# Kalman Filtering under Wide Band Noises

**Kanda Jamal Abuassba**

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mathematics

Eastern Mediterranean University
October 2018
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Assoc. Prof. Dr. Ali Hakan Ulusoy
Acting Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy in Mathematics.

Prof. Dr. Nazim Mahmudov
Chair, Department of Mathematics

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Doctor of Philosophy in Mathematics.

Prof. Dr. Agamirza Bashirov
Supervisor

Examining Committee

1. Prof. Dr. Agamirza Bashirov
2. Prof. Dr. Elman Hasanoğlu
3. Prof. Dr. Varga Kalantarov
4. Prof. Dr. Nazim Mahmudov
5. Asst. Prof. Dr. Nidai Şemi

# ABSTRACT

Kalman filtering is a powerful estimation method. One of its weaknesses is related to the white or colored nature of the disturbing noises in the Kalman filtering model. At the same time, real noises are rarely white or colored. They are mostly wide band. In this regard, white or colored noise Kalman filtering makes concessions on adequacy. This pushes system scientists to develop mathematical methods of estimation for systems corrupted by wide band noises. In applications, wide band noises are detected by their autocovariance and cross-covariance functions which do not allow modeling them uniquely. Therefore, it becomes important to develop estimation methods which are independent of a class of wide band noises, but dependent on the unique autocovariance and cross-covariance functions. Such results are called invariant results. In this paper, we prove a complete set of invariant equations for Kalman type filter for a linear signal-observation system corrupted by correlated wide band noises. This filter has a ready form to be used in applications, just respective numerical methods must be developed. We also discuss an application scenario for the proposed filter.



**Keywords:** Wiener process, white noise, wide band noise, Kalman filter.

# ÖZ

Güçlü bir tahmim metodu olan Kalman filtrelemesinin zayıf yönlerinden biri Kalman filtreleme metodundaki etkileyici gürültünün beyaz veya renkli doğası ile ilgilidir. Aynı zamanda gerçek gürültüler genellikle geniş bandlı olup nadiren beyaz veya renklidirler. Bu bağlamda renkli veya beyaz gürültü uygulamalarında Kalman filtrelemesi kullanmak yetersiz kalıyor. Bu, sistem bilimcileri geniş bandlı gürültü tarafından etkilenmiş sistemler için için matematiksel tahmin metodları geliştirmeye yitmiştir. Uygulamalarda, geniş band gürültüleri özdeğişim ve çapraz değişim fonksiyonları tarafından saptanır ki bu onların tek olarak modellenmesine imkan vermez. Bu nedenle geniş bandlı gürültü sınıfından bağımsız fakat tek özdeğişim ve çapraz değişim fonksiyonlarına bağımlı bir tahmin metodu geliştirmek önemli olur. Bu sonuçlar değişmez sonuçlar olarak adlandırılır. Bu tezde, geniş bandlı gürültüler tarafından etkilenmiş lineer sinyal-gözlem sistemleri için Kalman tipli filtrenin tüm denklemleri bulunmuş ve ispat edilmiştir. Matematiksel olarak bulunmuş bu filter uygulamalarda kullanılmaya hazır durumdadır. Sadece uygun nümerik yöntemlerin geliştirilmesi gerekmektedir.

**Anahtar Kelimeler:** Wiener süreci, beyaz gürültü, geniş bandlı gürültü, Kalman filtresi.

# DEDICATION

*Normally, one has two parents, but I have four.*

*To them, Hind, Ayda, Jamal and Muhammad, for their constant support and unconditional love.*

*Together with my husband and my beloved kids for the time that I have to spend with them and I couldn't because of this.*

# ACKNOWLEDGMENT

Firstly, I owe my deepest gratitude to my supervisor Prof. dr. Agamirza Bashirov for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. He stands not just as a teacher for me, but as a father within the years of my Ph.D. study.

Besides my supervisor, I would like to thank the rest of my thesis committee, for the time they spend and insightful comments and encouragement.

My sincere thanks also goes to all of the members of mathematics departmen at EMU, professors and assistants, who provided me the opportunity to join their team as intern, and surrounded me with a very warm enviroment.

Finally, I found no suitable words to express my gratitude to all of my sisters, brothers and friends. without they precious support it would not be possible to conduct this research.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| DE | Differential Equation, |
| FDE | Partial Differential Equation, |
| GRV | Gaussian Random Variable. |
| KF | Kalman Filter, |
| ODE | Ordinary Differential Equation, |
| SDE | Stochastic Differential Equation, |
| WBN | Wide Band Noise, |
| WN | White Noise, |

# Chapter 1

# INTRODUCTION

Kalman filtering (simply KF) [1, 2] is a powerful estimation method having great engineering applications. Its application areas include guidance, navigation, and control of aircrafts and space crafts [3], control of robotic motions [4], forecasting and analysis of time series in signal processing and econometrics [5], but not restricted to these.

Despite its great applications, KF has deficiencies. For example, it assumes linear state-observation system. This deficiency is removed in an extended KF by linearization of nonlinearities. Another deficiency which is related to the nature of the noise processes in the KF model is discussed. Overall applications of KF considers WN model (independent or correlated) of disturbing noises. In [8] the KF is modified to colored noises. But as noted in [9], the real noises behave as a WBN in which WN's are an ideal case. Therefore, to be more adequate there is a need in a modification of KF to WBN driven systems.

There are principally two approaches to WBN's. In [21] and references therein, WBN driven systems are investigated by a method of approximation. The other method through integral representation was considered in [22, 23] This leads to modeling WBN's as a distributed delay of WN's [24, 25] Indeed, for a function

$\Phi_{t,s}$ of two variables on $[0,\infty) \times [-\varepsilon, 0]$ and a Wiener process $w$, it can be calculated that the random process

$$\varphi_t = \int_{\max(0,t-\varepsilon)}^{t} \Phi_{t,s-t} dw_t \,,$$

has the autocovariance function $(\varphi_{t+\theta}, \varphi_t) = \Lambda_{t,\theta} \neq 0$ if $0 \le \theta < \varepsilon$, and $\Lambda_{t,\theta} = 0$ if $\theta \ge \varepsilon$. Therefore, $\varphi$ is a WBN. It becomes stationary on $[\varepsilon, \infty)$ if $\Phi$, which is called a relaxing (damping) function, is independent on its first variable $t$. In the stationary case

$$\Lambda_{t,\theta} = \int_{\max(-\varepsilon,\theta-\varepsilon)}^{0} \Phi_{s-\theta} \Phi_s^* ds \,,$$

where $\Phi^*$ is the transpose of $\Phi$ and it is seen that $\Lambda_{t,\theta} \equiv \Lambda_\theta$ if $t \ge \varepsilon$.

This representation is somehow universal because it covers WN's and point wise delays of them if $\Phi$ is selected as Dirac's delta-function [28, 29]. Moreover, in [30, 31] in the one-dimensional stationary case it is shown that for given positive definite function $\Lambda$ there are infinitely many relaxing functions $\Phi$ such that all them produce a WBN with the autocovariance function $\Lambda$. Therefore, it is important to obtain results which are independent on the infinite variations of $\Phi$, but dependent on the unique $\Lambda$ because in applied problems WBN's are detected by their autocovariance functions. Such results are called invariant results.

Some invariant results for WBN driven systems have already been obtained. In [32, 33] invariant maximum principle in the Pontryagin's form and controllability result are established for nonlinear systems. Invariant KF is also obtained in the signal noise is wide band but the observation noise is non-degenerate white [34], and when the signal noise is WBN but the observations are WN's [35]. In fact, this work generalizes these results to the case when both signal and observations systems are corrupted by the sum of white and WBN’s. We assume that there is a correlation between the WBN’s while WN’s are independent each other as well as on WBN’s. The main result of the this study expresses a complete set of equations for the best least square estimate in terms of autocovariance and cross-covariance functions of disturbing WBN’s.

To the best of our knowledge, the first record about the concept of a WBN, demonstrating its adequacy, appears in page 126 of the 1975 edition book [9] by Fleming and Rishel, although possible earlier discussions in the engineering literature are not excluded. Later, the adequacy of WBN’s was also prompted in [10]. In [11] WBN’s in speech signals are analyzed. Since then, the system scientists created just two baselines for handling the WBN driven systems. In [12, 13, 14, 15, 16, 17, 18, 19, 20, 21] WBN driven systems are discussed by a method of approximation. A different method by integral representation was suggested in [22, 23] which leads to modeling WBN’s as a distributed delay of WN’s [24, 25].

Unfortunately, both of them have been remained in the theoretical level, do not providing any efficient algorithm of applied nature. Instead, efficient applied algorithms have been created for WN driven systems on the basis of its strong mathematical baseline, provided by Ito’s stochastic calculus [26, 27]. Additionally, in some applied areas WN’s, being mathematically an ideal case of WBN’s, present an approximation to real noises up to a certain level of adequacy. Perhaps, these factors prompted WN’s as a widely accepted noise model. Withal, technology needs more delicate estimation methods. In this way, KF for WBN driven signal-observation system which can be realized as an algorithm of applied nature is seem to be necessary. In this study, the main result obtained is converting those theoretical results, obtained via the above mentioned method by integral representation, to invariant results for more general KF (corrupted by correlated WBN's). This result can be turned into practically realizable algorithm.

This study consists of five chapters. Chapter 1 consists of introduction to this work and some preliminaries.

In Chapter 2 Issues related to the definition of KF and several generalizations of kf are concerned with. Staring from the classical KF in section 2.1 followed by the discrete and the continuous KF in the sections 2.2 and 2.3, ending the chapter with selected generalizations of KF in section 2.4.

Chapter 3 introduces the concept of WBN in section 3.1 and some results for the KF under WBN. Different cases have been investigated starting from KF for WBN in signal or in measurements, in sections 3.2 and 3.3, respectively. The last section, investigate the KF for WBN's in both signal and measurements.

Chapter 4 starts with introducing the integral representation of the WBN and the invariance concept in section 4.1. In section 4.2 and 4.3 two cases of invariant KF corrupted by independent WBN and WN's in the signal and measurements, respectively. Followed by, invariant KF corrupted by a summation of independent WN and WBN's in the signal in section 4.4.

The main and completely new result is presented in chapter 5. A complete set of invariant results for KF for a linear signal-observation system under correlated WBN is introduced. In the first section WBN's are defined and an integral representation is motivated for them. Section 5.2 defines a very crucial concept of invariance of results for WBN driven systems. Section 5.3 sets a filtering problem under consideration. In Section 5.4 the proof of the invariant equations for KF for correlated WBN is finalized and discussed. Finally, in Section 5.5 numerical aspects are investigated to introduce discrete formulas for the resultant equations. Moreover, in many sections, LQG applications for the KF problem are investigated as well.

# Chapter 2

# THE KF AND SOME OF ITS GENERALIZATIONS

## 2.1 Classical KF

KF is one of the most essential results of filtering theory. It plays great role in space engineering , telecommunications, etc. In 1960-1961 Kalman [1] and Bucy [2] it was presented as a method for estimating linear systems disturbed by WN processes. Kalman build the construction of state estimation depending on probability theory, precisely, on the properties of conditional GRV's. The concept that he wanted to minimize is the state vector covariance norm, so, yielding to the classical form: the new estimation is given from the previous one by addition of correction term respective to the prediction error.

KF presents the formulas of the best estimate $\hat{x}(t)$ of $x(t)$ using the observation process $z(t), 0 \leq s \leq t$, consisting a partially observable system $(x, z)$. Here the state process satisfy the following linear stochastic DE

$$\begin{cases} x'(t) = Ax(t) + Bw(t), t > 0, \\ x(0) = x_0, \end{cases} \tag{2.1.1}$$

and $z(t)$ is the observation process, based on $x(t)$ in the following linear form

$$z(t) = Kx(t) + v(t), t > 0, \tag{2.1.2}$$

where $A, B$ and $K$ are assumed to be matrices, $x$ and $z$ are respective vector-valued, $w(t)$ and $v(t)$ are independent vector-valued WN's, it is also assumed that $x_0$ is Gaussian random vector with mean equals to zero and known covariance $P_0$, $x_0$ is independent with respect to the WN's $w(t)$ and $v(t)$. The previous two equations formulate the classical form of the KF problem.

The importance of KF comes from its presenting of the best estimate $\hat{x}(t)$ of the observation signal $x(t)$ as a dynamical process. It is presented as a solution of the following linear equation

$$\begin{cases} \hat{x}'(t) = A\hat{x}(t) + P(t)K^T\big(z(t) - K\hat{x}(t)\big), t > 0, \\ \hat{x}(0) = 0, \end{cases} \tag{2.1.3}$$

Where $K^T$ is the transpose of $K$ and $P$ is presented as a solution of the following deterministic matrix Riccati equation

$$\begin{cases} P\acute{(}t) = AP(t)A^T + BB^T - P(t)K^TKP(t), t > 0, \\ P(0) = \mathrm{cov}x_0 = P_0, \end{cases} \tag{2.1.4}$$

the linear transformation (3.1.3) is called KF. It presents the best estimate $\hat{x}(t)$ for the observation $z(t)$ at every value $t > 0$.

## 2.2 Discrete KF

### 2.2.1 Introduction and problem formulation

The KF problem is different from the linear regression problem by assuming that $x$ is a dynamical process. In the discrete KF problem this difference is clearly seen. The aim of this section is to obtain a continual estimator of a state vector whose values change discretely over time. Updating is obtaining based on a set of measurements or predictions $z(t)$which gives information about the state signal $x(t)$. Depending on those observations the estimator will provide an estimate $\hat{x}(t + \epsilon)$ at some time $t + \varepsilon$. If $\varepsilon > 0$, the problem will be a prediction filter, and, if $\varepsilon < 0$ it will be a smoothing filter, and if $\varepsilon = 0$ the problem is simply called filtering.

The description of regression problem in section 3.1 will be modified to this case by considering the state model as a discrete dynamical model, it could be described with

$$x_{k+1} = A_k x_k + B_k w_k, k = 0, 1, 2, \ldots, \tag{2.2.1}$$

considering that the initial state vector $x_0$ is Gaussian random vector with expectation equals to zero and $B_k$ is the noise transition matrix of respective dimension, $A_k \in R^{n \times n}$ and the state noise $w_k \in R^n$ is assumed to be white Gaussian with zero mean, $Q_k$ covariance i.e.

$$w_k \sim \aleph(0, Q_k),$$

and

$$Q_k = E(w_k, w_k^T).$$

Consider the discrete measurement process $z_1, z_2, \ldots$ are defined in the following manner

$$z_k = H_k x_k + G_k v_k, k = 1, 2, \ldots, \tag{2.2.2}$$

where $H_k \in R^{m \times n}$ and $G_k$ is an invertible square matrix of respective dimensions and the measurement noise $v_k \in R^m$ . Measurements noise, $v_k$, is white Gaussian with zero mean , $R_k$ covariance i.e.

$$v_k \sim \aleph(0, R_k),$$

and

$$R_k = E(v_k v_k^T).$$

Also, assume the state noise $w_k$ and measurements noise $v_k$ independent and uncorrelated.

$$E\left(w_i w_j^T\right) = Q_k \delta(i, j),$$

$$E\left(v_i v_j^T\right) = R_k \delta(i, j),$$

$$E\left(w_i v_j^T\right) = 0,$$

where, $\delta(i,j)$ is defined as

$$\delta(i,j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{2.2.3}$$

The discrete KF problem (2.2.1) and (2.2.2) yields a formulation for the best estimation of the state vector $x_k$ depending on the measurements $z_k$, $k = 1, 2 \ldots$.

In fact, the best estimate $\hat{x}_k$ of the state vector $x_k$ is the conditional expectation

$$\hat{x}_k = E(x_k | z_1, \ldots, z_{k-1}),$$

since both of the state vector $x_k$ and the measurements are linear combinations of independent and Gaussian vectors, so $x_0, x_1, \ldots, z_1, z_2, \ldots$ is a Gaussian system.

**2.2.2 Derivation of discrete KF**

Define a priori estimate error $e_k$ as

$$e_k = \hat{x}_k - x_k.$$

Let $x_k^*$ be the best linear estimate of $x_k$ on the base of the measurement $z_1$, $z_2$, . . ., $z_k$. A posteriori estimate error as

$$e_k^* = x_k^* - x_k, \tag{2.2.4}$$

a priori estimate error covariance $P_k$ is defined as

$$P_k = E(e_k \, e_k^T), \tag{2.2.5}$$

and a posteriori estimate error covariance $P_k^*$ is defined as

$$P_k^* = E(e_k^* e_k^{*^T}).$$

Now, substituting error expression in the formula of the priori error covariance gives

$$P_k^* = E\big((\hat{x}_k - x_k)(\hat{x}_k - x_k)^T\big). \tag{2.2.6}$$

The expression of the state estimator could be written as

$$x_k^* = \hat{x}_k + K_k(z_k - H_K\hat{x}_k), \tag{2.2.7}$$

where, $K_k$ is the blending factor, which is called as a Kalman gain matrix. In fact, to design KF an equation for $K_k$ must be derived. From the previous expression the covariance $P_k$ could be written as

$$P_k^* = E\big((x_k^* - x_k)(x_k^* - x_k)^T\big), \tag{2.2.8}$$

substituting the expression (3.2.7) in the expression (3.2.8),

$$P_k^* = E(\{\hat{x}_k + K_k(z_k - H_K\hat{x}_k) - x_k\}\{\hat{x}_k + K_k(z_k - H_K\hat{x}_k) - x_k\}^T), \tag{2.2.9}$$

then

$$P_k^* = (\{e_k + K_k(z_k - H_K\hat{x}_k)\}\{e_k + K_k(z_k - H_K\hat{x}_k)\}^T),$$

after simplifying,

$$P_k^* = E\big(e_k e_k^T + K_k z_k - H_K\hat{x}_k e_k^T + e_k (z_k - H_K\hat{x}_k)^T {K_k}^T + K_k(z_k - H_K\hat{x}_k)(z_k - H_K\hat{x}_k)^T {K_k}^T\big). \tag{2.2.10}$$

We already know that

$$z_k - H_K\hat{x}_k = H_K x_k + v_k - H_K\hat{x}_k = v_k - H_K e_k\,.$$

Hence the expression (3.2.20) becomes

$$P_k^* = \big(e_k e_k^T + K_k v_k e_k^T - K_k H_K e_k e_k^T + e_k v_k^T {K_k}^T - e_k e_k^T H_k^T {K_k}^T + K_k(v_k v_k^T - v_k e_k^T H_k^T - H_K e_k v_k^T + H_K e_k e_k^T H_k^T){K_k}^T\big),$$

implying

$$P_k^* = E(e_k e_k^T) + K_k E(v_k e_k^T) - K_k H_K(e_k e_k^T) + E(e_k v_k^T){K_k}^T - E(e_k e_k^T)H_k^T {K_k}^T + K_k\{E(v_k v_k^T) - E(v_k e_k^T)H_k^T - H_K E(e_k v_k^T) + H_K E(e_k e_k^T)H_k^T\}{K_k}^T,$$

since $e_k$ and $v_k$ are linearly independent the last formulae can be modified to

$$P_k^* = E(e_k e_k^T) - K_k H_K E(e_k e_k^T) - E(e_k e_k^T)H_k^T {K_k}^T + K_k\{E(v_k v_k^T) + H_K E(e_k e_k^T)H_k^T\}{K_k}^T,$$

or,

$$P_k^* = P_k - K_k H_K P_k - P_k H_k^T {K_k}^T + K_k (R_k + H_K P_k H_k^T) {K_k}^T.$$

Now, re-arranging the above equation will introduce

$$P_k^* = (I - K_k H_K) P_k (I - K_k H_K)^T + K_k R_k {K_k}^T, \tag{2.2.11}$$

where, $I$ is an identity matrix of respective dimension.

Expanding formulae (2.2.11) and regrouping terms as linear and quadratic give

$$P_k^* = P_k - \left(K_k H_K P_k + P_k H_k^T {K_k}^T\right) + K_k (H_K P_k H_k^T + R_k) {K_k}^T, \tag{2.2.12}$$

where $P_k$ is quadratic only when $(H_K P_k H_k^T + R_k)$ is symmetric and positive definite.

Assume that

$$\phi_k = H_K P_k H_k^T + R_k,$$

substitute the value of $\phi_k$ in the last expression of $P_k^*$

$$P_k^* = P_k - \left(K_k H_k P_k + P_k H_k^T {K_k}^T\right) + K_k \phi_k {K_k}^T.$$

The aim is to find the value of $K_k$ such that the error covariance $P_k^*$ will be minimized. So, differentiate $P_k^*$ with respect to $K_k$ and find the critical $K_k$

$$\frac{\partial P_k^*}{\partial K_k} = -(H_k P_k)^T - P_k H_k^T + 2 K_k \phi_k = 0,$$

where $P_k$ is symmetric matrix. Then

$$K_k = P_k H_k^T (H_K P_k H_k^T + R_k)^{-1}. \tag{2.2.13}$$

Now substituting the value of Kalman gain matrix $K_k$ in the expression of the error covariance $P_k^*$ gives what we call, minimum variance estimator

$$P_k^* = (I - K_k H_k)\, P_k. \tag{2.2.14}$$

For a next priori state estimate $\hat{x}_{k+1}$ using $x_k^*$, starting from (2.2.1) by ignoring the state noise since it has zero mean we obtain

$$\hat{x}_{k+1} = A_k x_k^*, \; k = 0, 1, 2 \ldots. \tag{2.2.15}$$

Next, find a priori error for the step k+1

$$e_{k+1} = \hat{x}_{k+1} - x_{k+1} = A_k x_k^* - A_k x_k - B_k w_k = A_k e_k^* - B_k w_k.$$

Assuming that the error $e_{k+1}$ has zero mean,

$$P_{k+1} = E(e_{k+1} {e_{k+1}}^T),$$

substituting the value of $e_{k+1}$ we obtain

$$P_{k+1} = E\big((A_k e_k^* - B_k w_k)(A_k e_k^* - B_k w_k)^T\big),$$

or,

$$P_{k+1} = E\big(A_k e_k^* {e_k^*}^T {A_k}^T - A_k e_k^* w_k^T B_k^T - B_k w_k {e_k^*}^T e_k^* {A_k}^T + B_k w_k w_k^T B_k^T\big),$$

second term and third term of the last equation are zeros since $e_k^*$ and $w_k$ are independent i.e.

$$P_{k+1} = A_k E\big(e_k^* {e_k^*}^T\big) {A_k}^T + B_k E(w_k w_k^T) B_k^T,$$

or,

$$P_{k+1} = A_k P_k {A_k}^T + B_k Q_k B_k^T. \tag{2.2.16}$$

**Theorem 2.2.1.** The best estimate $\hat{x}_k$ of $x_k$ under the previous conditions and notations and based on the set of measurements $\{z_1,\ z_2, \ldots,\ z_{k-1}\}$ must satisfy the following recurrence equation

$$\begin{cases} \hat{x}_{k+1} = A_k \hat{x}_k + A_k P_k H_k^T (H_k P_k H_k^T + D_k D_k^T)^{-1} (z_k - H_k \hat{x}_k), \\ x_0 = 0, k = 1, 2, 3, \ldots, \end{cases} \tag{2.2.17}$$

also, the error covariance must satisfy the following recurrence equation

$$\begin{cases} P_{k+!} = A_k P_k A_k^T - A_k P_k H_k^T (H_k P_k H_k^T + D_k D_k^T)^{-1} H_k P_k A_k^T + B_k Q_k B_k^T, \\ P_0 = E(x_0 x_0^T), k = 1, 2, 3, \ldots. \end{cases} \tag{2.2.18}$$

**Proof**. Equation (2.2.17) follows directly from (2.2.15),(2.2.7) and (2.2.13) , and $\hat{x}_0 = 0$ since the initial estimation based on no observation equals to the expectation of the initial state i.e. $\hat{x}_0 = E(x_0) = 0$. Then, substituting (2.2.14) in (2.2.16) and using (2.2.13), the recurrence equation (2.2.18) directly will be done. Here

$$P_0 = E(e_0 e_0^T) = E(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T = E x_0 x_0^T,$$

since $\hat{x}_0 = 0$.

Moreover, equation (2.2.18) is called a discrete Riccati equation.

## 2.3 Continuous KF

### 2.3.1 Introduction and problem formulation

In this case, the time parameter changes continuously and the filtering problem (2.2.1)-(2.2.2) could be modified to the following linear dynamical system disturbed by noise processes:

$$\begin{cases} \acute{x}(t) = A(t)x(t) + B(t)w(t), t > 0, \\ x(0) = x_0, \end{cases} \tag{2.3.1}$$

and

$$z(t) = H(t)x(t) + v(t), t \geq 0, \tag{2.3.2}$$

where $A$, $B$ and $H$ are continuous matrix-valued functions on $[0,\infty)$ and $x_0$ is a Gaussian random vector with zero mean, and $w$ and $v$ are independent vector-valued WN's. Similar to the discrete case $v$, $w$ and $x_0$ are pairwise independent. Consider that $w$ and $v$ are normally distributed with mean zero and covariance equal one.

The linear dynamical filtering problem (2.3.1)-(2.3.2) yields finding the formulae for the best estimate $\hat{x}(t)$ depending on the measurements $z(r), 0 < r < t$. Again, similar to the discrete case, if $\hat{x}(t)$ exists, then

$$\hat{x}(t) = E\big(x(t)\big|z(r)\big), 0 < r < t.$$

### 2.3.2 Derivation of continuous KF

Starting from the discrete KF, the continuous KF can be deduced considering the following sequence $0 = t_0 < t_1 < \ldots < t_{k+1} = t$ as a partition of the time interval $[0, t]$, this partition consists of $k + 1$ subintervals of the same length $\Delta t$, considering that $k$ is sufficiently large.

Start by assuming that $x(t_k) = x_k$, then by (3.2.1) $x_{k+1}$ can be expressed as

$$A_k = e^{\int_{t_k}^{t_{k+1}} A(s)ds} \text{ and } B_k w_k = \int_{t_k}^{t_{k+1}} e^{\int_{t_k}^{t_{k+1}} A(s)ds} B(r)w(r)dr.$$

Matrix exponent formula implies

$$A_k \approx I = \int_{t_k}^{t_{k+1}} A(s)ds \approx I + \Delta t A(t_k),$$

note that, any term containing a power of $\Delta t$ more than one is neglected.

Substituting the value of $A_k$ in the state equation implies

$$x_{k+1} \approx \big(I + \Delta t A(t_k)\big)x_k + B_k w_k\,, \tag{2.3.3}$$

where $\{Bw_k\}$ is a sequence of independent GRV's with mean equals to zero. Next, consider

$$\begin{aligned}
B_k B_k^T &= B_k E(w_k w_k^T) B_k^T = E(B_k w_k)(B_k w_k)^T \\
&= \int_{t_k}^{t_{k+1}} e^{\int_r^{t_{k+1}} A(s)ds} B(r)B(r)^T e^{\int_r^{t_{k+1}} A(s)^T ds} dr \\
&\approx \int_{t_k}^{t_{k+1}} (I + (t_{k+1} - r)A(r))B(r)B(r)^T(I + (t_{k+1} - r)A^T(r))dr \\
&\approx \Delta t(I + \Delta t A(t_k))B(t_k)B(t_k)^T(I + \Delta t A^T(t_k)) \\
&= B(t_k)B(t_k)^T\{\Delta t + (\Delta t)^2 A(t_k) + (\Delta t)^2 A^T(t_k) \\
&\quad + (\Delta t)^3 A(t_k)A^T(t_k)\} \approx \Delta t B(t_k)B(t_k)^T,
\end{aligned}$$

where the terms containing a second or higher order of $\Delta t$ are neglected.

Applying Mean Value theorem for integrals on the measurements $z_k$ implies

$$z_k = \frac{1}{\Delta t}\int_{t_k}^{t_{k+1}} z(r)dr,$$

and considering $H_k = H(t_k)$ one may obtain

$$z_k = \frac{1}{\Delta t}\int_{t_k}^{t_{k+1}} H(r)x(r)dr + \frac{1}{\Delta t}\int_{t_k}^{t_{k+1}} v(t)dt,$$

or,

$$z_k = H_k x_k + \Psi_k v_k, \tag{2.3.4}$$

knowing that

$$\Psi_k v_k = \frac{1}{\Delta t}\int_{t_k}^{t_{k+1}} v(t)dt,$$

and $\{\Psi_k v_k\}$ is a a sequence of independent GRV's of mean equals to zero i.e. $E(\Psi_k v_k) = 0$ for all $k$. And

$$\Psi_k \Psi_k^T = \Psi_k E(v_k v_k^T)\Psi_k^T = E(\Psi_k v_k)(\Psi_k v_k)^T$$

$$= \frac{1}{(\Delta t)^2} E\left[\int_{t_k}^{t_{k+1}} v(r)v^T(r)dr\right]$$

$$= \frac{1}{(\Delta t)^2}\int_{t_k}^{t_{k+1}} v(r)v^T(r)dr$$

$$= \frac{1}{(\Delta t)^2}\int_{t_k}^{t_{k+1}} Idr = \frac{\Delta t}{(\Delta t)^2} = \frac{1}{\Delta t}.$$

It is clear that $\Psi_k$ is invertible since $\Psi_k = \frac{1}{\sqrt{\Delta t}}$. Hereby, the continuous KF problem (2.3.1)-(2.3.2) has been modified to the discrete case and theorem (3.2.1) implies

$$P_{k+1} = \big(I + \Delta t A(\mathrm{t}_k)\big)\mathrm{P}_k\big(I + \Delta t A(\mathrm{t}_k)\big)^T + \Delta t B(t_k)B(t_k)^T$$

$$- \big(I + \Delta t A(\mathrm{t}_k)\big)\mathrm{P}_k H_k^T(\mathrm{H}_k\mathrm{P}_k H_k^T + \Delta t^{-1}I)^{-1}\mathrm{H}_k\mathrm{P}_k\big(I + \Delta t A(\mathrm{t}_k)\big)^T,$$

implying (2.3.5)

$$P_{k+1} = P_k + \Delta t A(t_k)P_k + \Delta t P_k A(t_k)^T + (\Delta t)^2 A(t_k)P_k A(t_k)^T + \Delta t B(t_k)B(t_k)^T$$

$$- \Delta t P_k H_k^T(\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1}\mathrm{H}_k P_k$$

$$- (\Delta t)^2 P_k H_k^T(\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1}\mathrm{H}_k P_k A(t_k)^T$$

$$- (\Delta t)^2 A(t_k)P_k H_k^T(\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1}\mathrm{H}_k P_k$$

$$- (\Delta t)^3 A(t_k)P_k H_k^T(\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1}\mathrm{H}_k P_k A(t_k)^T.$$

Moreover, for $\hat{x}_{k+1}$,

$\hat{x}_{k+1} = \left(I + \Delta t A(\mathrm{t}_k)\right)\left(\hat{x}_k + P_k H_k^T (\mathrm{H}_k \mathrm{P}_k H_k^T + \Delta t^{-1} I)^{-1} (z_k - \mathrm{H}_k \hat{x}_k)\right) = \hat{x}_k +$

$$\Delta t A(t_k)\hat{x}_k + \left(I + \Delta t A(\mathrm{t}_k)\right) P_k H_k^T (\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1} (z_k - \mathrm{H}_k \hat{x}_k).$$

(2.3.6)

**Theorem 2.3.2** Under the previous conditions, the best estimate $\hat{x}(t)$ of $x(t)$ depending on the measurements $z(r), 0 \leq r \leq t$ which described in the continuous KF problem (2.3.1)-(2.3.2) must satisfy the following equation

$$\begin{cases} \dot{\hat{x}}(t) = A(t)\hat{x}(t) + P(t)H(t)^T(z(t) - H(t)\hat{x}(t)), t > 0, \\ \hat{x}(0) = 0, \end{cases} \tag{2.3.7}$$

where $P(t)$ is the covariance of the error and it satisfies the following equation

$$\begin{cases} \dot{P}(t) = A(t)P(t) + P(t)A(t)^T + B(t)B(t)^T - P(t)H(t)^T H(t)P(t), t > 0, \\ P(0) = P_0 = E x_0 x_0^T. \end{cases} \tag{2.3.8}$$

**Proof.** From (2.3.5) one can easily obtain

$$\begin{aligned} \frac{P_{k+1} - P_k}{\Delta t} &= A(t_k)P_k + P_k A(t_k)^T + \Delta t A(t_k) P_k A(t_k)^T + B(t_k)B(t_k)^T \\ &- P_k H_k^T (\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1} \mathrm{H}_k P_k \\ &- \Delta t P_k H_k^T (\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1} \mathrm{H}_k P_k A(t_k)^T \\ &- \Delta t A(t_k) P_k H_k^T (\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1} \mathrm{H}_k P_k \\ &- (\Delta t)^2 A(t_k) P_k H_k^T (\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1} \mathrm{H}_k P_k A(t_k)^T. \end{aligned}$$

Considering that $\lim_{\Delta t \to 0} P_k = P(t)$. Take the limit of the last formulae as $\Delta t$ tends to zero for both sides, then the equation (2.3.7) will be clear.

Also, (2.3.6) implies

$$\frac{\hat{x}_{k+1} - \hat{x}_k}{\Delta t} = A(t_k)\hat{x}_k + \left(I + \Delta t A(\mathrm{t}_k)\right) P_k H_k^T (\Delta t \mathrm{H}_k P_k H_k^T + I)^{-1} (z_k - \mathrm{H}_k \hat{x}_k),$$

letting $\lim_{\Delta \to 0} \hat{x}_{k+1} = \hat{x}_k$. Now, taking the limit of both sides as $\Delta t$ tends to zero yields the equation (2.3.8).

The formulae (2.3.7 - (2.3.8) gives the KF equations for the continuous case. Equation (2.3.8) is called a continuous time Riccati equation.

## 2.4 Selected Generalizations of the KF

In this section some of the generalizations of the KF will be considered. Such generalizations obtained and derived based on different types of noises corrupting the state and the measurement and on the relation between those noises.

### 2.4.1 The KF When the State and Measurements Noises are White and Correlated

The Kalman formulae in theorem 3.3.2 were presented to describe the solution of the KF when the state and the measurements are disturbed by independent WN's. Independency of the WN's means that the source that affect the state and measurements systems are independent, but reality attains that may the same source affect both systems which means that the noises are correlated. KF has a generalization to this case. For this, the following systems describe this case

$$\begin{cases} x'(t|) = A(t)x(t) + B(t)w(t), t > 0, \\ x(0) = x_0, \end{cases} \tag{2.4.1}$$

$$z(t) = H(t)x(t) + v(t), t \geq 0, \tag{2.4.2}$$

where $A$, $B,$ and $H$ are matrix-valued functions on $[0,\infty)$, $x_0$ is a Gaussian random vector with zero mean, $w$ and $v$ are WN processes, all of them of respective dimensions. Also, consider that $x_0$ and $\{v, w\}$ are independent and $w$ and $v$ are correlated, for this

$$\rho = Ew(t)v(t)^T = \Re\delta(t),$$

where $\rho$ is the correlation coefficient of $w$ and $v$. Note that, the correlation coefficient is same as the covariance since both of $w$ and $v$ are WN processes. The continuous KF with independent noises is a particular case considering that

$$\Re = 0,$$

in this case, it is assumed that $\Re$ is nonzero in general and $w(t) = v(t)$. This attains that

$$\rho = Ew(t)v(t)^T = \Re\delta(t) = I\delta(t).$$

KF formulae for this case can be derived in a similar manner to the continuous Kalman case with independent WN processes. Hereby, the best estimate $\hat{x}(t)$ is presented as a solution for the following equation

$$\begin{cases} \hat{x}'(t) = A(t)\hat{x}(t) + (P(t)H(t)^T + B(t))(z(t) - H(t)\hat{x}(t)), t > 0, \\ \hat{x}(0) = 0, \end{cases} \tag{2.4.3}$$

and the covariance of the error is given as a solution for the following equation

$$\begin{cases} P'(t) = A(t)P(t) + P(t)A(t)^T + B(t)B(t)^T \\ \qquad (P(t)H(t)^T + B(t))(H(t)P(t) + B(t)^T), t > 0 \\ P(0) = P_0 = Ex_0x_0^T. \end{cases} \tag{2.4.4}$$

### 2.4.2 The KF When the Signal Noise is Colored

If the real noises are sufficiently close to WN the previous model and formulae work with certain adequacy. Generally, the real noises can significantly be away from the WN, in fact, engineers proved that the noises in reality are never white [9]. This means that the resulting estimation from the previous model is not always the best one, even the error might be very large. For this issue, and to reduce this error and give the best estimate under such kind of noises, engineers introduced the concept of colored noise [8] as it was defined in the previous chapter. Recall that colored noise was defined as a solution of stochastic DE with additive WN.

Now, consider the following state and measurements systems, respectively

$$\begin{cases} x'(t) = A(t)x(t) + \varphi(t), t > 0, \\ x(0) = x_0, \end{cases}$$

and

$$z(t) = H(t)x(t) + w(t), t \geq 0,$$

where $\varphi$ is the colored noise corrupting the state and it is defined as a solution of the following stochastic DE

$$\begin{cases}\varphi'(t) = F(t)\varphi(t) + \Phi(t)w(t), t > 0,\\ \varphi(0) = 0,\end{cases}$$

where $w$ is WN.

Now, combining the state formula with its noise formula, the following new model can be introduced

$$\begin{cases}\tilde{x}'(t) = \tilde{A}(t)\tilde{x}(t) + \tilde{\Phi}(t)w(t), t > 0,\\ \quad \tilde{x}(0) = \tilde{x}_0,\end{cases}$$

and

$$z(t) = \tilde{H}(t)\tilde{x}(t) + w(t), t \geq 0,$$

with

$$\tilde{x}(t) = \begin{bmatrix} x(t) \\ \varphi(t) \end{bmatrix}, \tilde{x}_0 = \begin{bmatrix} x_0 \\ 0 \end{bmatrix},$$

the new matrix-valued functions are given by

$$\tilde{A}(t) = \begin{bmatrix} A(t) & I \\ 0 & F(t) \end{bmatrix}, \tilde{\Phi}(t) = \begin{bmatrix} 0 \\ \Phi(t) \end{bmatrix}, \tilde{\mathrm{H}}(t) = [H(t) \quad 0].$$

Hereby, we modified the system with colored state noise to a system with correlated WN's. For this, the best estimate $\hat{\tilde{x}}(t)$ is the solution of the following equation

$$\begin{cases}\hat{\tilde{x}}'(t) = \tilde{A}(t)\hat{\tilde{x}}(t) + \left(\tilde{P}(t)\tilde{H}(t)^T + \tilde{\Phi}(t)\right)\left(z(t) - \tilde{\mathrm{H}}(t)\hat{\tilde{x}}(t)\right), t > 0,\\ \hat{\tilde{x}}(0) = 0,\end{cases}$$

and the error covariance is a solution of the following Riccati equation

$$\begin{cases}\tilde{\mathrm{P}}'^{(t)} = \tilde{\mathrm{P}}(t)\tilde{\mathrm{A}}(t)^T + \tilde{A}(t)\tilde{P}(t) + \tilde{\Phi}(t)\tilde{\Phi}(t)^T - \left(\tilde{\mathrm{P}}(t)\tilde{\mathrm{H}}(t)^T + \tilde{\Phi}(t)\right)\left(\tilde{\mathrm{H}}(t)\tilde{\mathrm{P}}(t) + \tilde{\Phi}(t)^T\right),\\ \tilde{\mathrm{P}}(0) = cov\tilde{x}_0.\end{cases}$$

These equations can be modified to system of equations using the following manner.

Let

$$\hat{\hat{x}}(t) = \begin{bmatrix} \hat{x}(t) \\ \psi(t) \end{bmatrix}, \widetilde{P}\,(t) = \begin{bmatrix} P(t) & K(t) \\ K(t)^T & S(t) \end{bmatrix}.$$

It is clear that $\tilde{P}(t)$ is symmetric. Then based on the equations for $\hat{\hat{x}}(t)$ and $\tilde{P}(t)$, the equation:

$$\begin{cases} \hat{x}'(t) = A(t)\hat{x}(t) + \psi(t) + P(t)H(t)^T\big(z(t) - H(t)\hat{x}(t)\big), \ \ t > 0, \\ \hat{x}(0) = 0, \end{cases}$$

presents the best estimate $\hat{x}(t)$, with $\psi(t)$, satisfying

$$\begin{cases} \psi'(t) = F(t)\psi(t) + \Big(K(t)^T H(t)^T + \Phi(t)\big(z(t) - H(t)\hat{x}(t)\big)\Big), t > 0, \\ \psi(0) = 0, \end{cases}$$

and equation for $P(t)$ is

$$\begin{cases} P'(t) = P(t)A(t)^T + A(t)P(t) + K(t) + K(t)^T - P(t)H(t)^T H(t)P(t), \ \ t > 0, \\ P(0) = cov\ x_0, \end{cases}$$

with $K(t)$ is given as a solution of the following equation:

$$\begin{cases} K'^{(t)} = K(t)F(t)^T + A(t)K(t) + S(t) - P(t)H(t)^T(H(t)K(t) + \Phi(t)^T, t > 0), \\ K(0) = 0, \end{cases}$$

and $S(t)$, satisfying

$$\begin{cases} S'(t) = S(t)F(t)^T + F(t)S(t) + \Phi(t)\Phi(t)^T \\ \qquad\qquad -\big(K(t)^T H(t)^T + \Phi(t)\big)(H(t)K(t) + \Phi(t)^T), t > 0, \\ S(0) = 0. \end{cases}$$

### 2.4.3 KF When the Measurements Noise is Colored

The observation system may be corrupted by colored noise. The KF problem for this case can be presented for the signal system

$$\begin{cases} x'(t) = A(t)x(t) + B(t)w(t), \ \ t > 0, \\ x(t) = x_0, \end{cases}$$

and the measurements system

$$z(t) = H(t)x(t) + \varphi(t) + w(t), t \geq 0,$$

where $\varphi(t)$ is again a colored noise given as a solution of the equation

$$\begin{cases} \varphi'(t) = F(t)\varphi(t) + \Phi(t)w(t), \ \ t > 0, \\ \varphi(0) = 0. \end{cases}$$

Note that, the appearance of the summation of WN and colored noises in the measurements system is accepted by the nature of the KF which always assumes a non-degenerate WN in the measurements system.

Next, a new signal process $\tilde{x}(t)$ can be introduced

$$\tilde{x}(t) = \begin{bmatrix} x(t) \\ \varphi(t) \end{bmatrix}, \tilde{x}_0 = \begin{bmatrix} x_0 \\ 0 \end{bmatrix},$$

and new matrix-valued functions

$$\tilde{A}(t) = \begin{bmatrix} A(t) & 0 \\ 0 & F(t) \end{bmatrix}, \widetilde{\Phi}(t) = \begin{bmatrix} B(t) \\ \Phi(t) \end{bmatrix}, \widetilde{H}(t) = [H(t) \quad I] .$$

Then the modified signal and observation systems can be presented as

$$\begin{cases} \tilde{x}'(t) = \tilde{A}(t)\tilde{x}(t) + \widetilde{\Phi}(t)w(t), \;\; t > 0. \\ \tilde{x}(0) = \tilde{x}_0, \end{cases}$$

and

$$z(t) = \widetilde{H}(t)\tilde{x}(t) + w(t), t \geq 0.$$

Hereby, the KF problem is reduced to a problem with correlated WN's. Then the best estimate $\hat{\tilde{x}}$ can be presented as a solution of the following equation

$$\begin{cases} \hat{\tilde{x}}'(t) = \tilde{A}(t)\hat{\tilde{x}}(t) + \left(\tilde{P}(t)\widetilde{H}(t)^T + \widetilde{\Phi}(t)\right)\left(z(t) - \widetilde{H}(t)\hat{\tilde{x}}(t)\right), t > 0, \\ \hat{\tilde{x}}(0) = 0, \end{cases}$$

where $\tilde{P}(t)$ is a solution of the following Riccati equation

$$\begin{cases} \tilde{P}'(t) = \tilde{P}(t)\tilde{A}(t)^T + \tilde{A}(t)\tilde{P}(t) + \widetilde{\Phi}(t)\widetilde{\Phi}(t)^T \\ \qquad\qquad - \left(\tilde{P}(t)\widetilde{H}(t)^T + \widetilde{\Phi}(t)\right)\left(\widetilde{H}(t)\tilde{P}(t) + \widetilde{\Phi}(t)^T\right), t > 0, \\ \tilde{P}(0) = cov\, \tilde{x}_0 . \end{cases}$$

Letting

$$\hat{\tilde{x}}(t) = \begin{bmatrix} \hat{x}(t) \\ \psi(t) \end{bmatrix}, \tilde{P}(t) = \begin{bmatrix} P(t) & K(t) \\ K(t)^T & S(t) \end{bmatrix},$$

now, $\hat{x}(t)$ can be obtained from

$$\begin{cases} \hat{x}'(t) = A(t)\hat{x}(t) + \big(P(t)H(t)^T + K(t) + B(t)\big)\big(z(t) - H(t)\hat{x}(t) - \psi(t)\big), \;\; t > 0, \\ \hat{x}(0) = 0, \end{cases}$$

with $\psi(t)$ satisfying

$$\begin{cases} \psi'(t) = F(t)\psi(t) + \big(K(t)^T H(t)^T + S(t) + \Phi(t)\big)\big(\big(z(t) - H(t)\hat{x}(t) - \psi(t)\big), t > 0, \\ \psi(0) = 0, \end{cases}$$

and the covariance of the error $P(t)$ satisfy

$$\begin{cases} P'(t) = \;\; P(t)A(t)^T + A(t)P(t) + B(t)B(t)^T \\ \qquad\qquad -(P(t)H(t)^T + K(t) + B(t))(H(t)P(t) + K(t)^T + B(t)^T) \quad , t > 0, \\ P(0) = cov x_0 = P_0, \end{cases}$$

where $K(t)$ and $S(t)$ satisfy the following tow equations, respectively,

$$\begin{cases} K'(t) = K(t)F(t)^T + A(t)K(t) + B(t)\Phi(t)^T \\ \qquad\qquad -\big(P(t)H(t)^T + K(t) + B(t)\big)(H(t)K(t) + S(t)^T + \Phi(t)^T), t > 0, \\ K(0) = 0, \end{cases}$$

and

$$\begin{cases} S'(t) = S(t)F(t)^T + F(t)S(t) + \Phi(t)\Phi(t)^T \\ \qquad\qquad -\big(K(t)^T H(t)^T + S(t) + \Phi(t)\big)(H(t)K(t) + S(t)^T + \Phi(t)^T), t > 0, \\ S(0) = 0. \end{cases}$$

### 2.4.4 KF When the Signal and Measurements Noises are Colored

Consider the following signal system

$$\begin{cases} x'(t) = A(t)x(t) + \varphi_1(t) + B(t)w(t), t > 0, \\ x(0) = x_0, \end{cases}$$

where $\varphi_1(t)$ stands for colored noise affects the signal system, satisfying

$$\begin{cases} \varphi_1'(t) = F_1(t)\varphi_1(t) + \Phi_1(t)w(t), t > 0, \\ \varphi_1(0) = 0, \end{cases}$$

also, consider the measurements system

$$z(t) = H(t)x(t) + w(t) + \varphi_2(t), t \geq 0,$$

where $\varphi_2(t)$ stands for colored noise affects the measurements system, satisfying

$$\begin{cases} \varphi_2'(t) = F_2(t)\varphi_2(t) + \Phi_2(t)w(t), t > 0, \\ \varphi_2(0) = 0. \end{cases}$$

A new signal and measurements systems can be deduced by introducing new signal processes

$$\tilde{x}(t) = \begin{bmatrix} x(t) \\ \varphi_1(t) \\ \varphi_2(t) \end{bmatrix},\ \tilde{x}_0 = \begin{bmatrix} x_0 \\ 0 \\ 0 \end{bmatrix},$$

and matrix-valued functions

$$\tilde{A}(t) = \begin{bmatrix} A(t) & I & 0 \\ 0 & F_1(t) & 0 \\ 0 & 0 & F_2(t) \end{bmatrix},\ \widetilde{\Phi}(t) = \begin{bmatrix} B(t) \\ \varphi_1(t) \\ \varphi_2(t) \end{bmatrix}, \widetilde{\mathrm{H}}(t) = [H(t) \quad 0 \quad I].$$

Then the given signal and measurements systems become as

$$\begin{cases} \tilde{x}'(t) = \tilde{A}(t)\tilde{x}(t) + \widetilde{\Phi}(t)w(t), t > 0, \\ x(0) = 0, \end{cases}$$

and

$$z(t) = \widetilde{\mathrm{H}}(t)\tilde{x}(t) + w(t) + \varphi_2(t), t \geq 0,$$

the new signal and measurements are in the form when the signal and measurements are corrupted by correlated WN's. So, the best estimate $\hat{\tilde{x}}$ is introduced as a solution of

$$\begin{cases} \hat{\tilde{x}}'(t) = \tilde{A}(t)\hat{\tilde{x}}(t) + (\tilde{P}(t)\widetilde{\mathrm{H}}(t)^T + \widetilde{\Phi}(t))(z(t) - \widetilde{\mathrm{H}}(t)\hat{\tilde{x}}(t)), t > 0, \\ \hat{\tilde{x}}(0) = 0, \end{cases}$$

the new covariance of the error $\tilde{P}(t)$ is a solution of the following Riccati equation

$$\begin{cases} \tilde{P}'(t) = \tilde{A}(t)\tilde{P}(t) + \tilde{P}(t)\tilde{A}(t)^T + \widetilde{\Phi}(t)\widetilde{\Phi}(t)^T \\ \qquad\qquad -(\tilde{P}(t)\widetilde{\mathrm{H}}(t)^T + \widetilde{\Phi}(t))(\widetilde{\mathrm{H}}(t)\tilde{P}(t) + \widetilde{\Phi}(t)^T), t > 0, \\ \tilde{P}(0) = E\tilde{x}_0\tilde{x}_0^T = cov\tilde{x}_0. \end{cases}$$

Then, following system of equations can be obtained

$$\hat{\tilde{x}}(t) = \begin{bmatrix} \hat{x}(t) \\ \psi_1(t) \\ \psi_2(t) \end{bmatrix}, \tilde{P}(t) = \begin{bmatrix} P(t) & K_1(t) & K_2(t) \\ K_1(t)^T & S_1(t) & R(t) \\ K_2(t)^T & R(t)^T & S_2(t) \end{bmatrix}.$$

$\hat{x}(t)$ is a solution of

$$\begin{cases} \hat{x}'(t) = A(t)\hat{x}(t) + \psi_1(t) \\ \qquad\quad +(P(t)H(t)^T + K_2(t) + B(t))(z(t) - H(t)\hat{x}(t) - \psi_2(t)), \quad t > 0, \\ \hat{x}(0) = 0, \end{cases}$$

with $\psi_1(t)$ and $\psi_2(t)$ satisfy the following equations

$$\begin{cases} \psi_1'(t) = F_1(t)\psi_1(t) + \left(K_1(t)^T H(t)^T + R(t) + \Phi_1(t)\right)\left(z(t) - H(t)\hat{x}(t) - \psi_1(t)\right), t > 0, \\ \psi_1(0) = 0, \end{cases}$$

and

$$\begin{cases} \psi_2'(t) = F_2(t)\psi_2(t) + \left(K_2(t)^T H(t)^T + S_2(t) + \Phi_2(t)\right)\left(z(t) - H(t)\hat{x}(t) - \psi_2(t)\right), t > 0, \\ \psi_2(0) = 0. \end{cases}$$

Also, the error covariance is given as a solution of the equation

$$\begin{cases} P'(t) = P(t)A(t)^T + A(t)P(t) + K_1(t) + K_1(t)^T + B(t)B(t)^T \\ \qquad\quad -\left(P(t)H(t)^T + K_2(t) + B(t)\right)(H(t)P(t) + K_2(t)^T + B(t)^T), t > 0, \\ P(0) = cov x_0 = P_0, \end{cases}$$

where $S_1(t), S_2(t), K_1(t), K_2(t)$ and $R(t)$ are obtained by the following equations

$$\begin{cases} S_1'(t) = S_1(t)F_1(t)^T + F_1(t)S_1(t) + \Phi_1(t)\Phi_1(t)^T \\ \qquad\quad -\left(\mathrm{K}_1(t)^T H(t)^T + R(t) + \Phi_1(t)\right)(H(t)K_1(t) + R(t)^T + \Phi_1(t)^T), t > 0, \\ S_1(0) = 0, \end{cases}$$

$$\begin{cases} S_2'(t) = S_2(t)F_2(t)^T + F_2(t)S_2(t) + \Phi_2(t)\Phi_2(t)^T \\ \qquad\quad -\left(\mathrm{K}_2(t)^T H(t)^T + S_2(t) + \Phi_2(t)\right)(H(t)K_2(t) + S_2(t)^T + \Phi_2(t)^T), t > 0, \\ S_2(0) = 0, \end{cases}$$

$$\begin{cases} K_1'(t) = K_1(t)F_1(t)^T + A(t)K_1(t) + S_1(t) + B(t) \\ \qquad\quad -\left(P(t)H(t)^T + K_2(t) + B(t)\right)(H(t)K_1(t) + R(t)^T + \Phi_1(t)^T), t > 0, \\ K_1(0) = 0, \end{cases}$$

$$\begin{cases} K_2'(t) = K_2(t)F_2(t)^T + A(t)K_2(t) + R(t) + B(t)\Phi_1(t)^T \\ \qquad\quad -\left(P(t)H(t)^T + K_2(t) + B(t)\right)(H(t)K_2(t) + S_2(t)^T + \Phi_2(t)^T), t > 0, \\ K_2(0) = 0, \end{cases}$$

and

$$\begin{cases} R'(t) = R(t)F_2(t)^T + F_1(t)R(t) + \Phi_1(t)\Phi_2(t)^T \\ \qquad\quad -\left(\mathrm{K}_1(t)^T H(t)^T + R(t) + \Phi_1(t)\right)(H(t)K_2(t) + S_2(t)^T + \Phi_2(t)^T), \\ R(0) = 0, t > 0. \end{cases}$$

### 2.4.5 KF when the systems are of infinite dimensional

Hilbert space concept is very important to investigate the infinite dimensional dynamical systems. This issue is helpful for KF under WBN's as well. Therefore, in this section we assume that $X$ is a Hilbert space.

- Abstract Cauchy Problem

A function $\mathfrak{A}$ from $[0,\infty)$ to $\mathcal{L}(X)$, where $\mathcal{L}(X)$ is the space of bounded linear operators on the space $X$, is said to be a strongly continuous semigroup if it the following conditions are satisfied

1. $\mathfrak{A}(0) = I$,
2. $\mathfrak{A}(t + s) = \mathfrak{A}(t)\mathfrak{A}(s), \forall\, t \geq 0 \; and \; s \geq 0$,
3. $\parallel \mathfrak{A}(t)x - x \parallel \to 0 \; as \; t \to 0, \forall\, x \in X$.

One can associate the linear operator $A: X \to X$ with the strongly continuous semigroup $\mathfrak{A}$ by

$$Ax = \lim_{x \to 0^+} \frac{\mathfrak{A}(t)x - x}{t},$$

for all $x \in X$, such that previous limit exists. Let $\mathfrak{D}$ be the collection of all such $x$, It is proved $\mathfrak{D}$ is dense in $X$ and $A: \mathfrak{D} \to X$ is a closed linear operator. $A$ is said to be infinitesimal generator of the semigroup $\mathfrak{A}$. It is also proved that the following relationship between $\mathfrak{A}$ and $A$ is satisfied

$$\frac{d}{dt}\mathfrak{A}(t)x = A\mathfrak{A}(t)x = \mathfrak{A}(t)Ax, x \in D. t > 0,$$

this can be written as

$$\mathfrak{A}(t)x = x + \int_0^t \mathfrak{A}(s)Axds$$

$$= x + \int_0^t A\mathfrak{A}(s)x\, ds, x \in \mathfrak{D},$$

note that, this integral is in Bochner sense. Based on this the function $x(t) = \mathfrak{A}(t)x_0$

is a solution of the abstract Cauchy problem

$$x'(t) = Ax(t), x(0) = x_0, t > 0,$$

if $x_0 \in \mathfrak{D}$. If $x_0$ belongs to the set $X$, then $x(t) = \mathfrak{A}(t)x_0$ is still defined (since the domain of $\mathfrak{A}$ equals to $X$), but it is possible that $x(t) \notin \mathfrak{D}$ (since $\mathfrak{D} \neq X$ generally, although $\overline{\mathfrak{D}} = X$). In the wider case $x(t) = \mathfrak{A}(t)x_0$ is called a mild solution of the stated abstract Cauchy problem. Depending on this, $\mathfrak{A}$ together with $A$ is also denoted by $e^{At}$ since it reduces to the matrix exponents if $X = \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$.

**Example 1.** If $A \in$, then similar to finite dimensional case it can be shown that

$$\mathfrak{A}(t) = e^{At} = \sum_{n=0}^{\infty} \frac{(At)^n}{n!}.$$

**Example 2.** In this example, the case when the semigroups are strongly but not uniformly continuous is discussed. In theory of PDF's using separation of variables method it is true that the solution of the following heat equation

$$u'_t(t,\theta) = u''_{\theta\theta}(t,\theta), u(0,\theta) = f(\theta), u(t,0) = u(t,1) = 0, 0 \le \theta \le 1, t \ge 0,$$

has the following representation

$$u(t,\theta) = \sum_{k=1}^{\infty} 2e^{k^2\pi^2 t} \sin(k\pi\theta) \int_0^1 f(\gamma) \sin(k\pi\gamma)\, d\gamma.$$

Therefore, letting $X = L_2(0,1)$ (the space of all square integrable real valued functions on the interval [0; 1]) and $A = d^2/d\theta^2$ with

$$\mathfrak{D} = \{h \in L_2(0,1) : h^k \in L_2(0,1), h(0) = h(1) = 0\},$$

then for every $h \in X$

$$[\, e^{At}h](\theta) = \sum_{k=1}^{\infty} 2e^{k^2\pi^2 t} \sin(k\pi\theta) \int_0^1 h(\gamma) \sin(k\pi\gamma)\, d\gamma,$$

where $0 \le \theta \le 1$, $t \ge 0$.

**Example 3.** (Left Translation Semigroub) Consider the one dimensional PDE

$$v_t'(t,\theta) = v_\theta'(t,\theta), v(0,\theta) = f(\theta), v(t,0) = 0, -\varepsilon \leq \theta \leq 0, t \geq 0,$$

where $\varepsilon > 0$ and $v(t,\theta) \in \mathbb{R}^k$. In theory of PDF's it is proved that the solution of the this PDE has the representation

$$v(t,\theta) = \begin{cases} h(\theta+t), \theta+t \leq 0, \\ 0, \qquad\qquad \theta+t > 0. \end{cases}$$

Letting $X = L_2^k(-\varepsilon, 0)$ (the space of all square integrable functions on the interval $[-\varepsilon, 0]$ with values is $\mathbb{R}^k$) and $A = d/d\theta$ with

$$\mathfrak{D} = \{f \in L_2^k(-\varepsilon,0): f' \in L_2^k(-\varepsilon,0), f(0) = 0\},$$

it can be obtained that $\forall f \in X$,

$$[\mathcal{S}(t)f](\theta) = [e^{At}f](\theta) = \begin{cases} f(\theta+t), \theta+t \leq 0, \\ 0, \qquad\qquad \theta+t > 0, \end{cases} \tag{3.4.1}$$

where $-\varepsilon \leq \theta \leq 0, t \geq 0$. This semigroub will be denoted by $\mathcal{S}$.

**Example 4.** (Right Translation Semigroub) It is known that if $A$ generates a strongly continuous semigroup $\mathfrak{U}$ , then $\mathfrak{U}^*$, the semigroup generated by $A^*$ is also strongly continuous, Therefore, the domain of $A = d/d\theta$

$$\mathfrak{D}^* = \{f \in L_2^k(-\varepsilon,0): f' \in L_2^k(-\varepsilon,0), f(-\varepsilon) = 0\},$$

with the adjoint $A^* = -d/d\theta$ from previous example generates the strongly continuous semigroub

$$[\mathcal{S}^*(t)f](\theta) = [e^{A^*t}f](\theta) = \begin{cases} f(\theta-t), \theta-t \leq -\varepsilon, \\ 0, \qquad\qquad \theta-t > -\varepsilon, \end{cases} \tag{2.4.2}$$

where $-\varepsilon \leq \theta \leq 0, t \geq 0$.

Considering the linear SDE

$$x'(t) = Ax(t) + \varphi(t) + Bw(t);\ t > t_0, x(t_0) = x_0, \tag{2.4.3}$$

where $A$ is a closed linear operator on the space X which generates a strongly continuous semigroup $\mathfrak{U}$, $B \in \mathcal{L}(\mathbb{R}^m, X)$, $x_0$ is $X$-valued Gaussian random vector, $\varphi$

is an $X$-valued random process and $w$ is an $\mathbb{R}^m$-valued WN process, $x(t)$ can be written in a variation of constant formula

$$x(t) = \mathfrak{A}(t)x_0 + \int_{t_0}^{t} \mathfrak{A}(t-s)\varphi(s)ds + \int_{t_0}^{t} \mathfrak{A}(t-s)Bw(s)ds, \qquad (2.4.4)$$

it is assumed that both of these integrals exist where the first integral is an ordinary integral of $X$-valued functions (defined at fixed samples and called a Bochner integral), the second entegral is a stochastic integral of $X$-valued random process. The random process $x(t)$ defined by (2.4.4) is called a mild solution of the equation (2.4.3). Similar to non-homogenous case, it is a strong solution, i.e., $x(t) \in \mathfrak{D}$ and equation (2.4.3) holds, if some other conditions valid. By the solution of (2.4.3) it is always denote a mild solution.

- Operator Riccati Equation

The KF needs a Riccati equation in the form (2.4.4). The form of this equation in a Hilbert space is (Note that in the infinite case $H^* = H^T$)

$$\begin{cases} P'(t) = AP(t) + P(t)A^* + BB^* - (P(t)H^* + B)(HP(t) + B^*), \\ P(0) = cov(x_0), t > 0, \end{cases} \qquad (2.4.5)$$

where $A$, $H$ and $B$ are chosen to be time independent to make the issue simpler. Here we assume that $A$ is a closed linear operator in the space $X$ on the dense domain $\mathfrak{D}$, generating a strongly continuous semigroup $\mathfrak{A}$, $B \in \mathcal{L}(\mathbb{R}^m, X)$ and $B \in \mathcal{L}(X, \mathbb{R}^m)$. Generally, the solution of the equation (2.4.5) can not be understood in the ordinary sense since $A$ is an unbounded. Denote by $\mathfrak{D}^*$ the domain of $A^*$ , also, $A^*$ is a densely defined on $X$. Then $P\,(t)$ is called a solution of the equation (2.4.5) in the scalar product sense if $P(0) = covx_0$ and

$$\langle P(t)x, y\rangle = \langle A^*x, P(t)y\rangle + \langle P(t)x, A^*y\rangle$$
$$+ \langle BB^* - (P(t)H^* + B)(HP(t) + B^*)x, y\rangle,$$

$\forall x, y \in \mathfrak{D}^*$ and $t > 0$, where $\langle .,. \rangle$ is a scalar product in $X$. It is proved that under the above conditions there exists a unique solution in the scalar product sense of the operator Riccati equation (2.4.5) and it has the properties $P^*(t) = P(t)$ and $\langle P(t)x, x \rangle \geq 0$ for every $x \in X$ and $t \geq 0$, i.e., $P(t)$ is a nonnegative operator for every $t \geq 0$.

- The KF for Infinite Dimensional Systems

Consider the following signal system

$$\begin{cases} x'(t) = Ax(t) + Bw(t), t > 0, \\ x(0) = x_0, \end{cases} \tag{2.4.6}$$

and the measurements system

$$z(t) = Hx(t) + w(t), t \geq 0, \tag{2.4.7}$$

where $X$ is a Hilbert space, $x_0$ is an $X$-valued Gaussian random vector with zero mean, A is defined densely as a closed linear operator on the Hilbert space $X$ generating the strongly continuous semigroup $\mathfrak{A}$, $H \in \mathcal{L}(X, \mathbb{R}^m)$and $B \in \mathcal{L}(\mathbb{R}^m, X)$ and $w$ is an $\mathbb{R}^m$-valued WN processes, all of them are of respective dimensions. Considering that $x_0$ and $w$ are independent. Then the best estimate $\hat{x}(t)$ is a mild solution of the following equation

$$\begin{cases} \hat{x}'(t) = A\hat{x}(t) + (P(t)H^* + B)\big(z(t) - H\hat{x}(t)\big), t > 0, \\ \hat{x}(0) = 0, \end{cases} \tag{2.4.8}$$

and the error covariance $P(t)$ is given as a solution of the following operator Riccati equation

$$\begin{cases} P'(t) = AP(t) + P(t)A^* + BB^* - (P(t)H^* + B)(HP(t) + B^*), \\ P(0) = cov x_0, t > 0. \end{cases} \tag{2.4.9}$$

For the full proofs and derivation of these formulae they are listed in [36].

# Chapter 3

# KF FOR WBN'S

## 3.1 Introduction to WBN

Basically, the KF was presented for the WN model. Later, engineers have observed that the WN model is an ideal model of noise processes. To deal with the situation the WBN introduced as follows [9]:

"WBN. Suppose that some physical process, if unaffected by random disturbances can be described by a (vector) ordinary DE $d\zeta = b\big(t, \zeta(t)\big)dt$. If, however, such disturbances enter the system in an additive way, then one might take as a model

$$d\zeta = b\big(t, \zeta(t)\big) + v(t),$$

where $v$ is some stationary process with mean $v$ and known autocovariance matrix $R(r)$:

$$R_{ij} = E\big\{v_i(t)v_j(t+r)\big\}, i, j = 1, \dots, n.$$

If R(r) is nearly 0 except in a small interval near $r = 0$, then $v$ is called WBN. WN corresponds to the ideal case when $R_{ij}$ is a constant $a_{ij}$times a Dirac delta function. Then $v(t)dt$ is replaced by $\sigma dw$, where $\sigma$ is a constant matrix such that $\sigma\sigma^T = a, a = a_{ij}$"

**Definition** An $\mathbb{R}^m$-valued random process $\varphi$ on $[0,\infty]$ is called a WBN if for some $\epsilon > 0$,

$$cov(\varphi(t),\varphi(s)) = \begin{cases} \Lambda(t,s), & |t-s| < \varepsilon, \\ 0, & |t-s| \geq \varepsilon, \end{cases}$$

where $\Lambda$ is an $\mathbb{R}^{m\times m}$-valued nonzero function. If $E\varphi(t) = 0$ and $\Lambda(t,s) = \Lambda(t-s)$ for $0 \leq t \leq s$, then $\varphi$ is said to be stationary ( in wide sense). $\Lambda$ is known as an autocovariance function of $\varphi$. Bashirov [36] presents the WBN $\varphi$ under the general conditions in the following form

$$\varphi(t) = \int_{\max(0,t-\varepsilon)}^{t} \Phi(t,\theta - t)w(\theta)d\theta,$$

where $w$ is a WN process and $\Phi$ is a deterministic function depending on the autocovariance function $\Lambda$ and $\varphi$. In addition, $\varphi$ is stationary on $[\varepsilon,\infty)$ if $\Phi$ is independent on the first variable. It is also proved that if $\Phi$ is differentiable and $\Phi(-\varepsilon) = 0$ the WBN given above can be presented and reduced in the following form

$$\varphi(t) = \int_{-\epsilon}^{0} \tilde{\varphi}(t,\theta)d\theta,$$

where $\tilde{\varphi}$ is treated as a solution of the following PDE

$$\frac{\partial\tilde{\varphi}(t,\theta)}{\partial t} = -\frac{\partial\tilde{\varphi}(t,\theta)}{\partial\theta} + \Phi'(\theta)w(t).$$

## 3.2 KF When the Signal Noise is Wide Band

Consider the following signal system

$$\begin{cases} x'(t) = Ax(t) + \varphi(t);\ t > 0, \\ x(0) = x_0, \end{cases} \tag{3.2.1}$$

which is corrupted by the WBN $\varphi(t)$ which is defined by

$$\varphi(t) = \int_{\max(0,t-\epsilon)}^{t} \Phi(t-s)w(s)ds,$$

and the measurements system

$$z(t) = Hx(t) + w(t),\ t \geq 0, \tag{3.2.2}$$

Here consider $\varepsilon > 0$, and $\Phi$ is a differentiable on the interval $[-\varepsilon, 0]$, satisfying $\Phi(-\varepsilon) = 0$. We also take $A$ and $H$ are considered to be constant matrices.

Let

$$\breve{\varphi}(t,\theta) = \int_{\max(0,t-\varepsilon-\theta)}^{t} \Phi'(s - t + \theta) w(s) ds,$$

where$-\varepsilon \leq \theta \leq 0$ and $t \geq 0$. Consider the integral operator $\Gamma$ which is defined as

$$\Gamma g = \int_{-\varepsilon}^{0} g(\theta) d\theta,$$

on the space $L_2^m(-\varepsilon, 0)$ of all square integrable functions on the interval $[-\varepsilon, 0]$ with the values in the signal space, then it is clear that

$$\Gamma \breve{\varphi}(t,.) = \int_{-\epsilon}^{0} \breve{\varphi}(t,\theta) = \varphi(t),$$

while

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right) \breve{\varphi}(t,\theta) = \Phi'(\theta) w(t).$$

The previous equation shows that $\breve{\varphi}$ is, in fact, a colored noise with $\infty$-dimensional. Thus, if we introduce a new $\infty$-dimensional signal process as

$$\breve{x}(t) = \begin{bmatrix} x(t) \\ \breve{\varphi}(t,.) \end{bmatrix}, \quad \breve{x}(0) = \begin{bmatrix} x_0 \\ 0 \end{bmatrix},$$

together with the linear transformations

$$\breve{A} = \begin{bmatrix} A & \Gamma \\ 0 & -d/d\theta \end{bmatrix}, \breve{H} = [H \quad 0], \qquad \breve{\Phi} = \begin{bmatrix} 0 \\ \Phi'(.) \end{bmatrix},$$

we have

$$\begin{cases} \breve{x}'^{(t)} = \breve{A}\breve{x}(t) + \breve{\Phi}(t) w(t), t > 0, \\ \breve{x}(0) = \breve{x}_0, \end{cases} \tag{3.2.3}$$

where $x(t)$ is the signal process of (3.2.1). So, the system (3.2.1) is modified to the system (3.2.2) with $\infty$-dimensional state process.

Letting

$$\hat{\hat{x}}(t) = \begin{bmatrix} \hat{x}(t) \\ \psi(t,.) \end{bmatrix},$$

be the best estimate of $\hat{x}(t)$ on the base of measurements $z(s)$, $0 \le s \le t$, for the system (3.2.2). Then its first component of $\hat{\hat{x}}(t)$ is the best estimate of $x(t)$ depending on the measurements $z(s)$, $0 \le s \le t$ , for the system (3.2.1). Therefore, KF results in Hilbert spaces can be used to obtain equations of the KF for a WBN driven signal system.

Indeed, note that $\breve{P}(t)$ of the parallel Riccati equation has the form

$$\breve{P}(t) = \begin{bmatrix} P(t) & \breve{K}(t) \\ \breve{K}^*(t) & \check{S}(t) \end{bmatrix},$$

where for every $t \ge 0$, $\breve{K}(t)$ is an operator from $L_2$ to the finite-dimensional signal space, $\breve{K}^*(t)$ is its adjoin operator and $\check{S}(t)$ is an operator defined on $L_2^m(-\varepsilon, 0)$. It is given that such operators are integral operators with respective kernels. Let $K(t,\theta)$ and $S(t,\theta,\tau)$ be the antiderivatives of these kernels with zero boundary conditions $K(t,-\varepsilon) = 0$ and $(t,-\varepsilon,\tau) = S(t,\theta,-\varepsilon) = 0$. Then, in terms of $\psi(t,\theta), P(t), K(t,\theta)$ and $S(t,\theta,\tau)$, the equations of KF for WBN signal system (3.2.1) and its WN driven measurements system are

$$\begin{cases} \hat{x}'(t) = A\hat{x}(t) + \psi(t,0) + P(t)H^T\big(z(t) - H\hat{x}(t)\big), -\epsilon \le \theta \le 0, t > 0, \\ \hat{x}(0) = 0, \end{cases}$$

where

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial \theta}\right)\psi(t,\theta) = \big(K(t,\theta)^T H^T + \Phi(\theta)\big)\big(z(t) - H\hat{x}(t)\big), \\ \psi(0,\theta) = \psi(t,-\epsilon) = 0, -\varepsilon \le \theta \le 0, t > 0. \end{cases}$$

Here, $P$ is a solution of the DE

$$\begin{cases} P'(t) = P(t)A^T + AP(t) + K(t,0) + K(t,0)^T - P(t)H^T HP(t), t > 0, \\ P(0) = P_0, \end{cases}$$

where $K$ satisfies

$$\begin{cases}\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}\right)K(t,\theta)=AK(t,\theta)+S(t,0,\theta)-P(t)H^T(HK(t,\theta)+\Phi(\theta)^{\mathrm{T}}),\\ K(0,\theta)=K(t,-\varepsilon)=0,-\varepsilon\le\theta\le 0,t>0,\end{cases}$$

and $S$ satisfies

$$\begin{cases}\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \tau}\right)S(t,\theta,\tau)=\Phi(\theta)\Phi(\theta)^{\mathrm{T}}-\left(K(t,\theta)^TH^T+\Phi(\theta)\right)(HK(t,\tau)+\Phi(\tau)^{\mathrm{T}}),\\ S(0,\theta,\tau)=K(t,-\varepsilon,\tau)=S(t,\theta,-\epsilon)=0,-\varepsilon\le\theta\le 0,-\varepsilon\le\tau\le 0,t>0.\end{cases}$$

So, the Riccati equation in the case of WBN is presented in three equations, the first equation is the decomposition of the Riccati equation for WN model, and other two equations are PDE's resulting from the distributed delay nature of the WBN's. Finally, the error of estimation is given in the equation

$$e(t)=\mathbf{E}\|x(t)-\hat{x}(t)\|^2=\operatorname{tr}P(t),$$

where $\operatorname{tr}P$ is the trace of the matrix $P$ and $\|\cdot\|$ is the Euclidean norm.

## 3.3 KF When the Measurements Noise is Wide Band

Now we will investigate the case when a WBN corrupts the measurements system. Let the signal system be

$$\begin{cases}x'(t)=Ax(t)+Fw(t),t>0,\\ x(0)=x_0,\end{cases}\tag{3.3.1}$$

which is corrupted by the WN, $w$, and let the measurements system be

$$z(t)=Cx(t)+w(t)+\varphi(t),t\ge 0,\tag{3.3.2}$$

which is corrupted by the summation of WN and WBN's. Note that, the nature of KF always assumes a non-degenerate WN in the measurements system. That's why a WBN in addition to WN are considered in the measurements system. Here, as in the previous section $\varphi(t)$ is a WBN defined by

$$\varphi(t)=\int_{max(0,t-\sigma)}^{t}\Phi(s-t)w(s)ds,$$

where $\Phi$ is a differentiable function on $[-\sigma, 0]$with values in the observation space satisfying $\Phi(-\sigma) = 0$. Letting again

$$\breve{\varphi}(t,\gamma) = \int_{\max(0,t-\sigma-\gamma)}^{t} \Phi'(s - t + \gamma)w(s)\mathrm{ds},$$

where $-\sigma \leq \gamma \leq 0$ and $t \geq 0$ and the integral operator $\Gamma$ which is defined as

$$\Gamma g = \int_{-\sigma}^{0} g(\gamma)d\gamma,$$

on the space $L_2^m(-\sigma, 0)$ of all square integrable functions on the interval $[-\gamma, 0]$ with the values in the measurements space, then it is clear that

$$\Gamma\breve{\varphi}(t,.) = \int_{-\sigma}^{0} \breve{\varphi}(t,\gamma) = \varphi(t),$$

while

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \gamma}\right)\breve{\varphi}(t,\gamma) = \Phi'(\gamma)\mathrm{w(t)}.$$

The previous equation shows that $\breve{\varphi}$ is a colored noise with $\infty$-dimensional. Thus, if we introduce a new $\infty$-dimensional signal process as

$$\breve{x}(t) = \begin{bmatrix} x(t) \\ \breve{\varphi}(t,.) \end{bmatrix}, \quad \breve{x}(0) = \begin{bmatrix} x_0 \\ 0 \end{bmatrix},$$

together with the linear transformations

$$\breve{A} = \begin{bmatrix} A & 0 \\ 0 & -d/d\gamma \end{bmatrix}, \qquad \breve{H} = [H \quad \Gamma], \qquad \breve{\Phi} = \begin{bmatrix} B \\ \Phi'(.) \end{bmatrix},$$

we have

$$\begin{cases} \breve{x}'^{(t)} = \breve{A}\breve{x}(t) + \breve{\Phi}(t)w(t), t > 0, \\ \breve{x}(0) = \breve{x}_0, \end{cases} \tag{3.3.3}$$

and

$$z(t) = \breve{H}\breve{x}(t) + w(t), t \geq 0, \tag{3.3.4}$$

where $x(t)$ is the signal process of (3.3.1). So, the system (3.3.1) is modified to the system (3.3.3) with both signal and measurements systems corrupted by WN's with $\infty$-dimensional state process.

Letting

$$\hat{\check{x}}(t) = \begin{bmatrix} \hat{x}(t) \\ \psi(t,.) \end{bmatrix},$$

be the best estimate of $\check{x}(t)$ on the base of measurements $z(s)$, $0 \le s \le t$, then the first component of $\hat{\check{x}}(t)$ is the best estimate of $x(t)$depending on the measurements $z(s), 0 \le s \le t$ . Similar to the case when the signal noise is WBN, KF equations for driven measurements system can be presented as

$$\begin{cases} \hat{x}'(t) = A\hat{x}(t) + (K(t,0) + B + P(t)H^T)(z(t) - H\hat{x}(t) - \psi(t,0)), t > 0, \\ \hat{x}(0) = 0, \end{cases}$$

where

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \gamma}\right)\psi(t,\gamma) = \left(K^T(t,\gamma)H^T + S(t,0,\gamma) + \Phi(\gamma)\right)(z(t) - H\hat{x}(t) - \psi(t,0), \\ \psi(0,\gamma) = \psi(t,-\sigma) = 0, -\sigma \le \gamma \le 0, t > 0. \end{cases}$$

Here, $P$ is a solution of the DE

$$\begin{cases} P'(t) = P(t)A^T + AP(t) + BB^T - (P(t)H^T + K(t,0) + B)(HP(t) + K(t,0)^T + B^T), t > 0, \\ P(0) = P_0, \end{cases}$$

where $K$ satisfies

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \gamma}\right)K(t,\gamma) = AK(t,\gamma) + B\Phi(\gamma)^{\mathrm{T}} \\ \qquad\qquad -(P(t)H^T + K(t,0) + F)(HK(t,\gamma) + S(t,0,\gamma) + \Phi(\gamma)^{\mathrm{T}}), \\ K(0,\gamma) = K(t,-\sigma) = 0, -\sigma \le \gamma \le 0, t > 0, \end{cases}$$

and $S$ satisfies

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \gamma} + \frac{\partial}{\partial \tau}\right)S(t,\gamma,\tau) = \Phi(\gamma)\Phi(\gamma)^{\mathrm{T}} \\ \qquad\qquad -\left(K(t,\gamma)^T H^T + S(t,0,\gamma)^T + \Phi(\gamma)\right)(HK(t,\tau) + S(t,0,\gamma) + \Phi(\tau)^{\mathrm{T}}), \\ \qquad\qquad\qquad , \\ S(0,\gamma,\tau) = K(t,-\sigma,\tau) = S(t,\gamma,-\sigma) = 0, -\sigma \le \gamma \le 0, -\sigma \le \tau \le 0, t > 0. \\ \qquad\qquad\qquad . \end{cases}$$

Here again, the error of estimation is given in the equation

$$e(t) = \mathbf{E}\|x(t) - \hat{x}(t)\|^2 = \operatorname{tr} P(t).$$

## 3.4 KF When the Signal and Measurements Noises are Wide Band

Now we will investigate the case when a WBN corrupts both signal and measurements systems. Let the signal system be

$$\begin{cases} x'(t) = Ax(t) + \varphi_1(t) + Bw(t), t > 0, \\ x(0) = x_0, \end{cases} \tag{3.4.1}$$

which is corrupted by the WBN $\varphi_1(t)$ and $\varphi_1(t)$ is defined by the integral

$$\varphi_1(t) \ = \int_{max(0,t-\epsilon)}^{t} \Phi_1(\theta - t)w(\theta)d\theta,$$

and let the measurements system be

$$z(t) = Cx(t) + w(t) + \varphi_2(t), t \geq 0, \tag{3.4.2}$$

which is corrupted by the WBN $\varphi_2(t)$ defined by

$$\varphi_2(t) \ = \int_{max(0,t-\eta)}^{t} \Phi_2(\gamma - t)w(\gamma)d\gamma,$$

where $\varepsilon > 0$ and $\eta > 0$ are constants and the functions $\Phi_1$ and $\Phi_2$ are differentiable on $[-\varepsilon, 0]$ and $[-\eta, 0]$, respectively, with values in the measurements space satisfying $\Phi_1(-\varepsilon) = 0$ and $\Phi_2(-\eta) = 0$. Letting again

$$\breve{\varphi}_1(t, \theta) = \int_{\max(0,t-\sigma-\theta)}^{t} \Phi'(s - t + \theta)w(s)ds,$$

where$-\varepsilon \leq \theta \leq 0$, $-\eta \leq \gamma \leq 0$ and $t \geq 0$ and the integral operators $\Gamma_1$ and $\Gamma_2$ which are defined as

$$\Gamma_1\breve{\varphi}_1(t,.) = \int_{-\varepsilon}^{0} \breve{\varphi}_1(t,\theta)d\theta = \varphi(t),$$

and

$$\Gamma_2\breve{\varphi}_2(t,.) = \int_{-\eta}^{0} \breve{\varphi}_2(t,\gamma)d\gamma = \varphi(t),$$

on the spaces $L_2^m(-\varepsilon, 0)$ and $L_2^m(-\eta, 0)$ of all square integrable functions on the intervals $[-\epsilon, 0]$ and $[-\eta, 0]$, respectively with the values in the measurements space, then it is clear that

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right) \breve{\varphi}_1(t, \theta) = \Phi'_1(\theta) w(t),$$

and

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \gamma}\right) \breve{\varphi}_2(t, \gamma) = \Phi'_2(\gamma) w(t).$$

Thus, if we introduce a new infinite dimensional signal process as

$$\breve{x}(t) = \begin{bmatrix} x(t) \\ \breve{\varphi}_1(t, .) \\ \breve{\varphi}_2(t, .) \end{bmatrix}, \quad \breve{x}(0) = \begin{bmatrix} x_0 \\ 0 \\ 0 \end{bmatrix},$$

together with the linear transformations

$$\breve{A} = \begin{bmatrix} A & \Gamma_2 & 0 \\ 0 & -d/d\theta & 0 \\ 0 & 0 & -d/d\gamma \end{bmatrix}, \qquad \breve{H} = [H \quad 0 \quad \Gamma_2], \qquad \breve{\Phi} = \begin{bmatrix} B \\ \Phi'_1(.) \\ \Phi'_1(.) \end{bmatrix},$$

we have

$$\begin{cases} \breve{x}'^{(t)} = \breve{A}\breve{x}(t) + \breve{\Phi}(t) w(t), t > 0, \\ \breve{x}(0) = \breve{x}_0, \end{cases} \tag{3.4.3}$$

and

$$z(t) = \breve{H}\breve{x}(t) + w(t), t \geq 0, \tag{3.4.4}$$

where $x(t)$ is the signal process of (3.4.1). So, the system (3.4.1) is modified to the system (3.4.3).

Letting

$$\hat{\breve{x}}(t) = \begin{bmatrix} \hat{x}(t) \\ \psi_1(t, .) \\ \psi_2(t, .) \end{bmatrix},$$

be the best estimate of $\breve{x}(t)$ on the base of measurements $z(s)$, $0 \leq s \leq t$, then the first component of $\hat{\breve{x}}(t)$ is the best estimate of $x(t)$depending on the measurements

$z(s), 0 \leq s \leq t$ . Now, decomposition of the respective Riccati equation is presented in the form

$$\breve{P}(t) = \begin{bmatrix} P(t) & \breve{K}_1(t) & \breve{K}_2(t) \\ \breve{K}_1{}^T(t) & \breve{S}_1(t) & \breve{R}(t) \\ \breve{K}_2{}^T(t) & \breve{R}^T(t) & \breve{S}_2(t) \end{bmatrix}.$$

Again, let $K_1(t,\theta)$, $K_2(t,\gamma)$, $S_1(t,\theta,\tau)$, $S_2(t,\gamma,\sigma)$ and $R(t,\theta,\gamma)$ be the anti-derivatives of the kernels of the integral operators. KF equations for this system can be presented as

$$\begin{cases} \hat{x}'(t) = A\hat{x}(t) + \psi_1(t,0) \\ \qquad\qquad +(K_2(t,0) + B + P(t)H^T)(z(t) - H\hat{x}(t) - \psi_2(t,0)), t > 0, \\ \hat{x}(0) = 0, \end{cases}$$

where

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial \theta}\right)\psi_1(t,\theta) = \left(K_1^T(t,\theta)H^T + R(t,\theta,0) + \Phi_1(\theta)\right)\left(z(t) - H\hat{x}(t) - \psi_2(t,0)\right), \\ \psi_1(0,\theta) = \psi_1(t,-\varepsilon) = 0, -\varepsilon \leq \theta \leq 0, t > 0, \end{cases}$$

and

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial \gamma}\right)\psi_2(t,\gamma) = \left(K_2^T(t,\gamma)H^T + S_2(t,\gamma,0) + \Phi_2(\gamma)\right)\left(z(t) - H\hat{x}(t) - \psi_2(t,0)\right), \\ \psi_2(0,\gamma) = \psi_2(t,-\eta) = 0, -\eta \leq \gamma \leq 0, t > 0. \end{cases}$$

Here, $P$ is a solution of the DE

$$\begin{cases} P'(t) = P(t)A^T + AP(t) + K_1(t,0) + K_2^T(t,0) + BB^T \\ \qquad\qquad -(P(t)H^T + K_2(t,0) + B)\left(HP(t) + K_2^T(t,0) \quad + B^T\right), \\ P(0) = P_0, t > 0, \end{cases}$$

where $K_1$ and $K_2$ satisfies the following equations , respectively,

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial \theta}\right) K_1(t,\theta) = A\, K_1(t,\theta) + \ S_1(t,0,\theta) + F\Phi_1^T(\theta) \\ \qquad\qquad -(P(t)H^T + \ K_2(t,0) + B)\left(H\, K_1(t,\theta) + R(t,\theta,0) + \Phi_1^T(\theta)\right), \\ K_1(0,\theta) = \ K_1(t,-\varepsilon) = 0, -\varepsilon \leq \theta \leq 0, t > 0, \end{cases}$$

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial \gamma}\right) K_2(t,\theta) = A\, K_2(t,\gamma) + \ S_1(t,0,\gamma) + F\Phi_2^T(\theta) \\ \qquad\qquad -(P(t)H^T + \ K_2(t,0) + B)\left(H\, K_2(t,\gamma) + \ S_2(t,\gamma,0) + \Phi_2^T(\gamma)\right), \\ K_2(0,\gamma) = \ K_2(t,-\eta) = 0, -\eta \leq \gamma \leq 0, t > 0, \end{cases}$$

and, $S_1$ and $S_2$ satisfies the following equations, respectively,

$$\begin{cases}\left(\dfrac{\partial}{\partial t}+\dfrac{\partial}{\partial \theta}+\dfrac{\partial}{\partial \tau}\right) S_1(t,\gamma,\tau) = \Phi_1(\theta)\Phi_1^T(\tau) \\ \quad -\left(K_1^T(t,\theta)H^T + R(t,\theta,0) + \Phi_1(\theta)\right) \times (H\, K_1(t,\tau) + R^T(t,\tau,0) + \Phi_1^T(\gamma)), \\ S_1(0,\theta,\tau) = S_1(t,-\varepsilon,\tau) = S_1(t,\theta,-\varepsilon) = 0, -\varepsilon \le \theta \le 0, -\varepsilon \le \tau \le 0, t > 0,\end{cases}$$

$$\begin{cases}\left(\dfrac{\partial}{\partial t}+\dfrac{\partial}{\partial \gamma}+\dfrac{\partial}{\partial \sigma}\right) S_2(t,\gamma,\sigma) = \Phi_2(\gamma)\Phi_2^T(\gamma) \\ \quad -\left(K_2^T(t,\gamma)H^T + S_2(t,\gamma,0) + \Phi_2(\gamma)\right)\left(H\, K_2(t,\sigma) + S_2(t,\sigma,0) + \Phi_2^T(\sigma)\right), \\ , \\ S_2(0,\gamma,\tau) = S_2(t,-\sigma,\tau) = S_2(t,\gamma,-\sigma) = 0, -\eta \le \gamma \le 0, -\eta \le \sigma \le 0, t > 0, \\ .\end{cases}$$

and, finally, $R$ satisfies

$$\begin{cases}\left(\dfrac{\partial}{\partial t}+\dfrac{\partial}{\partial \theta}+\dfrac{\partial}{\partial \gamma}\right) R(t,\theta,\gamma) = \Phi_1(\theta)\Phi_1^T(\gamma) \\ \quad -\left(K_1^T(t,\theta)H^T + R(t,\theta,0) + \Phi_1(\theta)\right)\left(H\, K_1(t,\gamma) + S_2(t,\gamma,0) + \Phi_2^T(\gamma)\right), \\ , \\ R(0,\theta,\gamma) = R(t,-\varepsilon,\gamma) = R(t,\theta,-\eta) = 0, -\eta \le \theta \le 0, -\varepsilon \le \tau \le 0, t > 0. \\ .\end{cases}$$

# Chapter 4

# INVARIANT KF FOR WBN

## 4.1 Introduction

### 4.1.1 Integral Representation of WBN

A WBN is a vector-valued random process $\varphi$ satisfying

$$cov(\varphi_{t+\sigma}, \varphi_t) = \begin{cases} 0, & \sigma \geq \varepsilon, \\ \Lambda_{t,\sigma}, 0 \leq & \sigma \leq \varepsilon, \end{cases}$$

where $cov(.,.)$ is a covariance matrix, $\varepsilon > 0$, and $\Lambda$ is a matrix-valued nonzero function. Here, $cov(.,.)$ and $\Lambda$ are consistent. If the mean of $\varphi(t)$ is zero and $\Lambda$ is free of its first argument i.e ($\Lambda_{t,\sigma} \equiv \Lambda_\sigma$), the WBN $\varphi$ is said to be stationary (in the wide sense).

For motivation of the WBN's and the integral representation for it, we start from the evidence that in some cases a replacement of WBN's by WN's produces mathematical results quite acceptable for applications in reality. What does make these results acceptable? To discuss this question assume that $w$ is a standard Wiener process (Wiener process is a natural model of Brownian motion. It describes a random, but continuous motion of a particle).for simplicity, it is chosen to be one-dimensional. It is known that its derivative does not exist (in the ordinary sense). But it is possible to force the existence of its derivative $w'$ and we call it as a WN. Therefore, informally

$$w'_t = \lim_{\varepsilon \to 0} \frac{w_{t+\varepsilon} - w_t}{\varepsilon}.$$

Here $\varepsilon$ is a positive or negative change of the present time t. Since at the present time only the past of $w$ can be observed let us treat the preceding informal limit as the left limit and get

$$w'_t = \lim_{\varepsilon\to 0^+} \frac{w_{t-\varepsilon} - w_t}{-\varepsilon} = \lim_{\varepsilon\to 0^+} \int_{t-\varepsilon}^{t} \frac{1}{\varepsilon} dw_s.$$

If we let

$$\varphi_t = \int_{t-\varepsilon}^{t} \frac{1}{\varepsilon} dw_s, \qquad (4.1.1)$$

then

$$\Lambda_\sigma = cov(\varphi_{t+\sigma}, \varphi_t) = \mathrm{E}(\varphi_{t+\sigma}\varphi_t) = \frac{I(\varepsilon-\sigma)}{\varepsilon^2} \neq 0 \; if \;\; 0 \le \sigma < \varepsilon\,,$$

and $cov(\varphi_{t+\sigma}, \varphi_t) = 0$ if $\sigma \ge \varepsilon$.Therefore, $\varphi$ is a WBN and $\Lambda$ is its autocovariance function.

This motivates us to consider WBN's in real processes as an "uncompleted derivative" of WP's in the form of (4.1.1). In the cases when $\varepsilon$ is a sufficiently small positive value, $\varphi$ and $w'$ are very close to each other and, respectively, mathematical methods for the WN $w'$ reflect the reality with more or less acceptable accuracy. But for more adequate mathematical results (for serving the issues such as precise tracking satellites for improvement of preciseness of **GPS**), control and filtering results for the WBN $\varphi$ must be developed.

Interestingly, equation (4.1.1) presents a WBN as a distributed delay of a WN. For some reasons, a WN has an effect to the system: the action of a WN at $t-\varepsilon$ continues acting on $[t-\varepsilon, t]$ and then it becomes negligible. Therefore, for study of WBN's in (4.1.1), principle of delay can be uses. Indeed, this is a basic idea of our approach to WBN's. Previously, this was stressed in Bashirov et al [24].

Equation (4.1.1) should be updated in the form

$$\varphi_t = \int_{max(0,t-\varepsilon)}^{t} \frac{1}{\varepsilon} dw_s,$$

because the Wiener process $w$ is observed starting some initial instant that is ordinarily taken to be $0$. The removed parts

$$\int_{t-\varepsilon}^{t} \frac{1}{\varepsilon} dw_s, 0 \le t \le \varepsilon\ ,$$

must be considered as a part of the initial value of a WBN driven system that is normally independent on disturbing noise processes. More generally, the constant integrand in (4.1) can be taken vector-valued and dependent on $t, s$ and also on random sample $\in \Omega$ . In the sequel we consider WBN's in the form

$$\varphi_t = \int_{max(0,t-\varepsilon)}^{t} \Phi_{t,s-t}\, dw_s, t \ge 0, \qquad (4.1.2)$$

where $\varepsilon > 0, w$ is a $k$-dimensional standard Wiener process, and $\Phi$ is an $\mathbb{R}^{n\times k}$-valued function (random or not) on $[0,\infty) \times [-\varepsilon, 0]$. We regard equation (4.1.2) as an integral representation of $\varphi$ and the function $\Phi$ in (4.1.2) as a relaxing (damping) function of $\varphi$.

### 4.1.2 Invariance

Bashirov raised an important problem. In applications WBN can be measured just by autocovariance function which is one of the difficulties of working with it since autocovariance function may give different WBN's. Let $\mathrm{W}(\Lambda)$ be the collection of all WBN's measured by autocovariance function $\Lambda$. Basically, distinct WBN's from $\mathrm{W}(\Lambda)$ create distinct best estimates and distinct optimal controls in the given estimation problems. It also possible that estimations are just depend on $\Lambda$ and independent on WBN's from $\mathrm{W}(\Lambda)$. So, the below problems can be suggested:

(a) Is the $\Lambda$ sufficient for construction of best estimates no matter what is the WBN's from $\mathrm{W}(\Lambda)$?

(b) If it is yes, what are the conditions?

(c) on the other hand, if $\Lambda$ is not sufficient, how the best estimate of the best estimates should be chosen?

(d) If it is difficult to answer the previous questions for $\mathrm{W}(\Lambda)$, is it possible to find $W_0(\Lambda) \subseteq \mathrm{W}\,(\Lambda)$ such that the above questions can be answered?

(e) How reasonable is $W_0(\Lambda)$?

These problems were partially presented and formulated and listed as unsolved.

Hereby, some theoretical and applied arguments somehow describing $W_0(\Lambda)$:

• The set $W_0(\Lambda)$ is wide. In [30] and [31] it is shown that $W_0(\Lambda)$ has infinitely many WBN processes.

• The WBN from $W_0(\Lambda)$ has a natural interpretation: it is distributed delay of WN. So, the WBN's in real systems can be easily understood as an after effect of WN's.

• The WBN's from $W_0(\Lambda)$ are manageable. They can be presented through linear delay SDE's.

•Finally, it is true that the WBN's corrupting the real systems are belonging to $W_0(\Lambda)$ , this is because of:

Consider the Wiener process $w_t$ and consider the ratio $\frac{w_{t+\varepsilon}-w_t}{\varepsilon}$. The previous ratio is a WBN's of the sort $W_0(\Lambda)$. In the ordinary sense the limit of this ratio as $\varepsilon \to 0$ is not defined, but we force it to be and call it a WN. The substitution of WBN's by WN's gives acceptable results. This allow us to stress on the fact that the above ratio can be argued as an " uncompleted derivative" of the Wiener process for sufficient small $\varepsilon > 0$ which behave as a noise process in real systems.

In the next sections, linear filtering problems and LQG for linear signal and measurement systems under WBN's and WN's are considered, respectively. A complete set of equations for the best estimate and the optimal filter in terms of only the autocovariance function are presented, the independence of such results on the relaxing function but dependence on autocovariance function leads to very useful results, these results are named invariant knowing that the autocovariance function must be known. It is important to obtain invariant results in terms of autocovariance function rather than relaxing function because such results construct optimal filters based just on the autocovariance function.

## 4.2 Invariant KF When the Signal Noise is WBN

In this section the first invariant KF is presented. Consider the $n$-dimensional WBN $\varphi: [0,\infty) \times \Omega \to \mathbb{R}^n$ with the autocovariance function

$$\Lambda_\sigma = cov(\varphi_{t+\sigma}, \varphi_t) = \begin{cases} 0, & \sigma \geq \varepsilon, \\ \Lambda_{t,\sigma,} & 0 \leq \sigma < \varepsilon, \end{cases}$$

Where $\varepsilon > 0$ and $\Lambda$ is an $\mathbb{R}^{n\times n}$-valued nonzero function. One can show` that the random process $\varphi$ given by (4.1.2), in which $\Phi$ is an $\mathbb{R}^{n\times k}$-valued relaxing function on $[0,\infty) \times [-\varepsilon, 0]$, and $w$ is a $k$-dimensional standard Wiener process, is an $n$-dimensional WBN with

$$cov(\varphi_{t+\sigma}, \varphi_t) = \int_{max(0,t+\sigma-\varepsilon)}^{t} \mathrm{E}\left(\Phi_{t+\sigma,s-t-\sigma}\Phi_{t,s-t}^T\right) ds,$$

if $0 \leq \sigma < \varepsilon$. If $\Phi$ is nonrandom and depends only on its second argument, then

$$cov(\varphi_{t+\sigma}, \varphi_t) = \int_{max(-t,\sigma-\varepsilon)}^{0} \Phi_{s-\sigma}\Phi_s^T, \qquad (4.2.1)$$

if $0 \leq \sigma < \varepsilon$ ,that is $\varphi$ becomes stationary for $t \geq \varepsilon$.

For a moment, consider a simplest case when $\varphi$ is a one-dimensional WBN and stationary since the instant $\varepsilon$ and having the autocovariance function $\Lambda: [0, \varepsilon] \rightarrow \mathbb{R}$. Then, by (4.2.1), in order to be represented as

$$\varphi_t = \int_{max(0,t-\varepsilon)}^{t} \Phi_{s-t}\, dw_s,$$

where $w$ is also one-dimensional, the function $\Phi: [-\varepsilon, 0] \rightarrow \mathbb{R}$ should satisfy the equation

$$\Lambda_\sigma = \int_{\sigma-\varepsilon}^{0} \Phi_{s-\sigma}\Phi_s\, ds.$$

This is a convolution equation. In Bashirov and Ugural [31, 32], it is known that if $\Lambda_\sigma$ is a $+$ definite and very general conditions hold, then this equation has an infinitely many solutions $\Phi \in L_2(-\varepsilon, 0; \mathbb{R})$, noticing that the $+$definiteness is a defining property of $\Lambda_\sigma$. This result can be extended to many random cases. Therefore, given $\Lambda_\sigma$, there are infinitely many damping functions $\Phi$ and, so, infinitely many WBN in the form of (4.1.2) for the same autocovariance function $\Lambda$.

Hereby, we consider the partially observable linear system

$$\begin{cases} x'_t = Ax_t + \varphi_t, \quad x_0 = \mu, t > 0, \\ dz_t = Bx_t dt + dv_t, z_0 = 0, t > 0, \end{cases} \tag{4.2.2}$$

where $x$ and $z$ are vector-valued signal and observation processes, $A$ and $B$ are matrices, $\varphi \in W_0(\Lambda)$ has an integrable representation (4.1.2)

$$\varphi_t = \int_{max(0,t-\varepsilon)}^{t} \Phi_{t,s-t}\, dw_s, t \geq 0,$$

given that $\Phi$ is a square integrable relaxing function. $\mu$ is a GRV with zero mean, $w$ and $v$ are independent Wiener processes. The signal system in (4.2.2) is given in terms of only derivative while the measurements system is given in terms of differential. By this, we stress on the idea that not like WN's, which are defined as derivatives of WP's and they do not exist in the ordinary sense, WBN's are well defined. Working under these general conditions, the best estimate $\hat{x}$ for the system (4.2.2) is uniquely determined as a solution of

$$\begin{cases} d\hat{x}_t = \left(A\hat{x}_t + \psi_{t,\theta}\right)dt + P_t B^T(dz_t - B\hat{x}_t dt), \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right)\psi_{t,\theta} dt = K_{t,\theta} B^T(dz_t - B\hat{x}_t dt), \\ \hat{x}_0 = 0, \psi_{0,\theta} = \psi_{t,-\epsilon}, -\epsilon \leq \theta \leq 0, t > 0, \end{cases} \tag{4.2.3}$$

where $P$ and $K$ are given as solutions of the following equations, respectively,

$$\begin{cases} P'_t = AP_t + P_t A^T + K_{t,0} + K_{t,0}^T - P_t B^T B P_t, \\ P_0 = cov\mu, t > 0, \end{cases} \tag{4.2.4}$$

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right) K_{t,\theta} = K_{t,\theta}\Lambda^T + \Lambda_{t,-\theta} - S_{t,\theta,0} - K_{t,\theta} B^T B P_t, \\ K_{0,\theta} = K_{t,-\epsilon} = 0, -\epsilon \leq \theta \leq 0, t > 0, \end{cases} \tag{4.2.5}$$

and S is given as a solution of the equation

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta} + \frac{\partial}{\partial r}\right) S_{t,\theta,r} = K_{t,\theta} B^T B K_{t,r}^T, \\ S_{0,\theta,r} = S_{t,-\epsilon,r} = S_{t,\theta,-\epsilon} = 0, -\epsilon \leq \theta, r \leq 0, t > 0, \end{cases} \tag{4.2.6}$$

in addition, the mean square error of the estimation is given as

$$\mathrm{e}_t = \boldsymbol{E}\|\hat{x}_{\boldsymbol{t}} - x_t\|^2 = trP_t.$$

The filter determined by the equations (4.2.3)-(4.2.6) is called a WBN filter. Note that the classical KF consists of just two equations for $\hat{x}$ and $P$ while the WBN filter contains also equations for $\psi$, $K$ and $S$. By solving the second equation of (4.2.3), $\psi$ has the representation

$$\psi_{t,\theta} = \int_{\max(0,t-\theta-\epsilon)}^{t} K_{\tau,\tau-t+\theta} B^T(dz_\tau - B\hat{x}_\tau d\tau),$$

this gives

$$\psi_{t,0} = \int_{\max(0,t-\epsilon)}^{t} K_{\tau,\tau-t} B^T(dz_\tau - B\hat{x}_\tau d\tau).$$

So, $\psi_{t,0}$ sets as a WBN with its relaxing function equals

$$\Psi_{t,\theta} = K_{t+\theta,\theta}B^T.$$

Let the autocovariance function for $\psi_{t,0}$ be denoted by $\Pi$. The function $K$ is an important factor for the relaxing function $\Psi$. The equation (4.2.5) which includes $S_{0,\theta,r}$ is satisfied by $K$. $S$ satisfies the equation (4.2.6) and has the representation

$$S_{t,\theta,\eta} = \int_{\max(0,t-\theta-\epsilon,t-\eta-\epsilon)}^{t} K_{\tau,\tau-t+\theta}B^T B K_{\tau,\tau-t+\eta}^T d\tau.$$

This implies

$$S_{t,\theta,0} = \int_{\max(0,t-\theta-\epsilon)}^{t} K_{\tau,\tau-t+\theta}B^T B K_{\tau,\tau-t}^T d\tau.$$

It is clear now that

$$S_{t,\theta,0} = \Pi_{t,-\theta}.$$

Therefore, the WBN filter (4.2.3)-(4.2.6) works in following manner

- Equation (4.2.6) gives the covariance function of the WBN $\psi_{t,0}$.
- Equation (4.2.5) gives an important factor for the relaxing function of the WBN $\psi_{t,0}$.
- Equation (4.2.4) is a modified Riccati equation from the classic KF.
- The second equation in (4.2.3) gives the WBN $\psi_{t,0}$.
- This makes the first equation in (4.2.3) to be driven by the sum of WN and WBN's. This equation presents the best estimate $\hat{x}$.

The equations (4.2.4)–(4.2.6) are deterministic and can be solved without need of (4.2.3) and the values of $P$ and $K$ stored in computer. The PDEs (4.2.4)–(4.2.6) can be simply solved by numerical methods. Then the WBN noise filter (4.2.4)–(4.2.6) acts as in figure 4.1, in which $\Gamma$ stands for an operator sending a function, on the interval $[-\varepsilon, 0]$ to its value $f_0$.

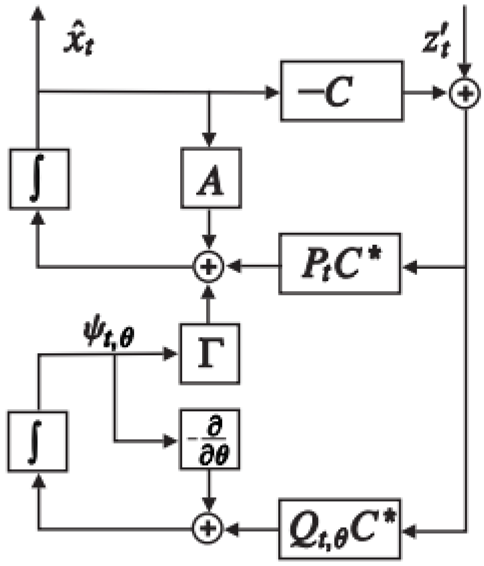


Figure 4.1. The WBN filter

An application of the WBN filter (4.2.3)–(4.2.6) to LQG problem can be presented. Consider LQG problem

$$\mathcal{L}(u) = E\left(\langle x_T, Bx_T\rangle + \int_0^T (\langle x_t, Kx_t\rangle + \langle u_t, Mu_t\rangle)dt\right). \tag{4.2.7}$$

Over the following partially observable system

$$\begin{cases} x'_t = Ax_t + Hu_t + \varphi_t, \quad x_0 = \delta, \ 0 < t \le T, \\ dz_t = Fx_t dt + dv_t, z_0 = 0, \quad 0 < t \le T, \end{cases} \tag{4.2.8}$$

where $E$ denotes the expectation. Assuming that the conditions of the previous filtering problem hold and, moreover, H, B, K and M are matrices in which B and K are considered to be nonnegative and M is positive. After all, the optimal control $u^*$ in the problem (4.2.7)–(4.2.8) is uniquely presented by

$$u_t^* = -G^{-1}H^T\left(\mathcal{V}_t\hat{x}_t^* + \int_t^{\min(T,t+\varepsilon)} \mathcal{Y}_{s,t}^* \mathcal{V}_s\psi_{t,t-s}ds\right), \tag{4.2.9}$$

giving that $\hat{x}_t^*$ is the best estimate of $x_t^*$ which is defined by (4.2.8) and corresponds to the optimal control $u = u^*$, $\psi$ is the associated process, both of them satisfying the following

$$\begin{cases} d\hat{x}_t^* = \left(A\hat{x}_t^* + \psi_{t,0} + Hu_t^*\right)dt + P_t C^T (dz_t^* - F\hat{x}_t^* dt), \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right)\psi_{t,\theta} dt = Q_{t,\theta}^T F^T (dz_t^* - F\hat{x}_t^* dt), \\ \hat{x}_t^* = 0, \psi_{0,\theta} = \psi_{t,-\varepsilon} = 0, -\varepsilon \le \theta \le 0, 0 < t \le T, \end{cases} \tag{4.2.10}$$

$z^*$ is denoting the measurements process which is defined by (4.2.8) and corresponds to the optimal control $u = u^*$, $\mathcal{V}$ is presented as a solution of the following Riccati equation

$$\begin{cases} \mathcal{V}_t' + \mathcal{V}_t A + A^T \mathcal{V}_t + K - \mathcal{V}_t H M^{-1} H^T \mathcal{V}_t = 0, \\ \mathcal{V}_T = B, 0 < t \le T, \end{cases} \tag{4.2.11}$$

$P, Q$ are solutions of the equations (4.2.4)-(4.2.5) and $\mathcal{Y}$ is a bounded perturbation of the transition matrix $e^{At}$ of $A$ by $-HM^{-1}H^*\mathcal{V}_t$. This result is presented in [37], pp. 224–225, for relaxing functions of special form. The previous filter stated before makes the optimal solution valid under given conditions.

Same as the WBN filter, the optimal control (4.2.9)–(4.2.11) is also independent of $\varphi \in W_0(\Lambda)$, just dependent on $\Lambda$. Which means it is also an invariant control result.

## 4.3 KF Filtering When the Signal and the measurements noises are uncorrelated WBN's and WN's, respectively.

In this section linear filtering and optimal control problems are investigated when the signal noise is a WBN, the measurements noise is a WN, and the cost functional is quadratic. Three theorems are proved.

Let us fix the autocovariance function $\Lambda$ and name the collection of all WBN's having the autocovariance function $\Lambda$ by $W(\Lambda)$, this is too wide class. According to the previously integral representation for the WBN, we are interested in those $\varphi \in W(\Lambda)$ which have an integral representation (4.2.2).

Depending on selections of $\Phi$, we can define the following subclasses of $W(\Lambda)$:

- Denote by $W_{L_2^{\mathcal{F}}}(\Lambda)$ the collection of all $\varphi \in W(\Lambda)$ such that $\varphi$ has the representation in (4.2.2) with $\Phi \in C\Big(0, \infty; L_2([-\varepsilon, 0] \times \Omega; \mathbb{R}^{n\times k})\Big)$ such that for all $t \geq 0$ and $max(-t; -\varepsilon) \leq \theta \leq 0, \Phi_{t,\theta}$ is $\mathcal{F}_{t+\theta}$ -measurable, where $\{\mathcal{F}_t\}$ is a complete and continuous filtration generated by $w$. Here the measurability condition surves the existence of stochastic integral in (4.2.2). This class is suitable for a study of control and estimation problems for stochastic systems disturbed by WBN's that are dependent on state or control. In such a way, in Bashirov [38] a stochastic maximum principle is proved for WBN driven nonlinear systems.
- Denote by $W_{W^{1,2}}(\Lambda)$ the collection of all $\varphi \in W(\Lambda)$ such that $\varphi$ has the representation in (4.2.2) with $\Phi \in C\Big(0, \infty; W^{1,2}(-\varepsilon, 0; \mathbb{R}^{n\times k})\Big)$. One can also de ne its subclass $W_{W_0^{1,2}}(\Lambda)$ of all $\varphi \in W_{W^{1,2}}$ with the integral representation in (4.2.2) where $\Phi_{t,-\varepsilon} = 0$. This class was employed in Bashirov [38].
- Denote by $W_{L_2}(\Lambda)$ the collection of all $\varphi \in W(\Lambda)$ such that $\varphi$ has the representation in (4.2.2) with $\Phi \in C(0, \infty; L_2\,(-\varepsilon, 0; \mathbb{R}^{n\times k}\,)\,)$. This class is our concern in this section.

Besides, two more classes can be defined in order to demonstrate that WBN's with integral representation cover white and colored noises as well.

- Denote by $W_{\sigma}(\Lambda)$ the collection of all $\varphi \in W(\Lambda)$ such that $\varphi$ has the representation in (4.2.2) with the relaxing function $\Phi$ in the form

$$\Phi_{t,\theta} = \sum_{i=1}^{m} F_i \delta_{\theta+t-\lambda_{i,t}}\,,$$

where $\delta$ is Dirac's delta-function, $0 \leq \varepsilon_1 < \cdots < \varepsilon_m \leq \varepsilon, \lambda_i$ satisfies the inequalities $t - \varepsilon_i \leq \lambda_{i,t} \leq t$, and $F_i \in \mathbb{R}^{n\times k}$ for all $i = 1, \dots, m$. Then

$$\varphi_t = \int_{max(0,t-\varepsilon)}^{t} \sum_{i=1}^{m} F_i \delta_{s-\lambda_{i,t}} dw_s = \sum_{i=1}^{m} F_i w'_{max(0,\lambda_{i,t})}.$$

Thus $\varphi$ becomes a delayed (multiply and time-dependent) WN. This kind of relaxing functions has been studied in Bashirov et al [28, 29] by approximation of them with relaxing functions from $C\left(0,\infty; W_0^{1,2}(-\varepsilon,0;\mathbb{R}^{n\times k})\right)$.

- Denote by $W_{e^{-At}}(\Lambda)$ the collection of all $\varphi \in W(\Lambda)$ such that $\varphi$ has the representation in (4.2.2) with

$$\Phi_{t,\theta} = e^{-A\theta} F_{t+\theta} ,$$

where $e^{-At}$ is a transition matrix of $-A$. Then

$$\varphi_t = \int_0^t e^{A(t-s)} F_s dw_s, 0 \le t \le \varepsilon ,$$

implying

$$d\varphi_t = A\varphi_t dt + F_t dw_t, \varphi_0 = 0, 0 < t \le \varepsilon.$$

Thus $\varphi$ becomes a colored noise.

Just for simplicity, below we consider filtering and LQG problems for a partially observable stationary linear system in finite-dimensional Euclidean spaces, assuming that the signal noise is WBN and the observation noise is white. A more general case when the signal process takes values in a Hilbert space and the system is non-stationary can be handled with minor changes. The WBN will be assumed to be non-stationary in general because the main object of discussion in this paper is the wide band nature of the signal noise. We will mainly concentrate on linear filtering problem. LQG problem will be considered as an application of the filtering result.

Throughout this section we assume:

$(F)$: $A \in \mathbb{R}^{n\times n}, C \in \mathbb{R}^{m\times n}, w$ and $v$ are $\mathbb{R}^k$ - and $\mathbb{R}^m$-valued standard Wiener processes, $\xi$ is an $\mathbb{R}^n$-valued GRV with $\mathrm{E}\xi = 0, (w, v)$ and $\xi$ are independent, $w$ and $v$ are correlated with $cov(w_t, v_s) = \mathrm{E}\, min(t, s)$.

Note that, $w$ and $v$ are assumed to be correlated just for generality and for a discussion of invariance in this case. The contributions of this section to these problems are in the case $\mathrm{E} = 0$, that is, when $w$ and $v$ are uncorrelated. This is equivalent to their independence because of Gaussian nature of the noises.

Consider the partially observable linear system

$$\begin{cases} x'_t = Ax_t + \varphi_t, \;\; x_0 = \xi, t > 0\,, \\ dz_t = Cx_t dt + dv_t, \;\; z_0 = 0, \quad t > 0, \end{cases} \tag{4.3.1}$$

where $x$ and $z$ are vector-valued signal and observation systems. We also assume:

$(W)$: $\varepsilon > 0$ and $\varphi$ is an $n$-dimensional WBN with the autocovariance function $cov(\varphi_{t+\sigma}, \varphi_t) = \Lambda_{t,\sigma}$ for $t \geq 0$ and $0 \leq \sigma \leq \varepsilon$, so that it has the integral representation in (4.2.2) for some $\Phi \in C(0, \infty; L_2\,(-\varepsilon, 0; \mathbb{R}^{n\times k}\,)\,)$ that is, $\varphi \in W_{L_2}(\Lambda)$.

The filtering problem for the system in (4.3.1) consists of finding equations for the best estimate $\hat{x}_t$ of $x_t$ based on the observations $z_s, 0 \leq s \leq t$, that is, for the conditional expectation $\hat{x}_t = \mathbf{E}(x_t | z_s,\; 0 \leq s \leq t)$.

Note that, the signal system in (4.3.1) is given in terms of derivative while the observation system in terms of differential which is the difference between this section and the previous one. By this, we stress on the fact that unlike WN's, which are generalized derivatives of WP and do not exist in the ordinary sense, WBN's are well-defined random processes. In condition $(W)$, the continuity of $\Phi$ in the first variable is not an essential restriction. It can be replaced by measurability and local boundedness. But here, it is essential for $\Phi$ being $L_2\ (-\varepsilon, 0; \mathbb{R}^{n\times k}\ )$- valued relaxing function.

A wide range of selection of relaxing functions creates a difficulty since the integral representation in (4.2.2), corresponding to given autocovariance function $\Lambda$ , is not unique. As it was mentioned previously, there are infinitely many relaxing functions $\Phi \in C(0, \infty; L_2\ (-\varepsilon, 0; \mathbb{R}^{n\times k}\ )\ )$ for which the WBN $\varphi$, represented in the form of (4.3.2), has the given autocovariance function , that is, $W_{L_2}(\Lambda)$ is an infinite set. This requires making a proper decision about selection one of $\varphi \in W_{L_2}(\Lambda)$ or one of $\Phi \in C(0, \infty; L_2\ (-\varepsilon, 0; \mathbb{R}^{n\times k}\ )\ )$. So, fix one of these relaxing functions and stick to the WBN $\varphi \in W_{L_2}(\Lambda)$, corresponding to this $\Phi$. The method adopted is a derivation of equations for the best estimate $\hat{x}$ for this $\varphi$ and getting these equations independent on $\Phi$ , just dependent on $\Lambda$ . Then all $\varphi \in W_{L_2}(\Lambda)$ became equivalent in the sense that $\hat{x}$ is independent of them, just depends on $\Lambda$.

In the coming theorem we obtain an optimal filter in the filtering problem for the system in (4.3.1) assuming that the WBN $\varphi$ is given by its relaxing function $\Phi$.

**Theorem 4.3.1.** Assuming that the conditions $(F)$ and $(W)$ are satisfied, the best estimate process $\hat{x}$ in the filtering problem for the system in (4.3.1) is uniquely presented as a solution of the equations

$$\begin{cases} d\hat{x}_t = \left(A\hat{x}_t + \psi_{t,0}\right)dt + P_t C^T (dz_t - C\hat{x}_t dt), \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right)\psi_{t,\theta} dt = \left(Q_{t,\theta} C^T + \Phi_{t-\theta,\theta} E\right)(dz_t - C\hat{x}_t dt), \\ \hat{x}_0 = 0, \psi_{0,\theta} = \psi_{t,-\varepsilon} = 0,\ -\varepsilon \le \theta \le 0, t > 0, \end{cases} \tag{4.3.2}$$

where $P, Q$ and $G$ are solutions of:

$$\begin{cases} P_t' = AP_t + P_t A^T + Q_{t,0} + Q_{t,0}^T - P_t C^T C P_t, \\ P_0 = cov\xi, t > 0, \end{cases} \tag{4.3.3}$$

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right) Q_{t,\theta} = Q_{t,\theta} A^T + G_{t,\theta,0} - \left(Q_{t,\theta} C^T + \Phi_{t-\theta,\theta} E\right) C P_t, \\ Q_{0,\theta} = Q_{t,-\varepsilon} = 0,\ -\varepsilon \le \theta \le 0, t > 0, \end{cases} \tag{4.3.4}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta} + \frac{\partial}{\partial \tau}\right) G_{t,\theta,\tau} = \Phi_{t-\theta,\theta} \Phi_{t-\tau,\tau}^T \\ \qquad\qquad - \left(Q_{t,\theta} C^T + \Phi_{t-\theta,\theta} E\right)\left(C Q_{t,\tau}^T + E^T \Phi_{t-\tau,\tau}^T\right), \\ G_{0,\theta,\tau} = G_{t,-\varepsilon,\tau} = G_{t,\theta,-\varepsilon} = 0, -\varepsilon \le \theta \le 0, -\varepsilon \le \tau \le 0\ , t > 0, \end{cases} \tag{4.3.5}$$

Moreover, the mean square error is equal to

$$e_t = \mathbf{E}\|\hat{x}_t - x_t\|^2 = \operatorname{tr} P_t\ .$$

**Proof.** The idea of the proof is as follows. Define the $L_2(-\varepsilon, 0; \mathbb{R}^n)$-valued random process $\phi$ by

$$[\phi_t]_\theta = \int_{max(0,t-\varepsilon-\theta)}^{t} \mathbf{\Psi}_{s,s-t+\theta} dw_s, -\varepsilon \le \theta \le 0, t \ge 0, \tag{4.3.6}$$

where

$$\mathbf{\Psi}_{t,\theta} = \Phi_{t-\theta,\theta}, -\varepsilon \le \theta \le 0, t \ge 0, \tag{4.3.7}$$

one can verify the equality

$$\Gamma_{\phi_t} = [\phi_t]_0 = \varphi_t, \tag{4.3.8}$$

for $\varphi$ defined by (4.2.2), where $\Gamma$ is a linear operator from $W^{1,2}(-\varepsilon, 0; \mathbb{R}^n)$ to $\mathbb{R}^n$, assigning to $h \in W^{1,2}(-\varepsilon, 0; \mathbb{R}^n)$ its value $h_0$ .Let $-d/d\theta$ be a differential operator on $L_2(-\varepsilon, 0; \mathbb{R}^n)$ with the domain

$$D(-d/d\theta) = \{h \in W^{1,2}(-\varepsilon, 0; \mathbb{R}^n): h_{-\varepsilon} = 0\},$$

noticing that $(-d/d\theta)^* = d/d\theta$ and

$$D(d/d\theta) = \{h \in W^{1,2}(-\varepsilon, 0; \mathbb{R}^n): h_0 = 0\}.$$

One can verify that $\phi$ is a mild solution of the linear SDE.

$$d\phi_t = (-d/d\theta)\phi_t dt + \Psi_t dw_t, \phi_0 = 0, t > 0. \tag{4.3.9}$$

Equations (4.3.7)-(4.3.9) lead to the reduction of the linear system in (4.3.1), driven by the WBN $\varphi$, to a linear system, driven by a WN, with an enlarged $\mathbb{R}^n \times L_2(-\varepsilon, 0; \mathbb{R}^n)$-valued signal process. Indeed, letting

$$\tilde{x}_t = \begin{bmatrix} x_t \\ \phi_t \end{bmatrix}, \tilde{\xi}_t = \begin{bmatrix} \xi \\ 0 \end{bmatrix},$$

and

$$\tilde{A} = \begin{bmatrix} A & \Gamma \\ 0 & -d/d\theta \end{bmatrix}, \widetilde{\Phi}_t = \begin{bmatrix} 0 \\ \Psi_t \end{bmatrix}, \tilde{C} = [C \quad 0],$$

we obtain that

$$\begin{cases} d\tilde{x}_t = \tilde{A}\tilde{x}_t dt + \widetilde{\Phi}_t dw_t, \tilde{x}_0 = \tilde{\xi}, t > 0, \\ dz_t = \tilde{C}\tilde{x}_t dt + dv_t,\ z_0 = 0, \quad t > 0, \end{cases} \tag{4.3.10}$$

obviously, the first component of $\hat{\tilde{x}}_t = \mathbf{E}(\tilde{x}_t | z_s,\ 0 \le s \le t)$ is the best estimate $\hat{x}_t$ for the system in (4.3.1). Therefore, it remains to find the equations for $\hat{\tilde{x}}_t$ which will be finalized by the methods of functional analysis.

In (4.3.10), $\tilde{A}$ is a densely defined closed linear operator on $\mathbb{R}^n \times L_2(-\varepsilon, 0; \mathbb{R}^n)$ with

$$D(\tilde{A}) = \mathbb{R}^n \times D(-d/d\theta),$$

generating a strongly continuous semigroup. According to linear filtering theory in Hilbert spaces,] the best estimate process $\hat{\tilde{x}}_t$ is a unique mild solution of the equation

$$\begin{cases} d\hat{\tilde{x}}_t = \tilde{A}\hat{\tilde{x}}_t dt + \left(\tilde{P}_t \tilde{C}^T + \widetilde{\Phi}_t \mathrm{E}\right)\left(dz_t - \tilde{C}\hat{\tilde{x}}_t dt\right), \\ \hat{\tilde{x}}_0 = 0,\ t > 0\,, \end{cases} \tag{4.3.11}$$

where $\tilde{P}$ is a scalar product solution of the operator Riccati equation

$$\begin{cases} \tilde{P}_t' = \tilde{A}\tilde{P}_t + \tilde{P}_t \tilde{A}^T + \widetilde{\Phi}_t \widetilde{\Phi}_t^T - \left(\tilde{P}_t \tilde{C}^T + \widetilde{\Phi}_t \mathrm{E}\right)\left(\tilde{C}\tilde{P}_t + \mathrm{E}^* \widetilde{\Phi}_t^T\right), \\ \tilde{P}_0 = cov\tilde{\xi},\ t > 0, \end{cases} \tag{4.3.12}$$

and

$$\mathbf{E}\left\|\tilde{x}_t - \hat{\tilde{x}}_t\right\|^2 = \operatorname{tr} \tilde{P}_t\,. \tag{4.3.13}$$

Here, the values of $\tilde{P}$ are self-adjoint Hilbert-Schmidt operators on the Hilbert space $\mathbb{R}^n \times L_2(-\varepsilon, 0; \mathbb{R}^n)$. Therefore, we can decompose $\tilde{P}_t$ as

$$\tilde{P}_t = \begin{bmatrix} P_t & \tilde{Q}_t^T \\ \tilde{Q}_t & \tilde{G}_t \end{bmatrix},$$

assuming that $\tilde{Q}_t$ and $\tilde{G}_t$ are linear integral operators from $\mathbb{R}^n$ and $L_2(-\varepsilon, 0; \mathbb{R}^n)$ to $L_2(-\varepsilon, 0; \mathbb{R}^n)$, respectively. Let $Q_{t,\theta}$ and $G_{t,\theta,\tau}$ be respective kernels, that is,

$$[\tilde{Q}_t x]_\theta = Q_{t,\theta,x}, -\varepsilon \le \theta \le 0, t \ge 0, x \in \mathbb{R}^n,$$

and

$$[\tilde{G}_t h]_\theta = \int_{-\varepsilon}^0 G_{t,\theta,\tau} h_\tau d\tau, -\varepsilon \le \theta \le 0, t \ge 0, h \in L_2(-\varepsilon, 0; \mathbb{R}^n).$$

We will deduce the equations for $P, Q$ and $G$ from (15) in the following way.

At first, note that

$$\tilde{A}^* = \begin{bmatrix} A^T & 0 \\ \Gamma^* & d/d\theta \end{bmatrix},$$

where $\Gamma^*$ is understood as

$$\int_{-\varepsilon}^0 \langle \Gamma^* x, h_\theta \rangle d\theta = \langle x, h_0 \rangle, x \in \mathbb{R}^n, h \in D(-d/d\theta).$$

Take arbitrary $(x, g), (y, h) \in \mathbb{R}^n \times D(d/d\theta)$, noticing that $g_0 = h_0 = 0$. Writing (4.3.12) for the component $\tilde{G}$ of $\tilde{P}$, we obtain

$$\tilde{G}_t' = (-d/d\theta)\tilde{G}_t + \tilde{G}_t(d/d\theta) + \Psi_t \Psi_t^T - (\tilde{Q}_t C^T + \Psi_t \mathrm{E})(C\tilde{Q}_t^T + \mathrm{E}^T \Psi_t^T),$$

or in scalar product

$$\langle \tilde{G}_t' g, h \rangle = \langle \tilde{G}_t g, \left(\frac{d}{d\theta}\right) h \rangle + \langle \tilde{G}_t \left(\frac{d}{d\theta}\right) g, h \rangle + \langle \Psi_t \Psi_t^T g, h \rangle$$

$$- \langle (\tilde{Q}_t C^T + \Psi_t \mathrm{E})(C\tilde{Q}_t^T + \mathrm{E}^T \Psi_t^T) g, h \rangle.$$

Here, the terms can be evaluated in the following way:

$$\langle \tilde{G}_t' g, h \rangle = \int_{-\varepsilon}^0 \int_{-\varepsilon}^0 \langle \frac{\partial}{\partial t} G_{t,\theta,\tau} g_\tau, h_\theta \rangle \, d\tau d\theta,$$

$$\langle \tilde{G}_t g, (d/d\theta) h \rangle = \int_{-\varepsilon}^0 \int_{-\varepsilon}^0 \langle G_{t,\theta,\tau} g_\tau, h'_\theta \rangle \, d\tau d\theta,$$

$$= -\int_{-\varepsilon}^{0} \langle G_{t,-\varepsilon,\tau} g_\tau, h_{-\varepsilon}\rangle \, d\tau - \int_{-\varepsilon}^{0}\int_{-\varepsilon}^{0} \langle \frac{\partial}{\partial\theta} G_{t,\theta,\tau} g_\tau, h_\theta\rangle \, d\tau d\theta,$$

$\langle \tilde{G}_t (d/d\theta) g, h\rangle = \int_{-\varepsilon}^{0}\int_{-\varepsilon}^{0} \langle G_{t,\theta,\tau} {g'}_\tau, h_\theta\rangle \, d\tau d\theta$

$$= -\int_{-\varepsilon}^{0} \langle G_{t,\theta,-\varepsilon} g_{-\varepsilon}, h_\theta\rangle \, d\theta - \int_{-\varepsilon}^{0}\int_{-\varepsilon}^{0} \langle \frac{\partial}{\partial\tau} G_{t,\theta,\tau} g_\tau, h_\theta\rangle \, d\tau d\theta,$$

$$\langle \Psi_t \Psi_t^T g, h\rangle = \int_{-\varepsilon}^{0}\int_{-\varepsilon}^{0} \langle \Phi_{t-\theta,\theta} \Phi_{t-\tau,\tau}^T g_\tau, h_\theta\rangle \, d\tau d\theta,$$

$$\langle \widetilde{W}_t \widetilde{W}_t^T g, h\rangle = \int_{-\varepsilon}^{0}\int_{-\varepsilon}^{0} \langle \mathrm{W}_{t,\theta} \mathrm{W}_{t,\tau}^T g_\tau, h_\theta\rangle \, d\tau d\theta,$$

where for brevity we denote

$$\widetilde{W}_t = \tilde{Q}_t C^T + \Psi_t E \text{ and } \mathrm{W}_{t,\theta} = \mathrm{Q}_{t,\theta} C^T + \Phi_{t-\theta,\theta} E.$$

Hence,

$$0 = \int_{-\varepsilon}^{0}\int_{-\varepsilon}^{0} \langle \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial\theta} + \frac{\partial}{\partial\tau}\right) G_{t,\theta,\tau} g_\theta, h_\tau\rangle \, d\tau d\theta$$

$$+ \int_{-\varepsilon}^{0}\int_{-\varepsilon}^{0} \langle \left(\mathrm{W}_{t,\theta} \mathrm{W}_{t,\tau}^T - \Phi_{t-\theta,\theta} \Phi_{t-\tau,\tau}^T\right) g_\theta, h_\tau\rangle \, d\tau d\theta$$

$$+ \int_{-\varepsilon}^{0} \langle G_{t,-\varepsilon,\tau} g_\tau, h_{-\varepsilon}\rangle \, d\tau + \int_{-\varepsilon}^{0} \langle G_{t,\theta,-\varepsilon} g_{-\varepsilon}, h_\theta\rangle \, d\theta.$$

Since $g, h \in D(d/d\theta)$, where $D(d/d\theta)$ is dense in $L_2(-\varepsilon, 0; \mathbb{R}^n)$, we can extend the last equality to all four-tuples $(g_{-\varepsilon}, g, h_{-\varepsilon}, h) \in \mathbb{R}^n \times L_2(-\varepsilon, 0; \mathbb{R}^n) \times \mathbb{R}^n \times L_2(-\varepsilon, 0; \mathbb{R}^n)$, treating $g_{-\varepsilon}$ and $h_{-\varepsilon}$ independently on $g$ and $h$. This implies that $G$ satisfies (8) with the zero initial and boundary conditions. Additionally, we also obtain that $G_{t,\cdot,\theta}, G_{t,\theta,\cdot} \in D(-d/d\theta)$.

In the same way, from (15), we derive the equation for $\tilde{Q}$ as

$$\tilde{Q}_t' = (-d/d\theta)\tilde{Q}_t A^T + \tilde{Q}_t^T \Gamma^T - \left(\tilde{Q}_t C^T + \Psi_t E\right) C \mathrm{P}_t,$$

or in scalar product

$$\langle \tilde{Q}_t' x, h\rangle = \langle \tilde{Q}_t x, (d/d\theta)h\rangle + \langle \tilde{Q}_t A^T x, h\rangle + \langle \Gamma^* x, \tilde{G}_t h\rangle - \langle \big(\tilde{Q}_t C^T + \Psi_t E\big) C \mathrm{P}_t x, h\rangle.$$

Here,

$$\langle \tilde{Q}_t' x, h\rangle = \int_{-\varepsilon}^{0} \langle \frac{\partial}{\partial t} Q_{t,\theta,} x, h_\theta\rangle \, d\theta,$$

$$\langle \tilde{Q}_t x, (d/d\theta)h\rangle = \int_{-\varepsilon}^{0} \langle Q_{t,\theta,} x, h_\theta'\rangle d\theta$$

$$= -\langle Q_{t,-\varepsilon} x, h_{-\varepsilon}\rangle - \int_{-\varepsilon}^{0} \langle \frac{\partial}{\partial \theta} Q_{t,\theta,} x, h_\theta\rangle \, d\theta,$$

$$\langle \tilde{Q}_t A^* x, h\rangle = \int_{-\varepsilon}^{0} \langle Q_{t,\theta,} A^T x, h_\theta\rangle d\theta,$$

$$\langle \Gamma^* x, \tilde{G}_t h\rangle = \langle x, \Gamma \tilde{G}_t h\rangle = \int_{-\varepsilon}^{0} \langle x, G_{t,0,\tau} h_\tau\rangle \, d\tau = \int_{-\varepsilon}^{0} \langle G_{t,0,\theta}^T x, h_\theta\rangle d\theta$$

$$= \int_{-\varepsilon}^{0} \langle G_{t,\theta,0} x, h_\theta\rangle d\theta,$$

$$\langle \widetilde{W}_t C P_t x, h\rangle = \int_{-\varepsilon}^{0} \langle W_{t,\theta,} C P_t x, h_\theta\rangle d\theta.$$

Hence,

$$0 = \int_{-\varepsilon}^{0} \langle \left( \left( \frac{\partial}{\partial t} + \frac{\partial}{\partial \theta} \right) Q_{t,\theta} - Q_{t,\theta} A^T - G_{t,\theta,0} \right) x, h_\theta\rangle \, d\theta$$

$$+ \int_{-\varepsilon}^{0} \langle W_{t,\theta,} C P_t x, h_\theta\rangle d\theta + \langle Q_{t,-\varepsilon} x, h_{-\varepsilon}\rangle.$$

In a similar way we can extend the last equality to all triples $(x, h_{-\varepsilon}, h) \in \mathbb{R}^n \times \mathbb{R}^n \times L_2(-\varepsilon, 0; \mathbb{R}^n)$, treating $h_{-\varepsilon}$ independently on $h$. This implies that $Q$ satisfies (4.3.4) with the zero initial and boundary conditions. Additionally, we obtain $Q_{t,\cdot}, Q_{t,\cdot}^* \in D(-d/d\theta)$.

Next, we concentrate on the equation for $P$. From Eq. (4.3.12), we deduce

$$P_t' = A P_t + P_t A^T + \tilde{Q}_t^T \Gamma^* + \Gamma \tilde{Q}_t - P_t C^T C P_t,$$

or in scalar product

$$\langle P_t' x, y\rangle = \langle P_t x, A^T y\rangle + \langle P_t A^T x, y\rangle + \langle \tilde{Q}_t^T \Gamma^* x, y\rangle + \langle \tilde{Q}_t x, \Gamma^* y\rangle - \langle P_t C^T C P_t x, y\rangle.$$

Here,

$$\left[\tilde{Q}_t x\right]_\theta = Q_{t,\theta} x, -\varepsilon \le \theta \le 0,$$

implying

$$\langle \tilde{Q}_t x, \Gamma^* y\rangle = \langle \Gamma \tilde{Q}_t x, y\rangle = \langle Q_{t,0} x, y\rangle.$$

Similarly,

$$\langle \tilde{Q}_t^T \Gamma^* x, y\rangle = \langle Q_{t,0}^T x, y\rangle.$$

Then

$$\langle \left(P_t' - AP_t - P_t A^T - Q_{t,0}^T - Q_{t,0} + P_t C^T C P_t\right) x, y\rangle = 0.$$

Since $x, y \in \mathbb{R}$ are arbitrary, we obtain the equation in (4.3.3) for $P$.

Now we consider (4.3.11). It produces two equations

$$d\hat{x}_t = A\hat{x}_t dt + \Gamma \psi_t dt + P_t C^T (dz_t - C\hat{x}_t dt),$$

and

$$d\psi_t = (-d/d\theta)\psi_t dt + \left(\tilde{Q}_t C^T + \psi_t E\right)(dz_t - C\hat{x}_t dt),$$

where we let $\psi = \widehat{\phi}$. It is not difficult to see that they produce the system in (4.3.5). Finally, the formula for the error $e_t$of estimation follows from (4.3.13). This completes the proof.

The classic KF contains two equations for $\hat{x}$ and $P$. But the filter from Theorem 4.3.1 includes the associated equations for, $Q$ and $G$. What is the meaning of them in the filter? The solution of the equation (4.3.2) for $\psi$ is presented as

$$\psi_{t,\theta} = \int_{max(0,t-\varepsilon)}^{t} \left(Q_{s,s-t+\theta} C^T + \Phi_{t-\theta,s-t+\theta} E\right)(dz_s - C\hat{x}_s ds),$$

implying

$$\psi_{t,0} = \int_{max(0,t-\varepsilon)}^{t} \big(Q_{s,s-t}C^T + \Phi_{t,s-t}E\big)(dz_s - C\hat{x}_s ds),$$

therefore, $\psi_{t,0}$in (4.3.2) acts as a WBN generated by the innovation process. The function $Q$, together with $C$, $\Phi$ and $E$, takes part in forming the relaxing function of $\psi_{t,0}$. It satisfies (4.3.4), that includes $G_{t,\theta,0}$. From the considerations of the next section it will follow that in fact $G_{t,\theta,0}$is the difference of the covariance functions of the WBN's $\varphi_t$of the system and $\psi_{t,0}$ of the filter. Therefore, the filter from Theorem 4.3.1 works in the following form:

- Equation (4.3.5) produces the difference of covariance functions of the system and filter WBN's.
- Equation (4.3.4) contributes to the relaxing function of the WBN filter.
- Equation (4.3.3) is a classic KF modified to Riccati equation.
- The second of equations in (4.3.2) produces the WBN's filter.
- Finally, the first of equations in (4.3.2) produces the best estimate.

It was mentioned that the WBN's cover delayed WN's as well. For this, the relaxing function $\Phi$ in (4.2.1) should be selected as $\Phi_{t,\theta} = F\delta_{\theta+t-\lambda_t}$ to achieve one single time dependent delay of a WN. Actually, Theorem 4.3.1 does not cover this case. To make the picture more complete, we present the following.

**Theorem 4.3.2**. Under the condition $(F)$ with $k = m$, $w = v$, and $E = I$ and for $\varepsilon > 0$, the best estimate process $\hat{x}$ in the filtering problem for the partially observable linear system

$$\begin{cases} dx_t = Ax_t dt + F dw_{max(0,\lambda_t)}, x_0 = \xi, t > 0, \\ dz_t = Cx_t dt + dw_t, z_0 = 0, t > 0, \end{cases}$$

where $\lambda$ is a strictly increasing differentiable function on $[0,\infty)$ and satisfies $t - \varepsilon \le \lambda_t \le t$ for all $t \ge 0$, is uniquely presented as a solution of the equations

$$\begin{cases} d\hat{x}_t = \left(A\hat{x}_t + \psi_{t,0}\right)dt + P_t C^T(dz_t - C\hat{x}_t dt), \\ \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial \theta}\right)\psi_{t,\theta} dt = Q_{t,\theta} C^T (dz_t - C\hat{x}_t dt), \\ \hat{x}_0 = 0, \psi_{0,\theta} = 0, d\,\psi_{t,t-\lambda_t^{-1}} = F(dz_t - C\hat{x}_t dt), -\lambda_0^{-1} \le \theta \le 0, t > 0, \end{cases}$$

where $P$, $Q$ and $G$ are solutions of

$$\begin{cases} P_t' = AP_t + P_t A^T + Q_{t,0} + Q_{t,0}^T - P_t C^T C P_t, \\ P_0 = cov\xi, t > 0, \end{cases}$$

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial \theta}\right) Q_{t,\theta} = Q_{t,\theta} A^T + G_{t,\theta,0} - Q_{t,\theta} C^* C P_t, \\ Q_{0,\theta} = 0, Q_{t,t-\lambda_t^{-1}} = -FCP_t, -\lambda_0^{-1} \le \theta \le 0, t > 0, \end{cases}$$

and

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial \theta} + \dfrac{\partial}{\partial \tau}\right) G_{t,\theta,\tau} = -Q_{t,\theta} C^T C Q_{t,\tau}^T, \\ G_{0,\theta,\tau} = 0, -\lambda_0^{-1} \le \theta \le 0, -\lambda_0^{-1} \le \tau \le 0\,, \\ G_{t,\theta,t-\lambda_t^{-1}} = -Q_{t,\theta} C^T F^T - FCQ_{t,t-\lambda_t^{-1}}^T, t - \lambda_t^{-1} \le \theta \le 0, t > 0, \\ G_{t,t-\lambda_t^{-1},\tau} = -Q_{t,t-\lambda_t^{-1}} C^T F^T - FCQ_{t,\tau}^T, t - \lambda_t^{-1} \le \tau \le 0, t > 0, \end{cases}$$

moreover, the mean square error of estimation is given by

$$e_t = \mathbf{E}\|\hat{x}_t - x_t\|^2 = \mathrm{tr}\, P_t\,.$$

**Proof.** This theorem is proved in Bashirov [28]. One can verify that Theorem 6.1 can be obtained from Theorem 4.3.2 by substitution $\Phi_{t,\theta} = F\delta_{\theta+t-\lambda_t}$ and using informal integral relations for delta-function.

In the case of independent noises, i.e., $E = 0$, Theorem 4.3.1 produces an exceptional result: the filter from Theorem 4.3.1 becomes independent on relaxing function $\Phi$ , depends just on the autocovariance function $\Lambda$.

**Theorem 4.3.3** Under the conditions $(F)$, $(W)$ and $E = 0$, the best estimate process $\hat{x}$ in the filtering problem for the system in (4.3.1) is uniquely presented as a solution of the equations

$$\begin{cases} d\hat{x}_t = \left(A\hat{x}_t + \psi_{t,0}\right)dt + P_t C^T(dz_t - C\hat{x}_t dt), \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right)\psi_{t,\theta} dt = Q_{t,\theta} C^T(dz_t - C\hat{x}_t dt), \\ \hat{x}_0 = 0, \psi_{0,\theta} = \psi_{t,-\varepsilon} = 0, -\varepsilon \le \theta \le 0, t > 0, \end{cases} \tag{4.3.14}$$

where $P$, $Q$ and $R$ are solutions of

$$\begin{cases} P'_t = AP_t + P_t A^T + Q_{t,0} + Q^T_{t,0} - P_t C^T C P_t, \\ P_0 = cov\xi, t > 0, \end{cases} \tag{4.3.15}$$

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right) Q_{t,\theta} = Q_{t,\theta} A^T + \Lambda_{t,-\theta} - R_{t,\theta,0} - Q_{t,\theta} C^T C P_t, \\ Q_{0,\theta} = Q_{t,-\varepsilon} = 0, -\varepsilon \le \theta \le 0, t > 0, \end{cases} \tag{4.3.16}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta} + \frac{\partial}{\partial \tau}\right) R_{t,\theta,\tau} = Q_{t,\theta} C^T C Q^T_{t,\tau}, \\ R_{0,\theta,\tau} = R_{t,-\varepsilon,\tau} = R_{t,\theta,-\varepsilon} = 0, -\varepsilon \le \theta \le 0, -\varepsilon \le \tau \le 0, t > 0. \end{cases} \tag{4.3.17}$$

**Proof.** Letting $E = 0$ in (4.3.2)-(4.3.5), we obtain (4.3.14) and (4.3.15) exactly, but the equations for $Q$ and $G$ become

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right) Q_{t,\theta} = Q_{t,\theta} A^T + \mathrm{G}_{t,\theta,0} - Q_{t,\theta} C^T C P_t, \\ Q_{0,\theta} = Q_{t,-\varepsilon} = 0, -\varepsilon \le \theta \le 0, t > 0, \end{cases} \tag{4.3.18}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta} + \frac{\partial}{\partial \tau}\right) G_{t,\theta,\tau} = \Phi_{t-\theta,\theta} \Phi^T_{t-\tau,\tau} - Q_{t,\theta} C^T C Q^T_{t,\tau}, \\ G_{0,\theta,\tau} = G_{t,-\varepsilon,\tau} = G_{t,\theta,-\varepsilon} = 0, -\varepsilon \le \theta \le 0, -\varepsilon \le \tau \le 0, t > 0, \end{cases} \tag{4.3.19}$$

the solution of Eq. (4.3.19) has the representation

$$G_{t,\theta,\tau} = \int_{max(0,t-\theta-\varepsilon,t-\tau-\varepsilon)}^{t} \left(\Phi_{t-\theta,s-t+\theta} \Phi^T_{t-\tau,s-t+\tau} - Q_{s,s-t+\theta} C^T C Q^T_{s,s-t+\tau}\right) ds.$$

then

$$G_{t,\theta,0} = \int_{max(0,t-\theta-\varepsilon)}^{t} \left(\Phi_{t-\theta,s-t+\theta} \Phi^T_{t,s-t} - Q_{s,s-t+\theta} C^T C Q^T_{s,s-t}\right) ds.$$

Using $\Lambda_{t,-\theta} = cov(\varphi_{t-\theta}, \varphi_t)$, one can derive

$$\Lambda_{t,-\theta} = \int_{max(0,t-\theta-\varepsilon)}^{t} \Phi_{t-\theta,s-t+\theta} \Phi^T_{t,s-t} ds.$$

This implies

$$G_{t,\theta,0} = \Lambda_{t,-\theta} - \int_{max(0,t-\theta-\varepsilon)}^{t} Q_{s,s-t+\theta} C^T C Q_{s,s-t}^T ds.$$

Therefore, we can introduce a function $R$ as a solution of (4.3.17) and write (4.3.19) in the form of (4.3.16). This proves the theorem.

Theorem 4.3.3 presents an invariant optimal filter while the filter from Theorem 4.3.1 is non- invariant. Why does it happen that in the case of independent noises the optimal filter does not depend on the relaxing functions and depends just on the autocovariance function? To answer this question we will make two simplifications in the conditions of Theorem 4.3.1.

At first, let $\xi = 0$. Then $P_0 = cov\xi = 0$ and the solution of (4.3.15) becomes identically zero if $Q_{t,0} = 0$. Secondly, let the autocovariance function $\Lambda$ be stationary (starting the instant $\varepsilon$). Then the relaxing functions $\Phi$, corresponding to $\Lambda$ , are independent on the first variable and we can write

$$\Lambda_\sigma = \int_{\sigma-\varepsilon}^{0} \Phi_{s-\sigma}\, \Phi_s^T ds.$$

If $\overline{\Phi}_\theta = \Phi_{-\theta}^T, 0 \le \theta \le \varepsilon$, then

$$\Lambda_\sigma = \int_{\sigma-\varepsilon}^{0} \Phi_{s-\sigma}\, \overline{\Phi}_{-s} ds = \int_{-\varepsilon}^{-\sigma} \Phi_s\, \overline{\Phi}_{\sigma-s} ds,$$

assuming that $\Phi$ and $\overline{\Phi}$ vanish outside of $[-\varepsilon, 0]$ and $[0, \varepsilon]$, respectively. Thus, $\Lambda$ (more correctly, the extension of $\Lambda$ to $[-\varepsilon, \varepsilon]$, defined by $\Lambda_\sigma = \Lambda_{-\sigma}$) is a convolution of $\Phi$ and $\overline{\Phi}$, i.e., $\Lambda = \Phi \star \overline{\Phi}$. Informally, regarding $\overline{\Phi}$ as a conjugate of $\Phi$ , very similar to complex numbers we can write$|\Phi\,|^2 = \Lambda$ . This equality can be treated as some sort of “circle” with the “squared radius” $\Lambda$ centered at the origin. Following to

Bashirov [39], we can wish to select that relaxing function $\Phi$ which answers to the autocovariance function and minimizes the error of estimation. In other words, we have to solve the constrained optimization problem

$$J(\Phi) = trP_t \to min, |\Phi|^2 = \Lambda, \quad (4.3.20)$$

where $P$ is defined by (4.3.15)-(4.3.17).

For a moment, remove the constraint from (4.3.20). Then $J(\Phi)$ is a quadratic functional of $\Phi$ in $L_2(-\varepsilon, 0; \mathbb{R}^{n\times k})$, (see the first term in the right hand side of (4.3.19)). By properties of Riccati equations, this functional is nonnegative. Moreover, at $\Phi = 0$, (4.4.16) and (4.4.17) have zero solutions $Q = 0$ and $R = 0$ by the uniqueness of solution of Riccati equations. This produces $J(0) = 0$. Therefore, we can treat $J(\Phi)$ as some sort of "upward oriented paraboloid" with the global minimum at zero. Turning back to the constrained optimization problem in (4.3.20), we can informally regard it as a minimization of the "paraboloid" $J(\Phi)$over the "circle" $|\Phi|^2 = \Lambda$ . Making analogy with elementary calculus, $J(\Phi)$ is a constant on $|\Phi|^2 = \Lambda$ if and only if $J(\Phi)$ is a "circular paraboloid". Thus, in the case of independent noises $J(\Phi)$ behaves as an "upward oriented circular paraboloid".

Now consider the constrained optimization problem in (4.3.20) under conditions of Theorem 4.3.1 (correlated noises). Again make the above simplifications. Then in a similar way $J(\Phi)$ can still be associated with an "upward oriented paraboloid" which takes its minimal value 0 at $\Phi = 0$. At the same time, $J(\Phi)$ is not constant on $|\Phi|^2 = \Lambda$ since (4.3.3)-(4.3.5) include the covariance of signal and observation noises. Therefore, now $J(\Phi)$ behaves as an "upward oriented elliptic paraboloid" and we can expect at least two points on the “circle” $|\Phi|^2 = \Lambda$ as a solution of the constrained minimization problem in (4.3.20).

This informal consideration is just a hint to a solution of the problem on selection of relaxing function that produces minimal error of estimation in the case of correlated noises (non-invariant filtering result). We left its strong justification still open at this stage of developments.

Now, an engineering applications investigation about how does the filter from Theorem 4.3.1 work if we try to implement it in? For this, using the representation

$$\psi_{t,0} = \int_{max(0,t-\varepsilon)}^{t} Q_{s,s-t} C^T (dz_s - C\hat{x}_s ds),$$

for the solution of the second equation in (4.3.14), write the first equation in (4.3.14) in the form

$$\hat{x}'_t = (A - P_t C^T C)\hat{x}_t$$
$$- \int_{max(0,t-\varepsilon)}^{t} Q_{s,s-t} C^T C \hat{x}_s ds + \int_{max(0,t-\varepsilon)}^{t} Q_{s,s-t} C^T z'_s ds, \qquad (4.3.21)$$

where we have employed the derivative notation for SDE instead of the differential notation. In this equation, $P$ and $Q$ (together with $R$) are solutions of deterministic PDEs (4.3.15)-(4.3.17). Therefore, they can be calculated beforehand by use of numerical methods for solution of PDE and stored somewhere in a computer.

Actually, $z'_t$ is an observation made at time $t$. Together with its $\varepsilon$-past it is an input of the filter. Therefore, the filter does not work without memorial manner. Storing the observation at instant $t$ up to the instant $t + \varepsilon$ is required. After the instant $t + \varepsilon$ the observation value $z'_t$ can be deleted from the memory. The term

$$f_t = f(z'_s; max(0, t - \varepsilon) < s \le t) = P_t C^T z'_t + \int_{max(0,t-\varepsilon)}^{t} Q_{s,s-t} C^T z'_s ds,$$

in the right hand side of (4.3.21) is formed on the base of the observation input data. So, the best estimate is a solution of the DE

$$\hat{x}_t' = (A - P_t C^T C)\hat{x}_t - \int_{t-\varepsilon}^{t} Q_{s,s-t} C^T C \hat{x}_s ds + f_t, t > 0, \tag{4.3.22}$$

with $\hat{x}_t = 0$ for $t \in [-\varepsilon, 0]$. This is a differential delay equation with a distributed delay of the best estimate. So, again calculated best estimate $\hat{x}_t'$ at the instant $t$ should be stored up to the instant $t + \varepsilon$. After the instant $t + \varepsilon$ it can be deleted.

Thus instead of ODE in the case of the KF, the filter from Theorem 4.3.1 is based on differential delay equation.

## 4.4 Invariant KF When the Signal and the measurements noises are sum of independent WN and WBN's.

In this section we present equations of KF when the signal and observation systems are corrupted by the sum of WN and WBN's. We assume that all the noise processes are independent and derive the equations depending on autocovariance function of the WBN's. As an application, it is used for synthesis of the optimal control in LQG problem under WBN's.

In accordance to the introductory section, the partially observable linear system will be considered

$$\begin{cases} dx_t = (Ax_t + \varphi_t^1)dt + Bdw_t, \;\; x_0 = \xi, t > 0\,, \\ dz_t = (Cx_t + \varphi_t^2)dt + dv_t, \;\; z_0 = 0 \;\; t > 0, \end{cases} \tag{4.4.1}$$

where $x$ and $z$ are vector-valued signal and mesurements processes, $A, B$ and $C$ are matrices, $\xi$ is a vector-valued GRV with $\boldsymbol{E}\,\xi = 0$, $w$ and $v$ are standard Wiener processes, $\varphi_t^1$ and $\varphi_t^2$ are stationary (starting the instants $t = \varepsilon > 0$ and $t = \delta > 0$) WBN's with the autocovariance functions $\Lambda_{t,\theta}$ and $\Sigma_{t,\alpha}$, respectively, and $\xi, w, v, \varphi_t^1$ , $\varphi_t^2$ are mutually independent. We will assume that $\varphi_t^1$ and $\varphi_t^2$ accept the integral representations

$$\varphi_t^1 = \int_{\max(0,t-\varepsilon)}^{t} \Phi_{s-t}^1 dw^1 \text{ and } \varphi_t^2 = \int_{\max(0,t-\varepsilon)}^{t} \Phi_{s-t}^2 dw^2, t \geq 0, \qquad (4.4.2)$$

where $w^1$ and $w^2$ are Wiener processes, $\Phi^1$ and $\Phi^2$ are absolutely continuous, $\Phi_0^1 = 0$, and $\Phi^2_{-\delta} = 0$. As it was mentioned in the introductory part of this chapter, the existence of such representations is already proved in the one-dimensional case although the number of functions $\Phi^1$ and $\Phi^2$ is infinite in general. To reach the independence of $\xi$, $w, v, \Phi^1, \Phi^2$ we assume that $\xi, w, v, w^1, w^2$ are mutually independent. The dimensions of all vectors and matrices in (4.4.1) and (4.4.2) are assumed to be consistent to each other. In this section, invariant equations for the best estimate $\hat{x}_t = E(x_t | z_s, 0 \leq s \leq t)$ will be deduced.

Considering the filtering problem (4.4.1), in [30] it is proved that the best estimate process $\hat{x}$ in the filtering problem (4.4.1) is uniquely presented as a solution of the equation

$$d\hat{x}_t = \left(A\hat{x}_t + \psi_{t,0}^1\right)dt + \left(P_t C^T + M_{t,0}^T\right)\left(dz_t - C\hat{x}_t dt - \psi_{t,0}^2 dt\right), \qquad (4.4.3)$$

for $t > 0$ with the initial condition $\hat{x}_0 = 0$, where $\psi^2$ and $\psi^2$ are unique solutions of

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta}\right)\psi_{t,\theta}^1 dt = \left(Q_{t,\theta} C^T + G_{t,\theta,0}^T\right)\left(dz_t - C\hat{x}_t - \psi_{t,0}^2 dt\right), \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \alpha}\right)\psi_{t,\alpha}^2 dt = \left(M_{t,\alpha} C^T + K_{t,\alpha,0}^T\right)\left(dz_t - C\hat{x}_t - \psi_{t,0}^2 dt\right), \end{cases} \qquad (4.4.4)$$

for $t > 0$, $-\varepsilon < \theta \leq 0$, $-\delta < \alpha \leq 0$ with the zero initial and boundary conditions along the lines $\theta = -\varepsilon$ and $\alpha = -\delta$. Here $P$ is a unique solution of the Riccati equation

$$P'_t = AP_t + P_t A^T + Q_{t,0} + Q_{t,0}^T + BB^T - \left(P_t C^T + M_{t,0}^T\right)\left(CP_t + M_{t,0}\right), \qquad (4.4.5)$$

for $t > 0$ and satisfies the initial condition $P_0 = cov\, \xi$. The supplemental functions $Q, M$ , $K$ and $G$ are solutions of

$$\begin{cases} \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}\right) Q_{t,\theta} = Q_{t,\theta}A^T + S_{t,\theta,0} - \left(Q_{t,\theta}C^T + G_{t,\theta,0}^T\right)\left(CP_t + M_{t,0}\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}\right) M_{t,\alpha} = M_{t,\alpha}A^T + G_{t,\alpha,0} - \left(M_{t,\alpha}C^T + K_{t,\alpha,0}\right)\left(CP_t + M_{t,0}\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \tau}\right) S_{t,\theta,\tau} = \Phi_\theta^1\Phi_\tau^{1T} - \left(Q_{t,\theta}C^T + G_{t,\theta,0}^T\right)\left(CQ_{t,\tau}^T + G_{t,\tau,0}\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}+\frac{\partial}{\partial \sigma}\right) K_{t,\alpha,\sigma} = \Phi_\alpha^2\Phi_\sigma^{2T} - \left(M_{t,\alpha}C^T + K_{t,\alpha.0}\right)\left(CM_{t,\sigma}^T + K_{t,\sigma,0}^T\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \alpha}\right) G_{t,\alpha,\theta} = -\left(M_{t,\alpha}C^T + K_{t,\alpha.0}\right)\left(CQ_{t,\theta}^T + G_{t,\theta,0}\right), \end{cases} \tag{4.4.6}$$

with the zero initial and boundary conditions. Moreover, $e_t = E\|\hat{x}_t - x_t\|^2 = trP_t$.

The method of the proof of equations (4.4.3)-(3.3.6) is a reduction of the system in (4.4.1) to an infinite dimensional linear system disturbed by only WN's, using the KF for the reduced system, and applying methods of functional analysis.

The equations in (4.4.3)-(3.3.6) are not invariant because they contain the relaxing functions $\Phi^1$ and $\Phi^2$. In the following theorem, these equations will be changed so that to make them independent on $\Phi^1$ and $\Phi^2$, dependent on $\Lambda$ and $\Sigma$ of the WBN's $\varphi_t^1$ and $\varphi_t^2$, respectively. This is stated in the following theorem.

**Theorem 4.4.1** Under the conditions stated in first section of this chapter, the best estimate process $\hat{x}$ in the filtering problem (4.4.1) is uniquely presented as a solution of the equation (4.4.3) for $t > 0$ with the initial condition $\hat{x}_0 = 0$, where $\psi^1$ and $\psi^2$ are unique solutions of

$$\begin{cases} \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}\right) \psi_{t,\theta}^1 dt = \left(Q_{t,\theta}C^T + G_{t,\theta,0}^T\right)\left(dz_t - C\hat{x}_t - \psi_{t,0}^2 dt\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}\right) \psi_{t,\alpha}^2 dt = \left(M_{t,\alpha}C^T + \Sigma_{t,-\alpha} - N_{t,\alpha,0}\right)\left(dz_t - C\hat{x}_t - \psi_{t,0}^2 dt\right), \end{cases} \tag{4.4.7}$$

for $t > 0, -\varepsilon < \theta \leq 0, -\delta < \alpha \leq 0$ with the zero initial and boundary conditions along the lines $\theta = -\varepsilon$ and $\alpha = -\delta$. Additionally, $P$ is a unique solution of the Riccati equation (4.4.5) for $t > 0$ with $P_0 = cov\,\xi$ and the supplemental functions $Q, M, R, N,$ and $G$ are solutions of

$$\begin{cases} \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}\right) Q_{t,\theta} = Q_{t,\theta}A^T + \Lambda_{t,-\theta} - R_{t,\theta,0} - \left(Q_{t,\theta}C^T - G^T_{t,\theta,0}\right)\left(CP_t + M_{t,0}\right) \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}\right) M_{t,\alpha} = M_{t,\alpha}A^T + G_{t,\alpha,0} - \left(M_{t,\alpha}C^* + \Sigma_{t,-\alpha} - N_{t,\alpha,0}\right)\left(CP_t + M_{t,0}\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \tau}\right) R_{t,\theta,\tau} = \left(Q_{t,\theta}C^T - G^T_{t,\theta,0}\right)\left(CQ^T_{t,\tau} - G_{t,\tau,0}\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}+\frac{\partial}{\partial \sigma}\right) N_{t,\alpha,\sigma} = \left(M_{t,\alpha}C^T + \Sigma_{t,-\alpha} - N_{t,\alpha.0}\right)\left(CM^T_{t,\sigma} + \Sigma^T_{t,-\sigma} - N^T_{t,\sigma,0}\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \alpha}\right) G_{t,\alpha,\theta} = \left(M_{t,\alpha}C^T + \Sigma_{t,-\alpha} - N_{t,\alpha.0}\right)\left(CQ^T_{t,\theta} + G_{t,\theta,0}\right), \end{cases}$$

…(4.4.8)

with zero initial and boundary conditions. Moreover, $e_t = E\|\hat{x}_t - x_t\|^2 = trP_t$.

Proof. As far as $\varphi^1_t$ and $\varphi^2_t$ are WBN's with the autocovariance functions $\Lambda$ and $\Sigma$, we can choose one of the relaxing functions $\Phi^1$ and $\Phi^2$ for them among infinitely many possible relaxing functions. Then the best estimate $\hat{x}$ in the problem (4.4.1)-(4.4.2) satisfies the set of equations (4.3.3)-(4.4.6). It remains to show that (4.4.4) and (4.4.6) can be written as (4.4.7) and (4.4.8), respectively. One can see that everywhere in (4.4.3)-(4.4.6), $G$ is replaced by $-G$. This is done just to make positive the sign of the right side of the last equation in (4.4.6).

Furthermore, letting $D_{t,\theta,\tau}$ be the solution of the equation

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \alpha}\right) D_{t,\theta,\tau} = \Phi^1_\theta \Phi^{1T}_\tau,$$

with the initial and boundary conditions $D_{0,\theta,\tau} = D_{t,\theta,-\varepsilon} = D_{t,-\varepsilon,\tau} = 0$, $-\varepsilon \le \theta \le 0$, $-\varepsilon \le \theta \le 0$, $t > 0$, we see that $D_{t,\theta,\tau}$ has the representation

$$D_{t,\theta,\tau} = \int_{\max(0,t-\theta-\varepsilon,t-\tau-\varepsilon)}^{t} \Phi^1_{s-t+\theta}\Phi^{1T}_{s-t+\tau}\, ds.$$

Using the definition of the autocovariance function, the previous equation becomes

$$D_{t,\theta,\tau} = \int_{\max(0,t-\theta-\varepsilon)}^{t} \Phi^1_{s-t+\theta}\Phi^{1T}_{s-t}\, ds = \Lambda_{t,-\theta}.$$

Letting $R_{t,\theta,\tau} = D_{t,\theta,\tau} - S_{t,\theta,\tau}$ , we see that $R_{t,\theta,\tau}$ satisfies the third equation in (4.4.8). Therefore, we can replace the equation for $S$ in (4.4.6) with the equation for $R$ in (4.4.8) with the substitution $S_{t,\theta,0} = \Lambda_{t,-\theta} - R_{t,\theta,0}$ everywhere. The same can be done with the equation for $K$ in (4.4.6), transforming it to the equation for $N$ in (4.4.8) and substituting $K_{t,\alpha,0} = \Sigma_{t,-\alpha} - \mathrm{N}_{t,\alpha,0}$ everywhere. This produces the equations of Theorem 4.4.1, which are valid for all relaxing functions corresponding to the autocovariance functions $\Lambda$ and $\Sigma$, which finalize the proof.

The process

$$\bar{z} = dz_t - C\hat{x}_t dt - \psi_{t,0}^2 dt, \bar{z}_0 = 0, t > 0,$$

is the innovation process in the filtering problem (4.4.1) and, therefore, it is a standard Wiener process. From the first equation in (4.4.7),

$$\psi_{t,\theta}^1 = \int_{\max(0,t-\theta-\varepsilon)}^{t} \left(Q_{s,s-t+\theta} C^T - G_{s,s-t+\theta,0}^T\right) d\bar{z}_s.$$

This implies

$$\psi_{t,0}^1 = \int_{\max(0,t-\varepsilon)}^{t} \left(Q_{s,s-t} C^T - G_{s,s-t,0}^T\right) d\bar{z}_s.$$

Comparing this with (4.4.2), one can see that $\psi_{t,0}^1$ is a WBN generated by the innovation process $\bar{z}$ and relaxing function

$$\Psi_{t,0}^1 = Q_{t+\theta,\theta} C^T - G_{t+\theta,\theta,0}^T.$$

Similarly, $\psi_{t,0}^2$ is a WBN generated by the innovation process $\bar{z}$ and relaxing function

$$\Psi_{t,0}^2 = M_{t+\alpha,\alpha} C^T + \Sigma_{t+\alpha,-\alpha} - M_{t+\alpha,\alpha,0}^T.$$

Therefore, the WBN KF from Theorem 4.4.1 works in the following way. The equations in (4.4.7) produce two WBN's $\psi_{t,0}^1$ and $\psi_{t,0}^2$ which effect to the equation in

(4.4.3) of the best estimate $\hat{x}_t$. One of them effects to the drift in (4.4.3), the other one forms the innovation process. The equation in (6) is a modification of the Riccati equation of the WN KF. The system of equations in (4.4.8) serves for relaxing and autocovariance functions of $\psi^1_{t,0}$ and $\psi^2_{t,0}$. The first two equations in (4.4.8) for $Q$ and $M$ produce components of the relaxing functions, the next two equations for $R$ and $N$ form the autocovariance functions, and the last equation for $G$ produces the joint autocovariance function of $\psi^1_{t,0}$ and $\psi^2_{t,0}$. Although $\varphi^1_t$ and $\varphi^1_t$ are independent WBN's in (4.4.1), the WBN's $\psi^1_{t,0}$ and $\psi^2_{t,0}$ of the filter are dependent because they are generated by the same innovation process. Therefore, the KF from Theorem 4.4.1 is a nontrivial extension of particular cases from [25, 28, 30] because it additionally adjusts the effect of the function G to the filter relating two correlated WBN's $\psi^1_{t,0}$ and $\psi^2_{t,0}$.

The WBN KF from Theorem 4.4.1 is independent on relaxing functions $\Phi^1$and $\Phi^2$, depends on autocovariance functions $\Lambda$ and $\Sigma$ which are exactly the parameters available in applications. Therefore, theoretically it is ready for applications if respective numerical methods are developed. In this way, no difficulty appears with the equations related to $\psi^1_{t,0}$. The way of overcoming the difficulties related to numerical solution of equations for $\psi^2_{t,0}$ is discussed somehow.

Theorem 4.4.1 has immediate impact to control problems for designing invariant optimal controls. As example, consider LQG problem for system corrupted by the sum of white and WBN's. Adding the first equation in (4.4.1) a control action $u$, consider LQG problem of minimizing the functional

$$\mathrm{J}(u) = E\left(\langle x^u_T, Lx^u_T\rangle + \int_0^T (\langle x^u_T, Fx^u_T\rangle + \langle u_t, Hu_t\rangle)dt\right), \tag{4.4.9}$$

over all square integrable vector-valued controls $u$ on $[0,T]$ subject to

$$\begin{cases} dx_t^u = (Ax_t^u + Du_t + \varphi_t^1)dt + Bdw_t, x_0^u = \xi, 0 < t \le T, \\ dz_t = (Cx_t^u + \varphi_t^2)dt + dv_t, z_0 = 0, 0 < t \le T, \end{cases} \tag{4.4.10}$$

where $E$ is the expectation. In addition to the conditions of the first section, we assume that $D$, $F$, $H$, and $L$ are matrices of respective dimensions such that $F$ and $L$ are nonnegative and $H$ is positive matrices.

**Theorem 4.4.2** (Application to LQG problem) Under the conditions stated in the first section of this chapter, the optimal control $u^*$ in the LQG problem (4.4.9)-(4.4.10) is uniquely determined by

$$u_t^* = -H^{-1}D^T \left( E_t \hat{x}_t^* - \int_{\max(-\varepsilon, t-T)}^{0} \mathcal{U}_{t-\theta}^* E_{t-\theta} \psi_{t,\theta}^1 \, d\theta \right),$$

where, $\hat{x}_t^*$ is the optimal trajectory satisfying

$$d\hat{x}_t^* = \left(A\hat{x}_t^* + Du_t^* + \psi_{t,0}^1\right)dt + (P_t C^T + M_{t,0}^T)(dz_t - C\hat{x}_t^* dt - \psi_{t,0}^2 dt,$$

for $0 < t \le T$ with the initial condition $\hat{x}_0^* = 0$, $E$ is a unique solution of the Riccati equation

$$E_t' + E_t A + A^T E_r + F - E_r D H^{-1} D^T E_t = 0, 0 < t \le T, E_T = L,$$

$\mathcal{U}_{t,s}$ is the transition matrix of $-DH^{-1}D^T E_t$, $(\psi^1, \psi^2)$ is a unique solution of the system in (4.4.7), $P$ is the unique solution of the Riccati equation in (4.4.5), and $(Q, M, R, N, G)$ is the unique solution of the system in (4.4.8).

**Proof.** The optimal control of this theorem was derived in terms of functions defined by (5)-(7) in [30], which are in non-invariant form. Applying the transition of the equations in (5) and (7) to equations (4.4.7) and (4.4.8), respectively, we easily obtain these formulae in the preceding invariant form.

# Chapter 5

# INVARIANT KF FOR CORRELATED WBN’S

## 5.1 Introduction and Motivation

In this chapter a complete set of invariant equations for KF for a linear signal-observation system corrupted by correlated WBN is introduced. In fact, the result here generalizes all cases in the previous chapter with some particular assumptions. This filter has a ready form to be used in application, just respective numerical methods must be developed.

In this chapter, we assume all the conditions in section 4.1, particularly, we define the WBN $\phi_t$ as

$$\phi_t = \int_{t-\varepsilon}^{t} \frac{1}{\varepsilon} dv_s, \tag{5.1.1}$$

we have

$$\Lambda_\theta = cov(\phi_{t+\theta}, \phi_t) = \mathrm{E}(\phi_{t+\theta}\phi_t) = \frac{I(\varepsilon-\theta)}{\varepsilon^2} \neq 0,$$

where $v$ is a Wiener process (for simplicity, one dimensional). If $0 < \theta \leq \varepsilon,$ and $cov(\phi_{t+\theta}, \phi_t) = 0$ if $\theta > \varepsilon$. Therefore, $\phi$ is a WBN and $\Lambda$ is its autocovariance function.

This motivates us to consider WBN’s in real processes as an “uncompleted derivative” in the form of (5.1.1) of Wiener processes. In the cases when ε is a sufficiently small, $\phi$ and $v'$are very close to each other and, respectively, mathematical methods for the WN $v'$ reflect the reality with more or less acceptable accuracy. But, for more adequate mathematical results (for example, serving the issues such as precise tracking satellites for improvement of GPS precision), WBN's with representation in (5.1.1) should be handled.

Equation (5.1.1) should be modified with replacement of $t-\varepsilon$ by $\max(0, t-\varepsilon\ )$ because WP's are observed starting some initial instant that is ordinarily taken to be zero. More generally, the constant integrand in (5.1.1) can be taken vector-valued and dependent on two time arguments. In the sequel we consider WBN's in the form

$$\varphi_t = \int_{max(0,t-\varepsilon)}^{t} \Phi_{t,s-t}\, dw_s, t \geq 0, \tag{5.1.2}$$

$w$ is a vector-valued standard Wiener process, and $\Phi$ is a matrix-valued non-random function on $[0,\infty] \times [-\varepsilon, 0]$. Here and below we do not specify the dimensions of vectors and matrices assuming that they are finite-dimensional and consistent to each other.

The covariance calculation formula for stochastic integrals implies that $\varphi$, defined by (5.1.2), has the autocovariance function

$$\Lambda_{t,\theta} = cov(\varphi_{t+\theta}, \varphi_t) = \int_{\max(0,t-\theta-\varepsilon)}^{t} \Phi_{t+\theta,s-t-\theta} {\Phi^T}_{t,s-t} ds \neq 0, \tag{5.1.3}$$

if $0 < \theta \leq \varepsilon$, and $\Lambda_{t,\theta} = 0$ if $\theta \geq 0$. Therefore, it is a WBN. We regard equation (5.1.2) as an integral representation of $\varphi$ and the function $\Phi$ as a relaxing (damping) function of $\varphi$. It is seen that $\varphi$ becomes stationary on $[\varepsilon, \infty)$ if $\Phi$ is independent on its first argument. In this case

$$\Lambda_{t,\theta} = \int_{\max(-t,\theta-\varepsilon)}^{0} \Phi_{s-\theta} \Phi_s^T ds, \tag{5.1.4}$$

implying $\Lambda_{t,\theta} = \Lambda_\theta$ if $t \geq \varepsilon$.

The formula in (5.1.2) presents the WBN $\varphi$ as a distributed delay of the WN $w'$. We can think that for some reasons, the WN $w'$ has an aftereffect to a system on time intervals of the length $\varepsilon$. The relaxing function $\Phi$ smoothly relaxes the aftereffect. Therefore, it becomes natural to assume that $\Phi$ is differentiable and $\Phi_{t,-\varepsilon} = 0$. In the sequel, these conditions will be imposed to relaxing functions.

A few scenarios of the aftereffects in specific areas are as follows:

- In a satellite communication, the basic sources creating noises are water vapors and radioactive particles in the higher layers of atmosphere. The density of the water vapor and the intensity of radioactive particles do not change momentarily, they are subjected to change during some time interval. Once affecting to a signal, they continue to affect till the atmospheric conditions and radioactivity change. This should create aftereffects of the initial effects.
- Financial markets are stochastic due to noise sources accumulated from unexpected changes in social and natural environments which do not change immediately. For example, the September 11 attacks had a significant impact to the stock markets over the world and had the aftereffect during some time interval.
- A quantization produces a partial loss of information, creating a quantization noise, commonly modeled as a WN. In a quantization of a continuously differentiable signals, a noise should have an aftereffect since, for example, a positive derivative at some quantization instant continues to be positive on some time interval and creates a correlation of the values of quantization noise within this time interval.

- The concept of a WN originates from the Brownian motion (a movement of a suspended particle in a liquid or a gas). This movement is probably an example of “mostly white” noise although it is not completely white. This is because a suspended particle has a constant speed within tiny time intervals between every two consequent collisions and, therefore, the derivative of the (real) Brownian motion is auto correlated within these tiny intervals.

These scenarios demonstrate that in a majority of cases the WN model of real noises is a simplified (although important) ideal version of a WBN model. The mathematical results for WN’s can be deduced from the results for WBN’s as a limiting case. This allows us to deduce a KF for systems with point wise delayed WN’s [28, 29] which is not covered by the classical KF.

## 5.2 Invariance

In modeling real processes, WBN’s are detected just by autocovariance and also cross- covariance functions. They are obtained as a result of estimation (commonly, by use of time series analysis) of large data. This is similar to estimation of covariance and correlation matrices in the WN KF model. Commonly, such estimations decrease the level of adequacy of the model. But, a discussion of this issue does not belong to the aim of this chapter, in which we assume that the respective autocovariance and cross-covariance functions are given readily.

There is another issue that also effects to the adequacy of the filtering model. In the one-dimensional stationary case, it is shown in [30, 31] that for a given positive definite function $\Lambda$, there are infinitely many relaxing functions $\Phi$ producing infinitely many WBN’s with the same autocovariance function $\Lambda$. If two WBN’s

$\varphi^{1}$ and $\varphi^{2}$ are given by autocovariance functions $\Lambda$ and $\Sigma$ and also by cross-covariance function $\Pi$, then still there are many pairs of relaxing functions $(\Phi^{1}, \Phi^{2})$. Therefore, modeling of WBN's in the integral form requires making a proper selection among infinitely many pairs of relaxing functions. This raises an importance of obtaining the results which are independent on the infinite variations of relaxing functions, but dependent on the unique autocovariance and cross covariance functions. This type of results are called invariant results.

## 5.3 Setting of the Problem

Consider the partially observable linear system

$$\begin{cases} dx_t = (Ax_t + \varphi_t^1)dt + Bdu_t, \ \ x_0 = \xi, t > 0\,, \\ dz_t = (Cx_t + \varphi_t^2)dt + dv_t, \ \ z_0 = 0 \ \ t > 0, \end{cases} \tag{5.3.1}$$

where $x$ and $z$ are vector-valued signal and observation processes, $A,\ B$ and $C$ are matrices, $\xi$ is a vector-valued GRV with $E\xi = 0$, $u$ and $v$ are standard Wiener processes. We assume that $\varphi^{1}$ and $\varphi^{2}$ are WBN's accepting the integral representations

$$\varphi_t^1 = \int_{\max(0,t-\varepsilon)}^{t} \Phi_{s-t}^1 dw_s\,, t \geq 0, \tag{5.3.2}$$

and

$$\varphi_t^2 = \int_{\max(0,t-\varepsilon)}^{t} \Phi_{s-t}^2 dw_s, \, t \geq 0, \tag{5.3.3}$$

where $\Phi^1$ and $\Phi^2$ are unknown differentiable relaxing functions with square integrable derivatives, satisfying $\Phi_{-\varepsilon}^1 = 0$ and $\Phi_{-\varepsilon}^2 = 0$. Instead, the autocovariance functions $\Lambda_{t,\theta}$ and $\Sigma_{t,\alpha}$ of $\varphi^1$ and $\varphi^2$, respectively, together with the cross-covariance function$\Pi_{t,\theta}$ are known. The random variable $\xi$ and processes $u, v$ and $w$ are assumed to be mutually independent. The dimensions of matrices and vectors are not specified considering that they are consistent. In this chapter, the aim is to derive the invariant results for the best least square estimate $E(x_t|z_s, 0 \le s \le t)$. This will be called the filtering problem (5.3.1).

According to (5.1.4), the following relations hold

$$\Lambda_{t,\theta} = cov\left(\varphi^1_{t+\theta}, \varphi^1_t\right) = \int_{\max(-t,\theta-\varepsilon)}^{0} \Phi^1_{s-\theta}\Phi^{1T}_s ds, \tag{5.3.4}$$

$$\Sigma_{t,\theta} = cov\left(\varphi^2_{t+\theta}, \varphi^2_t\right) = \int_{\max(-t,\theta-\varepsilon)}^{0} \Phi^2_{s-\theta}\Phi^{2T}_s ds, \tag{5.3.5}$$

$$\Pi_{t,\theta} = cov\left(\varphi^2_{t+\theta}, \varphi^1_t\right) = \int_{\max(-t,\theta-\varepsilon)}^{0} \Phi^2_{s-\theta}\Phi^{1T}_s ds, \tag{5.3.6}$$

where $t \ge 0$ and $-\varepsilon \le \theta \le 0$. As it was mentioned in section (5.2), the knowledge of $\Lambda, \Sigma$, and $\Pi$ does not imply a unique pair $(\Phi^1, \Phi^2)$. There are infinitely many such pairs. The essence of the invariant KF for the problem (5.3.1) consists of finding its equations in terms of $\Lambda, \Sigma$, and $\Pi$ and demonstrating that these equations do not change for different pairs $(\Phi^1, \Phi^2)$.

## 5.4 Invariant KF for a Linear Signal Measurements System Corrupted by Correlated WBN's

The following theorem states the main result of this chapter.

**Theorem 5.4.1** Under the conditions stated in the previous section, there exist a unique solution $(P, Q, M, R, N, S)$ of the system consisting of the Riccati equation

$$P'_t = AP_t + P_tA^T + Q_{t,0} + Q^T_{t,0} + BB^T - \left(P_tC^T + M^T_{t,0}\right)$$
$$\times \left(CP_t + M_{t,0}\right), P_0 = cov\,\xi, t > 0, \tag{5.4.1}$$

and the PDE's

$$
\begin{cases}
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}\right) Q_{t,\theta} = Q_{t,\theta}A^T + \Lambda_{t,-\theta} - R_{t,\theta,0} - \left(Q_{t,\theta}C^T + \Pi_{t,-\theta}^T - S_{t,\theta,0}^T\right)\left(CP_t + M_{t,0}\right), \\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}\right) M_{t,\alpha} = M_{t,\alpha}A^T + \Pi_{t,-\alpha} - S_{t,\alpha,0} - \left(M_{t,\alpha}C^T + \Sigma_{t,-\alpha} - N_{t,\alpha,0}\right)\left(CP_t + M_{t,0}\right), \\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \tau}\right) R_{t,\theta,\tau} = \left(Q_{t,\theta}C^T + \Pi_{t,-\theta}^T - S_{t,\theta,0}^T\right)\left(CQ_{t,\tau}^T + \Pi_{t,-\tau} - S_{t,\tau,0}\right), \\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}+\frac{\partial}{\partial \sigma}\right) N_{t,\alpha,\sigma} = \left(M_{t,\alpha}C^T + \Sigma_{t,-\alpha} - N_{t,\alpha.0}\right)\left(CM_{t,\sigma}^T + \Sigma_{t,-\sigma}^T - N_{t,\sigma,0}^T\right), \\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \alpha}\right) S_{t,\alpha,\theta} = \left(M_{t,\alpha}C^T + \Sigma_{t,-\alpha} - N_{t,\alpha.0}\right)\left(CQ_{t,\theta}^T + \Pi_{t,-\tau} - S_{t,\theta,0}\right),
\end{cases}
$$

… (5.4.2)

with the zero initial and boundary conditions

$$
\begin{cases}
Q_{0,\theta} = Q_{t,-\varepsilon} = 0, t \geq 0, \theta \in [-\varepsilon, 0], \\
M_{0,\alpha} = M_{t,-\varepsilon} = 0, t \geq 0, \alpha \in [-\varepsilon, 0], \\
R_{0,\theta,\tau} = R_{t,-\varepsilon,\tau} = R_{t,\theta,-\varepsilon} = 0, t \geq 0, \theta, \tau \in [-\varepsilon, 0], \\
N_{0,\alpha,\sigma} = N_{t,-\varepsilon,\sigma} = N_{t,\alpha,-\varepsilon} = 0, t \geq 0, \alpha, \sigma \in [-\varepsilon, 0], \\
S_{0,\alpha,\theta} = S_{t,-\varepsilon,\theta} = S_{t,\alpha,-\varepsilon} = 0, t \geq 0, \alpha, \theta \in [-\varepsilon, 0],
\end{cases}
$$

for this solution, there exist a unique solution $(\hat{x}, \psi^1, \psi^2)$ of the system consisting of the linear stochastic ordinary

$$
d\hat{x}_t = \left(A\hat{x}_t + \psi_{t,0}^1\right) + \left(P_tC^T + M_{t,0}^T\right)d\bar{z}_t, \hat{x}_0 = 0, t > 0, \tag{5.4.3}
$$

and partial

$$
\begin{cases}
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}\right) \psi_{t,\theta}^1 dt = \left(Q_{t,\theta}C^T + \Pi_{t,-\theta}^T - S_{t,\theta,0}^T\right)d\bar{z}_t, \\
\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}\right) \psi_{t,\alpha}^2 dt = \left(M_{t,\alpha}C^T + \Sigma_{t,-\alpha} - N_{t,\alpha,0}^T\right)d\bar{z}_t,
\end{cases} \tag{5.4.4}
$$

DE’s with zero initial and boundary conditions

$$
\begin{cases}
\psi_{0,\theta}^1 = \psi_{t,-\varepsilon}^1 = 0, \psi_{0,\alpha}^2 = \psi_{t,-\varepsilon}^2 = 0, \\
t \geq 0, \theta, \alpha \in [-\varepsilon, 0],
\end{cases}
$$

where $\bar{z}$ is defined by

$$
d\bar{z}_t = dz_t - C\hat{x}_t dt - \psi_{t,0}^2, t > 0, \bar{z}_0 = 0. \tag{5.4.5}
$$

The process $\hat{x}$ equals to the best least square estimate in the filtering problem (5.3.1) and has the error of estimation $e_t = E\|\hat{x}_t - x_t\| = trP_t$.

**Proof.** This consists of two major steps. The first and most time consuming step has already been done in [36]. Since the WBN's $\varphi^1$ and $\varphi^2$ accept integral representations, we can represent them as in (5.3.2)–(5.3.3) for some pair of relaxing functions $(\Phi^1, \Phi^2)$ which are related to the autocovariance and cross-covariance functions $\Lambda, \Sigma$, and $\Pi$ as in (5.3.4)–(5.3.6). In [36] it is proved that Theorem 5.4.1 holds if the system of equations (5.4.2) is replaced by

$$\begin{cases} \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}\right) Q_{t,\theta} = Q_{t,\theta}A^T + F_{t,\theta,0} - \left(Q_{t,\theta}C^T + G_{t,\theta,0}^T\right)\left(CP_t + M_{t,0}\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}\right) M_{t,\alpha} = M_{t,\alpha}A^T + G_{t,\alpha,0} - \left(M_{t,\alpha}C^T + K_{t,\alpha,0}\right)\left(CP_t + M_{t,0}\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \tau}\right) F_{t,\theta,\tau} = \Phi_\theta^1 \Phi_\tau^{1T} - \left(Q_{t,\theta}C^T + G_{t,\theta,0}^T\right)\left(CQ_{t,\tau}^T + G_{t,\tau,0}\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}+\frac{\partial}{\partial \sigma}\right) K_{t,\alpha,\sigma} = \Phi_\alpha^2 \Phi_\sigma^{2T} - \left(M_{t,\alpha}C^T + K_{t,\alpha.0}\right)\left(CM_{t,\sigma}^T + K_{t,\sigma,0}^T\right), \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}+\frac{\partial}{\partial \alpha}\right) S_{t,\alpha,\theta} = \Phi_\alpha^2 \Phi_\theta^{1T} - \left(M_{t,\alpha}C^T + K_{t,\alpha.0}\right)\left(CQ_{t,\theta}^T + G_{t,\theta,0}\right), \end{cases} \tag{5.4.6}$$

with the initial and boundary conditions

$$\begin{cases} Q_{0,\theta} = Q_{t,-\varepsilon} = 0, t \geq 0, \theta \in [-\varepsilon, 0], \\ M_{0,\alpha} = M_{t,-\varepsilon} = 0, t \geq 0, \alpha \in [-\varepsilon, 0], \\ F_{0,\theta,\tau} = F_{t,-\varepsilon,\tau} = F_{t,\theta,-\varepsilon} = 0, t \geq 0, \theta, \tau \in [-\varepsilon, 0], \\ K_{0,\alpha,\sigma} = K_{t,-\varepsilon,\sigma} = K_{t,\alpha,-\varepsilon} = 0, t \geq 0, \alpha, \sigma \in [-\varepsilon, 0], \\ G_{0,\alpha,\theta} = G_{t,-\varepsilon,\theta} = G_{t,\alpha,-\varepsilon} = 0, t \geq 0, \alpha, \theta \in [-\varepsilon, 0], \end{cases}$$

and (5.4.3) by

$$\begin{cases} \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \theta}\right) \psi_{t,\theta}^1 dt = \left(Q_{t,\theta}C^T + G_{t,\theta,0}^T\right) d\bar{z}_t, \\ \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \alpha}\right) \psi_{t,\alpha}^2 dt = \left(M_{t,\alpha}C^T + K_{t,\alpha,0}^T\right) d\bar{z}_t, \end{cases} \tag{5.4.7}$$

with the same zero initial and boundary conditions.

Shortly, this was proved by reduction of the system (5.3.1) to an $\infty$-dimensional system with the new state process

$$\tilde{x}_t = \begin{bmatrix} x_t \\ \tilde{\varphi}_t^1 \\ \tilde{\varphi}_t^2 \end{bmatrix},$$

corrupted only by WN's. A crucial role here are played by the integral representations in (5.3.2)–(5.3.4), which represent the WBN's $\varphi^1$ and $\varphi^2$ as a distributed delays of a WN. The equations in (5.4.3)–(5.4.4) are deduced on the basis of the equations for the components of $\hat{\tilde{x}}$. Since the state space is enlarged, the solution of the respective operator Riccati equation became in the form

$$\begin{bmatrix} P_t & \tilde{Q}_t^T & \tilde{M}_t^T \\ \tilde{Q}_t & \tilde{F}_t & \tilde{G}_t^T \\ \tilde{M}_t & \tilde{G}_t & \tilde{K}_t \end{bmatrix},$$

in which $\tilde{Q}_t^T$, $\tilde{M}_t^T$, $\tilde{F}_t$, $\tilde{K}_t$, and $\tilde{G}_t^T$ are integral operators. In fact, the solutions $Q$, $M$, $F$, $K$, and $G$ of the equations in (5.4.6) are the kernels (or adjoints of kernels) of these integral operators. The existence and uniqueness of solutions of these equations are immediate consequence of the same in the ∞-dimensional case. All these are realized in [36] by use of the WN KF in Hilbert spaces and methods of functional analysis. As a result, it produced KF for the problem (5.3.1) in the non-invariant form since (5.4.6) includes unknown relaxing functions $\Phi^1$ and $\Phi^2$.

The second step, which we are going to realize here, consists of demonstration that (5.4.6) can be replaced by (5.4.2) and (5.4.7) by (5.4.4). For this, introduce new functions $D$, $E$, and $H$ as the solutions of the equations

$$\begin{cases} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \theta} + \frac{\partial}{\partial \tau}\right) D_{t,\theta,\tau} = \Phi_\theta^1 \Phi_\tau^{1T}, \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \alpha} + \frac{\partial}{\partial \sigma}\right) E_{t,\alpha,\sigma} = \Phi_\alpha^2 \Phi_\sigma^{2T}, \\ \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \alpha} + \frac{\partial}{\partial \theta}\right) H_{t,\alpha,\theta} = \Phi_\alpha^2 \Phi_\theta^{1T}, \end{cases}$$

with the initial and boundary conditions

$$\begin{cases} D_{0,\theta,\tau} = D_{t,-\varepsilon,\tau} = D_{t,\theta,-\varepsilon} = 0, t \ge 0, \theta, \tau \in [-\varepsilon, 0], \\ E_{0,\alpha,\sigma} = E_{t,-\varepsilon,\sigma} = E_{t,\alpha,-\varepsilon} = 0, t \ge 0, \alpha, \sigma \in [-\varepsilon, 0], \\ H_{0,\alpha,\theta} = H_{t,-\varepsilon,\theta} = H_{t,\alpha,-\varepsilon} = 0, t \ge 0, \alpha, \theta \in [-\varepsilon, 0]. \end{cases}$$

It is seen that $D_{t,\theta,\tau}$ has the representation

$$D_{t,\theta,\tau} = \int_{\max(0,t-\theta-\varepsilon,t-\tau-\varepsilon)}^{t} \Phi^1_{s-t+\theta}\Phi^{1T}_{s-t+\tau}ds.$$

Using (5.3.4), we obtain

$$D_{t,\theta,0} = \int_{\max(0,t-\theta-\varepsilon)}^{t} \Phi^1_{s-t+\theta}\Phi^{1T}_{s-t}ds = \Lambda_{t,-\theta}.$$

Let $R_{t,\theta,\tau} = D_{t,\theta,\tau} - F_{t,\theta,\tau}$. Then

$$R_{t,\theta,0} = \Lambda_{t,-\theta} - F_{t,\theta,0}.$$

In similar manner, letting $N_{t,\alpha,\sigma} = E_{t,\alpha,\sigma} - K_{t,\alpha,\sigma}$, $S_{t,\alpha,\theta} = H_{t,\alpha,\theta} - G_{t,\alpha,\theta}$, and using (5.3.5)-(5.3.6), we obtain

$$N_{t,\alpha,0} = \Sigma_{t,-\alpha} - K_{t,\alpha,0},$$

and

$$S_{t,\alpha,0} = \Pi_{t,-\alpha} - G_{t,\alpha,0}.$$

So, changing $(F,K,G)$ to $(R,N,S)$, we see that the equations in (5.4.6) and (5.4.7) can be replaced by the equations in (5.4.2) and (5.4.4). Since the equations in (5.4.1)-(5.4.5) are independent on $\Phi^1$ and $\Phi^2$, $\hat{x}$ from (5.4.3) is the best estimate for all WBN's $\varphi^1$ and $\varphi^2$ which accept integral representations and have the autocovariance functions $\Lambda$ and $\Sigma$, respectively, and the cross-covariance function $\Pi$. This proves the theorem.

The WN KF consists of two equations for the best estimate $\hat{x}$ and the Riccati equation. In the case of additional WBN's as they appear in (5.3.1), these equations need adjustment to produce the best estimate. The additional equations in (5.4.4) and (5.4.2) are for this adjustment.

The process $\bar{z}$ defined by (5.4.5) is the innovation process in the filtering problem (5.3.1), therefore, it is a standard Wiener process. From the first equation in (5.4.4),

$$\psi^1_{t,\theta} = \int_{\max(0,t-\theta-\varepsilon)}^{t} \left(Q_{s,s-t+\theta}C^T + \Pi^T_{s,t-s-\theta} - S^T_{s,s-t+\theta,0}\right) d\bar{z}_s.$$

This implies

$$\psi^1_{t,0} = \int_{\max(0,t-\varepsilon)}^{t} \left(Q_{s,s-t}C^T + \Pi^T_{s,t-s} - S^T_{s,s-t,0}\right) d\bar{z}_s.$$

Comparing this with (4.4.2), one can see that $\psi^1_{t,0}$ is a non-stationary WBN generated by the innovation process $\bar{z}$ and relaxing function

$$\Psi^1_{t,\theta} = Q_{t+\theta,\theta}C^T + \Pi^T_{t+\theta,-\theta} - S^T_{t+\theta,\theta,0}.$$

Similarly, $\psi^2_{t,0}$ is a non-stationary WBN generated by the innovation process $\bar{z}$ and relaxing function

$$\Psi^2_{t,\alpha} = M_{t+\alpha,\alpha}C^T + \Sigma_{t+\alpha,-\alpha} - N^T_{t+\alpha,\alpha,0}.$$

Next, the third equation in (5.4.2) has the representation

$$R_{t,\theta,\tau} = \int_{\max(0,t-\theta-\varepsilon,t-\tau-\varepsilon)}^{t} \Psi^1_{t-\theta,s-t+\theta} \Psi^{1T}_{s-\tau,s-t+\tau}\, ds.$$

Using the definition of the autocovariance function, the previous equation becomes

$$R_{t,\theta,\tau} = \int_{\max(0,t-\theta-\varepsilon)}^{t} \Phi^1_{s-\theta,s-t+\theta} \Phi^{1T}_{t,s-t}\, ds = \Lambda_{t,-\theta}.$$

Comparing this with (5.1.3) we see that $R_{t,\theta,0}$ is the autocovariance function of the WBN $\psi^1_{t,0}$. In a similar way, $N_{t,\alpha,0}$ is the autocovariance function of the WBN $\psi^2_{t,0}$ and $S_{t,\alpha,0}$ is the cross-covariance function of the WBN's $\psi^2_{t,0}$ and $\psi^1_{t,0}$.

Resuming, the WBN KF from Theorem 5.4.1 works in the following way:

- The equations in (5.4.4) produce two WBN's $\psi^2_{t,0}$ and $\psi^1_{t,0}$ which effect to the equation in (5.4.3) of the best estimate $\hat{x}_t$. The first oneeffects to the drift in (5.4.3), the other one forms the innovation process given by (5.4.5).
- The equation in (5.4.1) is a modification of the Riccati equation of the WN KF and it provides the mean square error of estimation.
- The first two equations in (5.4.2) for $Q$ and $M$ produce components of the relaxing functions of $\psi^2_{t,0}$ and $\psi^1_{t,0}$.
- The next two equations in (5.4.2) for $R$ and $N$ form the autocovariance functions of $\psi^2_{t,0}$ and $\psi^1_{t,0}$ .
- The last equation in (5.4.2) for $S$ produces the cross-covariance function of $\psi^2_{t,0}$ and $\psi^1_{t,0}$ .

The WBN KF from Theorem 5.4.1 is independent on relaxing functions $\Phi^1$ and $\Phi^2$, it depends on autocovariance functions $\Lambda$ and $\Sigma$ and the cross-covariance function $\Pi$, which are exactly the parameters available in applications. This implies that the autocovariance and cross-covariance functions of disturbing WBN's are sufficient for the filter from Theorem 5.4.1.

Finally, some remarks about the wideness of the invariance of the filter from Theorem 5.4.1 are made. In [34] it is proved that the invariance of the signal WBN $\varphi^1$ can be extended (at least in a special case) to all square integrable relaxing functions $\Phi^1$. But its proof method is not extendable to the invariance of the observation WBN $\varphi^2$, which still remains within differentiable relaxing functions $\Phi^2$. with square integrable derivatives and $\Phi^2_{-\varepsilon} = 0$. While it is questionable, it may be possible the existence of WBN's which have a given autocovariance function but no integral representation. Mathematically, the invariance of the filter from Theorem

5.4.1 does not cover such WBN's. Although it is not proved yet, we think that Theorem 5.4.1 extends to them (if exits any) as well. But ordinary logic says that, if there is no any argument about selection of a WBN with a given autocovariance function, why not to choose that one (or ones) with which everything goes well.

## 5.5 Numerical Aspects

Theorem 5.4.1 carries into practice a long term mathematical investigations on estimation under WBN's. Therefore, it is appropriate sketching a skim, demonstrating how the filter from Theorem 5.4.1 can be numerically realized.

Once, it should be noted that the filter from Theorem 5.4.1 produces the best least square estimate. Therefore, any simplification of the WBN's $\varphi^1$ and $\varphi^2$ by some WN's with further application of the WN KF is obligated to produce a less adequate estimate.

Next, the filter from Theorem 5.4.1 is ready for applications since its equations depend on the system parameters $A$, $B$, $C$ as well as the autocovariance and cross-covariance functions $\Lambda$, $\Sigma$ and $\Pi$, which are available in applied problems. On the first hand, the equations in (5.4.1)–(5.4.5) may look complicated from computationally. But, they are a bit suitable for calculations. The equations in (5.4.1) and (5.4.2) are deterministic. Therefore, they can be solved numerically beforehand and stored somewhere in a computer. The equations in (5.4.3)–(5.4.5) are stochastic. These equations transform the observation process $z$ (input) to the best estimate process $\hat{x}$ (output). The processes $\psi^1$ and $\psi^2$ from (5.4.4) and $\bar{z}$ from (5.4.5) are just steps from the input toward to the output. Therefore, the main equation of the best estimate $\hat{x}$ in (5.4.3) could be solved depending on these stored data and timely available measurements.

In fact, the same numerical scheme can be applied to (5.4.3)–(5.4.5) and (5.4.1)–(5.4.2). The distinction is just in the number of equations and arguments. Therefore, we will just work on this for simple equations in (5.4.3)–(5.4.5).

convert the continuous argument *t to* discrete by considering

$$0 = t_0 < t_1 < \cdots < t_n < \cdots,$$

and do the same for $\theta$ and $\alpha$ by consisting

$$-\varepsilon = \theta_l < \theta_{l-1} < \cdots < \theta_m < \cdots < \theta_0 = 0,$$

$$-\varepsilon = \alpha_l < \alpha_{l-1} < \cdots < \alpha_k < \cdots < \alpha_0 = 0.$$

For simplicity, assume that

$$t_{n+1} - t_n = \theta_m - \theta_{m+1} = \alpha_k - \alpha_{k+1} = h,$$

for all $n = 0, 1, 2, \dots, m, k = 0, 1, \dots, l - 1$. Denote

$$U_t = P_t C^T + M_{t,0}^T,$$

$$V_{t,\theta} = Q_{t,\theta} C^T + \Pi_{t,-\theta}^T - S_{t,\theta,0}^T,$$

$$W_{t,\alpha} = M_{t,\alpha} C^T + \Sigma_{t,-\alpha} - N_{t,\alpha,0},$$

Assuming that $\Sigma$ and $\Pi$ are given autocovariance and cross-covariance functions, related to the WBN's $\varphi^1$ and $\varphi^2$, and $(P, Q, M, N, S)$ is numerically calculated as a solution of (5.4.1)-(5.4.2) and stored in a computer. Implicitly, $(U, V, W)$ depends on $\Lambda$ and $R$ as well while they do not appear in the preceding equations. Then (5.4.3) can be written as

$$\hat{x}_t' = (A - U_t C)\hat{x}_t + \psi_{t,0}^1 - U_t \psi_{t,0}^2 + U_t z_t'.$$

Again, for simplicity, let

$$\hat{x}_n = \hat{x}_{t_n}, z_n' = z_{t_n}', \psi_{n,m}^1 = \psi_{t_n,\theta_m}^1, \psi_{n,k}^2 = \psi_{t_n,\alpha_k}^2, U_n = U_{t_n}, V_{n,m} = V_{t_n,\theta_m},$$

$$W_{n,k} = W_{t_n,\alpha_k}.$$

Then using

$$\hat{x}'_{n+1} \approx \frac{\hat{x}_{n+1} - \hat{x}_n}{h},$$

the equation in (5.4.3) is converted to discrete form

$$\hat{x}_{n+1} = \hat{x}_n + h\left[(A - U_n C)\hat{x}_n + \psi^1_{n,0} - U_n \psi^2_{n,0} + U_n z'_n\right], \hat{x}_0 = 0. \qquad (5.4.8)$$

Here the input of the filter is $z'_n$. Therefore, we need only to determine formulae for $\psi^1_{t,0}$ and $\psi^2_{t,0}$ such that $\hat{x}_{n+1}$ can be calculated on the basis of $\hat{x}_n$. It can be fulfilled in $2l$ steps by discretization of the equations in (5.4.4), note that the number of those steps decreases to $2n$ if $0 \leq n \leq 1$.

For this, let $i$ and $j$ be unit vectors in the + directions of the $t$-axis and $\rho$-axis. Now, $\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \rho}\right)\psi_{t,\rho}$ is a directional derivative of $\psi$ in the direction of summation vector $i + j$ on the $t\rho$-plane. Therefore, we can let

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \rho}\right)\psi_{n,i} \approx \frac{\psi_{n,i} - \psi_{n-1,i+1}}{h\sqrt{2}},$$

where $\psi$ is either $\psi^1$ or $\psi^2$ , $\rho$ is either $\theta$ or $\alpha$, and $i$ is either $m$ or $k$. On this basis, the first equation in (5.4.4) can be discretized as

$$\psi^1_{n,m} = \psi^1_{n-1,m+1} + h\sqrt{2}V_{n-1,m+1}\left(z'_{n-1} - \hat{x}_{n-1} - \psi^2_{n-1,0}\right), \qquad (5.4.9)$$

and the second equation as

$$\psi^2_{n,k} = \psi^2_{n-1,k+1} + h\sqrt{2}W_{n-1,k+1}\left(z'_{n-1} - \hat{x}_{n-1} - \psi^2_{n-1,0}\right), \qquad (5.4.10)$$

which means that for going from calculation of $\hat{x}_n$ to calculation of $\hat{x}_{n+1}$ there are $\min(l, n)$ steps for finding $\psi^2_{n,0}$ and then the same number steps for $\psi^1_{n,0}$. Figure 5.1 demonstrated these steps, in which data at the black squares are taken from boundary conditions (all are 0), data at the black dots must be known before fining the values of $\hat{x}_{n+1}$, and data at the white dots must be calculated after the fining the values of $\hat{x}_{n+1}$. Data at the tails of the arrows must be calculated to find the data at the head points.

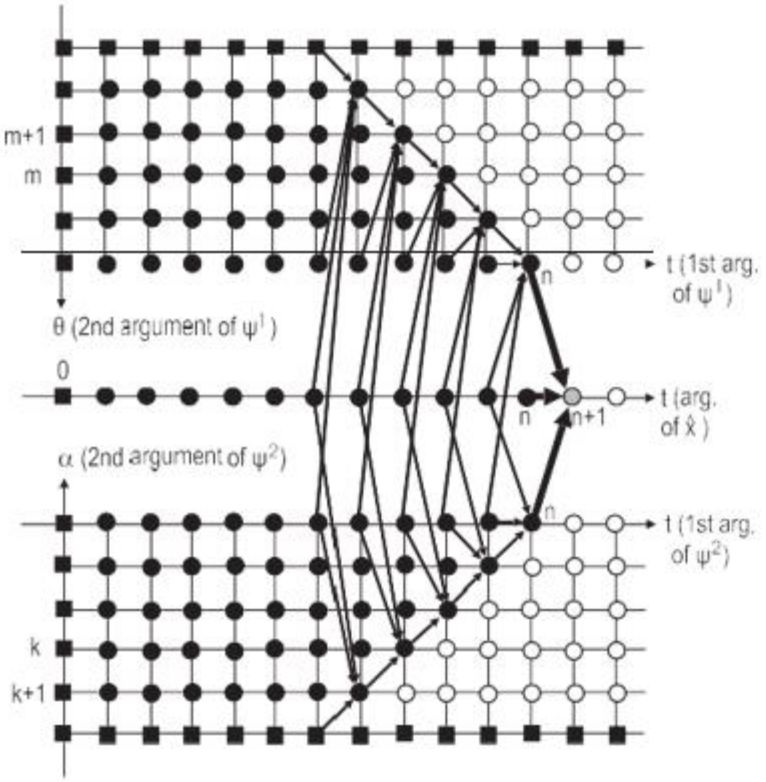


Figure 5.1. flow of steps for solution of (5.4.3)-(5.4.5)

# Chapter 6

# CONCLUSION

In this thesis invariant Kalman filter results under wide band noise are reviewed. The theory is contributed by the method involving integral representation of the wide band noise. This result can be turned into practically realizable algorithm. For this, numerical application is suggested.

# REFERENCES


[1] Kalman, R. E., A new approach to linear filtering and prediction problem”, *Trans. ASME (J. BasicEng.) Ser. D*, Vol. 82, pp. 35–45 (1960).

[2] Kalman, R. E. and R. S. Bucy, New results in linear filtering and prediction theory, Trans. ASME (J. Basic Eng.) Ser. D, Vol. 83, pp. 95–108 (1961).

[3] Crassides, J. L. and J. L. Junkins, Optimal estimation of dynamic systems, *Chapman and Hall*/CRC (2004).

[4] Lefebvre, T., H. Bruyninckx, and J. De Schutter, Nonlinear Kalman filtering for force-controlled robot tasks, *Springer-Verlag* (2005).

[5] Harv221ey, A. C., Forecasting, structural time series models and Kalman filter, *Cambridge University Press* (1989).

[6] Maitu, A. and R. Padhi, Robost control design of an air breathing engine for a supersonic vehicle using back stepping and UKF, *Asian J. Contr.*, Vol. 19, No. 5, pp. 1–12 (2017).

[7] Rana, M. M., L. Li, and S.W. Su, Distributed state estimation of smart grids with packet looses, *Asian J. Contr*., Vol. 19, No. 4, pp. 1306–1315 (2017).

[8] Bucy, R. S. and P. D. Joseph, Filtering for stochastic processes with application to guidance, *Wiley* (1968).

[9] Fleming, W. M. and R. W. Rishel, Deterministic and stochastic optimal control, *Springer-Verlag*(1975).

[10] Kushner, H. J. andW. J. Runggaldier, Filtering and control for wide bandwidth noise driven systems, IEEE Trans. *Automat. Control*, Vol. 32-AC, pp.123-133 (1987).

[11] Wang, W., D. Liu, and X. Wang, An improved wide band noise signal analysis method, *Computer and Information Science*, Vol. 3, No, 1, pp. 76-80 (2010).

[12] Al-Ansary, M. and H. Khalil, On the interplay of singular perturbations and wide band stochastic fluctuations, *SIAM J. Contr*., Vol. 24, pp. 83-98 (1986).

[13] Blankenship, G. B. and G. C. Papanicolaou, Stability and control of stochastic systems with wide band noise disturbances, *SIAM J. Appl. ath*. Vol. 34, pp 437-476 (1978).

[14] Kushner, H. J., Jump diffusion approximation for ordinary differential equations with wideband random right sides, *SIAM J. Contr. Optim*. Vol. 17, pp.729-744 (1979).

[15] Kushner, H. J., Diffusion approximations to output processes of nonlinear systems with wide-band inputs, and application, *IEEE Trans. Inform. Theory* Vol. IT-26, pp. 715-725 (1980).

[16] Kushner, H. J. and W. J. Runggaldier, Nearly optimal state feedback controls for stochastic systems with wideband noise disturbances, *SIAM J. Contr. Optim*. Vol. 25, No. 2, pp. 298-315 (1987).

[17] Kushner, H. J. and K. M. Ramachandran, Nearly optimal singular controls for wide band noise driven systems, *SIAM J. Contr. and Optim*., Vol. 26, No.3, pp. 569-591 (1988).

[18] Kushner, H. J., Weak convergence methods and singularly perturbed control and filtering problems, *Birkh¨auser, Boston* (1990).

[19] Kushner, H. J., W. J. Runggaldier, and M. Taksar, Diffusion approximation and optimal stochastic control, *Theory of Probability and Applications*, Vol. 44. pp. 669-698 (2000).

[20] Liptser, R. S., W. J. Runggaldier, and M. Taksar, Diffusion approximation and optimal stochastic control. *Theory of Probability and Applications*, 44, pp. 669-698 (2000).

[21] Kushner, H. J., Modelling and approximations for stochastic systems with state dependent singular controls and wide-band noise, *SIAM J. Contr. Optim*., Vol. 52, No. 1, pp. 311–338 (2014).

[22] Bashirov, A. E., H. Etikan, and N. S¸ emi, Filtering, smoothing and prediction for wide band noise driven systems, *J. Franklin Inst*., Vol. 334B , No.4, pp. 667–683 (1997).

[23] Bashirov, A. E., L. V. Eppelbaum, and L. R. Mishne, Improving E¨otv¨os correction by wide band noise Kalman filtering, *Geophys. J. Int.*, Vol. 108, No. 1, pp. 193–197 (1992).

[24] Bashirov, A. E., Z. Mazhar, H. Etikan, and S. Ert¨urk, Delay structure of wide band noises with application to filtering problems, *Optimal Control Appl. Methods,* Vol. 34, No. 1, pp. 69–79 (2013).

[25] Bashirov, A. E., Z. Mazhar, H. Etikan, and S. Ert¨urk, Boundary value problems arising in Kalman filtering, *Boundary Value Problems*, Vol. 2008, doi: 10.1155/2008/279410, 10 p. (2008).

[26] Ito, K., Stochastic integral, *Proceedings of the Imperial Academy*, Tokyo, Vol. 20, pp. 519-524(1944).

[27] Ito, K., On stochastic differential equations, *Memoirs of the American Mathematical Siciety*, Vol. 4, pp. 645-668 (1951).

[28] Bashirov, A. E., Filtering for linear systems with shifted noises, *Internat. J. Contr*., Vol. 78, No.7, pp. 521–529 (2005).

[29] Bashirov, A. E., Z. Mazhar, and S. Ert¨urk, Kalman type filter for systems with delay in observationnoise, *Appl. Comput. Math*., Vol. 12, No. 3, pp. 325–338 (2013).

[30] Bashirov, A. E. and S. Uˇgural, Analyzing wide band noise processes with application to control and filtering, *IEEE Trans. Automat. Control*, Vol. 47-AC, No. 2, pp. 323–327 (2002).

[31] Bashirov, A. E. and S. Uˇgural, Representation of systems disturbed by wide band noises, *Appl. Math. Lett*., Vol. 15, No. 5, pp. 607–613 (2002)

[32] Bashirov, A. E., Stochastic maximum principle in the Pontryagin's form for wide band noise driven systems, *Internat. J. Control*, Vol. 88, No. 3, pp. 461-468 2015).

[33] Bashirov, A. E. and N. Ghahramanlou, On partial S-controllability of semilinear partially observable systems, *Internat. J. Control*, Vol. 88, No. 5, pp. 969-982 (2015).

[34] Bashirov, A. E. and K. Abuassba, Invariant linear filtering and control results for wide band noise driven signal systems, *TWMS J. Appl. Eng. Math*., Vol. 8, No. 1.

[35] Bashirov, A. E., Linear filtering for wide bandnoise driven observation systems, Circuits Syst. *Signal Process*, Vol. 36, No. 3, pp. 1247-1263 (2017).

[36] Bashirov, A. E., Partially observable linear systems under dependent noises Systems & Control: Foundations & Applications, *Birkh¨auser, Basel* (2003).

[37] Bashirov, A. E., (1993), Control and filtering for wide band noise driven linear systems, AIAA J. *Guidance Control and Dynamics*, 16, pp. 983-985.

[38] Bashirov, A.E. Stochastic maximum principle in the Pontryagin's form for wide band noise driven systems, *International Journal of Control*, V.88, No.3, *2015*, pp.461-468.

[39] Bashirov, A.E. Problem 2.1: On error of estimation and minimum of cost for wide band noise driven systems, In: V.D. Blondel and A. Megretski (Eds.), *Unsolved Problems in Mathematical Systems and Control Theory*, Princeton University Press, Princeton and Cambridge, 2004, pp.67-73