A Dynamic Planning and Control System for Inventories of Raw Materials

by

G. Harhalakis and P. Sharma
A Dynamic Planning and Control System for Inventories of Raw Materials

G. Harhalakis¹, P. Sharma¹

Systems Research Center and Department of Mechanical Engineering

University of Maryland, College Park, MD 20742.

¹ Supported in part by the NSF under grant CDR-85-00108
"A Dynamic Planning and Control System for Inventories of Raw Materials"

ABSTRACT

Interfaces with operational MRP II systems can realize substantial benefits in warehousing management and control. This paper presents the design, development, programming and implementation of an inventory planning system, which, in co-ordination with MRP II, provides significant potential for increased stock turns, and reduced cash flow and storage space requirements. Finally, it presents a multi-objective ABC class analysis, combined with a method of calculating aggregate stock turns based on the turns of each item class.

Dr. G. Harhalakis\textsuperscript{1}, Mr. P. Sharma\textsuperscript{1}, Mr. W.S. Zachmann\textsuperscript{2} (CPIM)

\textsuperscript{1} Dept. of Mechanical Engineering, University of Maryland.

\textsuperscript{2} OHMEDA, Inc., Columbia, Maryland.
INTRODUCTION

Large numbers of end products, highly fluctuating production plans, and difficulties in obtaining and compiling large amounts of data sometimes cause problems in accurate estimation of storage space for future manufacturing needs in general, and inventory storage space in particular. The latter is one of the most critical resources, since 30 to 35 percent of the factory roofed space is assigned to the storage of raw materials, [8]. In addition, warehousing costs have climbed sharply and capital expenditures allotted to this function have failed to keep pace with other facets of the business. In a typical manufacturing outfit, warehousing costs for both raw materials and finished goods constitute 6 to 8 percent of the sales dollar, [2].

Also, there is a critical need for some kind of a decision making tool to aid inventory managers in identifying alternative solutions; for example, whether to lease or expand for additional warehousing space. Now that a large number of manufacturing firms have realized the importance of implementing MRP II systems, it is highly beneficial for these organizations to integrate the warehousing function with their MRP II system.Duplication of data gathering, entry and processing can be entirely eliminated, since most of the necessary data is already available in a typical MRP II system. The Master Schedule module of the MRP II system can provide direct input of the independent demand to the warehousing system.

The objective of this work has been to develop and implement a
dynamic inventory control system, which, in conjunction with MRP II, can reduce significantly storage space and cash flow requirements. In addition, it can increase stock turns and realize all of the advantages mentioned above.

PREVIOUS WORK

Work has been done in warehouse planning, but it has been treated as an isolated function. Also, most planning systems have not been designed to react to updated information in a dynamic way. In companies where the warehousing activity was analyzed [4,10], no thought was expressed as to the reduction of storage space and the increase of stock turns. There has been, however, an attempt [3] to optimize storage space by taking into account the physical size of individual components, but this did not interact with any other manufacturing modules.

THE APPROACH

In this section we describe the major input data elements.

End-Products. They are classified into families, based on their architectural or functional similarities, and assigned with product codes. This is done in order to reduce the number of items to be forecast and it is in agreement with typical Master Scheduling practices. The finalized realistic Master Schedule is used as an input to the proposed system.

Purchased Parts. The proposed model considers purchased parts as the only items to be stored, although it could be easily expanded to incorporate any intermediate components and subassemblies, if they are kept in stock for sometime. The source to identify
purchased parts is the Bill of Material (BOM) module of the MRP II system. Their source code is usually "P" for purchased. Each one of the end-products belonging to the product codes defined earlier is exploded to capture the parts which have a source code of "P". Parts with other source codes are ignored.

Once the purchased parts are identified and extracted, their quantity per assembly is stored. If the same item is used elsewhere in the same end-product, then the quantity per assembly is summed up to reflect the total quantity of the part required for that end-product. This procedure is iterated until all of the purchased parts in all products, in all class codes are identified. The output of this exercise comprises the purchased part number, the number of products making use of that part and the total quantity of the part used in all product codes.

**Volume and cost data.** The next step is to extract unit volume and unit cost data for the purchased parts identified in the previous process. These data can be extracted from the part master records maintained in the BOM module of MRP II. In the absence of part volume data, a physical volume measurement has to be done for all raw materials. The part unit volume can then be stored in an appropriate data field (either existing or specifically created for this purpose) of the MRP II system.

Unit costs and unit volumes are then multiplied by the quantity per unit of end-product. The results at this point include the part number, the total quantity per unit end-product, unit cost, unit volume, total cost and total volume. Finally these total
part costs and volumes are multiplied by the monthly production
requirements of the end-products, thus yielding the total monthly
requirements for storage volume and cash flow.

THE ALGORITHM

Two techniques were investigated before finalizing the
algorithm that was used in the system: The Lagrange multipliers
method [9], and the ABC multi-criteria analysis [1]. The former
was soon rejected, due to its complexity and its high
computational requirements.

In ABC analysis, items of a population are ranked in descending
order of some exhibited activity, usually dollar-usage, and then
appropriate techniques are developed to handle high activity "A"
items, and perhaps different techniques for medium activity "B"
items and low activity "C" items. Such a differentiation for
handling, controlling or managing different groups of a
population is much superior than treating both the important and
unimportant items in the same way.

In a classical sense, ABC groupings are based on cost factors;
i.e., high cost, high turnover items are grouped as "A" items, and
low cost, low turnover items are grouped as "C" items. Although
rare, the application of multiple criteria in ABC analysis has
proven to be very effective both in business and manufacturing
operations, [6]. For inventory items the criterion is often the
dollar-usage of the item, although sometimes it is just the item
unit cost. For many items, however, there may be other criteria
that need to be addressed, such as certainty of supply, rate of
obsolescence, impact of stockouts and so on, [5]. In our case, we considered volume usage and dollar usage, since our objective was to reduce storage space and cash flow requirements.

To provide considerable flexibility to the system, we assigned weighing factors to each of the two criteria. These factors are user defined every time the system is run. If total cost is given a higher factor (i.e. greater than 50%), the number of stock turns is increased, at the expense of storage space and storage efficiency and vice versa.

PART CLASSIFICATION PROCEDURE

Let us first introduce the following symbols:

\[ P_i = \text{Unit purchase cost of part } i, \quad Q_i = \text{Monthly usage of part } i \]
\[ V_i = \text{Unit volume of part } i, \quad n = \text{Total number of parts} \]
\[ WF_C = \text{Cost weighing factor}, \quad WF_V = \text{Volume weighing factor} \]

The average purchased part cost is given by the expression:

\[
\frac{\sum_{i}^{n} P_i Q_i}{n}
\]

Similarly, the average purchased part volume is given by the expression:

\[
\frac{\sum_{i}^{n} V_i Q_i}{n}
\]

Hence, the deviation of the cost of any given purchased part from the average part cost can be calculated. We define it as "Part Cost Ratio":

\[
\text{Part Cost Ratio (PCR)} = \frac{\sum_{i}^{n} P_i Q_i - \frac{\sum_{i}^{n} P_i Q_i}{n}}{\frac{\sum_{i}^{n} P_i Q_i}{n}}
\]
Similarly, the "Part Volume Ratio" is given by the expression:

\[ \text{Part Volume Ratio (PVR)} = \frac{\sum V_i Q_i}{\sum V_i Q_i/n} \]

We now introduce the "Part Combined Ratio", defined as:

\[ \text{Part Combined Ratio} = \text{PCR} \times \text{WF}_C + \text{PVR} \times \text{WF}_V \]

The parts combined ratios thus computed for all parts are then subjected to ABC analysis by ranking them in descending order. The first, say, 10% (user defined) are classified as class "A" items, the next 20% as class "B" and the rest as class "C" items.

**PURCHASING FREQUENCIES**

Management has now to decide on the number of months of stock to be carried for each class of parts. One example could be one month for class "A" items, three months for class "B" items and six months for class "C" items. After the purchasing policies are defined, the monthly volume usage and monthly cost usage of each item are multiplied by the respective number of months, to calculate the total part volume to be stored and the total cash commitment. A sample output of this exercise is shown on table 1.

<table>
<thead>
<tr>
<th>Comp Part #</th>
<th>Total Qty</th>
<th>Class</th>
<th>Freq</th>
<th>Tot Cost</th>
<th>Tot Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>0203-3715-46</td>
<td>1837</td>
<td>A</td>
<td>1</td>
<td>46914</td>
<td>118937</td>
</tr>
<tr>
<td>0203-4572-07</td>
<td>548</td>
<td>A</td>
<td>1</td>
<td>44861</td>
<td>119023</td>
</tr>
<tr>
<td>0203-3721-32</td>
<td>2543</td>
<td>B</td>
<td>3</td>
<td>22876</td>
<td>97099</td>
</tr>
<tr>
<td>0301-8976-31</td>
<td>1876</td>
<td>B</td>
<td>3</td>
<td>19765</td>
<td>88400</td>
</tr>
<tr>
<td>0402-3656-70</td>
<td>845</td>
<td>B</td>
<td>3</td>
<td>21908</td>
<td>77634</td>
</tr>
<tr>
<td>0504-2517-41</td>
<td>6070</td>
<td>C</td>
<td>6</td>
<td>12156</td>
<td>32960</td>
</tr>
<tr>
<td>0607-2940-05</td>
<td>2300</td>
<td>C</td>
<td>6</td>
<td>8310</td>
<td>12876</td>
</tr>
<tr>
<td>0301-2764-75</td>
<td>81</td>
<td>C</td>
<td>6</td>
<td>530</td>
<td>951</td>
</tr>
<tr>
<td>0909-2718-76</td>
<td>102</td>
<td>C</td>
<td>6</td>
<td>890</td>
<td>420</td>
</tr>
</tbody>
</table>

Table 1. Part classes, purchasing frequencies and total volumes and costs.
ASSIGNMENT OF BINS AND PALLETS

Given that purchased parts have different sizes and shapes, it is advantageous to introduce multiple bin sizes and/or pallet sizes for better storage efficiency. The total volume of each part to be stored is divided first by the volume of the smallest bin size. If the result exceeds a predetermined number of bins per part allowed (e.g. 5), the system continues with the volume of the next larger bin and so on, until a satisfactory number of bins (less than 5 in this case) can accommodate the part in question. If two or more different bin sizes can satisfy the maximum-number-of-bins constraint, the system selects the bin size which provides the maximum storage efficiency. The reason for starting with the smallest bin size is that this is the most promising for maximum volume utilization. Similar calculations are repeated for palleted items. Thus the total number of bins and pallets (or other storage devices) are calculated, together with the storage efficiency for each individual part.

EFFICIENCY INDICATORS

The efficiency of the system can be measured through the use of ratios which have to be compiled and evaluated over a period of time so that improvements can be sought for. Two such indicators are introduced here and described in the following two sections.

Overall storage efficiency. It can be computed as follows:

\[
E = \frac{TV}{\sum N_j W_j} \times 100\%
\]

Where,
\[ s = \text{Total number of bin sizes} \]
\[ TV = \text{Total volume of all parts to be stored in the period} \]
\[ N_j = \text{Total number of bins of size } j \]
\[ W_j = \text{Volume of one bin size } j \]

It is evident that the greater the efficiency, the lower the warehousing and storage costs.

**Aggregate stock turns.** Theoretically, inventory stock turns are defined as [7]:

\[
\text{Annual cost of sales} \]
\[ \text{Stock turns} = \frac{\text{Annual cost of sales}}{\text{Cost of average inventory on hand}} \]

This approach fails to measure the "performance" of each individual part and may result in wrong decisions. A technique is proposed here to compile the aggregate stock turns on the basis of stock turns for each of the A B C classes.

First, the total cash committed to all class "A" parts is obviously:

\[ a = \sum_{i} p_i Q_i, \quad \text{where:} \]
\[ i = \text{part belonging to class A} \]
\[ p_i = \text{Unit cost of part } i \]
\[ Q_i = \text{Number of parts } i \text{ needed per month} \]

Similarly, the parameters "b" and "c" are established for classes B and C respectively.

Let \[ T = a + b + c \]

Secondly, the cost proportions are calculated as follows:

\[ r_1 = a/T, \quad r_2 = b/T, \quad \text{and} \quad r_3 = c/T \]
Now, the number of stock turns in months can be calculated by class as follows:

For class "A": \( A_t = \frac{24}{\text{frequency of purchasing in months}} \)

Similarly, \( B_t \) and \( C_t \) can be calculated for items in class "B" and "C" respectively.

The combination of cost proportions and turns by class is given by:

\[
R_1 = \frac{r_1}{A_t}, \quad R_2 = \frac{r_2}{B_t}, \quad \text{and} \quad R_3 = \frac{r_3}{A_t}
\]

which are added to \( R = R_1 + R_2 + R_3 \).

Finally, the aggregate stock turns are given by

\[
A_{st} = \frac{1}{R}
\]

It is interesting to note at this point that the higher the stock turns, the lower will be the overall storage efficiency of the warehouse. This is due to the fact that stock turns reflect the frequency of purchasing; hence, the higher the frequency, the less stock is stored in the same storage area, which in turn reduces the warehousing efficiency.

**IMPLEMENTATION**

The company we implemented this system is part of a multinational organization and manufactures infant care equipment. They run over one hundred types of end-products, each comprising 300 to 500 component parts. The magnitude of parts to manage is high, despite that some of these parts are commonly used by a variety of products. The system we developed and implemented, based on the theoretical analysis of the previous sections is called **Inventory Requirements Analysis (IRA)**.
Main system. This part of the system maintains the default parameters which can be used to develop various "models", simulating different purchasing and inventory policies. Examples of these parameters include the cost/volume weighing factors, bin and pallet types and sizes, percentages of parts belonging to class A, B, and C, and the respective purchasing frequencies. The user may select to use these defaults when running the system, or modify their values for each individual run.

Model specific sub-system. The term "model" identifies a specific run of the system with a given set of parameters. The values of these parameters can be obtained from the main system (default), or overridden by the user. Thus, a number of "what-if" questions can be answered, using different combinations of input values. Each model is given a name and can be stored for future review. It can also be used as a basis for other models, by copying it to another model name and modifying some of the input values.

SYSTEM INTERFACES

IRA is interfaced with the MRP II system the company uses on a regular basis. Data is extracted from part master records, bills of materials, master production schedules, inventory control and purchasing. There was no need to interfere with the source code of MRP II, although it was available. It was sufficient to read the data from the system files associated with the above modules.

Part master records and BOMs. This is the source for identifying finished products to be exploded down to the purchased parts level. Periodic BOM explosions were necessary to account for BOM
updates performed by the product design group. The quantity of each component part per assembly unit is extracted and, for commonly used parts, accumulated against the part record. The process is iterated until all parts, in all end-products, in all product classes are accounted for.

**Independent demand.** The master production schedule provides data on sales forecasts and customer orders. The user defines the number of months to look ahead in terms of production commitments. Also, the user may define the way to compute the monthly production rates of end-products: (1) to select the arithmetic average of all projected monthly figures, (2) to adopt the maximum monthly production rate as a constant rate throughout the planning horizon, (3) to use the mean absolute deviation to discard any extremely high or low rates.

**INPUT PARAMETERS**

The following input parameters must be pre-defined:

- Part cost weighing factor (0...100%, default: 50%)
- Part volume weighing factor (0...100%, default: 50%)
- Part classification groups (Up to 10 part classes can be defined, default: A,B,C.)
- Percentage of parts to be held in each group (Default: 10%, 20%, and 70% respectively)
- Purchasing frequency for each parts group (Default: 1 month, 3 months, and 6 months respectively)
- Types, quantities, and dimensions of available bins
- End-product class codes to be considered (default: ALL)
-Computation method of monthly production rates of end-products (Default: Arithmetic average)
-User overrides. Examples include purchasing frequencies for certain parts under special vendor terms, unit part cost and lead time if supplied by a non-default vendor, etc.

Clearly, there is a considerable amount of flexibility in defining inventory and purchasing policies, in compliance with the ever changing manufacturing environment.

SYSTEM OUTPUTS

The results provided after each model run are both on-line and in printed form. They can be generally categorized as follows:

- A detailed breakdown of all purchased parts accounted for, with their part number, description, monthly cash commitment, monthly volume requirement, individual storage efficiency, and part combined ratio.
- A purchasing report, displaying the buyer, vendor, cost, lead time and monthly quantity requirement for each part.
- A storage report, displaying the aggregate storage requirement against storage volume availability, and percentage storage efficiency. (See figure 1).
- Stock turns, both by part class and aggregate.

Finally, a plotting of aggregate stock turns for different cost weighing factors (between 0% and 100%) can be obtained, as shown in figure 2. Clearly, the impact of high cash flow concerns on stock turns is positive.
BENEFITS

The outputs of the system are subject to review and can instigate appropriate corrective action. So far, a number of changes in the company’s purchasing and inventory policies have been experienced. Vendors have been approached with the aim to reduce lead times and minimum order quantities. Smaller size bins have been introduced for more efficient storage of small size components. More storage space has been leased in anticipation of growth by about 30% over the next 9 months. Finally, cash flow requirements have been reduced by 30%, while stock turns have increased by 40% (from 2.5 to 3.5).

CONCLUDING REMARKS

The idea of developing dynamic modeling and simulation systems, run in almost real time, and in conjunction with existing application modules, seems to open new directions in production and inventory control. The capabilities of MRP II can be enhanced to assist in "informed" decision making, by the development of systems run in parallel and using existing data. Thus, tremendous savings in data entry and duplication can be experienced. Finally, correct and fast interpretation of results, thoughtfully presented, leads to actions that ultimately improve the performance of the company.

ACKNOWLEDGEMENTS

The authors greatly appreciate the cooperation and support from the management and staff of OHMEDA, over the entire duration of the project. Special recognition is due to Mr. William Zachmann
for his wholehearted collaboration and useful advice. Thanks are also due to Mr. William Rhoad for his professional programming work in the IRA system.

REFERENCES


Figure 1. Storage space available and required
Figure 2. Stock turns for different cost weighing factors.
but need not satisfy any absolute continuity conditions with respect to $P$. Before proving this main result, let us prove a standard measurability result which we shall need later. Let $IN = \{1, 2, \ldots\}$.

**Lemma 2.1.** Let $(E, \mathcal{E})$ be a measurable space and $\{X_t\}_{t=1}^\infty$ be a collection of IR-valued $\mathcal{E}$-measurable mappings from $E$ to IR. Then

$$\sigma\{X_n; n = 1, 2, \ldots\} = \{\{(X_1, X_2, \ldots) \in A\} : A \in B(IR^N)\}.$$  \hspace{1cm} (2.35)

**Proof.** For convenience, define the $\sigma$-fields

$$\mathcal{G} := \sigma\{X_n; n \in IN\}$$ \hspace{1cm} (2.36)

$$\mathcal{H} := \{\{(X_1, X_2, \ldots) \in A\} : A \in B(IR^N)\}.$$ \hspace{1cm} (2.37)

Take $n$ in $IN$ and $D \in B(IR)$, and define $D^*$ by

$$D^* := \{(x_1, x_2, \ldots) : x_n \in D\}.$$ \hspace{1cm} (2.38)

Since

$$\{X_n \in D\} = \{(X_1, X_2, \ldots) \in D^*\} \in \mathcal{H},$$ \hspace{1cm} (2.39)

we see that $\mathcal{G} \subset \mathcal{H}$.

To prove the reverse inclusion, define the $\sigma$-field $\mathcal{J}$ by

$$\mathcal{J} := \{D^* \in B(IR^N) : \{(Y_1, Y_2, \ldots) \in D^*\} \in \mathcal{G}\}.$$ \hspace{1cm} (2.40)

Now let $\mathcal{L}$ be the class of cylinder sets in $B(IR^N)$. Take $D^* \in \mathcal{L}$. Then there exists a collection $\{n_1, n_2, \ldots, n_k\}$ of integers and a collection $\{A_i\}_{i=1}^k$ of sets in $B(IR)$ such that

$$D^* = \{(x_1, x_2, \ldots) : x_n, \in A_i; i = 1, 2, \ldots, k\}.$$ \hspace{1cm} (2.41)

Consequently,

$$\{(Y_1, Y_2, \ldots) \in D^*\} = \bigcap_{i=1}^k \{Y_{n_i} \in A_i\} \in \mathcal{H},$$ \hspace{1cm} (2.42)

so $\mathcal{L} \subset \mathcal{J}$, and $B(IR^N) = \sigma(\mathcal{L}) \subset \mathcal{J}$. Hence $\mathcal{H} \subset \mathcal{G}$.

We can now prove the following main result.
Theorem 2.4. Assume that the $\sigma$-field $\mathcal{F}_t$ is separable for $t = 0, 1, \ldots$. Then there is a probability measure $\hat{P}$ on $(\Omega, \mathcal{F}_\infty)$ enjoying properties (C.1) and (C.2).

Proof. The main thrust of the proof is to use the separability of $\{\mathcal{F}_t\}_{t=0}^\infty$ to generate a collection of finite-dimensional distributions from $\{\hat{P}_t\}_{t=0}^\infty$ and to then generate a probability measure $\hat{P}$ on $(IR^N, \mathcal{B}(IR^N))$ from these finite-dimensional distributions by the Daniell-Kolmogorov Extension Theorem. We shall then generate the probability measure $\hat{P}$ from the probability triple $(IR^N, \mathcal{B}(IR^N), \hat{P})$.

Fix $t = 0, 1, \ldots$. Since $\mathcal{F}_t$ is separable, there is a countable collection $\{X_{t,i}\}_{i=1}^\infty$ of $IR$-valued RV's such that $\mathcal{F}_t = \sigma\{X_{t,i}; i \in IN\}$. Now take any ordering of $\{(i, j); i \in IN \cup \{0\}; j \in IN\}$—any mappings $I : IN \to IN \cup \{0\}$ and $J : IN \to IN$ such that

$$\{(I(n), J(n)); n \in IN\} = \{(i, j); i = IN \cup \{0\}; j \in IN\}$$

and such that if $(I(n), J(n)) = (I(m), J(m))$, then $m = n$. Define the random variables $\{Y_n\}_{n=1}^\infty$ by

$$Y_n := X_{I(n), J(n)}. \quad n = IN \cup \{0\}$$

For any collection of positive integers $\{n_1, n_2, \ldots, n_k\}$, define the distribution function

$$F_{n_1, n_2, \ldots, n_k}(y_1, y_2, \ldots, y_k) = \hat{P}_{\max_{1 \leq i \leq k} I(n_i)}\{Y_n \leq y_i; i = 1, 2, \ldots, k\}$$

for $\{y_1, y_2, \ldots, y_k\}$ in $IR$.

Now take any collections $\{n_1, n_2, \ldots, n_k\}$ in $IN$ and $\{y_1, y_2, \ldots, y_k\}$ in $IR$. It is easy to see that

$$F_{n_1, n_2, \ldots, n_k}(y_{n_{i_1}}, y_{n_{i_2}}, \ldots, y_{n_{i_k}}) = F_{n_1, n_2, \ldots, n_k}(y_1, y_2, \ldots, y_k)$$

for any permutation $\{i_1, i_2, \ldots, i_k\}$ of $\{1, 2, \ldots, k\}$ and that

$$F_{n_1, n_2, \ldots, n_l}(y_1, y_2, \ldots, y_l) = F_{n_1, n_2, \ldots, n_k}(y_1, y_2, \ldots, y_l, \infty, \infty, \ldots, \infty)$$

for $l < k$ (assuming that $k \geq 2$). The Daniell-Kolmogorov Existence Theorem ([13, Thm 1.1.7]) then ensures the existence of a unique probability measure $\hat{P}$ on $(IR^N, \mathcal{B}(IR^N))$ with finite-dimensional distributions $\{F_{n_1, n_2, \ldots, n_k}\}$; for any collections $\{n_1, n_2, \ldots, n_k\}$ in $IN$ and $\{y_1, y_2, \ldots, y_k\}$ in $IR$,

$$\hat{P}\{(x_1, x_2, \ldots) : x_{n_i} \leq y_i \text{ for } i = 1, 2, \ldots, k\} = F_{n_1, n_2, \ldots, n_k}(y_1, y_2, \ldots, y_k).$$
Using Lemma 2.1, we can now define $\tilde{P}$. Take $A$ in $\mathcal{F}_\infty$. By Lemma 2.1, we can find $A^* \in \mathcal{B}(\mathbb{IR}^N)$ such that

$$A = \{(Y_1, Y_2, \ldots) \in A^*\},$$

(2.49)

so we define

$$\tilde{P}(A) := \tilde{P}(A^*).$$

(2.50)

Clearly $\tilde{P}$ is a probability measure on $(\Omega, \mathcal{F}_\infty)$. All that now remains is to show that $\tilde{P}$ enjoys property (C.2). Take $t = 0, 1, \ldots$ and define $IK \subset \mathbb{IN}$ by

$$IK := \{n \in \mathbb{IN} : I(n) = t\};$$

(2.51)

we have that $\mathcal{F}_t = \sigma\{Y_n; n \in IK\}$. Now define the $\pi$-system $\mathcal{D}$ by

$$\mathcal{D} := \{\{Y_{n_i} \leq y_i; i = 1, 2, \ldots, k\}; \{n_1, n_2, \ldots, n_k\} \subset IK, \{y_1, y_2, \ldots, y_k\} \subset \mathbb{IR}\},$$

(2.52)

and note that $\mathcal{F}_t = \sigma(\mathcal{D})$. By a well-known result [3, Thm 3.3], $\tilde{P}$ and $\hat{P}_t$ agree on $\mathcal{F}_t$ if $\tilde{P}(D) = \hat{P}_t(D)$ for all $D \in \mathcal{D}$; take

$$D = \{Y_{n_i} \leq y_i; i = 1, 2, \ldots, k\},$$

(2.53)

where $\{n_1, n_2, \ldots, n_k\} \subset IK$ and $\{y_1, y_2, \ldots, y_k\} \subset \mathbb{IR}$. By (2.45) and (2.46),

$$\tilde{P}(D) = F_{n_1, n_2, \ldots, n_k}(y_1, y_2, \ldots, y_k)$$

$$= \hat{P}_t\{Y_{n_i} \leq y_i; i = 1, 2, \ldots, k\}$$

$$= \hat{P}_t(D),$$

(2.54)

so that $\tilde{P}$ does in fact enjoy property (C.2), and the proof is complete.

\[\square\]
CHAPTER III: THE UNCORRELATED PROBLEM

III.1. Overview of the Uncorrelated Problem

We consider here the situation in which the processes \( \{W_{t+1}^o\}_0^\infty \) and \( \{V_{t+1}^o\}_0^\infty \) are independent and the GWN sequence \( \{V_{t+1}^o\}_0^\infty \) is standard; here

\[
\Gamma_{t+1} = \begin{pmatrix}
\Sigma_{t+1}^w & 0 \\
0 & I_k
\end{pmatrix}, \quad t = 0, 1, \ldots \quad (1.1)
\]

To standardize the notation, let \( \{\mathcal{Y}_t\}_0^\infty \) be a filtration of \( \mathcal{F} \) generated by the observation process \( \{Y_t\}_0^\infty \);

\[
\mathcal{Y}_t := \sigma\{Y_s; \ s = 0, 1, \ldots, t\}. \quad t = 0, 1, \ldots \quad (1.2)
\]

We thus seek \( E[\phi(X_{t+1}^o)|\mathcal{Y}_t] \) for each \( \phi \) in \( \mathcal{Z} \) and each \( t = 0, 1, \ldots \). We shall find representations for these conditional expectations by fixing an arbitrary \( \phi \) in \( \mathcal{Z} \) and finding formulæ for \( \{E[\phi(X_{t+1}^o)|\mathcal{Y}_t]\}_0^\infty \). We shall do this in two steps. In Section 2, we shall fix \( T = 0, 1, \ldots \) and find a representation for \( \{E[\phi(X_{t+1}^o)|\mathcal{Y}_t]\}_0^T \), as was done in [17] and [19]. In Section 3, we shall extend these representations to the infinite-horizon process \( \{E[\phi(X_{t+1}^o)|\mathcal{Y}_t]\}_0^\infty \). Finally, in Section 4, we shall discuss the methods and procedures used in Sections 2 and 3, in the hope of making the arguments of Chapter IV follow more naturally.
III.2. A Change of Measure and a Representation Result for a Finite Horizon

Fix an arbitrary $T = 0, 1, \ldots$ and an element $\phi$ in $\mathcal{Z}$. We here consider the problem of finding a representation for $\{E[\phi(X_{t+1}^\circ)|Y_t]\}_0^T$. Following [17] and [18], we first separate the process $\{X_t^\circ\}_0^\infty$ into two $IR^m$-valued processes $\{X_t\}_0^\infty$ and $\{Z_t\}_0^\infty$, the former representing the effects of the noise $\{W_{t+1}^\circ\}_0^\infty$ and the latter representing the effects of the initial condition $\xi$; define the processes $\{X_t\}_0^\infty$ and $\{Z_t\}_0^\infty$ according to the recursions

$$X_{t+1} = A_t X_t + W_{t+1}^\circ$$

$$X_0 = 0$$

and

$$Z_{t+1} = A_t Z_t$$

$$Z_0 = \xi.$$  

Using the state transition matrix $\Phi(\cdot, \cdot)$, we may represent (2.2) as

$$Z_t = \Phi(t, 0) \xi.$$  

We may also make the decomposition

$$X_t^\circ = X_t + Z_t,$$  

and if we define a new 'noise' process $\{V_{t+1}\}_0^\infty$ taking values in $IR^k$ by

$$V_{t+1} := V_{t+1}^\circ + H_t Z_t,$$  

the observation process $\{Y_t\}_0^\infty$ may then be represented as

$$Y_t = H_t X_t + V_{t+1}.$$  

As in [18], we next define a filtration $\{\mathcal{F}_t\}_0^\infty$ of $\mathcal{F}$ by

$$\mathcal{F}_0 := \sigma(\xi, W_{s+1}^\circ; s = 0, 1, \ldots)$$

$$\mathcal{F}_{t+1} := \mathcal{F}_0 \vee \sigma(V_{s+1}^\circ; s = 0, 1, \ldots, t).$$  

Clearly $\mathcal{Y}_t \subset \mathcal{F}_{t+1}$ for each $t = 0, 1, \ldots$ (In (2.1)-(2.7), we have anticipated the arguments of Section 3 by defining the various entities for $t = 0, 1, \ldots$ rather than for $t = 0, 1, \ldots, T$.)

24
Upon inspecting [17] and [19], it is evident that we seek a probability measure \( \hat{P}_{T+1} \) on \((\Omega, \mathcal{F})\) such that 

(E.1): the probability measures \( P \) and \( \hat{P}_{T+1} \) are mutually absolutely continuous,

(E.2): the probability measures \( P \) and \( \hat{P}_{T+1} \) agree on \( \mathcal{F}_0 \) (i.e., \( P(A) = \hat{P}_{T+1}(A) \) for all \( A \) in \( \mathcal{F}_0 \)) and

(E.3): the process \( \{V_{t+1}\}_0^T \) is a \((\mathcal{F}_t, \hat{P}_{T+1})\) zero-mean standard GWN sequence.

An examination of Chapter II—with special attention to (2.5)—leads us to a process \( \{L_t\}_0^\infty \) through which we can define a probability measure \( \hat{P}_{T+1} \) enjoying properties (E.1)-(E.3). Set

\[
L_{t+1} := \exp\left[-\sum_{s=0}^{t}[H_sZ_s]'V_{s+1}^0 - \frac{1}{2}\sum_{s=0}^{t}[H_sZ_s]'[H_sZ_s]\right] \quad t = 0, 1, \ldots \quad (2.8)
\]

\( L_0 := 1 \)

and define the probability measure \( \hat{P}_{T+1} \) by the Radon-Nikodym derivative

\[
\frac{d\hat{P}_{T+1}}{dP} = L_{T+1}. \quad (2.9)
\]

We see from Chapter II that \( \{L_t\}_0^\infty \) is an \((\mathcal{F}_t, P)\)-martingale. We may summarize the statistical nature of the RV’s \( \xi, \{W_{t+1}^0\}_0^T \) and \( \{V_{t+1}\}_0^T \) under \( \hat{P}_{T+1} \) by

(E.4): the RV’s \( \xi, \{W_{t+1}^0\}_0^T \) and \( \{V_{t+1}\}_0^T \) have the same joint statistics under \( \hat{P}_{T+1} \) as the RV’s \( \xi, \{W_{t+1}^0\}_0^T \) and \( \{V_{t+1}\}_0^T \) have under \( P \).

The representation of \( \{L_t\}_0^\infty \) given in (2.8) will not be entirely adequate for our purposes. We may simplify it by introducing the following notation. Define an \( IR^n \)-valued process \( \{B_t\}_0^\infty \) according to the recursion

\[
B_{t+1} = B_t + \Phi'(t,0)H_t'V_{t+1} \quad t = 0, 1, \ldots \quad (2.10)
\]

\( B_0 = 0 \)

and a deterministic sequence \( \{M_t\}_0^\infty \) in \( Q_\infty \) according to

\[
M_{t+1} = M_t + \Phi'(t,0)H_t'H_t\Phi(t,0) \quad t = 0, 1, \ldots \quad (2.11)
\]

\( M_0 = 0 \).

25
From (2.5) and (II.1.4), it is easy to verify that
\begin{equation}
- \sum_{s=0}^{t} [H_{s}Z_{s}]^{\prime}V_{s+1} - \frac{1}{2} \sum_{s=0}^{t} [H_{s}Z_{s}]^{\prime}[H_{s}Z_{s}] = -\xi^{\prime}B_{t+1} + \frac{1}{2} \xi^{\prime}M_{t+1} \xi, \quad t = 0, 1, \ldots \tag{2.12}
\end{equation}
so that
\begin{equation}
L_{t} = \exp[-\xi^{\prime}B_{t} + \frac{1}{2} \xi^{\prime}M_{t} \xi], \quad t = 0, 1, \ldots \tag{2.13}
\end{equation}

We now exploit the properties of \( P_{T+1} \) to solve the prediction problem for the finite horizon \( t = 0, 1, \ldots, T \). Let \( \bar{E}_{T+1}[\cdot] \) be the expectation operator associated with \( P_{T+1} \), and let the conditional expectation operators be similarly defined. For any bounded \( \mathcal{C} \)-valued \( \mathcal{F}_{T+1} \)-measurable RV \( \rho \), define
\begin{equation}
\sigma_{t}[\rho] := \bar{E}_{T+1}[\rho L_{T+1}^{-1} | \mathcal{Y}_{t}]. \quad t = 0, 1, \ldots \tag{2.14}
\end{equation}

Then it is well-known (see [16, Sec 27.4]) that
\begin{equation}
E[\rho | \mathcal{Y}_{t}] = \frac{\sigma_{t}[\rho]}{\sigma_{t}[1]} \quad \text{P-a.s.} \quad t = 0, 1, \ldots \tag{2.15}
\end{equation}

We can now use several properties of the conditional expectation operator to derive an alternate representation for \( \sigma_{t}[\phi(X_{t+1}^{o})]^{T} \). However, we first verify the following results, which are easy implications of (E.4). \( \Box \)

**Lemma 2.1.** Under \( P_{T+1} \), the process \( \{ (X_{t+1}, B_{t+1}) \}_{0}^{T} \) and \( \{ Y_{t} \}_{0}^{T} \) are jointly Gaussian and independent of the RV \( \xi \).

**Proof.** The dynamical equations (2.1) and (2.10) and the observation equation (2.6) imply that for each \( t = 0, 1, \ldots, T \), \( (X_{t+1}, B_{t+1}) \) and \( Y_{t} \) are linear functions of \( \{(W_{t+1}^{o}, V_{t+1})\}_{0}^{T} \). Thus \( \{(X_{t+1}, B_{t+1})\}_{0}^{T} \) and \( \{ Y_{t} \}_{0}^{T} \) are Gaussian under \( P_{T+1} \) and furthermore are measurable with respect to the \( \sigma \)-field \( \sigma\{W_{t+1}^{o}, V_{t+1}; t = 0, 1, \ldots, T\} \), which is independent of \( \sigma\{\xi\} \). \( \Box \)

Now use the Law of Iterated Conditioning and representations (2.4) and (2.13) to write
\begin{align}
\sigma_{t}[\phi(X_{t+1}^{o})] &= \bar{E}_{T+1} \left[ \bar{E}_{T+1} \left[ \phi(X_{t+1}^{o}) \exp[\xi^{\prime}B_{t+1} - \frac{1}{2} \xi^{\prime}M_{t+1} \xi] | \mathcal{Y}_{t} \vee \sigma(\xi) \right] | \mathcal{Y}_{t} \right] \\
&= \bar{E}_{T+1} \left[ \phi(X_{t+1} + Z_{t+1}) \exp[\xi^{\prime}B_{t+1} - \frac{1}{2} \xi^{\prime}M_{t+1} \xi] | \mathcal{Y}_{t} \vee \sigma(\xi) \right] \tag{2.16}
\end{align}
for each $t = 0, 1, \ldots, T$. Since the RV's $\{(X_{t+1}, B_{t+1})\}_0^T$ and $\{Y_t\}_0^T$ are jointly Gaussian under $\bar{P}_{T+1}$, the quantities

$$
\begin{align*}
\begin{pmatrix} \bar{X}_{t+1} \\
\bar{B}_{t+1} 
\end{pmatrix} &:= \bar{E}_{T+1} \begin{pmatrix} X_{t+1} \\
B_{t+1} \end{pmatrix} | Y_t \\
t & = 0, 1, \ldots, T 
\end{align*}
\quad (2.17)
$$

$$
\begin{align*}
\begin{pmatrix} \tilde{X}_{t+1} \\
\tilde{B}_{t+1} 
\end{pmatrix} &:= \begin{pmatrix} X_{t+1} \\
B_{t+1} \end{pmatrix} - \begin{pmatrix} \bar{X}_{t+1} \\
\bar{B}_{t+1} \end{pmatrix} \\
t & = 0, 1, \ldots, T 
\end{align*}
\quad (2.18)
$$

$$
\Sigma_{t+1} := \bar{E}_{T+1} \begin{pmatrix} \tilde{X}_{t+1} \\
\tilde{B}_{t+1} 
\end{pmatrix} \begin{pmatrix} \tilde{X}_{t+1} & \tilde{B}_{t+1} \end{pmatrix}' \\
t & = 0, 1, \ldots, T 
\quad (2.19)
$$

are easily generated using standard linear estimation theory [27, Sec 2.7]. By the $\bar{P}_{T+1}$-independence of the $\sigma$-fields $\sigma\{\xi\}$ and $\sigma\{X_{t+1}, B_{t+1}, Y_t; t = 0, 1, \ldots, T\}$, which we established in Lemma 2.1, it is then plain that ([13, Prop 6.1.7])

$$
\bar{E}_{T+1} \begin{pmatrix} X_{t+1} \\
B_{t+1} \end{pmatrix} | Y_t \vee \sigma\{\xi\} = \begin{pmatrix} \bar{X}_{t+1} \\
\bar{B}_{t+1} \end{pmatrix}, \\
t & = 0, 1, \ldots, T 
\quad (2.20)
$$

so that for each $t = 0, 1, \ldots, T$,

$$
\begin{align*}
\bar{E}_{T+1} \left[ \phi(X_{t+1} + Z_{t+1}) \exp[\xi' B_{t+1} - \frac{1}{2} \xi' M_{t+1} \xi] | Y_t \vee \sigma\{\xi\} \right] \\
= \bar{E}_{T+1} \left[ \phi(\tilde{X}_{t+1} + \tilde{X}_{t+1} + Z_{t+1}) \exp[\xi' \tilde{B}_{t+1} + \xi' \tilde{B}_{t+1} - \frac{1}{2} \xi' M_{t+1} \xi] | Y_t \vee \sigma\{\xi\} \right] \\
= \bar{E}_{T+1} \left[ \phi(\tilde{X}_{t+1} + (\tilde{X}_{t+1} + Z_{t+1})) \exp[\xi' \tilde{B}_{t+1}] | Y_t \vee \sigma\{\xi\} \right] \exp[\xi' \tilde{B}_{t+1} - \frac{1}{2} \xi' M_{t+1} \xi] \\
= \bar{E}_{T+1} \left[ \phi(\tilde{X}_{t+1} + u) \exp[\xi' \tilde{B}_{t+1}] \right]_{u = \tilde{X}_{t+1} + Z_{t+1}} \exp[\xi' \tilde{B}_{t+1} - \frac{1}{2} \xi' M_{t+1} \xi] \\
= T \phi(\tilde{X}_{t+1} + Z_{t+1}, \xi; \Sigma_{t+1}) \exp[\xi' \tilde{B}_{t+1} - \frac{1}{2} \xi' M_{t+1} \xi]; \\
\end{align*}
\quad (2.21)
$$

(2.22) follows from (2.21) since $\xi' \tilde{B}_{t+1} - \frac{1}{2} \xi' M_{t+1} \xi$ is $Y_t \vee \sigma\{\xi\}$-measurable.

From well-known results in estimation theory for Gaussian RV's (see [27, Sec 2.7]), $(\tilde{X}_{t+1}, \tilde{B}_{t+1})$ is independent of $Y_t$ under $\bar{P}_{T+1}$. Since the processes $\{(X_{t+1}, B_{t+1})\}_0^T$ and $\{(\tilde{X}_{t+1}, \tilde{B}_{t+1})\}_0^T$ are independent of the RV $\xi$ under $\bar{P}_{T+1}$ by Lemma 2.1, we see by (2.18) that the process $\{(\tilde{X}_{t+1}, \tilde{B}_{t+1})\}_0^T$ is also independent of $\sigma\{\xi\}$ under $\bar{P}_{T+1}$, and thus (2.23) follows from (2.22) (see [13, Prop. 6.1.15]). Eq. (2.24) is simply an application of the definition of $T \phi$. 

27
Recalling (2.16), we may now represent \( \sigma_t[\phi(X_t^{0+1})] \) as

\[
\sigma_t[\phi(X_t^{0+1})] = \bar{E}_{T+1} [T \phi[X_{t+1} + \Phi(t + 1, 0)\xi, \xi; \Sigma_{t+1}] \exp[\xi' B_t + \frac{1}{2} \xi' M_{t+1} \xi] | \mathcal{Y}_t]\tag{2.25}
\]

\[
= \bar{E}_{T+1} \left[ T \phi[x + \Phi(t + 1, 0)\xi, \xi; \Sigma_{t+1}] \exp[\xi'b - \frac{1}{2} \xi' M_{t+1} \xi] \right]_{b = B_t + 1, t + 1} \tag{2.26}
\]

\[
= U \phi[\bar{X}_{t+1}, B_{t+1}; M_{t+1}, \Phi(t + 1, 0); \Sigma_{t+1}] \tag{2.27}
\]

for each \( t = 0, 1, \ldots, T \). Here \((\bar{X}_{t+1}, \bar{B}_{t+1})\) is measurable with respect to the \( \sigma \)-field \( \mathcal{Y}_t \), which by Lemma 2.1 is \( \bar{P}_{T+1} \)-independent of \( \sigma\{\xi\} \), so that (2.25) implies (2.26) by standard arguments ([13, Prop. 6.15]). Recalling that the RV \( \xi \) has the same distribution under both \( P \) and \( \bar{P}_{T+1} \), we reach (2.27) from (2.26) by the definition of \( U \phi \).

We now have the representation of \( \{E[\phi(X_t^{0+1})|\mathcal{Y}_t]\}^T_0 \) that we sought. Upon combining (2.27) and (2.15), we get

\[
E[\phi(X_t^{0+1})|\mathcal{Y}_t] = \frac{U \phi[\bar{X}_{t+1}, B_{t+1}; M_{t+1}, \Phi(t + 1, 0); \Sigma_{t+1}]}{U \mathbb{I}[\bar{X}_{t+1}, B_{t+1}; M_{t+1}, \Phi(t + 1, 0); \Sigma_{t+1}].} \quad t = 0, 1, \ldots, T \tag{2.28}
\]

We summarize the efforts of this section in the following proposition.

**Theorem 2.1.** Consider the sequences \( \{X_t\}^\infty_0, \{Z_t\}^\infty_0, \{V_t\}^\infty_0, \{L_t\}^\infty_0, \{B_t\}^\infty_0 \) and \( \{M_t\}^\infty_0 \) as defined in this section and the probability measure \( \bar{P}_{T+1} \) on (\( \Omega, \mathcal{F} \)) as defined by (2.9). Then \( \bar{P}_{T+1} \) has properties (E.1)-(E.4). The process \( \{L_t\}^\infty_0 \) is an \( (\mathcal{F}_t, P) \)-martingale, and Lemma 2.1 holds. Also, let \( \{(\bar{X}_{t+1}, B_{t+1})\}^T_0 \) and \( \{\Sigma_{t+1}\}^T_0 \) be defined as in this section. Then for each \( \phi \) in \( \mathcal{Z} \), we have the representation

\[
E[\phi(X_t^{0+1})|\mathcal{Y}_t] = \frac{U \phi[\bar{X}_{t+1}, B_{t+1}; M_{t+1}, \Phi(t + 1, 0); \Sigma_{t+1}]}{U \mathbb{I}[\bar{X}_{t+1}, B_{t+1}; M_{t+1}, \Phi(t + 1, 0); \Sigma_{t+1}].} \quad t = 0, 1, \ldots, T \tag{2.29}
\]
III.3. The Infinite Horizon Problem

In this section, we extend the results of Section 2 to the problem of finding a representation for \( \{ E[\phi(X^o_{t+1})|\mathcal{Y}_t]\}_0^\infty \) for a specific \( \phi \) in \( \mathcal{Z} \), and then we specify methods for computing \( \{ (\tilde{X}_{t+1}^o, \tilde{B}_{t+1}) \}_0^\infty \) and \( \{ \Sigma_{t+1} \}_0^\infty \); we will then have a complete solution to the prediction problem. (Note that we have really only defined the process \( \{ (\tilde{X}_{t+1}^o, \tilde{B}_{t+1}) \}_0^T \) for a fixed but arbitrary \( T = 0, 1, \ldots \).

By combining (II.2.1), (II.2.10), and (II.2.6), we have the dynamical representation

\[
\begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix}
= \begin{pmatrix}
A_t & 0 \\
0 & I_n
\end{pmatrix}
\begin{pmatrix}
X_t \\
B_t
\end{pmatrix}
+ \begin{pmatrix}
I_n & 0 \\
0 & \Phi'(t, 0) H_t'
\end{pmatrix}
\begin{pmatrix}
W^o_{t+1} \\
V_{t+1}
\end{pmatrix}
\]

\[
\begin{pmatrix}
X_0 \\
B_0
\end{pmatrix}
= \begin{pmatrix}
0 \\
0
\end{pmatrix}
\quad t = 0, 1, \ldots
\quad (3.1)
\]

\[
Y_t = \begin{pmatrix}
H_t & 0
\end{pmatrix}
\begin{pmatrix}
X_t \\
B_t
\end{pmatrix}
+ \begin{pmatrix}
0 & I_k
\end{pmatrix}
\begin{pmatrix}
W^o_{t+1} \\
V_{t+1}
\end{pmatrix}.
\]

To extend the results of Section 2 to an infinite horizon, we define a sequence of probability measures \( \{ \bar{P}_t \}_0^\infty \) on \( (\Omega, \mathcal{F}) \) by

\[
\frac{d\bar{P}_t}{dP} = L_t, \quad t = 0, 1, \ldots
\quad (3.2)
\]

For each \( t = 0, 1, \ldots \), let \( \bar{E}_t \) be the expectation operator associated with \( \bar{P}_t \), and let the conditional expectation operators be similarly defined. The results of Section 3 clearly imply that the sequence \( \{ \bar{P}_t \}_0^\infty \) has the following properties:

(F.1): for \( t = 0, 1, \ldots \) the probability measures \( \bar{P}_t \) and \( \bar{P}_{t+1} \) agree on \( \mathcal{F}_t \) and

(F.2): for \( t = 0, 1, \ldots \) and \( s = 0, 1, \ldots, t \) and for any \( \sigma \)-field \( \mathcal{G} \subset \mathcal{F}_s \) and for any bounded \( \mathcal{F}_s \)-measurable RV \( \rho \), \( \bar{E}_t[\rho|\mathcal{G}] = \bar{E}_s[\rho|\mathcal{G}] \).

The extension of the results of Section 2 to an infinite horizon is performed by ‘past ing’ together the solutions for different finite horizons. We first define the conditional means and variances by

\[
\begin{pmatrix}
\bar{X}_{T+1}^T \\
\bar{B}_{T+1}^T
\end{pmatrix} := \bar{E}_{T+1}
\begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} \bigg| \mathcal{Y}_t
\]

\[
\Sigma_{T+1}^T := \bar{E}_{T+1}
\begin{pmatrix}
((X_{t+1} - \bar{X}_{T+1}^T) - (X_{t+1} - \bar{X}_{T+1}^T))((X_{t+1} - \bar{X}_{T+1}^T) - (X_{t+1} - \bar{X}_{T+1}^T))'
\end{pmatrix}
\quad (3.3)
\]

29
for each $T = 0, 1, \ldots$ and $t = 0, 1, \ldots, T$, and for notational convenience, we set

$$
\begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix}
:=
\begin{pmatrix}
\bar{X}_{t+1} \\
\bar{B}_{t+1}
\end{pmatrix}
\quad t = 0, 1, \ldots
$$

$$
\Sigma_{t+1} := \Sigma_{t+1}^t.
$$

We know from Theorem II.2.1 that

$$
E[\phi(X_{t+1})|\mathcal{Y}_t] = \frac{U\phi(X_{t+1}, B_{t+1}; M_{t+1}, \Phi(t + 1, 0); \Sigma_{t+1})}{U[\Phi(X_{t+1}, B_{t+1}; M_{t+1}, \Phi(t + 1, 0); \Sigma_{t+1})]}, \quad t = 0, 1, \ldots
$$

The following rather obvious result reveals a simple structure for $\{(X_{t+1}, B_{t+1})\}_0^\infty$ and $\{\Sigma_{t+1}\}_0^\infty$.

**Proposition 3.1.** The processes $\{(X_{t+1}, B_{t+1})\}_0^\infty$ and $\{\Sigma_{t+1}\}_0^\infty$ are related to the processes $\{(\bar{X}_{t+1}^T, \bar{B}_{t+1}^T); T = 0, 1, \ldots, t = 0, 1, \ldots, T \}$ and $\{\Sigma_{t+1}^T; T = 0, 1, \ldots, t = 0, 1, \ldots, T \}$ by

$$
\begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix}
= \begin{pmatrix}
\bar{X}_{t+1}^T \\
\bar{B}_{t+1}^T
\end{pmatrix}
\quad (3.6)
$$

$$
\Sigma_{t+1} = \Sigma_{t+1}^T
\quad (3.7)
$$

for each $T = 0, 1, \ldots$ and $t = 0, 1, \ldots, T$.

**Proof.** Fix $T = 0, 1, \ldots$ and $t = 0, 1, \ldots, T$, and note the inclusions

$$
\mathcal{Y}_t \subset \mathcal{F}_{t+1} \subset \mathcal{F}_{T+1}.
\quad (3.8)
$$

By property (F.2), we directly see that

$$
\begin{pmatrix}
\bar{X}_{t+1}^T \\
\bar{B}_{t+1}^T
\end{pmatrix}
:= \bar{E}_{T+1} \left[ \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} \mid \mathcal{Y}_t \right]
\quad (3.9)
$$

which proves (3.6). In view of (3.8) and the dynamical representation (3.1), it follows that $(X_{t+1}, B_{t+1})' - (\bar{X}_{t+1}^T, \bar{B}_{t+1}^T)'$ is $\mathcal{F}_{t+1}$-measurable, so by property (F.1), we conclude that

$$
\Sigma_{t+1} := \bar{E}_{T+1} \left[ \left( \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} - \begin{pmatrix}
\bar{X}_{t+1}^T \\
\bar{B}_{t+1}^T
\end{pmatrix}
\right) \left( \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} - \begin{pmatrix}
\bar{X}_{t+1}^T \\
\bar{B}_{t+1}^T
\end{pmatrix}
\right)' \right]
= \bar{E}_{t+1} \left[ \left( \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} - \begin{pmatrix}
\bar{X}_{t+1}^T \\
\bar{B}_{t+1}^T
\end{pmatrix}
\right) \left( \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} - \begin{pmatrix}
\bar{X}_{t+1}^T \\
\bar{B}_{t+1}^T
\end{pmatrix}
\right)' \right],
\quad (3.10)
$$

$$
= \Sigma_{t+1}
$$

30
which proves (3.7).

If we now turn to Appendix A to find expressions for \( \{\tilde{X}_{t+1}, \tilde{B}_{t+1}\}_0^\infty \) and \( \{\Sigma_{t+1}\}_0^\infty \) for some fixed \( T = 0, 1, \ldots \), we note the well-known fact that the Kalman filter is recursive; that we have the representation

\[
\Sigma_{t+1} = f(\Sigma_t; A_t, \Phi(t, 0), H_t, \Gamma_{t+1}) \\
\Sigma_0 = 0 \\
(X^T_{t+1}, B^T_{t+1}) = g((X^T_t, B^T_t), Y_t; \Sigma_t; A_t, \Phi(t, 0), H_t, \Gamma_{t+1}) \\
(\tilde{X}^T_0, \tilde{B}^T_0) = (0, 0)
\]

for some appropriate mappings \( f \) and \( g \) independent of our choice of \( T \). Because of this independence and (3.6)-(3.7), we see that

\[
\Sigma_{t+1} = f(\Sigma_t; A_t, \Phi(t, 0), H_t, \Gamma_{t+1}) \\
\Sigma_0 = 0 \\
(X_{t+1}, B_{t+1}) = g((X_t, B_t), Y_t; \Sigma_t; A_t, \Phi(t, 0), H_t, \Gamma_{t+1}) \\
(X_0, B_0) = (0, 0).
\]

Reviewing Proposition 3.1 and recalling the properties of the system (3.1) under the sequence of probability measures \( \{\tilde{P}_t\}_0^\infty \), we see that essentially, for the purposes of calculating \( \{X_{t+1}, B_{t+1}\}_0^\infty \) and \( \{\Sigma_{t+1}\}_0^\infty \), we may assume that there exists a probability measure \( Q \) on \( (\Omega, \mathcal{F}) \) under which the entire sequence \( \{(W^0_{t+1}, V^0_{t+1})\}_0^\infty \) is a GWN process with the same statistics as \( \{(W^0_{t+1}, V^0_{t+1})\}_0^\infty \) has under \( P \). Indeed, from Chapter II we can verify that such a probability measure \( Q \) exists, but that it need not satisfy property (E.1); moreover, from Theorem II.2.3, we have that if \( M_\infty := \lim_t M_t \), then \( Q \) will enjoy property (E.1) if \( E[\xi' M_\infty \xi] < \infty \). Observe from (2.11) that if the system is time-invariant, then we have the interesting fact that \( E[\xi' M_\infty \xi] < \infty \) if the matrix \( A \) is stable. Thus in the time-invariant case, the stability of \( A \) implies the existence of an infinite-horizon Girsanov measure transformation. We shall make no further comments on this point, but simply summarize the relevant Kalman filtering calculations under the probability measure \( Q \) in the following theorem.

**Theorem 3.1.** Let the deterministic sequences \( \{P_t\}_0^\infty \), \( \{Q_t\}_0^\infty \), and \( \{R_t\}_0^\infty \) in \( \mathcal{Q}_n, \mathcal{M}_n, \) and \( \mathcal{Q}_n \) (respectively) be such that

\[
\Sigma_{t+1} = \begin{pmatrix} P_{t+1} & Q_{t+1} \\ Q'_{t+1} & R_{t+1} \end{pmatrix}, \quad t = 0, 1, \ldots \quad (3.13)
\]
Then \( \{P_t\}_0^\infty \) propagates according to

\[
P_{t+1} = A_t P_t A_t' - A_t P_t H_t' [H_t P_t H_t' + I_k]^{-1} H_t P_t A_t' + \Sigma_{t+1}^w \\
R_0 = 0.
\]

Define an auxiliary deterministic sequence \( \{J_t\}_0^\infty \) taking values in \( Q_k \) by

\[
J_t := H_t P_t H_t' + I_k.
\]

Then \( \{Q_t\}_0^\infty \) propagates according to

\[
Q_{t+1} = A_t Q_t - A_t P_t H_t' J_t^{-1} H_t (Q_t + \Phi(t, 0)) \\
Q_0 = 0
\]

and \( \{R_t\}_0^\infty \) propagates according to

\[
R_{t+1} = R_t - (Q_t + \Phi(t, 0))' H_t' J_t^{-1} H_t (Q_t + \Phi(t, 0)) + \Phi'(t, 0) H_t' H_t \Phi(t, 0) \\
R_0 = 0.
\]

Furthermore, \( \{X_t\}_0^\infty \) and \( \{B_t\}_0^\infty \) evolve according to

\[
X_{t+1} = A_t [I_n - P_t H_t' J_t^{-1} H_t] X_t + A_t P_t H_t' J_t^{-1} Y_t \\
X_0 = 0
\]

and

\[
B_{t+1} = B_t - (Q_t + \Phi(t, 0))' H_t' J_t^{-1} H_t X_t + (Q_t + \Phi(t, 0))' H_t' J_t^{-1} Y_t \\
B_0 = 0.
\]

In view of (3.5), we have in fact solved the prediction problem in the uncorrelated case. We state the result as a theorem (see also [21, Theorem 1]).

**Theorem 3.2.** Let the processes \( \{(X_{t+1}, B_{t+1})\}_0^\infty \) and \( \{\Sigma_{t+1}\}_0^\infty \) be as in Theorem 3.1. For any \( \phi \) in \( \mathcal{Z} \) and any \( t = 0, 1, \ldots \), the representation

\[
E[\phi(X_{t+1}) | Y_t] = \frac{U \phi[X_{t+1}; B_{t+1}; M_{t+1}, \Phi(t + 1, 0); \Sigma_{t+1}]}{U [X_{t+1}; B_{t+1}; M_{t+1}, \Phi(t + 1, 0); \Sigma_{t+1}]} \tag{3.20}
\]

32
holds.

We can simplify (3.16)-(3.19) by introducing the auxiliary quantities

\[ Q^*_t := Q_t + \Phi(t, 0) \]
\[ R^*_t := M_t - R_t. \]

We then have the following representation.

**Proposition 3.2.** Define deterministic sequences \( \{Q^*_t\}^\infty_0 \) in \( M_n \) and \( \{R^*_t\}^\infty_0 \) in \( Q_n \) by the dynamical equations

\[ Q^*_{t+1} = A_t [I_n - P_t H_t^t J_t^{-1} H_t] Q^*_t \quad t = 0, 1, \ldots \quad (3.22) \]
\[ Q^*_0 = I_n \]

and

\[ R^*_{t+1} = R^*_t + Q^*_t H_t^t J_t^{-1} H_t Q^*_t \quad t = 0, 1, \ldots \quad (3.23) \]

\[ R^*_0 = 0. \]

Then

\[ \Sigma_{t+1} = \begin{pmatrix} P_{t+1} & Q^*_{t+1} - \Phi(t + 1, 0) \\ Q^*_t - \Phi'(t + 1, 0) & M_{t+1} - R^*_{t+1} \end{pmatrix}, \quad t = 0, 1, \ldots \quad (3.24) \]

and we may rewrite (3.19) as

\[ B_{t+1} = B_t - Q^*_t H_t^t J_t^{-1} H_t X_t + Q^*_t H_t^t J_t^{-1} Y_t \]
\[ B_0 = 0. \]

The usefulness of the representation of Proposition 3.2 will become evident in Chapter V in the consideration of the conditional mean.
III.4. Discussion—The Methodology for the Uncorrelated Case

In order to motivate the arguments used in Chapter IV, let us now consider the methods used in Sections 2 and 3 to solve the prediction problem.

We recall that we first fixed a \( \phi \) in \( Z \) and studied the process \( \{ E[\phi(X_{t+1}^0)|Y_t]\}^\infty_0 \). In Section 2, we attempted to solve the prediction problem over a finite horizon; we found \( \{ E[\phi(X_{t+1}^0)|Y_t]\}^T_0 \) for a fixed but arbitrary \( T = 0,1,\ldots \). The extension to an infinite horizon occurred in Section 3 and was relatively straightforward given the machinery we developed in Section 2. Fix \( T = 0,1,\ldots \), and consider the means we used to find \( \{ E[\phi(X_{t+1}^0)|Y_t]\}^T_0 \). The procedures of Section 2 fundamentally consisted of two steps (G.1) and (G.2), where

(G.1): a decomposition of the \( \{ X_t^0 \}^{T+1}_0 \) of the form

\[
X_t^0 = X_t + Z_t \quad t = 0,1,\ldots,T+1
\]  

where \( \{ X_t^0 \}^{T+1}_0 \) represented the effects of the noise process \( \{ W_t^0 \}^T_0 \) and \( \{ Z_t \}^{T+1}_0 \) represented the effects of the initial condition \( \xi \).

The most natural decomposition of this form is described by (2.1) and (2.2). Once we had this decomposition, the representation (2.6) followed from the definition of \( \{ V_{t+1} \}^T_0 \) given by (2.5). Note that if \( \{(W_{t+1}^0,V_{t+1})\}^T_0 \) were a GWN sequence under \( P \), the prediction problem for the system ((2.1), (2.6)) would be solved by Kalman filtering arguments. This observation led us to seek

(G.2): a probability measure \( \tilde{P} \) on \( (\Omega,\mathcal{F}) \), mutually absolutely continuous with \( P \) and agreeing with \( P \) on \( \sigma(\{\xi\}) \), such that under \( \tilde{P} \), \( \{(W_{t+1}^0,V_{t+1})\}^T_0 \) is a GWN process independent of the RV \( \xi \).

The filtering of the system ((2.1), (2.6)) under \( \tilde{P} \) could be accomplished simply under the assumption that \( \{(W_{t+1}^0,V_{t+1})\}^T_0 \) is a GWN sequence under \( \tilde{P} \), but the mutual absolute continuity of \( P \) and \( \tilde{P} \) and the requirements on the statistical nature of \( \xi \) ensured that there would be a correspondence between the solution of the original prediction problem (under \( P \)) and the solution of the prediction problem for ((2.1), (2.6)) under \( \tilde{P} \). This correspondence was based upon (2.15);

\[
E[\phi(X_{t+1}^0)|Y_t] = \frac{\tilde{E}[\phi(X_{t+1}^0)\frac{d\tilde{P}}{dP}|Y_t]}{\tilde{E}[\frac{d\tilde{P}}{dP}|Y_t]} \quad t = 0,1,\ldots,T
\]  

34
While the definition of $\hat{P}$ was naturally suggested by the Girsanov measure transformation outlined in Chapter II, we in fact needed $\frac{dP}{d\hat{P}}$ to have the special structure

$$
\frac{dP}{d\hat{P}} = C(\gamma, \xi)
$$

(4.3)

where $\gamma$ is an $IR^l$-valued RV which, under $\hat{P}$, is jointly Gaussian with $((W_{t+1}^0, V_{t+1}))^T_0$ and independent of $\xi$, and where $C : IR^l \times IR^n \rightarrow IR$ is some Borelian function; here $l$ is a fixed but unspecified positive integer. Given this form of $\frac{dP}{d\hat{P}}$, we were then able to arrive at representation (2.27) using several properties of conditional expectations. In a manner analogous to that of Section 2, we define

$$
\left( \begin{array}{c}
\tilde{X}_{t+1} \\
\gamma_t
\end{array} \right) = \mathbb{E} \left[ \left( \begin{array}{c}
X_{t+1} \\
\gamma_t
\end{array} \right) \mid \mathcal{Y}_t \right] 
$$

(4.4)

t = 0, 1, \ldots

and introduce the mappings $T_\phi : IR^n \times IR^l \times IR^n \times IR \rightarrow IR$ and $U_\phi : IR^n \times IR^l \times IR \rightarrow IR$ by

$$
T_\phi[u, v, w; t] := \mathbb{E}[\phi(X_{t+1}^0 + u)C(\tilde{\gamma}_t + v, w)]
$$

$$
U_\phi[x, u; t] := \mathbb{E}[T_\phi[x + \Phi(t + 1, 0)\xi, u, \xi; t]].
$$

(4.5)

With this notation, and by using exactly the same reasoning used to prove (2.24) from (2.21), we have

$$
\tilde{\mathbb{E}} \left[ \phi(X_{t+1}^0) \frac{dP}{d\hat{P}} \mid \mathcal{Y}_t \uplus \sigma\{\xi\} \right] 
$$

$$
= \tilde{\mathbb{E}} \left[ \phi(\tilde{X}_{t+1} + (\tilde{X}_{t+1} + Z_{t+1}))C(\tilde{\gamma}_t + \tilde{\gamma}_t, \xi) \mid \mathcal{Y}_t \uplus \sigma\{\xi\} \right] t = 0, 1, \ldots, T
$$

(4.6)

$$
= T_\phi[\tilde{X}_{t+1} + Z_{t+1}, \tilde{\gamma}_t, \xi; t],
$$

so that

$$
\tilde{\mathbb{E}} \left[ \phi(X_{t+1}^0) \frac{dP}{d\hat{P}} \mid \mathcal{Y}_t \right] = \tilde{\mathbb{E}}[T_\phi[\tilde{X}_{t+1} + \Phi(t + 1, 0)\xi, \tilde{\gamma}_t, \xi] \mid \mathcal{Y}_t] 
$$

(4.7)

$$
= U_\phi[\tilde{X}_{t+1}, \tilde{\gamma}_t; t],
$$

35
and upon invoking (4.2), we conclude that

$$\mathbb{E}[\phi(X_{t+1}^n)|\mathcal{Y}_t] = \frac{U_\phi[X_{t+1}, \tilde{\gamma}_t; t]}{U_1[X_{t+1}, \tilde{\gamma}_t; t]}, \quad t = 0, 1, \ldots, T \quad (4.8)$$

With these remarks as guidelines, the reasonings of Chapter IV will proceed more smoothly.
CHAPTER IV: THE CORRELATED PROBLEM

IV.1. Overview of the Correlated Problem

Guided by Chapter III, we now consider the situation where \(((W_{t+1}^0, V_{t+1}^0))_0^\infty\) is a general \(IR^{n+k}\)-valued zero-mean GWN process;

\[
\Gamma_{t+1} = \begin{pmatrix}
\Sigma_{t+1} & \Sigma_{t+1}^{uv} \\
\Sigma_{t+1}^{vu} & \Sigma_{t+1}
\end{pmatrix} \quad t = 0, 1, \ldots \quad (1.1)
\]

with no restrictions on \(\{\Gamma_{t+1}\}_0^\infty\) other than that the covariance matrix \(\Sigma_{t+1}\) be positive-definite for each \(t = 0, 1, \ldots\). As in Chapter III, we define \(\mathcal{Y}_t\) as the filtration of \(\mathcal{F}\) generated by the observations \(\{Y_t\}_0^\infty\);

\[
\mathcal{Y}_t := \sigma\{Y_s; s = 0, 1, \ldots, t\} \quad t = 0, 1, \ldots \quad (1.2)
\]

Our arguments will proceed as outlined in Section III.4. We first fix an element \(\phi\) in \(\mathcal{S}\) and a \(T = 0, 1, \ldots\) and study the process \(\{E[\phi(X_{t+1}^0)|\mathcal{Y}_t]\}_0^T\). In Section 2, we find an appropriate decomposition of \(\{X_t^0\}_0^{T+1}\) and identify the appropriate Girsanov measure transformation. Section 3 is devoted to finding a representation for \(\{E[\phi(X_{t+1}^0)|\mathcal{Y}_t]\}_0^T\) using this measure transformation. We then characterize the infinite-horizon process \(\{E[\phi(X_{t+1}^0)|\mathcal{Y}_t]\}_0^\infty\) in Section 4 by extending the representation of Section 3; a complete solution to the prediction problem in the correlated case is then at hand.
IV.2. A Decomposition and a Change of Measure

With the discussion of Section III.4 in mind, we first consider a general decomposition of \(\{X_t^\circ\}_0^{T+1}\) of the form

\[
X_t^\circ = X_t + Z_t \quad t = 0, 1, \ldots T + 1
\]  

(2.1)

where \(\{X_t\}_0^{T+1}\) represents the effects of the noise process \(\{W_{t+1}\}_0^T\) and \(\{Z_t\}_0^{T+1}\) represents the effects of the initial condition \(\xi\). Note that we do not as yet specify the evolution of either \(\{X_t\}_0^{T+1}\) or \(\{Z_t\}_0^{T+1}\), as was done by (III.2.1) and (III.2.2). We may define an \(IR^k\)-valued ‘noise’ process \(\{V_{t+1}\}_0^T\) by the relation

\[
V_{t+1} := V_{t+1}^\circ + H_t Z_t, \quad t = 0, 1, \ldots
\]  

(2.2)

so that the observation process \(\{Y_t\}_0^T\) may be represented by

\[
Y_t = H_t X_t + V_{t+1}. \quad t = 0, 1, \ldots
\]  

(2.3)

We next consider the discussion of Section III.4 concerning the change of measure. A close inspection of Section III.2 reveals that the independence of \(\{W_{t+1}\}_0^T\) and \(\{V_{t+1}\}_0^T\) was crucial to the measure transformation used there; because of it, we could find a probability measure \(\bar{P}\) on \((\Omega, \mathcal{F})\) under which property (B.4) held. Since \(W_{t+1}^\circ\) and \(V_{t+1}^\circ\) are no longer independent (i.e., uncorrelated) under \(P\) for \(t = 0, 1, \ldots\), we cannot hope to use the Girsanov transformation to find a probability measure \(\bar{P}\) on \((\Omega, \mathcal{F})\) enjoying properties (B.1)-(B.2) and under which \(\{(W_{t+1}, V_{t+1})_0^T\}\) is a GWN sequence. To make this more clear, note that in Section III.2, we defined the filtration \(\{\mathcal{F}_t\}_0^\infty\) of \(\mathcal{F}\) to ensure that \(\{V_{t+1}\}_0^\infty\) was a \((\mathcal{F}_t, P)\)-zero mean GWN sequence. In particular, \(V_{t+1}^\circ\) was independent of \(\mathcal{F}_0\) for each \(t = 0, 1, \ldots\). If we here define, in a manner similar to (III.2.7),

\[
\mathcal{F}_0 := \sigma\{\xi, W_{s+1}^\circ; s = 0, 1, \ldots\}
\]

\[
\mathcal{F}_{t+1} := \mathcal{F}_0 \lor \sigma\{V_{s+1}^\circ; s = 0, 1, \ldots, t\}, \quad t = 0, 1, \ldots
\]  

(2.4)

we see that \(\{V_{t+1}^\circ\}_0^T\) is not necessarily a \((\mathcal{F}_t, P)\)-zero mean GWN sequence, since the sequences \(\{W_{t+1}\}_0^T\) and \(\{V_{t+1}\}_0^T\) are not necessarily independent. Hence the Girsanov theory does not apply. This difficulty is easily overcome, however, if we consider a Girsanov transformation on the joint \(IR^{n+k}\)-valued process \(\{(W_{t+1}, V_{t+1})_0^T\}\).
In order to perform a Girsanov transformation on the joint process \( \{(W_{t+1}^0, V_{t+1}^0)\}_0^T \), we define the filtration \( \mathcal{F}_t \) by

\[
\mathcal{F}_t := \sigma\{\xi, W_{s+1}^0; s = 0, 1, \ldots\} \quad t = 0, 1, \ldots \quad (2.5)
\]

\[
\mathcal{F}_{t+1} := \mathcal{F}_t \lor \sigma\{V_{s+1}^0; s = 0, 1, \ldots, t\}
\]

instead of (2.4). Turning now to Chapter II, we see that if we define the \( \IR^{n+k} \)-valued process \( \{(W_{t+1}, V_{t+1})\}_0^T \) according to

\[
\begin{pmatrix}
W_{t+1} \\
V_{t+1}
\end{pmatrix}
= \begin{pmatrix}
W_{t+1}^0 \\
V_{t+1}^0
\end{pmatrix} - \begin{pmatrix}
\Sigma_{t+1}^{uu} & \Sigma_{t+1}^{uv} \\
\Sigma_{t+1}^{vu} & \Sigma_{t+1}^{vv}
\end{pmatrix}
\begin{pmatrix}
\varphi_t^w \\
\varphi_t^v
\end{pmatrix}
\quad t = 0, 1, \ldots, T \quad (2.6)
\]

where \( \{\varphi_t^w\}_0^T \) and \( \{\varphi_t^v\}_0^T \) are any two \( (\mathcal{F}_t) \)-adapted processes taking values in \( \IR^n \) and \( \IR^k \) (respectively), we can find a probability measure \( \bar{P} \) on \( (\Omega, \mathcal{F}) \) with the three properties:

(H.1): \( P \) and \( \bar{P} \) are mutually absolutely continuous,

(H.2): \( P \) and \( \bar{P}_{t+1} \) agree on \( \mathcal{F}_0 \) and

(H.3): the process \( \{(W_{t+1}, V_{t+1})\}_0^T \) is a \( (\mathcal{F}_t, \bar{P}) \) zero-mean GWN process with the same statistics under \( \bar{P} \) as \( \{(W_{t+1}^0, V_{t+1}^0)\}_0^T \) have under \( P \).

In a manner similar to that of Section III.2, note that if a probability measure \( \bar{P} \) on \( (\Omega, \mathcal{F}) \) enjoys properties (H.1)-(H.3), we may summarize the statistical nature of the RV's \( \xi \), \( \{W_{t+1}\}_0^T \) and \( \{V_{t+1}\}_0^T \) under \( \bar{P} \) by

(H.4): the RV's \( \xi \), \( \{W_{t+1}\}_0^T \) and \( \{V_{t+1}\}_0^T \) have the same joint statistics under \( \bar{P} \) as the RV's \( \xi \), \( \{W_{t+1}^0\}_0^T \) and \( \{V_{t+1}^0\}_0^T \) have under \( P \).

Combining (2.2) and (2.6), we note that in fact we need consider only those noise processes \( \{(W_{t+1}, V_{t+1})\}_0^T \) which are given by (2.6) with

\[
\begin{align*}
\varphi_t^w &= \varphi_t \\
\varphi_t^v &= - (\Sigma_{t+1}^v)^{-1} [\Sigma_{t+1}^{uv} \varphi_t + H_t Z_t]
\end{align*}
\quad t = 0, 1, \ldots, T \quad (2.7)
\]

where \( \{\varphi_t\}_0^T \) is an \( (\mathcal{F}_t) \)-adapted \( \IR^n \)-valued process. We shall henceforth assume that the processes \( \{(W_{t+1}, V_{t+1})\}_0^T \) and \( \{\varphi_t\}_0^T \) are related by (2.6) and (2.7). Consequently, the process \( \{W_{t+1}\}_0^T \) can be represented by

\[
W_{t+1} = W_{t+1}^0 + \Sigma_{t+1}^{uv} (\Sigma_{t+1}^u)^{-1} H_t Z_t \\
- [\Sigma_{t+1}^u - \Sigma_{t+1}^{uv} (\Sigma_{t+1}^u)^{-1} \Sigma_{t+1}^{vv}] \varphi_t.
\quad t = 0, 1, \ldots, T \quad (2.8)
\]
For completeness, we note that the probability measure with properties (E.1)-(E.4) suggested by the Girsanov theory is given by the Radon-Nikodym derivative

\[
\frac{d\tilde{P}}{dP} = \exp \left[ \sum_{s=0}^{T} \left[ \varphi_{s}^\prime [W_{s+1} - \Sigma_{s+1}^{uv} (\Sigma_{s+1}^{u})^{-1} V_{s+1}] - Z_{s} H_{s}^\prime (\Sigma_{s+1}^{u})^{-1} V_{s+1} \right] + \frac{1}{2} \sum_{s=0}^{T} \left[ \varphi_{s}^\prime [\Sigma_{s+1}^{u} - \Sigma_{s+1}^{uw} (\Sigma_{s+1}^{u})^{-1} \Sigma_{s+1}^{uw}] \varphi_{s} + Z_{s} H_{s}^\prime (\Sigma_{s+1}^{u})^{-1} H_{s} Z_{s} \right] \right].
\]  
(2.9)

To complete the arguments, we still must specify the dynamics of \( \{X_{t}\}_{0}^{T+1} \) and \( \{Z_{t}\}_{0}^{T+1} \) and select an \((\mathcal{F}_{t})\)-adapted process \( \{\varphi_{t}\}_{0}^{T} \). The decomposition equation (2.1) and the dynamics of \( \{X_{t}^{0}\}_{0}^{\infty} \) yield, via (2.8), the equalities

\[
X_{t+1} + Z_{t+1} = X_{t+1}^{0} = A_{t} X_{t}^{0} + W_{t+1}^{0} = A_{t} (X_{t} + Z_{t}) + W_{t+1} - \Sigma_{t+1}^{uw} (\Sigma_{t+1}^{u})^{-1} H_{t} Z_{t} + [\Sigma_{t+1}^{u} - \Sigma_{t+1}^{uw} (\Sigma_{t+1}^{u})^{-1} \Sigma_{t+1}^{uw}] \varphi_{t} + W_{t+1}
\]
\[
t = 0, 1, \ldots, T
\]  
(2.10)

Here we have represented \( \{W_{t+1}^{0}\}_{0}^{T} \) in terms of \( \{W_{t+1}\}_{0}^{T} \) since, as in Chapter III, we are concerned with the solution of the prediction problem under \( \tilde{P} \), and the process \( \{W_{t+1}^{0}\}_{0}^{T} \) is a not a GWN sequence under \( \tilde{P} \), while \( \{W_{t+1}\}_{0}^{T} \) is. Eq. (2.10) suggests a separation of the dynamics of \( \{X_{t+1}\}_{0}^{T} \) and \( \{Z_{t+1}\}_{0}^{T} \) of the form

\[
X_{t+1} = A_{t} X_{t} + W_{t+1} + [\Sigma_{t+1}^{uw} (\Sigma_{t+1}^{u})^{-1} \Sigma_{t+1}^{uw}] \varphi_{t} - \pi_{t}
\]
\[
t = 0, 1, \ldots, T
\]  
(2.11)

\[
X_{0} = \zeta
\]

and

\[
Z_{t+1} = [A_{t} - \Sigma_{t+1}^{uw} (\Sigma_{t+1}^{u})^{-1} H_{t}] Z_{t} + \pi_{t}
\]
\[
t = 0, 1, \ldots, T
\]  
(2.12)

where \( \zeta \) and \( \{\pi_{t}\}_{0}^{T} \) are unspecified RV’s taking values in \( IR^{n} \). We shall make no attempt to further analyze the different possible choices for \( \zeta, \{\pi_{t}\}_{0}^{T} \), and \( \{\varphi_{t}\}_{0}^{T} \) but shall simply

40
assume that
\[ \varphi_t = 0 \]
\[ \pi_t = 0 \quad t = 0, 1, \ldots, T \] (2.13)
\[ \zeta = 0. \]

With these assumptions, we can quantify the various relevant processes needed to solve the finite-horizon prediction problem. We summarize these quantities as follows:

- The Effect of the Initial Condition:
\[
Z_{t+1} = [A_t - \Sigma_{t+1}^{uv}(\Sigma_{t+1}^u)^{-1}H_t]Z_t \qquad t = 0, 1, \ldots, T \] (2.14)
\[ Z_0 = \xi. \]

- The Noise Processes:
\[
\begin{pmatrix} W_{t+1}^o \\ V_{t+1}^o \end{pmatrix} = \begin{pmatrix} W_{t+1}^o \\ V_{t+1}^o \end{pmatrix} - \begin{pmatrix} \Sigma_{t+1}^{uv} \\ \Sigma_{t+1}^{uw} \\ \Sigma_{t+1}^{uv} \\ \Sigma_{t+1}^{uw} \end{pmatrix} \begin{pmatrix} 0 \\ -(\Sigma_{t+1}^u)^{-1}H_tZ_t \end{pmatrix} \quad t = 0, 1, \ldots, T \] (2.15)
\[ = \begin{pmatrix} W_{t+1}^o + \Sigma_{t+1}^{uv}(\Sigma_{t+1}^u)^{-1}H_tZ_t \\ V_{t+1}^o + H_tZ_t \end{pmatrix}. \]

- The Auxiliary System:
\[
X_{t+1} = A_tX_t + W_{t+1} \]
\[ X_0 = 0 \quad t = 0, 1, \ldots, T \] (2.16)
\[ Y_t = H_tX_t + V_{t+1}. \]

- The Change of Measure:
\[
\frac{d\tilde{P}}{dP} = \exp \left[ -\sum_{s=0}^{T} Z_s^t H_s^t(\Sigma_{s+1}^u)^{-1}V_{s+1} + \frac{1}{2} \sum_{s=0}^{T} Z_s^t H_s^t(\Sigma_{s+1}^u)^{-1}H_sZ_s \right]. \] (2.17)

We then have the following result.

**Proposition 2.1.** Let the processes \( \{X_t\}_{0}^{T+1}, \{Z_t\}_{0}^{T+1} \) and \( \{(W_{t+1}, V_{t+1})\}_{0}^{T} \) be defined by (2.14)-(2.16) and the probability measure \( \tilde{P} \) be defined by (2.17). Then \( \tilde{P} \) enjoys properties (H.1)-(H.4).
IV.3. A Representation for the Finite-Horizon Predictor

We here use the decomposition and measure transformation introduced in Section 2 to find a representation for $\{E[\phi(X_{t_i+1})|\mathcal{Y}_t]\}_0^T$, where $T = 0, 1, \ldots$ is fixed but arbitrary, as in Section 2. We first fix some notation and compare what has been discovered so far in the correlated case to the analogous results in Chapter III. The arguments in this section will run parallel to those of Section III.2.

Let the filtration $\{\mathcal{F}_t\}_0^\infty$ and the processes $\{X_t\}_0^\infty$, $\{Z_t\}_0^\infty$, and $\{(W_{i+1}, V_{i+1})\}_0^\infty$ be given by (2.5) and (2.14)-(2.16) with an infinite horizon. We point out that the process $\{Z_t\}_0^\infty$ now is given by

$$Z_t = \Psi(t, 0)\xi, \quad t = 0, 1, \ldots \quad (3.1)$$

in contrast to (II.2.3), where $\Psi(\cdot, \cdot)$ was defined in Section I.3. Next, define the $(\mathcal{F}_t)$-adapted process $\{L_t\}_0^\infty$ by

$$L_t := \exp\left[ -\sum_{s=0}^{T} Z'_s H'_s (\Sigma_{s+1}^v)^{-1} V_{s+1} + \frac{1}{2} \sum_{s=0}^{T} Z'_s H'_s (\Sigma_{s+1}^v)^{-1} H_s Z_s \right] \quad t = 0, 1, \ldots \quad (3.2)$$

$$L_0 := 1,$$

and define the probability measure $\tilde{P}_{T+1}$ on $(\Omega, \mathcal{F})$ by the Radon-Nikodym derivative

$$\frac{d\tilde{P}_{T+1}}{dP} = L_{T+1}. \quad (3.3)$$

As in Section III.2, the representation of $\{L_t\}_0^\infty$ given by (3.2) is relatively cumbersome. To simplify it, define an $IR^n$-valued process $\{B_t\}_0^\infty$ by the recursion

$$B_{t+1} = B_t + \Psi'(t, 0) H_t' (\Sigma_{t+1}^v)^{-1} V_{t+1} \quad t = 0, 1, \ldots \quad (3.4)$$

$$B_0 = 0$$

and a deterministic sequence $\{M_t\}_0^\infty$ in $Q_n$ according to

$$M_{t+1} = M_t + \Psi'(t, 0) H_t' (\Sigma_{t+1}^v)^{-1} H_t \Psi(t, 0) \quad t = 0, 1, \ldots \quad (3.5)$$

$$M_0 = 0.$$

It is then a simple matter to check that

$$L_t = \exp[ -\xi'B_t + \frac{1}{2} \xi'M_t\xi]. \quad t = 0, 1, \ldots \quad (3.6)$$

42
Note that if \( W_{t+1}^0 \) and \( V_{t+1}^0 \) are uncorrelated for each \( t = 0, 1, \ldots \) and \( \{V_{t+1}^0\}_0^\infty \) is standard (i.e., \( \Sigma_{t+1}^{vv} = 0 \) and \( \Sigma_{t+1}^{v} = I_k \) for \( t = 0, 1, \ldots \)), then \( \Phi(\cdot, \cdot) \) and \( \Psi(\cdot, \cdot) \) are equivalent, and (3.1), (3.2), (3.4), and (3.5) reduce to (III.2.3), (III.2.8), (III.2.10), and (III.2.11), respectively. This specialization of the correlated case to the uncorrelated case was expected, and we shall see more of it.

As in Section III.2, for any bounded \( \mathcal{F}_{T+1} \)-measurable \( C \)-valued RV \( \rho \), set

\[
\sigma_t[\rho] := E_{T+1}[\rho L_{T+1}^{-1}\{Y_t\}], \quad t = 0, 1, \ldots \quad (3.7)
\]

so that

\[
E[\rho | Y_t] = \frac{\sigma_t[\rho]}{\sigma_t[1]} \quad \text{P.a.s.} \quad t = 0, 1, \ldots \quad (3.8)
\]

The following result should not then be too surprising.

**Lemma 3.1.** Under \( \bar{P}_{T+1} \), the processes \( \{(X_{t+1}, B_{t+1})\}_{0}^{T} \) and \( \{Y_t\}_{0}^{T} \) are jointly Gaussian and independent of the RV \( \xi \).

Now define the conditional means and variances under \( \bar{P}_{T+1} \) by

\[
\begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} := E_{T+1} \left[ \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} | Y_t \right], \quad t = 0, 1, \ldots, T \quad (3.9)
\]

\[
\begin{pmatrix}
\tilde{X}_{t+1} \\
\tilde{B}_{t+1}
\end{pmatrix} := \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} - \begin{pmatrix}
\tilde{X}_{t+1} \\
\tilde{B}_{t+1}
\end{pmatrix}, \quad t = 0, 1, \ldots, T \quad (3.10)
\]

\[
\Sigma_{t+1} := E_{T+1} \left[ \begin{pmatrix}
\tilde{X}_{t+1} \\
\tilde{B}_{t+1}
\end{pmatrix} \begin{pmatrix}
\tilde{X}_{t+1} \\
\tilde{B}_{t+1}
\end{pmatrix}^{'} \right], \quad t = 0, 1, \ldots, T \quad (3.11)
\]

For \( t = 0, 1, \ldots \), we observe that

\[
E_{T+1} \phi(X_{t+1}^0) L_{T+1}^{-1} \{Y_t \vee \sigma(\xi)\} = E_{T+1} \left[ \phi(\tilde{X}_{t+1} + (\tilde{X}_{t+1} + Z_{t+1})) \exp[\xi^{'} \tilde{B}_{t+1}] \right] | Y_t \vee \sigma(\xi) \exp[\xi^{'} \tilde{B}_{t+1} - \frac{1}{2} \xi^{'} M_{t+1} \xi] (3.12)
\]

\[
eq E_{T+1} \left[ \phi(\tilde{X}_{t+1} + u) \exp[v^{'} \tilde{B}_{t+1}] \right] \bigg|_{u = X_{t+1} + Z_{t+1}} \exp[\xi^{'} \tilde{B}_{t+1} - \frac{1}{2} \xi^{'} M_{t+1} \xi] (3.13)
\]

\[
= T \phi(\tilde{X}_{t+1} + Z_{t+1}, \xi; \Sigma_{t+1}) \exp[\xi^{'} \tilde{B}_{t+1} - \frac{1}{2} \xi^{'} M_{t+1} \xi] (3.14)
\]

and that

\[
\sigma_t[\phi(X_{t+1}^0)] = E_{T+1} \left[ T \phi(\tilde{X}_{t+1} + \Psi(t + 1, 0) \xi; \Sigma_{t+1}) \exp[\xi^{'} \tilde{B}_{t+1} - \frac{1}{2} \xi^{'} M_{t+1} \xi] \right] | Y_t \right) (3.15)
\]

\[
eq E_{T+1} \left[ T \phi(x + \Psi(t + 1, 0) \xi; \Sigma_{t+1}) \exp[\xi^{'} b - \frac{1}{2} \xi^{'} M_{t+1} \xi] \right] \bigg|_{b = \tilde{B}_{t+1}, x = \tilde{X}_{t+1}} (3.16)
\]

\[
= U \phi(\tilde{X}_{t+1}, \tilde{B}_{t+1}, M_{t+1}, \Psi(t + 1, 0); \Sigma_{t+1}), \quad (3.17)
\]
in exactly the same manner that we proved (III.2.27).

We thus have the following characterization for the solution of the finite horizon problem, which we may compare with Theorem III.2.1.

**Theorem 3.1.** Consider the sequences \( \{X_t\}_{0}^{\infty}, \{Z_t\}_{0}^{\infty}, \{(W_{t+1}, V_{t+1})\}_{0}^{\infty}, \{L_t\}_{0}^{\infty}, \{B_t\}_{0}^{\infty} \) and \( \{M_t\}_{0}^{\infty} \) as defined in this section and the probability measure \( \tilde{P}_{T+1} \) on \( (\Omega, \mathcal{F}) \) as defined by (3.3). Then \( \{L_t\}_{0}^{\infty} \) is an \((\mathcal{F}_t, P)\)-martingale, and \( \tilde{P}_{T+1} \) enjoys properties (H.1)-(H.4). Furthermore, Lemma 3.1 holds. Also, with \( \{(X_{t+1}, \tilde{B}_{t+1})\}_{0}^{T} \) and \( \{\Sigma_{t+1}\}_{0}^{T} \) as defined in this section, the representation

\[
E[\phi(X_{t+1}^\circ)|\mathcal{Y}_t] = \frac{U[\phi(X_{t+1}, \tilde{B}_{t+1}; M_{t+1}, \Psi(t + 1, 0); \Sigma_{t+1}]}{U[\Psi(X_{t+1}, \tilde{B}_{t+1}; M_{t+1}, \Psi(t + 1, 0); \Sigma_{t+1}]} \quad t = 0, 1, \ldots, T \tag{3.18}
\]

holds for each element \( \phi \) in \( \mathcal{Z} \).
IV.4. The Infinite Horizon Problem

A complete solution to the prediction problem in the correlated case is now within our reach. We shall use the results of the last section to find a representation for \( \{E[\phi(X_{t+1}^\infty)|Y_t]\}_0^\infty \) in terms of a collection of statistics \( \{(X_{t+1}, B_{t+1})\}_0^\infty \) and \( \{\Sigma_{t+1}\}_0^\infty \), which are the natural extension of the statistics \( \{(X_{t+1}, B_{t+1})\}_0^T \) and \( \{\Sigma_{t+1}\}_0^T \) introduced in the previous section, and then provide a method for computing these statistics. Our arguments are the ones given in Section III.3. It will be clear, upon inspection of the solution, that if \( W_t^\infty_t \) and \( V_t^\infty_t \) are uncorrelated for each \( t = 0, 1, \ldots \) and \( \{V_t^\infty_t\}_0^\infty \) is standard (i.e., \( \Sigma_{t+1} w = 0 \) and \( \Sigma_{t+1} v = I_k \) for \( t = 0, 1, \ldots \)), that this solution will reduce to the solution of Chapter III.

Note that, upon combining (2.16) and (3.4), we have the dynamical description

\[
\begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} = \begin{pmatrix}
A_t & 0 \\
0 & I_n
\end{pmatrix} \begin{pmatrix}
X_t \\
B_t
\end{pmatrix} + \begin{pmatrix}
I_n & 0 \\
0 & \Psi'(t, 0)H_t^v(\Sigma_{t+1}^v)^{-1}
\end{pmatrix} \begin{pmatrix}
W_{t+1} \\
V_{t+1}
\end{pmatrix}
\]

\[
\begin{pmatrix}
X_t \\
B_t
\end{pmatrix} = \begin{pmatrix} 0 \\
0 \end{pmatrix}
\]

\[
Y_t = \begin{pmatrix} H_t & 0 \end{pmatrix} \begin{pmatrix}
X_t \\
B_t
\end{pmatrix} + \begin{pmatrix} 0 & I_k \end{pmatrix} \begin{pmatrix}
W_{t+1} \\
V_{t+1}
\end{pmatrix}.
\]

\( t = 0, 1, \ldots \) \hspace{1cm} (4.1)

We define a sequence of probability measures \( \{\tilde{P}_t\}_0^\infty \) on \((\Omega, \mathcal{F})\) by the Radon-Nikodym derivatives

\[
\frac{d\tilde{P}_t}{dP} = L_t. \hspace{1cm} t = 0, 1, \ldots \hspace{1cm} (4.2)
\]

For each \( t = 0, 1, \ldots \), let \( \tilde{E}_t \) be the expectation operator associated with \( \tilde{P}_t \), and let conditional expectation operators be similarly defined. Define the conditional means and variances

\[
\begin{pmatrix}
\bar{X}_{T+1}^T \\
\bar{B}_{T+1}^T
\end{pmatrix} := \tilde{E}_{T+1} \left[ \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} \right| Y_t
\]

\[
\Sigma_{t+1}^T := \tilde{E}_{T+1} \left[ \left( \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} - \begin{pmatrix}
\bar{X}_{T+1}^T \\
\bar{B}_{T+1}^T
\end{pmatrix} \right) \left( \begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} - \begin{pmatrix}
\bar{X}_{T+1}^T \\
\bar{B}_{T+1}^T
\end{pmatrix} \right)' \right]
\]

\hspace{1cm} (4.3)
for each $T = 0, 1, \ldots$ and $t = 0, 1, \ldots, T$, and set

\[
\begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} := \begin{pmatrix}
\tilde{X}_{t+1} \\
\tilde{B}_{t+1}
\end{pmatrix} \\
\Sigma_{t+1} := \Sigma_{t+1}^t.
\]

Clearly

\[E[\phi(X_{t+1}^0)|Y_t] = \frac{U(\phi[X_{t+1}, B_{t+1}; M_{t+1}, \Psi(t+1, 0); \Sigma_{t+1}])}{U[[X_{t+1}, B_{t+1}; M_{t+1}, \Psi(t+1, 0); \Sigma_{t+1}^T]]}. \quad t = 0, 1, \ldots, T \quad (4.5)\]

The following result is proved in the same manner as Proposition III.3.1.

**Proposition 4.1.** The processes $\{(X_{t+1}, B_{t+1})\}_{0}^{T}$ and $\{\Sigma_{t+1}\}_{0}^{T}$ are related to the processes $\{(\tilde{X}_{t+1}, \tilde{B}_{t+1})_{T}^{t}; T = 0, 1, \ldots, t = 0, 1, \ldots, T \}$ and $\{\Sigma_{t+1}^{T}; T = 0, 1, \ldots, t = 0, 1, \ldots, T \}$ by

\[
\begin{pmatrix}
X_{t+1} \\
B_{t+1}
\end{pmatrix} := \begin{pmatrix}
\tilde{X}_{t+1} \\
\tilde{B}_{t+1}
\end{pmatrix} \\
\Sigma_{t+1} := \Sigma_{t+1}^T. \quad T = 0, 1, \ldots, t = 0, 1, \ldots, T \quad (4.6)
\]

As in Section III.2, we are, for an arbitrary $T = 0, 1, \ldots$, able to apply Appendix A to find a recursive form for $\{\Sigma_{t+1}^{T}\}_{0}^{T}$ and $\{(\tilde{X}_{t+1}, \tilde{B}_{t+1})\}_{0}^{T}$ which is independent of our choice of $T$; in view of Proposition 4.1, this recursion extends to the processes $\{(X_{t+1}, B_{t+1})\}_{0}^{\infty}$ and $\{\Sigma_{t+1}\}_{0}^{\infty}$. We shall again leave the details of the application of Appendix A to the reader, and simply state the result.

**Theorem 4.1.** Let the deterministic sequences $\{P_{t}\}_{0}^{\infty}$, $\{Q_{t}\}_{0}^{\infty}$, and $\{R_{t}\}_{0}^{\infty}$ in $Q_n$, $M_n$, and $Q_n$ (respectively) be such that

\[
\Sigma_{t+1} = \begin{pmatrix}
P_{t+1} & Q_{t+1} \\
Q_{t+1} & R_{t+1}
\end{pmatrix}. \quad t = 0, 1, \ldots \quad (4.7)
\]

Then $\{P_{t}\}_{0}^{\infty}$ propagates according to

\[
P_{t+1} = A_{t}P_{t}A_{t}^{\prime} - [A_{t}P_{t}H_{t}^{\prime} + \Sigma_{t+1}^{u}] [H_{t}P_{t}H_{t}^{\prime} + \Sigma_{t+1}^{u}]^{-1} [A_{t}P_{t}H_{t}^{\prime} + \Sigma_{t+1}^{u}]^{\prime} + \Sigma_{t+1}^{w} \]

\[P_{0} = 0. \quad t = 0, 1, \ldots \quad (4.8)\]
Define an auxiliary deterministic sequence \( \{J_t\}_0^\infty \) taking values in \( Q_k \) by

\[
J_t := H_tP_tH_t' + \Sigma_{t+1}^v_t, \quad t = 0, 1, \ldots \quad (4.9)
\]

Then \( \{Q_t\}_0^\infty \) propagates according to

\[
Q_{t+1} = A_tQ_t - [A_tP_tH_t' + \Sigma_{t+1}^{uv}_t]J_t^{-1}H_t(Q_t + \Psi(t, 0)) + \Sigma_{t+1}^{uv}(\Sigma_{t+1}^v)^{-1}H_t\Psi(t, 0)
\]

\[
Q_0 = 0
\]

and \( \{R_t\}_0^\infty \) propagates according to

\[
R_{t+1} = R_t - (Q_t + \Psi(t, 0))^tH_tJ_t^{-1}H_t(Q_t + \Psi(t, 0)) + \Psi'(t, 0)H_tH_t\Psi(t, 0)
\]

\[
R_0 = 0.
\]

Furthermore, the processes \( \{X_t\}_0^\infty \) and \( \{B_t\}_0^\infty \) evolve according to

\[
X_{t+1} = [A_t - [A_tP_tH_t' + \Sigma_{t+1}^{uv}_t]J_t^{-1}H_t]X_t + [A_tP_tH_t' + \Sigma_{t+1}^{uv}_t]J_t^{-1}Y_t
\]

\[
X_0 = 0
\]

and

\[
B_{t+1} = B_t - (Q_t + \Psi(t, 0))^tH_tJ_t^{-1}H_tX_t + (Q_t + \Psi(t, 0))^tH_tJ_t^{-1}Y_t
\]

\[
B_0 = 0.
\]

In view of (4.5), we have in fact solved the prediction problem in the general (i.e., correlated) case. The essential elements of this solution are summarized in

**Theorem 4.2.** Let the processes \( \{(X_{t+1}, B_{t+1})\}_0^\infty \) and \( \{\Sigma_{t+1}\}_0^\infty \) be as in Theorem 4.1.

Then for any \( \phi \) in \( \mathcal{Z} \) and any \( t = 0, 1, \ldots \), the representation

\[
E[\phi(X_{t+1}^\phi)|Y_t] = \frac{U\phi[X_{t+1}, B_{t+1}; M_{t+1}, \Psi(t+1, 0); \Sigma_{t+1}]}{U[I[X_{t+1}, B_{t+1}; M_{t+1}, \Psi(t+1, 0); \Sigma_{t+1}]]} \quad (4.14)
\]

holds.
In a manner similar to Proposition III.3.2, we can slightly simplify the representation of Theorem 4.1 by using the quantities

\[ Q_t^* := Q_t + \Psi(t, 0) \quad t = 0, 1, \ldots \quad (4.15) \]

\[ R_t^* := M_t - R_t. \]

We then have the following representation.

**Proposition 3.2.** Define deterministic sequences \( \{Q_t^*\}_0^\infty \) in \( \mathcal{M}_n \) and \( \{R_t^*\}_0^\infty \) in \( \mathbb{Q}_n \) by the dynamical equations

\[ Q_{t+1}^* = [A_t - [A_t P_t H_t^* + \Sigma_{t+1}^{uv}] J_t^{-1} H_t] Q_t^* \quad t = 0, 1, \ldots \quad (4.16) \]

\[ Q_0^* = I_n \]

and

\[ R_{t+1}^* = R_t^* + Q_t^* H_t J_t^{-1} H_t Q_t^* \quad t = 0, 1, \ldots \quad (4.17) \]

\[ R_0^* = 0. \]

Then

\[ \Sigma_{t+1} = \begin{pmatrix} P_{t+1} & Q_{t+1}^* - \Psi(t + 1, 0) \\ Q_{t+1}^* - \Psi'(t + 1, 0) & M_{t+1} - R_{t+1}^* \end{pmatrix}, \quad t = 0, 1, \ldots \quad (4.18) \]

and we may rewrite (4.13) as

\[ B_{t+1} = B_t - Q_t^* H_t J_t^{-1} H_t X_t + Q_t^* H_t J_t^{-1} Y_t \quad t = 0, 1, \ldots \quad (4.19) \]

\[ B_0 = 0. \]

Note that if \( W_{t+1}^\circ \) and \( V_{t+1}^\circ \) are uncorrelated for \( t = 0, 1, \ldots \) and \( \{V_{t+1}^\circ\}_0^\infty \) is standard (i.e., the uncorrelated case), the formulae of this section become identical to the corresponding formulae of Section III.2. Thus, without any fear of confusion, we shall, in the remainder of this thesis, assume (unless otherwise stated) that \( \{(W_{t+1}^\circ, V_{t+1}^\circ)\}_0^\infty \) is a general GWN process such that the covariance matrix \( \Sigma_{t+1}^\circ \) is positive definite for \( t = 0, 1, \ldots \) and assume that the relevant processes are as defined in this Chapter.
CHAPTER V: REPRESENTATIONS FOR \( \{\hat{X}_t\}^\infty_1, \{\hat{X}_t^K\}^\infty_1 \) AND \( \{\epsilon_t\}^\infty_1 \)

V.1. Representations for \( \{\hat{X}_t\}^\infty_1 \) and \( \{\hat{X}_t^K\}^\infty_1 \)

In this section, we apply the formulae of the previous section to find, for \( t = 0, 1, \ldots \), representations for the MMSE and LLSE estimates of \( X_{t+1}^\circ \) on the basis of \( \{Y_0, Y_1, \ldots, Y_t\} \). Our first step will be to find a representation for the conditional characteristic function \( E[\exp[i\theta' X_{t+1}^\circ]|\mathcal{Y}_t] \). Under the moment assumptions on \( \xi \), we then recover an expression for the conditional mean by differentiating with respect to \( \theta \). Finally, by substituting a Gaussian distribution for \( F \) in this representation for \( \hat{X}_{t+1} \), we obtain a formula for \( \hat{X}_{t+1}^K \).

To begin, we verify that \( \hat{X}_{t+1} \) and \( \hat{X}_{t+1}^K \) are well-defined for \( t = 0, 1, \ldots \). Indeed, we may write

\[
X_{t+1}^\circ = \Phi(t + 1, 0) \xi + \sum_{s=0}^{t} \Phi(t, s) W_{s+1}^\circ. \quad t = 0, 1, \ldots \tag{1.1}
\]

Since the RV’s \( \xi \) and \( \{(W_{t+1}^\circ, V_{t+1}^\circ)\}_0^\infty \) are all \( P \)-square-integrable, so is \( X_{t+1}^\circ \), and thus \( Y_t \), for \( t = 0, 1, \ldots \). Consequently, \( \hat{X}_{t+1} \) and \( \hat{X}_{t+1}^K \) are well-defined for \( t = 0, 1, \ldots \).

We now consider the conditional characteristic function \( E[\exp[i\theta' X_{t+1}^\circ]|\mathcal{Y}_t] \). For \( \theta \) in \( IR^n \), let \( \psi_\theta \) in \( \mathcal{E} \) be defined by

\[
\psi_\theta(x) := \exp[i\theta' x] \quad x \in IR^n \tag{1.2}
\]

and with the notation of Section I.3, for \( t = 0, 1, \ldots \) and \( \theta \) in \( IR^n \), define the RV \( q_{t+1}(\theta) \) by

\[
q_{t+1}(\theta) := U \psi_\theta[X_{t+1}, B_{t+1}; M_{t+1}, \Psi(t + 1, 0); \Sigma_{t+1}]. \tag{1.3}
\]

Noting that \( \psi_0 = \mathbb{I} \), we have

\[
q_{t+1}(0) := U \mathbb{I}[X_{t+1}, B_{t+1}; M_{t+1}, \Psi(t + 1, 0); \Sigma_{t+1}], \tag{1.4}
\]

so that

\[
E[\exp[i\theta' X_{t+1}^\circ]|\mathcal{Y}_t] = \frac{q_{t+1}(\theta)}{q_{t+1}(0)}. \quad t = 0, 1, \ldots \tag{1.5}
\]

Now for \( t = 0, 1, \ldots \), the RV \( X_{t+1}^\circ \) is \( P \)-integrable, so

\[
\hat{X}_{t+1} = \frac{1}{i} \nabla_\theta E[\exp[i\theta' X_{t+1}^\circ]|\mathcal{Y}_t] |_{\theta = 0} = \frac{1}{i} \frac{\nabla_\theta q_{t+1}(\theta)}{q_{t+1}(0)} |_{\theta = 0}. \quad t = 0, 1, \ldots \tag{1.6}
\]

49
Fix $t = 0, 1, \ldots$ and $\theta$ in $IR^n$. For $u$ and $v$ in $IR^n$, we easily check that

$$T\psi_\theta[u, v; \Sigma_{t+1}] = E'[\exp[i\theta' u] \exp[(i\theta' v) \left( \frac{X_{t+1}}{B_{t+1}} \right)]]$$

$$= \exp[i\theta' u] \exp[-\frac{1}{2} \theta' P_{t+1} \theta + iv' (Q_{t+1}^* - \Psi(t + 1, 0)) \theta]$$

$$\cdot \exp[\frac{1}{2} v' (M_{t+1} - R_{t+1}^*) v],$$

so that for all $x$ in $IR^n$,

$$T\psi_\theta[x + \Psi(t + 1, 0) \xi_{t+1}, \xi_{t+1}; \Sigma_{t+1}]$$

$$= \exp[i x' \theta - \frac{1}{2} \theta' P_{t+1} \theta + i \xi_{t+1} \xi_{t+1}^* Q_{t+1}^* \theta + \frac{1}{2} \xi_{t+1} \xi_{t+1}^* (M_{t+1} - R_{t+1}^*) \xi_{t+1}].$$

Consequently,

$$q_{t+1}(\theta) = E'[\exp[i \theta' x - \frac{1}{2} \theta' P_{t+1} \theta + i \xi_{t+1} \xi_{t+1}^* Q_{t+1}^* \theta + \frac{1}{2} \xi_{t+1} \xi_{t+1}^* (M_{t+1} - R_{t+1}^*) \xi_{t+1}]]$$

$$= \exp[i \theta' X_{t+1} - \frac{1}{2} \theta' P_{t+1} \theta] \int_{IR^n} \exp[z' (i Q_{t+1}^* \theta + B_{t+1}) - \frac{1}{2} z' R_{t+1}^* z] dF(z)$$

for $t = 0, 1, \ldots$ and $\theta$ in $IR^n$, and (1.5) now takes the form

$$E[\exp[i \theta' X_{t+1}]] = \exp[i \theta' X_{t+1} - \frac{1}{2} \theta' P_{t+1} \theta]$$

$$\frac{\int_{IR^n} \exp[z' (i Q_{t+1}^* \theta + B_{t+1}) - \frac{1}{2} z' R_{t+1}^* z] dF(z)}{\int_{IR^n} \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z)}.$$

(1.10)

Proceeding formally, from (1.10) we would then expect that for all $t = 0, 1, \ldots$

$$\dot{X}_{t+1} = X_{t+1} + Q_{t+1}^* \int_{IR^n} z \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z)$$

$$\int_{IR^n} \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z).$$

(1.11)

Clearly some justification is needed for (1.11). Specifically, it suffices to show that for all $t = 0, 1, \ldots$, the equality

$$\nabla_\theta \int_{IR^n} \exp[z' (i Q_{t+1}^* \theta + B_{t+1}) - \frac{1}{2} z' R_{t+1}^* z] dF(z) \bigg|_{\theta = 0}$$

$$= i Q_{t+1}^* \int_{IR^n} z \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z)$$

(1.12)
holds $P$-a.s.

The proof of (1.12) requires the following result.

**Theorem 1.1.** For any $t = 0, 1, \ldots$ and any IR-valued, nonnegative $\mathcal{Y}_t \vee \sigma\{\xi\}$-measurable RV $X$, the relation

$$E[X] = \bar{E}_{t+1}[X \exp[\xi' B_{t+1} - \frac{1}{2} \xi' \beta^*_t \xi]].$$

(1.13)

holds.

**Proof.** By definition,

$$E[X] = \bar{E}_{t+1}[X L^{-1}_{t+1}]$$

$$= \bar{E}_{t+1}[X \exp[\xi' B_{t+1} - \frac{1}{2} \xi' M_{t+1} \xi]]$$

$$= \bar{E}_{t+1}[\bar{E}_{t+1}[X \exp[\xi' B_{t+1} - \frac{1}{2} \xi' M_{t+1} \xi]|\mathcal{Y}_t \vee \sigma\{\xi\}]]$$

$$= \bar{E}_{t+1}[X \exp[-\frac{1}{2} \xi' M_{t+1} \xi]\bar{E}_{t+1}\left[\exp[\xi' B_{t+1}^-] | \mathcal{Y}_t \vee \sigma\{\xi\}\right],$$

(1.14)

by using the law of iterated conditioning and the measurability of $X \exp[-\frac{1}{2} \xi' M_{t+1} \xi]$ with respect to the $\sigma$-field $\mathcal{Y}_t \vee \sigma\{\xi\}$. But

$$\bar{E}_{t+1}[X \exp[\xi' B_{t+1}^-]|\mathcal{Y}_t \vee \sigma\{\xi\}] = \bar{E}_{t+1}\left[\exp[\xi' \tilde{B}_{t+1}^-] | \mathcal{Y}_t \vee \sigma\{\xi\}\right] X \exp[\xi' B_{t+1}^-],$$

(1.15)

where $\tilde{B}_{t+1} := B_{t+1} - B_{t+1}^-$ for $t = 0, 1, \ldots$. Now from Section IV.3, we know that under $\tilde{P}_{t+1}$, $\tilde{B}_{t+1}$ is normal with zero mean and covariance $M_{t+1} - \beta^*_t $ and is independent of the $\sigma$-field $\mathcal{Y}_t \vee \sigma\{\xi\}$. By techniques established in Section III.2, we then get

$$\left.\bar{E}_{t+1}[\exp[\xi' \tilde{B}_{t+1}^-]|\mathcal{Y}_t \vee \sigma\{\xi\}] \right|_{x = \xi} = \bar{E}_{t+1}[\exp[\xi' \tilde{B}_{t+1}^-]]$$

$$= \exp\left[\frac{1}{2} \xi' (M_{t+1} - \beta^*_t) \xi\right].$$

(1.16)

Collecting (1.14)-(1.16), (1.13) follows.

We are now in a position to prove (1.12), which in turn verifies representation (1.11).

**Proposition 1.1.** Equation (1.12) is true.
Proof. Fix \( t = 0, 1, \ldots \). From Theorem 1.1, we know that
\[
\tilde{E}_{t+1}[||\xi|| \exp[\xi' B_{t+1} - \frac{1}{2} \xi' R_{t+1}^* \xi]] = E[||\xi||] < \infty, \tag{1.17}
\]
while iterated conditioning implies that
\[
\tilde{E}_{t+1}[||\xi|| \exp[\xi' B_{t+1} - \frac{1}{2} \xi' R_{t+1}^* \xi]] = \tilde{E}_{t+1}[\tilde{E}_{t+1}[||\xi|| \exp[\xi' B_{t+1} - \frac{1}{2} \xi' R_{t+1}^* \xi]} | \mathcal{Y}_t] \tag{1.18}
\]
The techniques of Sections III.2 and IV.3 indicate that
\[
\tilde{E}_{t+1}[||\xi|| \exp[\xi' B_{t+1} - \frac{1}{2} \xi' R_{t+1}^* \xi]} | \mathcal{Y}_t] = \tilde{E}_{t+1}[||\xi|| \exp[\xi' B_{t+1} - \frac{1}{2} \xi' R_{t+1}^* \xi]} | b = B_{t+1}
\]
\[
= \int_{\mathbb{R}^n} \|z\| \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z), \tag{1.19}
\]
since by Lemma IV.3.1, \( B_{t+1} \) is \( \mathcal{Y}_t \)-measurable and \( \mathcal{Y}_t \) is \( \bar{P}_{t+1} \)-independent of \( \sigma \{ \xi \} \). Upon combining (1.17)-(1.19), we find that
\[
\tilde{E}_{t+1}[\int_{\mathbb{R}^n} \|z\| \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z)] < \infty. \tag{1.20}
\]
Consequently,
\[
\int_{\mathbb{R}^n} \|z\| \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z) < \infty \tag{1.21}
\]
\( \bar{P}_{t+1} \)-a.s., and thus (1.21) holds \( P \)-a.s., since \( P \) and \( \bar{P}_{t+1} \) are equivalent on \( \mathcal{F}_{t+1} \).

To prove (1.12), set
\[
f_\theta(z) := \frac{\exp[i z' Q^*_{t+1} \theta] - 1 - i \theta' Q^*_{t+1} z]}{|\theta|} \tag{1.22}
\]
for \( z \) in \( \mathbb{R}^n \) and \( \theta \) in \( \mathbb{R} \sim \{0\} \). By well-known results in complex analysis, \( 0 \leq f_\theta(z) \leq 2 ||Q^*_{t+1} z|| \leq 2 ||Q^*_{t+1}|| ||z|| \) for all \( z \) in \( \mathbb{R}^n \) and \( \theta \) in \( \mathbb{R}^n \sim \{0\} \), and \( \lim \theta \to 0 f_\theta(z) = 0 \) for all \( z \) in \( \mathbb{R}^n \). Consequently, by dominated convergence and the fact that
\[
2 ||Q^*_{t+1}|| \int_{\mathbb{R}^n} \|z\| \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z) < \infty \tag{1.23}
\]
\[ P.a.s., \text{ we find that} \]
\[
\lim_{\theta \to 0} \int_{\mathbb{R}^n} f_{\theta}(z) \exp[z' \mathcal{B}_{t+1} + \frac{1}{2} z' R_{t+1}^* z]dF(z) \\
= \int_{\mathbb{R}^n} \lim_{\theta \to 0} f_{\theta}(z) \exp[z' \mathcal{B}_{t+1} + \frac{1}{2} z' R_{t+1}^* z]dF(z) \\
= 0. \tag{1.24}
\]

As a result,
\[
\lim_{\theta \to 0} \frac{1}{|\theta|} \left\{ \int_{\mathbb{R}^n} \exp[iz' \mathcal{Q}_{t+1}^* \theta] \exp[z' \mathcal{B}_{t+1} + \frac{1}{2} z' R_{t+1}^* z]dF(z) \\
- \int_{\mathbb{R}^n} \exp[z' \mathcal{B}_{t+1} + \frac{1}{2} z' R_{t+1}^* z]dF(z) - i\theta' \mathcal{Q}_{t+1}^* \int_{\mathbb{R}^n} z \exp[z' \mathcal{B}_{t+1} + \frac{1}{2} z' R_{t+1}^* z]dF(z) \right\} \\
= 0 \tag{1.25}
\]

\[ P.a.s., \text{ or, equivalently,} \]
\[
\nabla_{\theta} \int_{\mathbb{R}^n} \exp[z' \mathcal{Q}_{t+1}^* \theta + \mathcal{B}_{t+1} + \frac{1}{2} z' R_{t+1}^* z]dF(z) \bigg|_{\theta=0} \\
= i\mathcal{Q}_{t+1}^* \int_{\mathbb{R}^n} z \exp[z' \mathcal{B}_{t+1} + \frac{1}{2} z' R_{t+1}^* z]dF(z) \tag{1.26}
\]

\[ P.a.s., \text{ which is (1.12). Note that we have implicitly used (1.21) and the fact (which naturally stems from representation IV.3.6) that P-a.s.} \]
\[
\int_{\mathbb{R}^n} \exp[z' \mathcal{B}_{t+1} + \frac{1}{2} z' R_{t+1}^* z]dF(z) < \infty \tag{1.27}
\]

in separating the integral
\[
\int_{\mathbb{R}^n} \left\{ \frac{\exp[iz' \mathcal{Q}_{t+1}^* \theta - 1 - i\theta' \mathcal{Q}_{t+1}^* z]}{|\theta|} \exp[z' \mathcal{B}_{t+1} + \frac{1}{2} z' R_{t+1}^* z]dF(z) \right\} \tag{1.28}
\]

This completes the proof of (1.12).

Since (1.12) holds, we then know that the representation (1.11) is true. Now let \( G \) be a Gaussian distribution with mean \( \mu \) and covariance \( \Delta \). A representation for \( \hat{X}_{t+1}^K \) for \( t = 0, 1, \ldots \) may be found by replacing \( F \) by \( G \) in (1.11) as follows.

For any two square-integrable RV's \( A \) and \( B \), let \( \hat{E}[A|B] \) be the wide-sense conditional expectation of \( A \) given \( B \). The classical Kalman filtering equations provide a means of
generating the wide-sense conditional expectations \( \{ \bar{E}[X_{t+1}^\circ | (Y_0, Y_1, \ldots, Y_t)] \}_{t=0}^\infty \). Now we may always enlarge \((\Omega, \mathcal{F}, P)\) to find an RV \( \xi \) with distribution \( G \) and which is independent of \( \{ W_{t+1}^\circ, V_{t+1}^\circ; t = 0, 1, \ldots \} \). Let the processes \( \{ X_t^\circ \}_{t=0}^\infty \) and \( \{ Y_t \}_{t=0}^\infty \) be generated by

\[
\begin{align*}
X_{t+1}^\circ &= A_t X_t^\circ + W_{t+1}^\circ \\
X_0^\circ &= \xi \\
Y_t &= H_t X_t^\circ + V_{t+1}^\circ.
\end{align*}
\tag{1.29}
\]

Note that \( \{ X_{t+1}^\circ, Y_t; t = 0, 1, \ldots \} \) and \( \{ X_t^\circ, Y_t; t = 0, 1, \ldots \} \) have the same second-order statistics. By definition of wide-sense expectations, we then know that if the mappings \( \{ \phi_t \}_{t=0}^\infty \) are such that

\[
\bar{E}[X_{t+1}^\circ | (Y_0, Y_1, \ldots, Y_t)] = \phi_t ((\bar{Y}_0, \bar{Y}_1, \ldots, \bar{Y}_t)),
\tag{1.30}
\]

then

\[
\begin{align*}
\hat{X}_{t+1}^K &= \hat{E}[X_{t+1}^\circ | (Y_0, Y_1, \ldots, Y_t)] = \phi_t ((Y_0, Y_1, \ldots, Y_t)).
\end{align*}
\tag{1.31}
\]

However, the statistics of \( \{ \bar{X}_{t+1}^\circ, \bar{Y}_t; t = 0, 1, \ldots \} \) are Gaussian, so in fact

\[
\phi_t ((\bar{Y}_0, \bar{Y}_1, \ldots, \bar{Y}_t)) = \bar{E}[\bar{X}_{t+1}^\circ | (\bar{Y}_0, \bar{Y}_1, \ldots, \bar{Y}_t)].
\tag{1.32}
\]

Defining the processes \( \{ \bar{X}_t \}_{t=0}^\infty \) and \( \{ \bar{X}_t^\circ \}_{t=0}^\infty \) through the recursions

\[
\begin{align*}
\bar{X}_{t+1} &= [A_t - [A_t P_t H_t' + \Sigma_{t+1}^{uu}] J_t^{-1} H_t] \bar{X}_t + [A_t P_t H_t' + \Sigma_{t+1}^{uu}] J_t^{-1} \bar{Y}_t \\
\bar{X}_0 &= 0
\end{align*}
\tag{1.33}
\]

and

\[
\begin{align*}
\bar{B}_{t+1} &= \bar{B}_t - Q_t^{uu} H_t' J_t^{-1} H_t \bar{X}_t + Q_t^{uu} H_t' J_t^{-1} \bar{Y}_t \\
\bar{B}_0 &= 0,
\end{align*}
\tag{1.34}
\]

we conclude that for \( t = 0, 1, \ldots \), the mapping \( \phi_t \) may be represented as the composition of the two transformations

\[
(\bar{Y}_0, \bar{Y}_1, \ldots, \bar{Y}_t) \mapsto (\bar{X}_{t+1}, \bar{B}_{t+1})
\tag{1.35}
\]
and
\[
(x, b) \mapsto z + Q_{t+1}^* \frac{\int_{\mathbb{R}^n} z \exp \left[ z' B_{t+1} + \frac{1}{2} z' R_{t+1}^* z \right] dG(z)}{\int_{\mathbb{R}^n} \exp \left[ z' B_{t+1} + \frac{1}{2} z' R_{t+1}^* z \right] dG(z)}.
\] (1.36)

Consequently, for \( t = 0, 1, \ldots \), \( \phi_t((Y_1, Y_2, \ldots, Y_t)) \) is the composition of
\[
(Y_0, Y_1, \ldots, Y_t) \mapsto (X_{t+1}, B_{t+1})
\] (1.37)

and the mapping (1.36); we may find \( \hat{X}_{t+1}^K \) by simply replacing \( F \) by \( G \) in (1.11).

Recalling assumption (A.4) that \( \Delta \) be positive-definite, by direct evaluation we find that
\[
\int_{\mathbb{R}^n} z \exp \left[ z' B_{t+1} + \frac{1}{2} z' R_{t+1}^* z \right] dG(z) = \frac{R_{t+1}^* + \Delta^{-1}}{\sqrt{\det(\Delta R_{t+1}^* + I)}} \cdot \exp \left[ -\frac{1}{2} \mu' \Delta^{-1} \mu + \frac{1}{2} (B_{t+1} + \Delta^{-1} \mu)' (R_{t+1}^* + \Delta^{-1})^{-1} (B_{t+1} + \Delta^{-1} \mu) \right],
\] (1.38)

and that
\[
\int_{\mathbb{R}^n} \exp \left[ z' B_{t+1} + \frac{1}{2} z' R_{t+1}^* z \right] dG(z) = \frac{1}{\sqrt{\det(\Delta R_{t+1}^* + I)}} \cdot \exp \left[ -\frac{1}{2} \mu' \Delta^{-1} \mu + \frac{1}{2} (B_{t+1} + \Delta^{-1} \mu)' (R_{t+1}^* + \Delta^{-1})^{-1} (B_{t+1} + \Delta^{-1} \mu) \right]
\] (1.39)
P-a.s. for \( t = 0, 1, \ldots \), so that
\[
\hat{X}_{t+1}^K = X_{t+1} + Q_{t+1}^* [R_{t+1}^* + \Delta^{-1}]^{-1} (B_{t+1} + \Delta^{-1} \mu)
\] (1.40)
P-a.s. We have consequently arrived at the the main result of this section.

**Theorem 1.2.** For all \( t = 0, 1, \ldots \), the representations
\[
\hat{X}_{t+1} = X_{t+1} + Q_{t+1}^* \frac{\int_{\mathbb{R}^n} z \exp \left[ z' B_{t+1} + \frac{1}{2} z' R_{t+1}^* z \right] dF(z)}{\int_{\mathbb{R}^n} \exp \left[ z' B_{t+1} + \frac{1}{2} z' R_{t+1}^* z \right] dF(z)}
\] (1.41)

and
\[
\hat{X}_{t+1}^K = X_{t+1} + Q_{t+1}^* [R_{t+1}^* + \Delta^{-1}]^{-1} (B_{t+1} + \Delta^{-1} \mu)
\] (1.42)
hold P-a.s.
Now (1.42) provides a non-standard representation for the Kalman filter associated with system (1.1.1). This representation is notable in that it explicitly displays the effects of the mean $\mu$ and covariance $\Delta$ of the initial condition $\xi$—the only dependence of the filtering formulae on $\mu$ and $\Delta$ is through the affine mapping $x \mapsto [R_{t+1}^* + \Delta^{-1}]^{-1}[x + \Delta^{-1} \mu]$. By Theorem IV.4.1 and Proposition IV.4.2, this is equivalent to the fact that the dynamics of $\{P_t\}_0^\infty$, $\{Q_t^*\}_0^\infty$, $\{R_t^*\}_0^\infty$, $\{X_t\}_0^\infty$ and $\{B_t\}_0^\infty$ do not depend upon either $\mu$ or $\Delta$. 
V.2. A Representation for $\{\epsilon_t\}_{t=1}^{\infty}$

This section is devoted to using the results of the previous section to characterize $\epsilon_{t+1} := E[||\hat{X}_{t+1} - \hat{X}_{t+1}^K||^2]$ for $t = 0, 1, \ldots$. Directly from Theorem 1.2, we have

$$\hat{X}_{t+1} - \hat{X}_{t+1}^K = Q_{t+1}^*$$

$$\int_{\mathbb{R}^n} \{z - [R_{t+1}^* + \Delta^{-1}]^{-1} [B_{t+1} + \Delta^{-1} \mu] \} \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z)$$

$$\int_{\mathbb{R}^n} \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z),$$

$$t = 0, 1, \ldots \quad (2.1)$$

and by utilizing Theorem 1.1, we get

$$\epsilon_{t+1} = E[||\hat{X}_{t+1} - \hat{X}_{t+1}^K||^2]$$

$$= E \left[ \left\{ \frac{1}{\int_{\mathbb{R}^n} \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z)} \right\}^2 \right] \quad (2.2)$$

$$\cdot \left\| Q_{t+1}^* \int_{\mathbb{R}^n} \{z - [R_{t+1}^* + \Delta^{-1}]^{-1} [B_{t+1} + \Delta^{-1} \mu] \} \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z) \right\|^2$$

$$= \tilde{E}_{t+1} \left[ \frac{1}{\int_{\mathbb{R}^n} \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z)} \right] \quad (2.3)$$

$$\cdot \left\| Q_{t+1}^* \int_{\mathbb{R}^n} \{z - [R_{t+1}^* + \Delta^{-1}]^{-1} [B_{t+1} + \Delta^{-1} \mu] \} \exp[z' B_{t+1} - \frac{1}{2} z' R_{t+1}^* z] dF(z) \right\|^2.$$

In (2.3), we are taking the $\tilde{E}_{t+1}$-expectation of a function of $B_{t+1}$; the following lemma provides us with the statistics of $B_{t+1}$ under $\tilde{E}_{t+1}$.

**Lemma 2.1.** For $t = 0, 1, \ldots$, $B_{t+1}$ is a normal RV with zero mean and covariance $R_{t+1}^*$ under $\tilde{E}_{t+1}$.

**Proof.** From Lemma IV.3.1, it is clear that $B_{t+1}$, $\tilde{B}_{t+1}$, and $\hat{B}_{t+1}$ have Gaussian distributions under $\tilde{E}_{t+1}$ for $t = 0, 1, \ldots$. Note that for all $t = 0, 1, \ldots$,

$$\tilde{E}_{t+1}[B_{t+1}] = \tilde{E}_{t+1}[\tilde{B}_{t+1}] + \tilde{E}_{t+1}[\hat{B}_{t+1}]$$

$$= \tilde{E}_{t+1}[\tilde{B}_{t+1}], \quad (2.4)$$

57
and

\[ E_{t+1}[B_{t+1}B'_{t+1}] = E_{t+1}\left[ (B_{t+1} + B'_{t+1}) E_{t+1}[B_{t+1}B'_{t+1}] \right] \]

\[ = E_{t+1}\left[ B_{t+1}B'_{t+1} \right] + E_{t+1}\left[ B_{t+1}B'_{t+1} \right] + E_{t+1}\left[ B_{t+1}B'_{t+1} \right] \]

\[ + E_{t+1}\left[ B_{t+1}B'_{t+1} \right] \]

\[ = E_{t+1}\left[ B_{t+1}B'_{t+1} \right] + E_{t+1}\left[ B_{t+1}B'_{t+1} \right], \tag{2.5} \]

since \( E_{t+1}[B_{t+1}B'_{t+1}] = 0 \) for \( t = 0, 1, \ldots \) by the orthogonality principle. By the recursion (IV.3.4), it is plain that

\[ E_{t+1}[B_{t+1}] = E_{t+1}[B_t] + \Psi'(t,0)H_t'(\Sigma_{t+1})^{-1}E_{t+1}[V_{t+1}] \]

\[ = E_{t+1}[B_t]. \tag{2.6} \]

The martingale property of \( \{L_t\} \) then implies that \( E_{t+1}[B_t] = E_t[B_t] \) for \( t = 0, 1, \ldots \), and so we have the recursion

\[ E_{t+1}[B_{t+1}] = E_t[B_t] \]

\[ E_0[B_0] = 0, \tag{2.7} \]

and obviously \( E_t[B_t] = 0 \) for \( t = 0, 1, \ldots \). By (IV.3.4), we also see that \( B_t \) is \( \tilde{P}_{t+1} \)-independent of \( V_{t+1} \) for \( t = 0, 1, \ldots \), so that

\[ E_{t+1}[B_{t+1}B'_{t+1}] = E_{t+1}[B_tB'_t] + \Psi'(t,0)H_t'(\Sigma_{t+1})^{-1}E_{t+1}[V_{t+1}V_{t+1}](\Sigma_{t+1})^{-1}H_t\Psi(t,0) \]

\[ = E_{t+1}[B_tB'_t] + \Psi'(t,0)H_t'(\Sigma_{t+1})^{-1}H_t\Psi(t,0). \tag{2.8} \]

Again, \( E_{t+1}[B_{t+1}B'_{t+1}] = E_t[B_{t+1}B'_{t+1}] \) for \( t = 0, 1, \ldots \), so we have the dynamical evolution

\[ E_{t+1}[B_{t+1}B'_{t+1}] = E_t[B_tB'_t] + \Psi(t,0)H_t'(\Sigma_{t+1})^{-1}H_t\Psi(t,0) \]

\[ E_0[B_0B'_0] = 0, \tag{2.9} \]

which is identical to (IV.3.5), and necessarily \( E_t[B_tB'_t] = M_t \) for \( t = 0, 1, \ldots \). Returning to (2.4) and (2.5), we see that

\[ E_{t+1}[B_{t+1}] = E_{t+1}[B_{t+1}] = 0 \]

\[ E_0[B_0] = 0, \tag{2.10} \]

58
and

\[ E_{t+1}[B_{t+1}B_{t+1}'] = \tilde{E}_{t+1}[B_{t+1}B_{t+1}'] - \tilde{E}_{t+1}[\tilde{B}_{t+1}\tilde{B}_{t+1}'] \]

\[ = M_{t+1} - \{ M_{t+1} - R_{t+1}^* \} \]

\[ = R_{t+1}^*. \quad t = 0, 1, \ldots \quad (2.11) \]

The lemma is verified.

For each \( \Lambda \) in \( Q_n \), let \( G_\Lambda \) be a normal distribution with zero mean and covariance \( \Lambda \); \( G_\Lambda \sim N(0, \Lambda) \). Combining (2.3) and Lemma 2.1, we have the following representation for \( \{ \epsilon_t \}_1^\infty \), which concludes this section.

**Theorem 2.1.** The sequence \( \{ \epsilon_{t+1} \}_0^\infty \) may be represented as

\[
\epsilon_{t+1} = \int_{\mathbb{R}^n} \left\| Q_{t+1}^* \int_{\mathbb{R}^n} \left\{ z - [R_{t+1}^* + \Delta^{-1}]^{-1}[b + \Delta^{-1} \mu] \right\} \exp\left[ z' b - \frac{1}{2} z' R_{t+1}^* z \right] dF(z) \right\|^2 \cdot \frac{1}{\int_{\mathbb{R}^n} \exp\left[ z' b - \frac{1}{2} z' R_{t+1}^* z \right] dF(z)} dG_{R_{t+1}^*}(b). \quad t = 0, 1, \ldots \quad (2.12)
\]
CHAPTER VI: THE ASYMPTOTIC BEHAVIOR OF $\{\epsilon_t\}_1^\infty$

VI.1. Overview

In this part of the thesis, we use the results of earlier sections to study the asymptotic behavior of $\epsilon_t$ as $t$ tends to infinity. We do this for the time-invariant version of system (I.1.1), where $A_t = A$, $H_t = H$, and $\Gamma_{t+1} = \Gamma$ for $t = 0, 1, \ldots$ for some $A$ in $M_{n \times n}$, $H$ in $M_{n \times k}$, and $\Gamma$ in $Q_{2n}$. Our analysis is based on the representation (V.2.12). Note, from (V.2.12), that for $t = 1, 2, \ldots$, $\epsilon_t$ depends upon the dynamics $(A, H, \Gamma)$ of (I.1.1) and the initial distribution $F$, so that we should in fact write

$$\epsilon_t = \epsilon_t((A, H, \Gamma), F). \quad t = 1, 2, \ldots \quad (1.1)$$

We are then interested in $\lim_t \epsilon_t((A, H, \Gamma), F)$ for an arbitrary system $(A, H, F)$ and an arbitrary distribution $F$.

Now observe that there is no loss in generality in assuming that $F$ has zero mean, that $\mu = \int_{\mathbb{R}^n} zdF(z) = 0$. Indeed, with the notation

$$\tilde{X}_t := X_t - \Phi(t, 0)\mu$$

$$\bar{Y}_t := Y_t - H\Phi(t, 0)\mu, \quad t = 0, 1, \ldots \quad (1.2)$$

we see that the RV's $\{\tilde{X}_t\}_0^\infty$ and $\{\bar{Y}_t\}_0^\infty$ obey the dynamics

$$\tilde{X}_{t+1} = A\tilde{X}_t + W_{t+1}$$

$$\tilde{X}_0 = \tilde{\xi}$$

$$\bar{Y}_t = H\tilde{X}_t + V_{t+1}$$

$$t = 0, 1, \ldots \quad (1.3)$$

with $\tilde{\xi} := \xi - \mu$ (and thus $E[\tilde{\xi}] = 0$). (This system is not related to the system V.1.29 with the same notation).

Let $\tilde{E}[A|B]$ be the LLSE estimate of $A$ on the basis of $B$ for any square-integrable random vectors $A$ and $B$. From basic principles, we see that

$$E[\tilde{X}_{t+1}|\bar{Y}_0, \bar{Y}_1, \ldots, \bar{Y}_t] = E[\tilde{X}_{t+1}|Y_0, Y_1, \ldots, Y_t]$$

$$= E[X_{t+1} - \Phi(t + 1, 0)\mu|Y_0, Y_1, \ldots, Y_t] \quad t = 0, 1, \ldots \quad (1.4)$$

$$= \tilde{X}_{t+1} - \Phi(t + 1, 0)\mu,$$
and that
\[ \mathcal{E}[\dot{X}_{t+1}^0|\dot{Y}_0, \dot{Y}_1, \ldots, \dot{Y}_t] = \dot{X}_{t+1}^K - \Phi(t + 1, 0)\mu, \quad t = 0, 1, \ldots \] (1.5)
so that
\[ \dot{X}_{t+1} - \dot{X}_{t+1}^K = \mathcal{E}[\dot{X}_{t+1}^0|\dot{Y}_0, \dot{Y}_1, \ldots, \dot{Y}_t] - \mathcal{E}[\dot{X}_{t+1}^0|\dot{Y}_0, \dot{Y}_1, \ldots, \dot{Y}_t]. \quad t = 0, 1, \ldots \] (1.6)
If we define the probability distribution function \( G \) on \( IR \) by
\[ G(x) := F(x + \mu), \quad x \in IR^n \] (1.7)
then by (1.6), we see that \( \epsilon_t((A, H, \Gamma), F) = \epsilon_t((A, H, \Gamma), G) \) for \( t = 1, 2, \ldots \). Hence we need only to consider \( \{\epsilon_t((A, H, \Gamma), F)\}^\infty_{t=1} \) for those distributions \( F \) such that
\[ \int_{IR^n} \|z\|^2 dF(z) < \infty \text{ and } \int_{IR^n} zdF(z) = 0 \text{ (the first condition is the square integrability condition of Assumption (A.3)).} \] For convenience, define \( \mathcal{E}_n \) as the class of distributions on \( (IR^n, \mathcal{B}(IR^n)) \) such that \( \int_{IR^n} \|z\|^2 dF(z) < \infty \), and let \( \mathcal{D}_n \) be defined by
\[ \mathcal{D}_n := \left\{ F \in \mathcal{E}_n : \int_{IR^n} zdF(z) = 0 \right\}. \] (1.8)
For \( F \) in \( \mathcal{D}_n \), then (V.2.12) becomes
\[ \epsilon_t = \int_{IR^n} \left\| Q_t^* \int_{IR^n} \left\{ z - [R_t^* + \Delta^{-1}]^{-1}b \right\} \exp[z'b - \frac{1}{2} z'R_t^*z]dF(z) \right\|^2 \]
\[ \frac{1}{\int_{IR^n} \exp[z'b - \frac{1}{2} z'R_t^*z]dF(z)} dG_t(b). \quad t = 1, 2, \ldots \] (1.9)
In Section 2, we provide a general stability condition on \( (A, H, \Gamma) \) such that for any \( F \) in \( \mathcal{E}_n \), \( \lim_t \epsilon_t((A, H, \Gamma), F) = 0 \). We consider the scalar, time-invariant version of (1.1.1) in Section 3; in this case, (1.9) reduces to
\[ \epsilon_t((A, H, \Gamma), F) = (q_t^*)^2 I_F(r_t^*) \quad t = 1, 2, \ldots \] (1.10)
with
\[ I_F(t) = \int_{IR} \left\{ \int_{IR^n} \frac{\left\{ \left\{ z - \frac{b}{t+1/\sigma^2} \right\} \exp[zb - z^2 t/2]dF(z) \right\}^2}{\int_{IR} \exp[zb - z^2 t/2]dF(z)} dG_t(b) \right\}. \quad t > 0 \] (1.11)
61
We shall discover that if the stability conditions of Section 2 are not met in this case, in general \( \lim_{t \to \infty} \epsilon_t ((A, H, \Gamma), F) \) will depend nontrivially on both \((A, H, \Gamma)\) and \(F\). We demonstrate this dependence by performing a complete analysis, as \((A, H, \Gamma)\) varies, of two special classes of distributions for \(F\):

Class 1: the distribution \(F\) admits a density with respect to \(\lambda_1\) which is the finite convex combination of a collection of nondegenerate Gaussian densities and

Class 2: the RV \(\xi\) has a finite support \(\{z_i\}_1^n\).

We note that in [1], [21] and [22], attention was given to the filtering and prediction problems for linear discrete-time systems with non-Gaussian noises and initial condition modelled with densities of Class 1. Finally, Section 4 will be devoted to an analysis of the results of Sections 2 and 3 and to the significance of the asymptotic agreement between \(\{\hat{X}_t\}_1^\infty\) and \(\{\hat{X}_t^K\}_1^\infty\).
VI.2. A Stability Criterion

In this section, we provide a general stability condition on the system \((A, H, \Gamma)\) under which \(\lim_{t \to \infty} \epsilon_t((A, H, \Gamma), F) = 0\) holds for an arbitrary initial distribution \(F\) in \(\mathcal{E}_n\). We start with a result of Caines and Mayne which describes the asymptotic behavior of the sequence \(\{P_t\}_0^\infty\) when \((A, H, \Gamma)\) satisfies certain detectability and controllability criterion, and then we use the dynamics (IV.4.16) and (IV.4.17) and representation (V.2.12) to show that \(\lim_{t \to \infty} \epsilon_t((A, H, \Gamma), F) = 0\) under these conditions. The strength of this result lies in the fact that the dynamics \((A, H, \Gamma)\), and not the initial distribution \(F\), enter into the stability criterion. We shall then present a partial result in the converse direction, stating that if the detectability and controllability criteria are met and \(\{\epsilon_t((A, H, \Gamma), F)\}_1^\infty\) decays fast enough, then \(F\) must necessarily be Gaussian.

We begin by defining the following two matrices \(\tilde{A}\) and \(\tilde{C}\) by

\[
\tilde{A} := A - \Sigma^{uw}(\Sigma^v)^{-1}\Sigma^{vw} \quad (2.1)
\]

\[
\tilde{C} := \Sigma^w - \Sigma^{uw}(\Sigma^v)^{-1}\Sigma^{vw}. \quad (2.2)
\]

We then have the following result.

**Theorem 2.1.** If the pair \((A, H)\) is detectable and the pair \((\tilde{A}, \tilde{C})\) is controllable, then

1. \(P_\infty := \lim_{t \to \infty} P_t\) is well-defined and positive-definite and
2. \(K_\infty := A - [AP_\infty H' + \Sigma^{uw}][HP_\infty H' + \Sigma^v]^{-1} H\) is stable.

**Proof.** It is not difficult to verify that

\[
\Sigma^w - \Sigma^{uw}(\Sigma^v)^{-1}\Sigma^{vw} = E \left[ [W_{t+1}^o - E[W_{t+1}^o||V_{t+1}^o]] [W_{t+1}^o - E[W_{t+1}^o||V_{t+1}^o]]' \right] \quad (2.3)
\]

so that \([\Sigma^w - \Sigma^{uw}(\Sigma^v)^{-1}\Sigma^{vw}]^{\frac{1}{2}}\), the unique matrix such that

\[
[\Sigma^w - \Sigma^{uw}(\Sigma^v)^{-1}\Sigma^{vw}]^{\frac{1}{2}}[\Sigma^w - \Sigma^{uw}(\Sigma^v)^{-1}\Sigma^{vw}]^{\frac{1}{2}} = [\Sigma^w - \Sigma^{uw}(\Sigma^v)^{-1}\Sigma^{vw}], \quad (2.4)
\]

is well-defined (see [7, Secs. VIII.6 and VIII.7]). Claim 1 is Appendix 1 in [4] and claim 2 is Theorem 5.1 in [4].

Describe the mappings \(\lambda_{\min} : \mathcal{M}_{n \times n} \mapsto IR\) and \(\lambda_{\max} : \mathcal{M}_{n \times n} \mapsto IR\) by

\[
\lambda_{\min}(K) := \max\{||\lambda|| : \lambda \text{ is an eigenvalue of } K\} \quad (2.5)
\]
and
\[ \lambda_{\text{min}}(K) := \min \{ |\lambda| : \lambda \text{ is an eigenvalue of } K \} \] (2.6)

for \( K \) in \( \mathcal{M}_{n \times n} \). Combining Theorem 2.1 with (IV.4.16) and (IV.4.17), we may prove

**Theorem 2.2.** Under the hypothesis of Theorem 2.1,

1) \( \lim_t Q_t^* = 0 \) with
\[ \limsup_t \frac{1}{t} \ln \lambda_{\text{max}}(Q_t^*Q_t^*) \leq 2 \ln \lambda_{\text{max}}(K_\infty), \] (2.7)

and if the matrix \( A - [AP_tH + \Sigma^w][HP_tH' + \Sigma^v]^{-1}H \) is invertible for each \( t = 0, 1, \ldots \), then
\[ \liminf_t \frac{1}{t} \ln \lambda_{\text{min}}(Q_t^*Q_t^*) \geq 2 \ln \lambda_{\text{min}}(K_\infty); \] (2.8)

and

2) \( R_\infty^* := \lim_t R_t^* \) is well-defined and finite.

**Proof.** Define the sequence \( \{K_t\}_0^\infty \) in \( \mathcal{M}_{n \times n} \) by
\[ K_t := A - [AP_tH + \Sigma^w][HP_tH' + \Sigma^v]^{-1}H, \quad t = 0, 1, \ldots \] (2.9)

and observe from the relation
\[ Q_{t+1}^* = K_tQ_t^* \quad t = 0, 1, \ldots \] (2.10)

that
\[ Q_{t+1}^* = K_tK_{t-1}K_{t-2} \ldots K_0. \] (2.11)

By Theorem 2.1, \( K_\infty = \lim_t K_t \) exists and is stable. Then from Appendix B, (2.7) and (2.8) follow, and since \( \lambda_{\text{max}}(K_\infty) < 1 \), claim 1 is immediate.

Note, from (IV.4.17), that \( R_{t+1}^* \geq R_t^* \) for \( t = 0, 1, \ldots \), so that \( R_\infty^* \) is always well-defined, though possibly infinite.

To prove the finiteness of \( R_\infty^* \), take \( t = 1, 2, \ldots \) and \( x \) in \( IR^n \) with \( \|x\| = 1 \); then
\[ x'R_t^*x = \sum_{s=0}^{t-1} x'Q_t^*H'[HP_sH' + \Sigma^v]^{-1}HQ_s^*x \] (2.12)
\[ \leq M \sum_{s=0}^{t-1} x'Q_t^*Q_s^*x \] (2.13)
\[ \leq \left( M \sum_{s=0}^{t-1} \lambda_{\text{max}}(Q_t^*Q_s^*) \right) \|x\|^2, \] (2.14)
where $M := \sup_t \lambda_{\max} (H [H \gamma H' + \Sigma^v]^{-1} H)$ and where $M < \infty$ by continuity arguments and the fact that $\lim_t P_t = P_\infty$. Thus

$$\| R^* \|^2_H \leq M \sum_{s=0}^{t-1} \lambda_{\max}(Q_s'^* Q_s^*),$$

(2.15)

where the Hilbert norm $\| \cdot \|_H$ on $Q_n$ is defined in the usual way, namely

$$\| R \|_H := \sup_{x; \| x \|=1} \sqrt{x^t R x}. \quad R \in Q_n$$

R \in Q_n \quad (2.16)

But by the first claim, we see that $\sum_{s=0}^{\infty} \lambda_{\max}(Q_s'^* Q_s^*) < \infty$ and the finiteness of $R^*_*$ follows.

Now that we know the behavior of $\{\Sigma_t \}_1^\infty$ when the hypotheses of Theorem 2.1 are in effect, we can proceed with our analysis of $\{\epsilon_t ((A, H, \Gamma), F) \}_1^\infty$. Before doing so, however, let us introduce some notation which will simplify our efforts.

Define the mapping $\gamma : IR^n \times IR^n \times Q_n \mapsto IR^n$ by

$$\gamma(z; b; R) := \exp[z^t b - \frac{1}{2} z^t R z]$$

(2.17)

and the mapping $\Upsilon_F : IR^n \times Q_n \mapsto IR \cup \{\infty\}$ by

$$\Upsilon_F(b; R) := \int_{R^n} \gamma(z; b; R) dF(z).$$

(2.18)

Next, define a family of probability measures $G^F_{b, R}$ on $IR^n$ parametrized by $b$ in $IR^n$ and $R$ in $Q_n$ through the Radon-Nikodym derivatives

$$\frac{dG^F_{b, R}}{dF}(z) := \begin{cases} \gamma(z; b; R)/\Upsilon_F(b; R) & \text{if } \Upsilon_F(b; R) < \infty \\ 1 & \text{if } \Upsilon_F(b; R) = \infty. \end{cases}$$

(2.19)

Returning to our auxiliary space $\Omega', \mathcal{F}', P'$, for $R$ in $Q_n$, let $B_R$ be an $(\Omega', \mathcal{F}')$ Gaussian RV with zero mean and covariance $R$ under $P'$. With this notation, set

$$I_F(R) := E' \left[ \left\| \int_{R^n} [z - [R + \Delta^{-1}]^{-1} B_R] dG_{B_R, R}(z) \right\|^2 \Upsilon_F(B_R; R) \right]$$

(2.20)
for every $R$ in $Q_n$ and every $F$ in $D_n$. Collecting the notation we have developed, it is easy to see that

$$
\lambda_{\min}(Q_t^* Q_t^*) I_F(R_t^*) \leq \epsilon(t ((A, H, \Gamma), F) \leq \lambda_{\max}(Q_t^* Q_t) I_F(R_t^*). \quad t = 1, 2, \ldots \quad (2.21)
$$

Motivated by (2.21), we now study the asymptotic behavior of $\{I(R_t^*)\}_1^\infty$ and have the following result.

**Proposition 2.1.** Under the conditions of Theorem 2.1, if $R_\infty^*$ is finite, then $\limsup_t I_F(R_t^*) < \infty$ for every $F$ in $D_n$.

**Proof.** Fix $t = 0, 1, \ldots$, and, for convenience, define $R := R_t^*$. By Jensen’s inequality, we immediately see that

$$
I_F(R) \leq I'_F(R) \quad t = 0, 1, \ldots \quad (2.22)
$$

where

$$
I'_F(R) := E' \left[ \int_{\mathbb{R}^n} \|z - [R + \Delta^{-1}]^{-1} B_R\|^2 dG_{B_R, R(z)}^E \right] \gamma_F(B_R; R) \quad (2.23)
$$

$$
= E' \left[ \int_{\mathbb{R}^n} \|z - [R + \Delta^{-1}]^{-1} B_R\|^2 \gamma(z; B_R; R) dF(z) \right].
$$

Now by Tonelli’s theorem,

$$
I'_F(R) := \int_{\mathbb{R}^n} \int_{\Omega'} \|z - [R + \Delta^{-1}]^{-1} B_R\|^2 \exp[z' B_R] dP' \exp[-\frac{1}{2} z' R z] dF(z) \quad (2.24)
$$

with

$$
\|z - [R + \Delta^{-1}] B_R\|^2 = z' z - 2 z' [R + \Delta^{-1}]^{-1} B_R + B_R' [R + \Delta^{-1}]^{-1} [R + \Delta^{-1}]^{-1} B_R, \quad (2.25)
$$

upon expanding the integrand in (2.24). Elementary results concerning Gaussian RV’s lead to the relations

$$
\int_{\Omega'} z' z \exp[z' B_R] dP' = z' z \exp[\frac{1}{2} z' R z], \quad (2.26)
$$

$$
\int_{\Omega'} z' [R + \Delta^{-1}]^{-1} B_R \exp[z' B_R] dP' = z' [R + \Delta^{-1}]^{-1} R z \exp[\frac{1}{2} z' R z] \quad (2.27)
$$
\[
\int_{\Omega'} B_R [R + \Delta^{-1}]^{-1} [R + \Delta^{-1}]^{-1} B_R \exp[z'B_R]dP' \n\]
\[
= \{ \text{tr} \left( [R + \Delta^{-1}]^{-1} R [R + \Delta^{-1}]^{-1} \right) + z'[R + \Delta^{-1}]^{-1} [R + \Delta^{-1}]^{-1} R z \} \exp\left[ \frac{1}{2} z' R z \right] \}; \quad (2.28)
\]
and so we obtain
\[
I'(R) = \text{tr} \left( [R + \Delta^{-1}]^{-1} R [R + \Delta^{-1}]^{-1} \right) 
+ \int_{R^n} z' \left[ I_n - 2[R + \Delta^{-1}]^{-1} R + [R + \Delta^{-1}]^{-1} [R + \Delta^{-1}]^{-1} R \right] z dF(z) 
= \text{tr} \left( [R + \Delta^{-1}]^{-1} R [R + \Delta^{-1}]^{-1} \right) 
+ \int_{R^n} z' \Delta^{-1} [R + \Delta^{-1}]^{-1} [R + \Delta^{-1}]^{-1} \Delta^{-1} z dF(z)
\]
and consequently,
\[
I'_p (R^*_t) = \text{tr} \left( [R^*_t + \Delta^{-1}]^{-1} R^*_t [R^*_t + \Delta^{-1}]^{-1} \right) 
+ \int_{R^n} z' \Delta^{-1} [R^*_t + \Delta^{-1}]^{-1} [R^*_t + \Delta^{-1}]^{-1} \Delta^{-1} z dF(z). \quad (2.30)
\]
Since \( \lim_t R^*_t = R^*_\infty \), there clearly exists a positive constant \( M \) such that for all \( z \) in \( IR^n \),
\[
z' \Delta^{-1} [R^*_t + \Delta^{-1}]^{-1} [R^*_t + \Delta^{-1}]^{-1} \Delta^{-1} z \leq M z' z, \quad t = 0, 1, \ldots \quad (2.31)
\]
and by fact that \( \int_{R^n} z' z dF(z) < \infty \), dominated convergence implies that
\[
\lim_t I'_p (R^*_t) = \text{tr} \left( [R^*_\infty + \Delta^{-1}]^{-1} R^*_\infty [R^*_\infty + \Delta^{-1}]^{-1} \right) 
+ \int_{R^n} z' \Delta^{-1} [R^*_\infty + \Delta^{-1}]^{-1} [R^*_\infty + \Delta^{-1}]^{-1} \Delta^{-1} z dF(z) < \infty. \quad (2.32)
\]
This and (2.22) complete the proof. \( \square \)

The following convergence result is now trivial.

**Theorem 2.3.** Assume that the pair \((A, H)\) is detectable and the pair \((\bar{A}, \bar{C})\) is controllable. Then for any distribution \(F\) in \(E_n\), \( \lim_t \epsilon_t ((A, H, \Gamma), F) = 0 \), and
\[
\limsup_t \frac{1}{t} \ln \epsilon_t ((A, H, \Gamma), F) \leq 2 \ln \lambda_{\max}(K_\infty) < 0. \quad (2.33)
\]
Proof. By combining Theorem 2.2, Proposition 2.1, and (2.21), we see that the claim holds for any distribution $F$ in $D_n$, so by the translation arguments of Section 1, the claim holds for any distribution $F$ in $E_n$.

We can also prove an interesting result in the converse direction. The result stems from the following proposition.

Proposition 2.2. If $F$ is in $D_n$ and $R^*_\infty$ is positive definite and finite and $\liminf_t I_F(R^*_t) = 0$, then $F$ is Gaussian.

Proof. Define a probability measure $\tilde{G}$ on $(IR^n, B(IR^n))$ by the Radon-Nikodym derivative

$$
\frac{d\tilde{G}}{dF}(z) := \frac{\exp\left[-\frac{1}{2}z' R^*_\infty z\right]}{\int_{IR^n} \exp\left[-\frac{1}{2}z' R^*_\infty z\right] dF(z)}, \quad z \in IR^n \tag{2.34}
$$

and let $N$ be the moment generating function of $\tilde{G}$, namely

$$
N(b) := \int_{IR^n} \exp[z'b]d\tilde{G}(z), \quad b \in IR^n \tag{2.35}
$$

Our procedure will be to show that if $\liminf_t I(R^*_t) = 0$, then $N$ satisfies the initial value problem

$$
\nabla_b N(b) = \left[R^*_\infty + \Delta^{-1}\right]^{-1} b N(b), \quad \begin{array}{c} b \in IR^n \tag{2.36} \\
N(0) = 1, \end{array}
$$

which has the unique solution

$$
N(b) = \exp\left[\frac{1}{2}b' \left[R^*_\infty + \Delta^{-1}\right]^{-1} b\right], \quad b \in IR^n \tag{2.37}
$$

which then implies that $\tilde{G}$ is Gaussian. By (2.34), it will then be possible to show that $F$ is Gaussian with zero mean and variance $\Delta$.

Since $R^*_\infty$ is positive definite, there exists $T > 0$ such that $R^*_t$ is also positive definite for $t = T, T + 1, \ldots$, so for $t = T, T + 1, \ldots$, $BR^*_T$ admits a density $f^T_t$ with respect to $\lambda_n$,

$$
f^T_t(b) := \frac{1}{\sqrt{(2\pi)^n\det(R^*_t)}} \exp\left[-\frac{1}{2}z'(R^*_t)^{-1}z\right], \quad b \in IR^n, \ t = T, T + 1, \ldots \tag{2.38}
$$
Define the sequences of mappings \( \{ f_t^1 \}_0^\infty \) and \( \{ f_t^2 \}_0^\infty \) from \( IR^n \) to \( IR \) by

\[
f_t^2(b) := \left\| \int_{IR^n} \left[ z - \left[ R_t^* + \Delta^{-1} \right]^{-1} \right] \gamma(z, b; R_t^*) dF(z) \right\|^2 \quad t = 0, 1, \ldots \quad (2.39)
\]

and

\[
f_t^2(b) := \int_{IR^n} \gamma(z, b; R_t^*) dF(z) \quad t = 0, 1, \ldots \quad (2.40)
\]

for all \( b \) in \( IR^n \). Consequently, under the assumptions made,

\[
\int_{IR^n} \liminf_t f_t^2(b) f_t^1(b) db \leq \liminf_t J(R_t^*) = 0 \quad (2.41)
\]

by Fatou’s Lemma, so that

\[
\liminf_t f_t^2(b) f_t^1(b) = 0 \quad (2.42)
\]

for \( \lambda_n \)-almost all \( b \) in \( IR^n \). Since

\[
\lim_t f_t^1(b) = \frac{1}{\sqrt{(2\pi)^n \det(R_{\infty}^*)}} \exp \left[ -\frac{1}{2} b' \left( R_{\infty}^* \right)^{-1} b \right], \quad b \in IR^n \quad (2.43)
\]

for \( \lambda_n \)-almost all \( b \) in \( IR^n \), we get

\[
\liminf_t f_t^2(b) f_t^1(b) = 0, \quad (2.44)
\]

and therefore,

\[
\frac{\liminf_t f_t^2(b)}{\limsup_t f_t^1(b)} = 0. \quad (2.45)
\]

Now another application of Fatou’s Lemma implies that for any \( b \) in \( IR^n \),

\[
\limsup_t f_t^3(b) \geq \liminf_t f_t^3(b) \geq \int_{IR^n} \gamma(z, b; R_{\infty}^*) dF(z) > 0, \quad (2.46)
\]

since \( R_{\infty}^* \) is finite. From (2.45), we consequently conclude that

\[
\liminf_t f_t^2(b) = 0 \quad (2.47)
\]

for \( \lambda_n \)-almost all \( b \) in \( IR^n \).
Theorem V.1.1 implies that for $\lambda_n$-almost all $b$ in $\mathbb{R}^n$ and all $t = T, T + 1, \ldots$ (i.e., when $B_{R_t^*}$ admits a nonvanishing density with respect to $\lambda_n$),

$$
\int_{\mathbb{R}^n} \|z\| \gamma(z, b; R_t^*) dF(z) < \infty
$$

(2.48)

and

$$
\int_{\mathbb{R}^n} \gamma(z, b; R_t^*) dF(z) < \infty.
$$

(2.49)

Clearly $t \mapsto \gamma(z, b; R_t^*)$ is monotone decreasing for each $z$ and $b$ in $\mathbb{R}^n$, so by dominated convergence, (2.48) and (2.49) imply that

$$
\int_{\mathbb{R}^n} \|z\| \gamma(z, b; R_\infty^*) dF(z) < \infty
$$

(2.50)

and

$$
\int_{\mathbb{R}^n} \gamma(z, b; R_\infty^*) dF(z) < \infty
$$

(2.51)

and

$$
\lim_t \int_{\mathbb{R}^n} \left[ z - [R_t^* + \Delta^{-1}]^{-1} b \right] \gamma(z, b; R_t^*) dF(z)
= \int_{\mathbb{R}^n} z \gamma(z, b; R_\infty^*) dF(z) - [R_\infty^* + \Delta^{-1}]^{-1} b \int_{\mathbb{R}^n} \gamma(z, b; R_\infty^*) dF(z)
$$

(2.52)

for $\lambda_n$-almost all $b$ in $\mathbb{R}^n$. The calculations of Appendix C and (2.47), (2.50), (2.51) and (2.52) imply that for all $b$ in $\mathbb{R}^n$,

$$
\nabla_b N(b) = [R_\infty^* + \Delta^{-1}]^{-1} b N(b).
$$

(2.53)

Note that Lemma C.3 in Appendix C ensures that (2.53) holds for all $b$ in $\mathbb{R}^n$, and not simply for $\lambda_n$-almost all $b$ in $\mathbb{R}^n$. As a result, (2.36) is established, since $N(0) = 1$ is automatic from (2.35).

From (2.36), we know that (2.37) holds, so that $\tilde{G}$ admits a density with respect to $\lambda_n$ given by

$$
\frac{d\tilde{G}}{d\lambda_n}(z) = \frac{1}{\sqrt{(2\pi)^n \det [R_\infty^* + \Delta^{-1}]^{-1}}} \exp \left[ -\frac{1}{2} z' \left[ R_\infty^* + \Delta^{-1} \right] z \right]. \quad z \in \mathbb{R}^n
$$

(2.54)
Since $F$ and $\tilde{G}$ are mutually absolutely continuous by definition (2.34), $F$ admits a density with respect to $\lambda_n$ given by

$$
\frac{dF}{d\lambda_n}(z) = \frac{dF}{d\tilde{G}}(z) \frac{d\tilde{G}}{d\lambda_n}(z) = \frac{\int_{\mathbb{R}^n} \exp \left[ -\frac{1}{2} z' R^*_{\infty} z \right] dF(z)}{\sqrt{(2\pi)^n \det [R^*_{\infty} + \Delta^{-1}]}^{-1}} \exp \left[ -\frac{1}{2} z' \Delta^{-1} z \right].
$$

$z \in \mathbb{R}^n$ (2.55)

Since $F$ is a probability distribution, we can immediately make the identification

$$
\frac{\int_{\mathbb{R}^n} \exp \left[ -\frac{1}{2} z' R^*_{\infty} z \right] dF(z)}{\sqrt{(2\pi)^n \det [R^*_{\infty} + \Delta^{-1}]}^{-1}} = \frac{1}{\sqrt{(2\pi)^n \det \Delta}}.
$$

so that

$$
\frac{dF}{d\lambda_n}(z) = \frac{1}{\sqrt{(2\pi)^n \det \Delta}} \exp \left[ -\frac{1}{2} z' \Delta^{-1} z \right];
$$

$z \in \mathbb{R}^n$ (2.57)

hence $F$ is Gaussian with zero mean and covariance $\Delta$.

The following claims are now trivial.

**Theorem 2.4.** If the pair $(A, H)$ is detectable, the pair $(\tilde{A}, \tilde{C})$ is controllable, $R^*_{\infty}$ is positive-definite, and the matrix $A - [AP_t H + \Sigma w'] [H P_t H' + \Sigma']^{-1} H$ is invertible for each $t = 0, 1, \ldots$, then

$$
\lim \inf_t \frac{1}{t} \ln \epsilon_t ((A, H, \Gamma), F) \geq 2 \ln \lambda_{\min}(K_{\infty})
$$

(2.58)

for any non-Gaussian distribution $F$ in $\mathcal{E}_n$.

**Proof.** The claim holds for any non-Gaussian distribution $F$ in $\mathcal{D}_n$, upon combining Theorem 2.2, Proposition 2.2, and (2.21), so the translation arguments of Section 1 ensure that the claim holds for any distribution $F$ in $\mathcal{E}_n$.

**Theorem 2.5.** If $(A, H)$ is detectable, $(\tilde{A}, \tilde{C})$ is controllable, $\lambda_{\min}(Q_t^* Q_t^*) > 0$ for all $t$ sufficiently large, $R^*_{\infty}$ is positive-definite and

$$
\lim_t \frac{\epsilon_t ((A, H, \Gamma), F)}{\lambda_{\min}(Q_t^* Q_t^*)} = 0,
$$

(2.59)

for some distribution $F$ in $\mathcal{E}_n$, then $F$ is Gaussian.
Proof. Suppose $F$ is in $E_n$ and (2.59) holds. By the transformation (1.7), then

$$\lim_{t} \frac{\epsilon_t ((A, H, \Gamma), G)}{\lambda_{\min}(Q_t^* Q_t^*)} \geq \lim_{t} I_G(t) = 0$$

(2.60)

so by Proposition 2.2, $G$ is Gaussian, and we have that $F$ is Gaussian by inverting the transformation (1.7).
VI.3. The Scalar Case

We now analyze the scalar time-invariant version of system (1.1.1), namely

\[
\begin{align*}
x_{t+1}^o &= ax_t^o + w_{t+1}^o \\
x_0^o &= \xi \\
y_t &= hx_t^o + v_{t+1}^o
\end{align*}
\tag{3.1}
\]

where \(a\) and \(h\) are scalars and

\[
\Gamma_{t+1} := \text{Cov} \begin{pmatrix} w_{t+1}^o \\ v_{t+1}^o \end{pmatrix} = \begin{pmatrix} \sigma_{w} & \sigma_{wv} \\ \sigma_{vw} & \sigma_v \end{pmatrix}.
\tag{3.2}
\]

Define the constants

\[
\bar{a} := a - \frac{\sigma_{wv}}{\sigma_v} h \tag{3.3}
\]

\[
\bar{c} := \sigma_w - \frac{(\sigma_{wv})^2}{\sigma_v} \tag{3.4}
\]

in a manner similar to (2.1) and (2.2).

We study the asymptotic behavior of \(\epsilon_t((a, h, \Gamma), F)\) for large \(t\) as a function of the dynamics \((a, h, \Gamma)\) and of the initial distribution \(F\). We shall find that if \((a, h, \Gamma)\) satisfies a certain stabilizability criterion, then \(\lim_{t} \epsilon_t((a, h, \Gamma), F) = 0\) for any distribution \(F\) in \(\mathcal{E}_1\), and, moreover, the rate at which \(\{\epsilon_t((a, h, \Gamma), F)\}_1^\infty\) decays is independent of \(F\) for any non-Gaussian distribution \(F\) in \(\mathcal{E}_1\). On the other hand, if a corresponding instability criterion is satisfied, then the asymptotic behavior of \(\{\epsilon_t((a, h, \Gamma), F)\}_1^\infty\) strongly depends upon \(F\).

Before proceeding with our analysis, we make the following definitions. The definitions are made for the multivariable case, but their specialization to the scalar case is obvious. We assume, as always, that \(A\) is in \(\mathcal{M}_{n \times n}\) and that \(H\) is in \(\mathcal{M}_{n \times k}\). The first definition is standard in control theory [12, Def. 6.5].

**Definition:** A system \((A, H)\) is said to be stabilizable if all unstable modes (i.e., all modes with eigenvalues of magnitude greater than or equal to 1) are in the controllable subspace.

The second definition is similar.
**Definition:** A system \((A, H)\) is said to be *marginally stabilizable* if all modes which are not stable or critically stable are in the controllable subspace; if all modes with eigenvalues of magnitude greater than 1 are in the controllable subspace.

The main result of this section is contained in the following proposition.

**Theorem 3.1.** The asymptotic behavior of \(\{\epsilon_t((a, h, \Gamma), F)\}_{0}^{\infty}\) is characterized as follows:

1) If the pair \((\bar{a}, \bar{c})\) is marginally stabilizable, then \(\lim_t \epsilon_t((a, h, \Gamma), F) = 0\) for any distribution \(F\) in \(\mathcal{E}_1\). If the pair \((\bar{a}, \bar{c})\) is not marginally stabilizable, then the asymptotic behavior of \(\{\epsilon_t((a, h, \Gamma), F)\}_{0}^{\infty}\) depends nontrivially upon \(F\) in \(\mathcal{E}_1\).

2) If \((\bar{a}, \bar{c})\) is stabilizable, then the sequence \(\{\epsilon_t((a, h, \Gamma), F)\}_{1}^{\infty}\) decays to 0 at an exponential rate independent of \(F\) for non-Gaussian \(F\) in \(\mathcal{E}_1\), whereas if the pair \((\bar{a}, \bar{c})\) is marginally stabilizable but not stabilizable, then this rate depends nontrivially upon \(F\).

Of course, in the second claim, we must specify that \(F\) in \(\mathcal{E}_1\) is non-Gaussian, since for Gaussian distributions \(F\), \(\epsilon_t((a, h, \Gamma), F) = 0\) for all \(t = 1, 2, \ldots\) and all triples \((a, h, \Gamma)\), as the LLSE and MMSE predictors coincide.

Theorem 3.1 is established through a series of technical lemmas which are now developed. Of critical importance in establishing Theorem 3.1 is the evolution of \(\{p_t\}_{0}^{\infty}\), \(\{q_t^*\}_{0}^{\infty}\) and \(\{r_t^*\}_{0}^{\infty}\) as given by (IV.4.8), (IV.4.16) and (IV.4.17), which we now reproduce in simplified scalar notation:

\[
p_{t+1} = a^2 p_t - \frac{(a \sigma + \sigma^w h)^2}{h^2 p_t + \sigma^v} + \sigma^w \quad t = 0, 1, \ldots \quad (3.5)
\]

\[
p_0 = 0,
\]

\[
q_{t+1}^* = \left( \frac{a \sigma^v - \sigma^w h}{h^2 p_t + \sigma^v} \right) q_t^* \quad t = 0, 1, \ldots \quad (3.6)
\]

\[
q_0^* = 1,
\]

and

\[
r_{t+1}^* = r_t^* + \frac{(q_t^*)^2 h^2}{h^2 p_t + \sigma^v} \quad t = 0, 1, \ldots \quad (3.7)
\]

\[
r_0^* = 0.
\]

We now proceed to verify the technical arguments leading to Theorem 3.1. Fix an arbitrary system \((a, h, \Gamma)\).
Lemma 3.1. If \( h = 0 \), then \( \epsilon_t((a, h, \Gamma), F) = 0 \) for all \( t = 1, 2, \ldots \) and for all \( F \) in \( \mathcal{E}_1 \).

Proof. Since \( h = 0 \), we see from (3.7) that \( r_t^* = 0 \) for all \( t = 0, 1, \ldots \) whence for any \( F \) in \( \mathcal{D}_1 \),

\[
\epsilon_t((a, h, \Gamma), F) = q_t^* I_F(0) \quad \text{for} \quad t = 1, 2, \ldots \quad (3.8)
\]

from (1.10). However,

\[
I_F(0) = E' \left[ \frac{\int_{IR} (z - B_0 \sigma^2) \exp[zB_0]dF(z)}{\int_{IR} \exp[zB_0]dF(z)} \right] = \frac{\int_{IR} (z - 0)(1) dF(z)}{\int_{IR} (1)dF(z)} = \frac{\left| \int_{IR} zdF(z) \right|^2}{\int_{IR} (1)dF(z)} = 0,
\]

where the last equality follows from the assumption that \( F \) has zero mean. Thus \( \epsilon_t((a, h, \Gamma), F) = 0 \) for all \( t = 1, 2, \ldots \) and for all \( F \) in \( \mathcal{D}_1 \), so that \( \epsilon_t((a, h, \Gamma), F) = 0 \) for all \( t = 1, 2, \ldots \) for all \( F \) in \( \mathcal{E}_1 \) by the translation arguments of Section 1.

Note that if \( \{w_t^0\}_0^\infty \) and \( \{v_t^0\}_0^\infty \) are independent and thus uncorrelated, Lemma 3.1 is trivial since \( \{x_t^0\}_0^\infty \) and \( \{y_t\}_0^\infty \) are independent, in which case, \( \hat{x}_t = \hat{x}_t^K = a^t\mu \) for all \( t = 1, 2, \ldots \).

Observe now that the pair \((\hat{a}, \hat{c})\) is controllable if and only if \( \sigma^w - \left( \frac{\sigma^w}{\sigma^v} \right)^2 \neq 0 \), or, equivalently, if and only if

\[
\sigma^w \sigma^v \neq (\sigma^{uv})^2. \quad (3.10)
\]

From the previous section we have the following result.

Lemma 3.2. If \( h \neq 0 \) and \((\hat{a}, \hat{c})\) is controllable (i.e., \( \sigma^w \sigma^v \neq (\sigma^{uv})^2 \)), then

\[
\lim_t \epsilon_t((a, h, \Gamma), F) = 0 \quad \text{for all} \quad F \quad \text{in} \quad \mathcal{E}_1. \quad \text{If} \quad \hat{a} = 0, \quad \text{then} \quad \epsilon_t((a, h, \Gamma), F) = 0 \quad \text{for all} \quad t = 1, 2, \ldots \quad \text{and for all} \quad F \quad \text{in} \quad \mathcal{E}_1, \quad \text{while if} \quad \hat{a} \neq 0, \quad \text{then}
\]

\[
\lim_{t} \frac{1}{t} \ln \epsilon_t((a, h, \Gamma), F) = 2 \ln \left| \hat{a} \left( \frac{\sigma^v}{h^2 \rho_{\infty} + \sigma^v} \right) \right| < 0 \quad (3.11)
\]

for all non-Gaussian distributions \( F \) in \( \mathcal{E}_1 \).
Proof. The claim follows directly from Theorems 2.3, 2.4, (1.10) and the fact that

\[ q_{t+1}^* = \bar{a} \frac{\sigma^v}{h^2 p_t + \sigma^v q_t^*} \quad t = 0, 1, \ldots \quad (3.12) \]

\[ q_0^* = 1 \]

which is an alternate representation for (3.6). If \( \bar{a} = 0 \), then \( q_t^* = 0 \) for all \( t = 1, 2, \ldots \), so by (1.10), \( \epsilon_t((a, h, \Gamma), F) = 0 \) for all \( t = 1, 2, \ldots \) and all distributions \( F \) in \( \mathcal{D}_1 \), so \( \epsilon_t((a, h, \Gamma), F) = 0 \) for all \( t = 1, 2, \ldots \) and all distributions \( F \) in \( \mathcal{E}_1 \) by the translation of (1.7). If \( \bar{a} \neq 0 \), then \( q_t^* \neq 0 \) for \( t = 0, 1, \ldots \), and since \( h \neq 0 \), we have \( r_{\infty}^* > 0 \). Thus, by combining Theorems 2.3 and 2.4, we have that (3.11) holds for all non-Gaussian distributions \( F \) in \( \mathcal{D}_1 \) and thus for all non-Gaussian distributions \( F \) in \( \mathcal{E}_1 \).

Now suppose that \( h \neq 0 \) and that the pair \((\bar{a}, \bar{c})\) is uncontrollable. We may then write (3.5) in the simplified form

\[ p_{t+1} = \frac{(a\sqrt{\sigma^v} - h\sqrt{\sigma^w})^2}{h^2 p_t + \sigma^v} p_t \quad t = 0, 1, \ldots \quad (3.13) \]

\[ p_0 = 0, \]

and therefore \( p_t = 0 \) for \( t = 0, 1, \ldots \). Equation (3.6) then yields

\[ q_t^* = \bar{a}^t \quad (3.14) \]

while (3.7) becomes

\[ r_{t+1}^* = \frac{h^2}{\sigma^v} \sum_{s=0}^{t} \bar{a}^{2s} \quad t = 0, 1, \ldots \quad (3.15) \]

\[ r_0^* = 0. \]

The following result is then true.

Lemma 3.3. If \( h \neq 0 \) and \((\bar{a}, \bar{c})\) is stabilizable but uncontrollable (i.e., \(|\bar{a}| < 1\) and \(\sigma^w \sigma^v = (\sigma^{wv})^2\)), then \( \lim_t \epsilon_t((a, h, \Gamma), F) = 0 \), and

\[ \lim_t \frac{1}{t} \ln \epsilon_t((a, h, \Gamma), F) = 2 \ln |\bar{a}| < 0 \quad (3.16) \]

for all non-Gaussian distributions \( F \) in \( \mathcal{E}_1 \).
Proof. From (3.15), we get
\[ r_t^* = \frac{h^2}{\sigma v} \frac{1 - \bar{a}^{2t}}{1 - \bar{a}^2}, \quad t = 0, 1, \ldots \] (3.17)

Since \(|\bar{a}| < 1\), \(\lim_t q_t^* = 0\) and
\[ \lim_t \frac{1}{t} \ln (q_{t+1}^*)^2 = 2 \ln |\bar{a}| < 0 \] (3.18)

and
\[ r_\infty^* = \frac{h^2}{\sigma v} \frac{1}{1 - \bar{a}^2} > 0. \] (3.19)

By Propositions 2.1 and 2.2, we see that \(\limsup_t I_F(r_t^*) < \infty\) and \(\liminf_t I_F(r_t^*) > 0\) for non-Gaussian distributions \(F\) in \(D_1\). Thus, it is straightforward to verify that for non-Gaussian distributions \(F\) in \(D_1\),
\[ \lim_t \epsilon_t ((a, h, \Gamma), F) = 0 \] (3.20)

and
\[ \lim_t \frac{1}{t} \ln \epsilon_t ((a, h, \Gamma), F) = 2 \ln |\bar{a}| < 0, \] (3.21)

and thus (3.21) holds for non-Gaussian \(F\) in \(E_1\), which proves the lemma.

We now turn to the cases where \((a, h, \Gamma)\) is neither marginally stabilizable nor controllable. We shall see that in these cases, the asymptotic behavior of \(\{\epsilon_t ((a, h, \Gamma), F)\}_{t=1}^\infty\) strongly depends upon the initial distribution \(F\). This will be demonstrated by considering the two classes of distribution functions introduced in Section 1 (We restrict ourselves to those distributions \(F\) in \(D_1\).)

Class 1: the distribution \(F\) admits a density with respect to \(\lambda_1\) given by
\[ \frac{dF}{d\lambda}(z) = \sum_{i=1}^n \alpha_i \exp \left[ -\frac{1}{2} \frac{(z - \mu_i)^2}{\rho^2} \right], \quad z \in IR \] (3.22)

where \(\rho > 0, 0 < \alpha_i \leq 1\) for \(i = 1, 2, \ldots, n\), \(\sum_{i=1}^n \alpha_i = 1\), and \(\sum_{i=1}^n \alpha_i \mu_i = 0\), and

Class 2: the RV \(\xi\) takes on a finite number number of values \(z_1 < z_2 \ldots < z_n\) with probabilities \(p_1, p_2, \ldots, p_n\) respectively with \(\sum_{i=1}^n p_i z_i = 0\).
From representation (1.10), we see that the asymptotic behavior of \( \{\epsilon_t((a, h, \Gamma), F)\}_1^\infty \) is determined by the asymptotic behavior of \( \{q_t^*\}_0^\infty \), which is governed by (3.14), and the asymptotic behavior of \( \{I_F(t)\}_0^\infty \). The next three propositions are motivated by this remark.

**Proposition 3.1.** For any distributions \( F \) in \( D_1 \), \( \lim_t I_F(t) = 0 \), and \( \limsup_t I_F(t) < \infty \).

**Proof.** Since \( \{I_F(t)\}_0^\infty \) is independent of the system dynamics \( (a, h, \Gamma) \), to compute its asymptotics, it is convenient to assume that

\[
\begin{pmatrix}
  a & h \\
  \Gamma \\
\end{pmatrix}
= \begin{pmatrix}
  1 & 1 \\
  0 & 0 \\
  0 & 1
\end{pmatrix}. 
\tag{3.23}
\]

With this system,

\[
x_t^o = \xi \\
y_t = \xi + v_{t+1}^o.
\tag{3.24}
\]

By (3.14) and (3.15), \( q_t^* = 1 \) and \( r_t^* = t \) for all \( t = 0, 1, \ldots \), so that

\[
I_F(t) = \epsilon_t((a, h, \Gamma), F). 
\tag{3.25}
\]

Define the IR-valued process \( \{\hat{x}_t\}_0^\infty \) by

\[
\hat{x}_{t+1} := \frac{1}{t+1} \sum_{s=0}^t y_s, \\
\hat{x}_0^o := 0,
\tag{3.26}
\]

and observe from the triangle inequality that for all \( t = 1, 2, \ldots \),

\[
\|\hat{x}_t - \hat{x}_t^K\|^2 \leq (\|\hat{x}_t - x_t^o\|^2 + \|\hat{x}_t^K - x_t^o\|^2)^2 \\
\leq 2 (\|\hat{x}_t - x_t^o\|^2 + \|\hat{x}_t^K - x_t^o\|^2). 
\tag{3.27}
\]

For all \( t = 1, 2, \ldots \), the RV \( \hat{x}_t \) is linear in \( (y_0, y_1, \ldots, y_{t-1}) \) so that

\[
E[\|\hat{x}_t - x_t^o\|^2] \leq E[\|\hat{x}_t^K - x_t^o\|^2] \leq E[\|\hat{x}_t - x_t^o\|^2] 
\tag{3.28}
\]
for $t = 1, 2, \ldots$, by definition of the MMSE and LLSE estimates $\{\hat{x}_t^K\}_{1}^{\infty}$ and $\{\hat{x}_t\}_{1}^{\infty}$. Consequently,

$$
\epsilon_t((a, h, \Gamma), F) \leq 4E \left[ \left\| \hat{x}_t - x_{\tau}^\circ \right\|^2 \right] \\
= 4E \left[ \left\| \frac{1}{t + 1} \sum_{s=0}^{t} u_{s+1}^\circ \right\|^2 \right] = \frac{4}{t + 1} \quad (3.29)
$$

for $t = 1, 2, \ldots$ by elementary calculations. As a result, $\lim_t I_F(t) = 0$ and $\lim \sup_t tI_F(t) < 4$, which proves the proposition.

**Proposition 3.2.** For distributions $F$ of Class 1,

$$
I_F(t) = \frac{K + o_1(\frac{1}{t})}{(\rho^2 t + 1)^2} \quad t > 0 \quad (3.30)
$$

where $\lim_t o_1(\frac{1}{t}) = 0$ and $K \geq 0$. Here $K = 0$ if and only if $F$ is Gaussian, in which case $I_F(t) = 0$ for all $t > 0$.

**Proof.** Appendix D.

**Proposition 3.3.** For distributions $F$ of Class 2,

$$
I_F(t) = \frac{1 + o_1(\frac{1}{t})}{t} \quad t > 0 \quad (3.31)
$$

where $\lim_t o_1(\frac{1}{t}) = 0$.

**Proof.** Appendix E.

We then have the following result.

**Lemma 3.4.** If $h \neq 0$ and $(\bar{a}, \bar{c})$ is marginally stabilizable, but neither stabilizable nor controllable (i.e., $|\bar{a}| = 1$ and $\sigma^w \sigma^v = (\sigma^uv)^2$), then $\lim_t \epsilon_t((a, h, \Gamma), F) = 0$ for any distribution $F$ in $E_1$, the rate of convergence depending nontrivially upon the distribution $F$.

**Proof.** By (3.14) and (3.15), $q_t^* = 1$ and $\tau_t^* = t$ for all $t = 0, 1, \ldots$. By Proposition 3.1, $\lim_t \epsilon_t((a, h, \Gamma), F) = 0$ for all distributions $F$ in $E_1$. The rate of convergence, however, depends upon the rate at which $\{I_F(t)\}_{0}^{\infty}$ converges to 0, and this rate depends upon $F$, as indicated by Propositions 3.2 and 3.3. Indeed, for non-Gaussian distributions $F$ in $D_1$ and Class 1, $\lim_t \ln (\epsilon_t((a, h, \Gamma), F)/\ln t) = -2$, whereas for distributions $F$ in $D_1$ and Class 2, we have that $\lim_t (\ln \epsilon_t((a, h, \Gamma), F)/\ln t) = -1$.  

79
Finally, we make the following claim for systems \((a, h, \Gamma)\) which are neither marginally stabilizable nor controllable.

**Lemma 3.5.** If \(h \neq 0\) and \((\bar{a}, \bar{c})\) is neither marginally stabilizable nor controllable (i.e., \(|\bar{a}| > 1\) and \(\sigma^w \sigma^v = (\sigma^{vw})^2\)), then \(\lim \sup_t \epsilon_t ((a, h, \Gamma), F) < \infty\) for all distributions \(F\) in \(\mathcal{E}_1\), the asymptotic behavior of \(\{\epsilon_t ((a, h, \Gamma), F)\}_{t=1}^{\infty}\) depending nontrivially upon \(F\).

**Proof.** Let \(G\) be given by the transformation (1.7). From (3.15),

\[
\frac{r_t^*}{\frac{\sigma^v}{\bar{a}^2 - 1}} = \frac{h^2}{\sigma^v} \frac{\bar{a}^{2t}}{\bar{a}^2 - 1}, \quad t = 0, 1, \ldots \quad (3.32)
\]

so that

\[
\lim_t \frac{(q_t^*)^2}{r_t^*} = \frac{\sigma^v}{h^2} \left\{ \bar{a}^2 - 1 \right\}. \quad (3.33)
\]

This suggests the representation

\[
\epsilon_t ((a, h, \Gamma), G) = \frac{(q_t^*)^2}{r_t^*} (r_t^* I_G(r_t^*)), \quad t = 0, 1, \ldots \quad (3.34)
\]

from which we conclude

\[
\lim \sup_t \epsilon_t ((a, h, \Gamma), G) = \frac{\sigma^v}{h^2} \left\{ \bar{a}^2 - 1 \right\} \lim \sup_t (r_t^* I_G(r_t^*)). \quad (3.35)
\]

By (3.33) and the second claim of Proposition 3.1, we have that \(\lim \sup_t \epsilon_t ((a, h, \Gamma), G) < \infty\), and so we see that \(\lim \sup_t \epsilon_t ((a, h, \Gamma), F) < \infty\). However, if \(F\) is non-Gaussian and in Class 1, then \(\lim_t \epsilon_t ((a, h, \Gamma), F) = 0\), whereas if \(F\) is in Class 2, then \(\lim_t \epsilon_t ((a, h, \Gamma), F) = 1\). Thus the dependency of the asymptotics of \(\{\epsilon_t ((a, h, \Gamma), F)\}_{t=1}^{\infty}\) on the initial distribution \(F\) is evident.

As we have established the asymptotic behavior of \(\{\epsilon_t ((a, h, \Gamma), F)\}_{t=1}^{\infty}\) for all systems \((a, h, \Gamma)\), Theorem 3.1 is clear.
VI.4. The Significance of the Results of Sections 2 and 3

We now make some observations about the asymptotic agreement of \( \{\hat{X}_t\}_1^\infty \) and \( \{\hat{X}^K_t\}_1^\infty \).

One of the main results of this thesis is Theorem 2.3, which gives sufficient conditions on \( (A, H, \Gamma) \) under which \( \lim_{\epsilon \rightarrow 0} \epsilon \epsilon_t ((A, H, \Gamma), F) = 0 \) for all distributions \( F \) in \( \mathcal{E}_n \). This is notable since by [4, Theorem 5.1] the mean square error \( E \left[ \| X^\infty_t - \hat{X}^\infty_t \|^2 \right] \) converges to a positive-definite limit under the detectability and controllability criterion of Theorem 2.3. In other words, the MMSE and LLSE predictors asymptotically agree with each other, but not with the true state vectors \( \{X^\infty_t\}_1^\infty \). We might conjecture that the controllability and detectability of \( (A, H) \) and \( (\tilde{A}, \tilde{C}) \) (resp.) would force \( \lim_{t \rightarrow \infty} E \left[ \| X^\infty_t - \hat{X}_t \|^2 \right] = 0 \), but Theorem 2.3 shows us that this is not the case; that the MMSE and LLSE estimates both have a mean square error with respect to the true state vector. Another interpretation is that in the \( L^2 \) sense, under the hypothesis of Theorem 2.3, the LLSE predictor is asymptotically optimal, since it asymptotically agrees with the MMSE predictor, which is by definition optimal in the \( L^2 \) sense; asymptotically, we do not gain anything by forming the true conditional expectations as opposed to the wide-sense conditional expectations. This result has obvious practical value, as the Kalman, or linear, estimates, are in general much less costly to generate than the true nonlinear conditional mean; Theorem 2.3 posits that under certain conditions on the plant, we gain nothing asymptotically by forming the nonlinear, as opposed to linear, estimates.

Note that Theorem 3.1 could be extended. If we could establish the form of the dependency of the asymptotic behavior of \( \{I_F(t)\}_{t \geq 0} \) upon the distribution \( F \) for any distribution \( F \) in \( \mathcal{D}_1 \), we would know the asymptotic behavior of \( \epsilon_t ((A, H, \Gamma), F) \) for any system \( (a, h, \Gamma) \) and any initial distribution \( F \) in \( \mathcal{E}_1 \).
Appendix A. The Kalman Filtering Equations

In this appendix, we provide the Kalman filtering equations necessary to generate the estimates \( \{(X_{t+1}^T, Y_{t+1}^T)\}_0^\infty \) from (III.3.1) and (IV.4.1). This appendix is taken directly from [12, Theorem 6.42].

Consider the system

\[
X_{t+1} = A_t X_t + W_{t+1} \\
X_0 = \xi \\
Y_t = H_t X_t + V_{t+1} 
\]

defined on a probability triple \((\Omega, \mathcal{F}, P)\) which carries the \( IR^n \)-valued plant process \( \{X_t\}_0^\infty \) and the \( IR^k \)-valued observation process \( \{Y_t\}_0^\infty \). Assume that \( \{(W_{t+1}, V_{t+1})\}_0^\infty \) is a zero-mean GWN process with covariance structure

\[
\Sigma_{t+1} = \text{Cov}\left( \begin{array}{c} W_{t+1} \\ V_{t+1} \end{array} \right) = \begin{pmatrix} \Sigma_{t+1}^{ww} & \Sigma_{t+1}^{wv} \\ \Sigma_{t+1}^{vw} & \Sigma_{t+1}^{vv} \end{pmatrix} 
\]

where \( \Sigma_{t+1}^{wv} \) is positive definite for \( t = 0, 1, \ldots \). Further assume that \( \xi \) is a Gaussian RV with mean \( \mu \) and covariance matrix \( \Delta \) and independent of the process \( \{(W_{t+1}, V_{t+1})\}_0^\infty \).

For \( t = 0, 1, \ldots \), define \( \hat{X}_{t+1} := E[X_{t+1}|Y_s; s = 0, 1, \ldots, t] \), \( E[\cdot | \cdot] \) being the conditional expectation operator associated with \( P \). Set \( \hat{X}_0 := \mu \) for convenience.

Define the deterministic processes \( \{P_t\}_0^\infty \) in \( Q_n \) by the recursion

\[
P_{t+1} = A_t P_t A'_t - [A_t P_t H'_t + \Sigma_{t+1}^{ww}][H_t P_t H'_t + \Sigma_{t+1}^{vv}]^{-1}[A_t P_t H'_t + \Sigma_{t+1}^{ww}]' + \Sigma_{t+1}^{wv} \\
P_0 = \Delta 
\]

\( t = 0, 1, \ldots \) (A.3)

(this is an alternate representation for [12, Eq. 6.435]). Then

\[
E\left[ (X_{t+1} - \hat{X}_{t+1})(X_{t+1} - \hat{X}_{t+1})' \right] = P_{t+1}, \\
t = 0, 1, \ldots \quad (A.4)
\]
and \( \{\hat{X}_t\}_0^\infty \) obeys the recursion

\[
\hat{X}_{t+1} = [A_t - [A_t P_t H'_t + \Sigma_{t+1}^w][H_t P_t H'_t + \Sigma_{t+1}^v]^{-1} H_t] \hat{X}_t \\
+ [A_t P_t H'_t + \Sigma_{t+1}^w][H_t P_t H'_t + \Sigma_{t+1}^v]^{-1} Y_t
\]

\( \hat{X}_0 = \mu. \)

t = 0, 1, \ldots \quad (A.5)
Appendix B. Equations (VI.2.7) and (VI.2.8)

In this appendix, we prove (VI.2.7) and (VI.2.8). Specifically, let the sequence \( \{Q_t\}_0^\infty \) in \( \mathcal{M}_{n \times n} \) be given by
\[
Q_{t+1} = K_tQ_t \quad t = 0, 1, \ldots \tag{B.1}
\]
where \( \{K_t\}_{-1}^\infty \) is some sequence in \( \mathcal{M}_{n \times n} \) and \( K_\infty := \lim K_t \) is well-defined. We then wish to show that
\[
\limsup_t \frac{1}{t} \ln \lambda_{\max}(Q_t') \leq 2 \ln \lambda_{\max}(K_\infty) \tag{B.2}
\]
and that if the matrix \( K_t \) is invertible for each \( t = -1, 0, 1, \ldots \), then
\[
\liminf_t \frac{1}{t} \ln \lambda_{\min}(Q_t') \geq 2 \ln \lambda_{\min}(K_\infty). \tag{B.3}
\]

We begin with an auxiliary result.

**Lemma B.1.** We have
\[
\limsup_t \frac{1}{t} \ln \|Q_t\|_o \leq \ln \lambda_{\max}(K_\infty), \tag{B.4}
\]
where \( \| \cdot \|_o \) is the operator norm on \( \mathcal{M}_{n \times n} \).

**Proof.** From an elementary result in functional analysis [28, Thm 8.2.3], it is known that
\[
\lim_n \|K^n\|_o^{1/n} = \lambda_{\max}(K). \tag{B.5}
\]
for any matrix \( K \) in \( \mathcal{M}_{n \times n} \). Now take \( \epsilon > 0 \). Then there exists an integer \( N \) such that
\[
\|K_\infty^N\|_o^{1/N} \leq \lambda_{\max}(K_\infty) + \epsilon/2, \tag{B.6}
\]
and by continuity arguments, there exists an integer \( l \) such that for \( m \geq 1N \),
\[
\|K_{N+m} \cdots K_{m+1}K_m\|_o^{1/N} \leq \lambda_{\max}(K_\infty) + \epsilon \tag{B.7}
\]
and
\[
\|K_m\|_o \leq \|K_\infty\|_o + \epsilon. \tag{B.8}
\]

84
For every integer $t = 0, 1, \ldots$ define $p_t := \lfloor t/N \rfloor - l$, where $\lfloor \cdot \rfloor$ is the integer floor function, and observe that for $t = l, l + 1, \ldots$,

$$
\|Q_{t+1}\|_{\circ} \leq \|K_t \cdots K_{(p_t+1)N+1} K_{(p_t+1)N}\|_{\circ}
 \left\{ \prod_{q=0}^{p_t-1} \|K_{(q+l+1)N-1} \cdots K_{(q+l)N+1} K_{(q+l)N}\|_{\circ} \right\} \|Q_{lN-1}\|_{\circ}
 \leq \left( \|K_{\infty}\|_{\circ} + \epsilon \right)^{t-(p_t+1)N+1}
 \left\{ \prod_{q=0}^{p_t-1} \|K_{(q+l+1)N-1} \cdots K_{(q+l)N+1} K_{(q+l)N}\|_{\circ} \right\} \|Q_{lN-1}\|_{\circ},
$$

(B.9)

and

$$
\frac{1}{t} \ln \|Q_t\|_{\circ} \leq \frac{t-(p_t+1)N+1}{t} \ln (\|K_{\infty}\|_{\circ} + \epsilon)
 + \frac{(p_t+1)N}{t} \frac{p_t N}{(p_t+1)N} \frac{1}{p_t} \sum_{q=0}^{p_t-1} \frac{1}{N} \ln \|K_{(q+l+1)N-1} \cdots K_{(q+l)N+1} K_{(q+l)N}\|_{\circ}
 + \frac{1}{t} \ln \|Q_{lN-1}\|_{\circ}
 \leq \frac{t-(p_t+1)N+1}{t} \ln (\|K_{\infty}\|_{\circ} + \epsilon)
 + \frac{(p_t+1)N}{t} \frac{p_t N}{(p_t+1)N} \ln (\lambda_{\max}(K_{\infty}) + \epsilon)
 + \frac{1}{t} \ln \|Q_{lN-1}\|_{\circ}.
$$

(B.11)

Consequently,

$$
\limsup_t \frac{1}{t} \ln \|Q_t\|_{\circ} \leq \ln \lambda_{\max}(\|K_{\infty}\|_{\circ} + \epsilon)
$$

(B.13)

which proves (B.4), since $\epsilon > 0$ is arbitrary.

By the equivalence of norms, (B.2) follows.

Theorem B.1. Equation (B.2) holds.

Proof. It is well-known that the function $\| \cdot \|_{i,2} : \mathcal{M}_{n \times n} \to IR$ given by

$$
\|K\|_{i,2} := \sqrt{\lambda_{\max}(K'K)} \quad K \in \mathcal{M}_{n \times n}
$$

(B.14)

is a norm on $\mathcal{M}_{n \times n}$ (see [26, Table 3.1]). Thus, by the equivalence of norms on $IR^{n \times n}$, there exists a positive constant $M$ such that $\|K\|_{i,2} \leq M \|K\|_{\circ}$ for all $K$ in $\mathcal{M}_{n \times n}$.
Consequently,
\[ \limsup_t \frac{1}{t} \ln \lambda_{\max}(Q'_t Q_t) \leq \limsup_t \frac{1}{t} \ln (M^2 \|Q_t\|_o^2) \]
\[ = \limsup_t \frac{1}{t} \ln \|Q_t\|_o^2 \]  
\[ \leq 2 \ln \lambda_{\max}(K_\infty), \]  
\[
(B.15)\]

using Lemma B.1.

We can now verify

**Theorem B.2.** Equation (B.3) holds if the matrix $K_t$ is invertible for each $t = -1, 0, 1, \ldots$.

**Proof.** Since $K_t$ is invertible for each $t = -1, 0, 1, \ldots$, then $Q_t$ is invertible for each $t = 0, 1, \ldots$. Thus we may form the sequence
\[ \hat{Q}_{t+1} = (K_t^{-1})' \hat{Q}_t, \]
\[ \hat{Q}_0 = (K_{-1}^{-1})', \]
\[
(B.16)\]
and observe that $\hat{Q}_t = (Q_t^{-1})'$ for all $t = 0, 1, \ldots$.

Now elementary considerations reveal that
\[ \lambda_{\max}(K) = \frac{1}{\lambda_{\min}(K^{-1})} \]
for any invertible matrix $K$ in $M_{n \times n}$ (consider $K$ in Jordan canonical form and invert by Cramer’s Rule). We thus have
\[ \ln \lambda_{\min}(Q'_t Q_t) = \ln \left( \frac{1}{\lambda_{\max}} ((Q'_t Q_t)^{-1}) \right) \]
\[ = -\ln \lambda_{\max} (Q_t^{-1} (Q_t^{-1})') \]
\[ = -\ln \lambda_{\max} (\hat{Q}_t' \hat{Q}_t) \]  
\[
(B.17)\]
so
\[ \liminf_t \frac{1}{t} \ln \lambda_{\min}(Q'_t Q_t) = -\limsup_t \frac{1}{t} \ln \lambda_{\max}(\hat{Q}_t' \hat{Q}_t) \]
\[ \geq -2 \ln \lambda_{\max}(K_\infty) \]
\[ = 2 \ln \lambda_{\min}(K_\infty) \]  
\[
(B.18)\]
by applying Lemma B.1 to $\{\hat{Q}_t\}_{0}^{\infty}$ and using (B.17).
Appendix C. Proposition VI.2.2

The proof of Theorem Proposition VI.2.2 requires several technical lemmas which we present here. In the forthcoming discussion, let $Q$ represent any fixed probability measure on $(IR^n, B(IR^n))$. Let $f : IR^n \mapsto [0, \infty]$ be a Borel-measurable mapping and define $M : IR^n \mapsto [0, \infty]$ by

$$M(b) := \int_{IR^n} f(z) \exp[z'b]dQ(z). \quad b \in IR^n \quad (C.1)$$

Then we can make the following claim.

**Lemma C.1.** If $M(b) < \infty$ for $\lambda_n$-almost all $b$ in $IR^n$, then $M(b) < \infty$ for all $b$ in $IR^n$.

**Proof.** Fix $b$ in $IR^n$. The hypotheses of the claim then ensure that there exists $v$ in $IR^n$ and positive constants $\alpha$ and $\beta$ such that $M(b + \alpha v) < \infty$ and $M(b - \beta v) < \infty$. Now define the sets $A$ and $B$ in $IR^n$ by $A := \{z : z'v \geq 0\}$ and $B := \{z : z'v < 0\}$. It is clear that for all $z$ in $IR^n$,

$$1_A(z) \exp[z'b] \leq 1_A(z) \exp[z'(b + \alpha v)] \leq \exp[z'(b + \alpha v)] \quad (C.2)$$

and

$$1_B(z) \exp[z'b] \leq 1_B(z) \exp[z'(b - \beta v)] \leq \exp[z'(b - \beta v)], \quad (C.3)$$

so that

$$M(b) = \int_A f(z) \exp[z'b]dQ(z) + \int_B f(z) \exp[z'b]dQ(z) \quad (C.4)$$

$$\leq M(b + \alpha v) + M(b - \beta v) < \infty,$$

and the result is proved.

The next proposition contains a similar result.

**Lemma C.2.** If $M(b) < \infty$ for all $b$ in $IR^n$, then

$$\int_{IR^n} f(z) \exp[\mu||z||]dQ(z) < \infty \quad (C.5)$$

for all $\mu \geq 0$.

87
**Proof.** Fix $\mu > 0$ and let $| \cdot |$ be the norm on $\mathbb{R}^n$ defined by

$$|z| := \sum_{j=1}^{n} |z_j|, \quad z = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n \quad (C.6)$$

so that

$$\int_{\mathbb{R}^n} f(z) \exp [\mu |z|] \, dQ(z) = \int_{\mathbb{R}^n} f(z) \exp \left[ \mu \sum_{j=1}^{n} z_j \text{sgn}(z_j) \right] \, dQ(z) \quad (C.7)$$

where for all $x$ in $\mathbb{R}^n$,

$$\text{sgn}(x) := \begin{cases} 1, & \text{if } x \geq 0 \\ -1, & \text{if } x < 0. \end{cases} \quad (C.8)$$

By partitioning $\mathbb{R}^n$, we have that, for all $z = (z_1, z_2, \ldots, z_n)$ in $\mathbb{R}^n$, the relation

$$1 = \sum_{p_1 \in \{-1,1\}} \sum_{p_2 \in \{-1,1\}} \cdots \sum_{p_n \in \{-1,1\}} \prod_{l=1}^{n} 1_{\{z_l \text{sgn}(z_l) = p_l\}} \quad (C.9)$$

holds, so that

$$\exp [\mu |z|] = \sum_{p_1 \in \{-1,1\}} \sum_{p_2 \in \{-1,1\}} \cdots \sum_{p_n \in \{-1,1\}} \left( \prod_{l=1}^{n} 1_{\{z_l \text{sgn}(z_l) = p_l\}}(z) \right) \exp \left[ \mu \sum_{k=1}^{n} z_k \text{sgn}(z_k) \right]$$

$$= \sum_{p_1 \in \{-1,1\}} \sum_{p_2 \in \{-1,1\}} \cdots \sum_{p_n \in \{-1,1\}} \left( \prod_{l=1}^{n} 1_{\{z_l \text{sgn}(z_l) = p_l\}}(z) \exp \left[ \mu \sum_{k=1}^{n} (p_k z_k) \right] \right)$$

$$\leq \sum_{p_1 \in \{-1,1\}} \sum_{p_2 \in \{-1,1\}} \cdots \sum_{p_n \in \{-1,1\}} \exp \left[ \mu \sum_{k=1}^{n} (p_k z_k) \right] \quad (C.10)$$

for all $z = (z_1, z_2, \ldots, z_n)$ in $\mathbb{R}^n$. Now the finiteness assumption implies that

$$\int_{\mathbb{R}^n} f(z) \exp [\mu |z|] \, dQ(z) \leq \sum_{p_1 \in \{-1,1\}} \sum_{p_2 \in \{-1,1\}} \cdots \sum_{p_n \in \{-1,1\}} \int_{\mathbb{R}^n} f(z) \exp [(\mu p)'z] \, dQ(z) < \infty \quad (C.11)$$

where $p = (p_1, p_2, \ldots, p_n)$. From the equivalence of norms on $\mathbb{R}^n$, we know that there exists a positive constant $N$ with the property that

$$||z|| \leq N|z| \quad (C.12)$$
for all $z$ in $\mathbb{R}^n$, whence

$$\int_{\mathbb{R}^n} f(z) \exp[\mu ||z||] \, dQ(z) \leq \int_{\mathbb{R}^n} \exp[\mu N ||z||] \, dQ(z) < \infty \quad (C.13)$$

by (C.11).

We now give conditions guaranteeing the continuity of the mapping $b \mapsto M(b)$.

**Lemma C.3.** If for all $\mu \geq 0$,

$$\int_{\mathbb{R}^n} f(z) \exp[\mu ||z||] \, dQ(z) < \infty, \quad (C.14)$$

then the mapping $b \mapsto M(b)$ from $\mathbb{R}^n$ to $\mathbb{R}$ is continuous.

**Proof.** Fix $b$ in $\mathbb{R}^n$ and take $\{v_m\}_0^\infty$ in $\mathbb{R}^n$ with $\lim_m v_m = 0$. For all $z$ in $\mathbb{R}^n$, it is plain that

$$0 \leq \sup_m [f(z) \exp[z'(b + v_m)]] \leq f(z) \exp[||z|| (||b|| + \sup_m ||v_m||)], \quad (C.15)$$

and $\lim_m M(b + v_m) = M(b)$ by dominated convergence, since

$$\int_{\mathbb{R}^n} f(z) \exp[||z|| (||b|| + \sup_m ||v_m||)] \, dQ(z) < \infty \quad (C.16)$$

by virtue of (C.14).

We now present simple conditions under which the moment generating function of $Q$ may be differentiated under the integral.

**Theorem C.1.** If for all $b$ in $\mathbb{R}^n$, the finiteness conditions

$$\int_{\mathbb{R}^n} \exp[z' b] \, dQ(z) < \infty \quad (C.17)$$

and

$$\int_{\mathbb{R}^n} ||z|| \exp[z' b] \, dQ(z) < \infty, \quad (C.18)$$

hold, then

$$\nabla_b \int_{\mathbb{R}^n} \exp[z' b] \, dQ(z) = \int_{\mathbb{R}^n} z \exp[z' b] \, dQ(z). \quad (C.19)$$
for all \( b \) in \( IR^n \).

**Proof.** For convenience, define the mappings \( N_1 : IR^n \mapsto IR \) and \( N_2 : IR^n \mapsto IR^n \) by

\[
N_1(b) := \int_{IR^n} \exp[z'b]dQ(z) \quad b \in IR^n \quad (C.20)
\]

and

\[
N_2(b) := \int_{IR^n} z \exp[z'b]dQ(z) \quad b \in IR^n \quad (C.21)
\]

Fix \( b \) in \( IR^n \). We must then show that

\[
\lim_{v \to 0} \left| \frac{N_1(b + v) - N_1(b) - v'N_2(b)}{\|v\|} \right| = 0, \quad (C.22)
\]

or, alternately, in view of the integrability conditions (C.17) and (C.18), that

\[
\lim_{v \to 0} \left| \int_{IR^n} f_v(z) \exp[z'b]dQ(z) \right| = 0 \quad (C.23)
\]

where

\[
f_v(z) := \frac{\exp[z'v] - 1 - z'v}{\|v\|} \quad z \in IR^n, \ v \in IR^n \sim \{0\} \quad (C.24)
\]

We shall in fact prove that

\[
\lim_{v \to 0} \int_{IR^n} |f_v(z)| \exp[z'b]dQ(z) = 0, \quad b \in IR^n \quad (C.25)
\]

which obviously implies (C.23).

By elementary results, the inequality

\[
1 - \exp(-x) \leq x \quad (C.26)
\]

holds for all \( x \geq 0 \). Consequently, for all \( v \neq 0 \) in \( IR^n \) and \( z \) in \( IR^n \),

\[
f_v(z) \leq \frac{\exp[\|z\|\|v\|] - 1 + \|z\|\|v\|}{\|v\|}
\]

\[
= \exp[\|z\|\|v\|] \frac{1 - \exp[-\|z\|\|v\|] + \|z\|\|v\| \exp[-\|z\|\|v\|]}{\|v\|}
\]

\[
\leq 2\|z\| \exp[\|z\|\|v\|],
\]

90
while

$$f_v(z) \geq \frac{\exp[-\|z\||v|] - 1 - \|z\||v|}{\|v\|} \geq -2\|z\|, \tag{C.28}$$

and therefore

$$|f_v(z)| \leq 2\|z\| \exp[\|z\||v||]. \tag{C.29}$$

Now for any \( v \neq 0 \) in \( IR^n \) such that \( \|v\| \leq 1 \), it is a simple matter to see that

$$2 \int_{IR^n} \|z\| \exp[\|z\||v||] \exp[z'b] dQ(z) \leq 2 \int_{IR^n} \|z\| \exp[\|z\|(|b| + 1)] dQ(z) < \infty \tag{C.30}$$

where the finiteness is obtained from Lemma C.2 (with \( f : z \mapsto \|z\| \) and (C.18)). Then (C.25) follows by dominated convergence.
Appendix D. Asymptotic Behavior of $I_F$ for Distributions $F$ in Class 1

In this appendix, we consider the asymptotic behavior of $\{I_F(t)\}_{t>0}$ where the distribution $F$ is in Class 1—where $F$ admits a density with respect to $\lambda_1$ given by

$$\frac{dF}{d\lambda_1}(z) = \sum_{i=1}^{n} \alpha_i \exp \left[ -\frac{1}{2} \left( \frac{z - \mu_i}{\rho} \right)^2 \right], \quad z \in IR \quad (D.1)$$

where $\rho > 0$, $0 < \alpha_i \leq 1$ for $i = 1, 2, \ldots, n$, $\sum_{i=1}^{n} \alpha_i = 1$, and $\sum_{i=1}^{n} \alpha_i \mu_i = 0$.

A straightforward calculation reveals that

$$\sigma^2 = \rho^2 + \sum_{i=1}^{n} \alpha_i \mu_i^2. \quad (D.2)$$

We first manipulate equation (VI.1.11) in order to facilitate the analysis. Define collections of mappings $\{N_t\}_{t>0}$ and $\{D_t\}_{t>0}$ from IR to IR by

$$N_t(b) := \int_{IR} \left( z - \frac{b}{t + 1/\sigma^2} \right) \exp[z b - z^2 t/2] dF(z) \quad (D.3)$$

$$D_t(b) := \int_{IR} \exp[z b - z^2 t/2] dF(z) \quad (D.4)$$

for all $b$ in IR and $t > 0$. Then (VI.1.11) may be written as

$$I_F(t) := E' \left[ \frac{N_t^2(B_t)}{D_t(B_t)} \right]$$

$$= E' \left[ \frac{N_t(B_t)}{D_t(B_t)} \right]^2 D_t(B_t) \quad (D.5)$$

$$= \int_{IR} \left| \frac{N_t(b)}{D_t(b)} \right|^2 D_t(b) \frac{\exp[-b^2/(2t)]}{\sqrt{2\pi t}} db$$

and, after a change of variables, we get

$$I_F(t) = \int_{IR} \left| \frac{N_t(u t)}{D_t(u t)} \right|^2 D_t(u t) \frac{\exp[-u^2 t/2]}{\sqrt{2\pi/t}} du \quad (D.6)$$

for $t > 0$. Now let the mappings $\{K_t\}_{t>0}$ and $\{M_t\}_{t>0}$ from $[0, \infty)$ to IR be given by

$$K_t(u) := \frac{N_t(u t)}{D_t(u t)} \quad u \in IR, t > 0 \quad (D.7)$$
and
\[ M_t(u) := D_t(ut) \frac{\exp \left(-u^2 t/2\right)}{\sqrt{2\pi t}}, \quad u \in IR, \ t > 0 \quad (D.8) \]
so that for all \( u \) in \( IR \) and \( t > 0 \), we have
\[ I_P(t) = \int_{IR} K_t^2(u)M_t(u)du. \quad (D.9) \]

Now by direct evaluation, it is easy to verify that for all \( b \) in \( IR \) and \( t > 0 \),
\[ N_t(b) = \frac{1}{(\rho^2 t + 1)^{\frac{n}{2}}} \sum_{i=1}^{n} \alpha_i \left[ b^2 \rho^2 - \sigma^2 \sigma^2 t + 1 + \mu_i \right] \exp \left[ \frac{1}{2} \frac{b^2}{\rho^2 t + 1} + \frac{b \mu_i}{\rho^2 t + 1} - \frac{1}{2} \frac{\mu_i^2}{\rho^2 + 1/t} \right] \quad (D.10) \]
and
\[ D_t(b) = \frac{1}{(\rho^2 t + 1)^{\frac{n}{2}}} \sum_{i=1}^{n} \alpha_i \exp \left[ \frac{1}{2} \frac{b^2}{\rho^2 t + 1} + \frac{b \mu_i}{\rho^2 t + 1} - \frac{1}{2} \frac{\mu_i^2}{\rho^2 + 1/t} \right]. \quad (D.11) \]

After some tedious algebraic manipulation, we find that for all \( u \) in \( IR \) and \( t > 0 \),
\[ K_t(u) = \frac{1}{\rho^2 t + 1} \left[ \frac{u^2}{\sigma^2 + 1/t} + J_t(u) \right] \quad (D.12) \]
and
\[ M_t(u) = \sum_{i=1}^{n} \frac{\alpha_i}{\sqrt{2\pi(\rho^2 + 1/t)}} \exp \left[ \frac{-1}{2} \frac{(u - \mu_i)^2}{\rho^2 + 1/t} \right] du, \quad (D.13) \]
with
\[ J_t(u) := \sum_{i=1}^{n} \alpha_i \mu_i \exp \left[ \frac{-1}{2} \frac{(u - \mu_i)^2}{\rho^2 + 1/t} \right]. \quad (D.14) \]

Motivated by the forms of (D.9) and (D.13), we define the mappings \( \{I_i^*\}_1^n \) from \([0, \infty)\) to \([0, \infty)\) by
\[ I_i^*(t) := \int_{IR} K_t^2(u) \frac{\exp \left[-\frac{1}{2} \frac{(u - \mu_i)^2}{\rho^2 + 1/t} \right]}{\sqrt{2\pi(\rho^2 + 1/t)}} du \]
\[ = \int_{IR} \left[ \frac{u^2}{\sigma^2 + 1/t} + J_t(u) \right]^2 \frac{\exp \left[-\frac{1}{2} \frac{(u - \mu_i)^2}{\rho^2 + 1/t} \right]}{\sqrt{2\pi(\rho^2 + 1/t)}} du, \quad i = 1, 2, \ldots, n, \ t > 0 \quad (D.15) \]
so that for all \( t > 0 \),
\[
I_F(t) = \frac{1}{(\rho^2 t + 1)^2} \sum_{i=1}^{n} \alpha_i I^i_F(t).
\] (D.16)

Fix \( i = 1, 2, \ldots, n \). By a change of variables, we see that for all \( t > 0 \),
\[
I^i_F(t) = \int_{\mathbb{R}} \left\{ (\rho^2 - \sigma^2) \frac{v^2 \rho^2 + 1/t + \mu_i}{\sigma^2 + 1/t} + J_i(v \sqrt{\rho^2 + 1/t + \mu_i}) \right\}^2 \frac{\exp[-v^2/2]}{\sqrt{2\pi}} dv;
\] (D.17)

we now verify that the integration and limiting operations may be exchanged in (D.17) in order to find \( \lim_t I^i_F(t) \).

Clearly \( J_\infty(u) := \lim_t J_t(u) \) is well-defined for all \( u \) in \( \mathbb{R} \), with
\[
J_\infty(u) = \frac{\sum_{i=1}^{n} \alpha_i \mu_i \exp \left[ -\frac{1}{2} \frac{(u - \mu_i)^2}{\rho^2} \right]}{\sum_{i=1}^{n} \alpha_i \exp \left[ -\frac{1}{2} \frac{(u - \mu_i)^2}{\rho^2} \right]} \] (D.18)

for all \( u \) in \( \mathbb{R} \). Also observe that the mapping \((u, s) \mapsto J_{1/s}(u)\) is jointly continuous in \( u \) and \( s \) on \( \{(u, s) : u \in \mathbb{R}, s \geq 0\} \), so that for all \( v \) in \( \mathbb{R} \),
\[
\lim_t J_t(v \sqrt{\rho^2 + 1/t + \mu_i}) = J_\infty(v \rho + \mu_i)
\] (D.19)

and
\[
\lim_t \left\{ (\rho^2 - \sigma^2) \frac{v \sqrt{\rho^2 + 1/t + \mu_i}}{\sigma^2 + 1/t} + J_i(v \sqrt{\rho^2 + 1/t + \mu_i}) \right\} = (\rho^2 - \sigma^2) \frac{v \rho + \mu_i}{\sigma^2} + J_\infty(v \rho + \mu_i).
\] (D.20)

Inspection of (D.14) reveals that
\[
\sup_{\substack{u \in \mathbb{R} \\text{ for } t > 0}} |J_t(u)| \leq B
\] (D.21)

where \( B \geq \max_{i=1,2,\ldots,n} |\mu_i| \), so that by elementary analysis, it is not difficult to verify that there exist positive numbers \( B_1 \) and \( B_2 \) such that for all \( v \) in \( \mathbb{R} \) and \( t > 0 \),
\[
|(\rho^2 - \sigma^2) \frac{v \sqrt{\rho^2 + 1/t + \mu_i}}{\sigma^2 + 1/t} + J_i(v \sqrt{\rho^2 + 1/t + \mu_i})| \leq B_1 |v| + B_2.
\] (D.22)
Clearly
\[ \int_{\mathbb{R}} (B_1|v| + B_2)^2 \frac{\exp[-v^2/2]}{\sqrt{2\pi}} dv < \infty, \quad (D.23) \]
so that by dominated convergence and (D.20), we get
\[ \lim_t I_F^i(t) = \int_{\mathbb{R}} \left\{ (\rho^2 - \sigma^2) \frac{\nu \rho + \mu_i}{\sigma^2} + J_\infty(\nu \rho + \mu_i) \right\}^2 \frac{\exp[-v^2/t]}{\sqrt{2\pi}} dv. \quad (D.24) \]
By a change of variables, (D.24) becomes
\[ \lim_t I_{F_1}^i(t) = \int_{\mathbb{R}} \left\{ u \frac{\rho^2 - \sigma^2}{\sigma^2} + J_\infty(u) \right\}^2 \frac{\exp[-(u - \mu_i)^2/(2\rho^2)]}{\sqrt{2\pi \rho^2}} du, \quad (D.25) \]
so that returning to (D.16), we see that
\[ \lim_t \{(\rho^2 t + 1)^2 I_F(t)\} = \int_{\mathbb{R}} \left\{ u \frac{\rho^2 - \sigma^2}{\sigma^2} + J_\infty(u) \right\}^2 \sum_{i=0}^{\infty} \alpha_i \frac{\exp[-(u - \mu_i)^2/(2\rho^2)]}{\sqrt{2\pi \rho^2}} du \]
\[ = \int_{\mathbb{R}} \left\{ u \frac{\rho^2 - \sigma^2}{\sigma^2} + J_\infty(u) \right\}^2 dF(u). \quad (D.26) \]
We state Equation (D.26) as

**Theorem D.1.** We have
\[ \lim_t \{(\rho^2 t + 1)^2 I_F(t)\} = \int_{\mathbb{R}} \left\{ u \frac{\rho^2 - \sigma^2}{\sigma^2} + J_\infty(u) \right\}^2 dF(u). \quad (D.27) \]

Clearly \( \lim_t \{(\rho^2 t + 1)^2 I_F(t)\} \) is nonnegative. We now must determine conditions under which \( \lim_t \{(\rho^2 t + 1)^2 I_F(t)\} = 0 \). This is easily accomplished in the following theorem. Recall that \( F \) is in \( D_1 \).

**Theorem D.2.** We have \( \lim_t \{(\rho^2 t + 1)^2 I_F(t)\} = 0 \) if and only if \( F \) is Gaussian—if and only if \( \mu_i = 0 \) for \( i = 1, 2, \ldots, n \).

**Proof.** **Sufficiency:** If \( F \) is Gaussian, then, since \( \mu_i = 0 \) for \( i = 1, 2, \ldots, n \), we have that \( J_i(u) = 0 \) for all \( u \) in \( \mathbb{R} \) and all \( t > 0 \) by inspection of (D.14). By virtue of (D.2), we also have \( \sigma^2 = \rho^2 \), so by (D.12), \( K_i(u) = 0 \) for all \( u \) in \( \mathbb{R} \) and \( t > 0 \). By (D.9), it
follows that $I_F(t) = 0$ for all $t > 0$, so $\lim_t \{(\rho^2 t + 1)^2 I_F(t)\} = 0$. Alternately, we may directly argue that if $F$ is Gaussian, then the MMSE and LLSE predictors coincide, so that $e_t((a, h, \Gamma), F) = 0$ for all systems $(a, h, \Gamma)$, and necessarily $I_F(t) = 0$ for all $t > 0$.

**Necessity:** Suppose $\lim_t \{(\rho^2 t + 1)^2 I_F(t)\} = 0$. From Theorem D.1, we see that by Fatou’s Lemma that

$$u \left( \frac{\rho^2 - \sigma^2}{\sigma^2} \right) + J_\infty(u) = 0$$  \hspace{1cm} (D.28)

for $\lambda_1$-almost all $u$ in $\mathcal{M}$. But we know from (D.21) that $J_\infty$ is bounded, so that necessarily $\rho^2 = \sigma^2$. From (D.2), we conclude that $\mu_i = 0$ for all $i = 1, 2, \ldots, n$, proving the necessity.

The results of this section may now be collected as a theorem.

**Theorem D.3.** We have

$$I_F(t) = \frac{K + o_1(\frac{1}{t})}{(\rho^2 t + 1)^2} \hspace{1cm} t > 0$$  \hspace{1cm} (D.29)

where $lim_t o_1(\frac{1}{t}) = 0$ and $K \geq 0$. Here $K = 0$ if and only if $F$ is Gaussian, in which case $I_F(t) = 0$ for all $t > 0$. 

96
Appendix E. Asymptotic Behavior of $I_F$ for Distributions $F$ in Class 2

In this appendix, we consider the asymptotic behavior of $\{I_F(t)\}_{t>0}$ where the distribution $F$ is in Class 2—where the RV $\xi$ takes on a finite number of values $z_1 < z_2, \ldots < z_n$ with probabilities $p_1, p_2, \ldots, p_n$ (resp.) and with $\sum_{i=1}^{n} p_i z_i = 0$.

We first manipulate (VI.1.11) in order to facilitate the analysis. By expanding the square term in (VI.1.11), we have for $t > 0$

\[ I_F(t) = I^1_F(t) + I^2_F(t) + I^3_F(t) \quad (E.1) \]

where

\[ I^1_F(t) := E' \left[ \frac{\{ \int_{\mathbb{R}} z \exp[z B_t - z^2 t/2] dF(z) \}^2}{\int_{\mathbb{R}} \exp[z B_t - z^2 t/2] dF(z)} \right] \]

\[ I^2_F(t) := \frac{-2}{t + 1/\sigma^2} E' \left[ B_t \int_{\mathbb{R}} z \exp[z B_t - z^2 t/2] dF(z) \right] \quad (E.2) \]

\[ I^3_F(t) := \frac{1}{(t + 1/\sigma^2)^2} E' \left[ B_t^2 \int_{\mathbb{R}} \exp[z B_t - z^2 t/2] dF(z) \right]. \]

Theorem V.1.1 provides us with the integrability conditions needed to separate the expectation of $I_F(t)$ into the three expectations $I^1_F(t)$, $I^2_F(t)$ and $I^3_F(t)$ for each $t > 0$. By making use of the Fubini-Tonelli Theorem, Theorem V.1.1 and elementary results on Gaussian random variables, we may write for $t > 0$

\[ I^2_F(t) = \frac{-2}{t + 1/\sigma^2} \int_{\mathbb{R}} z \exp[-z^2 t/2] E'[B_t \exp[z B_t]] dF(z) \]

\[ = \frac{-2}{t + 1/\sigma^2} \int_{\mathbb{R}} z^2 t dF(z) \quad (E.3) \]

\[ = \frac{-2 t \sigma^2}{t + 1/\sigma^2} \]

97
and
\[ I_F^2(t) = \frac{1}{(t + 1/\sigma^2)^2} \int_{\mathbb{R}} \exp[-z^2 t/2] E' \left[ B_t^2 \exp[z B_t] \right] dF(z) \]
\[ = \frac{1}{(t + 1/\sigma^2)^2} \int_{\mathbb{R}} t(1 + z^2 t) dF(z) \]
\[ = \frac{t(1 + \sigma^2 t)}{(t + 1/\sigma^2)^2} \]
\[ = \frac{\sigma^2 t}{t + 1/\sigma^2}, \quad (E.4) \]

so that
\[ I_F(t) = I_F^1(t) - \frac{\sigma^2}{1 + 1/(\sigma^2 t)}. \quad t > 0 \quad (E.5) \]

In (E.4) we have used the fact that \( F \) is in \( \mathcal{D}_1 \)—that \( E[\xi] = 0 \).

As in Appendix D, define collections of mappings \( \{N_t\}_{t>0} \) and \( \{D_t\}_{t>0} \) from \( \mathbb{R} \) to \( \mathbb{R} \) by
\[ N_t(b) := \int_{\mathbb{R}} z \exp[z b - z^2 t/2] dF(z) \]
\[ D_t(b) := \int_{\mathbb{R}} \exp[z b - z^2 t/2] dF(z) \quad (E.6) \]

for all \( b \in \mathbb{R} \) and all \( t > 0 \), so that
\[ I_F^1(t) = E' \left[ \frac{N_t(B_t)}{D_t(B_t)} \right]^2 D_t(B_t) \]
\[ = \int_{\mathbb{R}} \left| \frac{N_t(b)}{D_t(b)} \right|^2 D_t(b) \frac{\exp[-b^2/(2t)]}{\sqrt{2\pi t}} db \]
\[ = \int_{\mathbb{R}} \left| \frac{N_t(u t)}{D_t(u t)} \right|^2 D_t(u t) \frac{\exp[-u^2 t/2]}{\sqrt{2\pi / t}} du \]
\[ = \int_{\mathbb{R}} K_t^2(u) M_t(u) du \quad (E.7) \]

where for all \( u \) in \( \mathbb{R} \) and \( t > 0 \), we have set
\[ K_t(u) := \frac{N_t(u t)}{D_t(u t)}, \quad (E.8) \]
and

$$M_t(u) := D_t(ut) \frac{\exp[-u^2 t/2]}{\sqrt{2\pi t}}.$$  \hfill (E.9)

Direct evaluation of (E.8) and (E.9) yield for all \( u \) in \( IR \) and \( t > 0 \) that

$$K_t(u) = \frac{\sum_{j=1}^{n} p_j z_j \exp[z_j ut - z_j^2 t/2]}{\sum_{j=1}^{n} p_j \exp[z_j ut - z_j^2 t/2]}$$

$$= \frac{\sum_{j=1}^{n} p_j z_j \exp[-(u - z_j)^2 t/2]}{\sum_{j=1}^{n} p_j \exp[-(u - z_j)^2 t/2]}.$$  \hfill (E.10)

and that

$$M_t(u) = \sum_{j=1}^{n} p_j \frac{\exp[-(u - z_j)^2 t/2]}{\sqrt{2\pi t}}.$$  \hfill (E.11)

From (E.7) and (E.11), we conclude that for all \( t > 0 \)

$$I_t^k(t) = \sum_{i=1}^{n} p_i \tilde{I}_t^k(t)$$

where for \( i = 1, 2, \ldots, n \)

$$\tilde{I}_t^k(t) := \int_{IR} K_t^2(u) \frac{\exp[-(u - z_j)^2 t/2]}{\sqrt{2\pi t}} du$$

$$= \int_{IR} K_t^2(v/\sqrt{t} + z_j) \frac{\exp[-v^2/2]}{\sqrt{2\pi}} dv.$$  \hfill (E.13)

Now fix \( i = 1, 2, \ldots, n, \) and observe from (E.10) that

$$K_t^2(u) = \frac{\sum_{1 \leq j, k \leq n} p_j p_k z_j z_k \exp[\{(u - z_j)^2 + (u - z_k)^2\} t/2]}{\sum_{1 \leq j, k \leq n} p_j p_k \exp[\{(u - z_j)^2 + (u - z_k)^2\} t/2]}$$  \hfill (E.14)

for all \( u \) in \( IR \) and \( t > 0 \), so elementary algebraic manipulations reveal that

$$K_t^2(v/\sqrt{t} + z_i) = \frac{\sum_{1 \leq j, k \leq n} p_j p_k z_j z_k \exp[v\beta_{ij,k} \sqrt{t} - \alpha_{ij,k} t]}{\sum_{1 \leq j, k \leq n} p_j p_k \exp[v\beta_{ij,k} \sqrt{t} - \alpha_{ij,k} t]}$$  \hfill (E.15)

for all \( u \) in \( IR \) and \( t > 0 \), where

$$\beta_{ij,k} := (z_j - z_i) + (z_k - z_i)$$

$$\alpha_{ij,k} := \{(z_j - z_i)^2 + (z_k - z_i)^2\}/2.$$  \hfill (E.16)
for all $1 \leq i, j, k \leq n$. Upon expanding both the numerator and denominator of (E.15) about the $(j, k) = (i, i)$ term, and noting that $\alpha_{i, i, i} = \beta_{i, i, i} = 0$, we conclude that for all $v$ in $\mathbb{R}$ and $t > 0$,

$$K_i^2(v/\sqrt{t} + z_i) = \frac{z_i^2 + g_i^1(v; t)}{1 + g_i^2(v; t)},$$  \hspace{1cm} (E.17)

with

$$g_i^1(v; t) := \sum_{1 \leq j, k \leq n, (j, k) \neq (i, i)} z_j z_k \left( \frac{p_j p_k}{(p_i^2)} \right) \exp[v \beta_{i, j, k} \sqrt{t} - \alpha_{i, j, k} t]$$

$$g_i^2(v; t) := \sum_{1 \leq j, k \leq n, (j, k) \neq (i, i)} \left( \frac{p_j p_k}{(p_i^2)} \right) \exp[v \beta_{i, j, k} \sqrt{t} - \alpha_{i, j, k} t]$$  \hspace{1cm} (E.18)

for all $1 \leq i \leq n$. Now (E.17) may be rewritten as

$$K_i^2(v/\sqrt{t} + z_i) = z_i^2 + g_i^3(v; t) \hspace{1cm} v \in \mathbb{R}, t > 0 \hspace{1cm} (E.19)$$

where for all $v$ in $\mathbb{R}$ and $t > 0$ we have set

$$g_i^3(v; t) = \frac{g_i^1(v; t) - z_i^2 g_i^2(v; t)}{1 + g_i^2(v; t)}$$

$$= \frac{\sum_{1 \leq j, k \leq n, (j, k) \neq (i, i)} (z_j z_k - z_i^2) \left( \frac{p_j p_k}{(p_i^2)} \right) \exp[v \beta_{i, j, k} \sqrt{t} - \alpha_{i, j, k} t]}{1 + \sum_{1 \leq j, k \leq n, (j, k) \neq (i, i)} \left( \frac{p_j p_k}{(p_i^2)} \right) \exp[v \beta_{i, j, k} \sqrt{t} - \alpha_{i, j, k} t]}. \hspace{1cm} (E.20)$$

If for all $1 \leq i, j, k \leq n$ with $(j, k) \neq (i, i)$, we define

$$\tilde{I}_{\mathcal{F}}^{i, j, k}(v; t) := \frac{(p_j p_k/(p_i^2)) \exp[v \beta_{i, j, k} \sqrt{t} - \alpha_{i, j, k} t]}{1 + \sum_{1 \leq j, k \leq n, (j, k) \neq (i, i)} (p_j p_k/(p_i^2)) \exp[v \beta_{i, j, k} \sqrt{t} - \alpha_{i, j, k} t]} \hspace{1cm} (E.21)$$

for all $v$ in $\mathbb{R}$ and $t > 0$, then from (E.19)-(E.21) we have that for $i = 1, 2, \ldots, n$,

$$K_i^2(v/\sqrt{t} + z_i) = z_i^2 + \sum_{1 \leq j, k \leq n, (j, k) \neq (i, i)} (z_j z_k - z_i^2) \tilde{I}_{\mathcal{F}}^{i, j, k}(v; t) \hspace{1cm} (E.22)$$

for all $v$ in $\mathbb{R}$ and $t > 0$, so upon returning to (E.13), we conclude that

$$\tilde{I}_{\mathcal{F}}^i(t) = \int_{\mathbb{R}} \left\{ z_i^2 + \sum_{1 \leq j, k \leq n, (j, k) \neq (i, i)} (z_j z_k - z_i^2) \tilde{I}_{\mathcal{F}}^{i, j, k}(v; t) \right\} \frac{\exp[-v^2/2]}{\sqrt{2\pi}} dv$$

$$= z_i^2 + \sum_{1 \leq j, k \leq n, (j, k) \neq (i, i)} (z_j z_k - z_i^2) \int_{\mathbb{R}} \tilde{I}_{\mathcal{F}}^{i, j, k}(v; t) \frac{\exp[-v^2/2]}{\sqrt{2\pi}} dv \hspace{1cm} (E.23)$$

100
for \( i, j, k \leq n \) with \((j, k) \neq (i, i)\) and all \( t > 0 \), and observe from (E.12) and (E.24) that
\[
\int_F^{i;j;k}(t) = \sigma^2 + \sum_{1 \leq i, j, k \leq n \atop (j, k) \neq (i, i)} p_i(z_jz_k - z_i^2) \int_F^{i;j;k}(t), \quad t > 0 \tag{E.26}
\]
so that
\[
\int_F(t) = \frac{1}{t + 1/\sigma^2} + \sum_{1 \leq i, j, k \leq n \atop (j, k) \neq (i, i)} p_i(z_jz_k - z_i^2) \int_F^{i;j;k}(t). \quad t > 0 \tag{E.27}
\]

The problem is now to determine the asymptotics of \( \{\int_F^{i;j;k}(t)\}_{t > 0} \) for \( 1 \leq i, j, k \leq n \) with \((j, k) \neq (i, i)\). This analysis is accomplished using Large Deviations Theory. In what follows, we shall use the following extension of Varadhan’s Theorem for the asymptotics of integrals.

**Theorem E.1.** Let \( \{P_\epsilon\}_{\epsilon > 0} \) be a collection of probability measure on the Polish space \( \mathcal{X} \) equipped with its natural Borel \( \sigma \)-field. Suppose that \( \{P_\epsilon\}_{\epsilon > 0} \) satisfy a large deviations principle with rate function \( I \). If \( \{F_\epsilon\}_{\epsilon > 0} \) is a family of nonnegative Borel functions from \( \mathcal{X} \) to \( IR \) such that for some lower semicontinuous nonnegative function \( F : \mathcal{X} \to IR \),
\[
\lim_{\epsilon \to 0} F_\epsilon(y) = F(x) \quad \text{for} \quad x \in \mathcal{X}, \tag{E.28}
\]
then
\[
\lim_{\epsilon \to 0} \epsilon \ln \int_{\mathcal{X}} \exp \left[ -\frac{1}{\epsilon} F_\epsilon(x) \right] dP_\epsilon(x) = -\inf_{x \in \mathcal{X}} \{F(x) + I(x)\}. \tag{E.29}
\]

**Proof.** From [25, Theorem 2.3], we know that
\[
\limsup_{\epsilon \to 0} \epsilon \ln \int_{\mathcal{X}} \exp \left[ -\frac{1}{\epsilon} F_\epsilon(x) \right] dP_\epsilon(x) \leq -\inf_{x \in \mathcal{X}} \{F(x) + I(x)\}, \tag{E.30}
\]
so that we need only show that
\[
\liminf_{\epsilon \to 0} \epsilon \ln \int_{\mathcal{X}} \exp \left[ -\frac{1}{\epsilon} F_\epsilon(x) \right] dP_\epsilon(x) \geq -\inf_{x \in \mathcal{X}} \{F(x) + I(x)\}. \tag{E.31}
\]
Our proof is based on the second half of the proof of [25, Theorem 2.2]. For every $\delta > 0$, there exists an element $y$ in $X$ such that

$$F(y) + I(y) \leq \inf_{x \in X} \{F(x) + I(x)\} + \delta / 2. \quad (E.32)$$

By (E.28), there is an open neighborhood $U$ of $y$ and some $\hat{\epsilon} > 0$ such that for all $x$ in $U$ and $0 < \epsilon < \hat{\epsilon}$,

$$F_{\epsilon}(x) < F(y) + \delta / 2, \quad (E.33)$$

so that for all $0 < \epsilon < \hat{\epsilon}$,

$$\epsilon \ln \int_X \exp \left[ -\frac{1}{\epsilon} F_{\epsilon}(x) \right] dP_{\epsilon}(x) \geq \epsilon \ln \int_U \exp \left[ -\frac{1}{\epsilon} F_{\epsilon}(x) \right] dP_{\epsilon}(x)$$

$$\geq \epsilon \ln \int_U \exp \left[ -\frac{1}{\epsilon} (F(y) + \delta / 2) \right] dP_{\epsilon}(x) \quad (E.34)$$

$$= -F(y) - \delta / 2 + \epsilon \ln P_{\epsilon}(U).$$

Since $\{P_{\epsilon}\}_{\epsilon > 0}$ satisfies a large deviations principle with rate function $I$, we get

$$\liminf_{\epsilon \to 0} \epsilon \ln \int_X \exp \left[ -\frac{1}{\epsilon} F_{\epsilon}(x) \right] dP_{\epsilon}(x) \geq -F(y) - \delta / 2 + \liminf_{\epsilon \to 0} \epsilon \ln P_{\epsilon}(U)$$

$$\geq -F(y) - \delta / 2 - \inf_{x \in U} I(x). \quad (E.35)$$

Clearly $\inf_{x \in U} I(x) \leq I(y)$, so that

$$\liminf_{\epsilon \to 0} \epsilon \ln \int_X \exp \left[ -\frac{1}{\epsilon} F_{\epsilon}(x) \right] dP_{\epsilon}(x) \geq -F(y) - I(y) - \delta / 2$$

$$\geq - \inf_{x \in X} \{F(x) + I(x)\} - \delta \quad (E.36)$$

where we have made use of (E.32) in the last step. Thus we have verified (E.31), since $\delta > 0$ was arbitrary.

Fix $1 \leq j, k \leq n$ with $(j, k) \neq (i, i)$. In order to use Varadhan's Theorem on $\{I^1_F(t)\}_{t > 0}$, let us make the following definitions. Define a collection of probability measures $\{P_t\}_{t > 0}$ on $(IR, B(IR))$ by means of the Radon-Nikodym derivatives with respect to $\lambda_1$ given by

$$\frac{dP_t}{d\lambda_1}(u) := \frac{\exp[-u^2 t / 2]}{\sqrt{2\pi t}}. \quad u \in IR, \ t > 0 \quad (E.37)$$

102
By well-known results (see [23, p. 1]), \( \{P_t\}_{t > 0} \) satisfies a Large Deviations Principle with rate function \( I \) given by
\[
I(x) := \frac{x^2}{2}, \quad x \in IR \quad (E.38)
\]
It is plain, from (E.21), that \( 0 \leq \tilde{J}^{i;j,k}_{F}(v; t) \leq 1 \) for all \( v \) in \( IR \) and \( t > 0 \) and \( 1 \leq i, j, k \leq n \) with \( (j, k) \neq (i, i) \), so that we may define the collection of mappings \( \{S_t\}_{t > 0} \) from \( IR \) to \( IR \) by fixing \( 1 \leq i, j, k \leq n \) with \( (j, k) \neq (i, i) \) and setting
\[
S_t(u) := -\frac{1}{t} \ln \tilde{J}^{i;j,k}_{F}(u\sqrt{t}; t) \quad (E.39)
\]
for \( u \) in \( IR \) and \( t > 0 \). The relation
\[
\tilde{J}^{i;j,k}_{F}(t) = \int_{IR} \exp[-tS_t(u)]dP_t(u) \quad (E.40)
\]
follows for \( t > 0 \), where it is evident, from the equations (E.21) and (E.39), that for each \( t > 0 \), \( S_t \) is nonnegative.

Our next step is to find a lower-semicontinuous mapping \( S : IR \to IR \) such that for all \( v \) in \( IR \)
\[
\lim_{t \to \infty} S_t(v) = S(u). \quad (E.41)
\]
Before attempting to find such a function, let us define several auxiliary mappings. For \( 1 \leq j, k \leq n \), define the affine mapping \( f_{i;j,k} : IR \to IR \) by
\[
f_{i;j,k}(u) := u\beta_{i;j,k} - \alpha_{i;j,k}, \quad u \in IR \quad (E.42)
\]
and for \( i = 1, 2, \ldots, n \), define the convex mapping \( \hat{f}_i : IR \to IR \) by
\[
\hat{f}_i(u) := \max_{1 \leq j, k \leq n, (j,k) \neq (i,i)} f_{i;j,k}(u). \quad u \in IR, t > 0 \quad (E.43)
\]
Next, define the collection of mappings \( \{G_t\}_{t > 0} \) from \( IR \) to \( IR \) by
\[
G_t(u) := 1 + \sum_{1 \leq j, k \leq n, (j,k) \neq (i,i)} (p_j p_k/(p_i^2)) \exp[f_{i;j,k}(u)t], \quad u \in IR, t > 0 \quad (E.44)
\]
and note that
\[
G_t(u) \geq 1 \quad (E.45)
\]
103
for \( u \) in \( IR \) and \( t > 0 \). Then we may write for \( t > 0 \)

\[
S_t(u) = \frac{1}{t} \ln \frac{p_{i,j} p_k}{p_i^t} - f_{i,j,k}(u) + \frac{1}{t} \ln G_t(u). \quad u \in IR \tag{E.46}
\]

Also define the collection \( \{H_i\}_{i>0} \) of mappings from \( IR \) to \( IR \) by

\[
H_i(u) := \exp[ -\hat{f}_i(u) t ] G_i(u)
\]

\[
= \exp[-\hat{f}_i(u) t] + \sum_{i=1}^n \frac{p_{i,j} p_k}{p_i^t} \exp\{f_{i,j,k}(u) - \hat{f}_i(u)\}(t) \quad u \in IR \tag{E.47}
\]

for \( t > 0 \). We are now in a position to find a mapping \( S : IR \to IR \) satisfying (E.41).

**Lemma E.1.** The mapping \( S : IR \to IR \) given by

\[
S(u) := \begin{cases} 
-f_{i,j,k}(u) & \text{if } \hat{f}_i(u) \leq 0 \\
\hat{f}_i(u) - f_{i,j,k}(u) & \text{if } \hat{f}_i(u) > 0
\end{cases} \quad u \in IR \tag{E.48}
\]

satisfies (E.41).

**Proof.** Fix \( u \) in \( IR \). There are three cases that need to be considered.

**Case 1:** \( \hat{f}_i(u) < 0 \). Since \( \hat{f}_i \) is continuous, there is an open neighborhood \( U \) of \( u \) such that \( \hat{f}_i(v) < 0 \) for \( v \) in \( U \). Thus \( f_{i,j,k}(v) < 0 \) for \( v \) in \( U \) and all \( 1 \leq j, k \leq n \) with \( (j,k) \neq (i,i) \), and

\[
1 \leq G_t(v) \leq 1 + \sum_{1 \leq j,k \leq n} \frac{p_{i,j} p_k}{p_i^t}, \quad v \in U, \ t > 0 \tag{E.49}
\]

and as a result, we get that

\[
\lim_{t \to \infty} \frac{1}{t} \ln G_t(v) = 0, \tag{E.50}
\]

and

\[
\lim_{t \to \infty} S_t(v) = -f_{i,j,k}(u). \tag{E.51}
\]

**Case 2:** \( \hat{f}_i(u) > 0 \). As in Case 1, there is a neighborhood \( U \) of \( u \) such that for \( v \) in \( U \), \( \hat{f}_i(v) > 0 \). Since \( f_{i,j,k}(v) - \hat{f}_i(v) \leq 0 \) for all \( v \) in \( U \) and \( 1 \leq j, k \leq n \) with \( (j,k) \neq (i,i) \), we see that

\[
H_i(v) \leq 1 + \sum_{1 \leq j,k \leq n} \frac{p_{i,j} p_k}{p_i^t}, \quad v \in U, \ t > 0 \tag{E.52}
\]
Now for each \( v \) in \( U \), we have \( f_{i;j,k}(v) - \hat{f}_i(u) = 0 \) for some \( 1 \leq j, k \leq n \) with \( (j, k) \neq (i, i) \), so that

\[
H_t(v) \geq \min_{1 \leq j, k \leq n \atop (j, k) \neq (i, i)} \frac{p_j p_k}{p_i^t}, \quad v \in U, \ t > 0 \quad (E.53)
\]

Combining (E.52) and (E.53), we find that

\[
\lim_{t \to \infty} \frac{1}{t} \ln H_t(v) = 0, \quad (E.54)
\]

and

\[
\lim_{t \to \infty} \frac{1}{t} \ln G_t(v) = \lim_{t \to \infty} \frac{1}{t} \ln \exp[\hat{f}_i(v)t]H_t(v) = \hat{f}_i(u), \quad (E.55)
\]

so the conclusion

\[
\lim_{t \to \infty} S_t(v) = \hat{f}_i(u) - f_{i;j,k}(u) \quad (E.56)
\]

is obtained.

**Case 3:** \( \hat{f}_i(u) = 0 \). Using the procedures of Cases 1 and 2, we see that for all \( v \) in \( IR \) and \( t > 0 \),

\[
1 \leq \frac{1}{t} \ln G_t(v) \leq \{\hat{f}_i(v) + \frac{1}{t} \ln H_t(v)\} + \{\hat{f}_i(v) \leq 0\} \frac{1}{t} \ln \left(1 + \sum_{1 \leq j, k \leq n \atop (j, k) \neq (i, i)} \frac{p_j p_k}{p_i^t}\right), \quad (E.57)
\]

and

\[
\left\{\hat{f}_i(v) \geq 0\right\} \frac{1}{t} \ln \left(\min_{1 \leq j, k \leq n \atop (j, k) \neq (i, i)} \frac{p_j p_k}{p_i^t}\right) \leq \left\{\hat{f}_i(v) \geq 0\right\} \frac{1}{t} \ln H_t(v)
\]

\[
\leq \left\{\hat{f}_i(v) \geq 0\right\} \frac{1}{t} \ln \left(1 + \sum_{1 \leq j, k \leq n \atop (j, k) \neq (i, i)} \frac{p_j p_k}{p_i^t}\right). \quad (E.58)
\]

Consequently, combining (E.57) and (E.58), we see that

\[
0 \leq \liminf_{t \to \infty} \frac{1}{t} \ln G_t(v) \leq \limsup_{t \to \infty} \frac{1}{t} \ln G_t(v) \leq \limsup_{t \to \infty} \frac{1}{t} \ln \hat{f}_i(v) = 0, \quad (E.59)
\]
since \( \hat{f}_i \) is continuous and \( \hat{f}_i(u) = 0 \). Thus

\[
\lim_{t \to \infty} S_t(v) = -f_{i;j,k}(u), \quad (E.60)
\]

and \( S \) satisfies (E.41).

Note that in fact \( S \) is continuous, so it is surely lower semicontinuous.

We are now in a position to apply Theorem E.1 and conclude that

\[
\lim_{t \to \infty} \frac{1}{t} \ln \tilde{I}_F^{i;j,k}(t) = -\inf_{x \in \mathbb{R}} [I(x) + F(x)].
\]  

Define \( l := \inf_{x \in \mathbb{R}} [I(x) + S(x)] \). If we can now show that \( l > 0 \), then \( \lim_{t} \{ tI_f^{i;j,k}(t) \} = 0 \) for arbitrary \( 1 \leq i, j, k \leq n \) with \( (j, k) \neq (i, i) \), and it follows from (E.27) that \( \lim_{t} \{ tI_F(t) \} = 1 \).

**Lemma E.2.** We have \( l > 0 \).

**Proof.** The functions \( I \) and \( S \) are nonnegative, the nonnegativity of \( S \) following from the nonnegativity of \( \{ S_t \}_{t>0} \), so \( l \geq 0 \). Assume that \( l = 0 \). Then there is a sequence \( \{ x_n \}_{n=0}^{\infty} \) of real numbers such that

\[
\lim_{n} \left( \frac{x_n^2}{2} + S(x_n) \right) = 0
\]  

(E.62)

and consequently \( \lim_{n} x_n = 0 \), so

\[
S(0) = 0
\]  

(E.63)

holds from the continuity of \( S \). Now from (E.42) and (E.43), we have

\[
\hat{f}_i(0) = \max_{1 \leq j, k \leq n, (j, k) \neq (i, i)} (-\alpha_{i;j,k})
\]  

(E.64)

with \( \alpha_{i;j,k} > 0 \) for \( 1 \leq i, j, k \leq n \) with \( (j, k) \neq (i, i) \). Thus \( \hat{f}_i(0) < 0 \), and therefore

\[
S(0) = -f_{i;j,k}(0) = -\alpha_{i;j,k} < 0,
\]  

(E.65)

in contradiction with (E.63). Thus \( l > 0 \).

We now have the following characterization of the asymptotic behavior of \( \{ I_F(t) \}_{t>0} \).

**Theorem E.2.** We have

\[
I_F(t) = \frac{1 + o_1(1)}{t}, \quad t > 0
\]  

(E.66)
where \( \lim_{t \to 0} tI_F(t) = 0 \).

**Proof.** Equation (E.66) is equivalent to showing that \( \lim_{t \to 0} tI_F(t) = 1 \). From (E.27),

\[
tI_F(t) = \frac{1}{1 + 1/(\sigma^2 t)} + \sum_{1 \leq i,j,k \leq n, (j,k) \neq (i,i)} p_i (z_j z_k - z_i^2) tI^{i,j,k}(t) \quad (E.67)
\]

for all \( t > 0 \). From Lemma E.2 and (E.61) we have \( \lim_{t \to 0} \frac{1}{t} \ln I_F^{i,j,k}(t) < 0 \) for \( 1 \leq i, j, k \leq n \) with \( (j, k) \neq (i, i) \), so that \( \lim_{t \to 0} tI_F^{i,j,k}(t) = 0 \) for \( 1 \leq i, j, k \leq n \) such that \( (j, k) \neq (i, i) \). The sum in (E.67) being finite, we see that indeed \( \lim_{t \to 0} tI_F(t) = 1 \). \( \Box \)
BIBLIOGRAPHY


