Estimation and Control under Communication Network Constraints

A Thesis Presented
by
Bei Yan
to
The Department of Electrical and Computer Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Electrical Engineering
in the field of
Communication and Signal Processing
Northeastern University
Boston, Massachusetts
May 2010
Abstract

Motivated by distributed control and sensor network applications in Electric Energy Systems, we consider the problem of estimation via a communication network. When data is sent via communication channels in a large, wireless, multi-sensor network, the effect of communication constraints on estimation performance, such as communication delay and asynchronous irregular sampling, have to be considered. In this thesis we formulate a delay mitigation method based on a Kalman filter with time-stamping technology, which transmutes communication delay into increased estimation error, so that the closed-loop control system remains stable even in the presence of significant delay. The resulting signal to estimation error ratio (SEER), at any point in the estimation-control loop, is a monotone decreasing function of the average communication delay. When the sensor sampling patterns (SSP) are irregular and asynchronous, the SEER is a monotone decreasing function in each one of the average sampling intervals, with very minor dependence on higher order moments of this interval. The best performance is achieved when the individual SSPs are regular and aligned in such a way that the superposed set of sampling instants is as close to regular as possible. In particular, synchronized regular sampling (i.e., all sensors sampled regularly at the same time instants) is inferior to uniformly staggered regular sampling, in which the sampling instants of individual sensors are spaced evenly within a single sampling interval.
Acknowledgements

It took almost two years to write this thesis, and fortunately the support and guidance from many sources during this long time helped me finish it as well as possible. In my mind, the assistance has been academic and emotional. To these people I am thankful.

I would like to express my gratitude to my advisor, Professor Hanoch Lev-Ari, whose guidance and support from the initial to the final level enabled me to make progress on research and finish the thesis as well. I have a great fortune to work with him, and his patience and motivation helped me to overcome the many challenges I encountered in the process of carrying out my research and writing this thesis. I am really thankful for the time he spent on my research and thesis and all the help he offered.

Besides my advisor, I would like to thank Professor Aleksandar Stankovic who introduced me to a new and exciting research area: networked estimation and control, which became my research subject for a long time.

Lastly, I offer my regards and blessings to all my friends who supported me in any respect during the completion of the thesis.
Contents

Abstract .............................................................................................................................. i

Acknowledgements ......................................................................................................... ii

List of Figures .................................................................................................................. v

List of Tables ..................................................................................................................... vii

Chapter 1  Introduction ..................................................................................................... 1

  1.1  Spatially-Distributed Cyber-Physical Systems ......................................................... 3
  1.2  Continuous-Discrete Kalman Filter ......................................................................... 7
  1.3  Continuous-Discrete Controller .............................................................................. 12
  1.4  Thesis Contributions and Structure ......................................................................... 14

Chapter 2  Delay-Mitigating Estimation and Control ...................................................... 16

  2.1  Delay in Observation Path ...................................................................................... 17
  2.2  Delay-Mitigating Control ....................................................................................... 19
  2.3  Example: controlling a simple unstable system ....................................................... 21
  2.4  Concluding Remarks ............................................................................................. 23

Chapter 3  Effects of Sensor Sampling Patterns .............................................................. 25

  3.1  Sensor Sampling Pattern Classification ................................................................... 26
  3.2  The Effects of SSP on Kalman Filter Performance .................................................. 31

    3.2.1  The Dynamics of $P(t)$ ..................................................................................... 31
    3.2.2  Known Performance Bounds ............................................................................ 39
  3.3  Simulation Results: single-sensor estimation with irregular SSP ............................ 42

    3.3.1  Regular Sampling .............................................................................................. 43
    3.3.2  Irregular Sampling ............................................................................................ 45
3.3.3 The Effect of Sampling Interval Variance Index ..................................46
3.3.4 The Effect of Kurtosis of The Sampling Interval ..................................50
3.4 Simulation Results: two-sensor estimation with irregular SSP ..................53
  3.4.1 Effect of SSP Irregularity ..........................................................56
  3.4.2 Effect of Sensor Differences ......................................................58
3.5 Summary of Results ........................................................................61

Chapter 4 Concluding Remarks ............................................................64
  4.1 Summary of Results ........................................................................64
  4.2 Further Research ............................................................................66

Appendix A Averaged Time Update .......................................................68
Appendix B Diagonal Scaling Lemma ....................................................72
Appendix C The Characteristic Function and Statistics .........................73
Appendix D Perturbation Analysis of $P_{t,y}$ ........................................78

References ............................................................................................80
List of Figures

Fig. 1.1 Structure of shared-network control systems ........................................... 5
Fig. 1.2 The first approach of remote control system via a network .................. 6
Fig. 1.3 The second approach of remote control system via a network ........... 6
Fig. 1.4 The flow figure of remote control system via a network ...................... 7
Fig. 1.5 The evolution of $P(t)$ ............................................................................. 11
Fig. 1.6 State estimator based feedback control loop ........................................... 14
Fig. 2.1 Networked estimation/control configuration .......................................... 20
Fig. 2.2 Effect of delay mitigation on SEER, in the presence of observation noise 22
Fig. 2.3 Effect of delay mitigation on SEER, in the presence of system parameters inaccuracy .............................................................................. 23
Fig. 3.1 Periodically-variant SSP ........................................................................ 28
Fig. 3.2 Probability density function of uniform distribution ............................. 28
Fig. 3.3 Probability density function for the geometric distribution ................. 29
Fig. 3.4 $P_{+av}$ from equation vs $F_s$ in irregular SSP with scalar parameter..... 37
Fig. 3.5 $P_{+av}$ vs $F_s$ in regular SSP with unstable system ............................... 38
Fig. 3.6 SEER vs. $T_{av}$ (regular sampling) ............................................................ 44
Fig. 3.7 Effect of sensor choice on estimation performance (regular sampling) .... 45
Fig. 3.8 SEER vs. $T_{av}$ (uniform random distribution sampling with $\xi = 0.32$) ... 46
Fig. 3.9 Effect of relative variance index $\xi$ on SEER (top) and zoom-in (bottom):
periodically-variant SSP ................................................................................. 47
Fig. 3.10 Effect of relative variance index $\xi$ on SEER (top) and zoom-in (bottom):
uniformly-distributed random SSP .................................................................. 48
Fig. 3.11 Effect of relative variance index $\xi$ on SEER (top) and zoom-in (bottom):
comparing both SSP choices ............................................................................ 49
Fig. 3.12 Kurtosis of the mixed SSP ..................................................................... 51
Fig. 3.13  Effect of sampling interval kurtosis on SEER (top) and zoom-in (bottom):
convex combination of two periodically-variant SSP .................................... 52

Fig. 3.14  Estimation performance (\(\text{trace}(P_{\text{est}})\)) as a function of \(F_1\) and \(F_2\) ....... 54

Fig. 3.15  Contour plot of Fig. 3.14................................................................. 55

Fig. 3.16  Effect of relative shift on estimation performance: \(F_1 = F_2 = 10\) ........... 57

Fig. 3.17  Effect of relative shift on estimation performance: \(F_1 \neq F_2\) .................... 58

Fig. 3.18  The effect of sensor differences on estimation performance....................... 59

Fig. 3.19  The effect of sensor differences in another aspect................................. 61
List of Tables

Table 1.1 The continuous-discrete Kalman filter ................................................................. 11
Table 3.1 Statistics of the SSPs .................................................................................................. 30
Table 3.2 Characteristic functions of the SSPs ........................................................................ 30
Table 3.3 Kurtosis with changing $\lambda$ and fixed variance index $\xi = 0.02$ .................. 51
Table 3.4 Combination of sampling rate in sensor 1 and 2.................................................... 60
Chapter 1

Introduction

Process in communication technology has opened up the possibility of large scale control system in which the control task is distributed among several processors, sensors, estimators and controllers interconnected via communication channels. Such control systems may be distributed over long distances and may use a large number of actuators and sensors. The combination of a spatially-distributed physical (analog) system with a digital communication network imposes unique constraints that affect the design of such a cyber-physical systems (CPS). Sensor data need to be sampled and quantized so it can be transmitted over digital communication links with limited capacity. Communication delay [1] has significant effect on the quality of estimation, degrades the performance of networked controllers, and often results in loss of stability. Data packets may get lost, or in case of retransmission, are received with an added delay. Finally, the proposed use of large numbers of sensors in spatially-distributed cyber-physical systems may require accommodation of different, asynchronous, sampling rates.

Motivated by distributed control and Electric Energy Systems applications, we consider the problem of estimation and control via communication networks using Kalman filtering [2]. When data is sent via unreliable communication channels in a large, wireless, multi-hop sensor network, the effect of communication delays and loss of information in the control loop cannot be neglected. We analyze this problem using a continuous-time discrete-measurement Kalman filter, measure the delays with time-stamp technology and model the arrival of observations as a random process. We study the statistical convergence properties of the estimation error covariance, showing some results about the relationship between communication delay, sensor sampling pattern and estimation error covariance.

Numerous results have been published within the broad area of networked estimation and control. Most of these have focused on the effects of dropped communication packets or, more generally, the impact of irregular sensor sampling time patterns [3]. Only a few published results address the challenge of irregular observation delay and its effect on the quality of state estimation ([4], [5]). And even
fewer have considered the possibility of overcoming the destabilizing effects of delay in a networked control loop.

In Ch. 2 we analyze the effect of delay in electric power systems as a special application. Electric power systems are among most spatially extended engineered systems. Their stable operation in the presence of disturbances and outages is made possible by real-time control over communication networks. This mode of operation is likely to expand in the future, as main enablers for development of sustainable energy systems are in the information layer. Examples include sensing, estimation and control of components in the energy layer.

We propose a delay canceling approach based on time-stamping of transmitted signals, which can successfully mitigate the destabilizing effects of delay if we communicate state estimates rather than control signals from the estimation/control center to the actuators. By “time-stamping” we mean that every sensor sample $x_k(t)$ is transmitted over the network as an ordered triplet $(x,t,k)$, consisting of the sample value $'x'$, the corresponding sampling time instant $'t'$ and the identifier $'k'$ of the sensor. Time stamping is enabled by the availability of precise time signals derived from the Global Positioning System (GPS) in various nodes of the power system. Such signals are part of Phasor Measurement Units (PMUs) that are increasingly being deployed as a key ingredient of system-wide monitoring and control.

Our delay canceling method transmutes communication delay into increased estimation error, so that the average signal to noise ratio of any signal in the estimation-control loop is a monotone decreasing (and, in fact nearly linear) function of the average communication delay. We demonstrate the utility of our approach via an example. In particular, we show that our time-stamping-based delay canceling method makes it possible to overcome very large delays in the control path while avoiding adverse effects on the stability of the system. In our examples the delay can be 10-20 times higher than the smallest value that will cause instability in the original control path.

Moreover we analyze the effect of sensor sampling pattern (SSP) on the estimation performance. A regular SSP is completely characterized in terms of its sampling interval $T_s$. An irregular SSP has a time-varying sampling interval $T_s(n)$, which we choose to characterize in terms of its moments: the average and the variance of $T_s(n)$, as well as higher-order moments, as needed. We show in Ch. 3 that the average signal to estimation error ratio (SEER) depends primarily on the
average sampling interval, with a very minor dependence on the variance and a negligible dependence on higher-order moments of the time-varying sampling interval. The best performance (i.e., smallest estimation error) for a single-sensor system is achieved when the SSP is regular, so that all its higher-order moments vanish.

The same conclusions hold also for a Kalman filter estimation that employs multiple sensors: the average estimation error covariance is a monotone decreasing function in each one of the average sampling rates. In addition, the average estimation error covariance depends also on the relative alignment of the individual SSPs. The best performance is achieved when the individual SSPs are regular and aligned in such a way that the superposed set of sampling instants is as close to regular as possible.

1.1 Spatially-Distributed Cyber-Physical Systems

The broad subject of this thesis is estimation and control in Spatially-Distributed Cyber-Physical System. Cyber-Physical Systems (CPS) are integrations of computation with physical processes, such as the electrical power grid.

“Cyber-Physical” here means that the system consists of two layers, one is physical (energy flow) layer which is analog and the other is cyber (information flow) layer which is digital. The interface between the two layers is provided by sensors and actuators. Estimation and control algorithms are implemented digitally in the cyber layer [6].

“Spatially-Distributed” means that the system is spread over a large physical distance [7]. This necessitates to use a communication network as part of the cyber layer to connect between sensors, estimation/control center and actuators. It also means that the system, in general, would exhibit wave-propagation phenomena and may need to be described by partial differential equations. In this thesis, we only consider systems that can be described by a lumped, linear model. Specifically we assume a state space model as below

\[ \dot{x}(t) = Ax(t) + Bu(t) + Gw(t) \]  

where \( x(t) \), \( u(t) \) and \( w(t) \) are the state, input and process noise, respectively.

We shall assume that \( u(t) \) and \( w(t) \) are uncorrelated, and that \( w(t) \) is white.
Gaussian, so that \( E\{w(t)w'(\tau)\} = Q\delta(t-\tau) \).

Due to the spatial distribution, we can say the sensors and actuators are not co-located, and in general the A/D step (sampling) may have different, non-synchronous rates at different sensors. Furthermore, the communication network may cause occasional loss of information (packet drop), which means that some of the samples could be missing in the estimation step. Thus, in general, sampling patterns could be mildly irregular. We will denote the sampling instants for sensor \( k \) as \( t_n^{(k)} \), so that the measurements received at the central estimator are given by

\[
y_k(t_n^{(k)}) = C_k x(t_n^{(k)}) + \eta_k(n)
\]

So even if each sensor uses regular sampling, the stream of measurements received at the central estimator is still irregular, in general.

The synchrophasor standard (an evolving IEEE standard for power system) will force all samplers to use a GPS to get synchronized measurements, resulting in a regular, synchronized stream of observations from all sensors. Also using a TCP-IP protocol which involves retransmission of dropped packets will eliminate the phenomenon of packet dropping at the expense of increased communication delay. Nevertheless, one of our objectives in this thesis is to demonstrate that reliable state estimation can be obtained even in the presence of non-synchronous, irregular, multirate sensor sampling. In particular, we show that estimation performance depends mainly on the average sampling rate, and is only mildly affected by the irregularities in the sampling pattern.

Spatially-Distributed Cyber-Physical Systems are known in the control system literature as Networked Control Systems (NCS) [8]. General speaking, there are two major types of NCS architecture, shared-network control systems and remote control systems.

Using shared-network to transfer data from sensors to controllers and from controllers to actuators can greatly reduce the complexity of the whole system. This structure, as shown in Fig. 1.1, provides more flexibility in installation and troubleshooting. Furthermore this structure is especially useful when a control loop needs to exchange information with other control loops.
On the other hand, a remote control system is a system controlled by a controller located far away from it. There are two general approaches to designing a remote control system through a network. The first structure is to have several subsystems, in which each of the subsystem contains a set of sensors, a set of actuators, and a controller by itself. These system’s components are attached to the same control plant, as shown in Fig. 1.2. In this case, a subsystem controller receives a set point from the central controller. Another structure is to connect a set of sensors and a set of actuators to a network directly. Sensors and actuators in this case are attached to a plant (physical layer), while a central controller is separated from the plant via a network connection to perform a closed-loop control over the network, as shown in Fig. 1.3.
Both structures have their pros and cons. The first one (Fig. 1.2) is more modular. A control loop is simpler to be reconfigured. The second one (Fig. 1.3) has better interaction because data are transmitted to components directly. The main
difference of these two types of architecture is information layer, as the first one has the hierarchical information layer and the second one has the centralized. A controller in the second structure can observe and process every measurement, whereas a controller in the first structure may have to wait until the set point is satisfied to transfer the complete measurements, statue signals, or alarm signals [9]. For the sake of better performance, we use the second structure to realize our control system.

In addition to an irregular stream of observations, the use of a communication network within the cyber layer may cause delay both in the path from sensor to central controller and in the path from central controller to actuator, as shown in Fig. In this thesis, we focus on the effects of delay in the control path, and we show, in Ch. 2, that such effects can be mitigated by Kalman filtering with time-stamp technology which we will explain in detail in Ch. 2.

Fig. 1.4 The flow figure of remote control system via a network

1.2 Continuous-Discrete Kalman Filter

The state space model described in (1.1) and (1.2) is known as a continuous-discrete model in which the state equation is in continuous time and the measurement equation is in discrete time. The reason for a continuous-discrete model is that we
are interested in CPS with their characteristic two-layer structure:
1) A physical system that ‘lives’ in the real world, in continuous time.
2) An estimation and control system that ‘lives’ in ‘cyberspace’, namely is implemented in a computer/digital processor, in discrete time.

Sensors and actuators act as interfaces between the two layers, and they turn analog signals into a digital format and vice versa.

The Continuous-Discrete Kalman filter addresses the general problem of trying to estimate the states of a continuous time system from discrete-time measurements. We will use the continuous time linear state equation (1.1), as discussed in section 1.1. It will be easier to describe the operation of the continuous-discrete Kalman filter if we interpret the measurement equation (1.2) as an equivalent ‘single-sensor’ time variant measurement equation, viz., equation (1.3).

\[ y(t_n) = C(t_n)x(t_n) + \eta(n) \]

where \( \{t_n\} \) is the superposition [10] of the individual sampling sequences \( t^{(k)}_n \). A superposition of two ordered sequences is the ordered union of the two individual sets, along with a set of labels that indicated the origin of each element in the combined sequence. For instance, the superposition of \( S_1 = \{1,3,7,8\} \) and \( S_2 = \{2,5,6,9\} \) is the two-row matrix

\[
\begin{pmatrix}
1 & 2 & 3 & 5 & 6 & 7 & 8 & 9 \\
1 & 2 & 1 & 2 & 2 & 1 & 1 & 2
\end{pmatrix}
\]

The label below each element of the combined sequence tells us whether it came from \( S_1 \) or \( S_2 \). In case of a tie, when two or more elements in the first row have the same value, their labels can be arranged in any order. When we superpose the sampling sequences \( \{t^{(k)}_i; 1 \leq i < \infty\} \), we will refer to the elements of the combined sequence as \( \{t_n; 1 \leq n < \infty\} \) and to their labels as \( \{S_n \equiv S(t_n); 1 \leq n < \infty\} \). Also \( \eta(n) \) represents the measurement noise, assumed to be independent, white, with a Gaussian distribution, so that

\[ E(\eta(n)\eta'(n-k)) = R(n)\delta(k) \]

Since \( \eta(n) \) (in 1.3) represents the measurement noise at time instant \( t_n \), and each sensor may have a different level of measurement noise, \( R(n) \) depends on the sensor, namely \( R(n) = R_k \) when \( S(t_n) = k \). In other words, we assume
sensor-dependent but stationary measurement noise in our model.

The continuous-discrete state space model (1.1) and (1.3) gives rise to a standard continuous-discrete Kalman filter, as described in [11], which alternates between a time-update process (in continuous time) and a measurement-update step (in discrete time), as follows.

Measurement update:
When a measurement becomes available at \( t = t_n \), we carry out the update:

\[
x_+(t_n) = x_-(t_n) + K(t_n) v(t_n)
\]

Where \( x_+(t_n) \triangleq E[x(t_n) | \{ y(t_i), t_i \leq t_n \}] \) is the posterior estimate of \( x(t_n) \) from all measurements acquired up to (and including) the time instant \( t_n \), and \( x_-(t_n) \triangleq E[x(t_n) | \{ y(t_i), t_i < t_n \}] \) is the prior estimate of the same \( x(t_n) \) based on all measurements preceding the time instant \( t_n \). Also \( v(t_n) \) is the innovation process associated with the measurements \( y(t_n) \), and is determined via

\[
v(t_n) = y(t_n) - C(t_n) x_-(t_n).\]

Finally, the Kalman gain \( K(t_n) \) is given by

\[
K(t_n) = P(t_n) C^* (t_n) \psi^{-1} (t_n), \quad \text{where} \quad P(t_n) = E[[x(t_n) - x_-(t_n)][x(t_n) - x_-(t_n)]^*]
\]

and \( \psi(t_n) = E \{ v(t_n) v(t_n)^* \} = C(t_n) P(t_n) C^* (t_n) + R(n) \).

To complete the measurement update step, we update also the error covariance, viz.,

\[
P_+(t_n) = P(t_n) - K(t_n) \psi(t_n) K^*(t_n)
\]

where \( P_+(t_n) \triangleq E[[x(t_n) - x_+(t_n)][x(t_n) - x_+(t_n)]^*] \)

Notice that if several measurements occur at the same time instant, say \( t_{n-1} < t_n = t_{n+1} = \cdots = t_{n+m-1} < t_{n+m} \), then the Kalman filter has to carry out \( m \) consecutive measurement updates at each time instant \( \{ t_i ; n \leq i \leq n+m \} \) without any time updates in between.

Time update:
Between measurement updates, the Kalman filter carries out a continuous-time time update as follows:
\[
\dot{x}(t) = Ax(t) + Bu(t) \\
\dot{P}(t) = AP(t) + P(t)A^* + GQG^*
\]

where \( x(t) = E\{x(t) \mid y(t_n); t_n \leq t \} \) and \( P(t) = E\{[x(t) - x(t)][x(t) - x(t)]^* \} \). Notice that for any time \( t \) which is not a measurement instant, there is only one type of state estimate \( x(t) \equiv x_*(t) \equiv x_-(t) \). The differential equations (1.6) are initialized with \( x_*(t_n) \) and \( P_+(t_n) \) at time \( t_n \), and evolve in the interval \( t_n \leq t < t_{n+1} \). Notice that the time update and measurement update of \( P(t) \) cannot be computed in advance, because they depend on the locations of the sampling instants \( \{t_n\} \) and this information is not available apriori. This also means that the error covariance \( P(t) \) has to be interpreted as being conditional on the time pattern, especially when \( \{t_n\} \) are random. Thus, a more accurate definition of the state error covariance matrix \( P(t) \) is in terms of a conditional mean, namely

\[
P(t) = E\{[x(t) - x(t)][x(t) - x(t)]^* \mid \{t_n; t_n \leq t\}\}.
\]

For computational purposes, it is more convenient to replace (1.6) by equivalent integral expressions as follows:

\[
x(t) = e^{A(t-t_n)}x_*(t_n) + \int_0^{t-t_n} e^{A\tau}Bu(t-\tau)d\tau
\]

(1.7)

\[
P(t) = e^{A(t-t_n)}P_+(t_n)e^{A(t-t_n)} + \int_0^{t-t_0} e^{A\tau}GQG^*e^{A\tau}d\tau
\]

In particular, we will use these integral equations in our performance analysis in Ch. 3 to evaluate \( P_+(t_n) \) and

\[
P_-(t_{n+1}) = \lim_{\tau \to \infty} P(t) = e^{A(t_{n+1}-t_n)}P_+(t_n)e^{A(t_{n+1}-t_n)} + \int_0^{t_{n+1}-t_n} e^{A\tau}GQG^*e^{A\tau}d\tau.
\]

Notice that \( P(t) \), as given by (1.7), is a monotone increasing function of time once the initialization transients have died off. On the other hand, the measurement update step (1.5) reduces the error covariance from \( P_+(t_n) \) to \( P_-(t_n) \), as illustrated in Fig. 1.5.
A complete listing of the continuous-discrete Kalman filter equations is provided in Table 1.1. The measurement update steps are carried out at the time instants $\{t_n; 1 \leq n < \infty\}$. Each time update step starts with $t = t_n$, and evolves in the continuous interval $t_n \leq t < t_{n+1}$. Since there are no measurements in the interval $t_0 = 0 \leq t < t_1$, the Kalman filter starts its operation with a time update step, using the initialization $x_-(t_0) = 0$ and $P_-(t_0) = E\{x(t_0)x^*(t_0)\} \triangleq \Pi_0$.

Table 1.1 The continuous-discrete Kalman filter

<table>
<thead>
<tr>
<th>Measurement Update</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v(t_n) = y(t_n) - C(t_n)x_-(t_n)$</td>
</tr>
<tr>
<td>$\psi(t_n) = C(t_n)P_-(t_n)C^*(t_n) + R(n)$</td>
</tr>
<tr>
<td>$K(t_n) = P_-(t_n)C^*(t_n)\psi^{-1}(t_n)$</td>
</tr>
<tr>
<td>$x_+(t_n) = x_-(t_n) + K(t_n)v(t_n)$</td>
</tr>
<tr>
<td>$P_+(t_n) = P_-(t_n) - K(t_n)\psi(t_n)K^*(t_n)$</td>
</tr>
</tbody>
</table>
In Ch. 3, our main interest is to obtain a compact characterization of the relation between the sensor sampling pattern \( \{ t_n; 1 \leq n < \infty \} \) and the error covariance matrix \( P(t) \). For this purpose we focus only on the “steady-state” values of the two sequences \( \{ P_+(t_n); 1 \leq n < \infty \} \) and \( \{ P_-(t_n); 1 \leq n < \infty \} \), since these sequences provide lower and upper bounds, respectively, on the instantaneous \( P(t) \). In some cases (e.g., when the sampling pattern is regular) the sequence \( \{ P_+(t_n) \} \) converges to a limit value as \( n \to \infty \), and so does \( \{ P_-(t_n) \} \). However, in general, these sequences may have an oscillatory behavior even when \( n \to \infty \), so we choose to define

\[
P_{\pm, \text{av}} \triangleq \left\langle \frac{1}{N} \sum_{n=1}^{N} E(P_{\pm}(t_n)) \right\rangle \equiv \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} E(P_{\pm}(t_n))
\]

where \( \left\langle \right\rangle_n \) indicates a long-term time-average with respect to the discrete index \( n \) (a so called Cesaro limit [12]). Notice that the expectation operation \( E\{-\} \) is carried out with respect to the randomness in the sampling instants \( \{ t_n \} \). When the sampling pattern is deterministic, only the time average \( \left\langle \right\rangle_n \) is needed. The time-invariant \( P_{\pm, \text{av}} \) are used in Ch.3 to provide a meaningful measure of estimation performance with respect to sampling patterns.

### 1.3 Continuous-Discrete Controller

In control theory, one basic control law is feedback control. The control input \( u_c(t) \)
is assumed to be an instantaneous function of the plant state $x(t)$, and we could write the control law as $u_c(t) = k(x(t), t)$ to indicate the dependence of $u_c(t)$ on both $x(t)$ and $t$. As we assume a linear system model in this thesis, the optimal solution to the LQG (Linear Quadratic Gaussian) control problem is $u_c(t) = -K(t)\hat{x}(t)$, where $K(t)$ is the controller’s gain and $\hat{x}(t)$ is the optimal (in MSE sense) state estimation because we can’t get the state directly, and what we have is the estimated state instead of real state. Also in this thesis, because the system parameters are constant, the controller’s gain is time-invariant, as $K$. What makes the problem here different from LQG is that we introduce delay in the control path. Consequently, it is no longer obvious what the optimal control solution is. In any case, in this thesis, we aim to mitigate the effects of delay in the control path, so that the LQG optimal control solution $u_c(t) = -K(t)\hat{x}(t)$ is still applicable (optimal or nearly optimal). In reality, there will also be delay in the sensing path, and it will cause additional error in the state estimation $\hat{x}(t)$. However, in this thesis, we will assume that there is no delay in the sensing path, so that we can focus on the effect of delay in the control path and introduce our technique to mitigate this delay.

We also consider the question which signal we are going to send back from central estimator to actuator: the state estimate $\hat{x}(t)$, or the control signal $u_c(t) = -K(t)\hat{x}(t)$. There will be a significant difference between these two signals, especially in dimensionality. For a single actuator case, $u_c(t)$ is a scalar, but $\hat{x}(t)$ is a vector. In a complicated system like the power system, $\hat{x}(t)$ may consist of tens, even hundreds of states. However, we choose to transmit $\hat{x}(t)$ instead of $u_c(t)$.

First of all, our delay mitigation scheme relies on using $\hat{x}(t)$ and having a reasonable accurate system model. Secondly, the processing can be combined with the actuator’s D/A unit which converts digital transmission into a continuous-time control signal. The architecture of State-Estimator-Based Feedback Control is showed in Fig. 1.6.
The presence of delay in the control path degrades the quality of the control signal and may eventually lead to loss of stability. The delay mitigation technique that we propose in Ch. 2 makes it possible to overcome very large delays in the control path without loss of stability. In fact, we show that the SEER in the closed loop system decreases almost linearly (in dB/sec) as the average delay increases.

1.4 Thesis Contributions and Structure

In this thesis we propose a new method for mitigation of delay in the control path of a spatially-distributed cyber-physical system. In particular, this approach of delay canceling based on time-stamping of transmitted signals can successfully mitigate the destabilizing effects of delay. Actually the method transmutes communication delay into increased estimation error, so that the average SEER of any signal in the estimation-control loop is a monotone decreasing (and, in fact nearly linear) function of the average communication delay. We demonstrated the utility of our approach via a simple example. In particular, we showed that the value of delay in the control loop can be increased by a factor of 10 beyond the critical value \( \Delta_{cr} = 0.18s \) with only a moderate degradation (20 dB) in SEER. While longer delays (\( \Delta > 10\Delta_{cr} \)) can also be tolerated, the resulting degradation in estimator quality (\( \approx 10dB \)) may not be
acceptable in some applications. We also show that the relation between average SEER and average communication delay does not change when additional white noise is applied at sensor output and communication channel. In particular, 20dB noise brings around 20dB decrease in SEER for each average delay. Moreover, we obtain similar results if there is inaccuracy in the system parameter. Even more, we find that the SEER decrease in dB is linear with the percentage of inaccuracy of system parameters.

In addition, we also analyze the performance of a (Kalman filter) state estimator in the presence of irregular sensor sampling. In particular, for one sensor sampling, we find that the average SEER is a monotone decreasing function of the average sampling interval. More specifically, we show in detail that the average SEER mainly depends on the average sampling rate, with very weak dependence on higher order moments (i.e. variance and kurtosis) of the sampling interval. Furthermore, for the two-sensor case, we also find that the average error covariance is a monotone decreasing function of the average sampling rate of the two sensors. From above, we can say that the most significant factor that decides the estimation performance is the average sampling rate of each sensor.

This thesis is arranged in the following order: Ch. 2 proposes our delay canceling method that transmutes communication delay into increased estimation error. With this method, we improve the control performance when delay is short and stabilize the system when delay is long. Ch. 3 demonstrates the relation between estimation performance and the statistics of sampling pattern (regular and irregular). Finally, Ch. 4 is dedicated to conclusions and a brief discussion on possible directions for future research.
Chapter 2

Delay-Mitigating Estimation and Control

Communication delay is one of the unavoidable side-effects of using a computer network to interconnect sensors, actuators and control centers/nodes. The time to read a sensor measurement and to send a control signal to an actuator through the network depends on network characteristics such as topology, routing schemes and network condition. The overall performance of a network-based control system can be significantly affected by network delays. The severity of the delay problem is aggravated when data loss occurs during a transmission. Moreover, delays not only degrade the performance of a network-based control system, but can also result in instability.

In this chapter, we analyze the effect of delay in electric power systems as a special application. Electric power systems are among most spatially extended engineered systems. Their stable operation in the presence of disturbances and outages is made possible by real-time control over communication networks. This mode of operation is likely to expand in the future, as main enablers for development of sustainable energy systems are in the information layer. Examples include sensing, estimation and control of components in the energy layer.

We propose here a delay canceling approach based on time-stamping of transmitted signals, which can successfully mitigate the destabilizing effects of delay if we communicate state estimate rather than control signal from the estimation/control center to the actuators. By "time-stamping" we mean that every sensor sample $x_k(t)$ is transmitted over the network as an ordered triplet $(x,t,k)$, consisting of the sample value $x$, the corresponding sampling time instant $t$ and the identifier $k$ of the sensor. Time stamping is enabled by the availability of precise time signals derived from the Global Positioning System (GPS) in various nodes of the power system. Such signals are part of Phasor Measurement Units (PMUs) that are increasingly being deployed as a key ingredient of system-wide monitoring and control.

Our delay canceling method transmutes communication delay into increased
estimation error, so that the average signal to noise ratio of any signal in the estimation-control loop is a monotone decreasing (and, in fact nearly linear) function of the average communication delay. We demonstrate the utility of our approach via an example. In particular, we show that our time-stamping-based delay canceling method makes it possible to overcome very large delays in the control path while avoiding adverse effects on the stability of the system. In our examples the delay can be 10-20 times higher than the smallest value that will cause instability in the original control path.

In addition we examine the effects of modeling errors and additive noise on our delay mitigation scheme. Both contribute a further degradation in the overall SEER.

2.1 Delay in Observation Path

The optimum MSE estimate of the state \( x(t) \) is \( \hat{x}(t) \triangleq \hat{x}(t \mid y(t)) \), namely the projection of the state \( x(t) \) on the measurement subspace

\[
y(t) \triangleq \text{span}\{y_k(t_n^{(k)}); k = 1, 2, \ldots, K, t_n^{(k)} \leq t\}.
\]

In the presence of delay in the observation path, the measurement \( y_k(t_n^{(k)}) \) becomes available only at the time instant \( t_n^{(k)} + \Delta_n^{(k)} \), where \( \Delta_n^{(k)} \) is the (generally time-variant) delay experienced in the transmission of \( y_k(t_n^{(k)}) \) from the \( k \)-th sensor to the estimation/control center. As a result, the optimum MSE estimate of \( x(t) \) in the presence of known delay is \( \hat{x}^{(d)}(t) \triangleq \hat{x}(t \mid y^{(d)}(t)) \), namely the projection of the state \( x(t) \) on the delayed measurement subspace

\[
y^{(d)}(t) \triangleq \text{span}\{y_k(t_n^{(k)}); k = 1, 2, \ldots, K, t_n^{(k)} + \Delta_n^{(k)} \leq t\}.
\]

Since \( y^{(d)}(t) \) is a subspace of \( y(t) \), it follows that \( \hat{x}^{(d)}(t) \) has a larger error covariance, namely \( P^{(d)}(t) \geq P(t) \) for all \( t \).

We can make a more detailed statement about the relation between \( P^{(d)}(t) \) and \( P(t) \) by carefully examining the relation between \( \hat{x}^{(d)}(t) \) and \( \hat{x}(t) \). In
particular, when $\Delta_n$ is not too large, i.e., $t_n < t_n + \Delta_n < t_{n+1}$, we can explicitly solve the differential equation (2.1), viz.,

$$\frac{d}{dx} x_0^{(d)}(t) = Ax_0^{(d)}(t)$$

$$\frac{d}{dx} P_0^{(d)}(t) = AP_0^{(d)}(t) + P_0^{(d)}(t)A^* + BQB^*$$

So that

$$x_0^{(d)}(t_n + \Delta_n) = e^{A\Delta_n} x_0^{(d)}(t_n)$$

$$P_0^{(d)}(t_n + \Delta_n) = e^{A\Delta_n} P_0(t_n)e^{A\Delta_n} + \int_0^\Delta e^{A\tau} GQG^* e^{A\tau} d\tau$$

This result allows us to compare the average value of $P_0^{(d)}(\theta_n)$ with the average value of $P_0(t_n)$. As we mentioned in Sec. 1.2, $P_{\tau, av} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^N E[P_0(t_n)]$, namely, it is both time averaged as well as ensemble averaged with respect to possible randomness of the sampling time instants. Similarly, we define the average delayed error covariance $P_{\tau, av}^{(d)} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^N E[P_0^{(d)}(\theta_n)]$. To avoid unnecessary complications, we assume here that the matrix $A$ is diagonalizable, so that

$$A = T \begin{bmatrix} \lambda_1 & 0 \\ & \ddots \\ & 0 & \lambda_N \end{bmatrix} T^{-1},$$

where $\{\lambda_i\}$ are the eigenvalues of $A$ and $T$ is a non-singular (square) matrix. Using the approach described in App. A, we can obtain a simple relation between $P_{\tau, av}^{(d)}$ and $P_{\tau, av}$, viz.,

$$T^{-1}P_{\tau, av}^{(d)}T^{-*} = \Omega_{\Delta} \circ (T^{-1}P_{\tau, av}T^{-*}) + \mathcal{U}_{\Delta} \circ (T^{-1}GQG^*T^{-*})$$

where $\circ$ denotes the Schur-Hadamard product, and

$$\Omega_{\Delta} = [\phi_\Delta(\lambda_i + \lambda_m^*)]_{i,m=1}^M$$

$$\mathcal{U}_{\Delta} = \left[ \frac{1}{\lambda_i + \lambda_m^*} (\phi_\Delta(\lambda_i + \lambda_m^*) - 1) \right]_{i,m=1}^M$$

Here $\phi_\Delta(s)$ is the averaged characteristic function associated with the distribution
of the delay $\Delta_n$, namely $\phi_\Delta(s) = \langle E[e^{\Delta_n}] \rangle_n = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int_{-\infty}^{\infty} e^{\Delta_n} p(\Delta_n) d\Delta_n$, where we allow for both random and/or deterministic variation in $\Delta_n$.

The expression (2.4) provides an explicit relation between $P_{\text{av}}^{(d)}$, the average error covariance in the presence of delay, and $P_{\text{av}}$, the delay-free version of the same. Our preliminary qualitative analysis implies that $P_{\text{av}}^{(d)} \geq P_{\text{av}}$, regardless of the nature of the delay sequence $\{\Delta_n\}$. For instance, when the delay is constant (i.e., $\Delta_n = \Delta$) and relatively small, we find that $\phi_\Delta(s) = e^{s\Delta} \approx 1 + s\Delta$. With the condition $\max_i (\text{Re}(\lambda_i)\Delta)$, we have $\phi_\Delta(\lambda_i + \lambda_m^*) = 1 + (\lambda_i + \lambda_m^*)\Delta \approx 1$ and $\psi_\Delta(\lambda_i + \lambda_m^*) = \frac{1}{\lambda_i + \lambda_m^*} \approx \Delta$.

so that $\Omega_\Delta \approx \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 \end{pmatrix} \Delta \approx \Delta 1$ and $\mathcal{V}_\Delta = \Delta 1$.

This reduces (2.4) to $P_{\text{av}}^{(d)} = P_{\text{av}} + \Delta GQG^*$ (2.5)

which shows clearly that the added covariance error equals $\Delta GQG^*$, so that $P_{\text{av}}^{(d)}$ increases with increasing delay $\Delta$. As a matter of fact, $P_{\text{av}}^{(d)}$ will get close to $P_{\text{initial}}$ as the delay $\Delta$ increases if the system is stable, otherwise $P_{\text{av}}^{(d)}$ will increase to infinity if the system is unstable.

### 2.2 Delay-Mitigating Control

In this section, we consider the problem with delay in control path and a delay free in observation path. The network estimation and control configuration that we consider here consists of a single estimation center that collects observations from several sensors and provides estimated state estimation to one, or more, actuators. A communication network is used to transmit observations from the sensors to the central estimation center, and to deliver state estimate information from this entity.
to individual actuators. We assume that each actuator is equipped with a control module that translates the state estimate information received from the estimation center into an appropriate control signal, as shown in Fig. 2.1. For the sake of presentation simplicity we limit our discussion here to the case of a single actuator; however, our approach can be used with any number of sensors and actuators.

![Networked estimation/control configuration](image)

The estimation center transmits a sequence of state estimates, say $\hat{x}(\tau_n)$, to the control module attached to the actuator. The time instants $\{\tau_1, \tau_2, \ldots\}$ include, at least, all the times when $\hat{x}(t)$ underwent a measurement update. The control module receives, at the time instant $\theta_n = \tau_n + \Delta_n$, the pair $\{\hat{x}(\tau_n), \tau_n\}$, where $\Delta_n$ denotes the delay in the control path (Fig. 2.1). By comparing the time-stamp $\tau_n$ with the physical time of arrival $\theta_n$, the control module can accurately determine the value of the delay $\Delta_n = \theta_n - \tau_n$.

Notice that the estimate delivered to the control module at time $\theta_n$ refers to delay version of the state vector (i.e., at time $\tau_n$). To remedy this discrepancy, the
control module carries out a local time update, viz.,
\[
\hat{x}(\theta_n) = e^{A\Delta\theta_n} x(\tau_n) + \int_{\tau_n}^{\theta_n} e^{A(\theta_n-s)} Bu(s) ds
\]  
(2.6)
Moreover, for all \( \theta_n \leq t < \theta_{n+1} \), the state estimate used by the control unit is
\[
\hat{x}(t) = e^{A(t-\theta_n)} \hat{x}(\theta_n) + \int_{\theta_n}^{t} e^{A(t-s)} Bu(s) ds
\]  
(2.7)
It can be shown that for all \( \theta_n \leq t < \theta_{n+1} \), \( \hat{x}(t) = \hat{x}(t \mid y(\tau_n)) \), so that the control module forms the optimum MSE estimate of the current state \( x(t) \) from all the measurements acquired by the estimation center in the interval \([0, \tau_n]\). This local estimate is then used to construct the control signal \( u(t) = -K\hat{x}(t) \). The time-update equation (2.6), which maps \( x(\tau_n) \) into \( \hat{x}(\theta_n) \), induces a similar relation between the associated error covariance, viz.,
\[
\hat{P}(\theta_n) = e^{A\Delta\theta_n} P_v(\tau_n) e^{A\Delta\theta_n} + \int_{0}^{\Delta\theta_n} e^{A\theta} GQG^* e^{A\theta} ds
\]  
(2.8)
This is the same relation as (2.3), with \( \hat{P}(\theta_n) \), \( P_v(\tau_n) \) replacing \( P_v^{(d)}(\theta_n) \), \( P_v(\tau_n) \), respectively. Thus our delay-mitigation method results in increased error covariance, but at the same time, avoids the destabilizing effect of delay in the control path.

The presence of delay in the control path may results in loss of stability of the overall system. This is true, in particular, when one of the functions of the control loop is to stabilize an unstable open loop system. Thus the primary objective of our control module is to mitigate the effects of delay in the control path, so as to avoid loss of stability of the closed loop system. The method that we use will allow us to overcome very large delay (at least 10-20 times higher than the smallest destabilizing delay in the original control loop) at the cost of increased state estimation error.

### 2.3 Example: controlling a simple unstable system

We use a single-input single-output example to illustrate the capability of our delay-mitigation method. The system parameters we use in the simulation are listed as below.
This is an unstable system with (continuous-time) poles at $\pm \sqrt{5}$. We stabilize it by applying feedback control with $K = [17 \ 7]$, so that the poles of the closed-loop system are stable (located at $-4$ and $-3$). When delay is present in the control feedback path, the closed-loop system becomes unstable for $\Delta \geq 0.18s \triangleq \Delta_{cr}$. In contrast, using our delay mitigation technique, the closed-loop system remains stable, with delays as high as $\Delta = 10s$. However, the level of estimation error increases with the increase of delay, as shown in Fig. 2.2: the SEER decreases when delay increases. Notice that we still have $SEER \approx 18\text{dB}$ when $\Delta = 2s$, a delay value that is 10 times higher than the critical delay $\Delta_{cr} = 0.18s$, without delay mitigation (top curve in Fig. 2.2).

![Fig. 2.2](image) Effect of delay mitigation on SEER, in the presence of observation noise

The effect of additive noise at the sensor output or in the communication channel is captured by the lower two curves in Fig. 2.2. In one case (dashed line), white noise is added to the acquired observations $y(t)$. In the other (circle line),
noise is added to the transmitted estimate $x(t)$. The level of noise in each case is adjusted to achieve a local signal to noise ratio of 20dB. All three curves in Fig. 2.2 display the same dependence on the length of delay ($\Delta$), and illustrate the fact that our delay-mitigation technique transmutes delay into increased estimation error.

Our method for mitigating delay in the control path relies on availability of the system parameter A, B at the control module, so that (2.6) and (2.7) can be used to evaluate $\hat{x}(t)$. Inaccuracies in system parameters will result in further degradation in the quality of the estimation $\hat{x}(t)$, as demonstrated in Fig. 2.3. The two lower curves in this figure correspond to element wise inaccuracy in A and B of 1% (dash line) and 5% (circle line). Clearly, higher inaccuracy results in higher estimation error level (equivalently lower SEER), and this relation is approximately linear for small levels of inaccuracy.

![SEER vs Delay involving inaccuracy of A and B](image)

Fig. 2.3  Effect of delay mitigation on SEER, in the presence of system parameters inaccuracy

2.4 Concluding Remarks

In this chapter we have proposed a delay canceling approach based on
time-stamping of transmitted signals, which can successfully mitigate the destabilizing effects of delay. Time-stamping is enabled by the availability of precise time signals derived from the Global Positioning System (GPS).

Our delay canceling method transmutes communication delay into increased estimation error, so that the average SEER of any signal in the estimation-control loop is a monotone decreasing function of the average communication delay. We demonstrated the utility of our approach via a simple example. In particular, we have shown (see Fig. 2.2) that the value of delay in the control loop can be increased by a factor of 5-10 beyond the critical value $\Delta_{cr}$ with only a moderate degradation (10-20 dB) in the signal-to-estimation error ratio. While longer delays can also be tolerated, the resulting degradation in estimator quality may not be acceptable in some applications.

Finally we observe that our time-stamping-based approach can be used with any state-estimation method and any feedback-control strategy, not just Kalman filtering and LQG state-feedback control. In particular, robust estimation and control methods ([13], [14], [15]) may be required to accommodate significant uncertainty in system parameters (and models). Endowing such robust estimation and control techniques with delay mitigation capability will be the key to successful implementation in spatially-extended systems.
Chapter 3

Effects of Sensor Sampling Patterns

The need to transmit sensor measurements over digital communication channels makes it necessary to sample sensor data at a discrete-time sequence of sampling instants. In this chapter, we study the problem of state estimation of linear dynamic continuous system with discrete-time measurement. The sequence of sensor sampling times, which we call a sensor sampling pattern (SSP), could be irregular, in general.

It is apparent that the sampling rate is a critical parameter in the application of control system. It is also reasonable that as the sampling rate is increased, the performance of the control system improves. However, computation costs also increase because less time is available to do state estimation and the controller equation, and thus high performance processors must be used. Additionally, if the system uses A/D converters, a higher sampling rate requires faster A/D conversion speed which may also increase system costs. On the other hand, reducing the sampling rate for the sake of reducing costs may degrade the performance or even cause instability. In addition to performance and computational cost consideration, there are additional constraints imposed by the use of a communication network. In other words, we can no longer rely on the Nyquist criterion in selecting the sensor sampling rate. Rather, we need to take into account the (possibly time-varying) constraints imposed by the communication network, and the relation between sensor sampling patterns and the estimation performance. We may even consider changing the sampling pattern in response to changing network conditions.

A regular SSP is completely characterized in terms of its sampling interval $T_s$. An irregular SSP has a time-varying sampling interval $T_s(n)$, which we choose to characterize in terms of its moments: the average and the variance of $T_s(n)$, as well as higher-order moments, as needed. We show in this chapter that the average signal to estimation error ratio (SEER) depends primarily on the average sampling interval, with a very minor dependence on the variance and a negligible dependence
on higher-order moments of time-varying sampling interval. The best performance (i.e., smallest estimation error) for a single-sensor system is achieved when the SSP is regular, so that all its higher-order moments vanish.

The same conclusions hold also for a Kalman filter estimation that employs multiple sensors: the average estimation error covariance is a monotone decreasing function in each one of the average sampling rates. In addition, the average estimation error covariance depends also on the relative alignment of the individual SSPs. The best performance is achieved when the individual SSPs are regular and aligned in such a way that the superposed set of sampling instants is as close to regular as possible.

We consider the same dynamic model that we introduced in Ch. 1, viz.,

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t) + Gw(t) \\
y(t_n) &= C(t_n)x(t_n) + \eta(n)
\end{align*}
\]  

(3.1)

where \( E[w(t)w^*(\tau)] = Q(\tau)\delta(\tau - \tau), \ E[\eta(n)\eta^*(n-k)] = R(n)\delta(k) \) and \( \{t_n\}_{n=0,\infty} \) is the sequence of sampling time instants. Several types of sensor sampling patterns are described in Sec. 3.1.

The performance of a continuous-discrete Kalman filter, as described in Table 1.1, is a function of the SSPs we employ. A detailed analysis of the single-sensor case is provided in Sec. 3.2 and 3.3. The multiple sensor case is discussed in Sec. 3.4.

### 3.1 Sensor Sampling Pattern Classification

An SSP for a single sensor is the sequence \( \{t_n; n \geq 0\}_n \) of sampling times. We can assume without loss of generality that \( t_0 = 0 \), and concern ourselves only with the time variation of the sampling interval \( T_s(n) \triangleq t_n - t_{n-1} \) for \( n \geq 1 \). This variation could be deterministic, (e.g., periodically-variant) or random. To be specific, we will consider 5 distinct SSP, as follows:

**1) Regular sampling:**

\( T_s(n) = t_n - t_{n-1} = \Delta \) for all \( n \geq 1 \)

Regular sampling is the most commonly used sampling pattern. Samples are evenly spaced in time with a fixed interval \( \Delta \), so the average sampling rate:

\( F_{av} = 1/\Delta. \)
(2) Piecewise-regular sampling:

\[
t_n - t_{n-1} = \begin{cases} \Delta_1 & 1 \leq n \leq N_1 - 1 \\ \Delta_2 & N_1 \leq n \leq N_2 - 1 \\ \vdots & \vdots \\
\end{cases}
\]

This type of sampling pattern can occur when we intentionally adjust the sampling rate in response to significant changes in operating conditions. For instance, while a reasonable low sampling rate (under 3000 samples/sec) may be sufficient during steady-state operation of a power system, a much higher rate (more than 20K samples/sec) is needed when transients occur. From a short term perspective, the sampling is regular over the interval \( N_r \leq n \leq N_{r+1} - 1 \), with rate \( \frac{1}{\Delta_r} \). The average sampling rate is a long term property, and is given by

\[
F_{av} = \lim_{r \to \infty} \frac{N_1 + N_2 + \cdots + N_r}{N_1\Delta_1 + N_2\Delta_2 + \cdots + N_r\Delta_r}.
\]

In other words, we have regular sampling with an occasional step change in the sampling rate, in response to change in system status.

(3) Periodically-variant sampling:

The sample interval \( T_r(n) = \Delta(1 + \alpha \sin(\omega n)) \), as shown in Fig. 3.1, where \( \Delta \) is the average sample interval, \( \alpha \ (0 \leq \alpha < 1) \) is the relative amplitude of the \( T_r(n) \) fluctuation and \( \omega \) is the frequency of fluctuation. When \( \alpha = 0 \), \( T_r(n) = \Delta \), so that periodic sampling includes regular sampling as a special case.
(4) **Uniformly distributed random sampling:**

\[ T_s(n) = \Delta (1 + \alpha X) \]

where \( \Delta \) is the average sample interval, \( \{X\} \) is an i.i.d. sequence with each \( X_n \) uniformly distributed in the interval \([-1,1]\), and \( \alpha \) (\( 0 \leq \alpha < 1 \)) determines the variance of the fluctuation of \( T_s(n) \), as shown in Fig. 3.2. When \( \alpha = 0 \), \( T_s(n) = \Delta \), so that uniform distributed random sampling includes regular sampling as a special case.
(5) Bernoulli-drop random sampling

$T_s(n) = t_n - t_{n-1}$ is an i.i.d. sequence with each $T_s(n)$ geometrically distributed, namely, $P(T_s(n) = k\Delta) = (1 - p)p^{(k-1)}$, for $k \in \{1, 2, 3, \ldots\}$, as shown in Fig. 3.3, where $(0 \leq p < 1)$. This sampling pattern occurs when a regularly sampled sensor output suffers an occasional loss of a sample caused by a packet loss in a communication network. If the packet dropping process is Bernoulli with drop probability $p$ then the resulting distribution is known as geometric, as described above. The average sampling interval is $\frac{\Delta}{1 - p}$, so that the average sampling rate $F_{av} = \frac{1 - p}{\Delta}$. When $p = 0$, this case reduces to regular sampling with interval $\Delta$.

![Probability density function for the geometric distribution](image)

Fig. 3.3 Probability density function for the geometric distribution

In this thesis we consider mostly periodically-variant sampling and uniformly distributed random sampling to study the effects of SSP on the estimation error covariance of a (continuous-discrete) Kalman filter. Both patterns allow independent control of the mean value of $T_s(n)$ via the parameter $\Delta$, and of the variance of $T_s(n)$ via the parameter $\alpha$ as shown in Table 3.1.
Table 3.1 Statistics of the SSPs

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Variance</th>
<th>( \xi = \frac{\text{Var}(T_i(n))}{\text{Mean}(T_i(n))^2} )</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periodically variant</td>
<td>( \Delta )</td>
<td>( \frac{\Delta^2 \alpha^2}{2} )</td>
<td>( \frac{\alpha^2}{2} )</td>
<td>( -\frac{3}{2} )</td>
</tr>
<tr>
<td>Uniformly Distributed Random</td>
<td>( \frac{\Delta^2 \alpha^2}{3} )</td>
<td>( \frac{\alpha^2}{3} )</td>
<td></td>
<td>( -\frac{6}{5} )</td>
</tr>
<tr>
<td>Bernoulli drop Random</td>
<td>( \frac{\Delta}{1-p} )</td>
<td>( \frac{p}{(1-p)^2} \Delta^2 )</td>
<td>( p )</td>
<td>( \frac{(1-p)^2}{p} + 6 )</td>
</tr>
</tbody>
</table>

One convenient way to characterize a sensor sampling pattern is in terms of its averaged characteristic function \( \phi_{T_i}(s) \equiv \mathbb{E}\{e^{isT_i(n)}\}_n \). For instance, the characteristic function of the periodically-variant SSP is \( \phi_{T_i}(s) = e^{is\Delta} I_0(\alpha \Delta s) \), where \( I_0(*) \) is the modified zero-order Bessel function (see App. C for a detailed derivation). Various moments of the sampling interval can be conveniently determined from the moment-generating property of the characteristic function, viz.,

\[ \frac{d^k \phi_{T_i}(s)}{ds^k} \bigg|_{s=0} = \mathbb{E}\{T_i(n)^k\}_n \]

The characteristic functions of all the SSPs that interest us are given in Table 3.2 (see derivation in App. C).

Table 3.2 Characteristic functions of the SSPs

<table>
<thead>
<tr>
<th></th>
<th>characteristic function ( \phi_{T_i}(s) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periodically variant Sampling</td>
<td>( e^{is\Delta} I_0(\alpha \Delta s) )</td>
</tr>
<tr>
<td>Uniformly Distributed Random Sampling</td>
<td>( e^{is\Delta} \frac{\sinh(\alpha \Delta s)}{\alpha \Delta s} )</td>
</tr>
<tr>
<td>Bernoulli drop Random Sampling</td>
<td>( \frac{(1-p)e^{is\Delta}}{1-p e^{is\Delta}} )</td>
</tr>
</tbody>
</table>
3.2 The Effects of SSP on Kalman Filter Performance

The state error covariance matrix $P(t)$ of a Kalman filter decreases with every measurement update, and increases in between measurement update (recall Fig. 1.5). When the sampling pattern is regular, the post-update covariance $P_n(t_n)$ converges to a steady-state limit, and similarly for the pre-update covariance $P_{-}(t_n)$. However, when the SSP is irregular, the two sequences $\{P_n(t_n); n = 1, 2, \cdots\}$ and $\{P_{-}(t_n); n = 1, 2, \cdots\}$ may oscillate and not converge. This motivated us to introduce the average error covariance

$$P_{+\text{av}} \triangleq \langle E\{P_n(t_n)\} \rangle_n$$
$$P_{-\text{av}} \triangleq \langle E\{P_{-}(t_n)\} \rangle_n$$

as a measure of steady state performance. Our objective in this section is to obtain a simple characterization of $P_{+\text{av}}$ and $P_{-\text{av}}$ in terms of the system parameters $A, B, C, Q, R$ and the SSP.

3.2.1 The Dynamics of $P(t)$

The time update of $P(t)$ gives rise to an explicit relation between $P_{+\text{av}}$ and $P_{-\text{av}}$, similar to the relation we obtained in (2.3). When the sampling pattern is regular, we can also obtain a dual relation from the steady state version of the measurement update. With two matrix equation in the two (matrix) unknown $P_{+\text{av}}$ and $P_{-\text{av}}$, we can, in principle, obtain explicit expressions for the average error covariance in terms of the system parameter $A, G, C, Q, R$ and the characteristic function $\phi_T(s)$ of the sampling interval $T_c(n)$. However when the SSP is irregular we can only deduce an approximate relation from the measurement update, and only under the assumption of relatively small deviations from regularity in the SSP.
Average time update relation:

The time update relation in the sampling interval $[t_{n-1}, t_n)$ is (recall Table 1.1)

$$P_-(t_n) = e^{AT_n}(n)P_+(t_{n-1})e^{AT_n(n)} + \int_0^{T_{n}(n)} e^{AT_n(n)}GQGE^{AT_n(n)}d\tau$$  \hspace{1cm} (3.2)

where we use the fact that $T_{n}(n) = t_n - t_{n-1}$. This relation has the same form as (2.3), so that the analysis of App. A can be applied to deduce an average relation of the same form as (2.4). We obtain a relation between $P_{+,\text{av}}$ and $P_{-,\text{av}}$, namely

$$T^{-1}P_{-,\text{av}}T^+ = \Omega_{T_n} \circ (T^{-1}P_{+,\text{av}}T^+) + \mathcal{V}_{T_n} \circ (T^{-1}GQGT^+)$$  \hspace{1cm} (3.3)

where

$$\Omega_{T_n} = [\phi_{T_n}(\lambda_i + \lambda_m^*)]_{i,m=1}^M$$

$$\mathcal{V}_{T_n} = [\frac{1}{\lambda_i + \lambda_m^*}(\phi_{T_n}(\lambda_i + \lambda_m^*) - 1)]_{i,m=1}^M$$  \hspace{1cm} (3.4)

are determined by $\phi_{T_n}(s) = \left\{E[e^{iT_n(n)}]\right\}_n = \int_{-\infty}^{\infty} e^{iT_n(n)}p(T_n(n))dT_n(n)$, the characteristic function of sampling interval $T_n(n)$. Here $\circ$ denotes a Schur-Hadamard product, and $\lambda_i$ are the eigenvalues of the system matrix $A$. As in Sec. 2.1 (and App. A), we need to assume here that the matrix $A$ is diagonalizable, say $A = T\Lambda T^{-1}$. Notice that the relation (3.3) is static, in the sense that $G, Q, \Omega_{T_n}, \mathcal{V}_{T_n}$ and $P_{+,\text{av}}, P_{-,\text{av}}$ are all constant (time invariant) quantities.

The equation (3.4) depends on the specific SSP via the characteristic function $\phi_{T_n}(\lambda_i + \lambda_m^*)$. For instance, when the SSP is periodically-variant (see Table 3.2),

$$\phi_{T_n}(s) = e^{i\Delta}I_0(\alpha\Delta s) = e^{i\Delta}[1 + \left(\frac{\alpha\Delta s}{2}\right)^2 + \frac{1}{2!}\left(\frac{\alpha\Delta s}{2}\right)^4 + \cdots]$$

where we use the Taylor series expansion for the modified zero-order Bessel function $I_0(\cdot)$. Similarly, when the SSP is uniformly distributed random,

$$\phi_{T_n}(s) = e^{i\Delta} \frac{\sinh(\alpha\Delta s)}{\alpha\Delta s} = e^{i\Delta}[1 + \frac{1}{3!}(\alpha\Delta s)^2 + \frac{1}{5!}(\alpha\Delta s)^4 + \cdots]$$

When the fluctuation in $T_n(n)$ is not too large, i.e., when $\max_{i,m} \alpha\Delta |\lambda_i + \lambda_m^*| \ll 1$, then both types of SSP can be approximated by $\phi_{T_n}(s) = e^{i\Delta}[1 + \mu(\alpha\Delta s)^2]$.  \hspace{1cm}
where $\mu = \frac{1}{4}$ for the periodically-variant case and $\mu = \frac{1}{6}$ for the uniformly distributed random one. This expression shows that $\Omega_{\tau}$ and $\Theta_{\tau}$ depends mainly on $\Delta$ (the average value of $T_{\tau}(n)$), and to a much lesser extent on the fluctuation amplitude $\alpha$. We will verify this observation via numerical examples in Sec. 3.3.

**Average measurement update relation:**

To complete the analysis, we need to obtain an average version of measurement update relation

$$P_+(t) = P_-(t) - P_-(t)C^*\psi^{-1}(t)\psi(t)\psi^{-1}(t)CP_-(t)^*$$

$$= P_-(t) - P_-(t)C^*\psi^{-1}(t)CP_-(t)^*$$

$$= P_-(t) - P_-(t)C^*\psi^{-1}(t)CP_-(t)$$

$$= P_-(t) - P_-(t)C^*(CP_+(t)C^* + R)^{-1}CP_-(t)$$

This will provide a second equation relating the two unknown matrices, $P_{+, av}$ and $P_{-, av}$. Unfortunately, (3.5) is nonlinear in $P_+(t)$, so that it appears difficult to express the average $P_{+, av} = \left\{P(t_n) - P(t_n)C^*(CP_+(t_n)C^* + R)^{-1}CP_-(t_n)\right\}_n$ in terms of $P_{-, av} \triangleq \left\{P(t_n)\right\}_n$. A more compact form of this measurement update relation can be obtained by applying to (3.5) the matrix inversion lemma, viz.,

$$(A + BDC)^{-1} = A^{-1} - A^{-1}B(D^{-1} + CA^{-1}B)^{-1}CA^{-1}$$

(3.6)

with

$$\begin{align*}
A^{-1} &\rightarrow P_+(t) \\
B &\rightarrow C^* \\
C &\rightarrow C \\
D^{-1} &\rightarrow R
\end{align*}$$

We obtain

$$P_+(t_n) = (P_+(t_n))^{-1} + C^*R^{-1}C^{-1},$$

so that

$$P^{-1}_+(t_n) = P^{-1}_-(t_n) + C^*R^{-1}C$$

(3.7)

which also makes it evident that always $P_+(t_n) \leq P_-(t_n)$. Averaging results in

$$\left\{E[P^{-1}_+(t_n)]\right\}_n = \left\{E[P^{-1}_-(t_n)]\right\}_n + C^*R^{-1}C$$

(3.8)

Unfortunately, there is no simple relation between $\left\{E[P_+(t)]\right\}_n$ and $\left\{E[P^{-1}_+(t)]\right\}_n$,.
so that, in general, we cannot get a simple average version of the measurement update relation. Notice, however that when the SSP is regular we obtain (in steady state) $P_+(t_n) = P_{+av}$ and $P_-(t_n) = P_{-av}$, so that in this case (3.7) becomes the desired (static relation)

$$P_{+av}^{-1} = P_{-av}^{-1} + C^* R^{-1} C$$

(3.9)

A perturbation analysis reveals that this relation also holds, approximately, for mildly irregular SSPs (see App. D for details).

$$P_{+av}^{-1} \approx P_{-av}^{-1} + C^* R^{-1} C$$

(3.10)

In summary, for mildly irregular, we obtain both an average time update relation and an average measurement update relation, viz.,

$$\begin{cases}
T^{-1}P_{-av}T^x = \Omega_{\tau_x} \circ (T^{-1}P_{+av}T^\ominus) + \mathcal{V}_{\tau_x} \circ (T^{-1}GQ^*T^\ominus) \\
P_{+av}^{-1} \approx P_{-av}^{-1} + C^* R^{-1} C
\end{cases}$$

(3.11)

where

$$\Omega_{\tau_x} = [\phi_{\tau_x} (\lambda_i + \lambda^*_m)]_{i,m=1}^M$$

$$\mathcal{V}_{\tau_x} = [\frac{1}{\lambda_i + \lambda^*_m} (\phi_{\tau_x} (\lambda_i + \lambda^*_m) - 1)]_{i,m=1}^M$$

(3.12)

The accuracy of the approximation (3.10) depends on the level of fluctuation of $P_+(t_n)$ around its average value $P_{+av}$, as well as the fluctuation of $P_-(t_n)$ around its average value $P_{-av}$. For instance, if we use the uniformly-distributed random SSP, then the accuracy of (3.10) depends on the parameter $\alpha$. Perfect accuracy is achieved when $\alpha = 0$ (regular sampling), but as the value of $\alpha$ increases, the accuracy of (3.10) deteriorates. The following calculation of $P_{+av}$ and $P_{-av}$ is based on the approximation (3.10), which assumes a relatively small fluctuation of $P_\pm(t_n)$.

**Explicit expressions for** $P_{\pm av}$

Now, we turn to study the behavior of $P_{+av}$ as a function of average sampling rate, and the SSP under the assumption of mild irregularity. This will allow us to demonstrate that fluctuation of the sampling interval $T_s(n)$ have a relatively minor
effect on $P_{\text{av}}$ as compared with the dependence on the average sampling rate.

The set of nonlinear equations (3.11) provides an implicit characteristic of the average error covariance $P_{\text{av}}$ and $P_{\text{av}}^\pm$. This set of equations can be solved numerically for given system parameters $A, G, C, Q, R$ and a given SSP. However, our objective here is to obtain a qualitative understanding for the dependence of $P_{\text{av}}$ on these parameters. To this end we consider the simplest possible version of the problem, involving a single-state system, so that every quantity in (3.11) is a scalar (and $T=1$).

Now (3.11) reduces to be a quadratic equation in $P_{\text{av}}$, viz.,

$$R_s\Omega_T P_{\text{av}}^2 + (Q_s R_s \mathfrak{U}_T - \Omega_T + 1)P_{\text{av}} - Q_s \mathfrak{U}_T = 0$$

(3.13)

where $Q_s = GQ^*$, $R_s = C^*R^{-1}C$ and

$$\Omega_T = \phi_T(A + A^*) = \phi_T(2 \text{Re}(A)),$$

$$\mathfrak{U}_T = \frac{1}{(A + A^*)} (\phi_T(A + A^*) - 1) = \frac{\Omega_T - 1}{2 \text{Re}(A)}$$

(3.14)

Because of the property of characteristic function, we have $\Omega_T > 1$ when $\text{Re}(A) > 0$ (unstable) and $0 < \Omega_T < 1$ when $\text{Re}(A) < 0$ (stable). We also have $\mathfrak{U}_T$ always be positive from the above results.

Now we need to discuss the properties of this quadratic equation. The discriminant $b^2 - 4ac = (Q_s R_s \mathfrak{U}_T - \Omega_T + 1)^2 + 4R_s Q_s \Omega_T \mathfrak{U}_T \geq 0$, so that we have two real roots. Then we look at "$c", \ c = -Q_s \mathfrak{U}_T < 0$, so that we have one positive and one negative root. The positive root is

$$P_{\text{av}} = \frac{-(Q_s R_s \mathfrak{U}_T - \Omega_T + 1) + \sqrt{(Q_s R_s \mathfrak{U}_T - \Omega_T + 1)^2 + 4R_s Q_s \Omega_T \mathfrak{U}_T}}{2R_s \Omega_T}$$

(3.15)

so this is the expression for $P_{\text{av}}$ which only depends on $Q_s$, $R_s$, $\text{Re}(A)$ and $\phi_T(s)$. For instance, when the SSP is regular, we have the characteristic function $\phi_T(s) = e^{s\Delta}$, so that $P_{\text{av}}$ depends only in $\Delta$, the average sampling interval (for given $Q_s$, $R_s$ and $\text{Re}(A)$). When the SSP is mildly irregular we can use the
approximation $\phi_\tau(s) = e^{i\alpha}[1 + \mu(\alpha \Delta s)^2]$, where $\mu = \frac{1}{4}$ for the periodically-variant case and $\mu = \frac{1}{6}$ for the uniformly-distributed random case.

We can now plot $P_{r,\mathrm{av}}$ as a function of the average sampling rate $F_{\mathrm{av}} = \frac{1}{\Delta}$, for several values of $\alpha$ (the “amplitude” of $T_\tau(n)$ fluctuation). We set system parameters to the values $Q_r = 1 = R_r$, $A = -1$, and we consider a uniformly-distributed random $T_\tau(n)$, so that $\phi_\tau(s) \approx e^{i\alpha}(1 + \frac{1}{6}(\alpha \Delta s)^2)$. The resulting plot of $P_{r,\mathrm{av}}$ as a function of $F_{\mathrm{av}}$ is shown in Fig. 3.4. Clearly $\alpha$ has a very negligible effect over the range $0.1 \leq \alpha \leq 0.9$, as is evident from the top plot. The zoomed-in version (bottom plot) shows that $P_{r,\mathrm{av}}$ increases slightly with bigger $\alpha$.

There we use the fluctuation index $\xi = \frac{\left\langle \text{Var}[T_\tau(n)]\right\rangle}{\left\langle \text{E}[T_\tau(n)]\right\rangle^2}$, as a universal metric for comparison with other SSP choices. In the example of uniformly random distribution SSP, $\xi = \frac{\sigma^2}{\Delta^2} = \frac{\alpha^2}{3}$.

Although the measurement update expression (3.9) may be inaccurate for high values of $\alpha$, we will demonstrate via simulation in Sec. 3.3 that $P_{r,\mathrm{av}}$ is essentially unaffected by fluctuations in $T_\tau(n)$, and depends mainly on the average sampling interval $\left\langle \text{E}[T_\tau(n)]\right\rangle$. We will also confirm that $P_{r,\mathrm{av}}$ increase mildly as the fluctuation index $\xi$ increases.
Fig. 3.4 $P_{r,av}$ from equation vs $F_s$ in irregular SSP with scalar parameter

The top plot in Fig. 3.4 demonstrates the fact that, regardless of the SSP type, in this (stable) example, the average error covariance $P_{r,av}$ decreases as the average sampling rate $F_{av}$ increases. The curve is monotone decreasing and converges, as $F_{av} \to \infty$, to an asymptotic positive value that depends on the system parameters.
On the other hand, as $F_{av} \to 0$, the error covariance $P_{+,av}$ converges to the steady-state covariance $\Pi_0$, which is the solution of the algebraic equation:

$$A\Pi_0 + \Pi_0 A^* + GQG^* = 0.$$  

In our scalar example, this limit value is $\Pi_0 = \dfrac{GQG^*}{-2A} = 0.5$. The corresponding value of $P_{+,av}$ is given, via (3.9) as

$$P_{+,av} \big|_{F_{av}=0} = [\Pi_0^{-1} + C^* R^{-1} C]^{-1} = \frac{1}{3} \quad (3.16)$$

The preceding conclusions remain essentially unchanged when the underlying system is unstable. Plotting $P_{+,av}$ as a function of $F_{av}$ when $A=1$ results in Fig. 3.5 which is very similar to the top plot in Fig. 3.4. The main difference is that $P_{+,av} \big|_{F_{av}=0} = \frac{1}{3}$ when stable, but $P_{+,av} \big|_{F_{av}=0} = 1$ when unstable. Using (3.9) we can calculate $P_{+,av} \big|_{F_{av}=0} = \frac{1}{2} = \Pi_0$ when stable, but $P_{+,av} \big|_{F_{av}=0} = +\infty$ when unstable. So when there is no measurement update, the average error covariance of a stable system will still be the initial $\Pi_0$, while for an unstable system $P_{-,av}$ becomes infinite as $F_{av} \to 0$.

![Fig. 3.5](image)

**Fig. 3.5** $P_{+,av}$ vs $F_{av}$ in regular SSP with unstable system
3.2.2 Known Performance Bounds

Our preceding analysis highlighted the difficulties in obtaining closed form expressions for $P_{+,av}$ and $P_{-,av}$. Perhaps this is why some researchers have focused on obtaining upper and/or lower bounds on $P_{+,av}$, such as in [16]. The authors of this paper study the effects of random (Bernoulli) loss of measurements in a regular sampling scheme on the performance of a Kalman filter. They use a discrete-time system model based on the sampling interval used to acquire measurements. Each measurement is either delivered to the Kalman filter (with probability $1-p$) or lost (with probability $p$). This is equivalent, as explained in Sec. 3.1, to a random sampling scheme with $P(T_s(n)=k\Delta)=(1-p)p^{k-1}$ ($k \geq 1$). So that the average sampling interval $\langle E[T_s(n)] \rangle_n$ and the average sampling rate $F_{av} = \frac{1}{\langle E[T_s(n)] \rangle_n}$ both depend on $\Delta$ and $p$. Clearly, as $p$ increases the average sampling rate decreases and $P_{+,av}$ should increase. Notice that, since the SSP is a stationary random process, $E[P_+(t_n)]$ is asymptotically stationary, so that

$$P_{+,av} \triangleq \langle E[P_+(t_n)] \rangle_n = \lim_{n \to \infty} E[P_+(t_n)].$$

The main result of [16] for our purpose is the double bound

$$0 < S \leq \lim_{n \to \infty} E[P_+(t_n)] \leq V \quad (3.17)$$

The lower bound $S$ satisfies the Lyapunov equation $S = pFSF^* + \Delta BBQ^*$ where $p$ is the probability of dropping an observation, $\Delta = \frac{1}{F_s}$ is the regular sampling interval, and $F = e^{A\Delta}$. The upper bound $V$ is the solution of the Modified Algebraic Riccati Equation (MARE): $V = FVF^* + \Delta BBQ^* - (1-p)FV(CV^* + R)^{-1}CVF^*$. Since $F_{av} = F_s(1-p)$, we can numerically evaluate both $S$ and $V$ as a function of $F_{av}$, and obtain upper and
lower bounding curves for the desired curve of $P_{r,av}$ vs. $F_{av}$.

The results in [16] indicate that the Kalman filter breaks down, namely $P_r(t_n)$ diverges to infinity when the probability of losing a measurement exceeds a certain critical value $p_r$, where $p_r \leq e^{-2\sigma_{\max}^}\}$ and $\sigma_{\max} = \max\{\Re(\lambda_i(A))\}$. Clearly this bound is useful only when $A$ is unstable since then $\sigma_{\max} > 0$ and $p_r < 1$, so that $F_{cr} \triangleq F_r(1-p_r) > 0$. For stable systems we apparently have $F_{cr} < 0$, so that any average sampling rate can be used. However recall that $P_{r,av}$ will increase as average sampling rate goes down.

A different approach is used in [17] whose authors analyze the probability distribution of $\{P_r(t_n)\}$, assuming an i.i.d. sequence of sampling intervals $\{T_i(n)\}$ and a single-state, single-input and single-measurement system. The main result is expressed in term of exceedance probability, namely $\Pr\{P_r(t_n) \leq P\} > \alpha$ in steady state, where $\alpha \in (0,1)$, for all $t_n$ and all average sampling rates above a threshold $F_{th}$, which is specified in terms of $\alpha$, $P$ and the system parameters. Assuming an unstable system $A > 0$, we choose $P \in (\frac{1}{2} \frac{R}{C^2}, \frac{R}{C^2})$ arbitrarily. We shall have that $\Pr\{P_r(t_n) \leq P\} > \alpha$, for any choice of $F_{av} > F_{th}$, where

$$F_{th} = 2A \log(1-\alpha) \left\{ \log \left( \frac{\frac{R}{C^2} - P}{\frac{R}{C^2} + \frac{G^2Q}{2A}} \right) \right\}^{-1} \left[ \frac{\frac{G^2Q}{2A} P \left( \frac{R}{C^2} - \frac{G^2Q}{2A} \right)}{\left( \frac{R}{C^2} + \frac{G^2Q}{2A} \right)^2 + P \left( \frac{R}{C^2} - \frac{G^2Q}{2A} \right)} \right]$$

This result can be converted into a lower bound on $E\{P_r(t_n)\}$. Since the fundamental result holds for all $F_{av}$ above the threshold value $F_{th}$, we could imagine that at the given rate $F_{av}$ we should have an equality

$$\Pr\{P_r(t_n) \leq P\} = \alpha \quad \text{for all } t_n$$

and using the Markov inequality $\Pr\{|X| \geq a\} \leq \frac{E[|X|]}{a}$, we can deduce that at the
given rate $F_{av}$,

$$E[P_{\alpha}(t_n)] \geq P(1-\alpha) \Rightarrow P_{\alpha,av} \geq P(1-\alpha)$$  \hfill (3.18)

This provides another lower bound for the $P_{\alpha,av}$ vs. $F_{av}$ curve.

An alternative, non-statistical, set of bounds on estimation performance is introduced in [11], viz.,

$$[J_N(t_n) + G_N^{-1}(t_n)]^{-1} \leq P_{\alpha}(t_n) \leq J_N^{-1}(t_n) + G_N(t_n)$$  \hfill (3.19)

where $J_N(t_n)$ and $G_N(t_n)$ are the “complete observability’’ and “complete controllability” Grammians, also known as the “stochastic observability and controllability” Grammians. These Grammians are defined as

$$J_N(t_n) = \sum_{p=-N}^{N} \varphi^*(t_p,t_n)C^*(t_p)R^{-1}(t_p)C(t_p)\varphi(t_p,t_n)$$  \hfill (3.20)

and

$$G_N(t_n) = \int_{t_n-N}^{t_n} \varphi(t_n,t)GQG^*\varphi^*(t_n,t)dt$$  \hfill (3.21)

where $\varphi(t_p,t_n) = e^{A(t_p-t_n)}$. Notice that both Grammians depend explicitly on the sampling instants $\{t_n\}$. The bounds hold for every individual $P_{\alpha}(t_n)$ and for every choice of SSP, without any time- or ensemble-averaging.

In order to convert this result into a static bound on $P_{\alpha,av}$, we would need to average the upper and lower bounds in (3.19). Here we have the same challenge as in obtaining the averaged version of the measurement-update relation (3.7). One way to get around this difficulty is by considering the case of a uniformly completely observable and controllable state space model [11]. A state space model is considered “uniformly completely observable’’ if there exists an integer $N$ such that

$$0 < \alpha I \leq J_N(t_n) \leq \beta I$$  \hfill (3.22)

for all $n \geq N$, where $\alpha$ and $\beta$ are two positive constants. Similarly, a system is “uniformly completely controllable’’ if there exists an integer $N$ and two positive constants $\gamma$ and $\delta$ such that

$$0 < \gamma I \leq G_N(t_n) \leq \delta I$$  \hfill (3.23)
for all $n \geq N$, where $\gamma$ and $\delta$ are two positive constants. So, when the system is uniformly completely observable and controllable, we have
\[
(\beta + \gamma^{-1})^{-1} I \leq P_+(t_n) \leq (\alpha^{-1} + \delta) I
\]
for all $n \geq N$, so that $P_+(t)$ is “well-behaved” asymptotically, and after averaging we obtain the bounds
\[
(\beta + \gamma^{-1})^{-1} I \leq P_{+\text{av}} \leq (\alpha^{-1} + \delta) I
\]
The challenge, of course, is to determine the constants $\alpha$, $\beta$, $\gamma$ and $\delta$.

3.3 Simulation Results: single-sensor estimation with irregular SSP

Our theoretical analysis in Sec. 3.2 provided an approximate implicit characterization of $P_{+\text{av}}$ for any SSP with a mild fluctuation. In particular, in the simple case of a single-state system, we were able to reduce the coupled set of nonlinear matrix equations (3.11) to a single scalar quadratic equation. In this simple case we were able to show that $P_{+\text{av}}$ depends mainly on $F_{\text{av}} \triangleq \frac{1}{\langle E[T_s(n)] \rangle_n}$ and to a much lesser extend on the variance index $\zeta = \frac{\langle \text{Var}[T_s(n)] \rangle_n}{\langle E[T_s(n)] \rangle_n^2}$. Here we will carry out simulations to demonstrate that the same conclusions hold qualitatively for any number of states and any SSP.

To be specific, we will consider here a two-state, single-input, single-output, continuous-discrete system (recall (3.1)) with
\[
A = \begin{bmatrix} 0 & 1 \\ -12 & -7 \end{bmatrix}, B = G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \end{bmatrix}, Q = R = 1
\]
(3.26)
Since the eigenvalues of $A$ are $\lambda_1 = -3$ and $\lambda_2 = -4$, this is clearly a stable system. We will evaluate the performance of the continuous-discrete Kalman filter with several types of SSP, including the periodically-variant (deterministic) and uniformly-distributed (random) SSP. We will use the Signal to Estimation Error Ratio (SEER), viz.,
\( \text{SEER}(dB) = 10 \log \frac{\langle \| x(t_n) \|^2 \rangle_n}{\langle \| x(t_n) - x_+(t_n) \|^2 \rangle_n} \)

as a measure of state-estimation quality. Notice that the denominator of the SEER is, in fact, equal to \( tr(P_{\text{av}}) \), the sum of the diagonal elements (i.e., estimation error variance) of \( P_{\text{av}} \).

We shall present our results as plots of SEER vs. the average sampling interval \( T_{\text{av}} = \langle E[T_s(n)] \rangle_n \). The analysis of Sec 3.2 tells us that \( tr(P_{\text{av}}) \) is a monotone decreasing function of \( F_{\text{av}} = \frac{1}{T_{\text{av}}} \). Since the SEER is inversely proportional to \( tr(P_{\text{av}}) \), we conclude that, in general, SEER is a monotone decreasing function of \( T_{\text{av}} \). We shall demonstrate that the SEER depends primarily on the average sampling interval \( T_{\text{av}} \), with less dependence on the variance of \( T_s(n) \) and even smaller dependence on the kurtosis. Our simulation results are consistent with the analytical ones although here we use two-state, single-input, single-output system.

### 3.3.1 Regular Sampling

First of all, we confirm the relationship between Average Sampling Interval and SEER in the case of regular sampling. As Fig. 3.6 shows, indeed the SEER is a monotone decreasing function of the average sampling interval. The results make sense, because when the average sampling interval is small, the number of observations used to for \( \hat{x}(t) \), for any given \( t \), increases. Measurement update in Kalman filter can help reduce the error covariance, so the more measurements, the smaller the SEER. Also the results from Fig. 3.6 are consistent with Fig. 3.4, in which the average error covariance \( P_{\text{av}} \) decreases as the average sampling rate increases.
Notice that when $T_{av}$ decreases to zero, the discrete time sensing becomes indistinguishable from continuous time sensing. As in the above plot when the average sample interval is zero, we achieve a very high SEER, around 45dB, which is the same level as a continuous Kalman filter. On the other hand, when $T_{av}$ increases to infinity we lose all measurements, so that the estimated state $\hat{x}(t)$ becomes zero, the estimation error $x(t) - \hat{x}(t) = x(t)$ coincides with the state itself, and consequently the estimation SEER becomes 0 dB.

In general the performance of state estimation depends also on the choice of sensor, as shown in Fig. 3.7. For instance, choosing a sensor with $C = [2 \ 6]$ results in a 2-3 dB improvement in SEER across the whole range of $T_{av}$ values.
Irregular Sampling

Our analysis in section 3.2.1 suggested that estimation performance is not very sensitive to the choice of sampling pattern. In order to confirm this prediction we change the sampling pattern in our simulation from regular to periodically-variant sampling. We keep the fluctuation index constant at $\xi = 0.32$ by setting $\alpha = \sqrt{2\xi} = 0.8$, and we vary $\Delta = T_{av}$ to achieve the desired range of values for $T_{av}$. The resulting plot of SEER vs. $T_{av}$ (Fig. 3.8) is practically indistinguishable from the regular sampling case (Fig. 3.6), which confirms our original assertion that the performance of Kalman filter estimation depends mainly on $T_{av}$, regardless of the choice of sampling pattern.
3.3.3 The Effect of Sampling Interval Variance Index

We now turn to study the effect of the variance index $\xi = \frac{\langle \text{Var}[T_s(n)] \rangle_n}{\langle \text{E}[T_s(n)] \rangle_n^2}$ on the “SEER vs. Average Sampling Interval” curve. We consider the periodically-variant SSP and uniformly distributed random SSP, both of which allow for independent control of the mean and the variance index of $T_s(n)$. In fact $\xi \sim \alpha^2$ in both cases (see Table 3.1), so that specific values of the variance index $\xi$ are achieved by selecting appropriate values for $\alpha$.

We select $\xi = 0, 0.02, 0.125, 0.32$ (recall that $\xi = 0$ represents a regular SSP) and plot the corresponding SEER vs. $T_{av}$ curve in Fig. 3.9 (periodically-variant SSP) and Fig. 3.10 (uniformly distributed random SSP). In both cases, the effect of $\xi$ is minor, and can only be observed in the zoomed-in version of the plot. In both cases, increasing the variance index $\xi$ results in a reduction of the SEER. Furthermore the same values of $\xi$ produces essentially the same curve (Fig. 3.11). In summary,
higher-order moments of the distribution of the sampling interval $T_i(n)$ have a truly small effect on the SEER, as compared with the effect of the mean of $T_i(n)$ and that this small effect is almost entirely due to the variance of $T_i(n)$.

Fig. 3.9  Effect of relative variance index $\zeta$ on SEER (top) and zoom-in (bottom): periodically-variant SSP
Fig. 3.10  Effect of relative variance index $\xi$ on SEER (top) and zoom-in (bottom): uniformly-distributed random SSP
Fig. 3.11 Effect of relative variance index $\zeta$ on SEER (top) and zoom-in (bottom):
comparing both SSP choices

Also the results we get from Fig. 3.9 and Fig. 3.10 are consistent with Fig. 3.4,
which also indicated that the averaged error covariance increases with $\zeta$.

Thus we can draw the conclusion that in the Kalman filter state estimation, the
average sampling rate is the most significant factor in terms of estimation
performance. The variance index $\xi$ of the sampling interval has a much smaller impact on the SEER comparing to the average sampling rate. So the best performance of state estimation is always achieved with a regular sampling pattern.

3.3.4 The Effect of Kurtosis of The Sampling Interval

In view of our results so far we expect higher-order moments of the sampling interval distribution to play an even smaller role than the mean and variance. To validate this conjecture we now consider the kurtosis

$$K_T \triangleq \frac{\mu_4}{\sigma^4} - 3 = \frac{\left\{E(T_s(n) - T_{av})^4\right\}_n}{(\text{Var}(T_s(n)))^2} - 3 \quad (3.27)$$

which is a measure of “peakedness” of the probability distribution of $T_s(n)$. A Gaussian distribution has zero kurtosis. Positive kurtosis indicates longer tails and a flatter central peak, while negative kurtosis indicates shorter tails and a sharper central peak. Our objective is to study the effect of varying kurtosis on estimation performance while the mean and variance of the sampling interval $T_s(n)$ are kept constant.

In order to achieve our objective we need an SSP with three degrees of freedom that will allow us to control mean, variance and kurtosis independently. Since the distributions we considered so far have only two degrees of freedom, we introduce a mixed distribution by using a convex combination of two periodically-variant SSPs, viz.,

$$T_s(n) = \lambda \Delta (1 + \alpha \sin(\omega_1 n)) + (1 - \lambda) \Delta (1 + \alpha \sin(\omega_2 n)) \quad 0 \leq \lambda \leq 1$$

So we have the mean $T_{av} = \Delta$, the variance index $\xi = \frac{\alpha^2}{2} \left[ \lambda^2 + (1 - \lambda)^2 \right]$ and the kurtosis $K_T = \frac{3}{2} \frac{\left[ \lambda^2 + (1 - \lambda)^2 \right]^2 + 2 \lambda^2 (1 - \lambda)^2}{\left[ \lambda^2 + (1 - \lambda)^2 \right]^2} - 3$ (see App. C for derivation).

We can now control the value of $K_T$ by varying $\lambda$ (Fig. 3.12). This mixed distribution achieves kurtosis in the range [-1.5, -0.75]. In comparison, a single periodic distribution has a fixed kurtosis of -1.5.
Notice that the variance index $\xi$ depends on both $\alpha$ and $\lambda$, while the kurtosis only depends on $\lambda$. This allows us to vary the kurtosis by changing the parameter $\lambda$, and adjust $\alpha$ accordingly to maintain a fixed variance. In Table 3.3, we fix the variance index $\xi = \frac{\text{Var}(T_c)}{\text{Mean}(T_c)^2} = 0.02$, but change $\lambda$ from 0 to 0.5. Since we have $\xi = \frac{\alpha^2}{2} [\lambda^2 + (1 - \lambda)^2]$, this requires that we adjust $\alpha$ for every choice of $\lambda$ according to the relation $\alpha = \sqrt{\frac{2\xi}{\lambda^2 + (1 - \lambda)^2}}$.

Table 3.3  Kurtosis with changing $\lambda$ and fixed variance index $\xi = 0.02$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.2</td>
<td>0.221</td>
<td>0.243</td>
<td>0.263</td>
<td>0.277</td>
<td>0.283</td>
</tr>
<tr>
<td>$K_{T_c}$</td>
<td>-1.5</td>
<td>-1.46</td>
<td>-1.33</td>
<td>-1.1</td>
<td>-0.86</td>
<td>-0.75</td>
</tr>
</tbody>
</table>

From Fig. 3.13, we can see clearly that the kurtosis still has some effect on the quality of estimation, but much less compared to effect of the mean and variance. So we can draw the conclusion that the performance of state estimation using Kalman
filter depends primarily on the average sampling interval $T_{av}$, with less dependence on the relative variance index $\zeta$, and even smaller dependence on the kurtosis of the SSP.

Fig. 3.13 Effect of sampling interval kurtosis on SEER (top) and zoom-in (bottom): convex combination of two periodically-variant SSP
3.4 Simulation Results: two-sensor estimation with irregular SSP

Our previews results in section 3.3 show that, in single-sensor estimation, the average error covariance is a monotone decreasing function of average sampling rate $F_{av}$, and fluctuations in $T_i(n)$ have a small effect, with the best results obtained for regular sampling. In this section, two-sensor state estimation is analyzed as a special example of multi-sensor estimation. The principle we find in single-sensor case can also be applied to multiple sensor case, so we expect to see that the average error covariance is a monotone decreasing function of each average sampling rate $F_{av,i}$ ($i=1,2$), and the best results are achieved when the superposed sequence of sampling times is as close to regular as possible.

In the multiple-sensor case the choice of SSPs involves numerous degrees of freedom. In addition to the average sampling interval for each sensor, we need to consider higher-order moments, as well as the relative alignment of the SSPs. Since we have already established that higher-order moments have an almost negligible effect on estimation performance, we focus here on the case of regular sampling with arbitrary alignment, so that for two-sensor we have

$$t_i^{(1)} = t_0^{(1)} + iT_1, \quad t_i^{(2)} = t_0^{(2)} + iT_2$$

(3.28)

where $T_1$ and $T_2$ are the regular sampling intervals of sensor 1 and 2. We define

$\{t_n\}$ to be the superposition of the individual sampling sequences $t_i^{(k)}$ (recall (1.3)).

Estimation performance is quantified in terms of $\text{trace}(P_{av})$, where

$$P_{av} = \left\langle P_{av}(t_n) \right\rangle_n = \left\langle E \left\{ \left[ x(t_n) - x_v(t_n) \right] \left[ x(t_n) - x_v(t_n) \right]^T \right\} \right\rangle_n.$$

We can now plot $\text{trace}(P_{av})$ as a function of $F_{av,1}$ and $F_{av,2}$, the average sampling rates of sensor 1 and 2, respectively. Notice that since we restrict our attention to regular sampling for each sensor, we have $F_{av,1} = F_1 = \frac{1}{T_1}$ and $F_{av,2} = F_2 = \frac{1}{T_2}$.

Here we use the same example as in Sec. 3.3, namely

$$A = \begin{bmatrix} 0 & 1 \\ -12 & -7 \end{bmatrix}, \quad B = G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad Q = R = 1$$
but now we have two sensors with $C_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $C_2 = \begin{bmatrix} 2 & 6 \end{bmatrix}$.

The results are shown in Fig. 3.14, which confirms our basic observation from Sec. 3.3: the average error covariance is a monotone decreasing function of the average sampling rates $F_1$ and $F_2$. The same conclusion is also evident from the corresponding contour plot (Fig. 3.15). Since the average error covariance drops rapidly with increasing $F_1$ and $F_2$, we focus in the contour plot on small values of $\text{trace}(P_{\tau,\text{av}})$, namely $0.0001 \leq \text{trace}(P_{\tau,\text{av}}) \leq 0.0007$ for the top plot of Fig. 3.15, and $0.0007 \leq \text{trace}(P_{\tau,\text{av}}) \leq 0.0035$ for the bottom plot of Fig. 3.15.

The lines in the contour plot are almost straight with slope $\approx -1.1$, so that $\text{trace}(P_{\tau,\text{av}})$ is approximately a monotone decreasing function of the linear combination $F_1 + 1.1F_2$. This effect is a result of the difference between the two sensors: measurements from sensor 2 have a bigger impact on the value of $\text{trace}(P_{\tau,\text{av}})$ than measurements from sensor 1. We elaborate further on this difference in Sec. 3.4.2.

![Fig. 3.14](image_url)  

**Fig. 3.14** Estimation performance ($\text{trace}(P_{\tau,\text{av}})$) as a function of $F_1$ and $F_2$.
The alignment of the two SSPs can be described in terms of its effect on the regularity of the superposed sequence \( \{r_n\} \). We show in Sec. 3.4.1 that the best estimation performance is achieved when we adjust the relative alignment so as to reduce the variance of the superposed sequence \( \{r_n\} \).
3.4.1 Effect of SSP Irregularity

We now recall that in general irregularity in the sampling pattern has a minor effect on estimation performance. In the single-sensor case we use the relative variance index

$$\xi = \frac{\langle \text{Var}[T_s(n)] \rangle_n}{\left| \langle E[T_s(n)] \rangle_n \right|^2}$$

as a measure of irregularity. In the two-sensor case we propose to apply the same index to the superposed sequence \(\{t_n\}\), so that

$$\tilde{\xi} = \frac{\langle \text{Var}[t_n - t_{n-1}] \rangle_n}{\left| \langle E[t_n - t_{n-1}] \rangle_n \right|^2}.$$ 

Notice that even though each sensor is sampled regularly, the superposed sequence \(\{t_n\}\) is irregular. In our two-sensor example the relative variance index \(\tilde{\xi}\) depends on \(F_1, F_2\) and the relative initial shift \(t^{(2)}_0 - t^{(1)}_0\).

When \(F_1 = F_2\) (so that \(T_1 = T_2 = T_{av}\) in (3.28)), the index \(\tilde{\xi}\) depends only on the relative shift \(\delta \equiv \frac{t^{(2)}_0 - t^{(1)}_0}{T_{av}}\). The plot of \(\text{trace}(P_{av})\) as a function of \(\delta\) (Fig. 3.16) confirms our intuition: estimation performance improves as \(\delta\) approaches 0.5, which corresponds to \(\tilde{\xi} = 0\). On the other extreme, the highest value of \(\text{trace}(P_{av})\) is achieved when \(\delta = 0\), which corresponding to \(\tilde{\xi} = \xi_{max} = 1\), the highest possible value of \(\tilde{\xi}\) for this example. The range of \(\delta\) here is \(0 \leq \delta \leq 0.5\).

Notice that the value of \(\text{trace}(P_{av})\) varies by a factor of 1.86 over the range \(0 \leq \delta \leq 0.5\).
Fig. 3.16 Effect of relative shift on estimation performance: $F_1 = F_2 = 10$

Similar conclusions hold for the case $F_1 \neq F_2$, except that the range of possible $\xi$ values is usually narrower, and the variation of $\text{trace}(P_{r,m})$ is, consequently, smaller. For instance, when $F_1 = 4$ samples/s and $F_1 = 16$ samples/s, we find that $\text{trace}(P_{r,m})$ only varies by a factor of 1.33 (Fig. 3.17). Here we define the relative shift as $\delta \triangleq \frac{t_0^{(2)} - t_0^{(1)}}{\min(T_1, T_2)}$, and observe that the relative variance has a much smaller range, namely $0.09 \leq \xi \leq 0.25$. 
Fig. 3.17 Effect of relative shift on estimation performance: \( F_1 \neq F_2 \)

Fig. 3.16 and 3.17 confirm the results obtained in Sec. 3.3 and extend to two sensor estimation. \( \text{trace}(P_{\tau,\text{av}}) \) increases when the relative variance of the superposed sequence \( \{t_n\} \) increases and the best results are achieved when \( \xi = \xi_{\text{min}} \).

### 3.4.2 Effect of Sensor Differences

We have observed earlier that \( \text{trace}(P_{\tau,\text{av}}) \) is, approximately, a monotone decreasing function of the linear combination \( F_1 + \alpha F_2 \), where the coefficient \( \alpha \) captures the difference between the effects of each sensor on estimation performance. In our example \( \alpha \approx 1.1 \), so that sensor 2 is more effective in reducing \( \text{trace}(P_{\tau,\text{av}}) \). This is directly evident, for instance in Fig. 3.18, in which using sensor 2 alone results in a lower \( \text{trace}(P_{\tau,\text{av}}) \) than using sensor 1 alone. Notice that in this
plot we compare performance at the same sampling rate, i.e., $F_1 = F_2 = F_{av}$. When both sensors are used simultaneously, but with $F_1 = F_2 = F_{av}/2$, the resulting plot of $\text{trace}(P_{n,av})$ vs. $F_{av}$ lies between the plots associate with using only sensor 1 or only sensor 2 (Fig. 3.18).

Fig. 3.18 The effect of sensor differences on estimation performance
Another way to demonstrate the same effect is by plotting of $\text{trace}(P_{r,av})$ as a function of the ratio $\frac{F_1}{F_1 + F_2}$ for some predetermined choices of $F_1 + F_2$ (Fig. 3.19). As $F_1$ increases, with $F_1 + F_2$ kept constant, the quality of estimation gradually degrades ($\text{trace}(P_{r,av})$ increases), reflecting the fact that sensor 2 is more effective in reducing $\text{trace}(P_{r,av})$. The various traces in Fig. 3.19 (top) correspond to the 12 distinct $F_{av} = F_1 + F_2$ values shown in Table 3.4, ranging from $F_{av} = 1$ sample/s (top trace) to $F_{av} = 1000$ samples/s (lowest trace). The bottom plot in Fig. 3.19 provides a zoomed-in version of the 8 lowest traces.

Table 3.4  Combination of sampling rates in sensor 1 and 2

<table>
<thead>
<tr>
<th>Sensor</th>
<th>$F_{av} = F_1 + F_2 = {1,2,5,10,20,25,50,100,200,250,500,1000}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1$</td>
<td>0</td>
</tr>
<tr>
<td>$F_2$</td>
<td>$F_{av}$</td>
</tr>
</tbody>
</table>
3.5 **Summary of Results**

In this chapter, we discussed the effect of different SSPs on state estimation performance using a continuous-time discrete-measurement Kalman filter.

We were able to characterize $P_{\text{err}}$ by a pair of coupled matrix equations that
hold exactly for regular sampling and approximately for mildly irregular sampling. These equations imply that $P_{\pm,av}$ depends on the SSP via two matrices, $\Omega_{T_s}$ and $\mathcal{U}_{T_s}$, which are determined by the characteristic function $\phi_{T_s}(\bullet)$, and by the eigenvalues of the (continuous-time) system matrix $A$.

In the special case of a single-state system these equation provide closed-from expression for $P_{\pm,av}$. In particular, we were able to show that $P_{\pm,av}$ is mainly a function of $F_{av} = \frac{1}{T_{av}}$, with a very mild dependence on higher-order moments of the sampling interval $T_s(n)$.

The general case, when the state $x(t)$ is a vector and $P_{\pm}(t_n)$ are matrices, the coupled set of matrix equation (3.11) is difficult to analyze for arbitrary SSPs. Instead we carried out a numerical study, using the periodically-variant SSP and uniformly distributed random SSP as benchmarks for evaluating the estimation performance of the continuous discrete Kalman filter. In the single-sensor case, we used an example to demonstrate that the performance of state estimation (measured by SEER) depends primarily on the average sampling interval $T_{av}$, with less dependence on the relative variance index $\xi$ and even smaller dependence on the kurtosis of the SSP. So the best performance is achieved with a regular SSP.

Our findings hold also, qualitatively, in the multi-sensor case. In particular, we carried out a numerical study of the two-sensor case, demonstrate that $\text{trace}(P_{\pm,av})$ is monotone decreasing in both $F_1$ and $F_2$, as expected. In fact, it appears that $\text{trace}(P_{\pm,av})$ is, approximately, a monotone decreasing function of a linear combination of $F_1$ and $F_2$—in our example, this combination is $F_1 + 1.1F_2$ because of the different impact of the two sensors on performance.

Regularity of the overall (superposed) sampling pattern also plays a role in controlling the level of estimation error. For a single sensor, the best results (for a given average sampling rate) are achieved with a regular SSP. Similarly, in the two-sensor case, the level of state estimation error (for a given $F_{1,av}$ and $F_{2,av}$)
increases with the relative variance index $\zeta$ of superposed SSP. This index depends on the relative alignment of the two SSPs, and it achieves its minimal value when the superposed SSP is as close to regular as possible. In particular, when both sensors use the same rate, the worst case (highest error) occurs when the two SSPs are perfectly aligned. Performance improves as the relative shift of between the two SSPs increases toward to 0.5.
Chapter 4

Concluding Remarks

4.1 Summary of Results

In this thesis we have analyzed two problems associated with a spatially distributed cyber physical system: the effect of delay in the control path, and the effect of asynchronous irregular sensor sampling patterns on the quality of MSE estimation. In both cases, we assumed the presence of an imperfect communication network between a central estimator (a continuous discrete Kalman filter) and the spatially dispersed sensors and actuators. While the modern smart grid relies on perfectly synchronized sampling, our research considered the more challenging case in which sensor sampling patterns need not be coordinated, and samples may get lost or corrupted in the process of communication, thereby contributing to the irregularity of the sampling process.

In the problem of delay in the control path, we formulated a Kalman-filter-based delay mitigation method with time-stamping technology, which is enabled by the availability of precise time signals derived from the GPS. Our delay canceling method transmutes communication delay into increased estimation error, so that the average SEER of any signal in the estimation-control loop is a monotone decreasing function of the average communication delay. This method not only successfully mitigates the destabilizing effects of delay, but also improves the estimation performance when the delay is small. We have used a simple example to show that the time-stamping-based delay canceling method makes it possible to overcome significant delays in the estimation-control loop while avoiding adverse effects on the stability of the closed-loop system. In fact, the value of delay in the control loop can be increased by a factor of 5-10 beyond the critical value $\Delta_c$ with acceptable SEER (10-20 dB).

In our analysis of the effect of irregular SSPs on estimation performance, we found that, in the single-sensor case, $P_{r,av}$ is a monotone decreasing function of the
average sampling rate \( F_{\text{av}} \), which means that \( P_{+\text{av}} \) depends mostly on the average sampling rate \( F_{\text{av}} \). In addition, the fluctuations in the sampling interval \( T_s(n) = t_n - t_{n-1} \) result in a mild increase in estimation error, so that the more fluctuation, the worse the performance. The best results are obtained when \( \langle \text{var} T_s(n) \rangle_n = 0 \), which corresponds to regular sampling. In general, the average estimation error covariance \( P_{+\text{av}} \) depends mainly on the average sampling rate \( F_{\text{av}} = \langle E[T_s(n)] \rangle_n^{-1} \), with a minor dependence on \( \langle \text{Var}[T_s(n)] \rangle_n \) and a negligible dependence on the higher order moments of \( T_s(n) \), such as the kurtosis. These conclusions were confirmed analytically in the special case of single-sensor, single-state with almost regular sampling (see Fig. 3.4). The same conclusions hold also for \( P_{-\text{av}} \), the long-term average of \( P_{-}(t_n) \).

Our findings hold also, qualitatively, in the multi-sensor case. In particular, we carried out a numerical study of the two-sensor case. Thus \( \text{trace}(P_{+\text{av}}) \) is a monotone decreasing function of both average sampling rates, \( F_1 \) and \( F_2 \). Especially, in our example, \( \text{trace}(P_{+\text{av}}) \) is, approximately, a monotone decreasing function of the linear combination \( F_1 + 1.1F_2 \) because of the different impact of the two sensors on performance. Also \( \text{trace}(P_{+\text{av}}) \) increases when the variance of the superposed sequence \( \{t_n\} \) increases. Notice that with two SSPs, the fluctuation of the inter-sample interval in the superposed sequence \( \{t_n\} \) depends both on the individual SSPs and on the relative alignment. In particular, when both SSPs are regular with the same sampling rate \( F_1 = F_2 \), the best results are achieved when one SSP is shifted by half a sampling interval with respect to the other SSP, so that the superposed sequence \( \{t_n\} \) is regular.
4.2 Further Research

A number of research directions are suggested by the problems discussed in this thesis, such as:

- **Statistics of** $\mathbb{P}_\pm(t_n)$: we have explored the characteristics of $\mathbb{P}_{\pm,av}$, the long term time (and ensemble) average of $\mathbb{P}_\pm(t_n)$, respectively. Higher-order moments of $\mathbb{P}_\pm(t_n)$ need to be studied as well. In particular, we expect a simple (i.e., monotone increasing, nearly linear) relation between the average variance of $\mathbb{P}_\pm(t_n)$ (i.e. $\langle \text{Var } \mathbb{P}_\pm(t_n) \rangle_n$), and the average variance of $T_s(n)$ (i.e. $\langle \text{Var } T_s(n) \rangle_n$). More generally, we should determine the steady state pdf of $\mathbb{P}_\pm(t_n)$ in terms of the characteristic function of $T_s(n)$.

- **Multi-sensor SSP**: our preliminary results indicate that estimation performance depends on the statistics of the superposed sampling sequence $\{t_n\}$. This raises the question to what extent the statistics of $\mathbb{P}_\pm(t_n)$ can be determined from the statistics of the superposed SSP? In general, one needs to consider the joint statistics of the multi-sensor SSP sequence $\{t_{i(k)}; 1 \leq k \leq K, 0 \leq i < \infty\}$, so that working with a single (superposed) sequence has the potential to offer significant reduction in the computational and conceptual complexity. In particular, our results suggest that the average estimation error in a multi-sensor setting depends mainly on the individual average sampling rate and the relative variance index $\xi$ of the superposed SSP sequence.

- **Limited channel capacity**: we have assumed adequate channel capacity for high-accuracy (i.e., “infinite precision”) transmission of sensor measurements. When channel capacity is constrained we must rely on source coding in order to reduce the sensor data rate below the channel capacity limit. The simplest approach is to apply analog-to-digital conversion (i.e., quantization) to each sensor measurement. This results in increased estimation error, so that the number of bits per sample and the average sampling rate must be optimized jointly. A more advanced approach is to use information-theoretic multi-sensor coding schemes, such as the ones proposed in [18].
- **Optimal networked control**: in Ch. 2 we relied on the assumption that the classical optimal control law $u_c(t) = -K\hat{x}(t)$ is still optimal, or nearly optimal, in a networked configuration. This assumption needs closer scrutiny in the presence of arbitrary network delays and (possibly) irregular SSPs. In fact, it is very likely that even the celebrated separation principle does not hold in a networked setting.

- **Hierarchical network configuration**: we have carried out our analysis for the case when a single estimation/control center communicates with a number of spatially distributed sensors and actuators. A more realistic configuration requires at least two levels of estimation/control:
  (a) A local estimation/control hub (e.g. a substation in a power network) that communicates with sensors and actuators.
  (b) A top-level estimation/control center that communicates with all the local hubs, and provides a higher-level situation awareness.

Our centralized methodology will need to be redesigned for such a hierarchical setting.
Appendix A

Averaged Time Update

The time-update relation (2.3) between $P_{\tau}(t_n)$ and its delay-impacted counterpart $P_{\tau}^{(d)}(t_n + \Delta_n)$ induces a similar relation between the time-averaged version of these error covariances.

This relation is straightforward when the system matrix $A$ is a scalar (i.e. a single-state system). In this case, (2.3) reduces to

$$
P_{\tau}^{(d)}(t_n + \Delta_n) = e^{(A+\Delta)\Delta_n}P_{\tau}(t_n) + GQG^* \int_0^\Lambda e^{(A+\Delta)^*\tau} d\tau$$

$$= e^{2\text{Re}(A)\Delta_n}P_{\tau}(t_n) + \frac{GQG^*}{2\text{Re}(A)}(e^{2\text{Re}(A)\Delta_n} - 1)$$

and averaging (both in time and probabilistic ensemble) produces

$$P_{\tau,av}^{(d)} = \phi_A(2\text{Re}(A))P_{\tau,av} + \frac{GQG^*}{2\text{Re}(A)}[\phi_A(2\text{Re}(A)) - 1]$$

where $\phi_A(s)$ is the averaged characteristic function associated with the distribution of the delay sequence $\{\Delta_n\}$, namely

$$\phi_A(s) = \left\langle E[e^{i\Lambda_n}] \right\rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^N \left[ e^{i\Lambda_n} \cdot p(\Delta_n) \cdot d\Delta_n \right].$$

Also we assume the independence of $t_n$ and $\Delta_n$. In the general case when $A$ is a matrix, we apply the diagonal scaling lemma of App. B to a diagonalized version of the matrix $A$ to obtain a similar relation between $P_{\tau,av}^{(d)}$ and $P_{\tau,av}$.

Thus we assume that $A$ is diagonalizable, say $A = T A T^{-1}$, where $T$ is a nonsingular matrix and $\Lambda = \text{diag}\{\lambda_i\}$ is a diagonal matrix constructed from the eigenvalues of $A$. Consequently, for any matrix $G$, and for any scalar $\tau$, we have
\[ e^{A^\tau}Ge^{A^\tau} = T e^{A^\tau} T^{-1} GT^* e^{A^\tau} T^* = T \{ e^{A^\tau} \tilde{G} e^{A^\tau} \} T^* \] (A.1)

where \( \tilde{G} = T^{-1} GT^* \) and the asterisk \( (\cdot)^* \) denotes Hermitian transpose.

Using the diagonal scaling lemma of App. B, we can also rewrite \( e^{A^\tau} \tilde{G} e^{A^\tau} \) as

\[ e^{A^\tau} \tilde{G} e^{A^\tau} = \{ \text{vecd}(e^{A^\tau}) | \text{vecd}(e^{A^\tau}) \} \circ \tilde{G} \] (A.2)

where \( \circ \) denotes a Schur-Hadamard product, and

\[ \text{vecd}(e^{A^\tau}) = \begin{bmatrix} e^{\lambda_1^\tau} \\ e^{\lambda_2^\tau} \\ \vdots \\ e^{\lambda_n^\tau} \end{bmatrix} \]

is a column vector constructed from the diagonal elements of \( e^{A^\tau} \). Notice that

\[ \text{vecd}(e^{A^\tau})^T = [\text{vecd}(e^{A^\tau})]^* \], so that

\[ \text{vecd}(e^{A^\tau}) [\text{vecd}(e^{A^\tau})]^T = \text{vecd}(e^{A^\tau}) [\text{vecd}(e^{A^\tau})]^* = [e^{(\lambda_1^\tau + \lambda_m^\tau)^T}]_{m=1}^M \] (A.3)

Now, if \( \tau \) is random or time-variant, then

\[
\left\langle E \{ e^{A^\tau} Ge^{A^\tau} \} \right\rangle = T \left\langle E \{ e^{A^\tau} \tilde{G} e^{A^\tau} \} \right\rangle T^*
= T \left\langle E \left\{ \text{vecd}(e^{A^\tau}) | \text{vecd}(e^{A^\tau}) \} \circ \tilde{G} \right\} \right\rangle T^*
= T \left\langle E \left[ e^{(\lambda_1^\tau + \lambda_m^\tau)^T} \right]_{m=1}^M \circ \tilde{G} \right\} \right\rangle T^*
= T \left[ \phi_\tau(\lambda_1^\prime + \lambda_m^\prime) \right]_{m=1}^M \circ \tilde{G} \right\} \right\rangle T^* \] (A.4)

where we used the averaged characteristic function \( \phi_\tau(s) \) associated with the random variable \( \tau \), viz., \( \phi_\tau(s) \triangleq \left\langle E \{ e^{s \tau} \} \right\rangle \).

In summary, when \( A \) is diagonalizable, we have

\[
\left\langle E \{ e^{A^\tau} Ge^{A^\tau} \} \right\rangle = T \left[ \Omega_\tau \circ \left( T^{-1} GT^* \right) \right] T^* \] (A.5a)

where \( \Omega_\tau = \left[ \phi_\tau(\lambda_1^\prime + \lambda_m^\prime) \right]_{m=1}^M \) (A.5b)

Similarly, applying the diagonal scaling lemma to the integral

\[ \int_0^\tau e^{A^\tau} G Q G^* e^{A^\tau} ds \], we obtain
Applying time and probabilistic ensemble averaging, we obtain

\[
\left\langle E \left\{ \int_0^r e^{A_s T} GQG^* e^{A_s} \, ds \right\} \right\rangle = T \left[ \mathcal{U}_r \circ \left( T^{-1} GQG^* T^{-r} \right) \left| T^r \right| \right]
\]  

(A.6a)

where \( \mathcal{U}_r = \left[ \frac{1}{\lambda_j + \lambda_m} \left( e^{(\lambda_j + \lambda_m) \tau} - 1 \right) \right]_{j,m=1}^M \)  

(A.6b)

We can now apply (A.5) and (A.6) to (2.3), resulting in an expression for \( P_{+ \text{av}} \), viz.,

\[
P_{+ \text{av}}^{(d)} = T \left( \Omega_\Delta \circ T^{-1} P_{+ \text{av}} T^{-r} \right) T^r + T \left[ \mathcal{U}_\Delta \circ \left( T^{-1} GQG^* T^{-r} \right) \right] T^r
\]

\[
\Rightarrow T^{-1} P_{+ \text{av}}^{(d)} T^{-r} = \Omega_\Delta \circ \left( T^{-1} P_{+ \text{av}} T^{-r} \right) + \mathcal{U}_\Delta \circ \left( T^{-1} GQG^* T^{-r} \right)
\]

where \( \Omega_\Delta \) and \( \mathcal{U}_\Delta \) are determined from \( \phi_\Delta(s) \), the averaged characteristic function for the delay \( \Delta_n = \theta_n - t_n \). This establishes (2.4). Similarly, applying (A.5) and (A.6) to (3.2) results in an expression for \( P_{- \text{av}} \), viz.,
\[ T^{-1}P_{av} T^{-*} = \Omega_{T_i} \circ \left( T^{-1}P_{av} T^{-*} \right) + \Upsilon_{T_i} \circ \left( T^{-1}GQT T^{-*} \right) \]

where \( \Omega_{T_i} \) and \( \Upsilon_{T_i} \) are determined from \( \phi_{b_i}(s) \), the averaged characteristic function for the sampling interval \( T_i(n) = t_n - t_{n-1} \). This establishes (3.3).
Appendix B

Diagonal Scaling Lemma

Let $G$ be any $M \times N$ matrix and $D_1$, $D_2$ any diagonal matrices of size $M \times M$ and $N \times N$, respectively. Then we have

$$D_1 G D_2 = \{ \text{vecd}(D_1) \ [ \text{vecd}(D_2) ]^T \} \circ G$$

(B.1)

where $\text{vecd}(D_i)$ is a column vector consisting of the diagonal elements of the matrix $D_i$ (same as $\text{diag}(D_i)$ in Matlab), and $\circ$ denotes the Schur-Hadamard product (same as .* in Matlab).

The main utility of the result is in the case when $D_1$ and $D_2$ are random or time-variant, while $G$ is fixed. When $D_1$ and $D_2$ are real valued, then we have

$$\langle E[D_1GD_2^*] \rangle = R_{D_1,D_2} \circ G$$

(B.2)

where $R_{D_1,D_2}$ is the averaged cross covariance of $\text{vecd}(D_1)$ and $\text{vecd}(D_2)$, namely

$$R_{D_1,D_2} = \langle E[\text{vecd}(D_1)[\text{vecd}(D_2)]^T] \rangle$$

When $D_1$ and $D_2$ are complex valued, we have

$$E[D_1GD_2^*] = R_{D_1,D_2} \circ G, \quad R_{D_1,D_2} = \langle E[\text{vecd}(D_1)[\text{vecd}(D_2)]^*] \rangle$$

(B.3)
Appendix C

The Characteristic Function and Statistics

In probability theory and statistics, the characteristic function of any random variable completely defines its probability distribution. Similarly to the probability density function, the characteristic function provides an alternative way for describing the statistics of a random variable. Since our interest is mainly in time-averaged statistics of (possibly) non-stationary random sampling intervals, we refine the corresponding characteristic function as \( \phi_{T}(s) \triangleq \left\langle E\left\{ e^{iT(n)} \right\} \right\rangle_{n} \), where \( \left\langle \cdot \right\rangle_{n} \) denotes long-term time averaging with respect to the index "n", as defined in (1.8). Various averaged moments of the sampling interval \( \{T_{s}(n)\} \) can be determined by evaluating higher-order derivatives of \( \phi_{T}(s) \) at \( s=0 \). For instance, the first derivative of characteristic function, \( \frac{d\phi_{T}(s)}{ds}\bigg|_{s=0} \triangleq \left\langle E\left\{ T_{s}(n) \right\} \right\rangle_{n} \), equals the averaged mean of \( T_{s}(n) \). Also the second derivative, \( \frac{d^{2}\phi_{T}(s)}{ds^{2}}\bigg|_{s=0} \triangleq \left\langle E\left\{ T_{s}^{2}(n) \right\} \right\rangle_{n} \), equals the averaged second moment of \( T_{s}(n) \). We now turn to evaluate the characteristic function and some of the moments of several useful SSPs.

Periodically-variant Sampling

\( T_{s}(n) = \Delta(1+\alpha \sin(\omega n)) \), So the characteristic function is

\[
\phi_{T}(s) = \left\langle E\left\{ e^{i\Delta T(n)} \right\} \right\rangle_{n} = \left\langle E\left\{ e^{i(\Delta(1+\alpha \sin(\omega n)))} \right\} \right\rangle_{n} = e^{i\Delta} \left\langle E\left\{ e^{i\Delta \alpha \sin(\omega n)} \right\} \right\rangle_{n}
\]

Using the shorthand \( s\Delta = X \) and \( \omega n = \Theta \), we have

\[
e^{i\Delta \alpha \sin(\omega n)} = e^{X \sin \Theta} = \cosh(X \sin \Theta) + \sinh(X \sin \Theta) = \cos(jX \sin \Theta) - \sin(jX \sin \Theta) = [J_{0}(jX) + 2J_{2}(jX) \cos 2\Theta + 2J_{4}(jX) \cos 4\Theta \cdots] - j[2J_{1}(jX) \sin \Theta + 2J_{3}(jX) \sin 3\Theta + \cdots]
\]

where we used a Bessel series representation of trigonometric functions [19], viz.,
\[
cos(z \sin \theta) = J_0(z) + 2 \sum_{k=1}^{\infty} J_{2k}(z) \cos(2k\theta)
\]
\[
sin(z \sin \theta) = 2 \sum_{k=0}^{\infty} J_{2k+1}(z) \sin((2k+1)\theta)
\]

Now, averaging with respect to \( n \) eliminate all \( \theta \)-dependent terms of these series, so that
\[
\left\{ E\left\{ e^{i\Delta \alpha \sin(\omega n)} \right\} \right\}_n = J_0(jX) = J_0(js\Delta \alpha) = I_0(s\Delta \alpha)
\]

where \( I_0(s\Delta \alpha) \) is the zero-order modified Bessel function of order zero, and finally
\[
\phi_{T_s}(s) = e^{i\Delta} I_0(s\Delta \alpha).
\]

Mean: \( \text{Mean}(T_s(n)) = \frac{d\phi_{T_s}(s)}{ds} \bigg|_{s=0} = \Delta \)

Variance: \( \text{Var}(T_s(n)) = \frac{d^2\phi_{T_s}(s)}{ds^2} \bigg|_{s=0} - \text{Mean}(T_s(n))^2 = \frac{(\Delta \alpha)^2}{2} \)

Kurtosis: \( K_{T_s} = \frac{\left\{ (T_s(n) - \text{Mean}(T_s(n)))^4 \right\}_n}{\text{Var}(T_s(n))^2} - 3 \)

Since \( \left\{ \sin^4(\omega n) \right\}_n = \frac{3}{8} \), and \( \text{Var}(T_s(n)) = \frac{(\Delta \alpha)^2}{2} \)

So \( K_{T_s} = -\frac{3}{2} \)

**Uniformly Distributed Random Sampling**

\( T_s(n) = \Delta(1 + \alpha X_n) \), where \( X_n \) are i.i.d., uniformly distributed in the interval \([-1,1]\).

The characteristic function is
\[
\phi_{T_s}(s) = \left\{ E\left\{ e^{iT_s(n)} \right\} \right\}_n = \left\{ E\left\{ e^{i\Delta(1+\alpha X_n)} \right\} \right\}_n = e^{i\Delta} \left\{ E\left\{ e^{i\alpha X} \right\} \right\}_n
\]
\[
= \frac{1}{2} e^{i\Delta} \int_{-1}^{1} e^{i\alpha x} dx = \frac{1}{2} e^{i\Delta} \frac{e^{i\alpha} - e^{-i\alpha}}{i\alpha} = e^{i\Delta} \frac{\sinh(s\alpha\Delta)}{s\alpha\Delta}
\]
so that
\[
\phi_{T_s}(s) = e^{i\Delta} \frac{\sinh(s\alpha\Delta)}{s\alpha\Delta}
\]
Mean:  \[ \text{Mean}(T_s(n)) = \frac{d \phi_{T_s}(s)}{ds} \bigg|_{s=0} = \Delta \]

Variance:  \[ \text{Var}(T_s(n)) = \frac{d^2 \phi_{T_s}(s)}{ds^2} \bigg|_{s=0} - \text{Mean}(T_s(n))^2 = \frac{(\Delta \alpha)^2}{3} \]

Kurtosis:  \[ K_{T_s} = \frac{\left< E \left[ \alpha^4 \Delta X_n^4 \right] \right>}{(\text{Var}(T_s(n)))^2} - 3 = -\frac{6}{5} \]

where we used the fact that  \[ E[X_n^4] = \int_{-1/2}^{1/2} x^4 dx = \frac{9}{5}. \]

**Bernoulli Drop Random Sampling**

\[ P(T_s(n) = k \Delta) = (1 - p) p^{(k-1)} \]

So the characteristic function becomes

\[
\phi_{T_s}(s) = \left< e^{sT_s(n)} \right>_{T_s} = \left< e^{s \Delta k} \right>_{T_s} \\
= \sum_{k=1}^{\infty} e^{s \Delta k} (1 - p) p^{(k-1)} \\
= \frac{1 - p}{p} \sum_{k=1}^{\infty} \left( pe^{s \Delta} \right)^k \\
= \frac{1 - p}{p} \frac{pe^{s \Delta}}{1 - pe^{s \Delta}} \\
= \frac{(1 - p)e^{s \Delta}}{1 - pe^{s \Delta}}
\]  \( \text{C.3} \)

Mean:  \[ \text{Mean}(T_s(n)) = \frac{d \phi_{T_s}(s)}{ds} \bigg|_{s=0} = \frac{\Delta}{1 - p} \]

Variance:  \[ \text{Var}(T_s(n)) = \frac{d^2 \phi_{T_s}(s)}{ds^2} \bigg|_{s=0} - \text{Mean}(T_s(n))^2 = \frac{p}{(1 - p)^2} \Delta^2 \]

Kurtosis:  \[ K_{T_s} = \frac{\left< E \left[ \left( T_s(n) - \frac{\Delta}{1 - p} \right)^4 \right] \right>}{(\text{Var}(T_s(n)))^2} - 3 = \frac{(1 - p)^2}{p} + 6 \]

where we used the fact that  \[ E[T_s^4(n)] = \left( \frac{\Delta}{1 - p} \right)^4 \left[ 1 + 11p + 11p^2 + p^3 \right] \]
\[ E[T^3_s(n)] = \left( \frac{\Delta}{1-p} \right)^3 \left[ 1 + 4p + p^2 \right] \]

\[ E[T^2_s(n)] = \left( \frac{\Delta}{1-p} \right)^2 \left[ 1 + p \right] \]

**Convex Combination of Two Periodically-variant SSP**

\[ T_s(n) = \lambda \Delta (1 + \alpha \sin(\omega_1 n)) + (1 - \lambda) \Delta (1 + \alpha \sin(\omega_2 n)) \quad 0 \leq \lambda \leq 1 \]

**Mean:**

\[ \text{Mean}(T_s(n)) = \left\{ \lambda \Delta (1 + \alpha \sin(\omega_1 n)) + (1 - \lambda) \Delta (1 + \alpha \sin(\omega_2 n)) \right\}_n \]

\[ = \Delta + \left\{ \lambda \Delta \alpha \sin(\omega_1 n) + (1 - \lambda) \Delta \alpha \sin(\omega_2 n) \right\}_n \]

Since \( \left\langle \sin(\omega_1 n) \right\rangle_n = 0 \)

So \( \text{Mean}(T_s(n)) = \Delta \)

**Variance:**

\[ \text{Var}(T_s(n)) = \left\{ (T_s(n) - \Delta)^2 \right\}_n = \left\{ (\lambda \Delta \alpha \sin(\omega_1 n) + (1 - \lambda) \Delta \alpha \sin(\omega_2 n))^2 \right\}_n \]

\[ = \left\{ \lambda^2 \Delta^2 \alpha^2 \sin^2(\omega_1 n) + (1 - \lambda)^2 \Delta^2 \alpha^2 \sin^2(\omega_2 n) + 2\lambda(1 - \lambda) \Delta^2 \alpha^2 \sin(\omega_1 n)\sin(\omega_2 n) \right\}_n \]

Since \( \left\langle \sin^2(\omega_1 n) \right\rangle_n = \left( \frac{1 - \cos(2\omega_1 n)}{2} \right)_n = \frac{1}{2} \) and \( \left\langle \sin(\omega_1 n)\sin(\omega_2 n) \right\rangle_n = 0 \)

\[ \text{Var}(T_s(n)) = \frac{1}{2} \Delta^2 \alpha^2 \left[ \lambda^2 + (1 - \lambda)^2 \right] \quad (C.4) \]

with the relative variance index is \( \xi = \frac{\text{Var}(T_s(n))}{\text{Mean}(T_s(n))^2} = \frac{\alpha^2}{2} \left[ \lambda^2 + (1 - \lambda)^2 \right] \)

**Kurtosis:**

\[ K_T = \frac{\left\langle (T_s(n) - \text{Mean}(T_s(n)))^4 \right\rangle_n}{(\text{Var}(T_s(n))^2)} - 3 \]

\[ \left\langle (T_s(n) - \text{Mean}(T_s(n)))^4 \right\rangle_n = \left\langle (\lambda \Delta \alpha \sin(\omega_1 n) + (1 - \lambda) \Delta \alpha \sin(\omega_2 n))^4 \right\rangle_n \]

\[ = \left\{ \lambda^4 \Delta^4 \alpha^4 \sin^4(\omega_1 n) + (1 - \lambda)^4 \Delta^4 \alpha^4 \sin^4(\omega_2 n) + 6\lambda^2(1 - \lambda)^2 \Delta^4 \alpha^4 \sin^2(\omega_1 n)\sin^2(\omega_2 n) \right\}_n \]

\[ + 4\lambda^3(1 - \lambda) \Delta^4 \alpha^4 \sin^3(\omega_1 n)\sin(\omega_2 n) + \lambda(1 - \lambda)^3 \Delta^4 \alpha^4 \sin(\omega_1 n)\sin^3(\omega_2 n) \right\}_n \]

Since \( \left\langle \sin^4(\omega_1 n) \right\rangle_n = \frac{3}{8}, \left\langle \sin^2(\omega_1 n)\sin^2(\omega_2 n) \right\rangle_n = \frac{1}{4} \)
and \( \langle \sin^4(\omega_1 n)\sin(\omega_2 n) \rangle_n = \langle \sin(\omega_1 n)\sin^3(\omega_2 n) \rangle_n = 0 \)

So \( \langle (T_s(n) - \text{Mean}(T_s(n)))^4 \rangle_n = \frac{3}{8} \Delta^4 \alpha^4 \left[ \lambda^4 + (1 - \lambda)^4 + 4\lambda^2(1 - \lambda)^2 \right] \)

\[
K_{T_s} = \frac{\frac{3}{8} \Delta^4 \alpha^4 \left[ \lambda^4 + (1 - \lambda)^4 + 4\lambda^2(1 - \lambda)^2 \right]}{\frac{1}{4} \Delta^4 \alpha^4 \left[ \lambda^2 + (1 - \lambda)^2 \right]^2} - 3 = \frac{3}{2} \left[ \frac{\lambda^2 + (1 - \lambda)^2}{\lambda^2 + (1 - \lambda)^2} \right]^2 + 2\lambda^2(1 - \lambda)^2 - 3
\]
Appendix D

Perturbation Analysis of $P_{+,av}$

When the SSP is regular, the relation (3.9) holds exactly. When the SSP is mildly irregular, we should expect mild fluctuation in $P_+(t_n)$ (and similar for $P_-(t_n)$), so that $P_+(t_n) - P_{+,av}$ is small in comparison with $P_{+,av}$, and can be treated as a “perturbation”. Thus, let us define a relative perturbation $X_+$ via the expression

$$P_+(t_n) - P_{+,av} \triangleq P_{+,av}^{-\frac{1}{2}} X_+ P_{+,av}^{-\frac{1}{2}}$$

(E.1)

where $X_+$ is a Hermitian matrix corresponding to the relative deviation of $P_+(t_n)$ from $P_{+,av}$. Notice that $X_+$ is a function of “$n$”, and its average value is

$$\langle E[X_+] \rangle_n = P_{+,av}^{-\frac{1}{2}} \langle E[P_+(t_n) - P_{+,av}] \rangle_n P_{+,av}^{-\frac{1}{2}} = 0$$

Now, we can express $P_{+,av}^{-1}(t_n)$ in terms of $P_{+,av}$ and $X_+$, viz.,

$$P_{+,av}^{-1}(t_n) = P_{+,av}^{-\frac{1}{2}} [I + X_+]^{-1} P_{+,av}^{-\frac{1}{2}}$$

$$= P_{+,av}^{-\frac{1}{2}} [I - X_+ + X_+^2 - \ldots] P_{+,av}^{-\frac{1}{2}}$$

where we assume that $\|X_+\| < 1$, so that the identity $[I + X_+]^{-1} = \sum_{k=0}^{\infty} (-X_+)^k$ is valid.

Consequently,

$$\langle E[P_{+,av}^{-1}(t_n)] \rangle_n \approx P_{+,av}^{-\frac{1}{2}} \left[ I - \langle E[X_+] \rangle_n + \langle E[X_+^2] \rangle_n - \langle E[X_+^3] \rangle_n + \ldots \right] P_{+,av}^{-\frac{1}{2}}$$

Since $\langle E[X_+] \rangle_n = 0$, we get

$$\langle E[P_{+,av}^{-1}(t_n)] \rangle_n \approx P_{+,av}^{-\frac{1}{2}} \left[ I + \langle E[X_+^2] \rangle_n \right] P_{+,av}^{-\frac{1}{2}}$$

(E.2)
where we assume that \( \|X_r\| \) is small enough so that \( \|\mathbb{E} \{X_r^k\}_n\| \ll \mathbb{E} \{X_r^{2}\}_n \) for all \( k \geq 3 \). A similar result can be obtained for \( P_r(t_n) \). Applying (E.2) to (3.8) we get the (2\textsuperscript{nd} order) approximation

\[
P_{r,av}^{-\frac{1}{2}} \left[ I + \mathbb{E} \{X_r^{2}\}_n \right] P_{r,av}^{-\frac{1}{2}} = P_{r,av}^{-\frac{1}{2}} \left[ I + \mathbb{E} \{X_r^{2}\}_n \right] P_{r,av}^{-\frac{1}{2}} + C^* R^{-1} C \tag{E.3}
\]

Thus (3.10) is a reasonable approximation only when \( \mathbb{E} \{X_r^{2}\}_n \ll I \) and \( \mathbb{E} \{X_r^{2}\}_n \ll I \). Since both these average quantities are positive definite, the approximation (3.10) may hold with reasonable accuracy even when \( \|X_r\| \) is not negligible with respect to the identity matrix.
References


