.

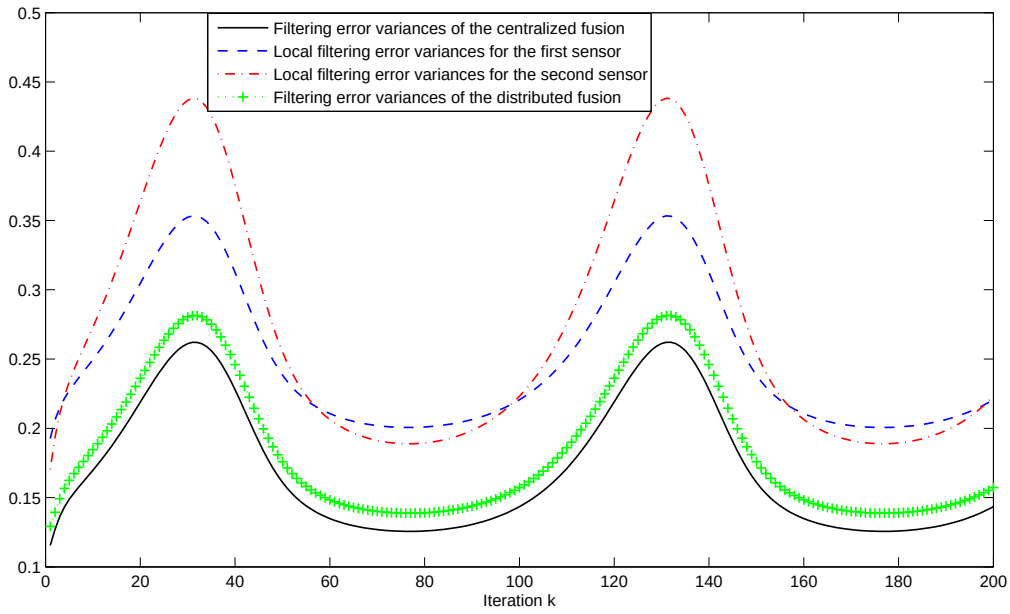


Figure 1.4: Filtering error variances for the centralized and distributed fusion methods for the first state component for $\gamma_1 = 0.1$, $\gamma_2 = 0.2$, when $m = 3$.

Also, analogous comments and conclusions to those made from Figures 1.2 and 1.3 in Example 1 are deduced for the first and second components of the state in this example. For this reason, the corresponding figures have not been included.

Finally, for $\gamma_1 = 0.2, \gamma_2 = 0.4$ the performance of the estimators is compared for different values of m at a fixed iteration; specifically, for $m = 2$, 3, 4, 5 at $k = 30$, the filtering error variances of both state components are shown in Table 1.1. From this table it is gathered that the estimators are more accurate as the values of m are lower. In other words, a greater distance between the instants at which the variables are correlated (which means that the state can be missing in more consecutive observations) yields worse estimators. As expected, this table also shows that the estimators obtained by the centralized and distributed fusion methods have a

very rough precision. It must be noticed that an analogous comparison has been performed in Example 1 and the results obtained are completely similar, so they have been omitted.

Component	Filtering error variances	$m = 2$	$m = 3$	$m = 4$	$m = 5$
First	Centralized	0.3310	0.3483	0.3628	0.3744
	Distributed	0.3561	0.3717	0.3827	0.3906
Second	Centralized	0.4014	0.4451	0.4835	0.5151
	Distributed	0.4340	0.4722	0.5007	0.5218

Table 1.1: Filtering error variances for the centralized and distributed fusion methods for $\gamma_1 = 0.2$, $\gamma_2 = 0.4$ at $k = 30$ when $m = 2$, 3, 4, 5.

1.6 Conclusions

For multi-sensor linear discrete-time systems with missing measurements, the LS linear estimation problem has been addressed. The main contributions of the current paper can be summarized as follows:

1. Using both centralized and distributed fusion methods to process the measurement data from the different sensors, recursive filtering and fixed-point smoothing algorithms are derived by an innovation approach.

2. At each sensor, the possibility of missing measurements or uncertain observations (that is, observations containing no information about the state but only noise) is modeled by binary variables taking the values one or zero (Bernoulli variables), depending on whether the state is present or missing in the corresponding observation. Such variables are assumed to be correlated at instants that differ m units of time.

3. The basic model in which the Bernoulli variables describing the uncertainty in the observations at each sensor are independent is a particular case of the proposed model, just making $K_{k,s}^{\theta^i} = 0$ for $|k - s| = m$. Also, the

model with correlation in consecutive sampling times is covered by the current study when $m = 1$. However, theses two assumptions can be unrealistic in many practical situations, and the estimation algorithms must be modified to incorporate the effect of different types of correlation. Specifically, the form of correlation considered in this paper is appropriate, in particular, to model situations where the state cannot be missing in $m + 1$ consecutive observations, as occurs, for instance, in sensor networks where sensor failures may happen and a failed sensor is substituted not immediately, but after m sampling times.

4. The multi-sensor system model considered in the current paper covers those situations where the additive observation noises and the Bernoulli variables involved are independent from sensor to sensor. This independence assumption simplifies the mathematical expression considerably and it is valid in a wide spectrum of applications, for example in wireless sensor networks which are characterized by sensor independence, limited storage capacity, lack of physical infrastructure and limited energy. Nevertheless, if such assumption is omitted, a similar technique to that used in this paper would allow us to extend the current study to this more general case with no difficulty, except for a greater complexity in the mathematical expressions.

5. Two numerical simulation examples illustrate the applicability of the current results to estimate a scalar state process generated by an AR model and a two-dimensional state, respectively, from uncertain observations coming from two sensors featuring correlation in the uncertainty. The results confirm that centralized and distributed fusion estimators have approximately the same accuracy. For different uncertainty probabilities and different values of m, both examples confirm the greater effectiveness of the fixed-point smoothing estimators in contrast to the filtering ones and conclude that more accurate estimations are obtained as the values of m are lower.

Acknowledgements

This research is supported by Ministerio de Ciencia e Innovación (Programa FPU and Grant No. MTM2011-24718).

A.1 Proof of Theorem 1.3.1

From the OPL, the LS linear estimators, $\widehat{x}_{k/L}$, expressed as combination linear of the innovations, are given by

$$\widehat{x}_{k/L} = \sum_{i=1}^{L} S_{k,i}\Pi_i^{-1}\nu_i. \tag{1.30}$$

where $\nu_i = y_i - \widehat{y}_{i/i-1}$ are the innovation vectors, with $\widehat{y}_{i/i-1}$ the one-stage observation predictor, $\Pi_i = E[\nu_i\nu_i^T]$, and $S_{k,i} = E[x_k\nu_i^T]$.

Using (1.30) with $L = k,\ k-1$, expression (1.5) for the filter is immediately derived. From (1.1) and OPL, expression (1.6) for the the state predictor is easily obtained.

Now we show expression (1.7) for the innovation, $\nu_k = y_k - \widehat{y}_{k/k-1}$, for which it is enough to obtain an expression for $\widehat{y}_{k/k-1}$. From the OPL, it follows that $\widehat{y}_{k/k-1}$ is given by

$$\widehat{y}_{k/k-1} = \sum_{i=1}^{k-1} T_{k,i}\Pi_i^{-1}\nu_i, \quad k \geq 2, \quad T_{k,i} = E\left[y_k\nu_i^T\right].$$

Hence, we start by calculating $T_{k,i}$, for $i \leq k-1$. From the observation equation (1.3) and the model assumptions, it is clear that $T_{k,i} = E\left[\Theta_k H_k x_k \nu_i^T\right]$, for $i \leq k-1$, and $T_{k,i} = \overline{\Theta}_k H_k S_{k,i}$, for $k \leq m$ or $k > m$ and $i < k-m$. So, after some manipulations, we obtain:

(a) For $k \leq m$, using (1.30) for $L = k-1$, we have $\widehat{y}_{k/k-1} = \overline{\Theta}_k H_k \widehat{x}_{k/k-1}$.

(b) For $k > m$, the following equality is easily deduced

$$\widehat{y}_{k/k-1} = \overline{\Theta}_k H_k \sum_{i=1}^{k-1} S_{k,i}\Pi_i^{-1}\nu_i + \sum_{i=1}^{m}(T_{k,k-i} - \overline{\Theta}_k H_k S_{k,k-i})\Pi_{k-i}^{-1}\nu_{k-i}, \tag{1.31}$$

where

$$T_{k,k-i} - \overline{\Theta}_k H_k S_{k,k-i} = E\left[(\Theta_k - \overline{\Theta}_k) H_k x_k \nu_{k-i}^T\right], \quad 1 \leq i \leq m, \tag{1.32}$$

or equivalently,

$$T_{k,k-i} - \overline{\Theta}_k H_k S_{k,k-i} = E\left[(\Theta_k - \overline{\Theta}_k) H_k x_k y_{k-i}^T\right] - E\left[(\Theta_k - \overline{\Theta}_k) H_k x_k \widehat{y}_{k-i/k-(i+1)}^T\right].$$

Using again (1.3) for y_{k-i}, property (1.4), and since from (1.1), $E[x_k x_{k-i}^T] = \mathbb{F}_{k,k-i} D_{k-i}$, it is concluded that

$$E\left[(\Theta_k - \overline{\Theta}_k) H_k x_k y_{k-i}^T\right] = K_{k,k-i}^{\theta} \circ \left(H_k \mathbb{F}_{k,k-i} D_{k-i} H_{k-i}^T\right),$$

where $D_k = E[x_k x_k^T]$ can be clearly obtained by the recursive formula (1.8).

Summarizing, we have that

$$\begin{aligned} T_{k,k-i} - \overline{\Theta}_k H_k S_{k,k-i} = {} & K_{k,k-i}^{\theta} \circ \left(H_k \mathbb{F}_{k,k-i} D_{k-i} H_{k-i}^T\right) \\ & - E\left[(\Theta_k - \overline{\Theta}_k) H_k x_k \widehat{y}_{k-i/k-(i+1)}^T\right], \quad 1 \leq i \leq m. \end{aligned} \tag{1.33}$$

On the one hand, for $i = m$, since Θ_k is independent of the innovations ν_i, for $i < k - m$, we have that $E\left[(\Theta_k - \overline{\Theta}_k) H_k x_k \widehat{y}_{k-m/k-(m+1)}^T\right] = 0$, and from (1.33)

$$T_{k,k-m} - \overline{\Theta}_k H_k S_{k,k-m} = K_{k,k-m}^{\theta} \circ \left(H_k \mathbb{F}_{k,k-m} D_{k-m} H_{k-m}^T\right). \tag{1.34}$$

On the other hand, for $i < m$, $K_{k,k-i}^{\theta} = 0$ and, hence, from (1.33)

$$T_{k,k-i} - \overline{\Theta}_k H_k S_{k,k-i} = -E\left[(\Theta_k - \overline{\Theta}_k) H_k x_k \widehat{y}_{k-i/k-(i+1)}^T\right].$$

Now, using again that Θ_k is independent of ν_i, for $i \neq k - m$, it is deduced that

$$T_{k,k-i} - \overline{\Theta}_k H_k S_{k,k-i} = -E[\left(\Theta_k - \overline{\Theta}_k\right) H_k x_k \nu_{k-m}^T] \Pi_{k-m}^{-1} T_{k-i,k-m}^T$$

or, equivalently, from (1.32) for $i = m$, (1.34) and denoting

$$\Psi_{k,k-m} = K_{k,k-m}^{\theta} \circ \left(H_k \mathbb{F}_{k,k-m} D_{k-m} H_{k-m}^T\right) \Pi_{k-m}^{-1},$$

we have that

$$T_{k,k-i} - \overline{\Theta}_k H_k S_{k,k-i} = -\Psi_{k,k-m} T^T_{k-i,k-m}, \quad i < m. \tag{1.35}$$

Next, substituting this expression into (1.31) and using (1.30) for $\widehat{x}_{k/k-1}$, expression (1.7) is deduced. Moreover, using (1.3) and (1.35) and taking into account that, from (1.1), $S_{k,k-i} = \mathbb{F}_{k,k-i} S_{k-i,k-i}$, the matrices (1.9) are obtained.

Now, expression (1.10) for the matrix $S_{k,k} = E[x_k y_k^T] - E[x_k \widehat{y}^T_{k/k-1}]$ is derived. From (1.3) and the independence assumption, it is clear that $E[x_k y_k^T] = D_k H_k^T \overline{\Theta}_k$, $\forall k \geq 1$. To calculate $E[x_k \widehat{y}^T_{k/k-1}]$, the correlation assumption of the random variables θ_k must be taken into account and hence two cases must be considered:

- For $k \leq m$, from (1.7) we obtain $E[x_k \widehat{y}^T_{k/k-1}] = E[x_k \widehat{x}^T_{k/k-1}] H_k^T \overline{\Theta}_k$. From the OPL, $E[x_k \widehat{x}^T_{k/k-1}] = D_k - P_{k/k-1}$ where $P_{k/k-1}$ is the prediction error covariance matrix, and hence

$$E[x_k \widehat{y}^T_{k/k-1}] = \left(D_k - P_{k/k-1}\right) H_k^T \overline{\Theta}_k, \quad k \leq m.$$

- For $k > m$, from (1.7) it follows that

$$\begin{aligned} E[x_k \widehat{y}^T_{k/k-1}] = {} & E[x_k \widehat{x}^T_{k/k-1}] H_k^T \overline{\Theta}_k + E[x_k \nu^T_{k-m}] \Psi^T_{k,k-m} \\ & - E\left[x_k \left(\sum_{i=1}^{m-1} T^T_{k-i,k-m} \Pi^{-1}_{k-i} \nu_{k-i} \right)^T \right] \Psi^T_{k,k-m}, \end{aligned}$$

  hence, using again the OPL and taking into account that $E[x_k \nu^T_{k-i}] = S_{k,k-i}$, for $1 \leq i \leq m$, it is deduced that

$$\begin{aligned} E[x_k \widehat{y}^T_{k/k-1}] = {} & \left(D_k - P_{k/k-1}\right) H_k^T \overline{\Theta}_k + S_{k,k-m} \Psi^T_{k,k-m} \\ & - \sum_{i=1}^{m-1} S_{k,k-i} \Pi^{-1}_{k-i} T_{k-i,k-m} \Psi^T_{k,k-m}, \quad k > m. \end{aligned}$$

From the above expectations, expression (1.10) for $S_{k,k}$ is clear.

From (1.1), the expression for the prediction error covariance matrix, $P_{k/k-1}$ is immediately clear and, from (1.5), the expression for the filtering error covariance matrix, $P_{k/k}$, is also obvious.

Finally, we prove expression (1.11) for the innovation covariance matrix $\Pi_k = E[y_k y_k^T] - E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T]$. From (1.3) and using (1.4), we have that

$$E[y_k y_k^T] = E[\theta_k \theta_k^T] \circ \left(H_k D_k H_k^T\right) + R_k, \quad k \geq 1.$$

Due to the correlation hypothesis of the Bernoulli variables θ_k, we need to distinguish two cases to calculate $E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T]$. For $k \leq m$, from (1.7), (1.4) and the OPL, we have

$$E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T] = \left(\overline{\theta}_k \overline{\theta}_k^T\right) \circ \left(H_k(D_k - P_{k/k-1})H_k^T\right).$$

For $k > m$, using an analogous reasoning, applying the OPL and after some manipulations, we deduce that

$$\begin{aligned} E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T] = & \left(\overline{\theta}_k \overline{\theta}_k^T\right) \circ \left(H_k(D_k - P_{k/k-1})H_k^T\right) + \Psi_{k,k-m}\Pi_{k-m}\Psi_{k,k-m}^T \\ & + \Psi_{k,k-m} \sum_{i=1}^{m-1} T_{k-i,k-m}^T \Pi_{k-i}^{-1} T_{k-i,k-m} \Psi_{k,k-m}^T \\ & - \overline{\Theta}_k H_k \left(S_{k,k} - P_{k/k-1}H_k^T \overline{\Theta}_k\right) - \left(S_{k,k}^T - \overline{\Theta}_k H_k P_{k/k-1}\right) H_k^T \overline{\Theta}_k. \end{aligned}$$

So, from the above expectations, expression (1.11) for the innovation covariance matrix Π_k is obtained. □

A.2 Proof of Theorem 1.3.2

From the general expression (1.30), for each fixed $k \geq 1$, the recursive relation (1.12) is immediately clear.

Next, to prove (1.13) for $S_{k,k+N} = E[x_k y_{k+N}^T] - E[x_k \widehat{y}_{k+N/k+N-1}^T]$, it is necessary to calculate both expectations.

On the one hand, from Equation (1.3), taking into account that $E[x_k x_{k+N}^T] = D_k \mathbb{F}_{k+N,k}^T$ and using that Θ_{k+N} and v_{k+N} are independent of x_k, we obtain

$$E[x_k y_{k+N}^T] = D_k \mathbb{F}_{k+N,k}^T H_{k+N}^T \overline{\Theta}_{k+N}, \quad N \geq 1.$$

On the other hand, based on expression (1.7) for ν_{k+N}, which is different depending on wether $k+N \leq m$ or $k+N > m$, two options must be considered:

– From (1.7) for $k \leq m-N$, using (1.6) for $\widehat{x}_{k+N/k+N-1}$, we have that

$$E\left[x_k \widehat{y}_{k+N/k+N-1}^T\right] = M_{k,k+N-1} F_{k+N-1}^T H_{k+N}^T \overline{\Theta}_{k+N},$$

where $M_{k,k+N-1} = E\left[x_k \widehat{x}_{k+N-1/k+N-1}^T\right]$.

– A similar reasoning to the above one, but starting from (1.7) for $k > m-N$, yields

$$E[x_k \widehat{y}_{k+N/k+N-1}^T] = M_{k,k+N-1} F_{k+N-1}^T H_{k+N}^T \overline{\Theta}_{k+N} + \left(S_{k,k+N-m} - \sum_{i=1}^{m-1} S_{k,k+N-i} \Pi_{k+N-i}^{-1} T_{k+N-i,k+N-m} \right) \Psi_{k+N,k+N-m}^T.$$

Then, the replacement of the above expectations in $S_{k,k+N}$ leads to expression (1.13).

The recursive relation (1.14) for $M_{k,k+N} = E\left[x_k \widehat{x}_{k+N/k+N}^T\right]$ is immediately clear from (1.5) for $\widehat{x}_{k+N/k+N}$ and its initial condition $M_{k,k} = E[x_k \widehat{x}_{k/k}]$ is calculated taking into account that, from the orthogonality, $E[x_k \widehat{x}_{k/k}^T] = E[\widehat{x}_{k/k} \widehat{x}_{k/k}^T] = D_k - P_{k/k}$.

Finally, since $P_{k/k+N} = E\left[x_k x_k^T\right] - E\left[\widehat{x}_{k/k+N} \widehat{x}_{k/k+N}^T\right]$, using (1.12) and taking into account that $\widehat{x}_{k/k+N-1}$ is uncorrelated with ν_{k+N}, expression (1.15) is deduced.

□

References

[1] J. Hespanha, P. Naghshtabrizi and Yonggang Xu, A survey of recent results in networked control systems, *Proceedings of the IEEE,* 95(1) (2007), 138–162.

[2] L. Ma, Z. Wang, J. Hu, Y. Bo and Z. Guo, Robust variance-constrained filtering for a class of nonlinear stochastic systems with missing measurements, *Signal Processing,* 90 (2010), 2060–2071.

[3] A. Hermoso-Carazo and J. Linares-Pérez, Nonlinear estimation applying an unscented transformation in systems with correlated uncertain observations, *Applied Mathematics and Computation,* 217 (2011), 7998–8009.

[4] S. Zhou and G. Feng, H_∞ filtering for discrete-time systems with randomly varying sensor delays, *Automatica,* 44 (2008), 1918–1922.

[5] S. Nakamori, A. Hermoso-Carazo and J. Linares-Pérez, Design of RLS Wiener estimators from randomly delayed observations in linear discrete-time stochastic systems, *Applied Mathematics and Computation,* 217 (2010), 3801–3815.

[6] S.L. Sun, Optimal estimators for systems with finite consecutive packet dropouts, *IEEE Signal Processing Letters,* 16(7) (2009), 557–560.

[7] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Recursive least-squares quadratic smoothing from measurements with packet dropouts, *Signal Processing,* 92 (2012), 931–938.

[8] S. Sun and W. Xiao, Optimal linear estimators for systems with multiple random measurement delays and packet dropouts, *International Journal of Systems Science,* 44(2) (2013), 358–370.

[9] M. Moayedi, Y.K. Foo and Y.C. Soh, Adaptive Kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing

measurements, *IEEE Transactions on Signal Processing,* 58(3) (2010), 1577–1588.

[10] J. Ma and S. Sun, Optimal linear estimators for systems with random sensor delays, multiple packet dropouts and uncertain observations, *IEEE Transactions on Signal Processing,* 59(11) (2011), 5181–5192.

[11] F.O. Hounkpevi and E.E. Yaz, Robust minimum variance linear state estimators for multiple sensors with different failure rates, *Automatica,* 43 (2007), 1274–1280.

[12] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Linear and quadratic estimation using uncertain observations from multiple sensors with correlated uncertainty, *Signal Processing,* 91 (2011), 330–337.

[13] F.O. Hounkpevi and E.E. Yaz, Minimum variance generalized state estimators for multiple sensors with different delay rates, *Signal Processing,* 87 (2007), 602–613.

[14] J. Linares-Pérez, A. Hermoso-Carazo, R. Caballero-Águila and J.D. Jiménez-López, Least-squares linear filtering using observations coming from multiple sensors with one- or two-step random delay, *Signal Processing,* 89 (2009), 2045–2052.

[15] J. Ma and S. Sun, Information fusion estimators for systems with multiple sensors of different packet dropout rates, *Information Fusion,* 12 (2011), 213–222.

[16] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Least-squares linear estimators using measurements transmitted by different sensors with packet dropouts, *Digital Signal Processing,* 22 (2012), 1118–1125.

[17] K.H. Kim, Development of track to track fusion algorithms, *American Control Conference,* 1994, 1037–1041.

[18] S. Sun and Z.L. Deng, Multi-sensor optimal information fusion Kalman filter, *Automatica,* 40 (2004), 1017–1023.

[19] J. Feng and M. Zeng, Descriptor recursive estimation for multiple sensors with different delay rates, *International Journal of Control,* 84(3) (2011), 584–596.

[20] M.J. García-Ligero, A. Hermoso-Carazo and J. Linares-Pérez, Distributed and centralized fusion estimation from multiple sensors with Markovian delays, *Applied Mathematics and Computation,* 219 (2012), 2932–2948.

[21] W.-A. Zhang, G. Feng and L. Yu, Multi-rate distributed fusion estimation for sensor networks with packet losses, *Automatica,* 48 (2012), 2016–2028.

[22] S. Sun and W. Xiao, Distributed weighted fusion estimators with random delays and packet dropping, *Circuits Systems Signal Processing,* 26(4) (2007), 591–605.

[23] J. Sun and C. Zhang, Distributed optimal fusion estimators for multi-sensor systems with bounded random measurement delays and packet dropouts, *Journal of Computational Information Systems,* 8(10) (2012), 4087–4094.

[24] B. Shen, Z. Wang and Y.S. Hung, Distributed H_∞-consensus filtering in sensor networks with multiple missing measurements: The finite-horizon case, *Automatica,* 46 (2010), 1682–1688.

[25] B. Chen, L. Yu, W.-A. Zhang and A. Liu, Distributed robust fusion Kalman filtering for uncertain stochastic systems with random delays and missing measurements, *Control Conference (ASCC), 2011 8th Asian,* (2011), 1288–1293.

[26] M. Liu, D. Qi, S. Zhang, M. Qiu and S. Zheng, Optimal H_∞ fusion filters for a class of discrete-time intelligent systems with time delays and missing measurement, *Neurocomputing,* 74 (2011), 3741–3752.

Chapter 2

Centralized fusion quadratic estimators in multi-sensor systems with correlated missing measurements

Reference

García-Garrido, I., Linares-Pérez, J. and Caballero-Águila, R. (2013).
Centralized fusion quadratic estimators in multi-sensor systems with correlated missing measurements.
Applied Mathematical Sciences, 57(7): 2795-2813.

Abstract

In this paper, the centralized fusion quadratic estimation problem in linear discrete-time stochastic systems with missing measurements coming from multiple sensors is addressed when the Bernoulli variables describing the phenomenon of missing measurements are correlated at instants that differ m sampling times. For this purpose, an appropriate augmented system is defined and the required quadratic estimators of the original state are obtained from the linear estimators of the

augmented state. By using an innovation approach, recursive algorithms for the least-squares linear filtering and fixed-point smoothing problems of the augmented system are derived. The performance of the proposed estimators is illustrated by a simulation example where centralized fusion linear and quadratic estimators are compared in terms of their error variances for different missing probabilities and values of m.

Contents

2.1 Introduction

In some real situations where sensor networks are used, the state estimation problem is addressed under the assumption that, at each sampling time, the available measurements always contain information about the current state. However, the unreliable network characteristics usually lead to problems such as accidental loss of some measurements, intermittent failures or random interruptions in the transmission mechanism, among others. These situations, called *missing measurements* or *uncertain observations*, are characterized by including in the measured output besides an additive noise, a multiplicative noise defined by a sequence of Bernoulli random variables, modelling the possibility that the system state vector may or may not be present in the corresponding measurement.

Due to this multiplicative noise component, systems with missing measurements

are not Gaussian in general (even if the additive noises are Gaussian) and the practical computation of the optimal least-squares (LS) estimator is not a simple task. This difficulty gives rise to the need of searching for suboptimal estimators, easier to derive, such as linear or even polynomial estimators which improve the extensively used linear ones. In the last years, the estimation problem in this kind of systems has been widely studied under different hypotheses and approaches on the processed involved (see e.g. [1] and [10] and references therein).

In the above papers, authors consider that the measurement data available for the estimation come either from a single sensor or from multiple sensors with identical uncertainty characteristics. Nevertheless, this is not a realistic assumption in several application fields, for instance, in networked communication systems involving heterogeneous measurement devices (see e.g. [6]); hence, multi-sensor systems featuring different sensor statistical properties are an increasing research challenge. To process the measured sensor data, centralized fusion method, which consists of considering that all measurement data coming from multiple sensors are transmitted to a fusion center for being processed, has been commonly used. In [5] and [9] centralized linear estimators are designed in the linear minimum variance sense by considering independent variables modelling the uncertainty in the observations. The optimal LS centralized linear and quadratic problems are also investigated in [2] for systems in which the phenomenon of missing measurements is modelled by Bernoulli variables correlated at consecutive sampling times, while Bernoulli variables correlated at instants that differ two units of time are considered in [4].

These types of correlation cover several practical situations, for example, in sensor networks where sensor failures may happen and a failed sensor is not replaced or replaced one or two sampling times after having failed. However, even if it is assumed that any failure in the transmission results from sensor failures, usually the failed sensor may not be replaced immediately but after m instants of time; to cover such situations, correlation among the random variables modelling the

uncertainty in the observations at times k and $k+m$ has been considered in [8], where recursive algorithms for the linear filter and fixed-point smoother have been derived.

In this paper, the centralized fusion quadratic estimation problem in linear discrete-time stochastic systems with missing measurements coming from multiple sensors is addressed, when, at each sensor, the random variables modelling the phenomenon of missing measurements are m-step autocorrelated. For this purpose, recursive algorithms for the LS quadratic filtering and fixed-point smoothing problems are presented, both yielding a significant improvement over the linear estimation problem.

To address the quadratic estimation problem, the technique proposed in [3] is used, which consists of augmenting the state and measurement vectors, by assembling the original vectors and their second-order Kronecker powers, thus the quadratic estimation problem for the original state is reduced to the linear estimation problem for the augmented state. The rest of the paper is organized as follows. In Section 2.2 the state-space model is described and the assumptions about the state and noise processes are presented. In Section 2.3 the quadratic estimation problem is formulated based on the augmented system and some properties about the initial state and noise processes involved in this augmented system. By using an innovation approach, the linear estimators of the augmented state are derived in Section 2.4, providing the required quadratic estimators. Finally, in Section 2.5, the effectiveness of the estimation algorithms is illustrated by a numerical simulation example where centralized fusion linear and quadratic estimators are compared in terms of their error variances.

2.2 Model assumptions

Our aim is to address the LS quadratic estimation problem in linear discrete-time stochastic systems with missing measurements coming from multiple sensors when

the Bernoulli variables describing the phenomenon of missing measurements are correlated at instants that differ m sampling times. In this section, the system model is described and the assumptions about the state and noise processes are stated.

Consider a class of discrete-time linear stochastic systems with missing measurements (that is, observations containing only noise) coming from r sensors, whose mathematical modelling is described by the following state and measurement equations:

$$x_k = F_{k-1}x_{k-1} + w_{k-1}, \quad k \geq 1, \tag{2.1}$$

$$y_k^i = \theta_k^i H_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2, \ldots, r, \tag{2.2}$$

where $x_k \in \mathbb{R}^n$ is the state and $y_k^i \in \mathbb{R}$, $i = 1, 2, \ldots, r$, is the measurement coming from sensor i at sampling time k. The additive noises $\{w_k;\, k \geq 0\}$ and $\{v_k^i;\, k \geq 1\}$, $i = 1, 2, \ldots, r$, are white sequences. F_k and H_k^i, $i = 1, 2, \ldots, r$, are known time-varying matrices with compatible dimensions. The multiplicative noises $\{\theta_k^i;\, k \geq 1\}$, $i = 1, 2, \ldots, r$, are sequences of Bernoulli random variables, which are used to model the phenomenon of missing measurements; their values –one or zero– indicate whether the state is present or missing in the corresponding measurement. The number of sensors is denoted by r and the i-th sensor by superscript i.

Our aim is to obtain the LS quadratic estimator of the state x_k based on the received measurements $\{y_1^i, \ldots, y_L^i\}$, $L \geq k$, $i = 1, 2, \ldots, r$, by using the centralized fusion method, which consists of stacking the observations coming from multiple sensors in a fusion center for being processed. For this purpose and to simplify the notation, the measurement equation (2.2) is rewritten in a compact form as:

$$y_k = \Theta_k H_k x_k + v_k, \quad k \geq 1,$$

where $y_k = (y_k^1, \ldots, y_k^r)^T$, $v_k = (v_k^1, \ldots, v_k^r)^T$, $H_k = (H_k^{1T}, \ldots, H_k^{rT})^T$ and $\Theta_k = Diag(\theta_k^1, \ldots, \theta_k^r)$.

In order to address the LS quadratic estimation problem of the state x_k from the observations $y_1, \ldots, y_L$, $L \geq k$, several considerations must be taken into account. As it is known, the LS quadratic estimator of x_k based on the observations $y_1, \ldots, y_L$, is the orthogonal projection of x_k onto the space of n-dimensional random variables obtained as linear transformations of $y_1, \ldots, y_L$ and their second-order powers, $y_1^{[2]}, \ldots, y_L^{[2]}$ (defined by the Kronecker product, $y_i^{[2]} = y_i \otimes y_i$). Hence, to address the LS quadratic estimation problem, the existence of the second-order moments of such vectors, $y_i^{[2]}$, is required. On the other hand, as indicated previously, this problem will be addressed under the assumption that the variables describing the uncertainty in the observations are m-step autocorrelated. Specifically, the following assumptions are assumed:

(A1) The initial state x_0 is a random vector with $E[x_0] = \overline{x}_0$, $Cov[x_0] = P_0$, $Cov[x_0, x_0^{[2]}] = P_0^{(3)}$ and $Cov[x_0^{[2]}] = P_0^{(4)}$.

(A2) The state noise $\{w_k;\, k \geq 0\}$ is a zero-mean white sequence with $Cov[w_k] = Q_k$, $Cov[w_k, w_k^{[2]}] = Q_k^{(3)}$ and $Cov[w_k^{[2]}] = Q_k^{(4)}$.

(A3) The measurement noise $\{v_k;\, k \geq 1\}$ is a zero-mean white sequence with $Cov[v_k] = R_k$, $Cov[v_k, v_k^{[2]}] = R_k^{(3)}$ and $Cov[v_k^{[2]}] = R_k^{(4)}$.

(A4) For $i = 1, 2, \ldots, r$, the multiplicative noises $\{\theta_k^i; k \geq 1\}$ are sequences of Bernoulli random variables with known probabilities $P[\theta_k^i = 1] = \overline{\theta}_k^i$. For $i, j = 1, 2, \ldots, r$, the variables θ_k^i and θ_s^j are independent for $|k - s| \neq 0, m$, and $Cov[\theta_k^i, \theta_s^j]$ are known for $|k - s| = 0, m$.

(A5) The initial state x_0 and the processes $\{w_k; k \geq 0\}$, $\{v_k;\, k \geq 1\}$ and $\{\theta_k = (\theta_k^1, \ldots, \theta_k^r)^T;\, k \geq 1\}$ are mutually independent.

Remark 1. Obviously, from the state equation (2.1) and under assumptions (A1), (A2) and (A5), the state expectation satisfies $\overline{x}_k = E[x_k] = F_{k-1}\overline{x}_{k-1}$ and

$E_k = E[x_k x_k^T]$ is recursively calculated as

$$E_k = F_{k-1} E_{k-1} F_{k-1}^T + Q_{k-1}, \quad k \geq 1, \quad E_0 = P_0 + \overline{x}_0 \overline{x}_0^T.$$

2.3 Quadratic estimation problem

Under the above assumptions (A1)-(A5), our purpose is to obtain the LS quadratic estimator, $\widehat{x}^q_{k/L}$, of the state x_k, based on the measurements $y_1,\ldots,y_L$, $L \geq k$. More specifically, our aim is to derive recursive algorithms for the filter, $\widehat{x}^q_{k/k}$ and the fixed-point smoother, $\widehat{x}^q_{k/k+N}$, $N \geq 1$.

In order to obtain these estimators the following augmented state and measurement vectors are defined by assembling the original vectors and their second-order Kronecker powers:

$$\mathcal{X}_k = \begin{pmatrix} x_k \\ x_k^{[2]} \end{pmatrix}, \qquad \mathcal{Y}_k = \begin{pmatrix} y_k \\ y_k^{[2]} \end{pmatrix}.$$

Note that the n-dimensional space of linear transformations of $\mathcal{Y}_1,\ldots,\mathcal{Y}_L$ is equal to the n-dimensional space of linear transformations of $y_1,\ldots,y_L$ and $y_1^{[2]},\ldots,y_L^{[2]}$. Therefore, it is clear that the LS quadratic estimator, $\widehat{x}^q_{k/L}$, is the LS linear estimator of x_k based on $\mathcal{Y}_1,\ldots,\mathcal{Y}_L$, which is obtained by extracting the first n entries of the LS linear estimator of $\mathcal{X}_k$ based on $\mathcal{Y}_1,\ldots,\mathcal{Y}_L$. So, the quadratic estimation problem for the original state is reduced to the linear estimation problem for the augmented state.

For simplicity, as in [2] and [4], to address the LS linear estimation problem of the augmented state, the centered augmented state and measurement vectors $X_k = \mathcal{X}_k - E[\mathcal{X}_k]$ and $Y_k = \mathcal{Y}_k - E[\mathcal{Y}_k]$, respectively, are considered; these vectors satisfy the following augmented system:

$$X_k = \mathcal{F}_{k-1} X_{k-1} + W_{k-1}, \quad k \geq 1, \tag{2.3}$$

$$Y_k = D_k^{\Theta} \mathcal{H}_k X_k + V_k, \quad k \geq 1 \tag{2.4}$$

where

$$\mathcal{F}_k = Diag(F_k,\, F_k^{[2]}), \quad \mathcal{H}_k = Diag(H_k,\, H_k^{[2]}), \quad D_k^{\Theta} = Diag(\Theta_k, \Theta_k^{[2]}),$$

$$W_k = \begin{pmatrix} w_k \\ (I+K)((F_k x_k) \otimes w_k) + w_k^{[2]} - vec(Q_k) \end{pmatrix},$$

(I and K denote the identity and commutation matrices of compatible dimensions, respectively),

$$V_k = \begin{pmatrix} v_k \\ (I+K)((\Theta_k H_k x_k) \otimes v_k) + v_k^{[2]} - vec(R_k) \end{pmatrix} + (D_k^{\Theta} - D_k^{\overline{\Theta}})\mathcal{H}_k E[\mathcal{X}_k],$$

where $D_k^{\overline{\Theta}} = E[D_k^{\Theta}]$ and $E[\mathcal{X}_k] = (\overline{x}_k, vec(E_k))^T$, with $\overline{x}_k$ and E_k given in Remark 1 ($vec(\cdot)$ denotes the 'vec' or 'stack' operator, which vectorizes a matrix).

It should be noted that the LS linear estimator of $\mathcal{X}_k$ based on $\mathcal{Y}_1, \ldots, \mathcal{Y}_k$ is obtained from the LS linear estimator of X_k based on $Y_1, \ldots, Y_k$, just adding the mean vector $E[\mathcal{X}_k]$. Hence, the required quadratic estimators $\widehat{x}_{k/k+N}$, $N \geq 0$, are obtained by adding the mean $\overline{x}_k$ to the vector constituted by the first n entries of the LS linear filter of X_k.

The following statistical properties about the initial state and noise processes involved in (2.3) and (2.4) are used to derive the LS linear estimation algorithms for this augmented system. The proof of these properties is analogous to those in [2] and [4] and, hence, it is omitted.

(i) The initial state X_0 is a zero-mean random vector with covariance matrix

$$P_0^* = \begin{pmatrix} P_0 & P_0^{(3)} \\ P_0^{(3)} & P_0^{(4)} \end{pmatrix}.$$

(ii) The noise $\{W_k;\ k \geq 0\}$ is a zero-mean white process with

$$E[W_k W_k^T] = Q_k^W = \begin{pmatrix} Q_k & Q_k^{12} \\ Q_k^{12T} & Q_k^{22} \end{pmatrix},$$

where

$$\begin{aligned}
Q_k^{12} &= ((F_k\overline{x}_k)^T \otimes Q_k)(I+K) + Q_k^{(3)},\\
Q_k^{22} &= (I+K)((F_kE_kF_k^T) \otimes Q_k)(I+K) + Q_k^{(4)}\\
&\quad + (I+K)((F_k\overline{x}_k) \otimes Q_k^{(3)}) + ((F_k\overline{x}_k) \otimes Q_k^{(3)})^T(I+K).
\end{aligned}$$

(*iii*) The noise $\{V_k;\ k \geq 1\}$ is a zero-mean process with

$$\begin{aligned}
&E[V_kV_s^T] = 0, \quad |k-s| \neq 0, m,\\
&E[V_kV_k^T] = R_{k,k}^V = \overline{R}_k + Cov[C_k^\Theta] \circ \left(\mathcal{H}_kE[\mathcal{X}_k]E[\mathcal{X}_k^T]\mathcal{H}_k^T\right),\\
&E[V_kV_{k-m}^T] = R_{k,k-m}^V = Cov[C_k^\Theta, C_{k-m}^\Theta] \circ \left(\mathcal{H}_kE[\mathcal{X}_k]E[\mathcal{X}_{k-m}^T]\mathcal{H}_{k-m}^T\right),
\end{aligned}$$

where

$$\overline{R}_k = \begin{pmatrix} R_k & R_k^{12} \\ R_k^{12T} & R_k^{22} \end{pmatrix} \quad \text{and} \quad C_k^\Theta = \left(\theta_k^T, \theta_k^{[2]T}\right)^T$$

with

$$\begin{aligned}
R_k^{12} &= ((\overline{\Theta}_kH_k\overline{x}_k)^T \otimes R_k)(I+K) + R_k^{(3)},\\
R_k^{22} &= (I+K)((E[\theta_k\theta_k^T] \circ \left(H_kE_kH_k^T\right)) \otimes R_k)\,(I+K) + R_k^{(4)}\\
&\quad + (I+K)((\overline{\Theta}_kH_k\overline{x}_k) \otimes R_k^{(3)}) + ((\overline{\Theta}_kH_k\overline{x}_k) \otimes R_k^{(3)})^T(I+K).
\end{aligned}$$

where $\overline{\Theta}_k = E[\Theta_k]$ and $\circ$ denotes de Hadamard product ($[A \circ B]_{ij} = A_{ij}B_{ij}$).

(*iv*) The initial state X_0 and the noises $\{W_k;\ k \geq 0\}$ and $\{V_k;\ k \geq 1\}$ are uncorrelated.

(*v*) The matrix D_k^Θ is independent of

$$\begin{aligned}
&\left(X_0, \{W_k;\ k \geq 0\}, V_1, \ldots, V_{k-(m+1)}, V_{k-(m-1)}, \ldots, V_{k-1},\right.\\
&\left. D_1^\Theta, \ldots, D_{k-(m+1)}^\Theta, D_{k-(m-1)}^\Theta, \ldots, D_{k-1}^\Theta\right).
\end{aligned}$$

Remark 2. For any random matrix G independent of $\{D_k^\Theta;\ k \geq 1\}$, using the Hadamard product properties it is easily derived that (see [5]):

$$E[(D_k^\Theta - D_k^{\overline{\Theta}})G(D_s^\Theta - D_s^{\overline{\Theta}})] = Cov[C_k^\Theta, C_s^\Theta] \circ E[G]. \tag{2.5}$$

Remark 3. From the augmented state equation (2.3) and properties (i), (ii) and (iv), it is easy to deduce that $\mathcal{E}_k = E[X_k X_k^T]$ is recursively calculated as:

$$\mathcal{E}_k = \mathcal{F}_{k-1}\mathcal{E}_{k-1}\mathcal{F}_{k-1}^T + Q_{k-1}^W, \quad k \geq 1; \quad \mathcal{E}_0 = P_0^*. \tag{2.6}$$

Also, it is easy to see that $E[X_k X_{k-i}^T] = \mathbb{F}_{k,k-i}\mathcal{E}_{k-i}$, $1 \leq i \leq m$, and $E[X_k X_{k+N}^T] = \mathcal{E}_k \mathbb{F}_{k+N,k}^T$, $N \geq 1$, where $\mathbb{F}_{k,i} = \mathcal{F}_{k-1} \cdots \mathcal{F}_i$.

2.4 LS quadratic estimation algorithms

Our aim in this section is to obtain the LS quadratic estimators of the state x_k based on the observations $y_1, \ldots, y_L$. For this purpose, the LS linear estimation problem of the augmented state X_k based on the augmented observations $Y_1, \ldots, Y_L$, $L \geq k$ is addressed by recursive algorithms.

For this purpose, and to simplify the derivation of the algorithms, an innovation approach will be used [7]. This approach is based on the Gram-Schmidt orthogonalization procedure which consists of transforming the observation process $\{Y_k;\, k \geq 1\}$ into an equivalent one of orthogonal vectors $\{\nu_k;\, k \geq 1\}$ named innovation process. Let ν_i be defined as $\nu_i = Y_i - \widehat{Y}_{i/i-1}$, with $\widehat{Y}_{i/i-1}$ the observation predictor, i.e., the LS linear estimator of Y_i based on the previous observations. The fact that the innovation process is uniquely determined by the observations allows us to state that the LS linear estimator of the augmented state based on the augmented observations is equal to the LS linear estimator based on the innovations. By applying the Orthogonal Projection Lemma (OPL) and taking into account that the innovations constitute a white process, it is easy to see that the estimators can be expressed as

$$\widehat{X}_{k/L} = \sum_{i=1}^{L} \mathcal{G}_{k,i}\Pi_i^{-1}\nu_i, \tag{2.7}$$

where $\Pi_i = E[\nu_i \nu_i^T]$ and $\mathcal{G}_{k,i} = E[X_k \nu_i^T]$.

In a similar way, the augmented observation predictor, $\widehat{Y}_{k/k-1}$, satisfies

$$\widehat{Y}_{k/k-1} = \sum_{i=1}^{k-1} \mathcal{T}_{k,i} \Pi_i^{-1} \nu_i, \quad \mathcal{T}_{k,i} = E[Y_k \nu_i^T]. \tag{2.8}$$

In view of the above comments and the properties of the augmented system established in Section 2.3, recursive algorithms for the linear filter, $\widehat{X}_{k/k}$, and fixed-point smoothers, $\widehat{X}_{k/k+N}$, $N \geq 1$, of the augmented state X_k are derived in the following theorems. So, as indicated previously, the required quadratic estimators, $\widehat{x}^q_{k/k}$ and $\widehat{x}^q_{k/k+N}$, $N \geq 1$, of the original state x_k are obtained by adding the mean $\overline{x}_k$ to the vector constituted by the first n entries of $\widehat{X}_{k/k+N}$, $N \geq 0$.

Theorem 2.4.1 *The LS quadratic filter, $\widehat{x}^q_{k/k}$, of the state x_k is given by*

$$\widehat{x}^q_{k/k} = \Upsilon \widehat{X}_{k/k} + \overline{x}_k, \quad k \geq 1,$$

where Υ is the operator which extracts the first n entries of $\widehat{X}_{k/k}$, the linear filter of the augmented state X_k, which is recursively obtained by

$$\widehat{X}_{k/k} = \widehat{X}_{k/k-1} + \mathcal{G}_{k,k}\ \Pi_k^{-1}\ \nu_k, \quad k \geq 1, \quad \widehat{X}_{0/0} = 0 \tag{2.9}$$

where the state predictor, $\widehat{X}_{k/k-1}$, is obtained by

$$\widehat{X}_{k/k-1} = \mathcal{F}_{k-1} \widehat{X}_{k-1/k-1}, \quad k \geq 1. \tag{2.10}$$

The innovation, ν_k, satisfies

$$\begin{aligned} \nu_k &= Y_k - D_k^{\overline{\Theta}} \mathcal{H}_k \widehat{X}_{k/k-1}, \quad k \leq m, \\ \nu_k &= Y_k - D_k^{\overline{\Theta}} \mathcal{H}_k \widehat{X}_{k/k-1} - \Psi_{k,k-m} \left[\nu_{k-m} - \sum_{i=1}^{m-1} \mathcal{T}^T_{k-i,k-m} \Pi^{-1}_{k-i} \nu_{k-i} \right], \quad k > m, \end{aligned} \tag{2.11}$$

where $\Psi_{k,k-m} = \left(Cov[C_k^{\Theta}, C_{k-m}^{\Theta}] \circ \left(\mathcal{H}_k \mathbb{F}_{k,k-m} \mathcal{E}_{k-m} \mathcal{H}_{k-m}^T \right) + R^V_{k,k-m} \right) \Pi^{-1}_{k-m}$, with $\mathcal{E}_k$ given by (2.6).

The matrices $\mathcal{T}_{k,k-i}$ are determined by

$$\begin{aligned}
&\mathcal{T}_{k,k-i} = D_k^{\overline{\Theta}}\mathcal{H}_k\mathbb{F}_{k,k-i}\mathcal{G}_{k-i,k-i}, \quad k \leq m, \quad 1 \leq i \leq k-1,\\
&\mathcal{T}_{k,k-i} = D_k^{\overline{\Theta}}\mathcal{H}_k\mathbb{F}_{k,k-i}\mathcal{G}_{k-i,k-i} - \Psi_{k,k-m}\mathcal{T}_{k-i,k-m}^T, \quad k > m, \quad 1 \leq i \leq m-1.
\end{aligned} \tag{2.12}$$

The matrix $\mathcal{G}_{k,k}$ is calculated by the following expression

$$\begin{aligned}
&\mathcal{G}_{k,k} = \Sigma_{k/k-1}\mathcal{H}_k^T D_k^{\overline{\Theta}}, \quad k \leq m,\\
&\mathcal{G}_{k,k} = \Sigma_{k/k-1}\mathcal{H}_k^T D_k^{\overline{\Theta}} - \Bigg(\mathbb{F}_{k,k-m}\mathcal{G}_{k-m,k-m}\\
&\qquad - \sum_{i=1}^{m-1}\mathbb{F}_{k,k-i}\mathcal{G}_{k-i,k-i}\Pi_{k-i}^{-1}\mathcal{T}_{k-i,k-m}\Bigg)\Psi_{k,k-m}^T, \quad k > m,
\end{aligned} \tag{2.13}$$

where $\Sigma_{k/k-1}$, the prediction error covariance matrix, is obtained by

$$\Sigma_{k/k-1} = \mathcal{F}_{k-1}\Sigma_{k-1/k-1}\mathcal{F}_{k-1}^T + Q_{k-1}^W, \quad k \geq 1,$$

with $\Sigma_{k/k}$, the filtering error covariance matrix, calculated as

$$\Sigma_{k/k} = \Sigma_{k/k-1} - \mathcal{G}_{k,k}\Pi_k^{-1}\mathcal{G}_{k,k}^T, \quad k \geq 1; \quad \Sigma_{0/0} = P_0^*.$$

The innovation covariance matrix, $\Pi_k = E[\nu_k\nu_k^T]$, satisfies

$$\begin{aligned}
&\Pi_k = Cov[C_k^{\Theta}] \circ \left(\mathcal{H}_k\mathcal{E}_k\mathcal{H}_k^T\right) + R_{k,k}^V + D_k^{\overline{\Theta}}\mathcal{H}_k\mathcal{G}_{k,k}, \quad k \leq m,\\
&\Pi_k = Cov[C_k^{\Theta}] \circ \left(\mathcal{H}_k\mathcal{E}_k\mathcal{H}_k^T\right) + R_{k,k}^V + D_k^{\overline{\Theta}}\mathcal{H}_k\mathcal{G}_{k,k}\\
&\qquad + \mathcal{G}_{k,k}^T\mathcal{H}_k^T D_k^{\overline{\Theta}} - D_k^{\overline{\Theta}}\mathcal{H}_k\Sigma_{k/k-1}\mathcal{H}_k^T D_k^{\overline{\Theta}}\\
&\qquad - \Psi_{k,k-m}\left(\Pi_{k-m} + \sum_{i=1}^{m-1}\mathcal{T}_{k-i,k-m}^T\Pi_{k-i}^{-1}\mathcal{T}_{k-i,k-m}\right)\Psi_{k,k-m}^T, \quad k > m.
\end{aligned} \tag{2.14}$$

Proof. From expression (2.7) for $L = k$, $k-1$, relation (2.9) for the filter is clear. Expression (2.10) for the state predictor is immediately obtained from (2.3) and the OPL.

Next, an explicit formula for the innovations, $\nu_k = Y_k - \widehat{Y}_{k/k-1}$, or, equivalently, for the one-stage predictor of Y_k, is deduced. From expression (2.8), we start by

calculating $\mathcal{T}_{k,i}$, for $i \leq k-1$. From (2.4) and considering the properties about the noise processes, it is clear that

$$\mathcal{T}_{k,i} = \begin{cases} E\left[D_k^{\Theta}\mathcal{H}_k X_k \nu_i^T\right], & i \leq k-1 \\ D_k^{\overline{\Theta}}\mathcal{H}_k \mathcal{G}_{k,i}, & k \leq m \text{ or } k > m,\ i < k-m. \end{cases}$$

After some manipulations, we obtain:

(a) For $k \leq m$, using (2.7) for $L = k-1$, we have

$$\widehat{Y}_{k/k-1} = D_k^{\overline{\Theta}}\mathcal{H}_k \widehat{X}_{k/k-1}. \tag{2.15}$$

(b) For $k > m$, is easily deduced that

$$\widehat{Y}_{k/k-1} = D_k^{\overline{\Theta}}\mathcal{H}_k \sum_{i=1}^{k-1} \mathcal{G}_{k,i}\Pi_i^{-1}\nu_i + \sum_{i=1}^{m}(\mathcal{T}_{k,k-i} - D_k^{\overline{\Theta}}\mathcal{H}_k\mathcal{G}_{k,k-i})\Pi_{k-i}^{-1}\nu_{k-i}, \tag{2.16}$$

where

$$\begin{aligned} \mathcal{T}_{k,k-i} - D_k^{\overline{\Theta}}\mathcal{H}_k\mathcal{G}_{k,k-i} &= E\left[(D_k^{\Theta} - D_k^{\overline{\Theta}})\mathcal{H}_k X_k \nu_{k-i}^T\right] + E[V_k \nu_{k-i}^T] \\ &= E\left[(D_k^{\Theta} - D_k^{\overline{\Theta}})\mathcal{H}_k X_k Y_{k-i}^T\right] - E\left[(D_k^{\Theta} - D_k^{\overline{\Theta}})\mathcal{H}_k X_k \widehat{Y}_{k-i/k-(i+1)}^T\right] \\ &\quad + E[V_k Y_{k-i}^T]. \end{aligned} \tag{2.17}$$

From expression (2.4) for Y_{k-i} and property (2.5), it is concluded that

$$\begin{aligned} \mathcal{T}_{k,k-i} - D_k^{\overline{\Theta}}\mathcal{H}_k\mathcal{G}_{k,k-i} = Cov[C_k^{\Theta}, C_{k-i}^{\Theta}] \circ \left(\mathcal{H}_k \mathbb{F}_{k,k-i}\mathcal{E}_{k-i}\mathcal{H}_{k-i}^T\right) + E[V_k V_{k-i}^T] \\ - E\left[(D_k^{\Theta} - D_k^{\overline{\Theta}})\mathcal{H}_k X_k \widehat{Y}_{k-i/k-(i+1)}^T\right], \quad 1 \leq i \leq m. \end{aligned}$$

– For $i = m$, since D_k^{Θ} is independent of the innovations ν_i, for $i < k-m$, $E\left[(D_k^{\Theta} - D_k^{\overline{\Theta}})\mathcal{H}_k X_k \widehat{Y}_{k-m/k-(m+1)}^T\right] = 0$ and from property (iii), we have that $E[V_k V_{k-m}^T] = R_{k,k-m}^V$. Therefore,

$$\mathcal{T}_{k,k-m} - D_k^{\overline{\Theta}}\mathcal{H}_k\mathcal{G}_{k,k-m} = Cov[C_k^{\Theta}, C_{k-m}^{\Theta}] \circ \left(\mathcal{H}_k \mathbb{F}_{k,k-m}\mathcal{E}_{k-m}\mathcal{H}_{k-m}^T\right) + R_{k,k-m}^V. \tag{2.18}$$

– For $i < m$, from the correlation assumption and property (iii), we have that $Cov[C_k^{\Theta}, C_{k-i}^{\Theta}]$ and $E[V_k V_{k-i}^T]$ are equal to zero. By using again that D_k^{Θ} is independent of ν_i, for $i \neq k-m$, it is deduced that

$$T_{k,k-i} - D_k^{\overline{\Theta}} \mathcal{H}_k \mathcal{G}_{k,k-i} = -\left(E[(D_k^{\Theta} - D_k^{\overline{\Theta}}) \mathcal{H}_k X_k \nu_{k-m}^T] + E[V_k \nu_{k-m}^T] \right) \times \Pi_{k-m}^{-1} \mathcal{T}_{k-i,k-m}^T,$$

or, equivalently, from (2.17) for $i = m$ and (2.18), we have that

$$\mathcal{T}_{k,k-i} - D_k^{\overline{\Theta}} \mathcal{H}_k \mathcal{G}_{k,k-i} = -\Psi_{k,k-m} \mathcal{T}_{k-i,k-m}^T, \quad i < m, \tag{2.19}$$

where $\Psi_{k,k-m} = (Cov[C_k^{\Theta}, C_{k-m}^{\Theta}] \circ \left(\mathcal{H}_k \mathbb{F}_{k,k-m} \mathcal{E}_{k-m} \mathcal{H}_{k-m}^T \right) + R_{k,k-m}^V) \Pi_{k-m}^{-1}$.

Hence, substituting (2.19) into (2.16) and using (2.7) for $\widehat{X}_{k/k-1}$, the following expression for the one-stage predictor is obtained

$$\widehat{Y}_{k/k-1} = D_k^{\overline{\Theta}} \mathcal{H}_k \widehat{X}_{k/k-1} + \Psi_{k,k-m} \left[\nu_{k-m} - \sum_{i=1}^{m-1} \mathcal{T}_{k-i,k-m}^T \Pi_{k-i}^{-1} \nu_{k-i} \right], \; k > m, \tag{2.20}$$

from which expression (2.11) is immediately deduced.

Formula (2.12) for the matrices $\mathcal{T}_{k,k-i}$ is obtained using (2.4), (2.19), and taking into account that $\mathcal{G}_{k,k-i} = \mathbb{F}_{k,k-i} \mathcal{G}_{k-i,k-i}$, which is immediately clear from (2.3).

Now, expression (2.13) for the matrix $\mathcal{G}_{k,k} = E[X_k Y_k^T] - E[X_k \widehat{Y}_{k/k-1}^T]$ is derived. From (2.4) and property (v), it is clear that $E[X_k Y_k^T] = \mathcal{E}_k \mathcal{H}_k^T D_k^{\overline{\Theta}}$, $\forall k \geq 1$. To calculate $E[X_k \widehat{Y}_{k/k-1}^T]$ we replace $\widehat{Y}_{k/k-1}$ by its expressions (2.15) and (2.20), thus considering two cases:

(a) For $k \leq m$, $E[X_k \widehat{Y}_{k/k-1}^T] = E[X_k \widehat{X}_{k/k-1}^T] \mathcal{H}_k^T D_k^{\overline{\Theta}}$ and by applying the OPL, $E[X_k \widehat{X}_{k/k-1}^T] = \mathcal{E}_k - \Sigma_{k/k-1}$ where $\Sigma_{k/k-1}$ is the prediction error covariance matrix. Hence

$$E[X_k \widehat{Y}_{k/k-1}^T] = \left(\mathcal{E}_k - \Sigma_{k/k-1} \right) \mathcal{H}_k^T D_k^{\overline{\Theta}}, \quad k \leq m.$$

(*b*) For $k > m$, it follows that

$$E[X_k\widehat{Y}_{k/k-1}^T] = E[X_k\widehat{X}_{k/k-1}^T]\mathcal{H}_k^T D_k^{\overline{\Theta}} + E[X_k\nu_{k-m}^T]\Psi_{k,k-m}^T - E\left[X_k\left(\sum_{i=1}^{m-1}\mathcal{T}_{k-i,k-m}^T\Pi_{k-i}^{-1}\nu_{k-i}\right)^T\right]\Psi_{k,k-m}^T,$$

hence, using again the OPL and since $E[X_k\nu_{k-i}^T] = \mathcal{G}_{k,k-i}$, for $1 \leq i \leq m$, it is deduced that

$$E[X_k\widehat{Y}_{k/k-1}^T] = \left(\mathcal{E}_k - \Sigma_{k/k-1}\right)\mathcal{H}_k^T D_k^{\overline{\Theta}} + \mathcal{G}_{k,k-m}\Psi_{k,k-m}^T - \sum_{i=1}^{m-1}\mathcal{G}_{k,k-i}\Pi_{k-i}^{-1}\mathcal{T}_{k-i,k-m}\Psi_{k,k-m}^T, \quad k > m.$$

By substraction of the above expectations, and taking into account that $\mathcal{G}_{k,k-i} = \mathbb{F}_{k,k-i}\mathcal{G}_{k-i,k-i}$ for $i \leq m$, expression (2.13) for $\mathcal{G}_{k,k}$ is obtained.

From (2.3), the expression for the prediction error covariance matrix, $\Sigma_{k/k-1}$ is immediately clear and, from (2.9), the expression for the filtering error covariance matrix, $\Sigma_{k/k}$, is also obvious.

Finally, we prove expression (2.14) for the innovation covariance matrix $\Pi_k = E[Y_kY_k^T] - E[\widehat{Y}_{k/k-1}\widehat{Y}_{k/k-1}^T]$. From (2.4) and using (2.5), we have that

$$E[Y_kY_k^T] = E[C_k^{\Theta}C_k^{\Theta T}] \circ \left(\mathcal{H}_k\mathcal{E}_k\mathcal{H}_k^T\right) + R_{k,k}^V, \quad k \geq 1.$$

Due to the correlation assumption of the Bernoulli variables θ_k, we need to distinguish two cases to calculate $E[\widehat{Y}_{k/k-1}\widehat{Y}_{k/k-1}^T]$. For this purpose, expressions (2.15) and (2.20) for $\widehat{Y}_{k/k-1}$, along with formula (2.5) and the OPL, are used:

(*a*) For $k \leq m$, the following identity holds

$$E[\widehat{Y}_{k/k-1}\widehat{Y}_{k/k-1}^T] = \left(E[C_k^{\Theta}]E[C_k^{\Theta T}]\right) \circ \left(\mathcal{H}_k(\mathcal{E}_k - \Sigma_{k/k-1})\mathcal{H}_k^T\right).$$

(*b*) For $k > m$, after some manipulations, we deduce that

$$\begin{aligned} E[\widehat{Y}_{k/k-1}\widehat{Y}_{k/k-1}^T] = &\left(E[C_k^{\Theta}]E[C_k^{\Theta T}]\right) \circ \left(\mathcal{H}_k(\mathcal{E}_k - \Sigma_{k/k-1})\mathcal{H}_k^T\right) \\ &+ \Psi_{k,k-m}\Pi_{k-m}\Psi_{k,k-m}^T \\ &+ \Psi_{k,k-m}\sum_{i=1}^{m-1}\mathcal{T}_{k-i,k-m}^T\Pi_{k-i}^{-1}\mathcal{T}_{k-i,k-m}\Psi_{k,k-m}^T \\ &- D_k^{\overline{\Theta}}\mathcal{H}_k\left(\mathcal{G}_{k,k} - \Sigma_{k/k-1}\mathcal{H}_k^T D_k^{\overline{\Theta}}\right) \\ &- \left(\mathcal{G}_{k,k}^T - D_k^{\overline{\Theta}}\mathcal{H}_k\Sigma_{k/k-1}\right)\mathcal{H}_k^T D_k^{\overline{\Theta}}. \end{aligned}$$

So, from the above expectations, expression (2.14) for the innovation covariance matrix Π_k is obtained. □

Theorem 2.4.2 *For each fixed $k \geq 1$, the quadratic fixed-point smoothers, $\widehat{x}_{k/k+N}^q$, of the state x_k are given by*

$$\widehat{x}_{k/k+N}^q = \Upsilon\widehat{X}_{k/k+N} + \overline{x}_k, \quad N \geq 1,$$

where Υ is the operator which extracts the first n entries of $\widehat{X}_{k/k+N}$, the linear fixed-point smoothers of the augmented state X_k, which are calculated as

$$\widehat{X}_{k/k+N} = \widehat{X}_{k/k+N-1} + \mathcal{G}_{k,k+N}\ \Pi_{k+N}^{-1}\ \nu_{k+N}, \quad N \geq 1, \tag{2.21}$$

whose initial condition is the filter $\widehat{X}_{k/k}$, given in Theorem 2.4.1.
The matrices $\mathcal{G}_{k,k+N}$ are obtained by

$$\begin{aligned} &\mathcal{G}_{k,k+N} = \left(\mathcal{E}_k\mathbb{F}_{k+N,k}^T - M_{k,k+N-1}\mathcal{F}_{k+N-1}^T\right)\mathcal{H}_{k+N}^T D_{k+N}^{\overline{\Theta}}, \quad k \leq m-N, \quad N \geq 1, \\ &\mathcal{G}_{k,k+N} = \left(\mathcal{E}_k\mathbb{F}_{k+N,k}^T - M_{k,k+N-1}\mathcal{F}_{k+N-1}^T\right)\mathcal{H}_{k+N}^T D_{k+N}^{\overline{\Theta}} \\ &\qquad - \left(\mathcal{G}_{k,k+N-m} - \sum_{i=1}^{m-1}\mathcal{G}_{k,k+N-i}\Pi_{k+N-i}^{-1}\mathcal{T}_{k+N-i,k+N-m}\right) \\ &\qquad \times \Psi_{k+N,k+N-m}^T, \quad k > m-N, \quad N \geq 1, \end{aligned} \tag{2.22}$$

where the matrices $M_{k,k+N}$ satisfy the following recursive formula:

$$\begin{aligned} &M_{k,k+N} = M_{k,k+N-1}\mathcal{F}_{k+N-1}^T + \mathcal{G}_{k,k+N}\Pi_{k+N}^{-1}\mathcal{G}_{k+N,k+N}^T, \quad N \geq 1, \\ &M_{k,k} = \mathcal{E}_k - \Sigma_{k/k}. \end{aligned} \tag{2.23}$$

The fixed-point smoothing error covariance matrix, $\Sigma_{k/k+N}$, satisfies

$$\Sigma_{k/k+N} = \Sigma_{k/k+N-1} - \mathcal{G}_{k,k+N}\Pi_{k+N}^{-1}\mathcal{G}_{k,k+N}^T, \quad N \geq 1, \tag{2.24}$$

with the filtering error covariance matrix, $\Sigma_{k/k}$, as initial condition.
The innovations ν_{k+N}, their covariance matrices Π_{k+N}, the matrices $\mathcal{T}_{k+N,k+N-i}$, $\Psi_{k+N,k+N-m}$, $\mathcal{E}_k$ and $\Sigma_{k/k}$ are given in Theorem 2.4.1.

Proof. From the general expression (2.7), for each fixed $k \geq 1$, the recursive relation (2.21) is immediately clear.

Next, to prove (2.22) for $\mathcal{G}_{k,k+N} = E[X_k Y_{k+N}^T] - E[X_k \widehat{Y}_{k+N/k+N-1}^T]$, it is necessary to calculate both expectations.

On the one hand, from Equation (2.4) and using that D_{k+N}^{Θ} and V_{k+N} are independent of X_k, we obtain

$$E[X_k Y_{k+N}^T] = \mathcal{E}_k \mathbb{F}_{k+N,k}^T \mathcal{H}_{k+N}^T D_{k+N}^{\overline{\Theta}}, \quad N \geq 1.$$

On the other hand, based on expression (2.11) for ν_{k+N}, two options must be considered:

(a) For $k \leq m - N$, using (2.10) for $\widehat{X}_{k+N/k+N-1}$, we have that

$$E\left[X_k \widehat{Y}_{k+N/k+N-1}^T\right] = M_{k,k+N-1}\mathcal{F}_{k+N-1}^T \mathcal{H}_{k+N}^T D_{k+N}^{\overline{\Theta}},$$

where $M_{k,k+N-1} = E\left[X_k \widehat{X}_{k+N-1/k+N-1}^T\right]$.

(b) For $k > m - N$, a similar reasoning to the above one leads to

$$\begin{aligned} E[X_k \widehat{Y}_{k+N/k+N-1}^T] = M_{k,k+N-1}\mathcal{F}_{k+N-1}^T \mathcal{H}_{k+N}^T D_{k+N}^{\overline{\Theta}} + \Bigg(&\mathcal{G}_{k,k+N-m} \\ &- \sum_{i=1}^{m-1} \mathcal{G}_{k,k+N-i}\Pi_{k+N-i}^{-1}\mathcal{T}_{k+N-i,k+N-m} \Bigg) \Psi_{k+N,k+N-m}^T. \end{aligned}$$

Then, the replacement of the above expectations in $\mathcal{G}_{k,k+N}$ leads to (2.22).

The recursive relation (2.23) for $M_{k,k+N} = E\left[X_k \widehat{X}^T_{k+N/k+N}\right]$ is immediately clear from (2.9) for $\widehat{X}_{k+N/k+N}$ and its initial condition $M_{k,k} = E[X_k \widehat{X}_{k/k}]$ is calculated taking into account that, from the orthogonality, $E[X_k \widehat{X}^T_{k/k}] = E[\widehat{X}_{k/k} \widehat{X}^T_{k/k}] = \mathcal{E}_k - \Sigma_{k/k}$.

Finally, since $\Sigma_{k/k+N} = E\left[X_k X_k^T\right] - E\left[\widehat{X}_{k/k+N} \widehat{X}^T_{k/k+N}\right]$, using (2.21) and taking into account that $\widehat{X}_{k/k+N-1}$ is uncorrelated with ν_{k+N}, expression (2.24) is deduced. □

Remark 4. The first $n \times n$ blocks of the error covariance matrices $\Sigma_{k/k+N}$, $N \geq 0$, provide the covariance matrices of the quadratic smoothing and filtering errors, thus providing a measure of the accuracy of the respective quadratic estimators.

2.5 Numerical simulation example

In this section, a numerical simulation example is shown to illustrate the feasibility of the quadratic estimation algorithms proposed in this paper. We ran a program in Matlab simulating at each iteration the state and the measured values and computing the linear [8] and quadratic filter and fixed-point smoothers, as well as the corresponding estimation error covariance matrices.

Consider a scalar state process, $\{x_k;\, k \geq 1\}$ generated by a first-order autoregressive model with missing measurements coming from two sensors and perturbed by additive and multiplicative noises. Specifically, the following model is considered:

$$\begin{aligned} x_k &= \left(1 + 0.2 \sin\left(\frac{(k-1)\pi}{50}\right)\right) 0.8 x_{k-1} + w_{k-1}, \quad k \geq 1 \\ y_k^i &= \theta_k^i x_k^i + v_k^i, \quad k \geq 1, \quad i = 1, 2, \end{aligned}$$

where the initial state x_0 is a zero-mean Gaussian variable with variance $P_0 = 0.1$, the process $\{w_k;\, k \geq 0\}$ is a zero-mean white Gaussian noise with variance $Q_k =$

0.36, $\forall k \geq 0$ and the noise processes $\{v_k^i;\, k \geq 1\}$, $i = 1, 2$ are zero-mean white sequences with the following probability distributions:

$$P[v_k^1 = -8] = \frac{1}{8}, \quad P[v_k^1 = \frac{8}{7}] = \frac{7}{8}, \quad \forall k \geq 1,$$
$$P[v_k^2 = 1] = \frac{15}{18}, \quad P[v_k^2 = -3] = \frac{2}{18}, \quad P[v_k^2 = -9] = \frac{1}{18}, \quad \forall k \geq 1,$$

and variances given by $R_k^1 = 64/7$ and $R_k^2 = 19/3$, $\forall k \geq 1$, respectively.

To describe the phenomenon of missing measurements according to our theoretical model (the uncertainty at time k depends only on the uncertainty at the previous time $k - m$), the variables $\{\theta_k^i;\, k \geq 1\}$, $i = 1, 2, \ldots, r$ are defined from two independent sequences of independent Bernoulli random variables, $\{\gamma_k^i;\, k \geq 1\}$, $i = 1, 2$ with $P[\gamma_k^i = 1] = \gamma_i$. Specifically, the variables θ_k^i are defined as follows:

$$\theta_k^i = 1 - \gamma_{k+m}^i(1 - \gamma_k^i), \quad i = 1, 2.$$

Note that $\theta_k^i = 0$ when $\gamma_{k+m}^i = 1$ and $\gamma_k^i = 0$, and consequently, $\theta_{k+m}^i = 1$; this fact implies that, if the state is missing at time k, after $k + m$ sampling times the output measurement necessarily contains the state. Therefore, there cannot be more than m consecutive measured data consisting of noise only.

Since the variables γ_k^i and γ_s^i are independent, θ_k^i and θ_s^i are also independent for $|k - s| \neq 0, m$. The mean of these variables is $\overline{\theta}^i = 1 - \gamma_i(1 - \gamma_i)$ and the covariance function is given by

$$E[(\theta_k^i - \overline{\theta}^i)(\theta_s^i - \overline{\theta}^i)] = \begin{cases} 0, & |k - s| \neq 0, m, \\ -(1 - \overline{\theta}^i)^2, & |k - s| = m, \\ \overline{\theta}^i(1 - \overline{\theta}^i), & |k - s| = 0. \end{cases}$$

To analyze and compare the effectiveness of the proposed estimators, two hundred iterations of the proposed algorithms have been performed, and the linear and quadratic estimation error variances have been calculated for different values of the uncertainty probability and several values of m. Note that, for $i = 1, 2$, the roles played by γ_i and $1 - \gamma_i$ can be interchanged without affecting the means

and covariance functions of the random variables θ_k^i; so, henceforth, only the case $\gamma_i \leq 0.5$ will be considered. In such case, note that $\overline{\theta}^i$ (the probability that the observations coming from sensor i contain the state) is a decreasing function of γ_i, for $i = 1, 2$, whose maximum value, $\overline{\theta}^i = 1$ (for $\gamma_i = 0$), corresponds to the case when all the observations coming from sensor i contain the state, i.e. there are no missing measurements.

Firstly, assuming that the Bernoulli variables θ_k^i, $i = 1, 2$, are correlated at instants that differ $m = 3$ units of time and considering fixed values of the probabilities $\gamma_1 = 0.2$ and $\gamma_2 = 0.1$, the linear and quadratic filtering and fixed-point smoothing error variances ($N = 2, 5$) are displayed in Fig. 2.1. This figure shows, on the one hand, that the quadratic estimation error variances are significantly smaller than the linear ones and, on the other hand, that the estimation accuracy of the smoothing error variances are higher than those of the filter and, in turn, that the performance of the fixed-point smoothers improves as the number of available observations increases.

Now, in order to show more precisely the dependence of the error variances on the values γ_1 and γ_2, Fig. 2.2 presents the linear and quadratic filtering error variances, at a fixed iteration (namely, $k = 200$) for $m = 3$, versus γ_1 (for constant values of γ_2). From this figure it is gathered that, as γ_1 increases (equivalently, as the probability that the observations contain the state decreases), the filtering error variances are greater and, hence, the performance of the filters is worse. Also, agreeing with the comments on the previous figure, we conclude again that the quadratic filtering estimators are better than the linear ones. A similar study about the filtering error variances versus γ_2 (for constant values of γ_1) leads to analogous comments, so it is omitted for brevity.

Finally, for $\gamma_1 = 0.2$ and $\gamma_2 = 0.4$, the performance of the quadratic estimators is compared for different values of m. Specifically, for $m = 2$, 3, 4, 5, 6, the linear and quadratic estimation error variances are shown in Table 2.1. From this table

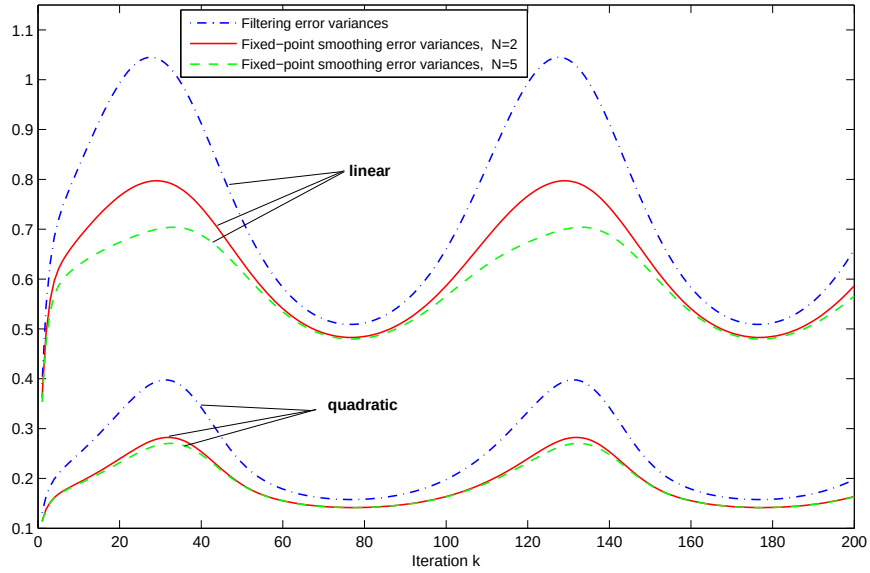


Figure 2.1: Linear and quadratic filtering and fixed-point smoothing error variances for $\gamma_1 = 0.2$, $\gamma_2 = 0.1$, when $m = 3$.

it is deduced that the estimators are more accurate as the values of m are lower, i.e., a greater distance between the instants at which the variables are correlated (which means that the state can be missing in more consecutive observations) yields worse estimators. Moreover, this table corroborates that the quadratic estimators perform quite better than the linear estimators.

Algorithm	Error variances	$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 6$
Linear	Filter	1.1568	1.1604	1.1633	1.1657	1.1675
	Smoother	0.9025	0.9085	0.9102	0.9116	0.9127
Quadratic	Filter	0.4925	0.5027	0.5110	0.5177	0.5229
	Smoother	0.3455	0.3577	0.3591	0.3619	0.3642

Table 2.1: Linear and quadratic filtering and fixed-point smoothing error variances ($N = 2$) for $\gamma_1 = 0.2$, $\gamma_2 = 0.4$ at $k = 30$ when $m = 2, 3, 4, 5, 6$.

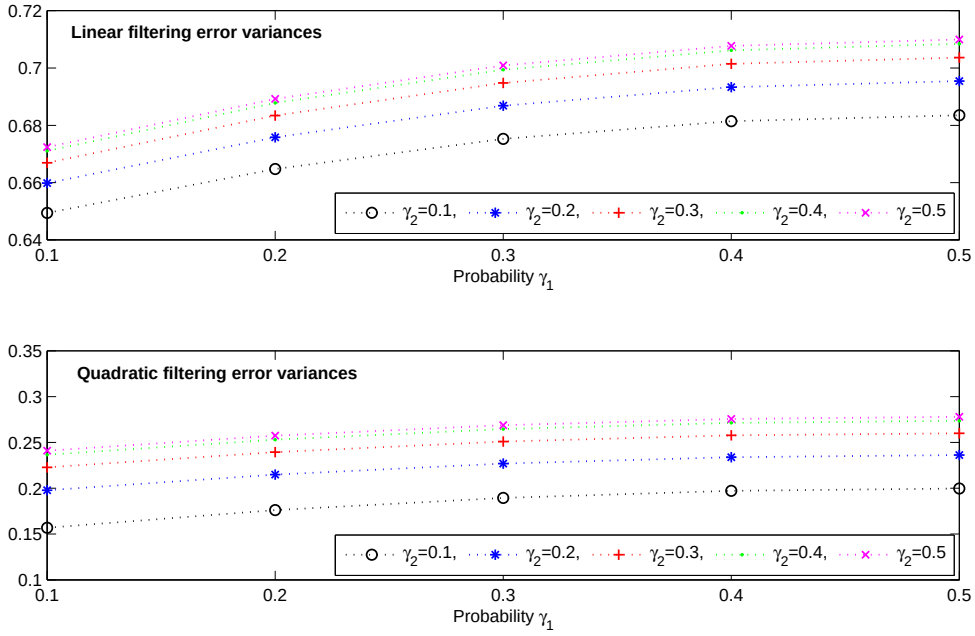


Figure 2.2: Linear and quadratic filtering error variances at $k = 200$ versus γ_1, with γ_2 varying from 0.1 to 0.5 when $m = 3$.

2.6 Conclusions

The least-squares quadratic filtering and fixed-point smoothing problems, for linear discrete-time stochastic systems with missing measurements coming from multiple sensors have been addressed by using centralized fusion method. At each sensor, the phenomenon of missing measurements is modelled by Bernoulli variables whose values –one or zero– indicate whether the state is present or missing in the corresponding measurement; such variables are assumed to be correlated at instants that differ m sampling times. Real applications with these features are, for example, signal transmission models in which any failure in the transmission is detected and the old sensor is replaced after m instants of time, thus avoiding the possibility of missing signal in $m + 1$ consecutive observations.

The quadratic estimation problem considered in the current paper provides a generalization of the result established in [4] and it is useful in several problems of signal processing, such as signal prediction, detection and control, as well as image restoration problems. The theoretical results are illustrated by a numerical simulation example, in which a scalar state process is generated by a first-order autoregressive model with missing measurements coming from two sensors. Linear and quadratic error variances are shown for different uncertainty probabilities and values of m. On the one hand, these results confirm that the quadratic estimators are more accurate than the linear ones and, on the other hand, that the fixed-point smoothing estimators are more effective than the filtering ones. The example also shows that, as the uncertainty probability decreases, the performance of the filters is better and finally that, as the values of m are lower, the estimators are more accurate since the state can be missing in less consecutive observations.

Acknowledgements

This research is supported by Ministerio de Ciencia e Innovación (Programa FPU and grant No. MTM2011-24718) and Junta de Andalucía (grant No. P07-FQM-02701).

References

[1] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Polynomial filtering with uncertain observations in stochastic linear systems, *International Journal of Modelling and Simulation,* 23 (2003), 22–28.

[2] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Linear and quadratic estimation using uncertain observations from multiple sensors with correlated uncertainty, *Signal Processing,* 91 (2011), 330–337.

[3] A. De Santis, A. Germani and M. Raimondi, Optimal quadratic filtering of linear discrete-time non-Gaussian systems, *IEEE Transactions on Automatic Control,* 40 (1995), 1274–1278.

[4] I. García-Garrido, J. Linares-Pérez, R. Caballero-Águila and A. Hermoso-Carazo, A Solution to the Filtering Problem for Stochastic Systems with Multi-Sensor Uncertain Observations, *International Mathematical Forum,* 7 (2012), 887–903.

[5] F.O. Hounkpevi and E.E. Yaz, Robust minimum variance linear state estimators for multiple sensors with different failure rates, *Automatica,* 43 (2007), 1274–1280.

[6] R. Huang and G.V. Záruba, Incorporating data from multiple sensors for localizing nodes in mobile ad hoc networks, *IEEE Transactions on Mobile Computing,* 6(9) (2007), 1090–1104.

[7] T. Kailath, A. Sayed and B. Hassibi, *Linear Estimation,* Prentice Hall, New Jersey, 2000.

[8] J. Linares-Pérez, R. Caballero-Águila and I. García-Garrido, Design of estimation algorithms from an innovation approach in linear discrete-time stochastic systems with uncertain observations, In: *Stochastic Modeling and Control* (pp. 1–22), InTech, Croatia, 2012. http://dx.doi.org/10.5772/2567

[9] J. Ma and S.L. Sun, Centralized fusion estimators for multi-sensor systems with multiplicative noises and missing measurements, *Journal of Networks,* 7 (10) (2012), 1538–1545.

[10] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo, J.D. Jiménez-López and J. Linares-Pérez, Least-squares ν th-order polynomial estimation of signals from observations affected by non-independent uncertainty, *Applied Mathematics and Computation,* 176 (2006), 642–653.

Chapter 3

Optimal fusion filtering in multi-sensor stochastic systems with missing measurements and correlated noises

Reference

Caballero-Águila, R., García-Garrido, I. and Linares-Pérez, J. (2013).
Optimal fusion filtering in multi-sensor stochastic systems with missing measurements and correlated noises.
Mathematical Problems in Engineering, vol. 2013, Article ID 418678, 14 pages.
DOI:10.1155/2013/418678

Abstract

The optimal least-squares linear estimation problem is addressed for a class of discrete-time multi-sensor linear stochastic systems with missing measurements and autocorrelated and cross-correlated noises. The stochastic uncertainties in the measurements coming from each sensor (missing measurements) are described by scalar random variables with arbitrary discrete probability distribution over the in-

terval [0,1]; hence, at each single sensor the information might be partially missed and the different sensors may have different missing probabilities. The noise correlation assumptions considered are: (i) the process noise and all the sensor noises are one-step autocorrelated; (ii) different sensor noises are one-step cross-correlated; and (iii) the process noise and each sensor noise are two-step cross-correlated. Under these assumptions and by an innovation approach, recursive algorithms for the optimal linear filter are derived by using the two basic estimation fusion structures; more specifically, both centralized and distributed fusion estimation algorithms are proposed. The accuracy of these estimators is measured by their error covariance matrices, which allow us to compare their performance in a numerical simulation example that illustrates the feasibility of the proposed filtering algorithms and shows a comparison with other existing filters.

Contents

3.1 Introduction

For a long time, the least-squares (LS) estimation problem in linear stochastic systems from measurements perturbed by additive noises has received considerable attention in the scientific community due to its wide applicability in many practical situations (for example, video and laser tracking systems, satellite navigation, radar and meteorological applications, etc. [1]). As it is well known, one of the major contributions made to solve this problem is the Kalman filter, which provides a recursive algorithm for the optimal LS estimator when the additive white noises and the initial state are Gaussian and mutually independent (or, equivalently, uncorrelated due to the gaussianity assumption) and, therefore, the optimal LS estimator is the optimal LS linear estimator. From the publication of the Kalman filter [2] in 1960, numerous results and several solution methods have been reported in the literature to address the state estimation problem from noisy observations, which depend on models representing possible relationships between the unknown state and the observable variables and also on the noise processes assumptions.

Specifically, during the past decades, there has been an increasing interest in the filtering problem in multi-sensor systems, where sensor networks are used to obtain the whole available information on the system state and its estimation must be carried out from the observations provided by all the sensors. A basic matter for this class of systems is how to fuse the measurement data from the different sensors to address the estimation problem. Commonly, two methods are used to process the measured data coming from multiple sensors: centralized and distributed fusion methods. In the centralized fusion method all the measured data from sensors are communicated to the fusion center for being processed; nevertheless, as is widely known, centralized estimators have many computational disadvantages, which motivate the research into other fusion methods. In the distributed fusion method, each sensor estimates the state based on its own single measurement data, and then it sends such estimate to the fusion center for fusion according to a certain

information fusion criterion. Although the use of sensor networks offers several advantages, the unreliable network characteristics usually cause problems during data transmission from sensors to the fusion center, such as missing measurements, random communication packet losses and/or delays. Taking into account these network uncertainties, the models representing the relationships between the state and measurements do not allow to apply the Kalman filter, and modifications of conventional estimation algorithms have been proposed (see for example, [3]-[9] and references therein).

As in the Kalman filter, independent white noises are considered in all the mentioned papers; however, this assumption may not be realistic and can be a limitation in many real-world problems in which noise correlation may be present. This problem arises, for example, when a target is taking an electronic countermeasure, e.g. noise jamming [10], or if the process noise and the sensor measurement noises are dependent on the system state, then there may be cross-correlation between different sensor noises and cross-correlation between process noise and sensor noises. Also, if all the sensors are observed in the same noisy environment, the measurement noises of different sensors are usually correlated.

For these reasons, the estimation problem in systems with correlated noises has received significant research interest in recent years. For example, the optimal Kalman filtering fusion problem in systems with cross-correlated sensor noises is addressed in [10], while [11] and [12] study the same problem in systems with cross-correlated process noises and measurement noises; in these papers correlated noises at the same sampling time are considered. In general, the assumption of correlation and cross-correlation of the noise process and measurement noises in different sampling times makes difficult the identification of optimal estimators; this limitation has encouraged a wider research into suboptimal Kalman-type estimation problems. In [13], a Kalman-type recursive filter is presented for systems with finite-step correlated process noises, and the filtering problem with multi-step

correlated process and measurement noises is investigated in [14]. The optimal robust non-fragile Kalman-type recursive filtering problem is studied in [15] for a class of uncertain systems with finite-step autocorrelated measurement noises and multiple packet dropouts. The problem of distributed weighted robust Kalman filter fusion is studied in [16] for a class of uncertain systems with autocorrelated and cross-correlated noises. In [17], a stochastic singular system with correlated noises at the same sampling time is transformed into an equivalent nonsingular system with correlated noises at the same and neighboring sampling times. Also, in [18], an augmented parameterized system with correlated noises at the same and neighboring sampling times is used to describe the sensor delay, packet dropout and uncertain observation phenomenons.

On the other hand, as noted above, the use of communication networks for transmitting measured data motivates the need of considering stochastic uncertainties. Missing measurements have been widely treated due to its applicability to model a large class of real-world problems, such as fading phenomena in propagation channels, target tracking or, in general, situations where there exist intermittent failures in the observation mechanism, accidental loss of some measurements, or inaccessibility of the data during certain times. The state estimation problem from missing measurement transmitted by multiple sensors has been studied based on the assumption that all the sensors are identical (see, e.g., [19]-[22]); however, this assumption can be unreasonable since some real systems usually involve multiple sensors with different characteristics. Recently, the filtering problem using missing measurements whose statistical properties are assumed not to be the same in all the sensors has been addressed by several authors under different approaches and hypotheses on the processes involved (see, e.g., [23]-[27]). In all the above papers, Bernoulli random variables are used to model the missing measurements phenomenon, and hence, it is assumed that the measurement signal is either completely lost (if the corresponding Bernoulli variable takes the value zero) or

successfully transferred (when the Bernoulli variable is equal to one). Recently, this missing measurement model has been generalized considering any discrete distribution on the interval $[0, 1]$, which allows to cover some practical applications where only partial information is missing (see [28], [29] and references therein).

Motivated by the above considerations, our attention is focused on investigating the optimal LS linear centralized and distributed fusion estimation problems in multi-sensor systems with missing measurements and autocorrelated and cross-correlated noises. In each sensor, the missing measurement phenomenon is governed by a scalar random variable with arbitrary discrete probability distribution over the interval [0,1], and the different sensors may have different missing probabilities. Assume that the process noise and all the sensor noises are one-step autocorrelated; different sensor noises are one-step cross-correlated; and the process noise and each sensor noise are two-step cross-correlated. *This paper makes a two-fold substantial novel contribution: (1) Unlike most previous results with correlated noises, in which suboptimal Kalman-type estimators are proposed, in this paper optimal LS linear estimators are obtained by using an innovation approach, which provides a simple derivation of the estimation algorithms due to the fact that the innovations constitute a white process; and (2) our missing measurement model considers at each sensor the possibility of observations containing only partial information about the state, or even only noise.*

The paper is organized as follows. In Section 3.2 the system model with autocorrelated and cross-correlated noises and missing measurements coming from multiple sensors is described. Also, the suitable properties on the state and noise processes are specified and a brief description of the innovation approach to the optimal LS linear estimation problem is included. In Section 3.3 a recursive algorithm for the centralized optimal linear filter is presented for the considered model (the derivation has been deferred to Appendix A.1). Next, in Section 3.4, the local LS linear filters and their corresponding error covariance matrices between any

two local estimates are provided, then the distributed optimal weighted fusion estimators and their error covariance matrices are obtained by applying the optimal information fusion criterion weighted by matrices in the linear minimum variance sense. Finally, in Section 3.5, a numerical simulation example is presented to show the effectiveness of the estimation algorithms proposed in the current paper, and some conclusions are drawn in Section 3.6.

Notation: The notation used throughout the paper is standard. For any matrix A, the notation symbols A^T and A^{-1} represent its transpose and inverse, respectively; $\mathbb{R}^n$ denotes the n-dimensional Euclidean space and $\mathbb{R}^{m\times n}$ is the set of all real matrices of dimension $m \times n$. The shorthand $Diag(a_1, \ldots, a_r)$ denotes a diagonal matrix whose diagonal entries are $a_1, \ldots, a_r$. If the dimensions of matrices are not explicitly stated, they are assumed to be compatible for algebraic operations. δ_{k-s} is the Kronecker delta function, which is equal to one, if $k = s$, and zero otherwise. Moreover, for arbitrary random vectors α and β, we will denote $\mathrm{Cov}[\alpha, \beta] = E\left[(\alpha - E[\alpha])\,(\beta - E[\beta])^T\right]$ and $\mathrm{Cov}[\alpha] = \mathrm{Cov}[\alpha, \alpha]$, where $E[\cdot]$ stands for the mathematical expectation operator. Finally, $\widehat{\alpha}$ denotes the estimator of α and $\widetilde{\alpha} = \alpha - \widehat{\alpha}$ the estimation error.

3.2 Problem formulation

Our aim is to obtain recursive algorithms for the optimal LS linear filtering problem in a class of discrete-time stochastic systems with missing measurements coming from multiple sensors, by using centralized and distributed fusion methods. In this section, firstly the system model and the assumptions about the state and noise processes are presented and, secondly, the optimal LS linear estimation problem is formulated using an innovation approach.

3.2.1 Stochastic system model

Consider a discrete-time linear stochastic system with autocorrelated and cross-correlated noises and missing measurements coming from r sensors. The phenomenon of missing measurements occurs randomly and, for each sensor, a different sequence of scalar random variables with discrete distribution over the interval $[0, 1]$ is used to model this phenomenon. Specifically, the following system is considered:

$$x_k = F_{k-1}x_{k-1} + w_{k-1}, \quad k \geq 1. \tag{3.1}$$

where $x_k \in \mathbb{R}^n$ is the state, $\{w_k;\ k \geq 0\}$ is the process noise, and F_k, for $k \geq 0$, are known matrices with compatible dimensions.

Consider r sensors which, at any time k, provide scalar measurements of the system state, perturbed by additive and multiplicative noises according to the following model:

$$y_k^i = \theta_k^i H_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2, \ldots, r \tag{3.2}$$

where $\{y_k^i;\ k \geq 1\}$ are the measured data; $\{v_k^i;\ k \geq 1\}$ are measurement noises; $\{\theta_k^i;\ k \geq 1\}$ are scalar random variables sequences; H_k^i, for $k \geq 1$, are known time-varying matrices with compatible dimensions; superscript i denotes the i-th sensor, and r is the number of sensors.

Next, the statistical properties assumed about the initial state and noise processes involved in (3.1) and (3.2) are specified:

(i) The initial state x_0 is a random vector with $E[x_0] = \overline{x}_0$ and $Cov[x_0] = P_0$.

(ii) The process noise, $\{w_k;\ k \geq 0\}$, and the measurement noises, $\{v_k^i;\ k \geq 1\}$, $i = 1, 2, \ldots, r$, are zero-mean sequences with covariances and cross-covariances:

$$Cov[w_k, w_s] = Q_{k,k}\delta_{k-s} + Q_{k,s}\delta_{k-s+1} + Q_{k,s}\delta_{k-s-1},$$

$$Cov[v_k^i, v_s^j] = R_{k,k}^{ij}\delta_{k-s} + R_{k,s}^{ij}\delta_{k-s+1} + R_{k,s}^{ij}\delta_{k-s-1},$$
$$Cov[w_k, v_s^i] = S_{k,k}^{i}\delta_{k-s} + S_{k,s}^{i}\delta_{k-s+1} + S_{k,s}^{i}\delta_{k-s+2}.$$

(iii) The multiplicative noises $\{\theta_k^i;\ k \geq 1\}$, $i = 1, 2, \ldots, r$, are white sequences of scalar variables with discrete distribution over the interval $[0, 1]$, with $E[\theta_k^i] = \overline{\theta}_k^i$ and $Var[\theta_k^i] = V_k^{\theta^i}$.

(iv) The initial state x_0 and the multiplicative noises $\{\theta_k^i;\ k \geq 1\}$, for $i = 1, 2, \ldots, r$, are mutually independent, and they are independent of the additive noises $\{w_k;\ k \geq 0\}$ and $\{v_k^i;\ k \geq 1\}$, for $i = 1, 2, \ldots, r$.

Remark 1. From assumption *(ii)* the following correlation properties of the additive noises are easily deduced:

- The noise vectors w_k and w_s are correlated at consecutive sampling times, $|k - s| = 1$, and independent otherwise; the covariance matrices of w_k with w_{k-1} and w_{k+1} are $Q_{k,k-1}$ and $Q_{k,k+1}$, respectively.

- For $i, j = 1, 2, \ldots, r$, the measurement noises v_k^i and v_s^j are cross-correlated at the same sampling time and at consecutive sampling times, $|k - s| = 0, 1$, and independent otherwise; the cross-covariances of v_k^i with v_k^j, v_{k-1}^j and v_{k+1}^j are $R_{k,k}^{ij}$, $R_{k,k-1}^{ij}$ and $R_{k,k+1}^{ij}$, respectively.

- For $i = 1, 2, \ldots, r$, the measurement noises v_k^i are correlated with the noise vectors w_s, for $s = k, k-1, k-2$, and independent otherwise; the cross-covariance matrices of v_k^i with w_k, w_{k-1} and w_{k-2} are $S_{k,k}^i$, $S_{k-1,k}^i$ and $S_{k-2,k}^i$, respectively.

The correlation conditions of the process noise and the measurement noises considered in this paper are the same as those in [16]. Systems with only finite-step correlated process noises or multi-step correlated process and measurement noises are considered in [13], [14] and [15], among others. The current study can be

extended to more general systems involving finite-step autocorrelated and cross-correlated noises with no difficulty, except for a greater complexity in the mathematical derivations.

Remark 2. From the state equation (3.1) and assumptions *(ii)* and *(iv)*, it is easy to deduce that $D_k = E[x_k x_k^T]$ is recursively calculated by:

$$\begin{aligned} &D_k = F_{k-1} D_{k-1} F_{k-1}^T + Q_{k-1,k-1} + F_{k-1} Q_{k-2,k-1} + Q_{k-1,k-2} F_{k-1}^T, \quad k \geq 2; \\ &D_1 = F_0 D_0 F_0^T + Q_{0,0}, \qquad D_0 = P_0 + \overline{x}_0 \overline{x}_0^T. \end{aligned} \tag{3.3}$$

Also, it is easy to see that the state x_k is correlated with the measurement noises v_k^i, for $i = 1, 2, \cdots, r$, and the expectations $E_k^i = E[x_k v_k^i]$ satisfy:

$$E_k^i = F_{k-1} S_{k-2,k}^i + S_{k-1,k}^i, \quad k \geq 2; \qquad E_1^i = S_{0,1}^i. \tag{3.4}$$

Remark 3. According to assumption *(iii)*, the scalar random variables θ_k^i take values over the interval $[0, 1]$ and they can satisfy any arbitrary discrete probability distribution over such interval, for instance, a Bernoulli distribution. Usually, Bernoulli random variables have been used to model the phenomenon of missing measurements (see, e.g., [25] and references therein), with $\theta_k^i = 1$ meaning that the state x_k is present in the measurement y_k^i coming from the i-th sensor at time k, while $\theta_k^i = 0$ means that the state is missing in the measured data at time k or, equivalently, that such observation only contains additive noise v_k^i. However, in practice, the information transmitted at a sampling time can usually be neither completely missing nor completely successful, but only part of the information can go through; in such situations, only partial information is missing and the proportion of missed data at one moment is a *fraction* other than 0 or 1 (see, e.g [28], [29] and references therein).

3.2.2 Stacked measurement equation

As noted above, our aim is to solve the optimal LS linear estimation problem of the state x_k based on the measurements $\{y_1^i, y_2^i, \ldots, y_k^i\}$, for $i = 1, 2, \ldots, r$, by using

centralized and distributed fusion methods to process the measured sensor data. The centralized fusion method considers that all the measurement data coming from r sensors are transmitted to a fusion center for being processed; for this purpose and to simplify the notation, the measurement equation (3.2) is rewritten in a stacked form as follows:

$$y_k = \Theta_k H_k x_k + v_k, \quad k \geq 1, \tag{3.5}$$

where $y_k = (y_k^1, \ldots, y_k^r)^T$, $v_k = (v_k^1, \ldots, v_k^r)^T$, $H_k = (H_k^{1T}, \ldots, H_k^{rT})^T$ and $\Theta_k = Diag(\theta_k^1, \ldots, \theta_k^r)$.

The following properties of the noises in (3.5) are easily inferred from the model assumptions *(ii)-(iv)* previously stated:

(I) The additive noise $\{v_k;\, k \geq 1\}$ is a zero-mean process satisfying:

$$Cov[v_k, v_s] = R_{k,k}\delta_{k-s} + R_{k,s}\delta_{k-s+1} + R_{k,s}\delta_{k-s-1},$$

$$Cov[w_k, v_s] = S_{k,k}\delta_{k-s} + S_{k,s}\delta_{k-s+1} + S_{k,s}\delta_{k-s+2},$$

where $R_{k,s} = \left(R_{k,s}^{ij}\right)_{i,j=1,2,\cdots,r}$ and $S_{k,s} = (S_{k,s}^1, \ldots, S_{k,s}^r)$.

(II) The state vector x_k and the measurement noise vector v_k are correlated with $E_k = E[x_k v_k^T]$ satisfying:

$$E_k = F_{k-1}S_{k-2,k} + S_{k-1,k}, \quad k \geq 2; \qquad E_1 = S_{0,1}. \tag{3.6}$$

(III) The random matrices $\{\Theta_k;\, k \geq 1\}$ satisfy $E[\Theta_k] = \overline{\Theta}_k = Diag(\overline{\theta}_k^1, \ldots, \overline{\theta}_k^r)$ and $E[(\Theta_k - \overline{\Theta}_k)^2] = Diag(V_k^{\theta^1}, \ldots, V_k^{\theta^r})$; also, denoting $\theta_k = (\theta_k^1, \ldots, \theta_k^r)^T$, it is clear that $Cov[\theta_k] = Diag(V_k^{\theta^1}, \ldots, V_k^{\theta^r})$.

Moreover, for any random matrix G independent of $\{\Theta_k;\, k \geq 1\}$, it is easily deduced that:

$$E[(\Theta_k - \overline{\Theta}_k)G(\Theta_k - \overline{\Theta}_k)] = Cov[\theta_k] \circ E[G], \tag{3.7}$$

where $\circ$ denotes the Hadamard product [23].

(IV) The initial state x_0 and $\{\Theta_k;\ k \geq 1\}$ are independent, and they are independent of $\{w_k;\ k \geq 0\}$ and $\{v_k;\ k \geq 1\}$.

3.2.3 Innovation approach to the optimal LS linear estimation problem

To address the optimal LS linear estimation problem of the state x_k based on the measurements $\{y_1^i, y_2^i, \ldots, y_k^i\}$, $i = 1, 2, \ldots, r$, the centralized and distributed fusion methods will be used. In both cases, recursive algorithms for the LS linear estimators will be established using an innovation approach and the orthogonal projection Lemma (OPL); more specifically we have the following.

Centralized fusion estimation problem. Our aim is to obtain the optimal LS linear filter, $\widehat{x}_{k/k}$, of the state x_k based on the measurements $\{y_1, y_2, \ldots, y_k\}$, given in (3.5), by recursive algorithms.

As known, the LS linear filter $\widehat{x}_{k/k}$ is the orthogonal projection of the state x_k over the linear space spanned by $\{y_1, y_2, \ldots, y_k\}$. These observations are generally non-orthogonal vectors, but the Gram-Schmidt orthogonalization procedure allows us to substitute them by a set of orthogonal vectors, called *innovations*, defined as the difference between each observation and its one-stage predictor. Due to the orthogonality property of the innovations and since the innovation process is uniquely determined by the observations, the LS linear filter, $\widehat{x}_{k/k}$, can be calculated as linear combination of the innovations; namely,

$$\widehat{x}_{k/k} = \sum_{s=1}^{k} \mathcal{X}_{k,s} \Pi_{s,s}^{-1} \mu_s, \quad k \geq 1, \tag{3.8}$$

where $\mu_s = y_s - \widehat{y}_{s/s-1}$ are the innovation vectors, with $\widehat{y}_{s/s-1}$ the one-stage observation predictor, $\Pi_{s,s} = E[\mu_s \mu_s^T]$, and $\mathcal{X}_{k,s} = E[x_k \mu_s^T]$.

Distributed fusion estimation problem. To address the distributed fusion estimation problem, firstly, recursive algorithms to obtain local LS linear filters, $\widehat{x}_{k/k}^i$,

for $i = 1, 2, \ldots, r$, and the error cross-covariance matrices between any two local estimates, are derived. Secondly, the distributed fusion filter, $\widehat{x}^D_{k/k}$, is established by applying the optimal information fusion criterion weighted by matrices in the linear minimum variance sense [30].

Analogously to (3.8), denoting $\mu^i_s = y^i_s - \widehat{y}^i_{s/s-1}$, $\Pi^{ii}_{s,s} = E[\mu^i_s \mu^i_s]$, and $\mathcal{X}^i_{k,s} = E[x_k \mu^i_s]$, the local filter $\widehat{x}^i_{k/k}$, is expressed as

$$\widehat{x}^i_{k/k} = \sum_{s=1}^{k} \mathcal{X}^i_{k,s} (\Pi^{ii}_{s,s})^{-1} \mu^i_s, \quad k \geq 1.$$

3.3 Optimal LS linear centralized fusion estimation

In this section a recursive algorithm for the centralized optimal (under the LS criterion) linear filter, $\widehat{x}_{k/k}$ is derived. Such algorithm is deduced using (3.8) and the OPL, and it is presented in Theorem 3.3.1. Firstly, in order to simplify the proof of Theorem 3.3.1, the following lemma is established.

Lemma 3.3.1 *Under assumptions (i)-(iv), the following results hold:*

$$\mathcal{W}_{k,k} = E\left[w_k \mu_k^T\right] = Q_{k,k-1} H_k^T \overline{\Theta}_k + S_{k,k}, \quad k \geq 1. \tag{3.9}$$

$$\mathcal{V}_{k,k-1} = E\left[v_k \mu_{k-1}^T\right] = S_{k-2,k}^T H_{k-1}^T \overline{\Theta}_{k-1} + R_{k,k-1}, \quad k \geq 2. \tag{3.10}$$

Proof. Since w_k is independent of $y_1, \ldots, y_{k-1}$, $E\left[w_k \widehat{y}^T_{k/k-1}\right] = 0$ and hence $\mathcal{W}_{k,k} = E\left[w_k y_k^T\right]$. Now, using (3.1) and (3.5), $\mathcal{W}_{k,k}$ can be calculated as follows:

$$\begin{aligned}
\mathcal{W}_{k,k} &= E\left[w_k \left(\Theta_k H_k x_k + v_k\right)^T\right] \\
&= E\left[w_k x_k^T\right] H_k^T \overline{\Theta}_k + S_{k,k} \\
&= E\left[w_k \left(F_{k-1} x_{k-1} + w_{k-1}\right)^T\right] H_k^T \overline{\Theta}_k + S_{k,k} \\
&= Q_{k,k-1} H_k^T \overline{\Theta}_k + S_{k,k}.
\end{aligned}$$

Taking into account that v_k is independent of $y_1, \ldots, y_{k-2}$, the calculation of $\mathcal{V}_{k,k-1}$ is similar to that of $\mathcal{W}_{k,k}$, and hence the proof is omitted. □

Theorem 3.3.1 *For the system model* (3.1) *and measurement model* (3.5)*, under assumptions (i)-(iv), the optimal LS linear filter* $\widehat{x}_{k/k}$ *is obtained as*

$$\widehat{x}_{k/k} = \widehat{x}_{k/k-1} + \mathcal{X}_{k,k}\Pi_{k,k}^{-1}\mu_k, \quad k \geq 1; \quad \widehat{x}_{0/0} = \overline{x}_0, \tag{3.11}$$

where the state predictor, $\widehat{x}_{k/k-1}$*, satisfies*

$$\widehat{x}_{k/k-1} = F_{k-1}\widehat{x}_{k-1/k-1} + \mathcal{W}_{k-1,k-1}\Pi_{k-1,k-1}^{-1}\mu_{k-1}, \quad k \geq 2; \quad \widehat{x}_{1/0} = F_0\widehat{x}_{0/0}. \tag{3.12}$$

The innovation, μ_k*, is given by*

$$\begin{aligned} \mu_k &= y_k - \overline{\Theta}_k H_k \widehat{x}_{k/k-1} - \mathcal{V}_{k,k-1}\Pi_{k-1,k-1}^{-1}\mu_{k-1}, \quad k \geq 2; \\ \mu_1 &= y_1 - \overline{\Theta}_1 H_1 \widehat{x}_{1/0}. \end{aligned} \tag{3.13}$$

The matrix $\mathcal{X}_{k,k} = E\left[x_k \mu_k^T\right]$ *is calculated by*

$$\begin{aligned} \mathcal{X}_{k,k} &= P_{k/k-1}H_k^T\overline{\Theta}_k + E_k - \mathcal{X}_{k,k-1}\Pi_{k-1,k-1}^{-1}\mathcal{V}_{k,k-1}^T, \quad k \geq 2; \\ \mathcal{X}_{1,1} &= P_{1/0}H_1^T\overline{\Theta}_1 + E_1, \end{aligned} \tag{3.14}$$

where $\mathcal{X}_{k,k-1} = E\left[x_k \mu_{k-1}^T\right]$ *satisfies*

$$\mathcal{X}_{k,k-1} = F_{k-1}\mathcal{X}_{k-1,k-1} + \mathcal{W}_{k-1,k-1}, \quad k \geq 2. \tag{3.15}$$

The prediction error covariance matrix, $P_{k/k-1}$*, is obtained by*

$$\begin{aligned} P_{k/k-1} &= F_{k-1}P_{k-1/k-1}F_{k-1}^T + Q_{k-1,k-1} + F_{k-1}\mathcal{J}_{k-1} + \mathcal{J}_{k-1}^T F_{k-1}^T \\ &\quad - \mathcal{W}_{k-1,k-1}\Pi_{k-1,k-1}^{-1}\mathcal{W}_{k-1,k-1}^T, \quad k \geq 2; \\ P_{1/0} &= F_0 P_{0/0} F_0^T + Q_{0,0}, \end{aligned} \tag{3.16}$$

where $\mathcal{J}_k = E\left[\widetilde{x}_{k/k} w_k^T\right]$ *is calculated by*

$$\mathcal{J}_k = Q_{k-1,k} - \mathcal{X}_{k,k}\Pi_{k,k}^{-1}\mathcal{W}_{k,k}^T, \quad k \geq 1. \tag{3.17}$$

The filtering error covariance matrix, $P_{k/k}$, is given by

$$P_{k/k} = P_{k/k-1} - \mathcal{X}_{k,k}\Pi_{k,k}^{-1}\mathcal{X}_{k,k}^T, \quad k \geq 1; \qquad P_{0/0} = P_0. \tag{3.18}$$

The innovation covariance matrix, $\Pi_{k,k}$, satisfies

$$\begin{aligned}\Pi_{k,k} &= Cov(\theta_k) \circ \left(H_k D_k H_k^T\right) + R_{k,k} + \overline{\Theta}_k H_k \mathcal{X}_{k,k} + \mathcal{X}_{k,k}^T H_k^T \overline{\Theta}_k \\ &\quad - \overline{\Theta}_k H_k P_{k/k-1} H_k^T \overline{\Theta}_k - \mathcal{V}_{k,k-1}\Pi_{k-1,k-1}^{-1}\mathcal{V}_{k,k-1}^T, \quad k \geq 2; \\ \Pi_{1,1} &= Cov(\theta_1) \circ \left(H_1 D_1 H_1^T\right) + R_{1,1} + \overline{\Theta}_1 H_1 \mathcal{X}_{1,1} + \mathcal{X}_{1,1}^T H_1^T \overline{\Theta}_1 - \overline{\Theta}_1 H_1 P_{1/0} H_1^T \overline{\Theta}_1.\end{aligned} \tag{3.19}$$

The matrices D_k, E_k, $\mathcal{W}_{k,k}$ and $\mathcal{V}_{k,k-1}$ are given in (3.3), (3.6), (3.9) and (3.10), respectively.

Proof. See Appendix A.1.

Remark 4. In conventional estimation problems in systems with missing measurements and uncorrelated additive white noises, the one-stage state and observation predictors are calculated as $\widehat{x}_{k/k-1} = F_{k-1}\widehat{x}_{k-1/k-1}$ and $\widehat{y}_{k/k-1} = \overline{\Theta}_k H_k \widehat{x}_{k/k-1}$, respectively. However, this is not true for the problem at hand since, due to the correlation assumption *(ii)*, the noise estimators $\widehat{w}_{k-1/k-1}$ and $\widehat{v}_{k/k-1}$ must be taken into account for the derivation of the predictors. Besides the fact of considering missing measurements, this is the main difference between the optimal estimators proposed in the current paper and the suboptimal Kalman-type ones proposed in [16], where the noise estimators are considered to be equal to zero.

3.4 Distributed fusion estimation

One of the main disadvantages of the centralized fusion estimators derived in Section 3.3 is that they may have a high computational cost due to augmentation. Moreover, as is widely known, the centralized approach has several other drawbacks, such as fault detection, isolation, poor reliability, etc. To overcome these

disadvantages, our aim in this section is to address the optimal distributed fusion estimation problem, in which each single sensor provides its local LS linear estimator and their estimation error covariance matrices, and then these local estimators along with the covariances and cross-covariance matrices of the estimation errors between any two sensors are sent to the fusion center for fusion based on the matrices-weighted fusion estimation criterion in the linear minimum variance sense [30].

3.4.1 Local LS linear filtering algorithms

For each single sensor subsystem of system (3.1)-(3.2), the following theorem provides recursive formulas for the local LS linear filters, $\widehat{x}^i_{k/k}$, and their corresponding error covariance matrices, $P^{ii}_{k/k}$.

Theorem 3.4.1 *For the i-th sensor subsystem of system* (3.1)-(3.2) *under assumptions (i)-(iv), the local LS linear filter,* $\widehat{x}^i_{k/k}$*, is calculated by*

$$\widehat{x}^i_{k/k} = \widehat{x}^i_{k/k-1} + \mathcal{X}^i_{k,k}\left(\Pi^{ii}_{k,k}\right)^{-1}\mu^i_k, \quad k \geq 1; \quad \widehat{x}^i_{0/0} = \overline{x}_0, \tag{3.20}$$

where the local LS linear predictor, $\widehat{x}^i_{k/k-1}$*, satisfies*

$$\begin{aligned} &\widehat{x}^i_{k/k-1} = F_{k-1}\widehat{x}^i_{k-1/k-1} + \mathcal{W}^i_{k-1,k-1}\left(\Pi^{ii}_{k-1,k-1}\right)^{-1}\mu^i_{k-1}, \quad k \geq 2; \\ &\widehat{x}^i_{1/0} = F_0\widehat{x}^i_{0/0}, \end{aligned} \tag{3.21}$$

with $\mathcal{W}^i_{k,k} = \overline{\theta}^i_k Q_{k,k-1}H^{iT}_k + S^i_{k,k},\ k \geq 1.$

The innovation, μ^i_k*, is given by*

$$\begin{aligned} &\mu^i_k = y^i_k - \overline{\theta}^i_k H^i_k \widehat{x}^i_{k/k-1} - \mathcal{V}^{ii}_{k,k-1}\left(\Pi^{ii}_{k-1,k-1}\right)^{-1}\mu^i_{k-1}, \quad k \geq 2; \\ &\mu^i_1 = y^i_1 - \overline{\theta}^i_1 H^i_1 \widehat{x}^i_{1/0}, \end{aligned} \tag{3.22}$$

with $\mathcal{V}^{ii}_{k,k-1} = \overline{\theta}^i_{k-1}S^{iT}_{k-2,k}H^{iT}_{k-1} + R^{ii}_{k,k-1},\ k \geq 2.$

The vector $\mathcal{X}_{k,k}^i = E\left[x_k \mu_k^i\right]$ *is calculated from the following expression*

$$\begin{aligned}\mathcal{X}_{k,k}^i &= \overline{\theta}_k^i P_{k/k-1}^{ii} H_k^{iT} + E_k^i - \mathcal{X}_{k,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \mathcal{V}_{k,k-1}^{ii}, \quad k \geq 2; \\ \mathcal{X}_{1,1}^i &= \overline{\theta}_1^i P_{1/0}^{ii} H_1^{iT} + E_1^i,\end{aligned}$$

where $\mathcal{X}_{k,k-1}^i = F_{k-1}\mathcal{X}_{k-1,k-1}^i + \mathcal{W}_{k-1,k-1}^i,\ k \geq 2.$

The local prediction error covariance matrix, $P_{k/k-1}^{ii}$*, is obtained by*

$$\begin{aligned}P_{k/k-1}^{ii} &= F_{k-1} P_{k-1/k-1}^{ii} F_{k-1}^T + Q_{k-1,k-1} + F_{k-1}\mathcal{J}_{k-1}^i + \mathcal{J}_{k-1}^{iT} F_{k-1}^T \\ &\quad - \mathcal{W}_{k-1,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \mathcal{W}_{k-1,k-1}^{iT}, \quad k \geq 2; \\ P_{1/0}^{ii} &= F_0 P_{0/0}^{ii} F_0^T + Q_{0,0},\end{aligned}$$

where $\mathcal{J}_k^i = Q_{k-1,k} - \mathcal{X}_{k,k}^i \left(\Pi_{k,k}^{ii}\right)^{-1} \mathcal{W}_{k,k}^{iT},\ k \geq 1,$ *and* $P_{k/k}^{ii}$*, the filtering error covariance matrix, is given by*

$$P_{k/k}^{ii} = P_{k/k-1}^{ii} - \mathcal{X}_{k,k}^i \left(\Pi_{k,k}^{ii}\right)^{-1} \mathcal{X}_{k,k}^{iT}, \quad k \geq 1; \qquad P_{0/0}^{ii} = P_0.$$

The innovation variance, $\Pi_{k,k}^{ii}$*, satisfies*

$$\begin{aligned}\Pi_{k,k}^{ii} &= V_k^{\theta^i} H_k^i D_k H_k^{iT} + R_{k,k}^{ii} + \overline{\theta}_k^i H_k^i \mathcal{X}_{k,k}^i + \overline{\theta}_k^i \mathcal{X}_{k,k}^{iT} H_k^{iT} \\ &\quad - (\overline{\theta}_k^i)^2 H_k^i P_{k/k-1}^{ii} H_k^{iT} - (\mathcal{V}_{k,k-1}^{ii})^2 \left(\Pi_{k-1,k-1}^{ii}\right)^{-1}, \quad k \geq 2; \\ \Pi_{1,1}^{ii} &= V_1^{\theta^i} H_1^i D_1 H_1^{iT} + R_{1,1}^{ii} + \overline{\theta}_1^i H_1^i \mathcal{X}_{1,1}^i + \overline{\theta}_1^i \mathcal{X}_{1,1}^{iT} H_1^{iT} - (\overline{\theta}_1^i)^2 H_1^i P_{1/0}^{ii} H_1^{iT}.\end{aligned}$$

The matrix D_k *and the vector* E_k^i *are given in (3.3) and (3.4), respectively.*

Proof. The proof, based on the innovation approach and the OPL, is omitted for being analogous to that of Theorem 3.3.1. Nevertheless, it should be indicated that, in this proof, the Hadamard product is not used since, instead of the diagonal stochastic matrix Θ_k, the scalar variable θ_k^i is now involved in the derivation of the estimators. □

Remark 5. As indicated in Remark 4 for the centralized estimators, it must be noted that, due to the correlation assumption *(ii)* of the additive noises $\{w_k\}$

and $\{v_k^i\}$, the estimators $\widehat{w}_{k-1/k-1}^i = \mathcal{W}_{k-1,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \mu_{k-1}^i$ and $\widehat{v}_{k/k-1}^i = \mathcal{V}_{k,k-1}^{ii} \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \mu_{k-1}^i$ are not equal to zero, and hence the optimal local state predictor, $\widehat{x}_{k/k-1}^i = F_{k-1}\widehat{x}_{k-1/k-1}^i + \widehat{w}_{k-1/k-1}^i$, and the observation predictor, $\widehat{y}_{k/k-1}^i = \overline{\theta}_k^i H_k^i \widehat{x}_{k/k-1}^i + \widehat{v}_{k/k-1}^i$, are quite different from conventional filtering algorithms with uncorrelated white noises. This issue, along with the consideration of missing measurements at each single sensor, constitutes the main difference between the current optimal local estimators and the suboptimal local estimators proposed in [16].

3.4.2 Cross-covariance matrices of local estimation errors

To apply the optimal fusion criterion weighted by matrices in the linear minimum variance sense, the filtering, $P_{k/k}^{ij}$, and prediction, $P_{k/k-1}^{ij}$, error cross-covariance matrices between local estimators of any two subsystems must be calculated.

For simplicity, besides the notation of Theorem 3.4.1, for $i \neq j$, $i, j = 1, 2 \ldots, r$, we introduce the following notation:

$$L_k^{ij} = E[\widehat{x}_{k/k-1}^i \mu_k^j], \quad \Pi_{k,s}^{ij} = E[\nu_k^i \nu_s^j], \quad \mathcal{V}_{k,k-1}^{ij} = E[v_k^i \mu_{k-1}^j].$$

Also, in order to simplify the calculation of the error cross-covariance matrices, the following lemmas are given.

Lemma 3.4.1 *Under assumptions (i)-(iv), the following results hold:*

a) The expectation $E[\widehat{x}_{k/k-1}^i \mu_{k-1}^j]$ satisfies

$$E[\widehat{x}_{k/k-1}^i \mu_{k-1}^j] = F_{k-1} L_{k-1}^{ij} + \mathcal{X}_{k,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \Pi_{k-1,k-1}^{ij}, \quad k \geq 2. \tag{3.23}$$

b) The expectation $E[\widehat{x}_{k/k-1}^i v_k^j]$ satisfies

$$E[\widehat{x}_{k/k-1}^i v_k^j] = \mathcal{X}_{k,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \mathcal{V}_{k,k-1}^{ji}, \quad k \geq 2, \tag{3.24}$$

where $\mathcal{V}_{k,k-1}^{ji} = \overline{\theta}_{k-1}^i H_{k-1}^i S_{k-2,k}^j + R_{k-1,k}^{ij}$.

c) The expectation $E[v_k^i \mu_k^j]$ satisfies

$$\begin{aligned} E[v_k^i \mu_k^j] = \overline{\theta}_k^j E_k^{iT} H_k^{jT} + R_{k,k}^{ij} - \mathcal{V}_{k,k-1}^{ij} \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \\ \times \left(\overline{\theta}_k^j H_k^j \mathcal{X}_{k,k-1}^j - \mathcal{V}_{k,k-1}^{jj}\right)^T, \; k \geq 2. \end{aligned} \tag{3.25}$$

Proof.

a) From (3.21) for $\widehat{x}_{k/k-1}^i$ and (3.20) for $\widehat{x}_{k-1/k-1}^i$, we have

$$\begin{aligned} E[\widehat{x}_{k/k-1}^i \mu_{k-1}^j] &= F_{k-1} E[\widehat{x}_{k-1/k-1}^i \mu_{k-1}^j] + \mathcal{W}_{k-1,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \Pi_{k-1,k-1}^{ij} \\ &= F_{k-1} E[\widehat{x}_{k-1/k-2}^i \mu_{k-1}^j] + F_{k-1} \mathcal{X}_{k-1,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \Pi_{k-1,k-1}^{ij} \\ &\quad + \mathcal{W}_{k-1,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \Pi_{k-1,k-1}^{ij} \\ &= F_{k-1} L_{k-1}^{ij} + \left(F_{k-1} \mathcal{X}_{k-1,k-1}^i + \mathcal{W}_{k-1,k-1}^i\right) \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \Pi_{k-1,k-1}^{ij}, \end{aligned}$$

and since $\mathcal{X}_{k,k-1}^i = F_{k-1} \mathcal{X}_{k-1,k-1}^i + \mathcal{W}_{k-1,k-1}^i$, expression (3.23) is proved.

b) Analogously, taking into account that $E[\widehat{x}_{k-1/k-2}^i v_k^j] = 0$, we have

$$\begin{aligned} E[\widehat{x}_{k/k-1}^i v_k^j] &= F_{k-1} E[\widehat{x}_{k-1/k-1}^i v_k^j] + \mathcal{W}_{k-1,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} E[\mu_{k-1}^i v_k^j] \\ &= \left(F_{k-1} \mathcal{X}_{k-1,k-1}^i + \mathcal{W}_{k-1,k-1}^i\right) \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \mathcal{V}_{k,k-1}^{ji}, \end{aligned}$$

and expression (3.24) is immediately obtained. Finally, the derivation of expression $\mathcal{V}_{k,k-1}^{ji} = \overline{\theta}_{k-1}^i H_{k-1}^i S_{k-2,k}^j + R_{k-1,k}^{ij}$ is similar to that of (3.10) and hence it is omitted.

c) Taking into account expression (3.22) for μ_k^j, with (3.2) for y_k^j, we have

$$E[v_k^i \mu_k^j] = \overline{\theta}_k^j E_k^{iT} H_k^{jT} + R_{k,k}^{ij} - \overline{\theta}_k^j E[v_k^i \widehat{x}_{k/k-1}^{jT}] H_k^{jT} - \mathcal{V}_{k,k-1}^{ij} \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \mathcal{V}_{k,k-1}^{jj}, \; k \geq 2,$$

and using (3.24) for $E[v_k^i \widehat{x}_{k/k-1}^{jT}]$, expression (3.25) is obtained. □

Lemma 3.4.2 *Under assumptions (i)-(iv), for $i \neq j$, $i, j = 1, 2 \ldots, r$, the expectations $L_k^{ij} = E[\widehat{x}_{k/k-1}^i \mu_k^j]$ are recursively obtained by*

$$\begin{aligned} L_k^{ij} &= \overline{\theta}_k^j \left(P_{k/k-1}^{jj} - P_{k/k-1}^{ij}\right) H_k^{jT} - F_{k-1} L_{k-1}^{ij} \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \mathcal{V}_{k,k-1}^{jj} \\ &\quad + \mathcal{X}_{k,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \left(\mathcal{V}_{k,k-1}^{ji} - \mathcal{V}_{k,k-1}^{jj} \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \Pi_{k-1,k-1}^{ij}\right), \quad k \geq 2, \end{aligned}$$

with initial condition $L_1^{ij} = 0$.

Proof. Taking into account expression (3.22) for μ_k^j, with (3.2) for y_k^j, we have

$$L_k^{ij} = \overline{\theta}_k^j E[\widehat{x}_{k/k-1}^i x_k^T] H_k^{jT} + E[\widehat{x}_{k/k-1}^i v_k^j] - \overline{\theta}_k^j E[\widehat{x}_{k/k-1}^i \widehat{x}_{k/k-1}^{jT}] H_k^{jT} - E[\widehat{x}_{k/k-1}^i \mu_{k-1}^j] \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \mathcal{V}_{k,k-1}^{jj}, \quad k \geq 2.$$

From the OPL, $E[\widehat{x}_{k/k-1}^i x_k^T] = E[\widehat{x}_{k/k-1}^i \widehat{x}_{k/k-1}^{iT}]$; then, taking into account (3.23) for $E[\widehat{x}_{k/k-1}^i \mu_{k-1}^j]$, and (3.24) for $E[\widehat{x}_{k/k-1}^i v_k^j]$, it is enough to prove that

$$E[\widehat{x}_{k/k-1}^i \widehat{x}_{k/k-1}^{iT}] - E[\widehat{x}_{k/k-1}^i \widehat{x}_{k/k-1}^{jT}] = P_{k/k-1}^{jj} - P_{k/k-1}^{ij},$$

which is easily deduced since

$$E[\widehat{x}_{k/k-1}^i \widehat{x}_{k/k-1}^{jT}] = P_{k/k-1}^{ij} - D_k + E[\widehat{x}_{k/k-1}^i \widehat{x}_{k/k-1}^{iT}] + E[\widehat{x}_{k/k-1}^j \widehat{x}_{k/k-1}^{jT}]$$

and $E[\widehat{x}_{k/k-1}^j \widehat{x}_{k/k-1}^{jT}] = D_k - P_{k/k-1}^{jj}$. □

Lemma 3.4.3 *Under assumptions (i)-(iv), for $i \neq j$, $i, j = 1, 2 \ldots, r$, the innovation cross-covariance $\Pi_{k,k}^{ij} = E[\mu_k^i \mu_k^j]$ satisfies*

$$\Pi_{k,k}^{ij} = \overline{\theta}_k^i H_k^i \left(\mathcal{X}_{k,k}^j - L_k^{ij}\right) + \overline{\theta}_k^j E_k^{iT} H_k^{jT} + R_{k,k}^{ij} - \mathcal{V}_{k,k-1}^{ii} \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \Pi_{k-1,k}^{ij} - \mathcal{V}_{k,k-1}^{ij} \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \left(\overline{\theta}_k^j H_k^j \mathcal{X}_{k,k-1}^j - \mathcal{V}_{k,k-1}^{jj}\right)^T, \quad k \geq 2;$$

$$\Pi_{1,1}^{ij} = \overline{\theta}_1^i H_1^i \mathcal{X}_{1,1}^j + \overline{\theta}_1^j E_1^i H_1^{jT} + R_{1,1}^{ij},$$

where $\Pi_{k-1,k}^{ij} = E[\mu_{k-1}^i \mu_k^j]$ is given by

$$\Pi_{k-1,k}^{ij} = \overline{\theta}_k^j \left(\mathcal{X}_{k,k-1}^i - F_{k-1} L_{k-1}^{ji} - \mathcal{X}_{k,k-1}^j \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \Pi_{k-1,k-1}^{ji}\right)^T H_k^{jT} + \mathcal{V}_{k,k-1}^{ji} - \Pi_{k-1,k-1}^{ij} \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \mathcal{V}_{k,k-1}^{jj}, \quad k \geq 2.$$

Proof. Taking into account expression (3.22) for μ_k^i, with (3.2) for y_k^i, we have

$$\Pi_{k,k}^{ij} = \overline{\theta}_k^i H_k^i E[x_k \mu_k^j] + E[v_k^i \mu_k^j] - \overline{\theta}_k^i H_k^i E[\widehat{x}_{k/k-1}^i \mu_k^j] - \mathcal{V}_{k,k-1}^{ii} \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} E[\mu_{k-1}^i \mu_k^j]$$
$$= \overline{\theta}_k^i H_k^i \left(\mathcal{X}_{k,k}^j - L_k^{ij}\right) + E[v_k^i \mu_k^j] - \mathcal{V}_{k,k-1}^{ii} \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \Pi_{k-1,k}^{ij}, \quad k \geq 2,$$

and, from (3.25) for $E[v_k^i \mu_k^j]$, expression for $\Pi_{k,k}^{ij}$ is clear.

Analogously, taking into account expression (3.22) for μ_k^j, with (3.2) for y_k^j, we have

$$\begin{aligned}\Pi_{k-1,k}^{ij} &= \overline{\theta}_k^j E[\mu_{k-1}^i x_k^T] H_k^{jT} + E[\mu_{k-1}^i v_k^j] - \overline{\theta}_k^j E[\mu_{k-1}^i \widehat{x}_{k/k-1}^{jT}] H_k^{jT} \\ &\quad - E[\mu_{k-1}^i \mu_{k-1}^j] \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \mathcal{V}_{k,k-1}^{jj}, \\ &= \overline{\theta}_k^j \left(\mathcal{X}_{k,k-1}^i - E[\mu_{k-1}^i \widehat{x}_{k/k-1}^{jT}]\right) H_k^{jT} + \mathcal{V}_{k,k-1}^{ji} \\ &\quad - \Pi_{k-1,k-1}^{ij} \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \mathcal{V}_{k,k-1}^{jj}, \quad k \geq 2,\end{aligned}$$

and, from (3.23) for $E[\mu_{k-1}^i \widehat{x}_{k/k-1}^{jT}]$, expression for $\Pi_{k-1,k}^{ij}$ is immediately derived.

□

In the following theorem, recursive formulas to calculate the filtering and prediction error cross-covariance matrices, $P_{k/k}^{ij}$ and $P_{k/k-1}^{ij}$, respectively, are derived.

Theorem 3.4.2 *Under assumptions (i)-(iv), the cross-covariance matrices, $P_{k/k}^{ij}$, of the filtering errors between the i-th and the j-th sensor subsystems are recursively computed by*

$$\begin{aligned}P_{k/k}^{ij} &= P_{k/k-1}^{ij} + \mathcal{X}_{k,k}^i (\Pi_{k,k}^{ii})^{-1} \Pi_{k,k}^{ij} (\Pi_{k,k}^{jj})^{-1} \mathcal{X}_{k,k}^{jT} \\ &\quad - \left(\mathcal{X}_{k,k}^j - L_k^{ij}\right) (\Pi_{k,k}^{jj})^{-1} \mathcal{X}_{k,k}^{jT} - \mathcal{X}_{k,k}^i (\Pi_{k,k}^{ii})^{-1} \left(\mathcal{X}_{k,k}^i - L_k^{ji}\right)^T, \quad k \geq 1; \\ P_{0/0}^{ij} &= P_0,\end{aligned}$$

where $P_{k/k-1}^{ij}$, the cross-covariance matrix of the prediction error between the i-th and the j-th sensor subsystems, satisfies

$$\begin{aligned}P_{k/k-1}^{ij} &= F_{k-1} P_{k-1/k-1}^{ij} F_{k-1}^T + Q_{k-1,k-1} + F_{k-1} \mathcal{J}_{k-1}^i + \mathcal{J}_{k-1}^{jT} F_{k-1}^T \\ &\quad + \mathcal{W}_{k-1,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \Pi_{k-1,k-1}^{ij} \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \mathcal{W}_{k-1,k-1}^{jT} \\ &\quad - \mathcal{G}_{k-1}^{ij} \left(\Pi_{k-1,k-1}^{jj}\right)^{-1} \mathcal{W}_{k-1,k-1}^{jT} - \mathcal{W}_{k-1,k-1}^i \left(\Pi_{k-1,k-1}^{ii}\right)^{-1} \mathcal{G}_{k-1}^{jiT} \quad k \geq 2; \\ P_{1/0}^{ij} &= F_0 P_{0/0}^{ij} F_0^T + Q_{0,0},\end{aligned}$$

where $\mathcal{G}_k^{ij} = \mathcal{W}_{k,k}^j + F_k \left(\mathcal{X}_{k,k}^j - L_k^{ij} - \mathcal{X}_{k,k}^i \left(\Pi_{k,k}^{ii}\right)^{-1} \Pi_{k,k}^{ij}\right)$, $k \geq 1$. The vectors L_k^{ij} and the innovation cross-covariances $\Pi_{k,k}^{ij}$ are given in Lemmas 3.4.2 and 3.4.3, respectively.

Proof. By using (3.20) for $\widehat{x}^i_{k/k}$ and $\widehat{x}^j_{k/k}$, we have

$$P^{ij}_{k/k} = P^{ij}_{k/k-1} + \mathcal{X}^i_{k,k}(\Pi^{ii}_{k,k})^{-1}\Pi^{ij}_{k,k}(\Pi^{jj}_{k,k})^{-1}\mathcal{X}^{jT}_{k,k} - E[(x_k - \widehat{x}^i_{k/k-1})\mu^{jT}_k](\Pi^{jj}_{k,k})^{-1}\mathcal{X}^{jT}_{k,k} - \mathcal{X}^i_{k,k}(\Pi^{ii}_{k,k})^{-1}E[\mu^i_k(x_k - \widehat{x}^j_{k/k-1})^T].$$

Taking into account that $E[x_k\mu^j_k] = \mathcal{X}^j_{k,k}$ and $E[\widehat{x}^i_{k/k-1}\mu^j_k] = L^{ij}_k$, the recursive expression for the cross-covariance matrices of the local filtering errors, is immediately deduced.

Following an analogous reasoning, using now (3.21) and taking into account that $E[(x_k - \widehat{x}^i_{k/k})w^T_k] = \mathcal{J}^i_k$ and $E[\mu^i_k w^T_k] = \mathcal{W}^{iT}_{k,k}$, it is easy to see that

$$\begin{aligned} P^{ij}_{k/k-1} = {} & F_{k-1}P^{ij}_{k-1/k-1}F^T_{k-1} + Q_{k-1,k-1} + F_{k-1}\mathcal{J}^i_{k-1} + \mathcal{J}^{jT}_{k-1}F^T_{k-1} \\ & + \mathcal{W}^i_{k-1,k-1}\left(\Pi^{ii}_{k-1,k-1}\right)^{-1}\Pi^{ij}_{k-1,k-1}\left(\Pi^{jj}_{k-1,k-1}\right)^{-1}\mathcal{W}^{jT}_{k-1,k-1} \\ & - \left(\mathcal{W}^j_{k-1,k-1} + F_{k-1}E[\widetilde{x}^i_{k-1/k-1}\mu^j_{k-1}]\right)\left(\Pi^{jj}_{k-1,k-1}\right)^{-1}\mathcal{W}^{jT}_{k-1,k-1} \\ & - \mathcal{W}^i_{k-1,k-1}\left(\Pi^{ii}_{k-1,k-1}\right)^{-1}\left(\mathcal{W}^i_{k-1,k-1} + F_{k-1}E[\widetilde{x}^j_{k/k}\mu^i_{k-1}]\right)^T. \end{aligned}$$

Finally, using again (3.20) for $\widehat{x}^i_{k-1/k-1}$, and since $E[\widehat{x}^i_{k-1/k-2}\mu^j_{k-1}] = L^{ij}_{k-1}$, we have

$$E[\widetilde{x}^i_{k-1/k-1}\mu^j_{k-1}] = \mathcal{X}^j_{k-1,k-1} - L^{ij}_{k-1} - \mathcal{X}^i_{k-1,k-1}\left(\Pi^{ii}_{k-1,k-1}\right)^{-1}\Pi^{ij}_{k-1,k-1}$$

and the expression for the cross-covariance matrices of the local prediction errors is easily obtained.

□

3.4.3 Distributed fusion filtering estimators

Once the local LS linear filtering estimators $\widehat{x}^i_{k/k}$ and their error covariance matrices $P^{ii}_{k/k}$, given in Theorem 3.4.1, along with the error cross-covariance matrices, $P^{ij}_{k/k}$, given in Theorem 3.4.2, are available, the distributed optimal weighted fusion estimators and their error covariance matrices are obtained by applying the optimal information fusion criterion weighted by matrices in the linear minimum variance sense [30].

Theorem 3.4.3 *For the system model* (3.1) *and measurement model* (3.2)*, under assumptions (i)-(iv), the distributed optimal fusion filter,* $\widehat{x}^D_{k/k}$*, is given by*

$$\widehat{x}^D_{k/k} = A^1_k \widehat{x}^1_{k/k} + \cdots + A^r_k \widehat{x}^r_{k/k}, \quad k \geq 0,$$

where the local estimators $\widehat{x}^i_{k/k}$*,* $k \geq 0$ $(i = 1, 2 \ldots, r)$ *are calculated by the recursive algorithm established in Theorem 3.4.1.*

The optimal matrix weights A^i_k $(i = 1, 2, \ldots, r)$ *are computed by*

$$A_k = \Sigma^{-1}_{k/k} e \left(e^T \Sigma^{-1}_{k/k} e \right)^{-1},$$

where the matrices $A_k = [A^1_k, \ldots, A^r_k]^T$ *and* $e = [I, \ldots, I]^T$ *are both* $nr \times n$ *matrices, and*

$$\Sigma_{k/k} = E\left[\left(\widetilde{x}^1_{k/k}, \ldots, \widetilde{x}^r_{k/k} \right) \left(\widetilde{x}^1_{k/k}, \ldots, \widetilde{x}^r_{k/k} \right)^T \right] = \left(P^{ij}_{k/k} \right)_{i,j=1,2,\ldots,r}$$

is an $nr \times nr$ *positive definite symmetric block matrix, whose* $n \times n$ *matrix entries* $P^{ij}_{k/k}$ *are given in Theorems 3.4.1 and 3.4.2.*

The error covariance matrices of the distributed weighted fusion filtering estimators are computed by

$$P^D_{k/k} = \left(e^T \Sigma^{-1}_{k/k} e \right)^{-1}, \quad k \geq 0,$$

and the following inequality holds: $P^D_{k/k} \leq P^{ii}_{k/k}$*,* $i = 1, 2, \ldots, r$*.*

Proof. The proof is omitted because it follows directly from the optimal information criterion weighted by matrices in the linear minimum variance sense [30].

□

Remark 6. The proposed distributed optimal LS linear fusion filter requires the computation of an $nr \times nr$ inverse matrix, with n the dimension of the system state and r the number of sensors. Consequently, the proposed distributed fusion method has a computational complexity of $O[(nr)^3]$, equal to that of the distributed

Kalman-type filter in [16] and less than that of the distributed fusion filters based on the state augmentation approach. Hence, our distributed fusion method is superior to the filter proposed in [16] (since it has the same computation burden but better accuracy) and also to the distributed fusion filters based on state augmentation (since it has less computational complexity).

3.5 Numerical simulation example

In this section, a numerical simulation example is presented to illustrate the effectiveness of the centralized and distributed filtering algorithms proposed in this paper. Consider a scalar first-order autoregressive model with missing measurements coming from two sensors with autocorrelated and cross-correlated noises. According to the proposed observation model, two different independent sequences of random variables with a certain probability distribution over the interval $[0, 1]$ are used to model the missing phenomenon. Specifically, the following model is considered:

$$\begin{aligned} x_k &= 0.95x_{k-1} + w_{k-1}, \quad k \geq 1 \\ y_k^i &= \theta_k^i \begin{bmatrix} 1 & 1 \end{bmatrix} x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2 \end{aligned}$$

where the initial state x_0 is a zero-mean Gaussian variable with variance $P_0 = 1$. The noise processes $\{w_k;\, k \geq 0\}$ and $\{v_k^i;\, k \geq 1\}$, $i = 1, 2$, are defined by

$$\begin{aligned} w_k &= 0.6(\eta_{k+1} + \eta_{k+2}) \\ v_k^i &= c_i(\eta_k + \eta_{k+1}), \quad i = 1, 2, \end{aligned}$$

where the sequence of variables $\{\eta_k;\, k \geq 1\}$ is a zero-mean Gaussian white process with variance 0.5. Clearly, according to assumption *(ii)*, the additive noises $\{w_k\}$ and $\{v_k^i\}$ are one-step autocorrelated and two-step cross-correlated with

$$Q_{k,k} = 0.36,\ Q_{k,k+1} = 0.18,\ S_{k,k}^i = 0.3c_i,\ S_{k-1,k}^i = 0.6c_i,\ S_{k-2,k}^i = 0.3c_i,$$

$$R_{k,k}^{ii} = c_i^2,\ R_{k,k+1}^{ii} = 0.5c_i^2,\ R_{k,k}^{ij} = c_ic_j,\ R_{k,k+1}^{ij} = 0.5c_ic_j.$$

The phenomenon of missing measurements for each sensor is described as follows:

- In the first sensor, a sequence of independent and identically distributed (i.i.d.) random variables, $\{\theta_k^1;\, k \geq 1\}$, is considered, with probability distribution given by

$$P[\theta_k^1 = 0] = 0.1, \quad P[\theta_k^1 = 0.5] = 0.5, \quad P[\theta_k^1 = 1] = 0.4.$$

  If $\theta_k^1 = 0$, which occurs with probability 0.1, the state x_k is missing and the observation y_k^1 contains only noise v_k^1; if $\theta_k^1 = 0.5$, only partial information of the state x_k is missing in such observation, which happens with probability 0.5; and, finally, the state is present in the observation y_k^1 with probability 0.4 when $\theta_k^1 = 1$. The mean and variance of these variables are easily calculated, being $\overline{\theta}_k^1 = 0.65$ and $V_k^{\theta^1} = 0.1025$, for all k.

- In the second sensor, a sequence of i.i.d. Bernoulli random variables, $\{\theta_k^2;\, k \geq 1\}$, is considered, with $P[\theta_k^2 = 1] = p$; in this case, if $\theta_k^2 = 1$ the state x_k is present in the measurement y_k^2 with probability p, whereas if $\theta_k^2 = 0$ such observation only contains additive noise, v_k^2, with probability $1 - p$. So, no partial missing information is considered in this sensor. Clearly, for all k, $\overline{\theta}_k^2 = p$ and $V_k^{\theta^2} = p(1-p)$.

To illustrate the feasibility and effectiveness of the proposed estimators we ran a program in MATLAB, in which fifty iterations of the proposed algorithms have been performed considering different values of c_i and p. Using simulated values of the state and the corresponding observations, both distributed and centralized filtering estimates of the state are calculated, as well as the corresponding error variances, which provide a measure of the estimation accuracy.

Firstly, for $p = 0.8$, the local, centralized and distributed filtering error variances are displayed in Figure 3.1 considering the values $c_1 = 1$ and $c_2 = 0.5$.

According to Theorem 3.4.3, this figure corroborates that the optimal fusion distributed filter performs quite better than each local filter, but lightly worse than the centralized filter. Nevertheless, although the distributed fusion filter has a bit lower accuracy than the centralized one, both filters perform similarly and provide good estimations. Moreover, this slight difference is compensated because the distributed fusion structure is in general more robust, reduces the computational cost and improves the reliability due to its parallel structure. For these reasons, the distributed filter is generally preferred in practice.

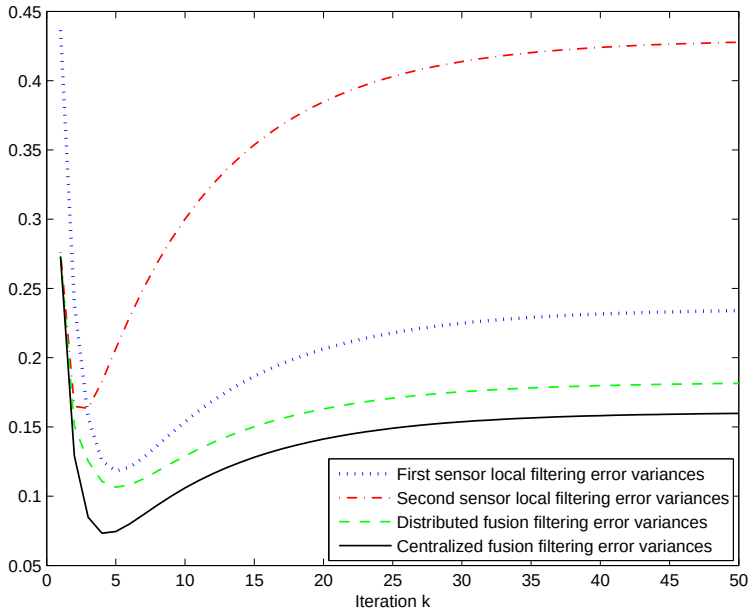


Figure 3.1: Local, centralized and distributed fusion filtering error variances.

Next, to analyze the performance of the proposed estimators versus the probability that the state x_k is present in the measurements of the second sensor, the centralized and distributed filtering error variances have been calculated for $c_1 = 1$, $c_2 = 0.5$ and different values of the probability $p = 0.2,\ 0.6$ and 0.8. The results are displayed in Figure 3.2; analysis of this figure reveals that as p increases (or, equiv-

alently, the probability $1-p$ that the state is missing in the observations from the second sensor decreases), the filtering error variances become smaller and, hence, better estimations are obtained. Also, this figure shows that, for all the considered probability values, the error variances corresponding to the centralized filter are always less than those of the distributed filter. Analogous results are obtained for other values of c_1, c_2 and the probability p.

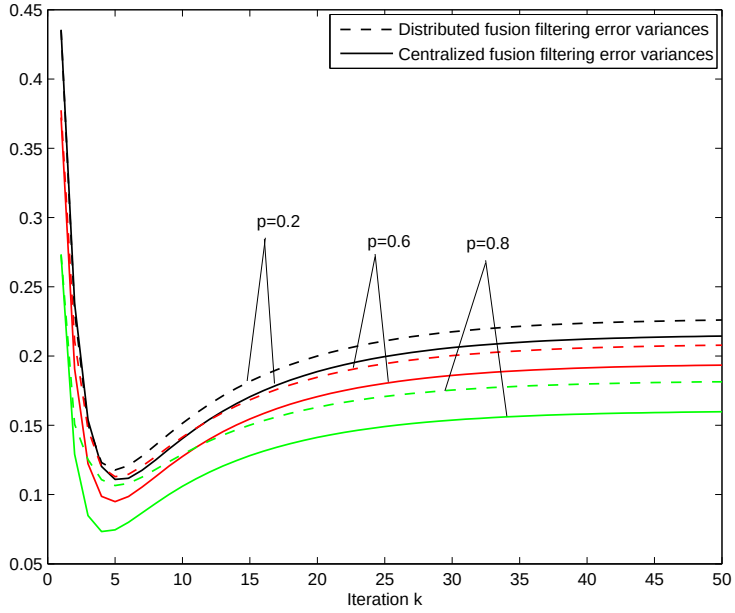


Figure 3.2: Centralized and distributed fusion filtering error variances for $p = 0.2, 0.6, 0.8$, when $c_1 = 1$, $c_2 = 0.5$.

On the other hand, to compare the performance of the estimators for different degrees of correlation between the state and the observation noises, the centralized and distributed filtering error variances have been calculated considering $c_2 = 0.5$, $p = 0.8$ and different values of c_1, specifically, $c_1 = 0.25,\ 0.5,\ 0.75$ and 1. These values provide different correlations between the noise process $\{w_k\}$ and the first sensor observation noise $\{v_k^1\}$ and, consequently, different correlations, E_k^1, between the state and the first sensor observation noise. The error variances are

displayed in Figure 3.3, from which it is inferred that the error variances are smaller (and, consequently, the performance of the estimators is better) as the value c_1 is greater; these results were expected, since the correlation between the state and observations increases with c_1. Analogous results are obtained for different values of c_2 and other values of the probability p.

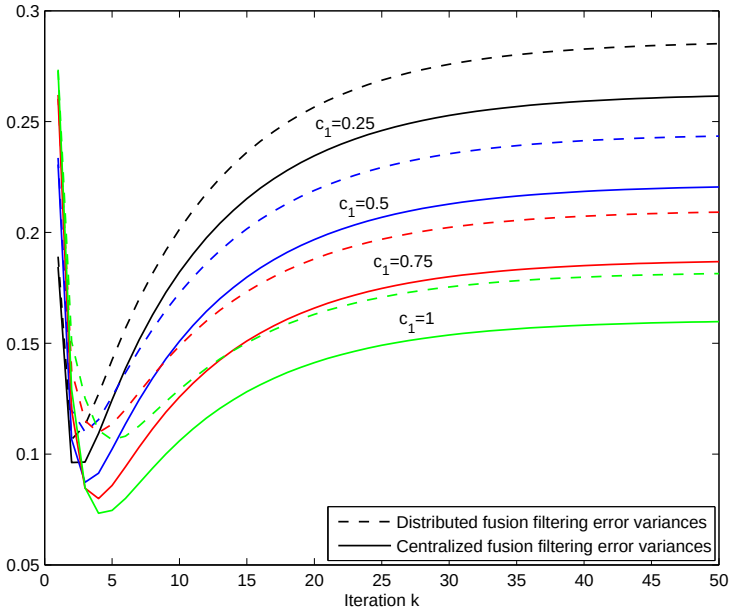


Figure 3.3: Centralized and distributed fusion filtering error variances for $c_1 = 0.25, 0.5, 0.75, 1$, when $c_2 = 0.5$ and $p = 0.8$.

Now, completing the results of the two previous figures, the performance of the filters is analyzed when $c_2 = 0.5$, the probability p is varied from 0.1 to 0.9, and the values $c_1 = 0.25,\ 0.5,\ 0.75,\ 1,\ 1.25$ and 1.5 are considered. It must be noted that in all the cases examined, the error variances present insignificant variation from a certain iteration on and, consequently, only the values at a specific iteration (namely, $k = 50$) are shown. The results are presented in Figure 3.4 which, for the sake of clarity, only displays the distributed filtering error variances. Agreeing with the comments about Figures 3.2 and 3.3, this figure shows that, for a fixed value of

c_1, the performance of the estimators improves as p becomes greater, and for a fixed value of p, also more accurate estimations are obtained as c_1 increases. Hence, from this figure it is gathered that, as c_1 or p decreases (which means that the correlation between the state and the first sensor observation noise decreases or the probability that the state is present in the second sensor measurements decreases, respectively), the filtering error variances become greater and, consequently, worse estimations are obtained.

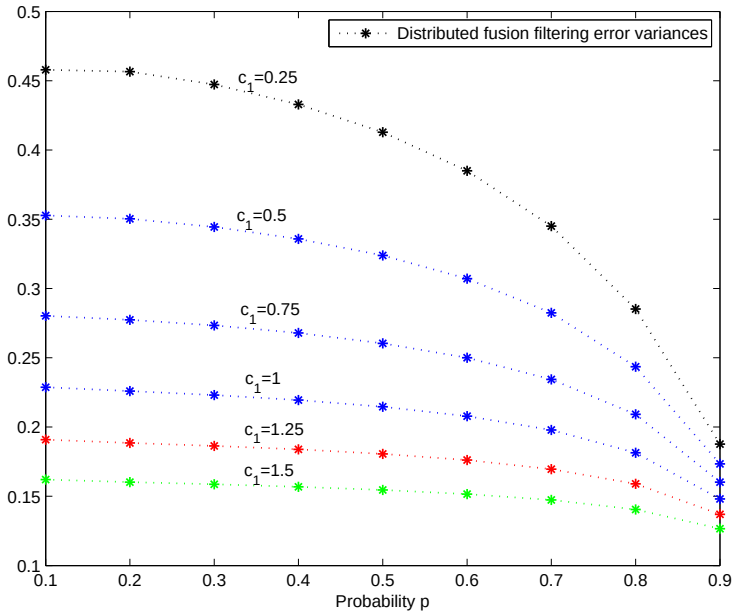


Figure 3.4: Distributed fusion filtering error variances versus p, with $c_1 = 0.25, 0.5, 0.75, 1, 1.25, 1.5$, when $c_2 = 0.5$.

Finally, a comparative analysis is presented between the classical Kalman filter [2], the Kalman-type filter with correlated and cross-correlated noises given in [16], the filter proposed in [23] for systems with different failure rates in multi-sensor networks, and the centralized and distributed filters proposed in this paper. For the comparison, the same parameter values as in Figure 3.1 are considered ($c_1 = 1$, $c_2 = 0.5$ and $p = 0.8$).

On the basis of one thousand independent simulations of the mentioned algorithms, a comparison between the different filtering estimates is performed using the mean square error (MSE) criteria. For $s = 1, \ldots, 1000$, let $\{x_k^{(s)},\ k = 1, \ldots, 50\}$ denote the s-th set of artificially simulated data (which is taken as the s-th set of true values of the state), and $\widehat{x}_{k/k}^{(s)}$ the filtering estimate at the sampling time k in the s-th simulation run. For each algorithm, the filtering MSE at time k is calculated by $\text{MSE}_k = \dfrac{1}{1000}\displaystyle\sum_{s=1}^{1000}(x_k^{(s)} - \widehat{x}_{k/k}^{(s)})^2$.

The values MSE_k, for $k = 1, \ldots, 50$, are displayed in Figure 3.5 which shows that, for all k, the proposed centralized and distributed filters have approximately the same MSE_k values, which in turn are smaller than the MSE_k values of the filter in [23] and considerably less than those of the filters [2] and [16]. Hence, we can conclude that, according to the MSE criterion, the proposed filtering estimates perform significantly better than other filters in the literature.

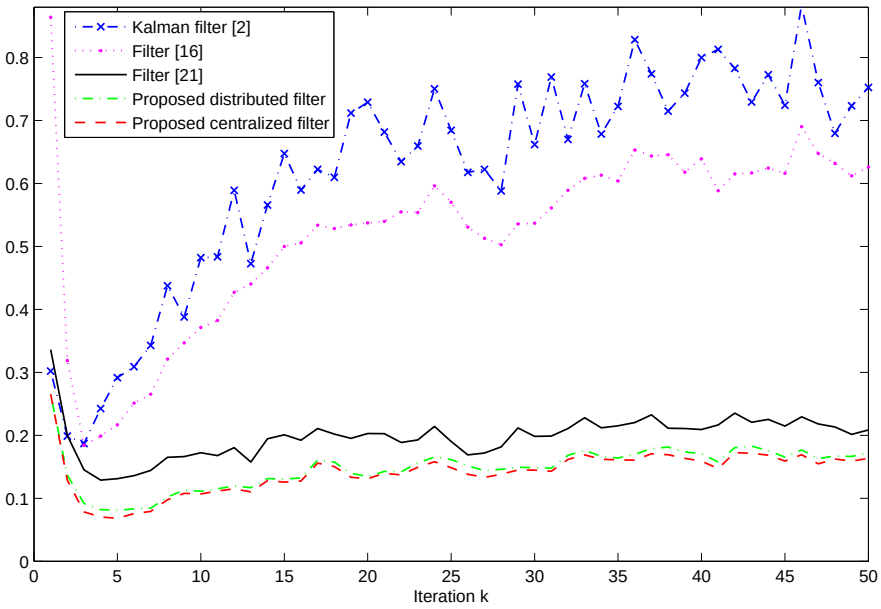


Figure 3.5: Comparison of MSE for different filters.

3.6 Conclusions

The LS linear estimation problem from missing measurements has been investigated for multi-sensor linear discrete-time systems with autocorrelated and cross-correlated noises. The main contributions are summarized as follows:

1. Using both centralized and distributed fusion methods to process the measurement data from the different sensors, recursive optimal LS linear filtering algorithms are derived by an innovation approach.

2. At each sensor, the possibility of missing measurements (that is, observations containing only partial information about the state or even only noise) is modelled by a sequence of independent random variables taking discrete values over the interval $[0, 1]$.

3. The multi-sensor system model considered in the current paper covers those situations where the sensor and process noises are one-step autocorrelated and two-step cross-correlated. Also, one-step cross-correlations between different sensor noises is considered. This correlation assumption is valid in a wide spectrum of applications, for example in target tracking systems with process and measurement noises dependent on the system state, or situations where a target is observed by multiple sensors and all of them operate in the same noisy environment. Nevertheless, the current study can be extended to more general systems involving finite-step autocorrelated and cross-correlated noises with no difficulty, except for a greater complexity in the mathematical expressions.

4. The applicability of the proposed centralized and distributed filtering algorithms is illustrated by a numerical simulation example, where a scalar state process generated by a first-order autoregressive model is estimated from missing measurements coming from two sensors with autocorrelated

and cross-correlated noises. The results confirm that centralized and distributed fusion estimators have approximately the same error variances, with a slight inferiority of the distributed one which is compensated by a reduced computational burden and reduced communication demands for the sensor networks. Also, compared with some existing estimation methods, the proposed algorithms provide better estimations in the mean square error sense.

Acknowledgements

This research is supported by Ministerio de Ciencia e Innovación (Programa FPU and grant No. MTM2011-24718) and Junta de Andalucía (grant No. P07-FQM-02701).

A.1 Proof of Theorem 3.3.1

From (3.8), expression (3.11) for the state filter $\widehat{x}_{k/k}$ in terms of the one-stage predictor $\widehat{x}_{k/k-1}$, is immediately clear.

Expression (3.12) for the state predictor $\widehat{x}_{k/k-1}$ is obtained as follows:

$$\begin{aligned}
\widehat{x}_{k/k-1} &= \sum_{s=1}^{k-1} E[x_k \mu_s^T] \Pi_{s,s}^{-1} \mu_s \\
&= \sum_{s=1}^{k-1} E[(F_{k-1} x_{k-1} + w_{k-1})\, \mu_s^T] \Pi_{s,s}^{-1} \mu_s \\
&= \sum_{s=1}^{k-1} F_{k-1} E[x_{k-1} \mu_s^T] \Pi_{s,s}^{-1} \mu_s + E[w_{k-1} \mu_{k-1}^T] \Pi_{k-1,k-1}^{-1} \mu_{k-1} \\
&= F_{k-1} \widehat{x}_{k-1/k-1} + \mathcal{W}_{k-1,k-1} \Pi_{k-1,k-1}^{-1} \mu_{k-1}, \quad k \geq 2,
\end{aligned}$$

and clearly, $\widehat{x}_{1/0} = E[x_1] = F_0 E[x_0] = F_0 \widehat{x}_{0/0}$.

Now we show expression (3.13) for the innovation, $\mu_k = y_k - \widehat{y}_{k/k-1}$, for which it is enough to obtain an expression for $\widehat{y}_{k/k-1}$. A similar reasoning to that used

to prove (3.12) leads to

$$\begin{aligned}\widehat{y}_{k/k-1} &= \sum_{s=1}^{k-1} E\left[y_k \mu_s^T\right] \Pi_{s,s}^{-1} \mu_s \\ &= \sum_{s=1}^{k-1} E\left[\left(\Theta_k H_k x_k + v_k\right) \mu_s^T\right] \Pi_{s,s}^{-1} \mu_s \\ &= \overline{\Theta}_k H_k \sum_{s=1}^{k-1} E\left[x_k \mu_s^T\right] \Pi_{s,s}^{-1} \mu_s + E\left[v_k \mu_{k-1}^T\right] \Pi_{k-1,k-1}^{-1} \mu_{k-1} \\ &= \overline{\Theta}_k H_k \widehat{x}_{k/k-1} + \mathcal{V}_{k,k-1} \Pi_{k-1,k-1}^{-1} \mu_{k-1}, \quad k \geq 2,\end{aligned}$$

with $\widehat{y}_{1/0} = E[y_1] = \overline{\Theta}_1 H_1 E[x_1] = \overline{\Theta}_1 H_1 \widehat{x}_{1/0}$. Hence,

$$\begin{aligned}&\widehat{y}_{k/k-1} = \overline{\Theta}_k H_k \widehat{x}_{k/k-1} + \mathcal{V}_{k,k-1} \Pi_{k-1,k-1}^{-1} \mu_{k-1}, \quad k \geq 2; \\ &\widehat{y}_{1/0} = \overline{\Theta}_1 H_1 \widehat{x}_{1/0}\end{aligned} \tag{3.26}$$

and expression (3.13) for the innovation is clear.

Next, expression (3.14) for the matrix $\mathcal{X}_{k,k} = E[x_k y_k^T] - E[x_k \widehat{y}_{k/k-1}^T]$ is derived. From (3.5) and the independence assumption, it is clear that

$$E[x_k y_k^T] = D_k H_k^T \overline{\Theta}_k + E_k, \quad k \geq 1.$$

From expression (3.26) for $\widehat{y}_{k/k-1}$ and the OPL, $E[x_k \widehat{y}_{k/k-1}^T]$ is calculated as follows:

$$\begin{aligned}E[x_k \widehat{y}_{k/k-1}^T] &= E[x_k \widehat{x}_{k/k-1}^T] H_k^T \overline{\Theta}_k + E[x_k \mu_{k-1}^T] \Pi_{k-1,k-1}^{-1} \mathcal{V}_{k,k-1}^T \\ &= E[\widehat{x}_{k/k-1} \widehat{x}_{k/k-1}^T] H_k^T \overline{\Theta}_k + \mathcal{X}_{k,k-1} \Pi_{k-1,k-1}^{-1} \mathcal{V}_{k,k-1}^T \\ &= \left(D_k - P_{k/k-1}\right) H_k^T \overline{\Theta}_k + \mathcal{X}_{k,k-1} \Pi_{k-1,k-1}^{-1} \mathcal{V}_{k,k-1}^T, \quad k \geq 2; \\ E[x_1 \widehat{y}_{1/0}^T] &= \left(D_1 - P_{1/0}\right) H_1^T \overline{\Theta}_1.\end{aligned}$$

By substraction of the above expectations, expression (3.14) for $\mathcal{X}_{k,k} = E[x_k y_k^T] - E[x_k \widehat{y}_{k/k-1}^T]$ is obtained. From (3.1), expression (3.15) for $\mathcal{X}_{k,k-1} = E\left[x_k \mu_{k-1}^T\right]$ is immediately clear.

Expression (3.16) for the prediction error covariance matrix, $P_{k/k-1}$ is easily obtained by using (3.1) and (3.12); and, from (3.1) and (3.11), expression (3.17)

for $\mathcal{J}_k = E\left[\widetilde{x}_{k/k} w_k^T\right] = E\left[x_k w_k^T\right] - E\left[\widehat{x}_{k/k} w_k^T\right]$ is also obvious. Expression (3.18) for the filtering error covariance matrix, $P_{k/k}$, is immediately derived by using (3.11).

Finally, we prove expression (3.19) for the innovation covariance matrix $\Pi_{k,k} = E[y_k y_k^T] - E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T]$. From (3.5) and using (3.7), we have that

$$E[y_k y_k^T] = E[\theta_k \theta_k^T] \circ \left(H_k D_k H_k^T\right) + R_{k,k} + \overline{\Theta}_k H_k E_k + E_k^T H_k^T \overline{\Theta}_k.$$

Using now (3.26) for $\widehat{y}_{k/k-1}$ and property (3.7), and taking into account that, from the OPL, $E[\widehat{x}_{k/k-1}\mu_{k-1}^T] = E[x_k \mu_{k-1}^T] = \mathcal{X}_{k,k-1}$, the following identity holds:

$$\begin{aligned} E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T] = &\left(\overline{\theta}_k \overline{\theta}_k^T\right) \circ \left(H_k(D_k - P_{k/k-1})H_k^T\right) + \mathcal{V}_{k,k-1}\Pi_{k-1,k-1}^{-1}\mathcal{V}_{k,k-1}^T \\ &+ \overline{\Theta}_k H_k \mathcal{X}_{k,k-1}\Pi_{k-1,k-1}^{-1}\mathcal{V}_{k,k-1}^T + \mathcal{V}_{k,k-1}\Pi_{k-1,k-1}^{-1}\mathcal{X}_{k,k-1}^T H_k^T \overline{\Theta}_k. \end{aligned}$$

From the above expectations and after some manipulations, expression (3.19) for the innovation covariance matrix $\Pi_{k,k}$ is obtained. □

References

[1] C.K. Chui and G. Chen, Kalman Filtering: with Real-Time Applications, *Springer,* Berlín, 1999.

[2] R. Kalman, A new approach to linear filtering and prediction problems, *Transactions of the ASME. Journal of Basic Engineering*, 82 (Series D) (1960), 35–45.

[3] J.D. Jiménez-López, J. Linares-Pérez, S. Nakamori, R. Caballero-Águila and A. Hermoso-Carazo, Signal estimation based on covariance information from observations featuring correlated uncertainty and coming from multiple sensors, *Signal Processing,* 88 (2008), 2998–3006.

[4] D.Y. Kim and M. Jeon, Robust Distributed Kalman Filter for Wireless Sensor Networks with Uncertain Communication Channels, *Mathematical Problems in Engineering,* vol. 2012, Article ID 238597, 12 pages.

[5] J. Sun and C. Zhang, Distributed optimal fusion estimators for multi-sensor systems with bounded random measurement delays and packet dropouts, *Journal of Computational Information Systems,* 8(10) (2012), 4087–4094.

[6] M.J. García-Ligero, A. Hermoso-Carazo and J. Linares-Pérez, Distributed and centralized fusion estimation from multiple sensors with Markovian delays, *Applied Mathematics and Computation,* 219 (2012), 2932–2948.

[7] W.-A. Zhang, G. Feng and L. Yu, Multi-rate distributed fusion estimation for sensor networks with packet losses, *Automatica,* 48 (2012), 2016–2028.

[8] J. Ma and S. Sun, Information fusion estimators for systems with multiple sensors of different packet dropout rates, *Information Fusion,* 12 (2011), 213–222.

[9] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Least-squares linear estimators using measurements transmitted by different sensors with packet dropouts, *Digital Signal Processing,* 22 (2012), 1118–1125.

[10] E. Song, Y. Zhu, J. Zhou and Z. You, Optimal Kalman filtering fusion with cross-correlated sensor noises, *Automatica,* 43(8) (2007), 1450–1456.

[11] X. Wang, S.L. Sun, K.H. Ding and J.Y. Xue, Weighted Measurement Fusion White Noise Deconvolution Filter with Correlated Noise for Multisensor Stochastic Systems, *Mathematical Problems in Engineering,* 2012, Article ID 257619, 16 pages.

[12] J. Feng and M. Zeng, Optimal distributed Kalman filtering fusion for a linear dynamic system with cross-correlated noises, *International Journal of Systems Science,* 43(2) (2012), 385–398.

[13] E. Song, Y. Zhu and Z. You, The Kalman type recursive state estimator with a finite-step correlated process noises, In: *Proceedings of the IEEE International Conference on Automation and Logistics,* pp. 196–200, September 2008.

[14] A. Fu, Y. Zhu and E. Song, The optimal Kalman type state estimator with multi-step correlated process and measurement noises, In *The 2008 International Conference on Embedded Software and Systems,* pp. 215–220, 2008.

[15] J. Feng, Z. Wang, and M. Zeng, Optimal robust non-fragile Kalman-type recursive filtering with finite-step autocorrelated noises and multiple packet dropouts, *Aerospace Science and Technology,* 15(6) (2011), 486–494.

[16] J. Feng, Z. Wang and M. Zeng, Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises, *Information Fusion,* 14 (2013), 78–86.

[17] S. Sun and J. Ma, Optimal filtering and smoothing for discrete-time stochastic singular systems, *Signal Processing*, 87(1) (2007), 189–201.

[18] J. Ma and S. Sun, Optimal Linear Estimators for Systems With Random Sensor Delays, Multiple Packet Dropouts and Uncertain Observations, *IEEE Transactions on Signal Processing*, 59(11) (2011), 5181–5192.

[19] Z.Wang, F. Yang, D.W.C. Ho and X. Liu, Robust finite-horizon filtering for stochastic systems with missing measurements, *IEEE Signal Processing Letters,* 12(6) (2005), 437–440.

[20] Z. Wang, F. Yang, D.W.C. Ho and X. Liu, Robust H_∞ filtering for stochastic time-delay systems with missing measurements, *IEEE Transaction on Signal Processing,* 54(7) (2006), 2579–2587.

[21] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo and J. Linares- Pérez, New recursive estimators from correlated interrupted observations using covari-

ance information, *International Journal of Systems Science,* 36(10) (2005), 617–629.

[22] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo and J. Linares- Pérez, Fixed-interval smoothing algorithm based on covariances with correlation in the uncertainty, *Digital Signal Processing,* 15 (2005), 207–221.

[23] F.O. Hounkpevi and E.E. Yaz, Robust minimum variance linear state estimators for multiple sensors with different failure rates, *Automatica,* 43 (2007), 1274–1280.

[24] B. Shen, Z.D. Wang and Y.S. Hung, Distributed H_∞-consensus filtering in sensor networks with multiple missing measurements: The finite-horizon case, *Automatica,* 46 (2010), 1682–1688.

[25] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Linear and quadratic estimation using uncertain observations from multiple sensors with correlated uncertainty, *Signal Processing,* 91(2) (2011), 330–337.

[26] M. Liu, D. Qi, S. Zhang, M. Qiu and S. Zheng, Optimal H_∞ fusion filters for a class of discrete-time intelligent systems with time delays and missing measurement, *Neurocomputing,* 74 (2011), 3741–3752.

[27] X. Qu and J. Zhou, The optimal robust finite-horizon Kalman filtering for multiple sensors with different stochastic failure rates, *Applied Mathematics Letters,* 26 (2013), 80–86.

[28] G. Wei, Z. Wang and H. Shu, Robust filtering with stochastic nonlinearities and multiple missing measurements, *Automatica,* 45(3) (2009), 836–841.

[29] J. Hu, Z. Wang, H. Gao and L.K. Stergiouslas, Extended Kalman filtering with stochastic nonlinearities and multiple missing measurements, *Automatica,* 48 (2012), 2007–2015.

[30] S.L. Sun and Z.L. Deng, Multi-sensor optimal information fusion Kalman filter, *Automatica,* 40 (2004), 1017–1023.

Chapter 4

Optimal linear filter design for systems with correlation in the measurement matrices and noises: recursive algorithm and applications

Reference

Linares-Pérez, J., Caballero-Águila, R. and García-Garrido, I. (2014).
Optimal linear filter design for systems with correlation in the measurement matrices and noises: recursive algorithm and applications.
International Journal of Systems Science, 45: 1548-1562.
DOI:10.1080/00207721.2014.909093

Abstract

This paper addresses the optimal least-squares linear estimation problem for a class of discrete-time stochastic systems with random parameter matrices and correlated additive noises. The system presents the following main features: (1) one-step correlated and cross-correlated random parameter matrices in the observation

equation are assumed; (2) the process and measurement noises are one-step auto-correlated and two-step cross-correlated. Using an innovation approach and these correlation assumptions, a recursive algorithm with a simple computational procedure is derived for the optimal linear filter. As a significant application of the proposed results, the optimal recursive filtering problem in multi-sensor systems with missing measurements and random delays can be addressed. Numerical simulation examples are used to demonstrate the feasibility of the proposed filtering algorithm, which is also compared with other filters that have been proposed.

Contents

4.1 Introduction

The least-squares (LS) state estimation problem in discrete-time linear systems from noise measurements has been widely considered, due to its applicability in many practical situations. The Kalman filter provides a recursive algorithm for the optimal LS estimator when the model parameter matrices are deterministic and the additive white noises and the initial state are Gaussian and mutually independent. However, many real systems do not meet these requirements and new filtering algorithms have been reported for models representing the relationship between the unknown state and the observable variables and under different assumptions for the noise processes.

In recent decades, the filtering problem in multi-sensor systems, where sensor networks are used to obtain all available information on the system state, has become an issue of great interest for researchers. In data transmission, unreliable network characteristics can produce random sensor delays, multiple packet dropouts and uncertain observations (missing measurements). Due to these random uncertainties, standard observation models are not appropriate and estimation algorithms cannot be derived directly from Kalman filter theory. Accordingly, new algorithms are needed, and many research papers have been presented concerning the state estimation problem in multi-sensor systems with some of the aforementioned uncertainties (see [1]-[3], among others).

In systems with uncertain observations, besides the usual additive noise, the observation equation includes a multiplicative noise; hence, these systems are a special case of random measurement matrices. Moreover, systems with random sensor delays or multiple packet dropouts are transformed into systems with random measurement matrices in [4]. Systems with random state transition matrices can be used, for example, to describe randomly variant dynamic systems with multiple models [5] or linear systems with state-dependent multiplicative noise [6]. Discrete-time systems with random state transition and measurement parameter

matrices also arise in areas such as the digital control of chemical processes, systems with human operators, economic systems and stochastically sampled digital control systems [7].

In [7] and [5], the optimal linear filtering problem in systems with independent random state transition and measurement matrices is addressed by transforming the original system into one with deterministic parameter matrices and state-dependent process and measurement noises, to which the Kalman filter is applied. Although [7] applies the Kalman filter without providing any theoretical justification, [5] shows that under mild conditions, the transformed system satisfies the Kalman filter requirements and, hence, optimal linear estimators are derived for systems with independent random parameter matrices.

However, in many practical situations the random parameter matrices are not independent but correlated; for example, when random sensor delays and/or multiple packet dropouts are converted into observation models with random measurement matrices [4], or when a nonlinear system is linearized around the random state estimate to apply the extended Kalman filter (for other realistic systems and backgrounds where the model parameter matrices are random and correlated, see [8], [9]). In [10] systems with deterministic transition matrices and one-step correlated measurement matrices are considered, and the optimal recursive state estimation is derived by converting the observation equation into one with deterministic measurement matrices and applying the optimal Kalman filter for the case of one-step correlated measurement noise. In addition, a specific class of systems, where both the state transition and the measurement matrices are one-step moving average matrix sequences driven by a common independent zero-mean parameter sequence, is considered in the latter paper.

In the above-mentioned papers, although the noises of the transformed system with deterministic matrices depend on the system state and therefore can be correlated, the original system noises are assumed to be independent white pro-

cesses. This assumption can be restrictive in many real-world problems in which correlation and cross-correlation of the noises may be present. In systems with deterministic (state transition and measurement) matrices and with correlated and cross-correlated noises, the estimation problem has aroused significant research interest recently (see [6], [11] and references therein).

In view of the above considerations, this study focuses on the optimal LS linear filtering problem in systems with random parameter matrices and autocorrelated and cross-correlated noises, assuming independent random state transition matrices and one-step correlated and cross-correlated random parameter matrices in the observation equation. The proposed optimal LS linear recursive filtering algorithm can be applied to two significant classes of systems: *(i)* multi-sensor systems with missing measurements when the missing measurement phenomenon in each sensor is described by different sequences of correlated scalar random variables with arbitrary discrete probability distribution over the interval [0,1] (see Section 4.4); and *(ii)* multi-sensor systems with correlated random delays in the observations (see Section 4.5). In both cases, correlated and cross-correlated noises are considered. *This paper makes a substantial and novel contribution in two respects: (1) Unlike most existing results with random parameter matrices, in which a system transformation is carried out, the proposed optimal LS linear recursive filtering algorithm is obtained by using an innovation approach, without requiring any system transformation and, moreover, in which noise correlation is considered; (2) multi-sensor systems with missing and randomly delayed measurements can be considered as particular cases of the current random measurement matrices model and, hence, the proposed filter can be applied to these kind of multi-sensor systems. Besides these advantages, the filtering algorithm described is very simple computationally.*

The rest of this paper is organized as follows. Section 4.2 describes the system model with random state transition and measurement matrices, and autocorrelated and cross-correlated noises. In addition, some properties of the state and noise

processes derived from the correlation assumptions are specified. In Section 4.3, by using an innovation approach, a recursive algorithm for the optimal LS linear filter is obtained. In Sections 4.4 and 4.5, applications to multi-sensor systems with missing and randomly delayed measurements, respectively, are considered. In both sections, a numerical simulation example is presented to show the effectiveness of the proposed recursive filtering algorithm. Finally, some conclusions are drawn in Section 4.6.

Notation: The notation used throughout the paper is standard. For any matrix A, the notation symbols A^T and A^{-1} represent its transpose and inverse, respectively; $\mathbb{R}^n$ denotes the n-dimensional Euclidean space and $\mathbb{R}^{m\times n}$ is the set of $m \times n$ real matrices. The shorthand $Diag(A_1, \ldots, A_m)$ denotes a block diagonal matrix with matrices $A_1, \ldots, A_m$, and $[A_1 \mid \cdots \mid A_m]$ denotes a partitioned matrix into sub-matrices $A_1, \ldots, A_m$. If a matrix dimension is not explicitly stated, it is assumed to be compatible for algebraic operations. I and 0 represent the identity and zero matrices of appropriate dimensions. $\delta_{k,s}$ is the Kronecker delta function, which is equal to one, if $k = s$, and zero otherwise. $\circ$ denotes the Hadamard product. Moreover, for arbitrary random vectors X and Y, we denote $Cov[X, Y] = E[(X - E[X])\,(Y - E[Y])^T]$ and $Cov[X] = Cov[X, X]$, where $E[\cdot]$ stands for the mathematical expectation operator.

4.2 Discrete-time system model with random parameter matrices

Our aim in this paper is to address the optimal LS linear filtering problem in a class of discrete-time stochastic systems with random parameter matrices (independent random transition matrices and one-step correlated and cross-correlated matrices in the observation equation) and autocorrelated and cross-correlated noises. In this section, the system model is described and the statistical properties of the initial state, the random parameter matrices and the noise processes are identified.

Consider a class of discrete-time linear stochastic systems whose $n-$dimensional state process, $\{x_k\}_{k\geq 0}$, is perturbed by $n \times n$ random parameter matrices $\{F_k\}_{k\geq 0}$ and by an additive process noise $\{w_k\}_{k\geq 0}$; specifically, the state evolution is given by:

$$x_{k+1} = F_k x_k + w_k, \quad k \geq 0. \tag{4.1}$$

The measurements of the state are described by the following observation equation:

$$y_k = H_k x_k + B_k v_k, \quad k \geq 1, \tag{4.2}$$

where $\{y_k\}_{k\geq 1}$ is the $r-$dimensional observation process; the measurement matrices, $\{H_k\}_{k\geq 1}$, are $r \times n$ random parameter matrices; $\{B_k\}_{k\geq 1}$ are $r \times m$ random parameter matrices and the additive measurement noise, $\{v_k\}_{k\geq 1}$, is an $m-$dimensional process.

It is known that, if the state x_k and the observations $y_1, \ldots, y_k$ have finite second-order moments, then the optimal LS linear filter of x_k is the orthogonal projection of the vector x_k onto $\mathcal{L}(y_1, \ldots, y_k)$, i.e., the space of n-dimensional random variables obtained as linear transformations of $y_1, \ldots, y_k$. The hypotheses about the processes in (4.1) and (4.2) that guarantee the existence of the second-order moments of the vectors $y_1, \ldots, y_k$, as well as the correlation assumptions of the noise processes and the random parameter matrices in the observation equation are as follows:

(a) The initial state x_0 is a random vector with $E[x_0] = \overline{x}_0$ and $Cov[x_0] = P_0$, and it is independent of the random parameter matrices and noise processes.

(b) The random parameter matrices $\{F_k\}_{k\geq 0}$, $\{H_k\}_{k\geq 1}$ and $\{B_k\}_{k\geq 1}$ satisfy:

$$\begin{array}{l} E[F_k] = \overline{F}_k, \quad E[H_k] = \overline{H}_k, \quad E[B_k] = \overline{B}_k, \\ Cov[f_{ij}^k, f_{pq}^s] = C_{f_{ij}^k f_{pq}^k} \delta_{k,s}, \\ Cov[h_{ij}^k, h_{pq}^s] = C_{h_{ij}^k h_{pq}^k} \delta_{k,s} + C_{h_{ij}^k h_{pq}^s} \delta_{k,s-1} + C_{h_{ij}^k h_{pq}^s} \delta_{k,s+1}, \\ Cov[b_{ij}^k, b_{pq}^s] = C_{b_{ij}^k b_{pq}^k} \delta_{k,s} + C_{b_{ij}^k b_{pq}^s} \delta_{k,s-1} + C_{b_{ij}^k b_{pq}^s} \delta_{k,s+1}, \\ Cov[h_{ij}^k, b_{pq}^s] = C_{h_{ij}^k b_{pq}^k} \delta_{k,s} + C_{h_{ij}^k b_{pq}^s} \delta_{k,s-1} + C_{h_{ij}^k b_{pq}^s} \delta_{k,s+1}, \end{array}$$

where f_{ij}^k, h_{ij}^k and b_{ij}^k denote the $(i,j)-$th entries of matrices F_k, H_k and B_k, respectively. The mean matrices $\overline{F}_k$, $\overline{H}_k$ and $\overline{B}_k$ are known matrices $\forall k$, and $C_{f_{ij}^k f_{pq}^k}$, $C_{h_{ij}^k h_{pq}^s}$, $C_{b_{ij}^k b_{pq}^s}$ and $C_{h_{ij}^k b_{pq}^s}$, the covariances of the entries of the system random parameter matrices, are also assumed to be known $\forall k$ and $\forall s = k-1, k, k+1$.

(c) The process noise, $\{w_k\}_{k\geq 0}$, and the measurement noise, $\{v_k\}_{k\geq 1}$, are zero-mean sequences with the following covariances and cross-covariances:

$$\begin{aligned} Cov[w_k, w_s] &= Q_{k,k}\delta_{k,s} + Q_{k,s}\delta_{k,s-1} + Q_{k,s}\delta_{k,s+1}, \\ Cov[v_k, v_s] &= R_{k,k}\delta_{k,s} + R_{k,s}\delta_{k,s-1} + R_{k,s}\delta_{k,s+1}, \\ Cov[w_k, v_s] &= S_{k,k}\delta_{k,s} + S_{k,s}\delta_{k,s-1} + S_{k,s}\delta_{k,s-2}. \end{aligned}$$

(d) Independence assumptions:

- $\{F_k\}_{k\geq 0}$ is independent of $(\{H_k\}_{k\geq 1}, \{B_k\}_{k\geq 1}, \{w_k\}_{k\geq 0}, \{v_k\}_{k\geq 1})$.
- $(\{H_k\}_{k\geq 1}, \{B_k\}_{k\geq 1})$ is independent of $(\{F_k\}_{k\geq 0}, \{w_k\}_{k\geq 0}, \{v_k\}_{k\geq 1})$.

Remark 1: correlation of the noise processes. The correlation hypothesis *(c)* of the process noise and the measurement noise is the same as those given in [6] and [11]. Specifically, both noise processes are correlated at consecutive sampling times and independent otherwise, and the measurement noise vector v_k is correlated with the noise vectors w_s, for $s = k, k-1, k-2$, and independent otherwise. Systems with only finite-step correlated process noise or multi-step correlated process and measurement noise are considered in [12]-[14], among others.

As a consequence of the noise correlation assumptions, it is easy to see that:

- The vectors w_k and y_k are correlated, with

$$\mathcal{W}_k := E\left[w_k y_k^T\right] = Q_{k,k-1}\overline{H}_k^T + S_{k,k}\overline{B}_k^T, \quad k \geq 1. \tag{4.3}$$

- The state vector x_k is correlated with the observation noise vectors v_k and v_{k-1}, with

$$\begin{aligned} \mathcal{E}_{k,k} &:= E[x_k v_k^T] = \overline{F}_{k-1} S_{k-2,k} + S_{k-1,k}, \quad k \geq 2; \qquad \mathcal{E}_{1,1} = S_{0,1}, \\ \mathcal{E}_{k,k-1} &:= E[x_k v_{k-1}^T] = \overline{F}_{k-1} \mathcal{E}_{k-1,k-1} + S_{k-1,k-1}, \quad k \geq 2. \end{aligned} \tag{4.4}$$

Remark 2: correlation of the random parameter matrices in the observation equation. Besides considering autocorrelated and cross-correlated noises (assumption (c)), the correlation assumption (b) of the random parameter matrices $\{H_k\}_{k\geq 1}$ and $\{B_k\}_{k\geq 1}$ in the observation equation is the main difference between the current model and the model in [5], where $\{H_k\}_{k\geq 1}$ is assumed to be a sequence of independent random parameter matrices, and the observation noise is not multiplied by random parameter matrices. The correlation of the measurement matrices $\{H_k\}_{k\geq 1}$ at consecutive sampling times allows us to apply the results proposed in this paper to multi-sensor systems with correlated missing measurements (see Section 4.4). In addition, the correlation and cross-correlation of the random parameter matrices $\{H_k\}_{k\geq 1}$ and $\{B_k\}_{k\geq 1}$ allow us to consider multi-sensor systems with correlated randomly delayed measurements as a particular case of the current study (see Section 4.5).

As a consequence of the correlation assumptions of the noises and random parameter matrices, using the conditional expectation properties, and denoting $\widetilde{H}_k = H_k - \overline{H}_k$ and $\widetilde{B}_k = B_k - \overline{B}_k$, it can be seen that the vector $B_k v_k$ is correlated with the observation vector y_{k-1}, and

$$\begin{aligned} \mathcal{V}_{k,k-1} := E\left[B_k v_k y_{k-1}^T\right] &= \overline{B}_k \left(\overline{H}_{k-1} S_{k-2,k} + \overline{B}_{k-1} R_{k-1,k}\right)^T \\ &\quad + E\left[\widetilde{B}_k S_{k-2,k}^T \widetilde{H}_{k-1}^T\right] + E\left[\widetilde{B}_k R_{k,k-1} \widetilde{B}_{k-1}^T\right], \quad k \geq 2. \end{aligned} \tag{4.5}$$

The matrix $E\left[\widetilde{B}_k S_{k-2,k}^T \widetilde{H}_{k-1}^T\right]$ is yielded by both the cross-correlation of the noise processes and that of the random parameter matrices $\{H_k\}_{k\geq 1}$ and $\{B_k\}_{k\geq 1}$, while the matrix $E\left[\widetilde{B}_k R_{k,k-1} \widetilde{B}_{k-1}^T\right]$ arises because of the one-step correlation of

the measurement noise and also that of the matrices $\{B_k\}_{k\geq 1}$. From assumption *(b)*, the $(p,q)-$th entries of these matrices are obtained by:

$$\left(E[\widetilde{B}_k S_{k-2,k}^T \widetilde{H}_{k-1}^T]\right)_{pq} = \sum_{j=1}^{m}\sum_{i=1}^{n} C_{b_{qj}^k h_{pi}^{k-1}} \left(S_{k-2,k}^T\right)_{ji} \quad (p,q=1,2,\ldots,r),$$
$$\left(E[\widetilde{B}_k R_{k,k-1} \widetilde{B}_{k-1}^T]\right)_{pq} = \sum_{j=1}^{m}\sum_{i=1}^{m} C_{b_{qj}^k b_{pi}^{k-1}} \left(R_{k,k-1}\right)_{ji} \quad (p,q=1,2,\ldots,r).$$

Remark 3: state transition equation. Linear discrete-time systems with random state transition matrices, $\{F_k\}_{k\geq 0}$, have important applications; for example, they can be used to describe randomly variant dynamic systems with multiple models [5] or linear systems with state-dependent multiplicative noise [6]. Furthermore, bilinear stochastic systems [15] can be reduced to models with random transition matrices.

After denoting $\widetilde{F}_k = F_k - \overline{F}_k$, from the state equation (4.1) and the conditional expectation properties, it is easy to deduce that $\mathcal{D}_{k+1} = E[x_{k+1}x_{k+1}^T]$ is recursively calculated by:

$$\begin{aligned} &\mathcal{D}_{k+1} = \overline{F}_k \mathcal{D}_k \overline{F}_k^T + E[\widetilde{F}_k \mathcal{D}_k \widetilde{F}_k^T] + Q_{k,k} + \overline{F}_k Q_{k-1,k} + Q_{k,k-1}\overline{F}_k^T, \quad k \geq 1; \\ &\mathcal{D}_1 = \overline{F}_0 \mathcal{D}_0 \overline{F}_0^T + E[\widetilde{F}_0 \mathcal{D}_0 \widetilde{F}_0^T] + Q_{0,0}, \qquad \mathcal{D}_0 = P_0 + \overline{x}_0 \overline{x}_0^T, \end{aligned} \tag{4.6}$$

where, from assumption *(b)*, the $(p,q)-$th entry of the matrix $E[\widetilde{F}_k \mathcal{D}_k \widetilde{F}_k^T]$ is obtained by

$$\left(E[\widetilde{F}_k \mathcal{D}_k \widetilde{F}_k^T]\right)_{pq} = \sum_{j=1}^{n}\sum_{i=1}^{n} C_{f_{qj}^k f_{pi}^k} \left(\mathcal{D}_k\right)_{ji} \quad (p,q=1,2,\ldots,n).$$

Also, from the state equation (4.1), it is immediately clear that $\mathcal{G}_{k+1,k} = E[x_{k+1}x_k^T]$ satisfies

$$\mathcal{G}_{k+1,k} = \overline{F}_k \mathcal{D}_k + Q_{k,k-1}, \quad k \geq 1; \qquad \mathcal{G}_{1,0} = \overline{F}_0 \mathcal{D}_0. \tag{4.7}$$

4.3 Optimal LS linear estimation problem

Given the observations up to time k, $\{y_1, \ldots, y_k\}$, our aim is to derive a recursive algorithm for the optimal LS linear filter, $\widehat{x}_{k/k}$, of the state x_k. Since $\widehat{x}_{k/k}$ is the orthogonal projection of x_k onto the space $\mathcal{L}(y_1, \ldots, y_k)$ of linear transformations of $y_1, \ldots, y_k$, and these observations are generally non-orthogonal vectors, an innovation approach will be used. This approach considerably simplifies the algorithm derivation, because the innovation process is a white noise. The innovation approach is based on the Gram-Schmidt orthogonalization procedure by means of which the observation process $\{y_k\}_{k\geq 1}$ is transformed into an equivalent process (*innovation process*) $\{\mu_k\}_{k\geq 1}$, equivalent in the sense that $\mathcal{L}(\mu_1, \ldots, \mu_k) = \mathcal{L}(y_1, \ldots, y_k)$; that is, each set $\{\mu_1, \ldots, \mu_k\}$ spans the same linear subspace as $\{y_1, \ldots, y_k\}$.

The innovation at time k is defined as $\mu_k = y_k - \widehat{y}_{k/k-1}$, where $\widehat{y}_{k/k-1}$, the one-stage LS linear predictor of y_k, is the projection of y_k onto $\mathcal{L}(\mu_1, \ldots, \mu_{k-1})$. The orthogonality property allows us to find the projection by projecting onto each of the previous orthogonal vectors separately; that is,

$$\widehat{y}_{k/k-1} = \sum_{i=1}^{k-1} E[y_k \mu_i^T](E[\mu_i \mu_i^T])^{-1}\mu_i, \quad k \geq 2; \qquad \widehat{y}_{1/0} = \overline{H}_1 \widehat{x}_{1/0}. \tag{4.8}$$

Similarly, by denoting $\mathcal{X}_{k,i} = E[x_k \mu_i^T]$ and $\Pi_i = E[\mu_i \mu_i^T]$, a general expression for the optimal LS linear filter, $\widehat{x}_{k/k}$, as a linear combination of the innovations is obtained; namely,

$$\widehat{x}_{k/k} = \sum_{i=1}^{k} \mathcal{X}_{k,i}\Pi_i^{-1}\mu_i, \quad k \geq 1; \quad \widehat{x}_{0/0} = \overline{x}_0,$$

and, the following expression for the filter, $\widehat{x}_{k/k}$, in terms of the predictor, $\widehat{x}_{k/k-1}$, is obvious:

$$\widehat{x}_{k/k} = \widehat{x}_{k/k-1} + \mathcal{X}_{k,k}\Pi_k^{-1}\mu_k, \quad k \geq 1; \quad \widehat{x}_{0/0} = \overline{x}_0. \tag{4.9}$$

Next, we obtain the state predictor $\widehat{x}_{k/k-1}$, the innovation μ_k and its covariance matrix Π_k, and the matrix $\mathcal{X}_{k,k}$, which together with (4.9) will constitute the proposed recursive linear filtering algorithm.

4.3.1 State predictor $\widehat{x}_{k/k-1}$

In systems with random parameter matrices and uncorrelated additive white noises [5], the one-stage state predictor is calculated as $\widehat{x}_{k/k-1} = \overline{F}_{k-1}\widehat{x}_{k-1/k-1}$; this is because the uncorrelation assumption of the noises guarantees that $\widehat{w}_{k-1/k-1} = 0$. However, this is not true for the problem at hand, where the noise estimator $\widehat{w}_{k-1/k-1}$ must be taken into account in order to derive the one-stage state predictor. From the Orthogonal Projection Lemma (OPL), we have

$$\widehat{x}_{k/k-1} = \overline{F}_{k-1}\widehat{x}_{k-1/k-1} + \widehat{w}_{k-1/k-1}, \quad k \geq 1,$$

and hence, an expression for the noise filter $\widehat{w}_{k/k}$ is necessary. Taking into account that w_k is independent of $\mu_1, \ldots, \mu_{k-1}$ and $\widehat{y}_{k/k-1}$, we have,

$$\widehat{w}_{k/k} = \sum_{i=1}^{k} E[w_k\mu_i^T]\Pi_i^{-1}\mu_i = E[w_k\mu_k^T]\Pi_k^{-1}\mu_k = E[w_k y_k^T]\Pi_k^{-1}\mu_k, \quad k \geq 1; \qquad \widehat{w}_{0/0} = 0.$$

Therefore, the state predictor, $\widehat{x}_{k/k-1}$, satisfies

$$\widehat{x}_{k/k-1} = \overline{F}_{k-1}\widehat{x}_{k-1/k-1} + \mathcal{W}_{k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k \geq 2; \qquad \widehat{x}_{1/0} = \overline{F}_0\widehat{x}_{0/0}, \tag{4.10}$$

where $\mathcal{W}_k$ is given by (4.3).

4.3.2 Prediction $P_{k/k-1}$ and filtering $P_{k/k}$ error covariance matrices

From (4.1) and (4.10), it is easy to see that the prediction error covariance matrix, $P_{k/k-1}$, satisfies

$$\begin{aligned} P_{k/k-1} &= \overline{F}_{k-1}P_{k-1/k-1}\overline{F}_{k-1}^T + E[\widetilde{F}_{k-1}\mathcal{D}_{k-1}\widetilde{F}_{k-1}^T] + Q_{k-1,k-1} \\ &\quad + \overline{F}_{k-1}\mathcal{J}_{k-1} + \mathcal{J}_{k-1}^T\overline{F}_{k-1}^T - \mathcal{W}_{k-1}\Pi_{k-1}^{-1}\mathcal{W}_{k-1}^T, \quad k \geq 2; \\ P_{1/0} &= \overline{F}_0 P_{0/0}\overline{F}_0^T + E[\widetilde{F}_0\mathcal{D}_0\widetilde{F}_0^T] + Q_{0,0}, \end{aligned} \tag{4.11}$$

where, using (4.9), it is clear that $\mathcal{J}_k = E\left[\left(x_k - \widehat{x}_{k/k}\right) w_k^T\right]$ is calculated by

$$\mathcal{J}_k = Q_{k-1,k} - \mathcal{X}_{k,k}\Pi_k^{-1}\mathcal{W}_k^T, \quad k \geq 1. \tag{4.12}$$

Again, from (4.9), the filtering error covariance matrix, $P_{k/k}$, is given by

$$P_{k/k} = P_{k/k-1} - \mathcal{X}_{k,k}\Pi_k^{-1}\mathcal{X}_{k,k}^T, \quad k \geq 1; \qquad P_{0/0} = P_0. \tag{4.13}$$

4.3.3 Innovation $\mu_k = y_k - \widehat{y}_{k/k-1}$

In [5], the one-stage observation predictor is calculated as $\widehat{y}_{k/k-1} = \overline{H}_k\widehat{x}_{k/k-1}$; this is because the uncorrelation assumption of the noises guarantees that $\widehat{v}_{k/k-1} = 0$. However, due to the correlation assumptions of the measurement matrices, *(b)*, and the noise processes, *(c)*, this is not true for the problem at hand and both the correlation of H_{k-1} and H_k and the noise estimator $\widehat{v}_{k/k-1}$, must be taken into account in deriving the predictor $\widehat{y}_{k/k-1}$.

Therefore, to obtain the innovation $\mu_k = y_k - \widehat{y}_{k/k-1}$, it is necessary to find a new expression for $\widehat{y}_{k/k-1}$. For this purpose, taking into account (4.8), we first calculate

$$E[y_k\mu_i^T] = E[H_k x_k \mu_i^T] + E[B_k v_k \mu_i^T] = \begin{cases} \overline{H}_k\mathcal{X}_{k,i}, & i \leq k-2, \\ E[H_k x_k \mu_{k-1}^T] + \mathcal{V}_{k,k-1}, & i = k-1, \end{cases} \tag{4.14}$$

and, substituting the expectations (4.14) in (4.8), we obtain

$$\widehat{y}_{k/k-1} = \overline{H}_k \sum_{i=1}^{k-1} \mathcal{X}_{k,i}\Pi_i^{-1}\mu_i + \left(E[H_k x_k \mu_{k-1}^T] + \mathcal{V}_{k,k-1} - \overline{H}_k\mathcal{X}_{k,k-1}\right)\Pi_{k-1}^{-1}\mu_{k-1}.$$

Now, from the conditional expectation properties, we obtain that

$$\begin{aligned} E[H_k x_k \mu_{k-1}^T] - \overline{H}_k\mathcal{X}_{k,k-1} &= E[\widetilde{H}_k x_k \mu_{k-1}^T] = E[\widetilde{H}_k x_k y_{k-1}^T] \\ &= E[\widetilde{H}_k\mathcal{G}_{k,k-1}\widetilde{H}_{k-1}^T] + E[\widetilde{H}_k\mathcal{E}_{k,k-1}\widetilde{B}_{k-1}^T], \end{aligned}$$

where $\mathcal{G}_{k,k-1}$ and $\mathcal{E}_{k,k-1}$ are given in (4.7) and (4.4), respectively.

Hence, it is concluded that the one-stage observation predictor satisfies

$$\widehat{y}_{k/k-1} = \overline{H}_k \widehat{x}_{k/k-1} + \Psi_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k \geq 1, \tag{4.15}$$

where

$$\Psi_{k,k-1} = E[\widetilde{H}_k \mathcal{G}_{k,k-1}\widetilde{H}_{k-1}^T] + E[\widetilde{H}_k \mathcal{E}_{k,k-1}\widetilde{B}_{k-1}^T] + \mathcal{V}_{k,k-1}, \quad k \geq 2; \quad \Psi_{1,0} = 0. \tag{4.16}$$

It can be observed that the matrices $E[\widetilde{H}_k \mathcal{G}_{k,k-1}\widetilde{H}_{k-1}^T]$ and $E[\widetilde{H}_k \mathcal{E}_{k,k-1}^T\widetilde{B}_{k-1}^T]$ are yielded by the correlation of the random parameter matrices of the observation equation at consecutive sampling times. From *(b)*, the $(p,q)-$th entries of these matrices are obtained by

$$\left(E[\widetilde{H}_k \mathcal{G}_{k,k-1}\widetilde{H}_{k-1}^T]\right)_{pq} = \sum_{j=1}^{n}\sum_{i=1}^{n} C_{h_{qj}^k h_{pi}^{k-1}} \left(\mathcal{G}_{k,k-1}\right)_{ji} \quad (p,q = 1,2,\ldots,r),$$

$$\left(E[\widetilde{H}_k \mathcal{E}_{k,k-1}^T\widetilde{B}_{k-1}^T]\right)_{pq} = \sum_{j=1}^{n}\sum_{i=1}^{m} C_{h_{qj}^k b_{pi}^{k-1}} \left(\mathcal{E}_{k,k-1}^T\right)_{ji} \quad (p,q = 1,2,\ldots,r).$$

Hence, the innovation μ_k is obtained as a linear combination of the new observation, the state predictor and the previous innovation:

$$\mu_k = y_k - \overline{H}_k \widehat{x}_{k/k-1} - \Psi_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k \geq 1. \tag{4.17}$$

4.3.4 Matrix $\mathcal{X}_{k,k} = E[x_k \mu_k^T]$

Next, an expression for the matrix $\mathcal{X}_{k,k} = E[x_k\mu_k^T] = E[x_k y_k^T] - E[x_k \widehat{y}_{k/k-1}^T]$ is derived. From (4.2) and (4.4), it is clear that

$$E[x_k y_k^T] = \mathcal{D}_k \overline{H}_k^T + \mathcal{E}_{k,k}\overline{B}_k^T, \quad k \geq 1.$$

From (4.15) and since, from the OPL, $E[x_k \widehat{x}_{k/k-1}^T] = \mathcal{D}_k - P_{k/k-1}$, we obtain:

$$E[x_k \widehat{y}_{k/k-1}^T] = \left(\mathcal{D}_k - P_{k/k-1}\right)\overline{H}_k^T + \mathcal{X}_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T, \quad k \geq 1,$$

where $\mathcal{X}_{k,k-1} = E\left[x_k\mu_{k-1}^T\right]$ satisfies

$$\mathcal{X}_{k,k-1} = \overline{F}_{k-1}\mathcal{X}_{k-1,k-1} + \mathcal{W}_{k-1}, \quad k \geq 2. \tag{4.18}$$

By subtracting the above expectations, the following expression for $\mathcal{X}_{k,k}$ is derived

$$\mathcal{X}_{k,k} = P_{k/k-1}\overline{H}_k^T + \mathcal{E}_{k,k}\overline{B}_k^T - \mathcal{X}_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T, \quad k \geq 1. \tag{4.19}$$

4.3.5 Innovation covariance matrix $\Pi_k = E[\mu_k \mu_k^T]$

Finally, we obtain an expression for $\Pi_k = E[\mu_k\mu_k^T] = E[y_k y_k^T] - E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T]$. From (4.2) and again using the conditional expectation properties, we have

$$\begin{aligned} E[y_k y_k^T] =& \overline{H}_k \mathcal{D}_k \overline{H}_k^T + E[\widetilde{H}_k \mathcal{D}_k \widetilde{H}_k^T] + \overline{B}_k R_{k,k} \overline{B}_k^T + E[\widetilde{B}_k R_{k,k} \widetilde{B}_k^T] \\ &+ \overline{H}_k \mathcal{E}_{k,k} \overline{B}_k^T + E[\widetilde{H}_k \mathcal{E}_{k,k} \widetilde{B}_k^T] + \overline{B}_k \mathcal{E}_{k,k}^T \overline{H}_k^T + E[\widetilde{B}_k \mathcal{E}_{k,k}^T \widetilde{H}_k^T], \end{aligned}$$

where $\mathcal{D}_k$ and $\mathcal{E}_{k,k}$ are given in (4.6) and (4.4), respectively.

Using (4.15) and since $E[\widehat{x}_{k/k-1}\mu_{k-1}^T] = E[x_k \mu_{k-1}^T] = \mathcal{X}_{k,k-1}$, the following identity holds:

$$\begin{aligned} E[\widehat{y}_{k/k-1}\widehat{y}_{k/k-1}^T] =& \overline{H}_k(\mathcal{D}_k - P_{k/k-1})\overline{H}_k^T + \Psi_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T \\ &+ \overline{H}_k \mathcal{X}_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T + \Psi_{k,k-1}\Pi_{k-1}^{-1}\mathcal{X}_{k,k-1}^T \overline{H}_k^T. \end{aligned}$$

From the above expectations, using (4.19) and after some manipulations, the following expression for the innovation covariance matrix Π_k is obtained:

$$\begin{aligned} \Pi_k =& E[\widetilde{H}_k \mathcal{D}_k \widetilde{H}_k^T] + E[\widetilde{B}_k R_{k,k} \widetilde{B}_k^T] + E[\widetilde{B}_k \mathcal{E}_{k,k}^T \widetilde{H}_k^T] + E[\widetilde{H}_k \mathcal{E}_{k,k} \widetilde{B}_k^T] + \overline{B}_k R_{k,k} \overline{B}_k^T \\ &+ \overline{H}_k \mathcal{X}_{k,k} + \mathcal{X}_{k,k}^T \overline{H}_k^T - \overline{H}_k P_{k/k-1} \overline{H}_k^T - \Psi_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T, \quad k \geq 1. \end{aligned} \tag{4.20}$$

It can be observed that the matrices $E[\widetilde{H}_k \mathcal{D}_k \widetilde{H}_k^T]$, $E[\widetilde{B}_k R_{k,k} \widetilde{B}_k^T]$ and $E[\widetilde{B}_k \mathcal{E}_{k,k}^T \widetilde{H}_k^T]$ are yielded by the correlation of the random matrices of the observation equation at the same time instant. From *(b)*, the $(p,q)-$th entries of these matrices are

obtained by

$$\left(E[\widetilde{H}_k\mathcal{D}_k\widetilde{H}_k^T]\right)_{pq} = \sum_{j=1}^{n}\sum_{i=1}^{n} C_{h_{qj}^k h_{pi}^k}\left(\mathcal{D}_k\right)_{ji} \quad (p,q=1,2,\ldots,r),$$

$$\left(E[\widetilde{B}_k R_{k,k}\widetilde{B}_k^T]\right)_{pq} = \sum_{j=1}^{m}\sum_{i=1}^{m} C_{b_{qj}^k b_{pi}^k}\left(R_{k,k}\right)_{ji} \quad (p,q=1,2,\ldots,r),$$

$$\left(E[\widetilde{B}_k\mathcal{E}_{k,k}^T\widetilde{H}_k^T]\right)_{pq} = \sum_{j=1}^{m}\sum_{i=1}^{n} C_{b_{qj}^k h_{pi}^k}\left(\mathcal{E}_{k,k}^T\right)_{ji} \quad (p,q=1,2,\ldots,r).$$

4.3.6 Filtering algorithm: computational procedure and advantages

The optimal LS linear filtering algorithm is constituted by equations (4.9)-(4.13) and (4.17)-(4.20), and the computational procedure can be summarized as follows:

i) The matrices $\mathcal{W}_k$, $\mathcal{E}_{k,k}$, $\mathcal{E}_{k,k-1}$ and $\mathcal{V}_{k,k-1}$ are computed by expressions (4.3)-(4.5). We then recursively compute $\mathcal{D}_k$ by (4.6) and thus $\mathcal{G}_{k,k-1}$ by (4.7); with the matrices $\mathcal{E}_{k,k-1}$, $\mathcal{V}_{k,k-1}$ and $\mathcal{G}_{k,k-1}$, we can compute $\Psi_{k,k-1}$ by (4.16). The matrices $E[\widetilde{B}_k\mathcal{E}_{k,k}^T\widetilde{H}_k^T]$, $E[\widetilde{H}_k\mathcal{D}_k\widetilde{H}_k^T]$ and $E[\widetilde{B}_k R_{k,k}\widetilde{B}_k^T]$ are also computed in order to obtain the innovation covariance matrix Π_k. Note that all these matrices depend only on the system model information and can be obtained before the observations are available.

ii) At the sampling time k, when the $(k-1)$th iteration is finished and the new observation y_k is available, the proposed filtering algorithm operates as follows (Figure 1):

1) By (4.18), we compute $\mathcal{X}_{k,k-1} = \overline{F}_{k-1}\mathcal{X}_{k-1,k-1} + \mathcal{W}_{k-1}$ and, from this, $\mathcal{X}_{k,k}$ by (4.19):

$$\mathcal{X}_{k,k} = P_{k/k-1}\overline{H}_k^T + \mathcal{E}_{k,k}\overline{B}_k^T - \mathcal{X}_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T.$$

2) The innovation μ_k and its covariance matrix Π_k are computed by (4.17) and (4.20), respectively:

$$\begin{aligned}\mu_k =& y_k - \overline{H}_k \widehat{x}_{k/k-1} - \Psi_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1},\\ \Pi_k =& E[\widetilde{H}_k \mathcal{D}_k \widetilde{H}_k^T] + E[\widetilde{B}_k R_{k,k} \widetilde{B}_k^T] + E[\widetilde{B}_k \mathcal{E}_{k,k}^T \widetilde{H}_k^T] + E[\widetilde{H}_k \mathcal{E}_{k,k} \widetilde{B}_k^T]\\ &+ \overline{B}_k R_{k,k} \overline{B}_k^T + \overline{H}_k \mathcal{X}_{k,k} + \mathcal{X}_{k,k}^T \overline{H}_k^T - \overline{H}_k P_{k/k-1} \overline{H}_k^T\\ &- \Psi_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T.\end{aligned}$$

3) The filter $\widehat{x}_{k/k}$ and the filtering error covariance matrix $P_{k/k}$ are computed by (4.9) and (4.13), respectively:

$$\begin{aligned}\widehat{x}_{k/k} &= \widehat{x}_{k/k-1} + \mathcal{X}_{k,k}\Pi_k^{-1}\mu_k,\\ P_{k/k} &= P_{k/k-1} - \mathcal{X}_{k,k}\Pi_k^{-1}\mathcal{X}_{k,k}^T.\end{aligned}$$

4) To implement the above steps at time $k+1$, we must:

- Compute the state predictor by (4.10): $\widehat{x}_{k+1/k} = \overline{F}_k \widehat{x}_{k/k} + \mathcal{W}_k \Pi_k^{-1}\mu_k$.
- Compute $\mathcal{J}_k = Q_{k-1,k} - \mathcal{X}_{k,k}\Pi_k^{-1}\mathcal{W}_k^T$ by (4.12), and from this, the prediction error covariance matrix $P_{k+1/k}$ by (4.11):

$$\begin{aligned}P_{k+1/k} =& \overline{F}_k P_{k/k} \overline{F}_k^T + E[\widetilde{F}_k \mathcal{D}_k \widetilde{F}_k^T] + Q_{k,k}\\ &+ \overline{F}_k \mathcal{J}_k + \mathcal{J}_k^T \overline{F}_k^T - \mathcal{W}_k \Pi_k^{-1} \mathcal{W}_k^T.\end{aligned}$$

The proposed algorithm has the following advantages: 1) the filter is globally optimal in the linear LS sense; 2) the filter structure is recursive, very simple computationally and suitable for online applications; 3) the algorithm takes into account both the influence of the correlation of the random parameter matrices and that of the noises; 4) the algorithm, obtained by an innovation approach, does not require any transformation of the original system into one with deterministic parameter matrices; 5) the proposed filter can be applied to multi-sensor systems with correlated missing measurements considering at each sensor the possibility of observations containing only partial information about the state (Section 4.4); 6) the proposed filter can be applied to multi-sensor systems with randomly delayed measurements correlated at consecutive sampling times (Section 4.5).

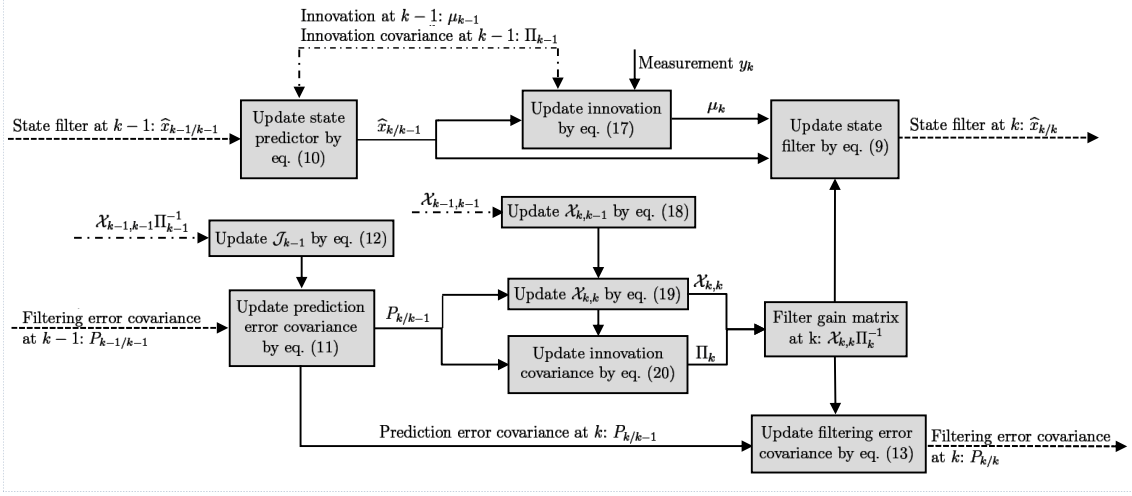


Figure 4.1: Optimal LS linear filtering algorithm.

4.4 Application to multi-sensor systems with missing measurements

Over the past few decades, considerable research has been carried out into multi-sensor systems with missing measurements, due to the importance of this question and its applicability to modelling a broad class of real-world problems. Most papers concerning systems with missing measurements transmitted by multiple sensors assume that the missing probabilities in all the sensors are identical (see [16], [17]). In recent years, however, this situation has been generalized to address missing measurements whose statistical properties are not assumed to be the same for all the sensors (see [18], [19]). Different missing probabilities have also been considered for some classes of nonlinear systems in [20], [21], where quantized H_∞ control and filtering problems are addressed, respectively. This is a realistic assumption in several application fields, for instance, in networked communication systems involving heterogeneous measurement devices. In [18], different sequences of independent Bernoulli random variables are used to describe the missing measurement phenomenon at each sensor. [22] subsequently generalized these results, by weak-

ening the independence restriction and considering sequences of Bernoulli random variables correlated at consecutive sampling times. This form of correlation covers practical situations where the state cannot be missing in two successive observations (for example, transmission models with stand-by sensors in which any failure in the transmission is detected immediately and the old sensor is then replaced). In all of the above papers, it is assumed that the state measurement is either completely lost or successfully transferred, and Bernoulli random variables are used to model the missing measurement phenomenon. In a more recent study, this missing measurement model was generalized to consider an arbitrary discrete distribution in the interval $[0, 1]$, thus covering some practical applications where only partial information is missing (see [11], [23] and references therein).

Our aim in this section is to show that the observation model in multi-sensor systems with missing measurements can be considered a special case of the observation model with random measurement matrices (4.2). Hence, the proposed optimal LS linear filtering algorithm can be applied to multiple missing measurement systems with correlated and cross-correlated noises, when the missing measurement phenomenon at each sensor is described by different sequences of correlated (at consecutive sampling times) scalar random variables with arbitrary discrete probability distribution over the interval [0,1]. In particular, the proposed optimal LS linear filtering algorithm extends the results in [5], [11] and [22], among others.

Accordingly, consider the state equation given by (4.1), with $\{F_k\}_{k\geq 0}$ and $\{w_k\}_{k\geq 0}$ verifying the hypotheses (b) and (c), and r sensors which, at any time k, provide scalar measurements of the state, perturbed by additive and multiplicative noises according to the following observation model:

$$y_k^i = \theta_k^i C_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2, \ldots, r, \tag{4.21}$$

where $\{y_k^i\}_{k\geq 1}$ are the measured data from the $i-$th sensor, $\{C_k^i\}_{k\geq 1}$, are known time-varying matrices with compatible dimensions, $\{v_k^i\}_{k\geq 1}$ are zero-mean measurement noises, and $\{\theta_k^i\}_{k\geq 1}$ are different sequences of scalar discrete-time random

variables over the interval $[0,1]$, with $E[\theta_k^i] = \overline{\theta}_k^i$. For $i,j = 1,\ldots,r$, the following noise correlation assumptions are made:

$$\begin{aligned} Cov[\theta_k^i, \theta_s^j] &= K_{k,k}^{\theta^{ij}}\delta_{k,s} + K_{k,s}^{\theta^{ij}}\delta_{k,s-1} + K_{k,s}^{\theta^{ij}}\delta_{k,s+1}, \\ Cov[v_k^i, v_s^j] &= R_{k,k}^{ij}\delta_{k,s} + R_{k,s}^{ij}\delta_{k,s-1} + R_{k,s}^{ij}\delta_{k,s+1}, \\ Cov[w_k, v_s^i] &= S_{k,k}^{i}\delta_{k,s} + S_{k,s}^{i}\delta_{k,s-1} + S_{k,s}^{i}\delta_{k,s-2}. \end{aligned}$$

The observation model (4.21) can be rewritten as follows:

$$y_k = H_k x_k + v_k, \quad k \geq 1,$$

where $y_k = (y_k^1,\ldots,y_k^r)^T$ is the $r-$dimensional observation vector, $H_k = \Theta_k C_k$, with $\Theta_k = Diag(\theta_k^1, \ldots, \theta_k^r)$ and $C_k = \left[C_k^{1T} \mid \cdots \mid C_k^{rT}\right]^T$, are $r \times n$ random parameter matrices, and $v_k = (v_k^1,\ldots,v_k^r)^T$ is the $r-$dimensional noise vector. Hence, the observation model (4.21) is a particular case of (4.2) with $B_k = I$, and clearly verifies that:

- the additive noise $\{v_k\}_{k\geq 1}$ is autocorrelated and cross-correlated with $\{w_k\}_{k\geq 1}$, with

$$\begin{aligned} Cov[v_k, v_s] &= R_{k,k}\delta_{k,s} + R_{k,s}\delta_{k,s-1} + R_{k,s}\delta_{k,s+1}, \\ Cov[w_k, v_s] &= S_{k,k}\delta_{k,s} + S_{k,s}\delta_{k,s-1} + S_{k,s}\delta_{k,s-2}, \end{aligned}$$

  where $R_{k,s} = \left(R_{k,s}^{ij}\right)_{i,j=1,\cdots,r}$ and $S_{k,s} = \left[S_{k,s}^1 \mid \cdots \mid S_{k,s}^r\right]$.

- the random parameter matrices $\{H_k\}_{k\geq 1}$ satisfy:
  - $E[H_k] = \overline{H}_k = \overline{\Theta}_k C_k$ where $\overline{\Theta}_k = E[\Theta_k] = Diag\left(\overline{\theta}_k^1, \ldots, \overline{\theta}_k^r\right)$.
  - Denoting $\theta_k = (\theta_k^1,\ldots,\theta_k^r)^T$, we have $Cov[\theta_k, \theta_s] = K_{k,k}^{\theta}\delta_{k,s} + K_{k,s}^{\theta}\delta_{k,s-1} + K_{k,s}^{\theta}\delta_{k,s+1}$, where $K_{k,s}^{\theta} = \left(K_{k,s}^{\theta^{ij}}\right)_{i,j=1,\cdots,r}$.
  - For any matrix $\mathcal{A} \in \mathbb{R}^{n\times n}$, we have

$$\begin{aligned} E[\widetilde{H}_k \mathcal{A} \widetilde{H}_s^T] &= E[(\Theta_k - \overline{\Theta}_k) C_k \mathcal{A} C_s^T (\Theta_s - \overline{\Theta}_s)] \\ &= K_{k,s}^{\theta} \circ \left(C_k \mathcal{A} C_s^T\right), \quad s = k,\ k-1. \end{aligned}$$

- since $B_k = I$, we have $\widetilde{B}_k = 0$, $\forall k$, and consequently all the expectations in which $\widetilde{B}_k$ or $\widetilde{B}_{k-1}$ appears, are zero.

Thus, the proposed optimal filtering algorithm in this case of multi-sensor systems with missing measurements is the following:

$$\begin{aligned}
&\widehat{x}_{k/k} = \widehat{x}_{k/k-1} + \mathcal{X}_{k,k}\Pi_k^{-1}\mu_k, \quad k \geq 1; \quad \widehat{x}_{0/0} = \overline{x}_0,\\
&\widehat{x}_{k/k-1} = \overline{F}_{k-1}\widehat{x}_{k-1/k-1} + \mathcal{W}_{k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k \geq 2; \qquad \widehat{x}_{1/0} = \overline{F}_0\widehat{x}_{0/0},\\
&\mathcal{X}_{k,k} = P_{k/k-1}C_k^T\overline{\Theta}_k + \mathcal{E}_{k,k} - \left(\overline{F}_{k-1}\mathcal{X}_{k-1,k-1} + \mathcal{W}_{k-1}\right)\Pi_{k-1}^{-1}\Psi_{k,k-1}^T, \quad k \geq 1,\\
&\mu_k = y_k - \overline{\Theta}_k C_k\widehat{x}_{k/k-1} - \Psi_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k \geq 1,\\
&\Pi_k = K_{k,k}^{\theta} \circ \left(C_k\mathcal{D}_k C_k^T\right) + R_{k,k} + \overline{\Theta}_k C_k\mathcal{X}_{k,k} + \mathcal{X}_{k,k}^T C_k^T\overline{\Theta}_k\\
&\qquad - \overline{\Theta}_k C_k P_{k/k-1}C_k^T\overline{\Theta}_k - \Psi_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T, \quad k \geq 1,
\end{aligned}$$

where $\mathcal{E}_{k,k}$, $\mathcal{D}_k$ and $P_{k/k-1}$ are given by (4.4), (4.6) and (4.11), respectively, and

$$\begin{aligned}
&\mathcal{W}_k = Q_{k,k-1}C_k^T\overline{\Theta}_k + S_{k,k}, \quad k \geq 1,\\
&\Psi_{k,k-1} = K_{k,k-1}^{\theta} \circ \left(C_k\left(\overline{F}_{k-1}\mathcal{D}_{k-1} + Q_{k-1,k-2}\right)C_{k-1}^T\right) + \mathcal{V}_{k,k-1}, \quad k \geq 2; \quad \Psi_{1,0} = 0,\\
&\mathcal{V}_{k,k-1} = S_{k-2,k}^T C_{k-1}^T\overline{\Theta}_{k-1} + R_{k,k-1}, \quad k \geq 2.
\end{aligned}$$

4.4.1 Numerical simulation example

Consider the following system with state-dependent multiplicative noise, and missing measurements from two sensors, with different missing characteristics and noise correlation:

$$\begin{aligned}
x_k &= (0.95 + 0.2\epsilon_{k-1})x_{k-1} + w_{k-1}, \quad k \geq 1,\\
y_k^i &= \theta_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2.
\end{aligned}$$

The initial state x_0 is a zero-mean Gaussian variable with $P_0 = 1$. The multiplicative state noise $\{\epsilon_k\}_{k\geq 0}$ is a zero-mean Gaussian white process with unit variance. The additive noise processes $\{w_k\}_{k\geq 0}$ and $\{v_k^i\}_{k\geq 1}$, $i = 1, 2$, are the same as those in [11], i.e., $w_k = 0.6(\eta_k + \eta_{k+1})$ and $v_k^i = c_k^i(\eta_{k-1} + \eta_k)$, $i = 1, 2$, where $c_k^1 = 1$, $c_k^2 = 0.5$, and $\{\eta_k\}_{k\geq 0}$ is a zero-mean Gaussian white process with variance 0.5.

Two different independent sequences of random variables with a probability distribution over the interval $[0, 1]$ are used to model the missing phenomenon:

- In the first sensor, the missing phenomenon is modelled by a sequence $\{\theta_k^1\}_{k\geq 1}$ of Bernoulli random variables correlated at consecutive sampling times; specifically, $\theta_k^1 = 1 - \beta_{k-1}(1 - \beta_k)$, where $\{\beta_k\}_{k\geq 0}$ is a sequence of independent Bernoulli random variables with $P[\beta_k = 1] = \beta$. Since the variables β_k and β_s are independent, it is clear that θ_k^1 and θ_s^1 are also independent for $|k - s| \geq 2$. Moreover, if $\theta_k^1 = 0$, then $\beta_{k-1} = 1$ and $\beta_k = 0$, and consequently $\theta_{k+1}^1 = 1$; hence, in the first sensor the state cannot be missing in two successive observations.

- In the second sensor, the missing phenomenon is modelled by a sequence $\{\theta_k^2\}_{k\geq 1}$ of independent and identically distributed random variables with the following probability distribution: $P[\theta_k^2 = 0] = 0.1$, $P[\theta_k^2 = 0.5] = 0.5$, $P[\theta_k^2 = 1] = 0.4$.

Under these assumptions, for all k, the mean $\overline{\Theta}_k$ and the covariances $K_{k,s}^{\theta}$, $s = k,\ k-1$, are given by

$$\overline{\Theta}_k = \begin{bmatrix} \overline{\theta}^1 & 0 \\ 0 & \overline{\theta}^2 \end{bmatrix} = \begin{bmatrix} 1-\beta(1-\beta) & 0 \\ 0 & 0.65 \end{bmatrix},$$

$$K_{k,k}^{\theta} = \begin{bmatrix} \overline{\theta}^1(1-\overline{\theta}^1) & 0 \\ 0 & 0.1025 \end{bmatrix} \quad \text{and} \quad K_{k,k-1}^{\theta} = \begin{bmatrix} -(1-\overline{\theta}^1)^2 & 0 \\ 0 & 0 \end{bmatrix}.$$

To analyze the effectiveness of the proposed estimator, one hundred iterations of the proposed filtering algorithm were performed and the filtering error variances were calculated for different values of the probability β, which provide different values of the probability $\overline{\theta}^1$ that the state is not missing from the observations of the first sensor. Since $\overline{\theta}^1$ is the same if the value $1 - \beta$ is considered instead of β, it is sufficient to consider $\beta \leq 0.5$ (note that, in this case, $\overline{\theta}^1$ is a decreasing

function of β). Specifically, the values $\beta = 0.1$, 0.2, 0.3, 0.4 and 0.5 (leading to $\overline{\theta}^1 = 0.91$, 0.84, 0.78, 0.76 and 0.75, respectively) are examined here.

Figure 4.2 shows that the filtering error variances become greater as β increases or, equivalently, as $\overline{\theta}^1$ decreases. This means that, as the probability of only noise measurements (false alarm probability) increases in the first sensor, worse estimations are obtained; note that for $\beta = 0.3$, 0.4, 0.5 the difference is smaller since the corresponding values of $\overline{\theta}^1$ are very close to each other.

Finally, we present a comparative analysis of four filters: the Kalman filter in systems with independent random parameter matrices and uncorrelated white noises [5]; the linear filter in systems with uncertain observations with correlated uncertainty and uncorrelated white noises [22]; the centralized filter in systems with missing measurements and correlated and cross-correlated noises [11]; and the filter proposed here. Using one thousand independent simulations of the mentioned algorithms, the different filtering estimates were compared using the mean square error (MSE) criterion. The filtering MSE at time k is calculated by $\text{MSE}_k = \dfrac{1}{1000}\displaystyle\sum_{s=1}^{1000}(x_k^{(s)} - \widehat{x}_{k/k}^{(s)})^2$, where $\{x_k^{(s)}\}_{1\leq k\leq 100}$ denote the s-th set of artificially simulated data and $\widehat{x}_{k/k}^{(s)}$ is the filter at the sampling time k in the s-th simulation run. These values are shown in Figure 4.3, where it can be seen that: 1) the performance of the filters with correlated uncertainty or with correlated and cross-correlated noises is better than that of the Kalman filter with independent random parameter matrices and uncorrelated white noises, since this filter ignores any correlation assumption; 2) the proposed filtering algorithm provides better estimations than other filtering algorithms reported previously.

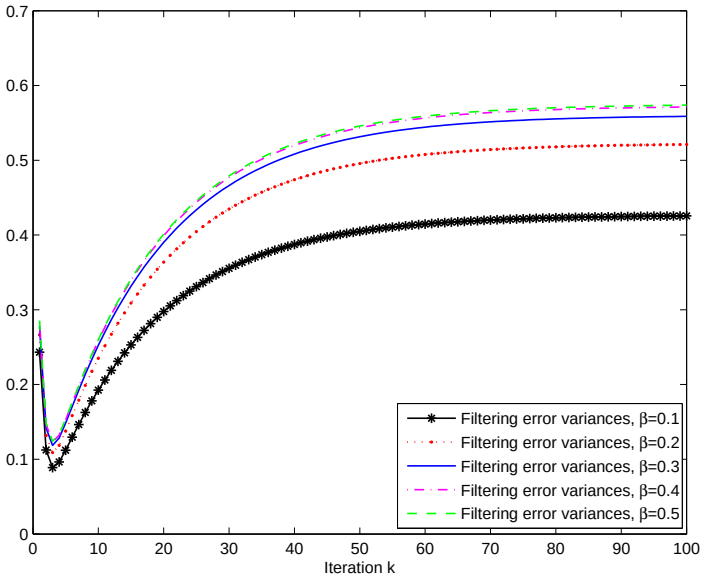


Figure 4.2: Filtering error variances for $\beta = 0.1$, 0.2, 0.3, 0.4 and 0.5.

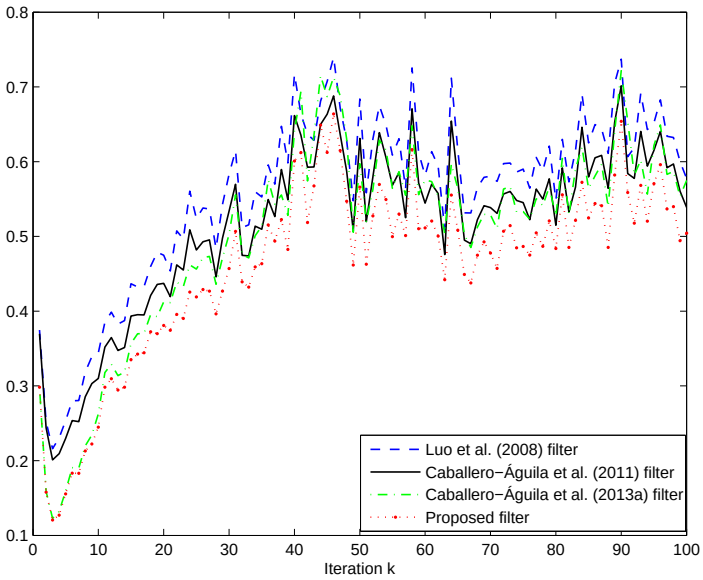


Figure 4.3: Comparison of MSE_k for different filters.

4.5 Application to multi-sensor systems with randomly delayed measurements

The estimation problem in multi-sensor systems with randomly delayed measurements is arousing increasing interest due to its broad scope of application. In networked systems, time delays are usually unavoidable, due to numerous causes including network congestion, random failures in the transmission mechanism or data inaccessibility at certain times. Since delays in measurement arrivals often occur randomly, the standard estimation algorithms are not applicable and several modifications have been proposed to incorporate the effects of randomly delayed measurements (see [24]-[26]).

Most papers on estimation in multi-sensor systems with randomly delayed observations assume that all the sensors have the same delay characteristics. Nevertheless, such an assumption is not realistic in many practical situations, where the information is gathered by an array of heterogeneous sensors, and the delay probability at each individual sensor can be different from the others. In recent years, this approach has been generalized in [27], [28] considering multiple delayed sensors with different delay characteristics and assuming that the delays are mutually independent. [29], recently weakened this assumption of independence by considering different sequences of Bernoulli variables correlated at consecutive sampling times to model the delays at each sensor. Similarly to the case of missing measurements, this correlation model avoids the possibility of two successive delayed observations, and so it can be applied to networked systems with stand-by sensors for the immediate replacement of a failed unit.

In this section we show that the current observation model with random measurement matrices (4.2) includes the observation model in multi-sensor systems with correlated random delays as a particular case; thus the current study generalizes the above results [27]-[29].

Assume that the state equation is given by (4.1), with $\{F_k\}_{k\geq 0}$ and $\{w_k\}_{k\geq 0}$

verifying hypotheses (b) and (c), and consider that z_k^i, $i = 1, \ldots, r$, are scalar sensor outputs perturbed by zero-mean additive noises v_k^i; namely,

$$z_k^i = c_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, \ldots, r. \tag{4.22}$$

For $i, j = 1, \ldots, r$, it is assumed that $Cov[v_k^i, v_s^j] = R_k^{ij}\delta_{k,s}$ and $Cov[w_k, v_s^i] = S_{k,k}^i\delta_{k,s} + S_{k,s}^i\delta_{k,s-1}$.

Consider that, at the initial time $k = 1$, the ith sensor outputs, z_1^i, are always available for the estimation but, at time $k \geq 2$, the ith sensor measurement, y_k^i, may be randomly delayed by one sampling time according to different delay characteristics, due to possible failures in data transmission. Therefore, the measurement model is described by

$$y_k^i = (1 - \gamma_k^i)z_k^i + \gamma_k^i z_{k-1}^i, \quad k \geq 2; \quad y_1^i = z_1^i, \quad i = 1, \ldots, r, \tag{4.23}$$

where $\{\gamma_k^i\}_{k\geq 2}$, $i = 1, \ldots, r$, denote sequences of Bernoulli variables with $P[\gamma_k^i = 1] = p_k^i$ and $Cov[\gamma_k^i, \gamma_s^j] = K_{k,k}^{\gamma^{ij}}\delta_{k,s} + K_{k,s}^{\gamma^{ij}}\delta_{k,s-1} + K_{k,s}^{\gamma^{ij}}\delta_{k,s+1}$.

By denoting $z_k = (z_k^1, \ldots, z_k^r)^T$, $C_k = \left[\begin{array}{c|c|c} c_k^{1T} & \cdots & c_k^{rT} \end{array}\right]^T$, $v_k = (v_k^1, \ldots, v_k^r)^T$, and $\Gamma_k = Diag\,(\gamma_k^1, \ldots, \gamma_k^r)$, (4.22) and (4.23) can be rewritten as:

$$\begin{aligned} z_k &= C_k x_k + v_k, \quad k \geq 1, \\ y_k &= (I - \Gamma_k)z_k + \Gamma_k z_{k-1}, \quad k \geq 2; \quad y_1 = z_1, \end{aligned} \tag{4.24}$$

and, from the correlation assumptions of the noises, it is clear that $Cov[v_k, v_s] = R_k\delta_{k,s}$ and $Cov[w_k, v_s] = S_{k,k}\delta_{k,s} + S_{k,s}\delta_{k,s-1}$, where $R_k = \left(R_k^{ij}\right)_{i,j=1,\cdots,r}$ and $S_{k,s} = \left[S_{k,s}^1 \mid \cdots \mid S_{k,s}^r\right]$. Moreover, by denoting $\gamma_k = (\gamma_k^1, \ldots, \gamma_k^r)^T$, we have $Cov[\gamma_k, \gamma_s] = K_{k,k}^{\gamma}\delta_{k,s} + K_{k,s}^{\gamma}\delta_{k,s-1} + K_{k,s}^{\gamma}\delta_{k,s+1}$, where $K_{k,s}^{\gamma} = \left(K_{k,s}^{\gamma^{ij}}\right)_{i,j=1,\cdots,r}$.

Now, as in [27], equations (4.1) and (4.24) are rewritten as follows, with random parameter matrices:

$$\begin{aligned} X_{k+1} &= \mathcal{F}_k X_k + W_k, \quad k \geq 1, \\ y_k &= H_k X_k + B_k V_k, \quad k \geq 2, \end{aligned} \tag{4.25}$$

where

$$X_k = \left[\frac{x_k}{x_{k-1}}\right], \quad W_k = \left[\frac{w_k}{0}\right], \quad V_k = \left[\frac{v_k}{v_{k-1}}\right], \quad \mathcal{F}_k = \left[\begin{array}{c|c} F_k & 0 \\ \hline I & 0 \end{array}\right],$$

$$H_k = \left[\begin{array}{c|c} (I-\Gamma_k)C_k & \Gamma_k C_{k-1} \end{array}\right], \quad B_k = \left[\begin{array}{c|c} I-\Gamma_k & \Gamma_k \end{array}\right].$$

It is clear that the random parameter matrices and noise processes of system (4.25) verify the hypotheses to apply the algorithm proposed in this paper. Specifically, we have:

- $E[\mathcal{F}_k] = \overline{\mathcal{F}}_k = \left[\begin{array}{c|c} \overline{F}_k & 0 \\ \hline I & 0 \end{array}\right]$, $\overline{H}_k = \left[\begin{array}{c|c} (I-\overline{\Gamma}_k)C_k & \overline{\Gamma}_k C_{k-1} \end{array}\right]$ and $\overline{B}_k = \left[\begin{array}{c|c} I-\overline{\Gamma}_k & \overline{\Gamma}_k \end{array}\right]$, where $\overline{\Gamma}_k = Diag\,(p_k^1, \ldots, p_k^r)$.

- The process noise, $\{W_k\}_{k\geq 1}$, and the measurement noise, $\{V_k\}_{k\geq 2}$, are zero-mean sequences with covariances and cross-covariances:

$$\begin{array}{l} Cov[W_k, W_s] = \mathbb{Q}_{k,k}\delta_{k,s} + \mathbb{Q}_{k,s}\delta_{k,s-1} + \mathbb{Q}_{k,s}\delta_{k,s+1}, \\ Cov[V_k, V_s] = \mathbb{R}_{k,k}\delta_{k,s} + \mathbb{R}_{k,s}\delta_{k,s-1} + \mathbb{R}_{k,s}\delta_{k,s+1}, \\ Cov[W_k, V_s] = \mathbb{S}_{k,k}\delta_{k,s} + \mathbb{S}_{k,s}\delta_{k,s-1} + \mathbb{S}_{k,s}\delta_{k,s-2}, \end{array}$$

  where

$$\mathbb{Q}_{k,k} = \left[\begin{array}{c|c} Q_{k,k} & 0 \\ \hline 0 & 0 \end{array}\right], \quad \mathbb{Q}_{k,k-1} = \left[\begin{array}{c|c} Q_{k,k-1} & 0 \\ \hline 0 & 0 \end{array}\right],$$

$$\mathbb{R}_{k,k} = \left[\begin{array}{c|c} R_k & 0 \\ \hline 0 & R_{k-1} \end{array}\right], \quad \mathbb{R}_{k,k-1} = \left[\begin{array}{c|c} 0 & 0 \\ \hline R_{k-1} & 0 \end{array}\right],$$

$$\mathbb{S}_{k,k} = \left[\begin{array}{c|c} S_{k,k} & 0 \\ \hline 0 & 0 \end{array}\right], \ \mathbb{S}_{k-1,k} = \left[\begin{array}{c|c} S_{k-1,k} & S_{k-1,k-1} \\ \hline 0 & 0 \end{array}\right], \ \mathbb{S}_{k-2,k} = \left[\begin{array}{c|c} 0 & S_{k-2,k-1} \\ \hline 0 & 0 \end{array}\right].$$

Then, it is clear that

- $\mathbb{D}_k = E[X_k X_k^T] = \left[\begin{array}{c|c} \mathcal{D}_k & \mathcal{G}_{k,k-1} \\ \hline \mathcal{G}_{k,k-1}^T & \mathcal{D}_{k-1} \end{array}\right]$, $E\left[\widetilde{\mathcal{F}}_k \mathbb{D}_k \widetilde{\mathcal{F}}_k^T\right] = \left[\begin{array}{c|c} E[\widetilde{F}_k \mathcal{D}_k \widetilde{F}_k^T] & 0 \\ \hline 0 & 0 \end{array}\right]$ and $\mathbb{G}_{k+1,k} = E[X_{k+1} X_k^T] = \left[\begin{array}{c|c} \mathcal{G}_{k+1,k}^T & \overline{F}_k \mathcal{G}_{k,k-1} \\ \hline \mathcal{D}_k & \mathcal{G}_{k,k-1} \end{array}\right]$, where $\mathcal{D}_k$ and $\mathcal{G}_{k,k-1}$ are given by (4.6) and (4.7), respectively.

– Analogously to (4.3) and (4.4), we have

$$\begin{aligned}
&\mathbb{W}_k = E[W_k y_k^T] = \mathbb{Q}_{k,k-1}\overline{H}_k^T + \mathbb{S}_{k,k}\overline{B}_k^T, \quad k \geq 1,\\
&\mathbb{E}_{k,k} = E[X_k v_k^T] = \overline{\mathcal{F}}_{k-1}\mathbb{S}_{k-2,k} + \mathbb{S}_{k-1,k}, \quad k \geq 2; \qquad \mathbb{E}_{1,1} = \mathbb{S}_{0,1},\\
&\mathbb{E}_{k,k-1} = E[X_k v_{k-1}^T] = \overline{\mathcal{F}}_{k-1}\mathbb{E}_{k-1,k-1} + \mathbb{S}_{k-1,k-1}, \quad k \geq 2.
\end{aligned}$$

– For arbitrary matrices $\mathcal{A}_1 \in \mathbb{R}^{2n\times 2n}$, $\mathcal{A}_2 \in \mathbb{R}^{2r\times 2r}$ and $\mathcal{A}_3 \in \mathbb{R}^{2r\times 2n}$, we have

$$\begin{aligned}
E[\widetilde{H}_k\mathcal{A}_1\widetilde{H}_s^T] &= K_{k,s}^{\gamma} \circ \left([-C_k \mid C_{k-1}]\,\mathcal{A}_1\,[-C_s \mid C_{s-1}]^T\right), \quad s = k,\ k-1,\\
E[\widetilde{B}_k\mathcal{A}_2\widetilde{B}_s^T] &= K_{k,s}^{\gamma} \circ \left([-I \mid I\ ]\,\mathcal{A}_2\,[-I \mid I\ ]^T\right), \quad s = k,\ k-1,\\
E[\widetilde{B}_k\mathcal{A}_3\widetilde{H}_s^T] &= K_{k,s}^{\gamma} \circ \left([-I \mid I\ ]\,\mathcal{A}_3\,[-C_s \mid C_{s-1}]^T\right), \quad s = k,\ k-1.
\end{aligned}$$

By applying the above expressions, we obtain $E[\widetilde{H}_k\mathbb{D}_k\widetilde{H}_k^T]$, $E[\widetilde{B}_k\mathbb{R}_{k,k}\widetilde{B}_k^T]$ and $E[\widetilde{B}_k\mathbb{E}_{k,k}^T\widetilde{H}_k^T]$, which are necessary to calculate the innovation covariance matrices. We also have

$$\begin{aligned}
\mathcal{V}_{k,k-1} &= \overline{\Gamma}_k\left(C_{k-1}S_{k-2,k-1} + R_{k-1}\right)^T\left(I - \overline{\Gamma}_{k-1}\right) - K_{k,k-1}^{\gamma} \circ \left(C_{k-1}S_{k-2,k-1} + R_{k-1}\right)^T,\\
\Psi_{k,k-1} &= K_{k,k-1}^{\gamma} \circ \left([-C_k \mid C_{k-1}]\left(\mathbb{G}_{k,k-1}\,[-C_{k-1} \mid C_{k-2}]^T + \mathbb{E}_{k,k-1}\,[-I \mid I\ ]^T\right)\right)\\
&\quad + \mathcal{V}_{k,k-1}.
\end{aligned}$$

Hence, the proposed optimal filtering algorithm for multi-sensor systems with randomly delayed measurements is:

$$\begin{aligned}
&\widehat{X}_{k/k} = \widehat{X}_{k/k-1} + \mathcal{X}_{k,k}\Pi_k^{-1}\mu_k, \quad k \geq 1,\\
&\widehat{X}_{k/k-1} = \overline{\mathcal{F}}_{k-1}\widehat{X}_{k-1/k-1} + \mathbb{W}_{k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k \geq 2; \qquad \widehat{X}_{1/0} = \left[\frac{\overline{F}_0\overline{x}_0}{\overline{x}_0}\right],\\
&\mathcal{X}_{k,k} = P_{k/k-1}\overline{H}_k^T + \mathbb{E}_{k,k}\overline{B}_k^T - \left(\overline{\mathcal{F}}_{k-1}\mathcal{X}_{k-1,k-1} + \mathbb{W}_{k-1}\right)\Pi_{k-1}^{-1}\Psi_{k,k-1}^T, \quad k \geq 1,\\
&\mu_k = y_k - \overline{H}_k\widehat{X}_{k/k-1} - \Psi_{k,k-1}\Pi_{k-1}^{-1}\mu_{k-1}, \quad k \geq 1,\\
&\Pi_k = E[\widetilde{H}_k\mathbb{D}_k\widetilde{H}_k^T] + E[\widetilde{B}_k\mathbb{R}_{k,k}\widetilde{B}_k^T] + E[\widetilde{B}_k\mathbb{E}_{k,k}^T\widetilde{H}_k^T] + E[\widetilde{H}_k\mathbb{E}_{k,k}\widetilde{B}_k^T] + \overline{B}_k\mathbb{R}_{k,k}\overline{B}_k^T\\
&\qquad + \overline{H}_k\mathcal{X}_{k,k} + \mathcal{X}_{k,k}^T\overline{H}_k^T - \overline{H}_kP_{k/k-1}\overline{H}_k^T - \Psi_{k,k-1}\Pi_{k-1}^{-1}\Psi_{k,k-1}^T,
\end{aligned}$$

$$P_{k/k} = P_{k/k-1} - \mathcal{X}_{k,k}\Pi_k^{-1}\mathcal{X}_{k,k}^T, \quad k \geq 1,$$

$$P_{k/k-1} = \overline{\mathcal{F}}_{k-1}P_{k-1/k-1}\overline{\mathcal{F}}_{k-1}^T + E[\widetilde{\mathcal{F}}_{k-1}\mathbb{D}_{k-1}\widetilde{\mathcal{F}}_{k-1}^T] + \mathbb{Q}_{k-1,k-1} + \overline{\mathcal{F}}_{k-1}\mathbb{J}_{k-1} + \mathbb{J}_{k-1}^T\overline{\mathcal{F}}_{k-1}^T - \mathbb{W}_{k-1}\Pi_{k-1}^{-1}\mathbb{W}_{k-1}^T, \quad k \geq 2;$$

$$P_{1/0} = \left[\begin{array}{c|c} \overline{F}_0 P_0 \overline{F}_0^T + E[\widetilde{F}_0\mathcal{D}_0\widetilde{F}_0^T] + Q_{0,0} & \overline{F}_0 P_0 \\ \hline P_0\overline{F}_0^T & P_0 \end{array}\right],$$

$$\mathbb{J}_k = \mathbb{Q}_{k-1,k} - \mathcal{X}_{k,k}\Pi_k^{-1}\mathbb{W}_k^T, \quad k \geq 1.$$

4.5.1 Numerical simulation example

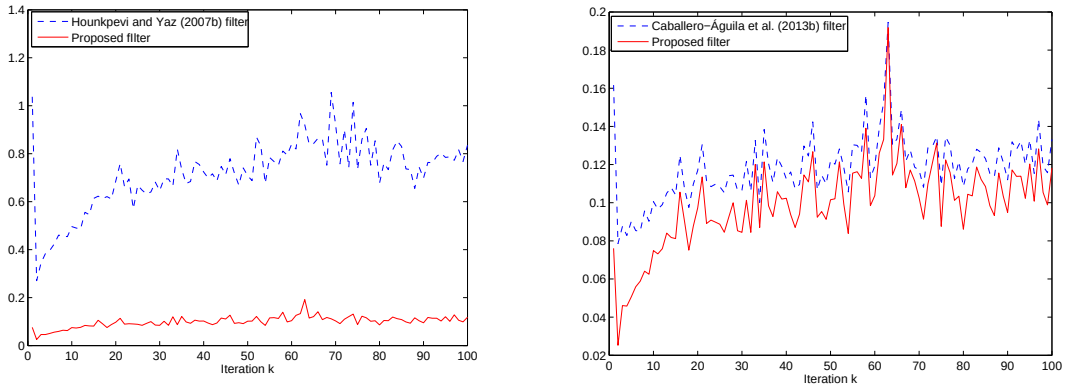
In this example, it is assumed that the state $\{x_k\}_{k\geq 0}$ is generated by the same model as that in Section 4.4.1, and we consider measured outputs coming from two sensors, $z_k^i = x_k + v_k^i,\ k \geq 1,\ i = 1, 2$, where the additive noises are defined by $v_k^1 = \eta_k$ and $v_k^2 = 0.5\eta_k$.

According to the proposed observation model, it is assumed that, at any sampling time $k \geq 2$, the measured output from the ith sensor, z_k^i, can be randomly delayed by one sampling period during network transmission; thus, the measurement model is described by

$$y_k^i = (1 - \gamma_k^i)z_k^i + \gamma_k^i z_{k-1}^i, \quad k \geq 2; \quad y_1^i = z_1^i, \quad i = 1, 2.$$

As in [29], it is assumed that the delays are correlated at consecutive sampling times, which guarantees that two successive observations cannot be delayed; specifically, the variables γ_k^i are defined by $\gamma_k^i = \alpha_{k+1}^i(1 - \alpha_k^i)$, where $\{\alpha_k^i\}_{k\geq 1}$, $i = 1, 2$, are two independent sequences of independent Bernoulli variables with probabilities $P[\alpha_k^1 = 1] = 0.5$ and $P[\alpha_k^2 = 1] = 0.1$, respectively.

To illustrate the accuracy of the proposed algorithm in comparison with other estimation methods that have been proposed, one thousand independent simulations were considered and one hundred iterations of each algorithm performed to compute the filtering MSE at each time instant k. A comparative analysis was carried out between the suboptimal Kalman-type filter for systems with independent random delays [27], the optimal linear filter using covariance information for



(a) (b)

Figure 4.4: (a) Proposed filter vs. Hounkpevi and Yaz (2007b) filter.
(b) Proposed filter vs. Caballero-Águila et al.(2013b) filter.

systems with one-step correlated random delays [29], and the current filter for multi-sensor systems with randomly delayed measurements. The results of this comparison are shown in Figure 4.4, where it can be seen that the proposed filter performs better than the other two. The difference with respect to [27] is greater since the correlation assumption on the delays and the noises is not taken into account and moreover the estimator in [27] is suboptimal.

4.6 Conclusions

This paper reports a study of the optimal LS linear filtering problem for discrete-time linear systems with random parameter matrices and correlated additive noise. The main contributions of this approach are:

1. The current system model includes independent random state transition matrices and one-step correlated and cross-correlated random parameter matrices in the observation equation. The process and measurement noises are assumed to be one-step autocorrelated and two-step cross-correlated.

2. An optimal LS linear recursive filtering algorithm with a simple computational procedure is derived by an innovation approach.

3. The proposed optimal LS linear filtering algorithm was applied to systems with multiple missing measurements with correlated and cross-correlated noises, when the missing measurement phenomenon in each sensor is described by different sequences of scalar random variables with arbitrary discrete probability distribution over the interval [0,1] correlated at consecutive sampling times. This kind of multi-sensor system is found in various real-world problems, such as transmission models with stand-by sensors or situations involving the partial loss of measurements.

4. Multi-sensor systems with randomly delayed measurements, correlated at consecutive sampling times, with correlated and cross-correlated noises are also treated as a particular case of the model described in this paper. These models cover the situations in which two successive observations cannot be delayed. This kind of delay frequently occurs, in situations such as network congestion, random failures in the transmission mechanism or data inaccessibility at certain times.

5. For both particular cases, the feasibility of the proposed filtering algorithm is analyzed by two numerical simulation examples, which show that the proposed filter performs better than others that have been reported.

6. A similar study to that performed in this paper would allow us to generalize the current results by considering correlation between random state transition matrices and the random matrices in the observation equation. This extension would cover systems with multiple packet dropouts as a particular case, and would constitute an interesting research topic.

7. Another interesting future direction would be to complement the current

study with a detailed analysis of the convergence and computational complexity of the proposed filtering algorithm.

8. The filtering methodology proposed in this paper can be applied to other, related problems, such as fault detection or control systems, which constitute interesting and challenging topics for future research ([20], [30], [31]).

Acknowledgements

This research is supported by Ministerio de Ciencia e Innovación (FPU programme and Grant No. MTM2011-24718).

References

[1] M. Moayedi, Y.K. Foo and Y.C. Soh, Adaptive Kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements, *IEEE Trans. Signal Process.,* 58(3) (2010), 1577–1588.

[2] J. Ma and S.L. Sun, Optimal linear estimators for systems with random sensor delays, multiple packet dropouts and uncertain observations, *IEEE Trans. Signal Process.,* 59(11) (2011), 5181–5192.

[3] S.L. Sun and W. Xiao, Optimal linear estimators for systems with multiple random measurement delays and packet dropouts, *International Journal of Systems Science,* 44(2) (2013), 358–370.

[4] M. Sahebsara, T. Chen and S.L. Shah, Optimal H_2 filtering with random sensor delay, multiple packet dropout and uncertain observations, *Int. J. Control* , 80 (2007), 292–301.

[5] Y. Luo, Y. Zhu, D. Luo, J. Zhou, E. Song and D. Wang, Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications, *Sensors,* 8(12) (2008), 8086–8103.

[6] J. Feng, Z. Wang and M. Zeng, Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises, *Information Fusion,* 14 (2013), 78–86.

[7] W.L. De Koning, Optimal estimation of linear discrete-time systems with stochastic parameters, *Automatica,* 20 (1984), 113–115.

[8] Y.T. Luo, Y.M. Zhu, X.J. Shen and E.B. Song, Novel Data Association Algorithm Based on Integrated Random Coefficient Matrices Kalman Filtering, *IEEE Transactions on Aerospace and Electronic Systems,* 48(1) (2012), 144–158.

[9] X.J. Shen, Y.T. Luo, Y.M. Zhu and E.B. Song, Globally optimal distributed Kalman filtering fusion, *Science China. Information Sciences,* 55(3) (2012), 512–529.

[10] X.J. Shen, Y.M. Zhu and Y.T. Luo, Optimal State Estimation of Linear Discrete-time Systems with Correlated Random Parameter Matrices, in *Proceedings of the 30th Chinese Control Conference,* (2011), (pp. 1488–1493). Yantai, China.

[11] R. Caballero-Águila, I. García-Garrido and J. Linares-Pérez, Optimal Fusion Filtering in Multisensor Stochastic Systems with Missing Measurements and Correlated Noises, *Mathematical Problems in Engineering,* Volume 2013, Article ID 418678, 14 pages.

[12] E. Song, Y. Zhu and Z. You, The Kalman type recursive state estimator with a finite-step correlated process noises, in *Proceedings of the IEEE International*

Conference on Automation and Logistics, (2008), (pp. 196–200). Qingdao, China.

[13] A. Fu, Y. Zhu and E. Song, The optimal Kalman type state estimator with multi-step correlated process and measurement noises, in *The 2008 International Conference on Embedded Software and Systems,* (2008), (pp. 215–220). Chengdu, China.

[14] J. Feng, Z. Wang and M. Zeng, Optimal robust non-fragile Kalman-type recursive filtering with finite-step autocorrelated noises and multiple packet dropouts, *Aerospace Science and Technology,* 15(6) (2011), 486–494.

[15] F. Carravetta, A. Germani and M. Raimondi, Polynomial Filtering of Discrete-Time Stochastic Linear Systems with Multiplicative State Noise, *IEEE Transactions on Automatic Control,* 42(8) (1997), 1106–1126.

[16] Z. Wang, F. Yang, D.W.C. Ho, and X. Liu, Robust H_∞ filtering for stochastic time-delay systems with missing measurements, *IEEE Transaction on Signal Processing,* 54(7) (2006), 2579–2587.

[17] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo and J. Linares- Pérez, New recursive estimators from correlated interrupted observations using covariance information, *International Journal of Systems Science,* 36(10) (2005), 617–629.

[18] F.O. Hounkpevi and E.E. Yaz, Robust minimum variance linear state estimators for multiple sensors with different failure rates, *Automatica,* 43 (2007), 1274–1280.

[19] X. Qu and J. Zhou, The optimal robust finite-horizon Kalman filtering for multiple sensors with different stochastic failure rates, *Applied Mathematics Letters,* 26 (2013), 80–86.

[20] Z. Wang, B. Shen, H. Shu and G. Wei, Quantized H_∞ control for nonlinear stochastic time-delay systems with missing measurements, *IEEE Transactions on Automatic Control,* 57(6) (2012), 1431–1444.

[21] J. Hu, Z. Wang, B. Shen and H. Gao, Quantized recursive filtering for a class of nonlinear systems with multiplicative noises and missing measurements, *International Journal of Control,* 86(4) (2013), 650–663.

[22] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Linear and quadratic estimation using uncertain observations from multiple sensors with correlated uncertainty, *Signal Processing,* 91 (2) (2011), 330–337.

[23] J. Hu, Z. Wang, H. Gao and L.K. Stergiouslas, Extended Kalman filtering with stochastic nonlinearities and multiple missing measurements, *Automatica,* 48 (2012), 2007–2015.

[24] A.S. Matveev and A.V. Savkin, The problem of state estimation via asynchronous communication channels with irregular transmission times, *IEEE Transactions on Automatic Control,* 48(4) (2003), 670–676.

[25] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan and S.S. Sastry, Kalman filtering with intermittent observations, *IEEE Transactions on Automatic Control,* 49 (9) (2004), 1453–1464.

[26] Z. Wang, D.W.C. Ho and X. Liu, Robust filtering under randomly varying sensor delay with variance constraints, *IEEE Transactions on Circuits and Systems, II Express Briefs,* 51(6) (2004), 320–326.

[27] F.O. Hounkpevi and E.E. Yaz, Minimum variance generalized state estimators for multiple sensors with different delay rates, *Signal Processing,* 87 (2007b), 602–613.

[28] R. Caballero-Águila, A. Hermoso-Carazo, J.D. Jiménez-López, J. Linares-Pérez and S. Nakamori, Signal estimation with multiple delayed sensors using covariance information, *Digital Signal Processing,* 20 (2010), 528–540.

[29] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Linear estimation based on covariances for networked systems featuring sensor correlated random delays, *International Journal of Systems Science,* 44(7) (2013b), 1233–1244.

[30] H. Dong, Z. Wang and H. Gao, Fault detection for Markovian jump systems with sensor saturations and randomly varying nonlinearities, *IEEE Transactions on Circuits and Systems - Part I,* 59(10) (2012a), 2354–2362.

[31] H. Dong, Z. Wang, J. Lam and H. Gao, Fuzzy-model-based robust fault detection with stochastic mixed time-delays and successive packet dropouts, *IEEE Transactions on Systems, Man, and Cybernetics - Part B,* 42(2) (2012b), 365–376.

Chapter 5

Distributed matrix-weighted fusion filter for sensor networks with random parameter matrices and noise correlation

Reference

Linares-Pérez, J., Caballero-Águila, R. and García-Garrido, I. (2015).
Distributed matrix-weighted fusion filter for sensor networks with random parameter matrices and noise correlation.
IEEE Sensors Journal, (submitted).

Abstract

In this paper, the distributed fusion state estimation problem is addressed for sensor network systems with random state transition matrix and random measurement matrices, which provide a unified framework to consider some network-induced random phenomena. The process noise and all the sensor measurement noises are assumed to be one-step autocorrelated and different sensor noises are one-step cross-correlated; also, the process noise and each sensor measurement noise

are two-step cross-correlated. These correlation assumptions cover many practical situations, were the classical independence hypothesis is not realistic. Using an innovation methodology, local least-squares linear filtering estimators are recursively obtained at each sensor. The distributed fusion method is then used to form the optimal matrix-weighted sum of these local filters according to the mean squared error criterion. A numerical simulation example shows the accuracy of the proposed distributed fusion filtering algorithm and illustrates some of the network-induced stochastic uncertainties that can be dealt with the current system model, such as sensor gain degradation, missing measurements and multiplicative noise.

Contents

5.1 Introduction

In recent years, information and communication technologies have experienced a fast development, making the use of sensor networks become very popular for mea-

surement acquisition and data processing, as they usually provide more information than traditional single-sensor communication systems. For this reason, the study of the estimation problem in sensor network stochastic systems has achieved great interest in many important research fields of engineering, computing and mathematics, mainly, by their broad scope of applications (target tracking, environment observation, habitat monitoring, animal tracking, communications, etc.).

Although fusion algorithms have been proposed according to different methods (see e.g. [1]-[3]), most existing results do not consider the new problems that arise inevitably in sensor network systems due to the restrictions of the physical equipment, mainly the limitations of bandwidth channels and uncertainties in the external environment, in both the modeling process and the transmission of information. These situations can dramatically worsen the quality of fusion estimators designed. Multiplicative noise uncertainties, random delays, packet dropouts and missing measurements, are some of the most common problems that motivated the need to develop new estimation algorithms. Therefore, it is not surprising that, in the past few years, the study of the state estimation problem in network systems with only one or several of the aforementioned uncertainties has become an active research area (see e.g. [4]-[10] and references therein).

Clearly, some of these situations with networked-induced phenomena are special cases of systems with transition and/or measurement random parameter matrices, which have important practical significance and arise in many application areas such as digital control of chemical processes, radar control, navigation systems or economic systems [11]. On the one hand, random state transition matrices arise in the context of systems with state-dependent multiplicative noise, of great interest for applications in aerospace systems, communication, processing images, etc. [12]. On the other hand, networked systems with random observation losses [13] or systems with stochastic observation multiplicative noises [14], clearly are special cases of systems with random parameter measurement matrices. Also,

networked systems with stochastic sensor gain degradation as those considered in [15] or the systems with state and measurement multiplicative noises and missing measurements in [4], can be rewritten by transition and measurement random parameter matrices. It must be mentioned that in many papers, see e.g. [8], systems with random delays and packet dropouts are transformed into systems with random parameter matrices. Consequently, these kind of systems can model a great variety of real situations and, for this reason, the estimation problem in this type of systems has gained a considerable interest in recent years (see e.g. [16]-[20] and references therein).

Furthermore, in the latest research on signal estimation, the fairly conservative assumption that the process and measurement noises are uncorrelated is commonly weakened, since in many practical situations, such noises are usually correlated. For example, when all the sensors operate in the same noisy environment, the sensor noises are usually correlated. Likewise when the process noise and the sensor noises are state dependent, there may be cross-correlation between them, as well as between different sensor noises. Also, the augmented systems used to describe random delays and measurement losses are systems with correlated noises, and discretized continuous-time systems have also inherently correlated noises. Hence, in both, systems with deterministic matrices and systems with random parameter matrices, the estimation problem with correlated and cross-correlated noises, has become a challenging research topic. In the first case, the optimal Kalman filtering fusion problem in systems with cross-correlated process noises and measurement noises at the same sampling time is addressed in [21] and [22], and at consecutive sampling times in [3]. Under different correlation assumptions of the noises, centralized and distributed fusion algorithms are obtained in [12], for systems with multiplicative noise in the state equation, and in [6], for systems where the measurements have partial information about the signal. For systems with random parameter matrices and autocorrelated and cross-correlated noises, the research

efforts have been devoted to the centralized fusion estimation problem ([17]-[20]). Centralized algorithms are based on a fusion centre able to receive all the measured data from sensors for being processed; they provide optimal estimators from the measurements of all the sensors and hence, when all the sensors work correctly, they have the best accuracy. Nevertheless, as it is known, the centralized approach has several drawbacks such as bad robustness, poor survivability, reliability, heavy communication and expensive computational cost, which can be overcome by using distributed approaches. In the distributed fusion method, each sensor estimates the state based on its own measurement data, and these local estimators are combined according to a certain information fusion criterion. To the best of the authors knowledge, the distributed fusion estimation problem in networked systems with both random parameter matrices and autocorrelated and cross-correlated noises has not been investigated.

Motivated by the above considerations, this paper deals with the distributed fusion estimation problem in sensor network systems including simultaneously random parameter matrices and correlated noises in the state-space model. The main contributions of our study can be highlighted as follows: (1) The network system model with random parameter matrices considered provides a unified framework to treat some network-induced phenomena, such as multiplicative noise uncertainties, missing measurements or sensor gain degradation, and, hence, the proposed distributed fusion filter has a wide applicability. (2) One-step autocorrelation of the noises and, also, two-step cross-correlation between the process noise and different sensor noises are considered. (3) The innovation technique is used to obtain algorithms for the local least-squares linear filtering estimators which are recursive and computationally simple. (4) The proposed distributed fusion filter is generated by a matrix-weighted linear combination of the local filtering estimators using the mean squared error as optimality criterion. It requires the cross-covariance matrices between any two local filters (and not the error cross-covariance matrices, as

in [1])

The rest of the paper is organized as follows. The system model with multiple sensors and random parameter matrices is presented in Section 5.2, including a brief description of the traditional centralized and distributed fusion estimation methods. The local least-squares linear filtering algorithms are derived in Section 5.3, using an innovation approach. In Section 5.4, the proposed distributed fusion filter is obtained by a matrix-weighted linear combination of the local filtering estimators using the mean squared error as optimality criterion. A simulation example is given in Section 5.5 to show the performance of the proposed estimation algorithms, and some conclusions are drawn in Section 5.6.

Notation: The notation used throughout the paper is standard. $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. A^T and A^{-1} denote the transpose and inverse of a matrix A, respectively. The shorthand $(A_1, \ldots, A_m)$ denotes a partitioned matrix into sub-matrices $A_1, \ldots, A_m$. If a matrix dimension is not explicitly stated, it is assumed to be compatible for algebraic operations. Moreover, for any function $G_{k,s}$, depending on the time instants k and s, we write $G_k = G_{k,k}$ for simplicity. Analogously, we write $K^{(i)} = K^{(ii)}$ for any function $K^{(ij)}$, depending on the sensors i and j. Finally, all the random vectors will be defined on the probabilistic space $(\Omega, \mathcal{A}, P)$, and for arbitrary random vectors X and Y, we denote $Cov[X, Y] = E[(X - E[X])\,(Y - E[Y])^T]$ and $Cov[X] = Cov[X, X]$, where $E[\cdot]$ stands for the mathematical expectation operator.

5.2 System formulation and problem statement

Consider a class of discrete-time linear stochastic systems whose n_x-dimensional state process, $\{x_k;\ k \geq 0\}$, is perturbed by an additive process noise, $\{w_k;\ k \geq 0\}$, and the state transition matrices, $\{F_k;\ k \geq 0\}$, are $n_x \times n_x$ random parameter matrices; specifically, the state evolution is given by

$$x_{k+1} = F_k x_k + w_k, \quad k \geq 0. \tag{5.1}$$

Assume that the state process is observed by m different sensors, and that the measurements provided by the i-th sensor are described as follows:

$$y_k^{(i)} = H_k^{(i)} x_k + v_k^{(i)}, \quad k \geq 1, \quad i = 1, 2, \ldots, m, \tag{5.2}$$

where $y_k^{(i)} \in \mathbb{R}^{n_y}$ is the output measured by sensor i at time k. For $i = 1, 2, \ldots, m$, $\{H_k^{(i)};\ k \geq 1\}$ is a sequence of random parameter matrices and $\{v_k^{(i)};\ k \geq 1\}$ is the measurement noise of the i-th sensor.

Model assumptions. The assumptions about the initial state and the noise processes involved in the system model (5.1)-(5.2), under which the fusion filtering problem will be addressed, are:

(i) The initial state x_0 is a random vector with $E[x_0] = \overline{x}_0$ and $Cov[x_0] = \Sigma_0$.

(ii) $\{F_k;\ k \geq 0\}$ and $\{H_k^{(i)};\ k \geq 1\}$ are sequences of independent random parameter matrices with known means, $E[F_k] = \overline{F}_k$, $E[H_k^{(i)}] = \overline{H}_k^{(i)}$, $i = 1, 2, \ldots, m$, and the covariances of their entries, $Cov[f_{pq}(k), f_{p'q'}(k)]$, $Cov[h_{pa}^{(i)}(k), h_{qb}^{(i)}(k)]$, are also assumed to be known. $f_{pq}(k)$ denotes the (p,q)-th entry of matrix F_k, for $p, q = 1, 2, \ldots, n_x$, and $h_{pq}^{(i)}(k)$ denotes the (p,q)-th entry of $H_k^{(i)}$, for $p = 1, 2, \ldots, n_x$ and $q = 1, 2, \ldots, n_y$.

(iii) The noises $\{w_k;\ k \geq 0\}$ and $\{v_k^{(i)};\ k \geq 1\}$, $i = 1, 2, \ldots, m$, are zero-mean sequences with known covariances and cross-covariances:

$$\begin{aligned} &Cov[w_k, w_s] = Q_{k,s}\left(\delta_{k-s} + \delta_{k-s+1} + \delta_{k-s-1}\right), \\ &Cov[v_k^{(i)}, v_s^{(j)}] = R_{k,s}^{(ij)}\left(\delta_{k-s} + \delta_{k-s+1} + \delta_{k-s-1}\right), \\ &Cov[w_k, v_s^{(i)}] = S_{k,s}^{(i)}\left(\delta_{k-s} + \delta_{k-s+1} + \delta_{k-s+2}\right). \end{aligned}$$

(iv) For $i = 1, 2, \ldots, m$, the initial state x_0 and the processes $\{F_k;\ k \geq 0\}$ and $\{H_k^{(i)};\ k \geq 1\}$ are mutually independent and they are independent of the additive noises $\{w_k;\ k \geq 0\}$ and $\{v_k^{(i)};\ k \geq 1\}$.

Remark 1. By denoting $\widetilde{F}_k = F_k - \overline{F}_k$, $\widetilde{H}_k^{(i)} = H_k^{(i)} - \overline{H}_k^{(i)}$, $i = 1, 2, \ldots, m$, and G an arbitrary deterministic matrix, the following identities hold for the (p, q)-th entries of the matrices $E[\widetilde{F}_k G \widetilde{F}_k^T]$ and $E[\widetilde{H}_k^{(i)} G \widetilde{H}_k^{(i)T}]$:

$$\Big(E[\widetilde{F}_k G \widetilde{F}_k^T]\Big)_{pq} = \sum_{a=1}^{n_x}\sum_{b=1}^{n_x} Cov[f_{pa}(k), f_{qb}(k)]G_{ab}, \quad p, q = 1, 2, \ldots, n_x.$$

$$\Big(E[\widetilde{H}_k^{(i)} G \widetilde{H}_k^{(i)T}]\Big)_{pq} = \sum_{a=1}^{n_x}\sum_{b=1}^{n_x} Cov[h_{pa}^{(i)}(k), h_{qb}^{(i)}(k)]G_{ab}, \quad p, q = 1, 2, \ldots, n_y.$$

Remark 2. Assumptions *(i)-(iii)* lead to the following recursive formula for $\mathcal{D}_k \equiv E[x_k x_k^T]$, the correlation matrix of the state vector x_k (see, e.g. [17]):

$$\begin{array}{l} \mathcal{D}_{k+1} = \overline{F}_k \mathcal{D}_k \overline{F}_k^T + E[\widetilde{F}_k \mathcal{D}_k \widetilde{F}_k^T] + Q_k + \overline{F}_k Q_{k-1,k} + Q_{k,k-1}\overline{F}_k^T, \quad k \geq 1; \\ \mathcal{D}_1 = \overline{F}_0 \mathcal{D}_0 \overline{F}_0^T + E[\widetilde{F}_0 \mathcal{D}_0 \widetilde{F}_0^T] + Q_0, \\ \mathcal{D}_0 = \Sigma_0 + \overline{x}_0 \overline{x}_0^T. \end{array} \tag{5.3}$$

Our aim is to address the optimal least-squares (LS) linear filtering problem of state x_k by fusing effectively the observations $y_1^{(i)}, \ldots, y_k^{(i)}$, $i = 1, 2, \ldots, m$; specifically, we use the traditional centralized and distributed fusion methods. As is known, the main drawbacks of the first are the expensive computational cost, poor robustness and flexibility. The latter overcomes these disadvantages and provides greater accuracy than local estimators. The centralized fusion method use all measurement data coming from m sensors directly in the fusion center for state estimation, while in the distributed fusion method the observations in the fusion center are replaced with estimates that have been locally computed.

Centralized fusion algorithm. Combining m measurement equations of (5.2) by setting $y_k = \left(y_k^{(1)T}, \ldots, y_k^{(m)T}\right)^T$, the discrete-time multi-sensor system with random parameter matrices and correlated additive noises (5.1)-(5.2) considered in this paper, is a special case of the discrete-time stochastic system with random parameter matrices and correlated additive noises considered in [19]. Hence, the

optimal centralized fusion filter could be obtained by the optimal LS lineal filtering algorithm in [19].

Distributed fusion algorithm. The distributed fusion method computes, at each sensor, a local optimal LS linear state filter using its own measurement data, and, subsequently, the fusion center computes the LS matrix-weighted linear combination of the local filtering estimators. Hence, the distributed fusion filtering algorithm is performed in two steps. In the first one (Section 5.3), for each $i = 1, 2, \ldots, m$, a local LS linear estimator of the signal x_k, denoted by $\widehat{x}_{k/k}^{(i)}$, is produced using the measurements $y_1^{(i)}, \ldots, y_k^{(i)}$, by a recursive algorithm. In the second step (Section 5.4), a fusion distributed estimator, $\widehat{x}_{k/k}^{(D)}$, is generated by a matrix-weighted linear combination of the local estimators, $\widehat{x}_{k/k}^{(i)},\ i = 1, 2, \ldots, m$, using the mean squared error as optimality criterion.

5.3 Local LS linear filtering algorithm

This section is concerned with the problem of obtaining, for each $i = 1, 2, \ldots, m$, a recursive algorithm for the local LS linear filter, $\widehat{x}_{k/k}^{(i)}$, which will be derived by using an innovation approach. For the i-th sensor, the innovation at time k is defined as $\mu_k^{(i)} = y_k^{(i)} - \widehat{y}_{k/k-1}^{(i)}$, where $\widehat{y}_{k/k-1}^{(i)}$ is the LS linear estimator of $y_k^{(i)}$ based on measurements $y_s^{(i)},\ s \leq k-1$. Since the innovation process is uniquely determined by the observations, and the innovations are orthogonal vectors, replacing the observation process by the innovation one, the LS linear estimator, $\widehat{z}_{k/L}^{(i)}$, of a random vector z_k based on the observations $y_1^{(i)}, \ldots, y_L^{(i)}$, can be calculated as linear combination of the innovations $\mu_1^{(i)}, \ldots, \mu_L^{(i)}$; namely,

$$\widehat{z}_{k/L}^{(i)} = \sum_{s=1}^{L} E[z_k \mu_s^{(i)T}](E[\mu_s^{(i)} \mu_s^{(i)T}])^{-1} \mu_s^{(i)}, \quad k \geq 1. \tag{5.4}$$

From (5.4), by denoting $\mathcal{X}_k^{(i)} = E[x_k \mu_k^{(i)T}]$ and $\Pi_k^{(i)} = E[\mu_k^{(i)} \mu_k^{(i)T}]$, the following

expression for the filter, $\widehat{x}_{k/k}^{(i)}$, in terms of the predictor, $\widehat{x}_{k/k-1}^{(i)}$, is obvious:

$$\begin{array}{l} \widehat{x}_{k/k}^{(i)} = \widehat{x}_{k/k-1}^{(i)} + \mathcal{X}_k^{(i)} \Pi_k^{(i)-1} \mu_k^{(i)}, \quad k \geq 1; \\ \widehat{x}_{0/0}^{(i)} = \overline{x}_0, \end{array} \tag{5.5}$$

and, from this relation and the Orthogonal Projection Lemma (OPL), the filtering error covariance matrix, $\Sigma_{k/k}^{(i)}$, is given by

$$\begin{array}{l} \Sigma_{k/k}^{(i)} = \Sigma_{k/k-1}^{(i)} - \mathcal{X}_k^{(i)} \Pi_k^{(i)-1} \mathcal{X}_k^{(i)T}, \quad k \geq 1; \\ \Sigma_{0/0}^{(i)} = \Sigma_0. \end{array} \tag{5.6}$$

Next, we obtain the state predictor $\widehat{x}_{k/k-1}^{(i)}$, the matrix $\mathcal{X}_k^{(i)}$ and the innovation $\mu_k^{(i)}$ or, equivalently, the observation predictor $\widehat{y}_{k/k-1}^{(i)}$.

From the system equations (5.1)-(5.2) and again the OPL, it is clear that the state and observation one-stage predictors verify:

$$\widehat{x}_{k/k-1}^{(i)} = \overline{F}_{k-1} \widehat{x}_{k-1/k-1}^{(i)} + \widehat{w}_{k-1/k-1}^{(i)}, \quad k \geq 1, \tag{5.7}$$

$$\widehat{y}_{k/k-1}^{(i)} = \overline{H}_k^{(i)} \widehat{x}_{k/k-1}^{(i)} + \widehat{v}_{k/k-1}^{(i)}, \quad k \geq 1. \tag{5.8}$$

Because of the correlation Assumption *(iii)*, the noise filter $\widehat{w}_{k/k}^{(i)}$ and the one-stage predictor $\widehat{v}_{k/k-1}^{(i)}$ are not zero and, hence, expressions for such estimators are necessary.

5.3.1 LS linear noise estimators $\widehat{w}_{k/k}^{(i)}$ and $\widehat{v}_{k/k-1}^{(i)}$

The following correlation properties of the vector noises w_k and $v_k^{(i)}$ are easily inferred from the assumptions *(iii)-(iv)*:

- For $i = 1, 2, \ldots, m$, the noise vector w_k is uncorrelated with the innovations $\mu_1^{(i)}, \ldots, \mu_{k-1}^{(i)}$, and correlated with $\mu_k^{(i)}$, with

$$\mathcal{W}_k^{(i)} \equiv E[w_k \mu_k^{(i)T}] = Q_{k,k-1} \overline{H}_k^{(i)T} + S_k^{(i)}, \ k \geq 1. \tag{5.9}$$

- For $i, j = 1, 2, \ldots, m$, the noise vector $v_k^{(i)}$ is uncorrelated with the innovations $\mu_1^{(j)}, \ldots, \mu_{k-2}^{(j)}$, and correlated with $\mu_{k-1}^{(j)}$, with

$$\mathcal{V}_{k,k-1}^{(ij)} \equiv E[v_k^{(i)} \mu_{k-1}^{(j)T}] = S_{k-2,k}^{(i)T} \overline{H}_{k-1}^{(j)T} + R_{k,k-1}^{(ij)}, \ k \geq 2. \tag{5.10}$$

From the general expression for the estimators (5.4) and the above properties, we have that the noise filter, $\widehat{w}_{k/k}^{(i)}$, and the one-stage predictor, $\widehat{v}_{k/k-1}^{(i)}$, satisfy:

$$\widehat{w}_{k/k}^{(i)} = \mathcal{W}_k^{(i)} \Pi_k^{(i)-1} \mu_k^{(i)}, \quad k \geq 1; \quad \widehat{w}_{0/0}^{(i)} = 0. \tag{5.11}$$

$$\widehat{v}_{k/k-1}^{(i)} = \mathcal{V}_{k,k-1}^{(i)} \Pi_{k-1}^{(i)-1} \mu_{k-1}^{(i)}, \quad k \geq 2; \quad \widehat{v}_{1/0}^{(i)} = 0. \tag{5.12}$$

5.3.2 State predictor, $\widehat{x}_{k/k-1}^{(i)}$, and matrix $\mathcal{X}_k^{(i)} = E[x_k \mu_k^{(i)T}]$

From (5.7) and (5.11), the following expression for the state predictor in terms of the filter is immediate:

$$\begin{array}{l} \widehat{x}_{k/k-1}^{(i)} = \overline{F}_{k-1} \widehat{x}_{k-1/k-1}^{(i)} + \mathcal{W}_{k-1}^{(i)} \Pi_{k-1}^{(i)-1} \mu_{k-1}^{(i)}, \quad k \geq 2; \\ \widehat{x}_{1/0}^{(i)} = \overline{F}_0 \overline{x}_0, \end{array} \tag{5.13}$$

with $\mathcal{W}_k^{(i)}$ satisfying (5.9). This expression together with the state equation (5.1) lead to the following formula for $\Sigma_{k/k-1}^{(i)}$, the prediction error covariance matrix:

$$\begin{array}{l} \Sigma_{k/k-1}^{(i)} = \overline{F}_{k-1} \Sigma_{k-1/k-1}^{(i)} \overline{F}_{k-1}^T + E[\widetilde{F}_{k-1} \mathcal{D}_{k-1} \widetilde{F}_{k-1}^T] + Q_{k-1} + \overline{F}_{k-1} \mathcal{J}_{k-1}^{(i)} \\ \qquad + \mathcal{J}_{k-1}^{(i)T} \overline{F}_{k-1}^T - \mathcal{W}_{k-1}^{(i)} \Pi_{k-1}^{(i)-1} \mathcal{W}_{k-1}^{(i)T}, \quad k \geq 2; \\ \Sigma_{1/0}^{(i)} = \overline{F}_0 \Sigma_{0/0} \overline{F}_0^T + E[\widetilde{F}_0 \mathcal{D}_0 \widetilde{F}_0^T] + Q_0, \end{array} \tag{5.14}$$

where $\mathcal{D}_k$ is given by (5.3), and $\mathcal{J}_k^{(i)} = E[(x_k - \widehat{x}_{k/k}^{(i)}) w_k^T]$ satisfies

$$\mathcal{J}_k^{(i)} = Q_{k-1,k} - \mathcal{X}_k^{(i)} \Pi_k^{(i)-1} \mathcal{W}_k^{(i)T}, \quad k \geq 1. \tag{5.15}$$

From (5.13) and (5.5), the following recursive expression for the state predictor is obtained:

$$\widehat{x}_{k/k-1}^{(i)} = \overline{F}_{k-1} \widehat{x}_{k-1/k-2}^{(i)} + \mathcal{X}_{k,k-1}^{(i)} \Pi_{k-1}^{(i)-1} \mu_{k-1}^{(i)}, \quad k \geq 2, \tag{5.16}$$

where

$$\mathcal{X}_{k,k-1}^{(i)} = E[x_k \mu_{k-1}^{(i)T}] = \overline{F}_{k-1}\mathcal{X}_{k-1}^{(i)} + \mathcal{W}_{k-1}^{(i)}, \quad k \geq 2. \tag{5.17}$$

Using relation (5.16) and the state equation (5.1), we obtain that the correlation between the prediction error and the measurement noise vector, $\mathcal{M}_k^{(ij)} \equiv E[(x_k - \widehat{x}_{k/k-1}^{(i)})v_k^{(j)T}]$, is given by

$$\begin{array}{l} \mathcal{M}_k^{(ij)} = \overline{F}_{k-1}S_{k-2,k}^{(j)} + S_{k-1,k}^{(j)} - \mathcal{X}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mathcal{V}_{k,k-1}^{(ji)T}, \ k \geq 2; \\ \mathcal{M}_1^{(ij)} = S_{0,1}^{(j)}. \end{array} \tag{5.18}$$

This correlation property allows us to obtain easily an expression for the matrix $\mathcal{X}_k^{(i)} = E[x_k \mu_k^{(i)T}]$. Indeed, by the OPL,

$$\mathcal{X}_k^{(i)} = E[(x_k - \widehat{x}_{k/k-1}^{(i)})\mu_k^{(i)T}] = E[(x_k - \widehat{x}_{k/k-1}^{(i)})y_k^{(i)T}]$$

Now, from (5.2) and again the OPL, we have

$$\mathcal{X}_k^{(i)} = \Sigma_{k/k-1}^{(i)}\overline{H}_k^{(i)T} + \mathcal{M}_k^{(i)}, \quad k \geq 1, \tag{5.19}$$

where $\mathcal{M}_k^{(i)}$ is given by (5.18).

5.3.3 Innovation, $\mu_k^{(i)}$, and its covariance matrix, $\Pi_k^{(i)}$

From (5.8) and (5.12), the innovation, $\mu_k^{(i)} = y_k^{(i)} - \widehat{y}_{k/k-1}^{(i)}$, is given by:

$$\begin{array}{l} \mu_k^{(i)} = y_k^{(i)} - \overline{H}_k^{(i)}\widehat{x}_{k/k-1}^{(i)} - \mathcal{V}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mu_{k-1}^{(i)}, \quad k \geq 2; \\ \mu_1^{(i)} = y_1^{(i)} - \overline{H}_1^{(i)}\widehat{x}_{1/0}^{(i)}. \end{array} \tag{5.20}$$

with $\mathcal{V}_{k,k-1}^{(i)}$ satisfying (5.10).

Next, an expression for $\Pi_k^{(i)} = E[\mu_k^{(i)}\mu_k^{(i)T}]$ is derived. From the OPL and (5.2), it is clear that

$$\Pi_k^{(i)} = E[y_k^{(i)}\mu_k^{(i)T}] = \overline{H}_k^{(i)}E[x_k\mu_k^{(i)T}] + E[\widetilde{H}_k^{(i)}x_k\mu_k^{(i)T}] + E[v_k^{(i)}\mu_k^{(i)T}].$$

– From (5.2) and the conditional expectation properties,

$$E[\widetilde{H}_k^{(i)} x_k \mu_k^{(i)T}] = E[\widetilde{H}_k^{(i)} x_k x_k^T \widetilde{H}_k^{(i)T}] = E[\widetilde{H}_k^{(i)} \mathcal{D}_k \widetilde{H}_k^{(i)T}], \quad k \geq 1.$$

– Using (5.20) with (5.2) for $y_k^{(i)}$, we obtain that

$$\begin{aligned} E[v_k^{(i)} \mu_k^{(i)T}] &= \mathcal{M}_k^{(i)T} \overline{H}_k^{(i)T} + R_k^{(i)} - \mathcal{V}_{k,k-1}^{(i)} \Pi_{k-1}^{(i)-1} \mathcal{V}_{k,k-1}^{(i)T}, \quad k \geq 2; \\ E[v_1^{(i)} \mu_1^{(i)T}] &= \mathcal{M}_1^{(i)T} \overline{H}_1^{(i)T} + R_1^{(i)}. \end{aligned}$$

Therefore, $\Pi_k^{(i)}$ is determined as follows:

$$\begin{aligned} \Pi_k^{(i)} &= E[\widetilde{H}_k^{(i)} \mathcal{D}_k \widetilde{H}_k^{(i)T}] + \overline{H}_k^{(i)} \mathcal{X}_k^{(i)} + \mathcal{M}_k^{(i)T} \overline{H}_k^{(i)T} + R_k^{(i)} \\ &\quad - \mathcal{V}_{k,k-1}^{(i)} \Pi_{k-1}^{(i)-1} \mathcal{V}_{k,k-1}^{(i)T}, \quad k \geq 2; \\ \Pi_1^{(i)} &= E[\widetilde{H}_1^{(i)} \mathcal{D}_1 \widetilde{H}_1^{(i)T}] + \overline{H}_1^{(i)} \mathcal{X}_1^{(i)} + \mathcal{M}_1^{(i)T} \overline{H}_1^{(i)T} + R_1^{(i)}. \end{aligned} \tag{5.21}$$

5.3.4 Computational procedure

The computational procedure of the proposed local LS linear filtering algorithm can be summarized as follows:

The matrices $\mathcal{W}_k^{(i)}$ and $\mathcal{V}_{k,k-1}^{(i)}$ are computed by expressions (5.9) and (5.10), respectively. We obtain $\mathcal{D}_k$ recursively by (5.3), where the matrix $E[\widetilde{F}_{k-1} \mathcal{D}_{k-1} \widetilde{F}_{k-1}^T]$ necessary to compute $\mathcal{D}_k$ and $\Sigma_{k/k-1}^{(i)}$ is obtained as indicated in *Remark 1*. The matrix $E[\widetilde{H}_k \mathcal{D}_k \widetilde{H}_k^T]$ is also computed in order to obtain the innovation covariance matrix $\Pi_k^{(i)}$. Note that all these matrices depend only on the system model information and can be obtained before the observations are available.

At the sampling time k, starting with the prior knowledge including $\mathcal{X}_{k-1}^{(i)}$, $\mu_{k-1}^{(i)}$, $\Pi_{k-1}^{(i)}$, $\widehat{x}_{k/k-1}^{(i)}$, $\Sigma_{k/k-1}^{(i)}$; that is, once the $(k-1)$-th iteration is finished, then when the new observation $y_k^{(i)}$ is available, the proposed filtering algorithm operates by the following steps:

Step 1: By (5.17), we compute $\mathcal{X}_{k,k-1}^{(i)}$ and, from this, $\mathcal{M}_k^{(i)}$ by (5.18), and then we compute $\mathcal{X}_k^{(i)}$ by (5.19):

$$\begin{aligned} \mathcal{X}_{k,k-1}^{(i)} &= \overline{F}_{k-1} \mathcal{X}_{k-1}^{(i)} + \mathcal{W}_{k-1}^{(i)}, \\ \mathcal{M}_k^{(i)} &= \overline{F}_{k-1} S_{k-2,k}^{(i)} + S_{k-1,k}^{(i)} - \mathcal{X}_{k,k-1}^{(i)} \Pi_{k-1}^{(i)-1} \mathcal{V}_{k,k-1}^{(i)T}, \\ \mathcal{X}_k^{(i)} &= \Sigma_{k/k-1}^{(i)} \overline{H}_k^{(i)T} + \mathcal{M}_k^{(i)}. \end{aligned}$$

Step 2: The innovation $\mu_k^{(i)}$ and its covariance matrix $\Pi_k^{(i)}$ are computed by (5.20) and (5.21), respectively:

$$\mu_k^{(i)} = y_k^{(i)} - \overline{H}_k^{(i)}\widehat{x}_{k/k-1}^{(i)} - \mathcal{V}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mu_{k-1}^{(i)},$$
$$\Pi_k^{(i)} = E[\widetilde{H}_k^{(i)}\mathcal{D}_k\widetilde{H}_k^{(i)T}] + \overline{H}_k^{(i)}\mathcal{X}_k^{(i)} + \mathcal{M}_k^{(i)T}\overline{H}_k^{(i)T} + R_k^{(i)} - \mathcal{V}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mathcal{V}_{k,k-1}^{(i)T}.$$

Step 3: The filter, $\widehat{x}_{k/k}^{(i)}$, and the filtering error covariance matrix, $\Sigma_{k/k}^{(i)}$, are computed by (5.5) and (5.6), respectively:

$$\widehat{x}_{k/k}^{(i)} = \widehat{x}_{k/k-1}^{(i)} + \mathcal{X}_k^{(i)}\Pi_k^{(i)-1}\mu_k^{(i)},$$
$$\Sigma_{k/k}^{(i)} = \Sigma_{k/k-1}^{(i)} - \mathcal{X}_k^{(i)}\Pi_k^{(i)-1}\mathcal{X}_k^{(i)T}.$$

Step 4: To implement the above steps at time $k+1$, we must:

a) Compute the state predictor, $\widehat{x}_{k+1/k}^{(i)}$, by (5.13):

$$\widehat{x}_{k+1/k}^{(i)} = \overline{F}_k\widehat{x}_{k/k}^{(i)} + \mathcal{W}_k^{(i)}\Pi_k^{(i)-1}\mu_k^{(i)}.$$

b) Compute $\mathcal{J}_k^{(i)}$ by (5.15):

$$\mathcal{J}_k^{(i)} = Q_{k-1,k} - \mathcal{X}_k^{(i)}\Pi_k^{(i)-1}\mathcal{W}_k^{(i)T},$$

and from this, the prediction error covariance matrix, $\Sigma_{k+1/k}^{(i)}$, by (5.14):

$$\Sigma_{k+1/k}^{(i)} = \overline{F}_k\Sigma_{k/k}^{(i)}\overline{F}_k^T + E[\widetilde{F}_k\mathcal{D}_k\widetilde{F}_k^T] + Q_k$$
$$+\overline{F}_k\mathcal{J}_k^{(i)} + \mathcal{J}_k^{(i)T}\overline{F}_k^T - \mathcal{W}_k^{(i)}\Pi_k^{(i)-1}\mathcal{W}_k^{(i)T}.$$

5.4 Distributed fusion filtering estimators

Once the local LS linear filters, $\widehat{x}_{k/k}^{(i)}$ for each sensor $i = 1, 2, \ldots, m$, have been obtained, our objective in this section is to design a distributed fusion filter, $\widehat{x}_{k/k}^{(D)}$, by a matrix-weighted linear combination of such estimators, which minimizes the mean squared estimation error. To simplify the derivation of the proposed fusion

estimators, we previously present some useful lemmas that provide the expectations $K_{k/k-1}^{(ij)} = E[\widehat{x}_{k/k-1}^{(i)}\widehat{x}_{k/k-1}^{(j)T}]$, $L_k^{(ij)} = E[\widehat{x}_{k/k-1}^{(i)}\mu_k^{(j)T}]$ and $\Pi_k^{(ij)} = E[\mu_k^{(i)}\mu_k^{(j)}]$, necessary for subsequent calculations. The assumptions and notation in these lemmas are those of Section 5.3.

5.4.1 Previous results

Lemma 5.4.1 *For $i, j = 1, 2, \ldots, m$, the cross-covariance matrix between any two local state predictors, $K_{k/k-1}^{(ij)} = E[\widehat{x}_{k/k-1}^{(i)}\widehat{x}_{k/k-1}^{(j)T}]$, satisfies*

$$\begin{aligned}
K_{k/k-1}^{(ij)} &= \overline{F}_{k-1}K_{k-1/k-2}^{(ij)}\overline{F}_{k-1}^{T} + \overline{F}_{k-1}L_{k-1}^{(ij)}\Pi_{k-1}^{(j)-1}\mathcal{X}_{k,k-1}^{(j)T} \\
&\quad + \mathcal{X}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\Pi_{k-1}^{(ij)}\Pi_{k-1}^{(j)-1}\mathcal{X}_{k,k-1}^{(j)T} + \mathcal{X}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}L_{k-1}^{(ji)T}\overline{F}_{k-1}^{T}, \quad k \geq 2; \quad i \neq j, \\
K_{1/0}^{(ij)} &= \overline{F}_0\overline{x}_0\overline{x}_0^T\overline{F}_0^T, \\
K_{k/k-1}^{(i)} &= \mathcal{D}_k - \Sigma_{k/k-1}^{(i)}, \quad k \geq 1.
\end{aligned}$$

Proof. Using (5.16) and that $L_k^{(ij)} = E[\widehat{x}_{k/k-1}^{(i)}\mu_k^{(j)T}]$ and $\Pi_k^{(ij)} = E[\mu_k^{(i)}\mu_k^{(j)}]$, the proof of this lemma is immediately clear. □

Lemma 5.4.2 *For $i, j = 1, 2, \ldots, m$ and $i \neq j$, the expectation $L_k^{(ij)} = E[\widehat{x}_{k/k-1}^{(i)}\mu_k^{(j)T}]$ satisfies*

$$\begin{aligned}
L_k^{(ij)} &= \left(K_{k/k-1}^{(i)} - K_{k/k-1}^{(ij)}\right)\overline{H}_k^{(j)T} + \mathcal{X}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mathcal{V}_{k,k-1}^{(ji)T} - L_{k,k-1}^{(ij)}\Pi_{k-1}^{(j)-1}\mathcal{V}_{k,k-1}^{(j)T}, \quad k \geq 2; \\
L_1^{(ij)} &= 0,
\end{aligned}$$

where $L_{k,k-1}^{(ij)} = E[\widehat{x}_{k/k-1}^{(i)}\mu_{k-1}^{(j)T}]$ is given by

$$L_{k,k-1}^{(ij)} = \overline{F}_{k-1}L_{k-1}^{(ij)} + \mathcal{X}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\Pi_{k-1}^{(ij)}, \quad k \geq 2.$$

Proof. Taking into account expression (5.20) for the innovation $\mu_k^{(j)}$, we have that

$$L_k^{(ij)} = E[\widehat{x}_{k/k-1}^{(i)}y_k^{(j)T}] - K_{k/k-1}^{(ij)}\overline{H}_k^{(j)T} - L_{k,k-1}^{(ij)}\Pi_{k-1}^{(j)-1}\mathcal{V}_{k,k-1}^{(j)T}.$$

Now, using (5.2) for $y_k^{(j)}$, (5.16) for $\widehat{x}_{k/k-1}^{(i)}$, and the OPL, we obtain

$$E[\widehat{x}_{k/k-1}^{(i)}y_k^{(j)T}] = K_{k/k-1}^{(i)}\overline{H}_k^{(j)T} + \mathcal{X}_{k,k-1}^{(i)}\Pi_{k-1}^{(i)-1}\mathcal{V}_{k,k-1}^{(ji)T}.$$

From above relations, the expression for $L_k^{(ij)}$ is immediately derived.

Using again (5.16) for $\widehat{x}_{k/k-1}^{(i)}$, expression for $L_{k,k-1}^{(ij)}$ is also immediately clear, and the proof is completed. □

Lemma 5.4.3 *For $i, j = 1, 2, \ldots, m$ and $i \neq j$, the innovation cross-covariance matrix, $\Pi_k^{(ij)} = E[\mu_k^{(i)} \mu_k^{(j)T}]$, satisfies*

$$\begin{array}{l} \Pi_k^{(ij)} = \overline{H}_k^{(i)} \left(\mathcal{X}_k^{(j)} - L_k^{(ij)} \right) + \mathcal{M}_k^{(ji)T} \overline{H}_k^{(j)T} + R_k^{(ij)} \\ \qquad -\mathcal{V}_{k,k-1}^{(ij)} \Pi_{k-1}^{(j)-1} \mathcal{V}_{k,k-1}^{(j)T} - \mathcal{V}_{k,k-1}^{(i)} \Pi_{k-1}^{(i)-1} \Pi_{k,k-1}^{(ji)T}, \ k \geq 2; \\ \Pi_1^{(ij)} = \overline{H}_1^{(i)} \left(\mathcal{X}_1^{(j)} - L_1^{(ij)} \right) + \mathcal{M}_1^{(ji)T} \overline{H}_1^{(j)T} + R_1^{(ij)}, \end{array}$$

where $\Pi_{k,k-1}^{(ij)} = E[\mu_k^{(i)} \mu_{k-1}^{(j)T}]$ is given by

$$\Pi_{k,k-1}^{(ij)} = \overline{H}_k^{(i)} \left(\mathcal{X}_{k,k-1}^{(j)} - L_{k,k-1}^{(ij)} \right) + \mathcal{V}_{k,k-1}^{(ij)} - \mathcal{V}_{k,k-1}^{(i)} \Pi_{k-1}^{(i)-1} \Pi_{k-1}^{(ij)}, \quad k \geq 2.$$

Proof. Using (5.20) for the innovation $\mu_k^{(i)}$, it is clear that

$$\Pi_k^{(ij)} = E[y_k^{(i)} \mu_k^{(j)T}] - \overline{H}_k^{(i)} L_k^{(ij)} - \mathcal{V}_{k,k-1}^{(i)} \Pi_{k-1}^{(i)-1} E[\mu_{k-1}^{(i)} \mu_k^{(j)T}].$$

A similar reasoning to that used to obtain (5.21) yields

$$\begin{array}{l} E[y_k^{(i)} \mu_k^{(j)T}] = \overline{H}_k^{(i)} \mathcal{X}_k^{(j)} + \mathcal{M}_k^{(ji)T} \overline{H}_k^{(j)T} + R_k^{(ij)} - \mathcal{V}_{k,k-1}^{(ij)} \Pi_{k-1}^{(j)-1} \mathcal{V}_{k,k-1}^{(j)T}, \quad k \geq 2; \\ E[y_1^{(i)} \mu_1^{(j)T}] = \overline{H}_1^{(i)} \mathcal{X}_1^{(i)} + \mathcal{M}_1^{(ji)T} \overline{H}_1^{(j)T} + R_1^{(ij)}. \end{array}$$

So, the expression for $\Pi_k^{(ij)}$ is immediately derived. The expression for $\Pi_{k,k-1}^{(ij)}$ is obtained by an analogous reasoning. □

5.4.2 Design of distributed fusion filter

As we already indicated, our goal is to obtain a distributed fusion filter, $\widehat{x}_{k/k}^{(D)}$, generated by a weighted sum of the local estimators, $\sum_{i=1}^{m} F_k^{(i)} \widehat{x}_{k/k}^{(i)}$, in which the weight matrices, $F_k^{(i)}$, $i = 1, 2, \ldots, m$, are computed to minimize the mean squared estimation error.

So, by denoting $\widehat{X}_{k/k} = \left(\widehat{x}_{k/k}^{(1)T}, \ldots, \widehat{x}_{k/k}^{(m)T}\right)^T$ and $\mathcal{F}_k = \left(F_k^{(1)}, \ldots, F_k^{(m)}\right)$, the aim is finding $\mathcal{F}_k$ such that the estimator $\mathcal{F}_k\widehat{X}_{k/k}$ minimizes

$$E\left[(x_k - \mathcal{F}_k\widehat{X}_{k/k})(x_k - \mathcal{F}_k\widehat{X}_{k/k})^T\right]. \tag{5.22}$$

It is straightforward to check that (5.22) can be expressed as

$$E\left[x_k x_k^T\right] + [\mathcal{F}_k - \mathcal{G}_k]\, E\left[\widehat{X}_{k/k}\widehat{X}_{k/k}^T\right][\mathcal{F}_k - \mathcal{G}_k]^T - \mathcal{G}_k E\left[\widehat{X}_{k/k}\widehat{X}_{k/k}^T\right]\mathcal{G}_k^T,$$

where $\mathcal{G}_k = E\left[x_k\widehat{X}_{k/k}^T\right]\left(E\left[\widehat{X}_{k/k}\widehat{X}_{k/k}^T\right]\right)^{-1}$ and, therefore, (5.22) is minimum when the estimator is performed by the matrix

$$\mathcal{F}_k^{opt} = E\left[x_k\widehat{X}_{k/k}^T\right]\left(E\left[\widehat{X}_{k/k}\widehat{X}_{k/k}^T\right]\right)^{-1}, \quad k \geq 1. \tag{5.23}$$

In the following theorem, the proposed distributed fusion filtering estimators, $\widehat{x}_{k/k}^{(D)}$, and their error covariance matrices, $\Sigma_{k/k}^{(D)}$, are established.

Theorem 5.4.1 *Let $\widehat{X}_{k/k} = \left(\widehat{x}_{k/k}^{(1)T}, \ldots, \widehat{x}_{k/k}^{(m)T}\right)^T$ be the vector formed by the local LS filtering estimators calculated in Section 5.3. Then, the distributed fusion filter is given by*

$$\widehat{x}_{k/k}^{(D)} = \Xi_{k/k}K_{k/k}^{-1}\widehat{X}_{k/k}, \quad k \geq 1, \tag{5.24}$$

with

$$K_{k/k} = \left(K_{k/k}^{(ij)}\right)_{i,j=1,\ldots,m} \quad \textit{and} \quad \Xi_{k/k} = \left(K_{k/k}^{(1)}, \ldots, K_{k/k}^{(m)}\right),$$

where $K_{k/k}^{(ij)} = E\left[\widehat{x}_{k/k}^{(i)}\widehat{x}_{k/k}^{(j)T}\right]$, $i, j = 1, 2, \ldots, m$, are computed by

$$\begin{aligned} K_{k/k}^{(ij)} &= K_{k/k-1}^{(ij)} + L_k^{(ij)}\Pi_k^{(j)-1}\mathcal{X}_k^{(j)T} + \mathcal{X}_k^{(i)}\Pi_k^{(i)-1}L_k^{(ji)T} \\ &\quad + \mathcal{X}_k^{(i)}\Pi_k^{(i)-1}\Pi_k^{(ij)}\Pi_k^{(j)-1}\mathcal{X}_k^{(j)T}, \quad k \geq 1; \quad i \neq j \\ K_{k/k}^{(i)} &= \mathcal{D}_k - \Sigma_{k/k}^{(i)}, \quad k \geq 1, \end{aligned} \tag{5.25}$$

with $K_{k/k-1}^{(ij)}$, $L_k^{(ij)}$ and $\Pi_k^{(ij)}$ given in lemmas 5.4.1, 5.4.2 and 5.4.3, respectively.

The error covariance matrices of the distributed fusion filtering estimators are computed by

$$\Sigma_{k/k}^{(D)} = \mathcal{D}_k - \Xi_{k/k}K_{k/k}^{-1}\Xi_{k/k}^T, \quad k \geq 1. \tag{5.26}$$

Proof. Expressions (5.24) and (5.26) for the distributed estimators and their error covariance matrices, respectively, are immediately derived from (5.23). Expression (5.25) for the cross-covariance matrices between local filtering estimators follows easily using expression (5.5) of such estimators. □

5.5 Numerical simulation example

Consider the following discrete-time linear networked system with state-dependent multiplicative noise, and scalar measurements from four sensors:

$$x_k = (0.95 + 0.2\epsilon_{k-1})x_{k-1} + w_{k-1}, \quad k \geq 1,$$
$$y_k^{(i)} = H_k^{(i)} x_k + v_k^{(i)}, \quad k \geq 1, \quad i = 1, 2, 3, 4$$

where $\{\epsilon_k;\ k \geq 0\}$ is a zero-mean Gaussian white process with unit variance. The additive noises are defined as $w_k = 0.6(\eta_k + \eta_{k+1})$ and $v_k^{(i)} = c_k^{(i)}(\eta_{k-1} + \eta_k),\ i = 1, 2, 3, 4$, where $c_k^{(1)} = 1$, $c_k^{(2)} = 0.5$, $c_k^{(3)} = 0.75$, $c_k^{(4)} = 0.85$, and $\{\eta_k;\ k \geq 0\}$ is a zero-mean Gaussian white process with variance 0.5.

For $i = 1, 2, 3, 4$, the random parameter matrices $H_k^{(i)}$ are defined as follows:

- $H_k^{(1)} = 0.82\lambda_k^{(1)}$, where $\{\lambda_k^{(1)};\ k \geq 1\}$, is a sequence of independent and identically distributed (iid) random variables uniformly distributed over $[0.3, 0.7]$.
- $H_k^{(2)} = 0.75\lambda_k^{(2)}$, where $\{\lambda_k^{(2)};\ k \geq 1\}$, is a sequence of iid discrete random variables with $P[\lambda_k^{(2)} = 0] = 0.1, \quad P[\lambda_k^{(2)} = 0.5] = 0.5, \quad P[\lambda_k^{(2)} = 1] = 0.4$.
- $H_k^{(3)} = 0.74\lambda_k^{(3)}$, where $\{\lambda_k^{(3)};\ k \geq 1\}$, is a Bernoulli process with $P[\lambda_k^{(3)} = 1] = p^{(3)}, \forall k$.
- $H_k^{(4)} = \lambda_k^{(4)}\,(0.75 + 0.95\xi_k)$, where $\{\lambda_k^{(4)};\ k \geq 1\}$ is a Bernoulli process with $P[\lambda_k^{(4)} = 1] = p^{(4)}, \forall k$, and $\{\xi_k;\ k \geq 0\}$ is a zero-mean Gaussian white process with unit variance.

Note that the random parameter matrices at each sensor, $H_k^{(i)}$, allow to model different types of uncertainty. Namely, in sensors 1 and 2, as in [15] and [6], the scalar random variables $\lambda_k^{(i)}$ take values over the interval $[0, 1]$ and represent continuous and discrete stochastic sensor gain degradations, respectively. In sensor 3, $\lambda_k^{(3)}$ are Bernoulli random variables, thus covering the phenomenon of missing measurements, with $\lambda_k^{(3)} = 1$ meaning that the signal x_k is present in the measurement $y_k^{(3)}$ coming from the third sensor at time k, while $\lambda_k^{(i)} = 0$ means that the signal is missing in the measured data at time k or, equivalently, that such observation is only noise $v_k^{(3)}$. Finally, as in [4], both missing measurements and multiplicative noise are considered in sensor 4.

To illustrate the feasibility and effectiveness of the proposed algorithms, they were implemented in MATLAB, and one hundred iterations of the algorithms were run. Using simulated values of the signal and the corresponding observations, both centralized and distributed filtering estimates were calculated, as well as the corresponding error variances, in order to measure the estimation accuracy.

The error variances of the local, centralized and distributed fusion filters are compared considering fixed values of the probabilities $p^{(3)}$ and $p^{(4)}$; specifically, $p^{(3)} = 0.5$, $p^{(4)} = 0.75$. In Fig.5.1, we can see that the error variances of the distributed fusion filter are significantly smaller than those of every local filter, but lightly greater than those of the centralized filter. Nevertheless, although the centralized fusion filter outperforms the distributed one, this difference is slight and both filters perform similarly and provide good estimations. Besides, this slight difference is compensated by the fact that the distributed fusion structure reduces the computational cost and has the advantage of better robustness and fault tolerance. For example, assuming that the fourth sensor is faulty and the measurement equation is given by $y_k^{(4)} = H_k^{(4)} x_k + v_k^{(4)} + a_k$, where $a_k = 0.5k$ for $50 \leq k \leq 70$ and $a_k = 0$ otherwise, Fig.5.2 displays the corresponding filtering mean square errors (MSE) of one thousand independent simulations at each sampling time k, show-

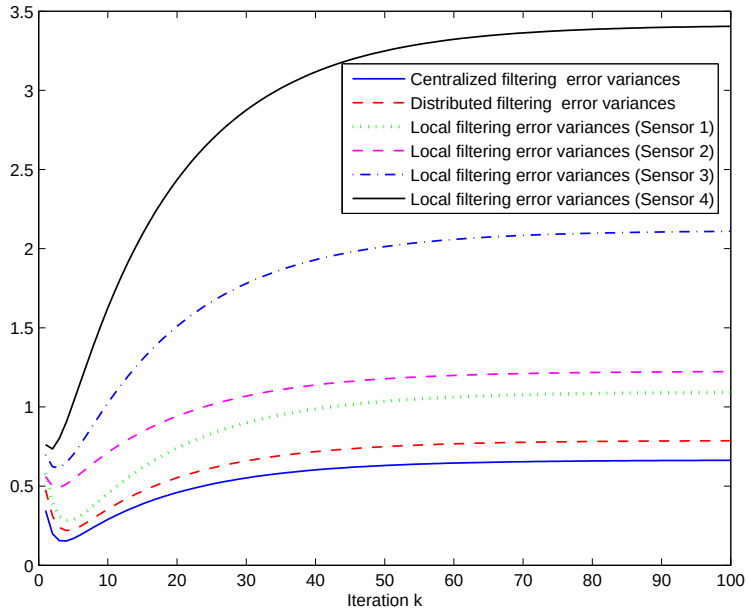


Figure 5.1: Error variance comparison of the centralized, distributed and local filtering estimators.

ing that the distributed fusion method has better fault-tolerance abilities than the centralized one.

Finally, we analyze the centralized and distributed filtering accuracy in function of the probabilities $p^{(3)}$ and $p^{(4)}$ of the Bernoulli variables that model uncertainties of the observations coming from sensors 3 and 4, respectively. Specifically, the filter performances are analyzed when $p^{(3)}$ is varied from 0.1 to 0.9, and different values of $p^{(4)}$ are considered. Since the behaviour of the error variances is analogous in all the iterations, only the results of a specific iteration ($k = 100$) are shown here. In Figure 5.3 the centralized and distributed filtering error variances are displayed versus $p^{(3)}$, for $p^{(4)} = 0.3,\ 0.5,\ 0.7$ and 0.9. As expected, from this figure it is concluded that, as $p^{(3)}$ or $p^{(4)}$ increase, the centralized and distributed filtering error variances both become smaller and, hence, the performance of the centralized and distributed filters improve as these probabilities increase.

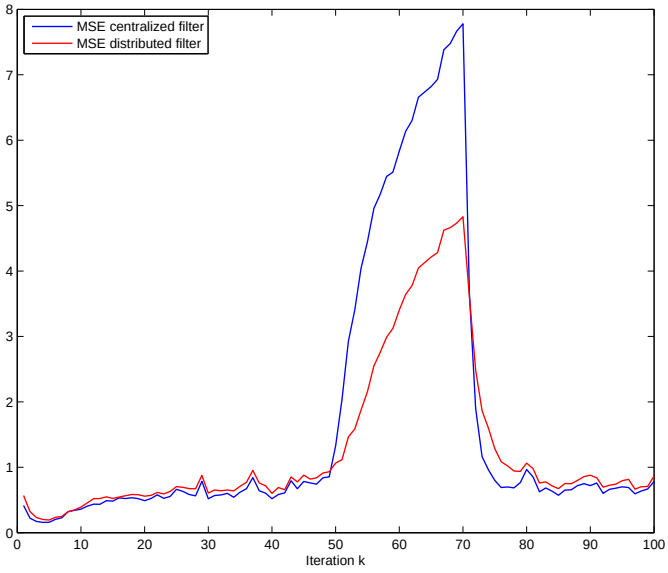


Figure 5.2: Centralized and distributed filtering mean square errors.

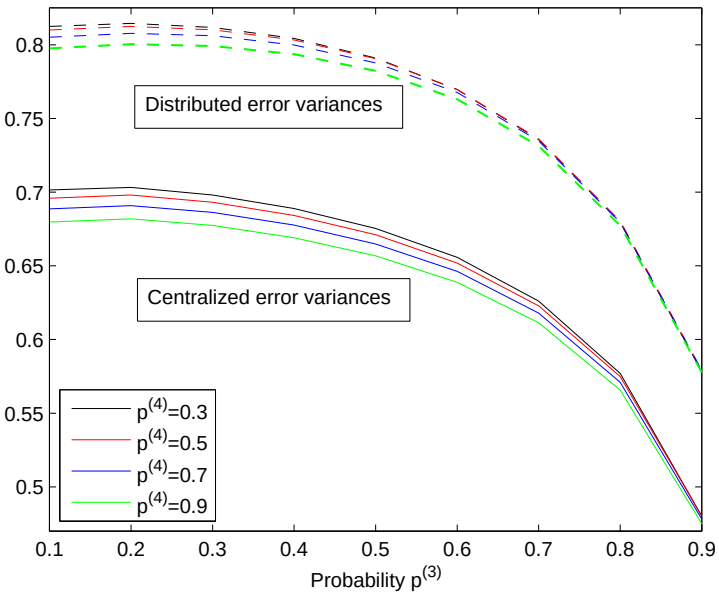


Figure 5.3: Centralized and distributed filtering error variances at $k = 100$ versus $p^{(3)}$, when $p^{(4)} = 0.3,\ 0.5,\ 0.7$ and 0.9

5.6 Conclusion

The distributed fusion filtering problem has been investigated for multi-sensor stochastic systems with random parameter matrices and correlated noises. The main outcomes and results can be summarized as follows:

- Recursive algorithms for the local LS linear filters of the system state based on the measured output data coming from each sensor have been designed by an innovation approach. The computational procedure of these local filtering algorithms is very simple and suitable for online applications. To measure the accuracy of the local estimators, recursive formulas for the local filtering error covariance matrices have been also established.

- Once the local filters have been obtained, a distributed fusion filter has been designed as the matrix-weighted sum of such local estimators that minimizes the mean-squared estimation error. The error covariance matrices of such distributed fusion filter have been also derived.

- A numerical simulation example has illustrated the usefulness of the proposed results. Error variance comparison has shown that both the centralized and the distributed filters outperform the local ones; this example has also shown that the slight superiority of the centralized filter over the distributed filter is compensated by better robustness and fault-tolerance abilities of the latter. This example has also highlighted the applicability of the proposed algorithm for a great variety of multi-sensor systems featuring network-induced stochastic uncertainties, such as sensor gain degradation, missing measurements or multiplicative observation noises, which can be dealt with the observation model considered in this paper.

A challenging further research topic is to address the estimation problem for this kind of systems with random parameter matrices, considering a sensor network

whose nodes are distributed according to a given topology, characterized by a directed graph. Also, an interesting future research topic is to consider other kinds of stochastic uncertainties which often appear in networked systems, such as random delays and packet dropouts.

Acknowledgements

This work was supported by the Ministerio de Economía y Competitividad [Number MTM2014-52291-P and FPU programme].

References

[1] S. L. Sun and Z. L. Deng, Multi-sensor optimal information fusion Kalman filter, *Automatica,* 40(6) (2004), 1017–1023.

[2] X. Sun and G. Yan, Distributed State Fusion Kalman Predictor Based on WLS Algorithm, *Procedia Engineerin,* 29 (2012), 216–222.

[3] L. Yan, X. Rong Li, Y. Xia and M. Fu, Optimal sequential and distributed fusion for state estimation in cross-correlated noise, *Automatica,* 49 (2013), 3607–3612.

[4] J. Ma and S. Sun, Centralized Fusion Estimators for Multi-sensor Systems with Multiplicative Noises and Missing Measurements, *Journal of Networks,* 7(10) (2012), 1538–1545.

[5] J. Ma and S. Sun, Centralized fusion estimators for multi-sensor systems with random sensor delays, multiple packet dropouts and uncertain observations, *IEEE Sensors Journal,* 13 (2013), 1228–1235.

[6] J. Linares-Pérez, R. Caballero-Águila and I. García-Garrido, Optimal Fusion Filtering in Multisensor Stochastic Systems with Missing Measurements and

Correlated Noises, *Mathematical Problems in Engineering,* vol. 2013, Article ID 418678, 14 pages.

[7] L. Li and Y. Q. Xia, Unscented Kalman filter over unreliable communication networks with Markovian packet dropouts, *IEEE Transactions on Automatic Control,* 58(12) (2013), 3224–3230.

[8] S. Sun and J. Ma, Linear estimation for networked control systems with random transmission delays and packet dropouts, *Information Sciences,* 269 (2014), 349–365.

[9] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Covariance-based estimation algorithms in networked systems with mixed uncertainties in the observations, *Signal Processing,* 94 (2014a), 163–173.

[10] B. Chen, W. Zhang and L. Yu, Distributed Fusion Estimation With Missing Measurements, Random Transmission Delays and Packet Dropouts, *IEEE Transactions on Automatic Control,* 59(7) (2014), 1961–1967.

[11] J. Hu, Z. Wang, H. Dong and H. Gao, Recent Advances on Recursive Filtering and Sliding Mode Design for Networked Nonlinear Stochastic Systems: A Survey, *Mathematical Problems in Engineering,* vol. 2013, Article ID 646059, 12 pages.

[12] J. Feng, Z. Wang and M. Zeng, Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises, *Information Fusion,* 14 (2013), 78–86.

[13] S. Gao and P. Chen, Suboptimal Filtering of Networked Discrete-Time Systems with Random Observation Losses, *Mathematical Problems in Engineering,* vol. 2014, Article ID 151836, 8 pages.

[14] F. Peng and S. Sun, Distributed fusion estimation for multisensor multirate systems with stochastic observation multiplicative noises, *Mathematical Problems in Engineering,* vol. 2014, Article ID 373270, 8 pages.

[15] Y. Liu, X., He, Z. Wang and D. Zhou, Optimal filtering for networked systems with stochastic sensor gain degradation, *Automatica,* 50(5) (2014), 1521–1525.

[16] Y. Luo, Y. Zhu, D. Luo, J. Zhou, E. Song and D. Wang, Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications, *Sensors,* 8(12) (2008), 8086–8103.

[17] J. Hu, Z. Wang and H. Gao, Recursive filtering with random parameter matrices, multiple fading measurements and correlated noises, *Automatica,* 49 (2013), 3440–3448.

[18] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Covariance-Based Estimation from Multisensor Delayed Measurements with Random Parameter Matrices and Correlated Noises, *Mathematical Problems in Engineering,* vol. 2014, Article ID 958474, 13 pages.

[19] J. Linares-Pérez, R. Caballero-Águila and I. García-Garrido, Optimal linear filter design for systems with correlation in the measurement matrices and noises: recursive algorithm and applications, *International Journal of Systems Science,* 45(7) (2014), 1548–1562.

[20] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Optimal state estimation for networked systems with random parameter matrices, correlated noises and delayed measurements, *International Journal of General Systems,* 44(2) (2015), 142–154.

[21] X. Wang, S.L. Sun, K.H. Ding and J.Y. Xue, Weighted Measurement Fusion White Noise Deconvolution Filter with Correlated Noise for Multisensor

Stochastic Systems, *Mathematical Problems in Engineering,* vol. 2012, Article ID 257619, 16 pages.

[22] J. Feng and M. Zeng, Optimal distributed Kalman filtering fusion for a linear dynamic system with cross-correlated noises, *International Journal of Systems Science,* 43(2) (2012), 385–398.

Chapter 6

Quadratic estimation problem in discrete-time stochastic systems with random parameter matrices

Reference

Caballero-Águila, R., García-Garrido, I. and Linares-Pérez, J. (2015).
Quadratic estimation problem in discrete-time stochastic systems with random parameter matrices.
Applied Mathematics and Computation, (submitted).

Abstract

This paper addresses the least-squares quadratic filtering problem in discrete-time stochastic systems with random parameter matrices in both the state and measurement equations. Defining a suitable augmented system, this problem is reduced to the least-squares linear filtering problem of the augmented state based on the augmented observations. Under the assumption that the moments, up to the fourth-order one, of the original state and measurement vectors are known, a recursive algorithm for the optimal linear filter of the augmented state is designed, from which the optimal quadratic filter of the original state is obtained. As a particular

case, the proposed results are applied to multi-sensor systems with state-dependent multiplicative noise and fading measurements and, finally, a numerical simulation example illustrates the performance of the proposed quadratic filter in comparison with the linear one and also with other quadratic filters in the existing literature.

Contents

6.1 Introduction

A basic assumption in classical estimation theory for linear stochastic systems is the knowledge of the model parameter matrices; also, the additive noises and the initial state are assumed to be Gaussian and mutually independent. As it is well known, under these conditions, the systems are Gaussian and the Kalman filter provides the conditional expectation of the state given the observations, that is, the optimal least-squares (LS) estimator. However, there exists a considerable number of situations in which the joint distribution of the state and the observations is not Gaussian and the Kalman filter provides only the linear LS estimator. In these cases, the optimal LS estimator is not a linear function of the observations and, generally, it is not easy to be obtained; this fact has motivated the necessity of

looking for suboptimal estimators which are computationally easier, such as linear estimators or, more generally, polynomial estimators.

In systems where the usual assumption of Gaussian noises must be removed in order to obtain a more realistic statistical description of the random processes involved, De Santis et *al.* [1] were the first to obtain a recursive algorithm for the quadratic LS filter, by improving the widely used linear filter. A more general study is carried out in [2], where the arbitrary-order polynomial LS estimation problem is addressed.

Systems with multiplicative noises in the state and/or observation equations constitute another kind of non-Gaussian systems in which the Kalman filter does not provide the optimal LS estimator and, hence, it is necessary to look for suboptimal estimators. This class of systems has been receiving great attention in the last years, mainly due to the fact that this kind of formulation arises in many applications, as image processing problems and communication systems. Therefore, under different hypotheses and performance criterions, the study of the linear LS estimation problem in systems with random multiplicative noises has become an active research area in the last years (see e.g. [3]-[5], and reference therein).

Because of its important applications, it is worth noting especially some classes of systems where the influence of multiplicative noises affects only the measurements of the model; for example, in cases where there are intermittent failures in the observation mechanism, fading phenomena in propagation channels, accidental loss of some measurements, or data inaccessibility during certain times. This kind of systems, named systems with uncertain observations or missing measurements, are modeled including in the observation equation, besides the additive noise, a multiplicative noise component described by a sequence of Bernoulli random variables. Under different hypotheses on the Bernoulli variables and the additive noises involved in the system equations, the linear and polynomial estimation problems have been widely studied in such systems (see e.g. [6]-[12], and references therein).

Recently, this missing measurement model, described by Bernoulli variables, has been generalized considering any random variables with arbitrary probability distribution over the interval [0, 1], which allows us to cover some practical applications where only partial information is missing. In this situation, considering also different assumptions on the system noises, the linear LS estimation problem has been treated in [13] and [14].

The above-mentioned systems are a special case of systems with random parameter matrices which clearly are non-Gaussian systems, even under the assumption that the additive noises are Gaussian. Also, systems with random delays and packet dropouts can be transformed into an equivalent stochastic parameterized system [15]. Due to the numerous realistic situations and practical applications in which both state transition and measurement are random parameter matrices, such as digital control of chemical processes, systems with human operators, economic systems, and stochastically sampled digital control systems (see e.g. [16]-[19], among others), the linear estimation problem in this type of systems has gained significant research interest in recent years (see e.g. [13], [20]-[22] and references therein). Considering scalar measurements with random observation matrices, the quadratic LS filtering problem has been addressed in [23] by applying the Kalman filter to a suitably augmented system with deterministic observation matrices.

Despite the importance of this kind of systems and the significant improvement that the quadratic LS estimators provide over the linear ones, to the best of the authors knowledge, the quadratic LS estimation problem in systems with both random parameter state transition and measurement matrices has not been investigated. *This paper makes the following contributions: (1) random parameter matrices in both the state and observation equations are considered simultaneously in the system state-space model, thus providing a unified framework to treat some classes of uncertainties, such as multiplicative noises or missing and fading measurements, and, hence, the proposed quadratic LS filter outperforms the linear LS*

estimators derived in the existing literature for systems with such uncertainties; (2) unlike [23], where deterministic state transition matrices and scalar measurements are assumed, we consider random state transition matrices and multidimensional observations; hence, the proposed estimators can be applied to multi-sensor systems and, furthermore, different uncertainty characteristics in the sensors can be considered; specifically, an application to multi-sensor systems with state-dependent multiplicative noise and fading measurements is presented; (3) also, unlike [23], the proposed quadratic filtering algorithm is obtained without requiring the original system transformation into one with deterministic observation matrices.

The rest of the paper is organized as follows. In Section 6.2, we present the system model with random parameter matrices to be considered and the assumptions and properties under which the quadratic LS estimation problem is addressed. The augmented system is constructed in Section 6.3 using the technique proposed by [1], consisting of augmenting the state and observation vectors with their second-order Kronecker powers. Also, in this section, the statistical properties of the augmented processes are analyzed. The proposed methodology reduces the quadratic estimation problem to the linear estimation problem in the augmented system, and the recursive algorithm for the linear LS filter of the augmented state is derived in Section 6.4. The application to multi-sensor systems with state-dependent multiplicative noise and fading measurements, together with a numerical simulation example which shows the effectiveness of the proposed quadratic estimators in contrast to the linear ones are both presented in Section 6.5. Finally, some conclusions are drawn in Section 6.6.

Notation: The notation used throughout this paper is standard. $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ are the n-dimensional Euclidean space and the set of all $m \times n$ real matrices, respectively. For any matrix A, A^T and A^{-1} denote its transpose and inverse, respectively. The shorthand $Diag(A_1, \ldots, A_m)$ denotes a block diagonal matrix whose diagonal blocks are the matrices $A_1, \ldots, A_m$, and $[A_1 \mid \cdots \mid A_m]$ represents

a partitioned matrix into sub-matrices $A_1, \ldots, A_m$. I and K denote the identity and commutation matrices, respectively, of appropriate dimensions. $A^{[2]} = A \otimes A$ where $\otimes$ denotes the Kronecker product. $vec(\cdot)$ stands for the 'vec' or 'stack' operator. $\delta_{k,s}$ is the Kronecker delta function and $\circ$ denotes the Hadamard product.

Moreover, for any random vector or matrix M, we denote $\overline{M} = E[M]$ and $\widetilde{M} = M - \overline{M}$, where $E[\cdot]$ stands for the expectation operator. For arbitrary random vectors β and γ, we denote $Cov[\beta, \gamma] = E[(\beta - E[\beta])\,(\gamma - E[\gamma])^T]$ and $Cov[\beta] = Cov[\beta, \beta]$.

6.2 Problem formulation

Consider a class of discrete-time linear stochastic systems and denote $x_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}^r$ the state vector and its measurement at time k, respectively. The evolution of the state and its measurements are given by the following equations:

$$x_k = F_{k-1} x_{k-1} + w_{k-1}, \quad k \geq 1, \tag{6.1}$$

$$y_k = H_k x_k + v_k, \quad k \geq 1, \tag{6.2}$$

where $\{F_k\}_{k\geq 0}$ and $\{H_k\}_{k\geq 1}$ are sequences of random parameter matrices, $\{w_k\}_{k\geq 0}$ is the process noise and $\{v_k\}_{k\geq 1}$ is the measurement noise.

Our aim is to obtain the least-squares (LS) quadratic estimator of the state x_k based on the measurements $\{y_1, \ldots, y_k\}$. As it is known, this estimator is its orthogonal projection onto the space of n-dimensional random variables obtained as linear transformations of $y_1, \ldots, y_k$ and their second-order powers, $y_1^{[2]}, \ldots, y_k^{[2]}$. To address the LS quadratic estimation problem, it is necessary that $E[y_i^{[2]T} y_i^{[2]}] < \infty$, and therefore, the fourth-order moments of vectors y_i, $i = 1, \ldots, k$ must be finite. Specifically, the following assumptions are required:

(A1) The initial state x_0 is a random vector whose moments up to the fourth-order one are known. We will denote $\overline{x}_0 = E[x_0]$, $P_0 = Cov[x_0]$, $P_0^{(3)} = Cov[x_0, x_0^{[2]}]$ and $P_0^{(4)} = Cov[x_0^{[2]}]$.

(A2) $\{F_k\}_{k\geq 0}$ and $\{H_k\}_{k\geq 1}$ are sequences of independent random parameter matrices with known mean matrices $\overline{F}_k$ and $\overline{H}_k$. The covariances and cross-covariances between the entries of the matrices F_k and $F_k^{[2]}$, as well as between the entries of the matrices H_k and $H_k^{[2]}$, are also assumed to be known.

(A3) The noise processes $\{w_k\}_{k\geq 0}$ and $\{v_k\}_{k\geq 1}$ are zero-mean white sequences with known moments, up to the fourth-order ones. We will denote

$$\begin{aligned} Q_k &= Cov[w_k], \quad Q_k^{(3)} = Cov[w_k, w_k^{[2]}], \quad Q_k^{(4)} = Cov[w_k^{[2]}], \\ R_k &= Cov[v_k], \quad R_k^{(3)} = Cov[v_k, v_k^{[2]}], \quad R_k^{(4)} = Cov[v_k^{[2]}]. \end{aligned}$$

(A4) The initial state x_0, the random parameter matrices $\{F_k\}_{k\geq 0}$, $\{H_k\}_{k\geq 1}$ and the processes $\{w_k\}_{k\geq 0}$, $\{v_k\}_{k\geq 1}$ are mutually independent.

Remark 1. Hereafter, it will be necessary to calculate different expectations associated with the random parameter matrices F_k, $F_k^{[2]}$, H_k and $H_k^{[2]}$. For this purpose, the following property is used:

Let $A = \left(a_{ij}\right)_{\substack{i=1,\ldots,N_1 \\ j=1,\ldots,N_2}}$ *and* $B = \left(b_{ij}\right)_{\substack{i=1,\ldots,M_1 \\ j=1,\ldots,M_2}}$ *be random parameter matrices, then for any deterministic matrix* $C = \left(c_{ij}\right)_{\substack{i=1,\ldots,N_2 \\ j=1,\ldots,M_2}}$, *the* (p,q)*-th entry of the matrix* $E[\widetilde{A}C\widetilde{B}^T]$ *is given by*

$$\left(E[\widetilde{A}C\widetilde{B}^T]\right)_{pq} = \sum_{i=1}^{N_2}\sum_{j=1}^{M_2} Cov(a_{pi}, b_{qj})c_{ij}, \quad p = 1,\ldots,N_1,\ q = 1,\ldots,M_1. \tag{6.3}$$

6.3 Quadratic estimation problem statement

Given the system model (6.1)-(6.2) under assumptions (A1)-(A4), the problem at hand is to find the LS quadratic estimator, $\widehat{x}_{k/k}^q$, of the state x_k based on the measurements until time k. For this purpose, the following augmented state

and measurement vectors are defined by assembling the original vectors and their second-order Kronecker powers:

$$\mathcal{X}_k = \begin{pmatrix} x_k \\ x_k^{[2]} \end{pmatrix}, \quad \mathcal{Y}_k = \begin{pmatrix} y_k \\ y_k^{[2]} \end{pmatrix}.$$

Since the space of linear transformations of $\mathcal{Y}_1, \ldots, \mathcal{Y}_k$ is equal to the space of linear transformations of $y_1, \ldots, y_k$ and $y_1^{[2]}, \ldots, y_k^{[2]}$, the LS quadratic filter, $\widehat{x}_{k/k}^q$, is the LS linear estimator of x_k based on $\mathcal{Y}_1, \ldots, \mathcal{Y}_k$. This estimator is obtained by extracting the first n entries of the LS linear estimator of $\mathcal{X}_k$ based on $\mathcal{Y}_1, \ldots, \mathcal{Y}_k$. Therefore, the quadratic estimation problem for the original state is reduced to the linear estimation problem for the augmented state.

In order to address the LS linear estimation problem of the augmented state based on the augmented measurements, the evolution of the vectors $\mathcal{X}_k$ and $\mathcal{Y}_k$ is analyzed. Using the Kronecker product properties [24], the evolution of the the second-order powers, $x_k^{[2]}$ and $y_k^{[2]}$, is given by (see [1]):

$$\begin{aligned} x_k^{[2]} &= F_{k-1}^{[2]} x_{k-1}^{[2]} + \Phi_{k-1}, \quad k \geq 1, \\ y_k^{[2]} &= H_k^{[2]} x_k^{[2]} + \Psi_k, \quad k \geq 1, \end{aligned}$$

where $\Phi_k = (I + K)((F_k x_k) \otimes w_k) + w_k^{[2]}$ and $\Psi_k = (I + K)((H_k x_k) \otimes v_k) + v_k^{[2]}$.

Then, the augmented vectors $\mathcal{X}_k$ and $\mathcal{Y}_k$ satisfy the following equations:

$$\begin{aligned} \mathcal{X}_k &= \mathcal{F}_{k-1} \mathcal{X}_{k-1} + \mathcal{W}_{k-1}, \quad k \geq 1, \\ \mathcal{Y}_k &= \mathcal{H}_k \mathcal{X}_k + \mathcal{V}_k, \quad k \geq 1, \end{aligned}$$

where

$$\mathcal{F}_k = Diag(F_k,\, F_k^{[2]}), \quad \mathcal{H}_k = Diag(H_k,\, H_k^{[2]}), \quad \mathcal{W}_k = \begin{pmatrix} w_k \\ \Phi_k \end{pmatrix}, \quad \mathcal{V}_k = \begin{pmatrix} v_k \\ \Psi_k \end{pmatrix}.$$

For simplicity, since the additive noises of this new model, $\mathcal{W}_k$ and $\mathcal{V}_k$, are non-zero mean vectors, with

$$\overline{\mathcal{W}}_k = \begin{pmatrix} 0 \\ vec(Q_k) \end{pmatrix}, \quad \overline{\mathcal{V}}_k = \begin{pmatrix} 0 \\ vec(R_k) \end{pmatrix},$$

the above equations are rewritten in terms of the centered augmented vectors, $X_k = \mathcal{X}_k - \overline{\mathcal{X}}_k$ and $Y_k = \mathcal{Y}_k - \overline{\mathcal{Y}}_k$, obtaining the following augmented system:

$$X_k = \mathcal{F}_{k-1} X_{k-1} + W_{k-1}, \quad k \geq 1, \tag{6.4}$$

$$Y_k = \mathcal{H}_k X_k + V_k, \quad k \geq 1, \tag{6.5}$$

where $W_k = \widetilde{\mathcal{F}}_k \overline{\mathcal{X}}_k + \widetilde{\mathcal{W}}_k$ and $V_k = \widetilde{\mathcal{H}}_k \overline{\mathcal{X}}_k + \widetilde{\mathcal{V}}_k$, being $\overline{\mathcal{X}}_k = \begin{pmatrix} \overline{x}_k \\ vec(D_k) \end{pmatrix}$ with $D_k = E[x_k x_k^T]$.

Taking into account the state equation (6.1) and under assumptions (A1)-(A4), the mean vector $\overline{x}_k$ and the correlation matrix D_k are recursively calculated by

$$\overline{x}_k = \overline{F}_{k-1} \overline{x}_{k-1}, \quad k \geq 1,$$
$$D_k = \overline{F}_{k-1} D_{k-1} \overline{F}_{k-1}^T + E[\widetilde{F}_{k-1} D_{k-1} \widetilde{F}_{k-1}^T] + Q_{k-1},\ k \geq 1; \quad D_0 = P_0 + \overline{x}_0 \overline{x}_0^T,$$

where the (p,q)-th entry of the matrix $E[\widetilde{F}_k D_k \widetilde{F}_k^T]$ is obtained as in (6.3).

It should be mentioned that the LS linear estimator of X_k based on $Y_1, \ldots, Y_k$ provides the LS linear estimator of $\mathcal{X}_k$ based on $\mathcal{Y}_1, \ldots, \mathcal{Y}_k$, adding the mean vector $\overline{\mathcal{X}}_k$. Therefore, the required quadratic filter $\widehat{x}_{k/k}^q$ is obtained by adding the mean $\overline{x}_k$ to the vector constituted by the first n entries of the LS linear filter of X_k.

In order to obtain the LS linear filter of X_k, the properties of the processes involved in the system (6.4)-(6.5) are required.

Clearly, the initial state X_0 is a zero-mean random vector with covariance matrix given by

$$P_0^* = \begin{pmatrix} P_0 & P_0^{(3)} \\ P_0^{(3)} & P_0^{(4)} \end{pmatrix}.$$

Moreover, it is easy to show that X_0 and $\{W_k\}_{k\geq 0}$, $\{V_k\}_{k\geq 1}$, $\{\mathcal{F}_k\}_{k\geq 0}$, $\{\mathcal{H}_k\}_{k\geq 1}$ are uncorrelated.

Next, the second-order statistical properties of the noise processes $\{W_k\}_{k\geq 0}$ and $\{V_k\}_{k\geq 1}$ are established in propositions 6.3.1 and 6.3.2, respectively.

Proposition 6.3.1 *The noise* $\{W_k\}_{k\geq 0}$ *is a zero-mean white process with covariance matrix,* $E[W_k W_k^T] = Q_k^W$*, given by*

$$Q_k^W = \begin{pmatrix} Q_k & Q_k^{12} \\ Q_k^{12T} & Q_k^{22} \end{pmatrix} + E[\widetilde{\mathcal{F}}_k \overline{\mathcal{X}}_k \overline{\mathcal{X}}_k^T \widetilde{\mathcal{F}}_k^T],$$

where

$$\begin{aligned} Q_k^{12} &= ((\overline{F}_k \overline{x}_k)^T \otimes Q_k)(I+K) + Q_k^{(3)}, \\ Q_k^{22} &= (I+K)\left(\left(\overline{F}_k D_k \overline{F}_k^T + E[\widetilde{F}_k D_k \widetilde{F}_k^T]\right) \otimes Q_k\right)(I+K) + Q_k^{(4)} \\ &\quad + (I+K)\left((\overline{F}_k \overline{x}_k) \otimes Q_k^{(3)}\right) + \left((\overline{F}_k \overline{x}_k) \otimes Q_k^{(3)}\right)^T (I+K) \end{aligned}$$

and

$$E[\widetilde{\mathcal{F}}_k \overline{\mathcal{X}}_k \overline{\mathcal{X}}_k^T \widetilde{\mathcal{F}}_k^T] = \begin{pmatrix} E[\widetilde{F}_k \overline{x}_k \overline{x}_k^T \widetilde{F}_k^T] & E[\widetilde{F}_k \overline{x}_k vec(D_k)^T \widetilde{F}_k^{[2]T}] \\ E[\widetilde{F}_k^{[2]} vec(D_k) \overline{x}_k^T \widetilde{F}_k^T] & E[\widetilde{F}_k^{[2]} vec(D_k) vec(D_k)^T \widetilde{F}_k^{[2]T}] \end{pmatrix},$$

whose blocks are calculated as in (6.3)*.*

Proof. Clearly, $\forall k \geq 0$, $E[W_k] = 0$. Now, taking into account the mutual independence between $\{w_k\}_{k\geq 0}$, $\{F_k\}_{k\geq 0}$ and the initial state x_0, it is easy to prove that $E[\widetilde{\mathcal{W}}_k \overline{\mathcal{X}}_s^T \widetilde{\mathcal{F}}_s^T] = 0$, $E[\widetilde{\mathcal{F}}_k \overline{\mathcal{X}}_k \overline{\mathcal{X}}_s^T \widetilde{\mathcal{F}}_s^T] = E[\widetilde{\mathcal{F}}_k \overline{\mathcal{X}}_k \overline{\mathcal{X}}_k^T \widetilde{\mathcal{F}}_k^T]\delta_{k,s}$, $\forall k, s \geq 0$, and, hence

$$E[W_k W_s^T] = E[\widetilde{\mathcal{W}}_k \widetilde{\mathcal{W}}_s^T] + E[\widetilde{\mathcal{F}}_k \overline{\mathcal{X}}_k \overline{\mathcal{X}}_k^T \widetilde{\mathcal{F}}_k^T]\delta_{k,s}.$$

Then, we only need to prove that $\forall k, s \geq 0$,

$$E[\widetilde{\mathcal{W}}_k \widetilde{\mathcal{W}}_s^T] := \begin{pmatrix} Q_{k,s}^{11} & Q_{k,s}^{12} \\ Q_{k,s}^{12T} & Q_{k,s}^{22} \end{pmatrix} = \begin{pmatrix} Q_k & Q_k^{12} \\ Q_k^{12T} & Q_k^{22} \end{pmatrix} \delta_{k,s}.$$

- Since $\{w_k\}_{k\geq 0}$ is a zero-mean white sequence with covariances $Q_k, \forall k \geq 0$, it is immediately clear that $Q_{k,s}^{11} = E[w_k w_s^T] = Q_k \delta_{k,s}$.
- Using the Kronecker product properties, Assumption (A3) and since $E[F_k x_k] = \overline{F}_k \overline{x}_k$, it is easy to obtain that $Q_{k,s}^{12} = E[w_k \Phi_s^T] = Q_k^{12} \delta_{k,s}$.

- From the conditional expectation properties we have $E\left[F_k x_k x_k^T F_k^T\right] = \overline{F}_k D_k \overline{F}_k^T + E[\widetilde{F}_k D_k \widetilde{F}_k^T]$, then, using again Assumption (A3) and the Kronecker product properties, we get $Q_{k,s}^{22} = E[\Phi_k^T \Phi_s^T] = Q_k^{22}\delta_{k,s}$. □

Proposition 6.3.2 *The noise $\{V_k\}_{k\geq 1}$ is a zero-mean white process with covariance matrix, $E[V_k V_k^T] = R_k^V$, given by*

$$R_k^V = \begin{pmatrix} R_k & R_k^{12} \\ R_k^{12T} & R_k^{22} \end{pmatrix} + E[\widetilde{\mathcal{H}}_k \overline{\mathcal{X}}_k \overline{\mathcal{X}}_k^T \widetilde{\mathcal{H}}_k^T],$$

where

$$\begin{aligned} R_k^{12} &= \left((\overline{H}_k \overline{x}_k)^T \otimes R_k\right)(I+K) + R_k^{(3)}, \\ R_k^{22} &= (I+K)\left(\left(\overline{H}_k D_k \overline{H}_k^T + E[\widetilde{H}_k D_k \widetilde{H}_k^T]\right) \otimes R_k\right)(I+K) + R_k^{(4)} \\ &\quad + (I+K)\left((\overline{H}_k \overline{x}_k) \otimes R_k^{(3)}\right) + \left((\overline{H}_k \overline{x}_k) \otimes R_k^{(3)}\right)^T (I+K) \end{aligned}$$

and

$$E[\widetilde{\mathcal{H}}_k \overline{\mathcal{X}}_k \overline{\mathcal{X}}_k^T \widetilde{\mathcal{H}}_k^T] = \begin{pmatrix} E[\widetilde{H}_k \overline{x}_k \overline{x}_k^T \widetilde{H}_k^T] & E[\widetilde{H}_k \overline{x}_k vec(D_k)^T \widetilde{H}_k^{[2]T}] \\ E[\widetilde{H}_k^{[2]} vec(D_k) \overline{x}_k^T \widetilde{H}_k^T] & E[\widetilde{H}_k^{[2]} vec(D_k) vec(D_k)^T \widetilde{H}_k^{[2]T}] \end{pmatrix},$$

whose blocks are calculated as in (6.3).

Proof. This proof is analogous to that of Proposition 6.3.1 and, hence, it is omitted.

□

Remark 2. From the augmented state equation (6.4) and Proposition 6.3.1, the following recursive equation for the matrix $\mathcal{D}_k = E[X_k X_k^T]$ holds:

$$\mathcal{D}_k = \overline{\mathcal{F}}_{k-1} \mathcal{D}_{k-1} \overline{\mathcal{F}}_{k-1}^T + E[\widetilde{\mathcal{F}}_{k-1} \mathcal{D}_{k-1} \widetilde{\mathcal{F}}_{k-1}^T] + Q_{k-1}^W, \quad k \geq 1; \quad \mathcal{D}_0 = P_0^*, \tag{6.6}$$

where

$$E[\widetilde{\mathcal{F}}_k \mathcal{D}_k \widetilde{\mathcal{F}}_k^T] = \begin{pmatrix} E[\widetilde{F}_k D_k \widetilde{F}_k^T] & E[\widetilde{F}_k D_k^{(3)} \widetilde{F}_k^{[2]T}] \\ E[\widetilde{F}_k^{[2]} D_k^{(3)T} \widetilde{F}_k^T] & E[\widetilde{F}_k^{[2]} D_k^{(4)} \widetilde{F}_k^{[2]T}] \end{pmatrix},$$

with $D_k^{(3)} = E[x_k x_k^{[2]T}]$ and $D_k^{(4)} = E[x_k^{[2]} x_k^{[2]T}]$ the blocks of the matrix $\mathcal{D}_k$.

6.4 LS quadratic estimator

To address the LS linear estimation problem of X_k based on $Y_1, \ldots, Y_k$, an innovation approach is used. Since the measurements are non-orthogonal vectors, this procedure consists of transforming the measurement process $\{Y_k;\, k \geq 1\}$ into an equivalent one of orthogonal vectors $\{\nu_k;\, k \geq 1\}$ called *innovations*. The innovation at time k is defined as $\nu_k = Y_k - \widehat{Y}_{k/k-1}$, where $\widehat{Y}_{k/k-1}$ is the one-stage linear predictor of Y_k. Therefore, the LS linear filter of the augmented state, $\widehat{X}_{k/k}$, can be calculated as a linear combination of the innovations, as follows:

$$\widehat{X}_{k/k} = \sum_{i=1}^{k} E[X_k \nu_i^T] \Pi_i^{-1} \nu_i, \quad k \geq 1, \tag{6.7}$$

where $\Pi_i = E[\nu_i \nu_i^T]$.

Next, a recursive algorithm for the optimal LS linear filter of the augmented state is derived.

Theorem 6.4.1 *The linear filter of the augmented state is recursively obtained by the following relation*

$$\widehat{X}_{k/k} = \widehat{X}_{k/k-1} + \mathcal{G}_k \Pi_k^{-1} \nu_k, \quad k \geq 1; \quad \widehat{X}_{0/0} = 0, \tag{6.8}$$

where the state predictor, $\widehat{X}_{k/k-1}$, is calculated by

$$\widehat{X}_{k/k-1} = \overline{\mathcal{F}}_{k-1} \widehat{X}_{k-1/k-1}, \quad k \geq 1. \tag{6.9}$$

The innovation, ν_k, satisfies

$$\nu_k = Y_k - \overline{\mathcal{H}}_k \widehat{X}_{k/k-1}, \quad k \geq 2; \quad \nu_1 = Y_1, \tag{6.10}$$

The matrix, $\mathcal{G}_k = E[X_k \nu_k^T]$ is determined by

$$\mathcal{G}_k = \Sigma_{k/k-1} \overline{\mathcal{H}}_k^T, \quad k \geq 1, \tag{6.11}$$

where the prediction error covariance matrix, $\Sigma_{k/k-1}$, is obtained by

$$\Sigma_{k/k-1} = \overline{\mathcal{F}}_{k-1}\Sigma_{k-1/k-1}\overline{\mathcal{F}}_{k-1}^T + E[\widetilde{\mathcal{F}}_{k-1}\mathcal{D}_{k-1}\widetilde{\mathcal{F}}_{k-1}^T] + Q_{k-1}^W, \quad k \geq 1, \tag{6.12}$$

with $\mathcal{D}_k$ given in (6.6) *and $\Sigma_{k/k}$, the filtering error covariance matrix, calculated by*

$$\Sigma_{k/k} = \Sigma_{k/k-1} - \mathcal{G}_k\Pi_k^{-1}\mathcal{G}_k^T, \quad k \geq 1; \quad \Sigma_{0/0} = P_0^*. \tag{6.13}$$

The innovation covariance matrix, Π_k, satisfies

$$\Pi_k = E[\widetilde{\mathcal{H}}_k\mathcal{D}_k\widetilde{\mathcal{H}}_k^T] + \overline{\mathcal{H}}_k\mathcal{G}_k + R_k^V, \quad k \geq 1, \tag{6.14}$$

where

$$E[\widetilde{\mathcal{H}}_k\mathcal{D}_k\widetilde{\mathcal{H}}_k^T] = \begin{pmatrix} E[\widetilde{H}_k D_k \widetilde{H}_k^T] & E[\widetilde{H}_k D_k^{(3)} \widetilde{H}_k^{[2]T}] \\ E[\widetilde{H}_k^{[2]} D_k^{(3)T} \widetilde{H}_k^T] & E[\widetilde{H}_k^{[2]} D_k^{(4)} \widetilde{H}_k^{[2]T}] \end{pmatrix},$$

whose blocks are calculated as in (6.3).

Proof. From expression (6.7), relation (6.8) for the state filter, $\widehat{X}_{k/k}$, in terms of the one-stage predictor, $\widehat{X}_{k/k-1}$, is directly derived.

Expression (6.9) for the state predictor is immediately obtained from (6.4) and the Orthogonal Projection Lemma (OPL).

Obtaining an explicit formula for the innovation, $\nu_k = Y_k - \widehat{Y}_{k/k-1}$, is equivalent to calculate $\widehat{Y}_{k/k-1}$, which, from (6.5) and using again the OPL, can be expressed as

$$\widehat{Y}_{k/k-1} = \overline{\mathcal{H}}_k\widehat{X}_{k/k-1}, \ k \geq 1. \tag{6.15}$$

Next, identity (6.11) for $\mathcal{G}_k$ is deduced. Applying the OPL, it is clear that

$$E[X_k\widehat{X}_{k/k-1}^T] = E[\widehat{X}_{k/k-1}\widehat{X}_{k/k-1}^T] = \mathcal{D}_k - \Sigma_{k/k-1}, \quad k \geq 1,$$

therefore,

$$\mathcal{G}_k = E[X_k\nu_k^T] = E[X_k(X_k - \widehat{X}_{k/k-1})^T]\mathcal{H}_k^T = \Sigma_{k/k-1}\overline{\mathcal{H}}_k^T, \quad k \geq 1.$$

Since $\Sigma_{k/k-1} = E[X_k X_k^T] - E[\widehat{X}_{k/k-1}\widehat{X}_{k/k-1}^T]$, using (6.6) for $E[X_k X_k^T]$ and (6.9) for $\widehat{X}_{k/k-1}$, expression (6.12) is easily deduced, taking into account that $E[\widehat{X}_{k-1/k-1}\widehat{X}_{k-1/k-1}^T] = \mathcal{D}_{k-1} - \Sigma_{k-1/k-1}$, $k \geq 1$.

Similarly, $\Sigma_{k/k} = E[X_k X_k^T] - E[\widehat{X}_{k/k}\widehat{X}_{k/k}^T]$ and, therefore, by using (6.8) for $\widehat{X}_{k/k}$, formula (6.13) is obtained.

Finally, we prove expression (6.14) for the innovation covariance matrix $\Pi_k = E[Y_k Y_k^T] - E[\widehat{Y}_{k/k-1}\widehat{Y}_{k/k-1}^T]$. On the one hand, from (6.5),we have

$$E[Y_k Y_k^T] = E[\mathcal{H}_k X_k X_k^T \mathcal{H}_k^T] + R_k^V$$

where, by considering the conditional expectation properties, it is satisfied that $E[\mathcal{H}_k X_k X_k^T \mathcal{H}_k^T] = \overline{\mathcal{H}}_k \mathcal{D}_k \overline{\mathcal{H}}_k^T + E[\widetilde{\mathcal{H}}_k \mathcal{D}_k \widetilde{\mathcal{H}}_k^T]$. On the other hand, using (6.15) and the OPL, it is deduced that $E[\widehat{Y}_{k/k-1}\widehat{Y}_{k/k-1}^T] = \overline{\mathcal{H}}_k(\mathcal{D}_k - \Sigma_{k/k-1})\overline{\mathcal{H}}_k^T$, $k \geq 1$. Then, the innovation covariance (6.14) is proved. □

Remark 3. As mentioned in Section 6.3, the LS quadratic filter of the original state x_k, is obtained by adding the mean $\overline{x}_k$ to the vector constituted by the first n entries of $\widehat{X}_{k/k}$. Specifically, the quadratic filter, $\widehat{x}_{k/k}^q$, is given by

$$\widehat{x}_{k/k}^q = \Upsilon \widehat{X}_{k/k} + \overline{x}_k, \quad k \geq 1,$$

where Υ is the operator which extracts the first n entries of $\widehat{X}_{k/k}$.

6.5 Application in multi-sensor systems with fading measurements

In this section, the optimal LS quadratic filter obtained in Section 6.4 is applied to linear discrete-time stochastic systems with fading measurements coming from multiple sensors. The phenomenon of measurement fading occurs in a random way and it is described by different sequences of scalar random variables with a certain probability distribution over the interval $[0, 1]$. Moreover, a simulation example is given to illustrate the effectiveness of the proposed recursive filtering algorithm.

6.5.1 Multi-sensor model and filtering algorithm

Consider the state equation given in (6.1) satisfying assumptions (A1)-(A3), and r sensors whose measurements of the state are described by the following observation equations:

$$y_k^i = \theta_k^i C_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2, \ldots, r, \tag{6.16}$$

where $y_k^i \in \mathbb{R}$, is the measured output provided by sensor i at the sampling time k, $\{C_k^i\}_{k\geq 1}$, are random parameter matrices with compatible dimensions, $\{v_k^i\}_{k\geq 1}$ are the measurement noises, and $\{\theta_k^i\}_{k\geq 1}$ are scalar random variables which model the fading phenomenon of the i-th sensor. In order to apply Theorem 6.4.1, the following assumptions of the noise processes and the random parameter matrices are considered:

(*i*) For $i = 1, 2, \ldots, r$, the sensor additive noises, $\{v_k^i\}_{k\geq 1}$, are zero-mean white processes. By denoting $v_k = (v_k^1, \ldots, v_k^r)^T$, it is supposed that its moments, up to the fourth-order one, are known.

(*ii*) For $i = 1, 2, \ldots, r$, $\{C_k^i\}_{k\geq 1}$ are white sequences of random parameter matrices. By denoting $C_k = \left[C_k^{1T} \mid \cdots \mid C_k^{rT}\right]^T$, its mean, $\overline{C}_k$, is known and the covariances and cross-covariances between the entries of the matrices C_k and $C_k^{[2]}$, are also assumed to be known.

(*iii*) For $i = 1, 2, \ldots, r$, the noises $\{\theta_k^i\}_{k\geq 1}$ are white sequences of scalar random variables over the interval $[0, 1]$. By denoting $\theta_k = (\theta_k^1, \ldots, \theta_k^r)^T$, it is supposed that its moments up to the fourth one are known. We will denote

$$K_k^{\theta} = Cov[\theta_k], \quad K_k^{\theta(3)} = Cov[\theta_k, \theta_k^{[2]}], \quad K_k^{\theta(4)} = Cov[\theta_k^{[2]}].$$

(*iv*) x_0, $\{F_k\}_{k\geq 0}$, $\{\theta_k\}_{k\geq 1}$, $\{C_k\}_{k\geq 1}$, $\{w_k\}_{k\geq 0}$ and $\{v_k\}_{k\geq 1}$ are mutually independent.

The observation model (6.16) can be rewritten in a compact form as follows:

$$y_k = \Theta_k C_k x_k + v_k, \quad k \geq 1,$$

where $y_k = (y_k^1, \ldots, y_k^r)^T$ is the measurement vector and $\Theta_k = Diag(\theta_k^1, \ldots, \theta_k^r)$. Accordingly, this observation model is a particular case of (6.2) with $H_k = \Theta_k C_k$, and clearly verifies the assumptions given in Section 6.2.

The corresponding augmented measurement equation is given by

$$Y_k = \mathcal{T}_k \mathcal{C}_k X_k + V_k, \quad k \geq 1,$$

where $\mathcal{T}_k = Diag(\Theta_k, \Theta_k^{[2]})$, $\mathcal{C}_k = Diag(C_k, C_k^{[2]})$ and $V_k = (\mathcal{T}_k \widetilde{\mathcal{C}}_k + \widetilde{\mathcal{T}}_k \overline{\mathcal{C}}_k)\overline{\mathcal{X}}_k + \widetilde{\mathcal{V}}_k$. This measurement equation is a particular case of (6.5) with $\mathcal{H}_k = \mathcal{T}_k \mathcal{C}_k$, and it is immediately clear that $\widetilde{\mathcal{H}}_k = \mathcal{T}_k \widetilde{\mathcal{C}}_k + \widetilde{\mathcal{T}}_k \overline{\mathcal{C}}_k$.

By applying the Hadamard product properties, for any deterministic matrix $\mathcal{A} \in \mathbb{R}^{(n+n^2)\times(n+n^2)}$, it is easy to see that

$$\begin{aligned} &E[(\mathcal{T}_k \widetilde{\mathcal{C}}_k + \widetilde{\mathcal{T}}_k \overline{\mathcal{C}}_k)\mathcal{A}(\mathcal{T}_k \widetilde{\mathcal{C}}_k + \widetilde{\mathcal{T}}_k \overline{\mathcal{C}}_k)^T] \\ &\qquad = E[J_k^\theta J_k^{\theta T}] \circ E[\widetilde{\mathcal{C}}_k \mathcal{A} \widetilde{\mathcal{C}}_k^T] + Cov[J_k^\theta] \circ (\overline{\mathcal{C}}_k \mathcal{A} \overline{\mathcal{C}}_k^T), \quad k \geq 1, \end{aligned}$$

where $J_k^\theta = \left(\theta_k^T, \theta_k^{[2]T}\right)^T$ and $Cov[J_k^\theta] = \begin{pmatrix} K_k^\theta & K_k^{\theta(3)} \\ K_k^{\theta(3)T} & K_k^{\theta(4)} \end{pmatrix}$.

Hence, taking into account this property and Proposition 6.3.2, we obtain that the covariance matrix of the noise process $\{V_k\}_{k\geq 1}$ is given by

$$R_k^V = \begin{pmatrix} R_k & R_k^{12} \\ R_k^{12T} & R_k^{22} \end{pmatrix} + E[J_k^\theta J_k^{\theta T}] \circ E[\widetilde{\mathcal{C}}_k \overline{\mathcal{X}}_k \overline{\mathcal{X}}_k^T \widetilde{\mathcal{C}}_k^T] + Cov[J_k^\theta] \circ (\overline{\mathcal{C}}_k \overline{\mathcal{X}}_k \overline{\mathcal{X}}_k^T \overline{\mathcal{C}}_k^T)$$

where

$$\begin{aligned} R_k^{12} &= \left((\overline{\Theta}_k \overline{C}_k \overline{x}_k)^T \otimes R_k\right)(I + K) + R_k^{(3)}, \\ R_k^{22} &= (I + K)\left(\left(E[\theta_k \theta_k^T] \circ \left(\overline{C}_k D_k \overline{C}_k^T + E[\widetilde{C}_k D_k \widetilde{C}_k^T]\right)\right) \otimes R_k\right)(I + K) + R_k^{(4)} \\ &\quad + (I + K)\left((\overline{\Theta}_k \overline{C}_k \overline{x}_k) \otimes R_k^{(3)}\right) + \left((\overline{\Theta}_k \overline{C}_k \overline{x}_k) \otimes R_k^{(3)}\right)^T (I + K). \end{aligned}$$

Thus, starting from the linear filter $\widehat{X}_{k/k}$ given by (6.8) with $\widehat{X}_{k/k-1}$ the state predictor determined by (6.9), a recursive optimal linear filtering algorithm is obtained by calculating the innovation ν_k, its covariance matrix Π_k, and the matrix $\mathcal{G}_k$ as follows:

$$\begin{aligned}
&\nu_k = Y_k - \overline{\mathcal{T}}_k \overline{\mathcal{C}}_k \widehat{X}_{k/k-1}, \quad k \geq 2; \quad \nu_1 = Y_1, \\
&\mathcal{G}_k = \Sigma_{k/k-1} \overline{\mathcal{C}}_k^T \overline{\mathcal{T}}_k, \quad k \geq 1, \\
&\Pi_k = E[J_k^\theta J_k^{\theta T}] \circ E[\widetilde{\mathcal{C}}_k \mathcal{D}_k \widetilde{\mathcal{C}}_k^T] + Cov[J_k^\theta] \circ (\overline{\mathcal{C}}_k \mathcal{D}_k \overline{\mathcal{C}}_k^T) + \overline{\mathcal{T}}_k \overline{\mathcal{C}}_k \mathcal{G}_k + R_k^V, \quad k \geq 1,
\end{aligned}$$

with $\mathcal{D}_k$ and $\Sigma_{k/k-1}$ given in (6.6) and (6.12), respectively, and

$$E[\widetilde{\mathcal{C}}_k \mathcal{D}_k \widetilde{\mathcal{C}}_k^T] = \begin{pmatrix} E[\widetilde{C}_k D_k \widetilde{C}_k^T] & E[\widetilde{C}_k D_k^{(3)} \widetilde{C}_k^{[2]T}] \\ E[\widetilde{C}_k^{[2]} D_k^{(3)T} \widetilde{C}_k^T] & E[\widetilde{C}_k^{[2]} D_k^{(4)} \widetilde{C}_k^{[2]T}] \end{pmatrix},$$

whose blocks are calculated as in (6.3).

As mentioned in the previous sections, the quadratic filter of the original state is formed by the first n entries of $\widehat{X}_{k/k}$ plus the mean $\overline{x}_k$.

6.5.2 Numerical simulation example

Consider the following uncertain discrete-time system with fading measurements coming from two sensors:

$$\begin{aligned}
x_k &= (0.95 + 0.1\epsilon_{k-1}) x_{k-1} + w_{k-1}, \quad k \geq 1, \\
y_k^i &= \theta_k^i C_k^i x_k + v_k^i, \quad k \geq 1, \quad i = 1, 2,
\end{aligned}$$

where $\{\epsilon_k\}_{k\geq 0}$ is a zero-mean Gaussian white process with unit variance. $\{\theta_k^i\}_{k\geq 1}$, $i = 1, 2$, are independent sequences of discrete-time random variables with the following probability distributions over the interval $[0, 1]$:

- In the first sensor, $\{\theta_k^1\}_{k\geq 1}$ is a sequence of independent and identically distributed (i.i.d.) Bernoulli variables with $P[\theta_k^1 = 1] = p, \ \forall k \geq 1$.

- In the second sensor, $\{\theta_k^2\}_{k\geq 1}$ is a sequence of i.i.d. random variables with $P[\theta_k^2 = 0] = 0.2, \quad P[\theta_k^2 = 0.5] = 0.6, \quad P[\theta_k^2 = 1] = 0.2, \ \forall k \geq 1.$

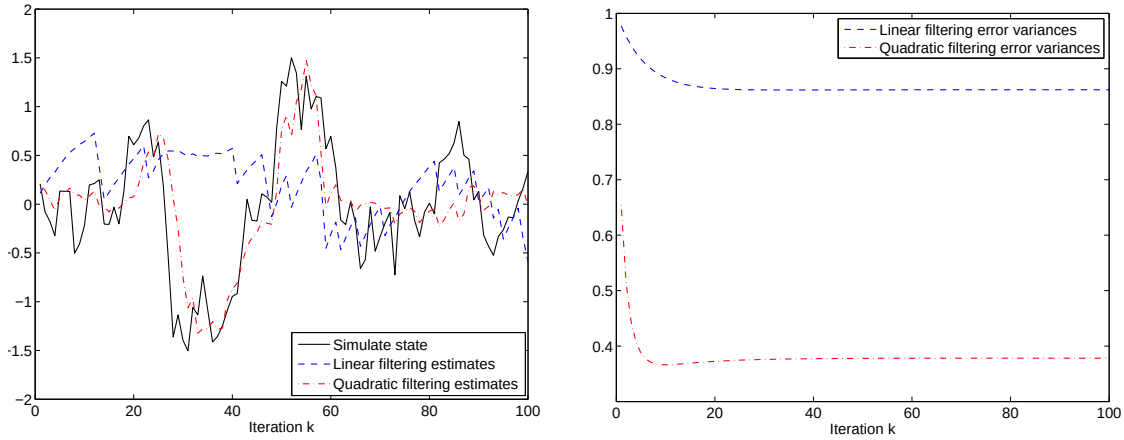
The matrices C_k^i, $i = 1, 2$, are defined as $C_k^1 = 0.5 + 0.4\zeta_k^1$ and $C_k^2 = 0.6 + 0.4\zeta_k^2$, where $\{\zeta_k^i\}_{k\geq 1}$, $i = 1, 2$, are independent zero-mean Gaussian white processes with unit variance. The initial state x_0 is a zero-mean Gaussian variable with $P_0 = 1$. The noise $\{w_k\}_{k\geq 0}$ is a zero-mean Gaussian white process with variance $Q_k = 0.1$, for all k, and $\{v_k^i\}_{k\geq 1}$, $i = 1, 2$, are independent zero-mean white processes with the following probability distributions:

$$P[v_k^1 = -8] = \frac{1}{8}, \quad P[v_k^1 = \frac{8}{7}] = \frac{7}{8}, \quad \forall k \geq 1,$$
$$P[v_k^2 = 1] = \frac{15}{18}, \quad P[v_k^2 = -3] = \frac{2}{18}, \quad P[v_k^2 = -9] = \frac{1}{18}, \quad \forall k \geq 1.$$

To analyze the performance of the proposed quadratic estimator, we ran a program in MATLAB, in which one hundred iterations of the linear filtering algorithm ([21]) and the proposed quadratic filtering algorithm have been carried out, considering different values of p. Linear and quadratic filters of the state are calculated, as well as the corresponding error variances, which provide a measure of the estimation accuracy.

Firstly, for $p = 0.9$, the performance of the linear and quadratic filtering estimators has been compared on the basis of both the estimates obtained from the corresponding simulated observations of the state, and the filtering error variances. The results are displayed in Figure 6.1. From Figure 6.1a, it is deduced that the quadratic filter follows the state evolution better than the linear one, fact which is confirmed in Figure 6.1b, where it is observed that the error variances of the quadratic filter are significantly less than those of the linear filter and, consequently, the quadratic filter outperforms the linear one.

Next, taking into account that the filtering error variances have insignificant variation from a certain iteration onwards, Figure 6.2 shows the linear and quadratic



(a) (b)

Figure 6.1: (a) Simulated state and linear and quadratic filtering estimates. (b) Linear and quadratic filtering error variances.

filtering error variances at a fixed iteration (namely, $k = 100$), when the probability p that the state x_k is present in the measurements of the first sensor is varied from 0.1 to 0.9. From this figure it is gathered that the filtering error variances become smaller, and hence better estimations are obtained, as p increases. Also, for all the different values of p, the quadratic filtering error variances decrease more quickly than those of the linear filter and their values are smaller, thus confirming again that the quadratic filtering estimators outperform the linear ones.

Finally, a comparative analysis is presented between the proposed filter and other quadratic filters in the existing literature; specifically, the proposed filter is compared with the quadratic filter for systems with non-random parameter matrices ([1]) and the quadratic filter for multi-sensor systems with uncertain observations ([9]). For this purpose, the filtering mean-square error (MSE) at each time instant k is calculated from one thousand independent simulations of the mentioned algorithms considering the probability $p = 0.9$. The MSE criterion consists

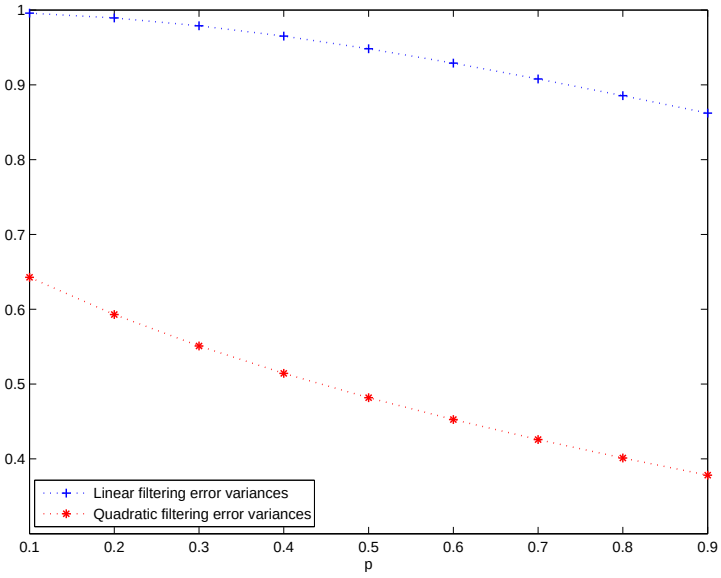


Figure 6.2: Linear and quadratic filtering error variances versus p.

in computing $\text{MSE}_k = \dfrac{1}{1000}\displaystyle\sum_{s=1}^{1000}(x_k^{(s)} - \widehat{x}_{k/k}^{(s)})^2$, where $\{x_k^{(s)}\}_{1\leq k\leq 100}$ denote the s-th set of artificially simulated data and $\widehat{x}_{k/k}^{(s)}$ is the filter at the sampling time k in the s-th simulation run. The results of this comparison are displayed in Figure 6.3, which shows that, for all k, the proposed filter performs better than the other two. The worst estimations are provided by the filtering algorithm in [1] since randomness is ignored in both state transition and measurement matrices. The filter in [9] performs better, since it takes into account the multiplicative noise in the measurement equation, but it is also improved significantly by the quadratic filter proposed in this paper.

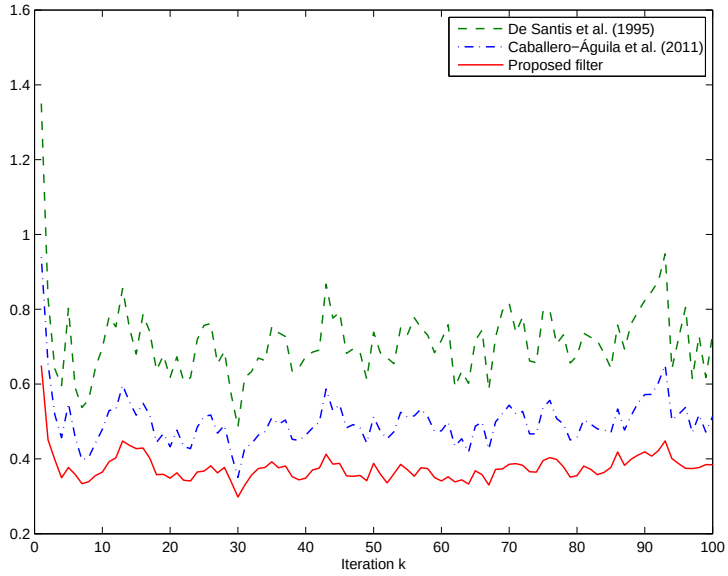


Figure 6.3: Comparison of MSE_k for three filters.

6.6 Conclusions

The LS quadratic estimation problem has been investigated for discrete-time linear stochastic systems with random parameter matrices. The main contributions are summarized as follows:

- Using the technique proposed by [1], consisting of augmenting the state and observation vectors with their second-order Kronecker powers, an augmented system with random parameter matrices has been constructed. A recursive algorithm for the linear LS filter of the augmented state based on the augmented measurements has been obtained by an innovation approach, and the quadratic LS filter of the original state is derived from the linear LS filter of the augmented state.

- The proposed quadratic filter has been applied to systems with fading measurements coming from multiple sensors, when the fading measurement phe-

nomenon in each sensor is described by different sequences of scalar random variables with arbitrary discrete probability distribution over the interval [0,1]. This kind of multi-sensor systems is found in various real-world situations, such as transmission models involving partial loss of measurements.

- The usefulness of the proposed results has been illustrated by a numerical simulation example. Error variance comparison has shown that the quadratic filters outperform the lineal ones. Furthermore, a comparative analysis of the proposed filter and other quadratic filters in the existing literature has been carried out. This example has also highlighted the applicability of the proposed algorithm in multi-sensor systems with state-dependent multiplicative noise and fading measurements, which can be addressed by the system model with random parameter matrices considered in this paper.

Since the noise independence assumption can be restrictive in many real-world problems, a challenging further research topic is to consider systems featuring autocorrelation and cross-correlation between the process noise and the measurement noises. Also, an interesting future research topic is to generalize the current results by considering correlation between random state transition and measurement matrices, which would cover systems with randomly delayed measurements or multiple packet dropouts as a particular case.

Acknowledgements

This research is supported by Ministerio de Economía y Competitividad (Grant No. MTM2014-52291-P and FPU programme).

References

[1] A. De Santis, A. Germani and M. Raimondi, Optimal quadratic filtering of

linear discrete-time non-Gaussian systems, *IEEE Trans. Automatic Control,* 40 (1995), 1274–1278.

[2] F. Carravetta, A. Germani and M. Raimondi, Polynomial filtering for linear discrete time non-Gaussian systems, *SIAM J. Control Optim.,* 34 (1996), 1666–1690.

[3] O.L.V. Costa and G.R.A.M Benites, Linear minimum mean square filter for discrete-time linear systems with Markov jumps and multiplicative noises, *Automatica,* 47 (2011), 466–476.

[4] J. Feng, Z. Wang and M. Zeng, Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises, *Inf. Fusion,* 14 (2013), 78–86.

[5] F. Peng and S. Sun, Distributed fusion estimation for multisensor multirate systems with stochastic observation multiplicative noises, *Math. Probl. Eng.,* vol. 2014, Article ID 373270, 8 pages.

[6] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo andJ. Linares-Pérez, Second-order polynomial estimators from uncertain observations using covariance information, *Appl. Math. Comput.,* 143 (2003), 319–338.

[7] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo, J.D. Jiménez-López and J. Linares-Pérez, Least-squares th-order polynomial estimation of signals from observations affected by non-independent uncertainty, *Appl. Math. Comput.,* 176 (2006), 642–653.

[8] F.O. Hounkpevi and E.E. Yaz, Robust minimum variance linear state estimators for multiple sensors with different failure rates, *Automatica,* 43 (2007), 1274–1280.

[9] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Linear and quadratic estimation using uncertain observations from multiple sensors with correlated uncertainty, *Signal Process.,* 91 (2011), 330–337.

[10] X. Qu and J. Zhou, The optimal robust finite-horizon Kalman filtering for multiple sensors with different stochastic failure rates, *Appl. Math. Lett.,* 26 (2013), 80–86.

[11] I. García-Garrido, J. Linares-Pérez and R. Caballero-Águila, Centralized fusion quadratic estimators in multi-sensor systems with correlated missing measurements, *Appl. Math. Sci.,* 57 (2013), 2795–2813.

[12] R. Caballero-Águila, I. García-Garrido and J. Linares-Pérez, Information fusion algorithms for state estimation in multi-sensor systems with correlated missing measurements, *Appl. Math. Comput.,* 226 (2014), 548–563.

[13] J. Hu, Z. Wang and H. Gao, Recursive filtering with random parameter matrices, multiple fading measurements and correlated noises, *Automatica,* 49 (2013), 3440–3448.

[14] Y. Liu, X. He, Z. Wang and D. Zhou, Optimal filtering for networked systems with stochastic sensor gain degradation, *Automatica,* 50 (2014), 1521–1525.

[15] S. Sun and J. Ma, Linear estimation for networked control systems with random transmission delays and packet dropouts, *Inf. Sci.,* 269 (2014), 349–365.

[16] W.L. De Koning, Optimal estimation of linear discrete-time systems with stochastic parameters, *Automatica,* 20 (1984), 113–115.

[17] Y. Luo, Y. Zhu, D. Luo, J. Zhou, E. Song and D. Wang, Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications, *Sensors,* 8 (2008), 8086–8103.

[18] Y.T. Luo, Y.M. Zhu, X.J. Shen and E.B. Song, Novel data association algorithm based on integrated random coefficient matrices Kalman filtering, *IEEE Trans. Aerosp. Electron. Syst.,* 48 (2012), 144–158.

[19] X.J. Shen, Y.T. Luo, Y.M. Zhu and E.B. Song, Globally optimal distributed Kalman filtering fusion, *Sci. China Inf. Sci.,* 55 (2012), 512–529.

[20] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Covariance-based estimation from multisensor delayed measurements with random parameter matrices and correlated noises, *Math. Probl. Eng.,* vol. 2014, Article ID 958474, 13 pages.

[21] J. Linares-Pérez, R. Caballero-Águila and I. García-Garrido, Optimal linear filter design for systems with correlation in the measurement matrices and noises: recursive algorithm and applications, *Int. J. Syst. Sci.,* 45 (2014), 1548–1562.

[22] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Optimal state estimation for networked systems with random parameter matrices, correlated noises and delayed measurements, *Int. J. Gen. Syst.,* 44 (2015), 142–154.

[23] F. Cacace, A. Fasano and A. Germani, Quadratic filtering of non-Gaussian linear systems with random observation matrices, In *the 19th World Congr. The Int. Fed. Automatic Control,* (2014), 2370–2375.

[24] J.R. Magnus, H. Neudecker, Matrix differential calculus with applications in statistics and econometrics, *second ed., Wiley & Sons, New York,* 1999.

Results and conclusions

This PhD thesis considers the problem of LS estimation for discrete-time linear stochastic systems that contain noisy measurements derived from multiple sensors and are affected by random parameters modeling different kinds of failure in the measurement data. As detailed in the methodology section, information fusion methods are used to process the measurements from the different sensors, and recursive estimation algorithms are derived under the LS optimality criterion by means of an innovation approach. Specifically, centralized and distributed fusion methods are used to obtain linear and quadratic filtering and fixed-point estimators, together with the corresponding error covariance matrices which measure the accuracy of these estimators. The main results and conclusions are detailed below.

For *sensor network systems with missing measurements*, we first assume that at each sensor, the possibility of missing measurements (that is, of observations containing no information about the state but only noise) is modeled by binary variables taking the values one or zero (Bernoulli variables), depending on whether the state is present or missing in the corresponding observation. Such variables are assumed to be correlated at instants that differ by m units of time. For this kind of systems, using both centralized and distributed fusion methods to process the measurement data from the different sensors, recursive linear filtering and fixed-point smoothing algorithms are derived in **Chapter 1** by an innovation approach, and quadratic filtering and fixed-point smoothing estimators are obtained in **Chapter 2** using the centralized fusion method.

In Chapter 1, two numerical simulation examples are given to illustrate the applicability of the results obtained to estimate a scalar state process generated by an AR model and a two-dimensional state, respectively, from uncertain observations derived from two sensors featuring correlation in the uncertainty. These results confirm that centralized and distributed fusion estimators have approximately the same accuracy. For different uncertainty probabilities and different values of m, both examples confirm the greater effectiveness of the fixed-point smoothing estimators in contrast to the filtering ones and thus we conclude that more accurate estimations are obtained when the values of m are lower. The theoretical results obtained in Chapter 2 are also illustrated by a numerical simulation example. In this case, a scalar state process is generated by a first-order autoregressive model with missing measurements derived from two sensors. Linear and quadratic error variances are obtained for different uncertainty probabilities and values of m. On the one hand, these results confirm that the quadratic estimators are more accurate than the linear ones and, on the other hand, that the fixed-point smoothing estimators are more effective than the filtering ones. This example also shows that, as the probability of uncertainty decreases, the performance of the filters improves; finally, it shows that, as the values of m decline, the estimators become more accurate, since the state is missing in fewer consecutive observations.

It should be noted that the basic model in which the Bernoulli variables describing the uncertainty in the observations at each sensor are independent is a particular case of the proposed model, in which the assumption is made that the correlation of the Bernoulli variables is zero for any two sampling times. The model with correlation at consecutive sampling times is represented by this study when $m = 1$. However, these two assumptions may be unrealistic in many practical situations, and then the estimation algorithms must be modified to incorporate the effects of different types of correlation. Specifically, the form of correlation considered in the first two chapters is appropriate for modeling situations in which the

state cannot be missing in $m+1$ consecutive observations, as occurs, for instance, in sensor networks where sensor failures may happen and a failed sensor is not substituted immediately, but after m sampling times. It should also be noted that the multi-sensor system model considered in these chapters covers those situations in which the additive observation noises and the Bernoulli variables involved are independent from one sensor to another. This independence assumption simplifies the mathematical expressions considerably and is valid in a wide variety of applications, for example in wireless sensor networks which are characterized by sensor independence, limited storage capacity, lack of physical infrastructure and limited energy. Nevertheless, if this assumption were omitted, a similar technique to that used in these chapters would allow us to extend the current study to a more general case of this type, with no difficulty except the presence of greater complexity in the mathematical expressions.

Continuing *sensor network systems with missing measurements*, in the next stage of this study, we consider, at each sensor, the possibility of missing measurements containing only partial information about the state or even only noise. This situation is modeled by a sequence of independent random variables taking discrete values over the interval $[0, 1]$. For this kind of multi-sensor systems, we also assume that the system noises are autocorrelated and cross-correlated. Under these assumptions and by an innovation approach, recursive algorithms for the optimal linear filter are derived in **Chapter 3** using the two basic estimation fusion structures; specifically, the centralized and the distributed fusion estimation algorithms. The accuracy of these estimators is measured according to their error covariance matrices. This approach allows us to compare the performance of these algorithms in a numerical simulation example that illustrates the feasibility of the proposed filtering algorithms and enables their comparison with other filters that have been proposed. Thus, the applicability of the results obtained in Chapter 3 is illustrated by a numerical simulation example, in which a scalar state process generated by a

first-order autoregressive model is estimated from missing measurements derived from two sensors with autocorrelated and cross-correlated noises and, in accordance with the proposed observation model, two different independent sequences of random variables with a certain probability distribution over the interval $[0, 1]$ are used to model the missing phenomenon. The results obtained confirm, as in Chapter 1, that centralized and distributed fusion estimators have approximately the same error variances, which are slightly higher in the case of the distributed estimator; however, this deficit is compensated by the lighter computational burden and the reduced communication demand for the sensor networks. Moreover, compared with other estimation methods, the proposed algorithms provide better estimations in the mean square error sense.

With regard to the results obtained in this chapter, it should be stressed that the multi-sensor system models considered cover situations in which the sensor and process noises are one-step autocorrelated and two-step cross-correlated, and also those in which there are one-step cross-correlations between different sensor noises. This correlation assumption is valid in a broad spectrum of applications, for example, in target tracking systems with process and measurement noises that are dependent on the system state, or situations in which a target is observed by multiple sensors all operating in the same noisy environment. It should also be noted that the results obtained can be readily extended to more general systems involving finite-step autocorrelated and cross-correlated noises, the only difficulty being the greater complexity in the mathematical expressions. Finally, we observe that, unlike most previous results with correlated noises, in which suboptimal Kalman-type estimators are proposed, this chapter describes how optimal LS linear estimators are obtained.

In view of the results described in the first three chapters of this thesis, and to model a more general kind of failure in the measurements, our research in the remaining chapters is focused on *sensor network systems with random parameter*

matrices. For this class of systems, we first consider independent random state transition matrices and one-step correlated and cross-correlated random parameter matrices in the observation equation. The process and measurement noises are assumed to be one-step autocorrelated and two-step cross-correlated. Under these assumptions and taking into account certain properties of the state and noise processes derived from them, an optimal LS linear recursive filtering algorithm with a simple computational procedure is derived in **Chapter 4**, under an innovation approach. The results obtained are then applied to multi-sensor systems with missing and randomly delayed measurements. Specifically, as particular cases of the model analyzed in this chapter, the following systems are considered:

- Multi-sensor systems with missing measurements, with correlated and cross-correlated noises, when the missing measurement phenomenon in each sensor is described by different sequences of scalar random variables, with an arbitrary discrete probability distribution over the interval $[0, 1]$ correlated at consecutive sampling times.

- Multi-sensor systems with randomly delayed measurements, correlated at consecutive sampling times, with correlated and cross-correlated noises.

For both particular cases, the feasibility of the resulting filtering algorithm is analyzed, using the filtering error variances, in two numerical simulation examples, in which the state process generated by a first-order autoregressive model with state-dependent multiplicative noise is estimated from the missing measurements and the delayed measurements, respectively, derived from two sensors with different characteristics of random failures and noise correlation. In addition, using the mean square error criterion, a comparative analysis is performed of other estimation algorithms that have been reported. These comparisons show that the filtering algorithm obtained in this chapter performs better than those proposed elsewhere.

It should be noted that the multi-sensor systems with missing measurements that are studied in this chapter, as a particular case of the proposed model, are

in fact present in various real-world problems, such as transmission models with stand-by sensors or situations involving a partial loss of measurements. Additionally, the multi-sensor systems with randomly delayed measurements, which are also treated as a special case of the systems considered in Chapter 4, cover the situation in which two successive observations cannot be delayed. This kind of delay is usual in situations such as network congestion, random failures in the transmission mechanism or data inaccessibility at certain times. Moreover, unlike most existing results with random parameter matrices, in which the estimation problem is addressed via a system transformation, the proposed optimal LS linear recursive filtering algorithm is obtained without requiring any transformation of the original system into one with deterministic parameter matrices.

Continuing our study of *sensor network systems with random parameter matrices*, the estimation problem is addressed in **Chapter 5** for systems with both independent random parameter state transition and measurement matrices, under the same correlation assumptions about the noises as are considered in Chapter 4. Specifically, for this kind of systems, we investigate the distributed fusion filtering problem, doing so using an innovation approach to design recursive algorithms for the local LS linear filters of the system state, based on the measured output data derived from each sensor. To measure the accuracy of the local estimators, recursive formulas for the local filtering error covariance matrices are established. After obtaining the local filters, a distributed fusion filter is designed as the matrix-weighted sum of the local estimators that minimize the mean-squared estimation error. The distributed fusion filtering error covariance matrices is also derived.

The usefulness of the results obtained is illustrated by analysis of a numerical simulation example, consisting of a discrete-time linear networked system with state-dependent multiplicative noise, together with scalar measurements from four sensors with different uncertainty characteristics. Error variance comparison shows that both the centralized and the distributed filters outperform the local ones; this

example also shows that the slight superiority of the centralized filter over the distributed filter is counteracted by the greater robustness and enhanced fault-tolerance abilities of the latter.

We emphasize that the network system with random parameter matrices that is considered in this chapter provides a unified framework to address certain network-induced phenomena, such as multiplicative noise uncertainties, missing measurements or sensor gain degradation, and hence the distributed fusion filter obtained has wide applicability. Moreover, the computational procedure of the recursive algorithms for the local LS linear filtering estimators is very simple and suitable for online applications. Also, unlike alternative approaches, the proposed distributed fusion filter does not require the error cross-covariance matrices between any two local filters, but only their cross-covariance matrices.

Finally, our most recent contribution to the estimation problem in *sensor network systems with random parameter matrices* is the study of the LS quadratic estimation problem in systems with independent random parameter matrices and noises. In this study, presented in **Chapter 6**, the quadratic filter is also applied to systems with missing measurements derived from multiple sensors, when the missing measurement phenomenon in each sensor is described by different sequences of scalar random variables with arbitrary discrete probability distribution over the interval $[0, 1]$. The usefulness of the obtained results is illustrated by a numerical simulation example in which a two-sensor system with state-dependent multiplicative noise and missing measurements is considered. As in Chapter 2, error variance comparison shows that the quadratic filters outperform the linear ones. A comparative analysis also highlights the superiority of this quadratic filter over others that have been reported in the literature.

The system model with random parameter matrices considered in this chapter covers different uncertainties that commonly arise in sensor network systems, as multiplicative noises or as missing measurements. Consequently, the LS quadratic

filter presented here is superior to the LS linear estimators derived previously for systems with such uncertainties. Moreover, since to date the LS quadratic estimation problem with random parameter matrices has only been investigated for systems with scalar measurements and random observation matrices, our contribution provides an important generalization, as random state-transition matrices and multidimensional observations are commonly found in many real-world situations. In particular, the estimators obtained can be applied to multi-sensor systems with different uncertainty characteristics in the sensors, such as those with state-dependent multiplicative noise and missing measurements that are considered in the simulation example.

Resultados y conclusiones

En esta tesis doctoral, se estudian problemas de estimación mínimo cuadrática en sistemas estocásticos lineales en tiempo discreto con observaciones ruidosas procedentes de múltiples sensores, afectadas por parámetros aleatorios que modelizan diferentes tipos de fallo en las medidas. Según se detalla en la sección de metodología, utilizando un tratamiento por innovaciones y métodos de fusión para combinar la información procedente de los diferentes sensores, se obtienen algoritmos recursivos de estimación mínimo cuadrática. Específicamente, empleando los métodos de fusión centralizado y distribuido, se establecen algoritmos recursivos para los estimadores de filtrado y suavizamiento punto fijo, tanto lineales como cuadráticos, junto con sus matrices de covarianzas del error que miden la bondad de dichos estimadores. A continuación se detallan los principales resultados y conclusiones de este estudio.

En una primera etapa, se consideran *sistemas de redes de sensores con observaciones inciertas*, (es decir, observaciones que pueden no contener información sobre el estado y ser sólo ruido), situación que se modeliza mediante variables binarias con valores uno o cero (variables de Bernoulli), dependiendo de si el estado está presente o ausente en la observación correspondiente. Suponiendo que dichas variables están correladas en instantes que se diferencian m unidades de tiempo, en el **Capítulo 1**, mediante los métodos de fusión centralizado y distribuido, se establecen algoritmos recursivos de filtrado y suavizamiento punto fijo lineales y, en el **Capítulo 2**, se obtienen estimadores cuadráticos utilizando el método de

fusión centralizado.

Los resultados obtenidos en el Capítulo 1 se ilustran mediante dos ejemplos de simulación numérica en los que se consideran sistemas lineales cuyo estado escalar y bidimensional, respectivamente, es estimado a partir de observaciones inciertas procedentes de dos sensores, bajo la hipótesis de que las variables que modelizan la incertidumbre son correladas en instantes que se diferencian m unidades de tiempo. En ambos ejemplos se observa que la precisión de los estimadores centralizado y distribuido es muy similar. También, al comparar las varianzas de los errores de estimación para diferentes probabilidades de incertidumbre y distintos valores de m, se confirma que los suavizadores punto fijo son más precisos que los filtros y, que conforme disminuyen los valores de m, el comportamiento de los estimadores mejora. Los resultados teóricos del Capítulo 2, se ilustran mediante la estimación del estado de un modelo autorregresivo de primer orden, a partir de observaciones inciertas procedentes de dos sensores. Comparando las varianzas de los errores de estimación lineales y cuadráticos para diferentes probabilidades de incertidumbre y valores de m, se concluye, que los estimadores cuadráticos son bastante más precisos que los lineales. También, como en el caso de los estimadores lineales, para el problema de estimación cuadrática, se tiene que los suavizadores punto fijo son más efectivos que los filtros y que a medida que disminuyen la probabilidad de incertidumbre y/o los valores de m, se produce una mejora en la precisión de los estimadores.

Es importante señalar que el modelo básico en que la incertidumbre de las observaciones está descrita por variables de Bernoulli independientes, es un caso particular del modelo propuesto, suponiendo que la correlación de las variables en dos instantes cualesquiera de tiempo es cero. También el caso en el que hay correlación en instantes consecutivos queda cubierto por el estudio realizado, considerando $m = 1$. Sin embargo, en diversas situaciones prácticas estas dos hipótesis pueden no ser realistas, y los algoritmos de estimación deben modificarse para incorporar

el efecto de los diferentes tipos de correlación exigidos. Concretamente, la hipótesis de correlación considerada en los dos primeros capítulos es apropiada para describir situaciones en las que el estado no puede faltar en $m+1$ observaciones consecutivas, como por ejemplo, en redes de sensores en las que éstos pueden fallar y no son sustituidos inmediatamente, sino m instantes después de haber fallado. También hay que señalar que el modelo con múltiples sensores considerado en estos capítulos, es adecuado en situaciones en las que los ruidos aditivos de las observaciones y las variables de Bernoulli que modelizan la incertidumbre son independientes de un sensor a otro. Esta hipótesis de independencia simplifica considerablemente las expresiones matemáticas y es válida en una amplia gama de aplicaciones, por ejemplo, en redes de sensores inalámbricas caracterizadas por sensores independientes, capacidad de almacenamiento restringida, falta de infraestructuras físicas y energía limitada. No obstante, si se debilita dicha hipótesis, un procedimiento similar al empleado en estos capítulos permitirá ampliar este estudio a modelos más generales sin más dificultad que la mayor complejidad de las expresiones matemáticas involucradas en los desarrollos.

En una segunda etapa, continuando con sistemas de redes de sensores con fallos en las medidas, se considera que las observaciones de los distintos sensores pueden contener sólo información parcial del estado. Este tipo de fallo está descrito por una sucesión de variables discretas independientes con distribución en el intervalo [0,1] y, se supone además que los ruidos aditivos del sistema son autocorrelados y con correlación cruzada. Bajo estas hipótesis, en el **Capítulo 3**, utilizando los métodos de fusión centralizado y distribuido, se establecen algoritmos recursivos de filtrado lineal. Para medir la bondad de los estimadores proporcionados por cada algoritmo se calculan las matrices de covarianzas del error de estimación. Con la finalidad de ilustrar los resultados obtenidos, se presenta un ejemplo de simulación numérica en el que se considera un proceso estado escalar generado por un modelo autorregresivo de primer orden, y se estima a partir de medidas procedentes de

dos sensores con información parcial sobre el estado y correlación de los ruidos como se ha indicado antes. Al igual que en el Capítulo 1, se observa que las varianzas de los errores de estimación centralizados y distribuidos son similares, siendo ligeramente peor el estimador distribuido, lo cual se compensa por el menor coste computacional. Finalmente, los algoritmos propuestos proporcionan mejores estimaciones, en el sentido de menor error cuadrático medio, al compararlos con otros algoritmos de estimación existentes en la literatura.

En relación con los resultados obtenidos en este capítulo, hay que destacar que los sistemas considerados abarcan tanto situaciones en las que los ruidos del estado y de las observaciones son autocorrelados en instantes consecutivos, como aquellas en las que existe correlación cruzada de los ruidos en instantes que se diferencian dos unidades de tiempo. Además, los ruidos de las observaciones procedentes de los distintos sensores pueden tener correlación cruzada en instantes consecutivos. Estas hipótesis de correlación son apropiadas en una gran variedad de aplicaciones, por ejemplo, en sistemas de localización de objetivos con ruidos dependientes del estado, o en situaciones en las cuales el objetivo es observado mediante múltiples sensores y todos actúan en el mismo entorno ruidoso. Es importante mencionar que este estudio se puede generalizar a sistemas que incluyan ruidos autocorrelados y con correlación cruzada en un número finito de instantes de tiempo, sin más dificultad que la mayor complejidad de las expresiones matemáticas. Por otra parte, los estimadores de mínimos cuadrados obtenidos en este capítulo son óptimos, a diferencia de otros resultados previos con ruidos correlados en los cuales sólo se proponen estimadores subóptimos de tipo Kalman.

A la vista de los resultados obtenidos en los tres primeros capítulos, y con objeto de modelizar otros tipos más generales de fallos en las medidas, los capítulos restantes se centran en *sistemas de redes de sensores con matrices aleatorias*. En este tipo de sistemas, el estudio se inicia considerando matrices de transición del estado independientes y matrices de las observaciones correladas y con correlación

cruzada en instantes consecutivos. En relación con los ruidos que intervienen en el sistema, éstos se suponen autocorrelados y con correlación cruzada. Bajo estas hipótesis, en el **Capítulo 4**, mediante un tratamiento por innovaciones, se obtiene un algoritmo recursivo para el filtro lineal óptimo de mínimos cuadrados. Los resultados establecidos en este capítulo se aplican a los siguientes sistemas:

- Sistemas con fallos en las medidas procedentes de múltiples sensores con ruidos correlados, cuando los fallos están descritos por diferentes sucesiones de variables aleatorias escalares con distribución discreta en el intervalo $[0, 1]$, correladas en instantes consecutivos.

- Sistemas con observaciones retrasadas procedentes de múltiples sensores, cuando los retrasos están modelizados por variables de Bernoulli correladas en instantes consecutivos y los ruidos aditivos son correlados.

Para estos dos casos particulares, la bondad de los estimadores es analizada mediante las varianzas del error de filtrado, en sendos ejemplos de simulación numérica, en los cuales se estima el proceso estado de un modelo autorregresivo de primer orden con ruido multiplicativo, a partir de observaciones inciertas y retrasadas, respectivamente, procedentes de dos sensores con diferentes características de los fallos aleatorios y distintas hipótesis de correlación de los ruidos. Utilizando el criterio de menor error cuadrático medio, se observa un mejor comportamiento del algoritmo de filtrado propuesto, con respecto a otros algoritmos de estimación publicados anteriormente.

Los sistemas con fallos en las medidas procedentes de múltiples sensores considerados en este capítulo, como caso particular del modelo propuesto, aparecen en diversas situaciones reales, como por ejemplo, en sistemas de transmisión con sensores *stand-by* o situaciones que conllevan una pérdida parcial de las observaciones. También los sistemas con retraso aleatorio en las observaciones procedentes de múltiples sensores, tratados como caso especial del modelo propuesto, son adecuados en situaciones en las que dos observaciones consecutivas no pueden estar

retrasadas. Este tipo de retraso es adecuado en situaciones como, por ejemplo, congestión en las redes de comunicación o inaccesibilidad de datos durante ciertos instantes de tiempo por cualquier tipo de fallo en el mecanismo de transmisión. Por otra parte, hay que destacar que el algoritmo recursivo de filtrado lineal propuesto se obtiene sin necesidad de transformar el sistema original en uno con matrices determinísticas, mientras que en la mayoría de resultados existentes con matrices aleatorias es necesaria dicha transformación del sistema para abordar el problema de estimación.

Continuando con el estudio de *sistemas de redes de sensores con matrices aleatorias*, en el **Capítulo 5** se aborda el problema de filtrado fusión distribuido en sistemas con matrices aleatorias independientes, tanto en la ecuación del estado como en las de las observaciones, considerando las mismas hipótesis de correlación sobre los ruidos que en el Capítulo 4. Con este propósito, a partir de las observaciones procedentes de cada sensor, se desarrollan algoritmos recursivos para la obtención de los filtros lineales locales del estado, junto con sus matrices de covarianzas del error que miden la precisión de dichos estimadores. Una vez que los estimadores de filtrado local están disponibles, se obtiene el filtro fusión óptimo como combinación lineal de dichos estimadores, ponderada por matrices, verificando el criterio de mínima varianza, junto con sus matrices de covarianzas del error.

Para ilustrar la aplicación de los resultados obtenidos, se presenta un ejemplo de simulación numérica, en el que se considera un sistema lineal en tiempo discreto con ruido multiplicativo en la ecuación del estado, y medidas escalares procedentes de cuatro sensores con diversas características de incertidumbre. Realizando comparaciones entre las varianzas de los errores de estimación, se concluye que tanto el filtro centralizado como el distribuido son más precisos que los locales, con una ligera inferioridad del filtro distribuido con respecto al centralizado, la cual se compensa por una mayor robustez y mejor tolerancia a los fallos del filtro fusión distribuido.

Es importante resaltar que el modelo de redes de sensores con matrices aleatorias, considerado en este capítulo, es adecuado para tratar de manera unificada algunos fenómenos inducidos por las redes, tales como, incertidumbres descritas por ruidos multiplicativos o medidas faltantes y, por tanto, es válido en diversas aplicaciones prácticas. Por otra parte, el procedimiento computacional de los algoritmos recursivos de los filtros lineales locales es muy sencillo y adecuado para realizar aplicaciones *on line*. Para la obtención del filtro fusión distribuido, se utilizan las matrices de covarianza cruzadas entre dos cualesquiera filtros locales sin necesidad de calcular las matrices de covarianzas cruzadas de los errores de estimación, como se describe en otros trabajos.

Por último, nuestra contribución más reciente al problema de estimación en *sistemas de redes de sensores con matrices aleatorias* es el estudio del problema de estimación cuadrática en sistemas con matrices aleatorias y ruidos independientes. En el **Capítulo 6** se aplica el filtro cuadrático obtenido a sistemas con fallos aleatorios en las medidas, descritos mediante diferentes sucesiones de variables aleatorias escalares con cualquier distribución discreta en el intervalo $[0, 1]$. La aplicabilidad de los resultados obtenidos se ilustra mediante un ejemplo numérico en el que se considera un sistema con ruido multiplicativo en la ecuación del estado y observaciones inciertas procedentes de dos sensores. Al igual que en el Capítulo 2, realizando comparaciones entre las varianzas del error de estimación, se comprueba que los filtros cuadráticos son mejores que los lineales. El filtro cuadrático presentado en este capítulo muestra una mayor precisión con respecto a otros filtros cuadráticos existentes en la literatura.

El modelo con matrices aleatorias considerado abarca diversos tipos de incertidumbre que suelen ocurrir en sistemas de redes de sensores, tales como, observaciones inciertas o ruidos multiplicativos. Por tanto, el filtro cuadrático obtenido mejora a los lineales obtenidos para otros sistemas con dichas características de incertidumbre. Teniendo en cuenta que hasta ahora el problema de estimación

cuadrática con matrices aleatorias sólo ha sido abordado para sistemas con medidas escalares y matrices aleatorias en la ecuación de observación, este estudio proporciona una generalización importante debido a que matrices de transición del estado aleatorias y observaciones multidimensionales se encuentran frecuentemente en muchas situaciones reales.