Akademie věd České republiky
Ústav teorie informace a automatizace

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

Miroslav Kárný, Ivan Nagy

DYNAMIC BAYESIAN DECISION-MAKING: part I.

No. 1971 December 1999

ÚTIA AV ČR, P. O. Box 18, 182 08 Prague, Czech Republic
Fax: (+42)(2)688 4903
E-mail: utia@utia.cas.cz

Contents

Chapter 1

Introduction

This text should support learning theory of Bayesian decision-making under uncertainty presented at the course of "Adaptive Systems". There are good textbooks for learning both decision-making theory, e.g. [1, 2] and adaptive control, e.g. [3, 4]. The view presented in this text, namely, viewing adaptive systems as an important branch of *dynamic* decision-making under uncertainty is dissipated in various textbooks and numerous papers. The classical text [5, 6] is probably closest to this one. Permanently growing computer power available, novel solutions of various theoretical and algorithmic subtasks justify this attempt to update the description of the theory.

1.1 Aim of the work

Theory of statistical decision-making under uncertainty [2, 7] provides a unified logical structure for solving problems ranging from parameter estimation, prediction, pattern recognition, learning, testing of hypothesis etc. up to feedback control. The application width is also enormous: inspection, maintenance and control of technological as well as economic processes; signal processing; information processing in medicine, physics, etc. Huge amount of available particular results, techniques and methods hide the common logical structure and leave often newcomers without a guide.

This *text aims to recall the relatively simple common structure of the mentioned decision tasks and to help the user in focusing on those ingredients that are specific to his/her problem.*

Naturally, the reaching of a satisfactory solution of a particular decision making problem remains to be an iterative process [8]. The number of trial solutions can be, however, significantly reduced by using the advocated logical structure. This is the main practical outcome of mastering the presented theory.

The logical structure of the decision-making under uncertainty is now well understood. The number of open problems, related especially to computational and implementation aspects, is, however, large. Some of them are formulated throughout the text. In this way, we try to *turn attention of (prospective) researchers to challenging important topics.*

1.2 Structure

The aim of the work dictates the top-down style of presentation: the logically simple but relatively abstract problem formulation and solution are gradually specialised. The logical skeleton of the theory is filled up by practical details later on in Chapters dealing with particular examples.

Chapter 2 summarises the theory of the design. Basic analytical results are in Chapter 3. Usual decision-making tasks are discussed, formulated and solved in Chapter 4. Common evaluation techniques are characterised in Chapter 5.

The rest of the text (contained in the part II) serves for illustration of the theory. The structure of these chapters is (more or less) uniform. For specific construction elements, they illustrate majority of decision-making tasks that can be related to the considered case. For instance, Chapter 6 deals with a bit academic problem of the controlled tossing of a coin. Simplicity of the calculus related to it helps us to focus on the logical structure to be mastered. Other examples are gradually less and less academic and become more and more of practical importance.

1.3 Readers

It is natural that authors of this text hope to have a wide range of readers. It is also natural to expect them to be very individual in their attitude, background and interests. So we can roughly distinguish just two groups:

Research inclined who are interested in the state and problems of the area and do not worry of relatively abstract notions. They are expected to read the text linearly with occasional skips and complementary reading of references.

Application oriented who are searching for a flexible tool-set for their real-life problem. They are supposed to start with demos supplied to particular examples in order to judge whether the presented text is of interest for them at all. Any example which is judged to be close to the considered application may serve to this purpose. If, as we hope, they will find it worth of the energy put in study they are expected to go through Chapter 4 in order to learn the structure of the problem formulation and its solution. Chapter 5 then serves for inspiration on practical evaluation techniques. Both of them contain pointers to the necessary auxiliary notions in remaining "theoretical" Chapters.

All readers are warned that:

- this text is not mathematical one in spite of all mathematical machinery (mis)used: mathematical purity and technical details are sacrificed in order to present basic explanations as directly as possible;

- the presentation adapts and sometimes creates symbols and terminology with the aim to balance generality and expressiveness: for instance, notion of probability space with its σ-algebras is replaced by notions like ignorance, experience etc.;

- the attention focuses on problems that can be quantified and potentially solved in as algorithmic way as possible: the relatively straightforward extension to non-numerical problems are not considered, for sake of conciseness.

1.4 Acknowledgment

This work has two authors but it could not be created without our predecessors A. Halousková, A. Peréz and V. Peterka, strong personalities of the field T. Bohlin, D.W. Clark, E. Mosca, P. Wellstead and K.J. Aström, colleagues J. Bůcha, J. Böhm, L. Jirsa, R. Kulhavý, J. Kadlec, P. Nedoma and J. Šindelář, applicators P. Ettler, J. Fessl, J. Heřmanská, A. Lízr, A. Lízrová and others, visiting members of our Department K. Hangos, K. Warwick and PhD students L. Berec, H. Gao, T.V. Guy, P. Ivanova, N. Khailova, L. Tesař, J. Rojíček, and M. Valečková.

The long-term support of Institute of Information Theory and Automation as well as grants of GA AV ČR No A 2147701 and K 1075601, GAČR No 102/97/0118 and 102/99/1564 have made possible to work in this interesting research area with a great application potential and to create finally this text.

Chapter 2

Design

At a general level, the presented theory should help the decision maker (typically, human being) to select one of the available options (decisions). These options relate to a part of the real world, to a system. They concern with descriptions of the system and/or its influencing.

This chapter summarises the design principles and tools used throughout the text. It starts with conventions and notions used throughout. The considered framework covers a broad range of problems. Inevitably, the adopted symbols and notions have more specific meanings in specific application fields. The reader is kindly asked to be patient (especially) in this respect.

Then, from the inspected principle of an optimal decision-making under uncertainty, we get the following important conclusion: incomplete knowledge and randomness have the same operational consequences for decision-making and should be treated in the same way: labelled as *Bayesian decision making*.

Then, the way for obtaining both the optimal decision rules and the strategies is derived. This design works with models that are usually obtained through Bayesian learning whose description closes the chapter.

A similarly formulated control design is presented in [9]. For a detailed explanation of Bayesian learning see [10].

2.1 General conventions

The conventions presented here are mostly respected in this work. If some exception is necessary it is named just at the place of its validity. If need be, the brief characterisation is complemented by explanatory comments (printed differently).

Quantity *is a mapping with a numerical range.*

The domain and form of the quantity are mostly unused and unspecified.

Realisation *is a value of the quantity.*

Often, the quantity and its realisation are not distiguished, as usual. The proper meaning is determined by the context.

f *is the letter reserved for probability (density) functions (p(d)f).*

The meaning of the p(d)f is given through the identifier of its argmument.

x^* *denotes the range of the quantity* x.

$\bar{x}$ *denotes cardinality of the set* x^*.

$\equiv$ *means the equality by definition.*

If some verbal notions are introduced within bodies of Propositions, Remarks etc., then they are emphasised by the print used and they appear in Index.

x_t *is the quantity* x *at the discrete time labelled by* $t \in t^* \equiv \{1, \ldots, \bar{t}\}$.

$\bar{t} \leq \infty$ *is called (decision, prediction, control) horizon.*

$x_{i;t}$ *is an* i*th entry of the array* x *at time* t.

The semicolon in the subscript indicates that the symbol following it has the meaning of a time index.

$x(k\_l)$ *denotes the sequence between time moments* $k \leq l$, i.e. $x(k\_l) \equiv x_k, \ldots, x_l$.

$x(t) \equiv x(1\_t)$.

$x(t)$ *with* $t < 1$ *is an empty sequence and thus an empty information source.*

$\mathrm{supp}\,[\,f(x)]$ *is the support of the function* $f : x^* \rightarrow [0, \infty]$, *i.e. the subset of* x^* *on which* $f(x) > 0$.

2.2 Basic notions and notations

System *is the part of the world that is of interest for a decision maker who should either describe or influence this part.*

Decision maker might be a person, group of persons or mechanisms.

The system is specified with respect to the aim that the decision maker wants to reach and with respect to the tools he/she has available. In other words, the (penetrable) boundaries of the system are implied by the decision task considered.

Decision $a \in a^*$ *is the quantity whose value can be directly chosen by the decision maker for reaching his/her aims.*

A decision task makes sense only when the decision maker faces a real option, i.e. when $\bar{a} > 1$.

(Decision) experience $\mathcal{P}_a \in \mathcal{P}_a^*$ *contains knowledge about the system available to the decision maker for the choice of the value of* a.

If data values D are available for constructing the estimate $\hat{\Theta}$ of an unkown quantity $\Theta \in \Theta^*$ then its experience is $\mathcal{P}_{\hat{\Theta}} \equiv D$. Often, experience includes the *past data observed.*

(Decision) ignorance $\mathcal{F}_a \in \mathcal{F}_a^*$ *specifies knowledge about the system unavailable to the decision maker for the choice of the value of* a.

The unknown value of the quantity Θ to be estimated is a part of the ignorance $\mathcal{F}_{\hat{\Theta}}$ of its estimate $\hat{\Theta}$. Often, ignorance contains *future still unobserved data.*

(System) behaviour $\mathcal{Q}^*$ *consists of all possible realisations (of trajectories)* $\mathcal{Q}$, *i.e. values of all quantities considered by the decision maker within the time span determined by the horizon of interest.*

The behaviour and its realisations can be split with respect to any action $a \in a^*$ into the relevant experience $\mathcal{P}_a$ and ignorance $\mathcal{F}_a$. Formally, $\mathcal{Q} = (\mathcal{P}_a, a, \mathcal{F}_a)$.

Note that a single behaviour can be split differently with respect to a pair of decisions, say a, $\tilde{a}$ with different experience $\mathcal{P}_a \neq \mathcal{P}_{\tilde{a}}$ and different ignorance $\mathcal{F}_a \neq \mathcal{F}_{\tilde{a}}$. $\mathcal{Q} = (\mathcal{P}_a, a, \mathcal{F}_a) = (\mathcal{P}_{\tilde{a}}, \tilde{a}, \mathcal{F}_{\tilde{a}})$.

(System) input $u \in u^*$ *is a decision which is supposed to influence the ignorance part* $\mathcal{F}_u$ *of the system behaviour.*

A manipulable valve position influencing a fluid flow is the system input. A point estimate $\hat{\Theta}$ of an unknown (realisation of) quantity Θ is an instance of the decision that is not the system input. It is chosen with the aim to describe the system and exhibits no influence on it.

(System) output $y \in y^*$ *is an observable quantity that informs the decision maker on the behaviour of the system.*

A pressure measured in an isolated heated system is an instance of the output. A pressure applied on a closed system is an instance of the input.

To be or not to be an input is a relative property.

Innovation $\Delta_{a_t} \in \Delta^*_{a_t}$ *contains quantities that are not included in* $\mathcal{P}_{a_t}$ *but are included in* $\mathcal{P}_{a_{t+1}} \setminus a_t$.

Often, $\Delta_{a_t} = y_t =$ the system output at time t.

Decision rule $\mathcal{R} : \mathcal{Q}^* \to a^*$ *is a mapping that assigns a decision* $a \in a^*(\supset$ *range) to the behaviour* $\mathcal{Q} \in \mathcal{Q}^*(\supset$ *domain).*

Causal decision rule $\mathcal{R} : \mathcal{P}^*_a \to a^*$ *is a mapping that assigns the decision* $a \in a^*$ *to its experience* $\mathcal{P}_a \in \mathcal{P}^*_a$.

In other words, the decision a made by a causal decision rule is not influenced by the related ignorance $\mathcal{F}_a$.

Mostly, we deal with the causal decision rules so that the term "causal" is often dropped.

Estimator is an instance of the causal decision rule $\mathcal{R} : \mathcal{P}^*_{\hat{\Theta}} \to \hat{\Theta}^*$ that assigns an estimate $\hat{\Theta}$ of the unknown quantity $\Theta \in \Theta^*$ to the available experience $\mathcal{P}_{\hat{\Theta}}$.

Strategy *is a sequence of decision rules* $\{\mathcal{R}_t : \mathcal{Q}^* \to a^*_t\}$.

Causal strategy $\{\mathcal{R}_t : \mathcal{P}^*_{a_t} \to a^*_t\}_{t \in t^*}$ *is made of causal decision rules.*

Again, we deal mostly with causal strategies so that the term "causal" is often dropped.

Controller *is a causal strategy assigning inputs* u_t *to experience* $\mathcal{P}_{u_t}$ $\forall t \in t^*$.

The proportional controller given by the proportionality constant p is an example of the causal control strategy $\{y^*_{t-1} \to u^*_t : u_t = p y_{t-1}\}_{t \in t^*}$.

Design *selects the decision rule or strategy.*

The design selecting a single rule is called *static design*. The choice of the strategy is *dynamic design*.

The person (group) who makes the selection is *designer*. Authors and readers of this text are supposed to be designers. In that sense, the term *we* used within the text should (mostly) be read "*we designers*".

Agreement 2.2.1 (Interface between the design & reality) *Physical connections of the design domain the to the real world (sensors, transmission lines, actuators etc.) are here always taken as a part of the system. Consequently, all quantities and mappings considered can be and are taken as mathematical entities living in an abstract calculating machine.*

Uncertain behaviour *related to a static design is the behaviour whose realisation $\mathcal{Q}^*$ that can be decomposed into*

- $\mathcal{Q}^*_{\mathcal{R}} \equiv$ *the part that is unambigously determined by the considered decision rule $\mathcal{R} \in \mathcal{R}^*$,*
- uncertainty $\mathcal{N}^*$ *belonging to the ignorance of decisions generated by $\mathcal{R}$ and uninfluenced even indirectly them.*

With an abuse of notation, we write the corresponding decomposition of realisations $\mathcal{Q} = (\mathcal{Q}_{\mathcal{R}}, \mathcal{N}) = (\cdot, uncertainty)$.

Uncertain behaviour related to dynamic design is encountered if any of its rules faces uncertainty.

Incomplete knowledge of (the realisation of) a considered quantity $\Theta \in \Theta^*$ makes the behaviour uncertain.

External unobserved noise influencing the system makes its behaviour uncertain.

Uncertainty expresses both incomplete knowledge and randomness.

Uncertain behaviour in the dynamic design is defined in a similar way by replacing rules with strategies. The related cumbersome notation is avoided in our presentation.

Decision-making *means design and application of a decision rule (strategy).*

Admissible strategy *is a strategy $\{\mathcal{R}_t\}_{t\in t^*}$ that*

- *is causal* $\Leftrightarrow$ $\{\mathcal{R}_t\}_{t\in t^*} \equiv \{\mathcal{R}_t : \mathcal{P}^*_{a_t} \to a^*_t\}_{t\in t^*}$ *and*
- *meets technological restrictions* $\Leftrightarrow$ *the ranges of its decision rules are in prespecified subsets of respective decision sets.*

Loss function , $\mathcal{Z} : \mathcal{Q}^* \to [0,\infty]$, *quantifies the degree of achievement of the designer's aim.*

The loss function measures quality of the behaviour $\mathcal{Q}$ and indirectly (partially) orders admissible strategies influencing the behaviour.

The smaller the value of $\mathcal{Z}(\mathcal{Q})$ the better.

Negative value of profit might be a loss function in economic applications.

Quadratic deviation $(\hat{\Theta} - \Theta)'(\hat{\Theta} - \Theta)$ ($'$ is transposition) of the estimate $\hat{\Theta}$ and of an unknown vector Θ is a wide spread loss function used for ranking of estimators ($\equiv$ decision rules generating estimates).

The important case of multi-valued loss function [11] is out of the scope of this text but the approach discussed in Section 2.3 could be relatively simply extended to it by embedding the index of the "decisive" entry of the loss funcion into uncertainty.

"Expected" loss $\tilde{\mathcal{E}}(\mathcal{Z})$ *assigns to the considered loss function* $\mathcal{Z}$ *and strategy a value in* $[0, \infty]$ *that is independent of the realisation uncertainty.*

Optimal design *selects an admissible strategy that leads to the smallest value of the "expected" loss function.*

Practically admissible strategy *is an admissible strategy that respects restrictions limiting complexity of decision-making.*

The complexity is considered with respect to the computational sources available at the design and application times. The majority of discussed problems in which the complexity restriction play a role are computationally hard in terms of computer sciences []. An intuitive understanding of the computational complexity is sufficient to our purposes.

Practically optimal design *selects a practically admissible strategy giving the smallest values of the "expected" loss function.*

The presented optimal design provides optimal admissible strategies and can be simply adapted to provide strategies of a pre-specified complexity by optimising over a set of simple decision rules (e.g. over proportional controllers only).

Operational formal tools for practically optimal design are not available. It is not known how to make the optimal design of a pre-specified complexity.

We never know whether the selection of proportionality constant made with ten operations is really best one possible among all selections that are allowed to use ten operations.

This is the main barrier of the applicability of the theory describing the optimal design. The optimal design becomes a practical tool by employing sound engineering heuristics. The practical optimality is not guaranteed. This fact is stressed by using the term *suboptimal design* .

2.3 Decision-making under uncertainty

Agreement 2.3.1 (Uncertainty in decision-making) *Decision-making under uncertainty arises if the optimal decision-making is to be performed and*

- *at least a pair of different decisions can be made,* $\bar{a} > 1$,
- *the considered loss function* $\mathcal{Z}(\mathcal{Q}) \equiv \mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \mathcal{N})$ *depends on non-void uncertainty* $\mathcal{N} \in \mathcal{N}^*$.

Under uncertainty, the function $\mathcal{Z}_{\mathcal{R}}(\mathcal{N}) \equiv \mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \mathcal{N})$ *of uncertainty* $\mathcal{N}$ *(the part of behaviour in ignorance uninfluenced by* $\mathcal{R}$*) is assigned to each considered strategy* $\mathcal{R} \in \mathcal{R}^*$. *The collection of such function is denoted* $\mathcal{Z}_{\mathcal{R}^*}$

$$\mathcal{Z}_{\mathcal{R}^*} \equiv \{\mathcal{Z}_{\mathcal{R}} : \mathcal{N}^* \to [0,\infty],\ \mathcal{Z}_{\mathcal{R}}(\mathcal{N}) \equiv \mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \mathcal{N})\}_{\mathcal{R}\in\mathcal{R}^*} \tag{2.1}$$

Under uncertainty, the loss function is insufficient for a *complete* ordering (comparing) of admissible rules (and consequently strategies) in spite of the fact that its values are in the fully ordered set. Each $\mathcal{R}$ have assigned a *function* in the set (2.1) and not a single number.

For instance, consider a pair of estimates $\hat{\Theta}_1 \neq \hat{\Theta}_2$ of an unknown scalar quantity $\Theta \in \Theta^* \equiv (-\infty, \infty)$. They cannot be compared with the help of the quadratic loss function $(\Theta - \hat{\Theta})^2$ as we do not know whether Θ is in that part of Θ^* where $(\Theta - \hat{\Theta}_1)^2 \leq (\Theta - \hat{\Theta}_2)^2$ or in its complement.

2.3.1 Ordering of strategies

Here, we describe a general way how to face uncertainty, i.e. incomplete ordering of strategies. In order to avoid a cumbersome notation, we formulate the adopted design principle, related requirements and their consequences for a static design, i.e. the design of a single *not necessarily causal* decision rule. The obtained results are also valid for the dynamic design, i.e. the choice of decision strategies.

We rely on the following systematic choice of a good decision rule.

Agreement 2.3.2 (Design principle)

- *A functional* $\tilde{\mathcal{E}}$ *is selected by designer. It assigns, to functions in (2.1), determined by the the considered loss function* $\mathcal{Z}$ *and indexed by inspected decision rules* $\mathcal{R} \in \mathcal{R}^*$, *a non-negative* "expected" loss

$$\begin{aligned} \tilde{\mathcal{E}} &: \mathcal{Z}_{\mathcal{R}^*} \to [0,\infty] \\ \mathcal{Z}_{\mathcal{R}^*} &\equiv \{\mathcal{Z}_{\mathcal{R}} : \mathcal{N}^* \to [0,\infty],\ \mathcal{Z}_{\mathcal{R}}(\mathcal{N}) \equiv \mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \mathcal{N})\}_{\mathcal{R}\in\mathcal{R}^*} \end{aligned} \tag{2.2}$$

- *The minimiser of* $\tilde{\mathcal{E}}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \mathcal{N})] \equiv \tilde{\mathcal{E}}[\mathcal{Z}_{\mathcal{R}}]$ *found in* $\mathcal{R}^*$ *is taken as the optimal decision rule.*

The outcome of this design depends on the introduced "expectation" $\tilde{\mathcal{E}}$. It has to avoid (at least) unequivocally bad rules. Such bad rules are identified here with dominated decision rules.

Agreement 2.3.3 (Dominated decision rules; Isotonic "expectation") *Let a loss function* $\mathcal{Z}$ *measure the quality of the behaviour. The decision rule* $\mathcal{R} : \mathcal{Q}^* \to a^*$ *is called* dominated *iff there is another decision rule* $\tilde{\mathcal{R}} : \mathcal{Q}^* \to a^*$ *such that for*

$$\mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \mathcal{N}) \geq \mathcal{Z}(\mathcal{Q}_{\tilde{\mathcal{R}}}, \mathcal{N}),\ \forall \mathcal{N} \in \mathcal{N}^*. \tag{2.3}$$

The decision rule is called strictly dominated *iff there is a non-trivial subset of* $\mathcal{N}^*$ *on which the inequality (2.3) is strict.*

The "expectation" $\tilde{\mathcal{E}}$ *(2.2) is said to be strictly isotonic if for a decision rule* $\mathcal{R}$ *strictly dominated by a decision rule* $\tilde{\mathcal{R}}$ *it holds*

$$\tilde{\mathcal{E}}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}}, \mathcal{N})] > \tilde{\mathcal{E}}[\mathcal{Z}(\mathcal{Q}_{\tilde{\mathcal{R}}}, \mathcal{N})]$$

Now we can specify the decision rules to be surely avoided.

Requirement 2.3.1 (Inadmissibility of dominated decision rules) *The considered "expectation"– minimisation design, Agreement 2.3.2, must not lead to the strictly dominated decision rule.*

We want to select the "expectation" $\tilde{\mathcal{E}}$ so that the decision rules are ordered as objectively as possible. It means that the ordering should weakly depend on the inspected set od strategies $\mathcal{R}^*$, on the loss function $\mathcal{Z}$ considered and on the designer using $\tilde{\mathcal{E}}$.

Requirement 2.3.2 (Independence of $\mathcal{R}^*$) *The chosen "expectation" should fulfill Requirement 2.3.1 even if we restrict the set of possible decision rules $\mathcal{R}^*$ to any of its non-trivial subsets that contain at least two different rules while at least one of them gives a finite expected loss.*

Proposition 2.3.1 (Isotonic ordering) *Assume that there is a rule in $\mathcal{R}^*$ for which the "expected" loss is finite. Then, Requirements 2.3.1, 2.3.2 are fulfilled iff the "expectation" is strictly isotonic, see Agreement 2.3.3.*

Proof:

1. We prove by contradiction that with strictly isotonic "expectation" $\tilde{\mathcal{E}}$ the minimiser in cannot be strictly dominated.

   Let $\tilde{\mathcal{E}}[\mathcal{Z}]$ be strictly isotonic on its domain $\mathcal{Z}_{\mathcal{R}^*}$ (2.2) and $\mathcal{R}_o \in \mathcal{R}^*$ be a minimiser of the "expected" loss. The minimiser gives necessarily a finite value of the corresponding $\tilde{\mathcal{E}}[\mathcal{Z}]$. Let $\mathcal{R}_d \in \mathcal{R}^*$ dominate it strictly. Then, because of the construction of $\mathcal{R}_o$, strict dominance and strict isotonic nature of $\tilde{\mathcal{E}}$, we get the following contradictory inequality

   $$\tilde{\mathcal{E}}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_o},\mathcal{N})] \underbrace{\leq}_{minimum} \tilde{\mathcal{E}}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_d},\mathcal{N})] \underbrace{<}_{isotonic} \tilde{\mathcal{E}}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_o},\mathcal{N})].$$

2. We prove by contradiction that use of "expectation" $\tilde{\mathcal{E}}$ that is not strictly isotonic leads to violation of Requirement 2.3.1 when Requirement 2.3.2 holds.

   If $\tilde{\mathcal{E}}[\mathcal{Z}]$ is not strictly isotonic on its domain $\mathcal{Z}_{\mathcal{R}^*}$ (2.2) then there is rule $\mathcal{R}_1 \in \mathcal{R}^*$ strictly dominated by a decision rule $\mathcal{R}_d \in \mathcal{R}^*$ such that

   $$\tilde{\mathcal{E}}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_d},\mathcal{N})] \geq \tilde{\mathcal{E}}[\mathcal{Z}(\mathcal{Q}_{\mathcal{R}_1},\mathcal{N})].$$

   If we restrict the set of decision rules $\mathcal{R}^*$ to the pair of rules $\{\mathcal{R}_1, \mathcal{R}_d\}$ then $\mathcal{R}_1$ can always be taken as the optimal decision rule. Thus, under Requirement 2.3.2, Requirement 2.3.1 is not fulfilled with such an "expectation" $\tilde{\mathcal{E}}$.

□

Requirement 2.3.2 guarantees suitability of the "expectation" to a wide range of decision rules. The "full objectivity" of ordering of decision rules is approached if $\tilde{\mathcal{E}}$ is made (almost) independent of the loss function considered. We have to adopt rather technical conditions on the set of such loss functions. Essentially, applicability to a very smooth functions and restricted version of addditive of $\tilde{\mathcal{E}}$ are required.

Requirement 2.3.3 (Independence of loss function) *Let us consider various loss functions $\mathcal{Z} \in \mathcal{Z}^*$. The "expectation" $\tilde{\mathcal{E}}$ acts on union of the set $\mathcal{Z}_{\mathcal{R}^*}$ (2.1) of functions with a common uncertainty $\mathcal{N}^*$. The union is taken over a wide set of loss functions $\mathcal{Z} \in \mathcal{Z}^*$*

$$\mathcal{Z}^*_{\mathcal{R}^*} \equiv \cup_{\mathcal{Z} \in \mathcal{Z}^*} \mathcal{Z}_{\mathcal{R}^*} \tag{2.4}$$

Suppose that $\mathcal{Z}^*_{\mathcal{R}^*}$ contains a subset of test loss functions *that are zero out of a compact non-empty subset Ω of $\mathcal{Q}^*$ and continuous on Ω.*

The "expectation" is an isotonic, sequentially continuous and boundedly uniformly continuous functional on $\mathcal{Z}^*_{\mathcal{R}^*}$ that is additive on loss functions with non-overlapping supports

$$\tilde{\mathcal{E}}[\mathcal{Z}_1 + \mathcal{Z}_2] = \tilde{\mathcal{E}}[\mathcal{Z}_1] + \tilde{\mathcal{E}}[\mathcal{Z}_2] \text{ if } \mathcal{Z}_1\mathcal{Z}_2 = 0,\ \mathcal{Z}_1,\ \mathcal{Z}_2 \in \mathcal{Z}^*_{\mathcal{R}^*}.$$

This technical Requirement allows us to get an integral representation of of the "expectation" searched for. Its proof as well as definitions of non-common terms can be found in [12].

Proposition 2.3.2 (Integral form of "expectation") *Under Requirement 2.3.3, the "expectation" $\tilde{\mathcal{E}}$ has the form*

$$\tilde{\mathcal{E}}[\mathcal{Z}] = \int_{\Omega} \mathcal{U}(\mathcal{Z}(\mathcal{Q}), \mathcal{Q})\, \mu(d\mathcal{Q}), \quad \textit{where} \tag{2.5}$$

μ is a finite regular non-negative Borel measure on Ω. The utility function *$\mathcal{U}$ satisfies $\mathcal{U}(0, \mathcal{Q}) = 0$. It is continuous in values of $\mathcal{Z}(\cdot)$ almost everywhere (a.e.) on Ω, bounded a.e. on Ω for each $\mathcal{Z}$ in the set of the test functions.*

Remarks 2.3.1

1. *The test loss functions are widely applicable and their consideration implies no practical restriction. The continuity requirements on $\tilde{\mathcal{E}}$ are also widely acceptable.*

2. *The additivity of $\tilde{\mathcal{E}}$ on functions with non-overlapping support seems to be sound. Any loss function $\mathcal{Z} \in \mathcal{Z}^*_{\mathcal{R}^*}$ can be written as $\mathcal{Z} = \mathcal{Z}\chi_\omega + \mathcal{Z}(1 - \chi_\omega) \equiv \mathcal{Z}_1 + \mathcal{Z}_2$, $\mathcal{Z}_1 Z_2 = 0$ with χ_ω denoting an* indicator of a set *$\omega \subset \Omega \subset \mathcal{Q}^*$ (χ equals 1 on ω and it is zero outside of it). The loss "expected" on the set ω and its complement should sum to the loss "expected" on the whole set of arguments.*

3. *The utility function $\mathcal{U}$ allows the designer to express his/her attitude to the design consequences: he/she might be risk aware, risk prone or risk indifferent [11].*

4. *The utility function $\mathcal{U}$ and the non-negative measure μ are universal for the whole set of test functions. $\mathcal{U}$ and μ are "objective", i.e. suitable for a wide range of decision tasks facing the same uncertainty.*

We formulate now our final objectivity oriented Requirement.

Requirement 2.3.4 (Indifference of the designer)

- *The designer is risk indifferent, i.e.* $\mathcal{U}$ *is the identity mapping.*
- *The expectation preserves any constant loss* $\tilde{\mathcal{E}}[constant] = constant$.
- *The involved measure* μ *has Radon-Nikodým derivative* $f(\mathcal{Q})$ *with respect to a dominating measure denoted* $d\mathcal{Q}$.

  In treated cases, $\mathrm{d}(\cdot)$ *is either Lebesgue or counting measure.*

Adopting Requirement 2.3.4, we get the basic representation Proposition.

Proposition 2.3.3 (Objective expectation) *Under Requirement 2.3.4, the "expectation"* $\tilde{\mathcal{E}}$ *(2.5) is formally identical with a* (mathematical) expectation. *The Radon Nikodým's derivative [12]* f *has all properties of the* joint probability (density) function *(*p(d)f*) on* $\mathcal{Q}^*$.

Proof: It is sufficient to observe that the preservation of constants implies that μ is a probabilistic measure, i.e. $\mu \geq 0,\ \mu(\mathcal{Q}^*) = 1$. □

Remarks 2.3.2

1. *The (mathematical) expectation is distinguished by dropping the sign* ˜

$$\mathcal{E}[\mathcal{Z}] \equiv \int_{\mathcal{Q}^*} \mathcal{Z}(\mathcal{Q}) f(\mathcal{Q})\, d\mathcal{Q} \equiv \int \mathcal{Z}(\mathcal{Q}) f(\mathcal{Q})\, d\mathcal{Q}. \tag{2.6}$$

2. *The first item in Requirement 2.3.4 is the last step in the objectivity direction we are following: we support objective designers.*

3. *The last item in Requirement 2.3.4 is unnecessary but it covers all cases we are here interested in. It helps us to deal with simpler objects, namely, with probability density functions (pdf) or probability functions (pf).*

   Mostly, we make no formal distinction between these cases and we use notations related to pdfs. Only when necessary, we underline that we deal with a pf and write integrals as sums.

Problem 2.3.1 (Is a partial ordering of strategies sufficient?) *It is intuitively obvious that the use of the complete ordering of strategies is quite demanding. It makes at least the related modelling hard. It is worthwhile to consider possibilities of focusing on a complete ordering "around" the optimal strategy.*

2.3.2 Calculus with pdfs

The joint pdf f on $\mathcal{Q} \equiv (\alpha, \beta, \gamma)$ is analysed/synthesised using several pdfs related to it. Let us recall meaning of pdfs derived from $f(\mathcal{Q})$.

Agreement 2.3.4 (Nomenclature of pdfs; Independence) *We work with*

Name	Meaning
joint pdf $f(\alpha,\beta\vert\gamma)$ *of* α,β conditioned on γ	*a pdf on* $(\alpha,\beta)^*$ *restricting* $f(\mathcal{Q})$ *on the cross-section of* $\mathcal{Q}^*$ *given by a fixed* γ
marginal pdf $f(\alpha\vert\gamma)$ *of* α conditioned on γ	*a pdf on* α^* *restricting* $f(\mathcal{Q})$ *on the cross-section of* $\mathcal{Q}^*$ *given by a fixed* γ *with no information on* β
marginal pdf $f(\beta\vert\alpha,\gamma)$ *of* β conditioned on α,γ	*a pdf on* β^* *restricting* $f(\mathcal{Q})$ *on the cross-section of* $\mathcal{Q}^*$ *given by a fixed* α,γ

The conditioning symbol $|$ *is dropped if just trivial conditions are considered. The pdf* $f(\alpha,\beta)$ *is the* lower dimensional joint pdf *(with respect to the pdf* $f(\alpha,\beta,\gamma)$*),* $f(\beta)$ *is the* marginal pdf.
Quantities α *and* β *are* conditionally independent *under the condition* γ *iff*

$$f(\alpha,\beta|\gamma) = f(\alpha|\gamma)f(\beta|\gamma). \tag{2.7}$$

Our manipulations with the introduced pdf rely on the following calculus.

Proposition 2.3.4 (Calculus with pdfs) *For generic* $(\alpha,\beta,\gamma) \in (\alpha,\beta,\gamma)^*$ *it holds:*

Non-negativity $f(\alpha,\beta|\gamma),\ f(\alpha|\beta,\gamma),\ f(\beta|\alpha,\gamma),\ f(\beta|\gamma) \geq 0.$

Normalisation $\int f(\alpha,\beta|\gamma)\, d\alpha d\beta = \int f(\alpha|\beta,\gamma)\, d\alpha = \int f(\beta|\alpha,\gamma)\, d\beta = 1.$

Chain rule $f(\alpha,\beta|\gamma) = f(\alpha|\beta,\gamma)f(\beta|\gamma) = f(\beta|\alpha,\gamma)f(\alpha|\gamma).$

Marginalisation $f(\beta|\gamma) = \int f(\alpha,\beta|\gamma)\, d\alpha,\ f(\alpha|\gamma) = \int f(\alpha,\beta|\gamma)\, d\beta.$

Bayes rule

$$f(\beta|\alpha,\gamma) = \frac{f(\alpha|\beta,\gamma)f(\beta|\gamma)}{f(\alpha|\gamma)} = \frac{f(\alpha|\beta,\gamma)f(\beta|\gamma)}{\int f(\alpha|\beta,\gamma)f(\beta|\gamma)\, d\beta} \propto f(\alpha|\beta,\gamma)f(\beta|\gamma). \tag{2.8}$$

> *The* proportionality sign $\propto$ *means that, in the equality represented, the factor independent of* β*, uniquely determined by the normalisation, is not explicitly written.*

Independence equivalents

$$f(\alpha,\beta|\gamma) = f(\alpha|\gamma)f(\beta|\gamma) \Leftrightarrow f(\alpha|\beta,\gamma) = f(\alpha|\gamma)\ \textit{or}\ f(\beta|\alpha,\gamma) = f(\beta|\gamma) \tag{2.9}$$

Proof: For demonstration see [10], a more precise and more technical treatment exploits the measure theory [12]. □

Remarks 2.3.3

1. *Duplicities in the presented formulae stress symmetry of many of them.*

2. *The technical fact that the identities hold only almost everywhere is mostly omitted in the subsequent explanations.*

3. *The Bayes rule is a simple consequence of previous formulae. Its importance in this text justifies the explicit presentation in various forms.*

4. *Identity (2.9) can be interpreted that β does not influence our description of α if α and β are conditionally independent for a given γ.*

Often, a pdf $f(\alpha)$ of a (multivariate) variable $\alpha \in \alpha^*$ is given and its image β is determined by a mapping $T : \alpha^* \to \beta^* \equiv T(\alpha^*)$. We need to specify the corresponding pdf $f(\beta)$. Rules for determining the pdf $f(\beta)$ are implied by the obvious need to preserve the expectation.

Proposition 2.3.5 (Pdfs of transformed quantities) *Let the expectation $\mathcal{E}$, acting on functions $\mathcal{B} : \beta^* \to (-\infty, \infty)$, be specified by the pdf $f_T(\beta)$, i.e.*

$$\mathcal{E}_T[\mathcal{B}] = \int \mathcal{B}(\beta) f_T(\beta)\, d\beta.$$

This functional expresses the same expectation as $\mathcal{E}[\mathcal{B}] = \int \mathcal{B}(T(\alpha)) f(\alpha)\, d\alpha$ iff

$$\int_{T(A)} f_T(T(\alpha))\, dT(\alpha) = \int_A f(\alpha)\, d\alpha \tag{2.10}$$

for all (measurable) sets $A \subset \alpha^*$.

Let α be a real vector and T bijection (one-to-one mapping) with finite continuous partial derivatives a.e. on α^*

$$J_{ij}(\alpha) \equiv \frac{\partial T_i(\alpha)}{\partial \alpha_j} \tag{2.11}$$

for all entries T_i of T and entries α_j of α. Then,

$$f_T(T(\alpha))|J(\alpha)| = f(\alpha) \tag{2.12}$$

where $|\cdot|$ denotes determinant.

Proof: Proposition describes substitutions in multivariate integrals, see [13]. □

It is useful to summarise basic properties of expectation which help us to simplify formal manipulations.

Proposition 2.3.6 (Basic properties of $\mathcal{E}$) *For arbitrary functions $\mathcal{Z}_1(\cdot)$, $\mathcal{Z}_2(\cdot)$ on which $\mathcal{E}[\cdot|\gamma]$ is well defined, the expectation $\mathcal{E}[\cdot|\gamma]$ has the following properties:*

Isotonic nature of $\mathcal{E}[\cdot|\gamma]$: $\mathcal{Z}_1 \leq \mathcal{Z}_2 \Rightarrow \mathcal{E}[\mathcal{Z}_1|\gamma] \leq \mathcal{E}[\mathcal{Z}_2|\gamma]$;

Linearity of $\mathcal{E}[\cdot|\gamma]$: $\mathcal{E}[A(\gamma)\mathcal{Z}_1 + B(\gamma)Z_2|\gamma] = A(\gamma)\mathcal{E}[\mathcal{Z}_1|\gamma] + B(\gamma)\mathcal{E}[\mathcal{Z}_2|\gamma]$
for arbitrary coefficients A, B depending at most on the condition γ;

Chain rule for expectation: $\mathcal{E}\left[\mathcal{E}[\cdot|\gamma,\zeta]|\gamma\right] = \mathcal{E}[\cdot|\gamma]$
for an arbitrary additional condition ζ.

Conditional covariance *of a vector* α

$$\mathrm{cov}[\alpha|\gamma] \equiv \mathcal{E}\left[(\alpha - \mathcal{E}[\alpha|\gamma])(\alpha - \mathcal{E}[\alpha|\gamma])'|\gamma\right]$$

is related to the non-central moments through the formula

$$\mathrm{cov}[\alpha|\gamma] = \mathcal{E}[\alpha\alpha'|\gamma] - \mathcal{E}[\alpha|\gamma]\mathcal{E}[\alpha'|\gamma], \quad (' \text{ is transposition}). \tag{2.13}$$

Jensen inequality *bounds expectation of a convex function* $T_\gamma : \alpha^* \rightarrow (-\infty, \infty)$

$$\mathcal{E}[T_\gamma(\alpha)|\gamma] \geq T_\gamma\left(\mathcal{E}[\alpha|\gamma]\right). \tag{2.14}$$

Proof: All statements can be verified by using the integral expression (2.6) of the expectation. Proof of Jensen inequality can be found e.g. in [14]. □

Remarks 2.3.4

1. *Proposition is formulated for the conditional expectation. The unconditional case is formally obtained by omitting the condition used.*
2. *Note that whenever the expectation is applied to a vector (matrix) function* V *it should be understood as the vector (matrix) of expectations*

$$\left[\mathcal{E}(V)\right]_i \equiv \mathcal{E}(V_i). \tag{2.15}$$

Problem 2.3.2 (How to respect uncertainties of pdfs?) *Formulae combining pdfs are deterministic. At the same time, we are always uncertain whether our options of individual elements creating them are "correct". Thus, a systematic way how to respect this uncertainty seems to be desirable. The use of upper and lower probabilities [15] or belief functions [16] represents advanced attempts to respect this situation. In our opinion, they are not still matured enough in addressing the discussed problem, at least within the considered context.*

2.4 Static design

The optimal selection of *admissible* decision rules relies on the following key proposition that converts minimisation over mappings to "ordinary" minimisation.

Proposition 2.4.1 (Basic decision-making lemma) *The optimal* admissible *decision rule* $\mathcal{R}^o$

$$\mathcal{R}^o(\mathcal{P}_a) \equiv a^o(\mathcal{P}_a), \ \forall \mathcal{P}_a \in \mathcal{P}_a^*$$

minimising the expected loss function (2.6) can be constructed value-wise. To each $\mathcal{P}_a \in \mathcal{P}_a^*$*, a minimising argument* $a^o(\mathcal{P}_a)$ *in*

$$\min_{a \in a^*} \mathcal{E}[\mathcal{Z}(\mathcal{P}_a, a, \mathcal{F}_a)|\mathcal{P}_a, a] \tag{2.16}$$

is assigned to it.

The obtained minimum has the value

$$\min_{\mathcal{R}:\, \mathcal{P}_a^* \rightarrow a^*} \mathcal{E}[\mathcal{Z}(\mathcal{P}_a, a, \mathcal{F}_a)] = \mathcal{E}\{\min_{a \in a^*} \mathcal{E}[\mathcal{Z}(\mathcal{P}_a, a, \mathcal{F}_a)|\mathcal{P}_a, a]\}. \tag{2.17}$$

Proof: Let us fix an arbitrary $\mathcal{P}_a \in \mathcal{P}_a^*$. The definition of minimum implies

$$\mathcal{E}[\mathcal{Z}(\mathcal{P}_a, \mathcal{R}^o(\mathcal{P}_a), \mathcal{F}_a)|\mathcal{P}_a, \mathcal{R}^o(\mathcal{P}_a))] \le \mathcal{E}[\mathcal{Z}(\mathcal{P}_a, a, \mathcal{F}_a)|\mathcal{P}_a, a], \ \forall a \in a^*.$$

Let an admissible rule $\mathcal{R} : \mathcal{P}_a^* \to a^*$ assign a decision $a \in a^*$ to the considered $\mathcal{P}_a$. Then, the previous inequality can be written

$$\mathcal{E}[\mathcal{Z}(\mathcal{P}_a, \mathcal{R}^o(\mathcal{P}_a), \mathcal{F}_a)|\mathcal{P}_a, \mathcal{R}^o(\mathcal{P}_a)] \le \mathcal{E}[\mathcal{Z}(\mathcal{P}_a, \mathcal{R}(\mathcal{P}_a), \mathcal{F}_a)|\mathcal{P}_a, \mathcal{R}(\mathcal{P}_a)].$$

We apply unconditional expectation $\mathcal{E}[\cdot]$ acting on functions of $\mathcal{P}_a$ to this inequality. Due to the isotonic nature of $\mathcal{E}$, the inequality is preserved. The chain rule for expectations, see Proposition 2.3.6, implies that on the left hand side of the resulting inequality the value of the unconditional expected loss for $\mathcal{R}^o$ is found. $\mathcal{R}^o$ assigns to each $\mathcal{P}_a \in \mathcal{P}_a^*$ the decision $a^o(\mathcal{P}_a)$. On the right hand side of the discussed inequality, we geet the unconditional expected loss corresponding to an arbitrarily chosen $\mathcal{R} : \mathcal{P}_a^* \to a^*$. $\square$

Proposition and its proof imply no preferences if there is more absolutely minimising arguments $a^o(\mathcal{P}_a)$. We can use any of them or switch between them in a random manner whenever the pdf $f(a_t|\mathcal{P}_{a_t})$ has its support concentrated on them. This is an example where a *randomised causal strategy* may occur. We specify it formally as it is used later on more extensively.

Agreement 2.4.1 (Outer model of strategy) *The collection of pdfs* $\{f(a_t|\mathcal{P}_{a_t})\}_{t \in t^*}$ *forms the* outer model of the decision strategy.

The decision strategy is called randomised *if the support of some* $f(a_t|\mathcal{P}_{a_t})$ *contains at least two values of* a_t.

Remarks 2.4.1

1. *We do not enter the technical game with ε-optimality: the existence of the indicated minima is supposed.*

2. *It is worth stressing that the optimal decision rule is constructed value-wise. Formally, the minimisation should be performed for all possible instances of experience $\mathcal{P}_a \in \mathcal{P}_a^*$ in order to get the decision rule.*

   Often, we are interested in the optimal decision for a given, say observed, experience. Then just a single minimisation is necessary. This is typically the case of the estimation problem. This possibility makes the main distinction from the dynamic design, when optimal strategy, a sequence of decision rules, is searched for like in control design. In this case, discussed in next Section, the construction of decision rules is necessary. This makes the dynamic design substantially harder and mostly infeasible [5, 6].

2.5 Dynamic design

We are searching for the optimal admissible strategy. We consider the usual case, when its individual rules use a non-decreasing sequence of experience. The extending experience models increasing amount of data available for the decision-making.

2.5.1 Dynamic programming

The optimal admissible strategy can be found by using a stochastic version of celebrated *dynamic programming* [17]. It is nothing but a repetitive application of Proposition 2.4.1.

Proposition 2.5.1 (Stochastic dynamic programming) *The optimal causal strategy* $\{\mathcal{R}^o : \mathcal{P}^*_{a_t} \rightarrow a^*_t\}_{t\in t^*} \in \{\mathcal{R}_t : \mathcal{P}^*_{a_t} \rightarrow a^*_t\}_{t\in t^*}$ *acting on non-decreasing sequence of experience* $\mathcal{P}_{a_t} \subset \mathcal{P}_{a_{t+1}}$ *and minimising the expected loss function* $\mathcal{E}[\mathcal{Z}(\mathcal{Q})]$ *can be constructed in a value-wise way. For every* $t \in t^*$ *and each* $\mathcal{P}_{a_t} \in \mathcal{P}^*_{a_t}$*, it is sufficient to take a minimising argument* $a^o(\mathcal{P}_{a_t})$ *in*

$$\mathcal{V}(\mathcal{P}_{a_t}) = \min_{a_t\in a^*_t} \mathcal{E}[\mathcal{V}(\mathcal{P}_{a_{t+1}})|\mathcal{P}_{a_t}, a_t],\ t \in t^*. \tag{2.18}$$

as the decision generated by the the t-th rule of the optimal strategy, i.e. $a^o(\mathcal{P}_{a_t}) = \mathcal{R}^o(\mathcal{P}_{a_t})$.

The recursion (2.18) is performed in the backward manner against the course given by the increasing experience. It starts with

$$\mathcal{V}(\mathcal{P}_{a_{\bar{t}+1}}) \equiv \mathcal{Z}(\mathcal{Q}). \tag{2.19}$$

The reached minimum has the value

$$\min_{\{\mathcal{R}_t:\,\mathcal{P}^*_{a_t}\rightarrow a^*_t\}_{t\in t^*}} \mathcal{E}[\mathcal{Z}(\mathcal{Q})] = \mathcal{E}[\mathcal{V}(\mathcal{P}_{a_1})].$$

Proof: Let us define $\mathcal{P}_{a_{\bar{t}+1}} \equiv \mathcal{Q}$ in order to get a unform notation It is legitime as $a_{\bar{t}+1}$ is not optimised. The definition of minimum and Proposition 2.4.1 imply

$$\min_{\{\mathcal{R}_t:\,\mathcal{P}^*_{a_t}\rightarrow a^*_t\}_{t\in t^*}} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a_{\bar{t}+1}})] = \min_{\{\mathcal{R}_t:\,\mathcal{P}^*_{a_t}\rightarrow a^*_t\}_{t<\bar{t}}} \min_{\mathcal{R}_{\bar{t}}:\,\mathcal{P}^*_{a_{\bar{t}}}\rightarrow a^*_{\bar{t}}} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a_{\bar{t}+1}})] \underbrace{=}_{(2.17)}$$

$$= \min_{\{\mathcal{R}_t:\,\mathcal{P}^*_{a_t}\rightarrow a^*_t\}_{t<\bar{t}}} \mathcal{E}\left[\min_{a_{\bar{t}}\in a^*_{\bar{t}}} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a_{\bar{t}+1}})|\mathcal{P}_{a_{\bar{t}}}, a_{\bar{t}}]\right].$$

Denoting

$$\mathcal{V}(\mathcal{P}_{a_{\bar{t}}}) \equiv \min_{a_{\bar{t}}\in a^*_{\bar{t}}} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a_{\bar{t}+1}})|\mathcal{P}_{a_{\bar{t}}}, a_{\bar{t}}]$$

we proved the first step of the recursion and specified the start (2.19). The following step becomes

$$\min_{\{\mathcal{R}_t:\,\mathcal{P}^*_{a_t}\rightarrow a^*_t\}_{t<\bar{t}}} \mathcal{E}\left[\mathcal{V}(\mathcal{P}_{a_{\bar{t}}})\right].$$

We face the identical situation as above with the horizon decreased by one and the procedure can be repeated till the initial rule $\mathcal{R}_1$ is constructed. □

The optimisation relies on our ability to evaluate the expectations

$$\mathcal{E}[\mathcal{V}(\mathcal{P}_{a_{t+1}})|\mathcal{P}_{a_t}, a_t] = \int_{\Delta^*_{a_t}} \mathcal{V}(\mathcal{P}_{a_t}, a_t, \Delta_{a_t}) f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)\, d\Delta_{a_t}.$$

The introduced *innovation* Δ_{a_t} contains those observable quantities that cannot be used for the choice of a_t. They are not in $\mathcal{P}_{a_t}$ but they belong to $\mathcal{P}_{a_{t+1}} \setminus a_t$. The pdfs $\{f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)\}_{t\in t^*}$ model the relationships of Δ_{a_t} to $\mathcal{P}_t$ and a_t.

Agreement 2.5.1 (Outer model of the system) *The collection of pdfs*

$$\{f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)\}_{t\in t^*}, \tag{2.20}$$

required for the optimal design, is called outer model of the system.

Remarks 2.5.1

1. *The term outer model of the system is abbreviated to model of the system or even to model. The exact meaning is clear from the specific context.*

2. *Often, the innovation* $\Delta_{a_t} = y_t =$ *observable output of the system. If the set point to which the output should be driven by the chosen input* u_t *is uncertain it has to be included into* Δ_{u_t}.

The following agreement is used in a presentation of the most common version of dynamic programming.

Agreement 2.5.2 (Data driven design; Internal quantities) *The design is called* data driven *if realisations* $\mathcal{Q}$ *consist of optional and potentially observable quantities, i.e. with ignorance consisting of unobserved data only*

$$\mathcal{Q} \equiv (\Delta(\bar{t}), a(\bar{t})) \equiv (\mathcal{P}_a, a, \mathcal{F}_a). \tag{2.21}$$

Quantities in $\mathcal{Q}$ *that are never observed directly are called* internal quantities.

Proposition 2.5.2 (Dynamic programming for additive loss function) *Let us consider data driven design and search for the optimal admissible strategy* $\{\mathcal{R}_t : \mathcal{P}^*_{a_t} \to a^*_t\}_{t\in t^*}$ *acting on a non-decreasing sequence of experience* $\{\mathcal{P}_{a_t}\}_{t\in t^*}$.

Then, the optimal strategy $\{\mathcal{R}^o : \mathcal{P}^*_{a_t} \to a^*_t\}_{t\in t^*}$ *minimising the expected additive loss function*

$$\mathcal{E}\left[\sum_{t\in t^*} z(\Delta(t), a(t))\right] \tag{2.22}$$

can be constructed value-wise. For all $t \in t^*$ *and to each* $\mathcal{P}_{a_t} \in \mathcal{P}^*_{a_t}$, *a minimising argument* $a^o(\mathcal{P}_{a_t}) = \mathcal{R}^o(\mathcal{P}_{a_t})$ *in*

$$\mathcal{V}(\mathcal{P}_{a_t}) = \min_{a_t \in a^*_t} \mathcal{E}[z(\Delta(t), a(t)) + \mathcal{V}(\mathcal{P}_{a_{t+1}})|\mathcal{P}_{a_t}, a_t], \ t \in t^*. \tag{2.23}$$

is assigned.

The recursion (2.23) is performed in the backward manner against the course given by the increasing experience, starting from

$$\mathcal{V}(\mathcal{P}_{a_{\bar{t}+1}}) \equiv 0. \tag{2.24}$$

The reached minimum has the value

$$\min_{\{\mathcal{R}_t : \mathcal{P}^*_{a_t} \to a^*_t\}_{t\in t^*}} \mathcal{E}[\mathcal{Z}(\mathcal{P}_{a_{\bar{t}+1}})] = \mathcal{E}[\mathcal{V}(\mathcal{P}_{a_1})].$$

Proof: It follows exactly the line of Proposition 2.5.1 with a modified definition of the function $\mathcal{V}(\cdot)$:

$$\mathcal{V}(\mathcal{P}_{a_t}) \equiv \min_{\{\mathcal{R}_\tau:\,\mathcal{P}^*_{a_\tau}\to a^*_\tau\}_{\tau\geq t}} \mathcal{E}\left[\sum_{\tau\geq t} z(\Delta(\tau),a(\tau))|\mathcal{P}_{a_t},a_t=\mathcal{R}_t(\mathcal{P}_{a_t})\right]. \quad (2.25)$$

□

Agreement 2.5.3 (Bellman function) *The function* $z(\Delta(t),a(t))$ *is called the* partial loss. *The function* $\mathcal{V}(\cdot)$ *occurring in dynamic programming is called* Bellman function. *Bellman function* $\mathcal{V}(\cdot)$ *(2.25) is called* optimal cost-to-go.

Problem 2.5.1 (How to restrict the search space?) *The value-wise construction of the optimal strategy is quite demanding in the dynamic design as it has to be done for all possible realisations of experience. A reduction to neighbourhoods of real realisations would simplify computations much. There are well-established techniques in particular cases [18] but a systematic efficient approach is still missing (if it can be found at all).*

2.5.2 Fully probabilistic design

A specific but rather general design is formulated and solved here that is believed to form a bridge between optimal and *practically* optimal design.

Here, and in many other places, we need the notion of *Kullback-Leibler distance* [19] that measures well proximity of a pair of pdfs.

Agreement 2.5.4 (Kullback-Leibler distance) *Let* f,g *be a pair of pdfs acting on a common set* x^*. *Then, Kullback-Leibler distance* $\mathcal{D}(f||g)$ *is defined by the formula*

$$\mathcal{D}(f||g) \equiv \int_{x^*} f(x)\ln\left(\frac{f(x)}{g(x)}\right)dx. \quad (2.26)$$

Proposition 2.5.3 (Basic properties of Kullback-Leibler distance) *Let* f,g *be a pair of pdfs acting on a same set. It holds*

$\mathcal{D}(f||g) \geq 0$,

$\mathcal{D}(f||g) = 0$ *iff* $f = g$ *(a.e.),*

$\mathcal{D}(f||g) = \infty$ *iff on a set of a positive measure* $f > 0$ *and* $g = 0$,

$\mathcal{D}(f||g) \not\equiv \mathcal{D}(g||f)$; *Kulback-Leibler distance* does not obey *triangle inequality.*

Proof: See, for instance, [14] □

Now we are ready to formulate and solve the fully probabilistic design problem. A simple version is presented here considering a simpler data driven design, see Agreement 2.5.2. In this case, the joint pdf $f(\mathcal{Q}) \equiv f(\Delta(\bar{t}),a(\bar{t}))$ describing observable quantities of interest can be factorised by a repetitive use of the chain rule, see Proposition 2.3.4,

$$f(\Delta(\bar{t}),a(\bar{t})) = \prod_{t\in t^*} f(\Delta_{a_t}|\mathcal{P}_{a_t},a_t)f(a_t|\mathcal{P}_{a_t}). \quad (2.27)$$

The first factors $\{f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)\}_{t\in t^*}$ under the product sign describe possible reactions of the system on the decision a_t and under the situation reflected in experience $\mathcal{P}_{a_t}$. These pdfs form the outer model of the system, see Agreement 2.5.1. Similarly, $\{f(a_t|\mathcal{P}_{a_t})\}_{t\in t^*}$ represent an outer model of a randomised decision strategy to be chosen, see Agreement 2.4.1. Looking at the joint pdf (2.27), it seems to be "natural" to formulate the design as an attempt to make this pdf as close as possible to some "ideal" joint pdf.

Agreement 2.5.5 (Fully probabilistic design) *The fully probabilistic data driven design specifies its target through an* ideal pdf ${}^{[I]}f$

$$ {}^{[I]}f(\mathcal{Q}) = \prod_{t\in t^*} {}^{[I]}f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)\, {}^{[I]}f(a_t|\mathcal{P}_{a_t}). $$

The optimal admissible, possibly randomised, decision strategy is defined as a minimiser of Kullback-Leibler distance (2.26) of $f(\Delta(\bar{t}), a(\bar{t}))$ and ${}^{[I]}f(\Delta(\bar{t}), a(\bar{t}))$

$$ \mathcal{D}(f||\,{}^{[I]}f) \equiv \int f(\Delta(\bar{t}), a(\bar{t})) \ln\left(\frac{f(\Delta(\bar{t}), a(\bar{t}))}{{}^{[I]}f(\Delta(\bar{t}), a(\bar{t}))}\right) d(\Delta(\bar{t}), a(\bar{t})). \tag{2.28} $$

Proposition 2.5.4 (Solution of fully probabilistic design) *The optimal strategy minimising Kullback-Leibler distance (2.28) has the form*

$$ \begin{aligned} f(a_t|\mathcal{P}_{a_t}) &= {}^{[I]}f(a_t|\mathcal{P}_{a_t}) \frac{\exp[-\omega(\mathcal{P}_{a_t}, a_t)]}{\gamma(\mathcal{P}_{a_t})}, \quad \textit{where} && (2.29) \\ \gamma(\mathcal{P}_{a_t}) &\equiv \int_{a_t^*} {}^{[I]}f(a_t|\mathcal{P}_{a_t}) \exp[-\omega(\mathcal{P}_{a_t}, a_t)]\, da_t \\ \omega(\mathcal{P}_{a_t}, a_t) &\equiv \int_{\Delta^*_{a_t}} f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t) \ln\left(\frac{f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)}{{}^{[I]}f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)\gamma(\mathcal{P}_{a_{t+1}})}\right) d\Delta_{a_t} \\ \gamma(\mathcal{P}_{a_{\bar{t}+1}}) &= 1. && (2.30) \end{aligned} $$

The solution is performed against the time course, starting at $t = \bar{t}$.

Proof: Proof uses essentially dynamic programming for a special form of additive loss and simplest results from a variation calculus. It is, however, a bit lengthy, for details see in [20, 21]. □

Remarks 2.5.2

1. *At a descriptive level, the usual dynamic programming consists of a sequence of evaluation pairs*

   $$ (\textit{take conditional expectation, minimise}). $$

   Except of a few numerically solvable cases, some approximation techniques have to be employed. The complexity of the approximated optimum prevents a systematic use of the standard approximation theory and various heuristic schemes are adopted, see Chapter 5.

   The fully probabilistic design finds minimisers explicitly and reduces the design to a sequence of conceptually feasible multivariate integrations.

2. *The found optimal strategy is randomised and obviously causal one. The technological restriction are met trivially if the chosen ideal strategy meets them, if* $\mathrm{supp}\left[\, {}^{[I]}f(a_t|\mathcal{P}_{a_t})\right] \subset a_t^*$, *cf. (2.29).*

Problem 2.5.2 (How to approximate probabilistic design?) *A good methodology of an (approximate) evaluation of the recursion (2.30) has to be developed. The need do store (interpolate) functions of many variables makes exact solution computationally hard, see discussion in Section 5.1.*

Problem 2.5.3 (Is there a design compatible with general situation?) *The presented design is data driven. It combines well-fitting elements, namely, the chain rule (Proposition 2.3.4) that involves products of conditional pdfs and Kullback-Leibler distance (2.26), i.e. the expected value of the logarithm of a ratio of products defining the joint and the ideal joint pdfs of data. A similar compatibility would be desirable when we want to influence time varying inner states, when a more complex composition rules, see Section 2.6.1, are involved.*

2.6 Learning

Generally, considered behaviour $\mathcal{Q}^*$ contains quantities that are never observed directly and, in spite of this, we want to describe them or influence them. We called them internal quantities, see Agreement 2.5.2. Even in this case, the optimal decision-making needs to get the outer model (2.20), see Proposition 2.5.1, that should reflect them. Here we describe how to get it. The presented construction is of an independent interest as a good *formal model of learning*.

2.6.1 Bayesian filtration

This section deals with a general case when the construction of the outer model relies on filtering, i.e. learning of *time variant* internal (directly unobservable) quantities. It requires the following elements.

Requirement 2.6.1 (Models; natural conditions of decision-making)

1. *The innovations* Δ_{a_t} *are related to experience* $\mathcal{P}_{a_t}$ *and decisions* a_t *through the* observation model

$$\{f(\Delta_{a_t}|\Theta_t, \mathcal{P}_{a_t}, a_t)\}_{t\in t^*} \tag{2.31}$$

that is known up to unknown internal quantities $\Theta_t \in \Theta_t^* \subset \mathcal{F}_{a_\tau}$, $\forall \tau \in t^*$.

2. *The evolution of the quantities* $\Theta(\bar{t}) \in \Theta^*(\bar{t})$ *is described by the known collection of pdfs called the* time evolution model

$$\{f(\Theta_t|\Theta_{t-1}, \mathcal{P}_{a_t}, a_t)\}_{t\in t^*}. \tag{2.32}$$

3. *The quantities* $\Theta(\bar{t})$ *are unknown to the strategies considered. The* natural conditions of decision-making *(a slight generalisation of natural conditions of control [10]) express it formally. They postulate independence of* a_t *and* Θ_t *when conditioned on* $\mathcal{P}_{a_t}$

$$f(a_t|\Theta_t, \mathcal{P}_{a_t}) = f(a_t|\mathcal{P}_{a_t}) \quad \underbrace{\Leftrightarrow}_{Proposition\ 2.3.4} \quad f(\Theta_t|\mathcal{P}_{a_t}, a_t) = f(\Theta_t|\mathcal{P}_{a_t}). \tag{2.33}$$

4. *The initial values of* Θ_0, $\mathcal{P}_{a_1}$ *add nothing new to the prior information so that the* prior pdf

$$f(\Theta_1) \equiv f(\Theta_1|\mathcal{P}_{a_1}, a_1) \underbrace{=}_{(2.33)} f(\Theta_1|\mathcal{P}_{a_1}) \qquad (2.34)$$

is the first term in the sequence of pdfs (2.32).

Remarks 2.6.1

1. *Often, the unknown quantities* Θ_t *together with the decision* a_t *are assumed to describe the involved conditional pdfs fully, i.e.* $\mathcal{P}_{a_t}$ *can be omitted. Then, these quantities can be identified with the* information state.

2. *The natural conditions of decision-making express the assumption that* $\Theta_t \notin \mathcal{P}_{a_\tau}$ $\forall \tau \in t^*$. *Thus, values of* Θ_t *cannot be used by the decision rules forming admissible strategy. Alternatively, see (2.33), we cannot gain information about* Θ_t *from the decision* a_t *if the corresponding innovation* Δ_{a_t} *(the corresponding reaction of the system) is not available.*

   The natural conditions of decision-making are "naturally" fulfilled by strategies we are designing. They have to be assumed when data influenced by "externally chosen" strategy are processed.

Proposition 2.6.1 (Generalised Bayesian filtering) *Let Requirement 2.6.1 be met. Then, the outer model of the system (2.20) is given by the formula*

$$f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t) = \int_{\Theta_t^*} f(\Delta_{a_t}|\Theta_t, \mathcal{P}_{a_t}, a_t) f(\Theta_t|\mathcal{P}_{a_t})\, d\Theta_t. \qquad (2.35)$$

The evolution of the pdf $f(\Theta_t|\mathcal{P}_{a_t})$, *called (generalised Bayesian)* filtration of unknown quantities, *is described by the following recursion that starts from the prior pdf* $f(\Theta_1)$:

- *Data updating*

$$f(\Theta_t|\mathcal{P}_{a_{t+1}}) = \frac{f(\Delta_{a_t}|\Theta_t, \mathcal{P}_{a_t}, a_t) f(\Theta_t|\mathcal{P}_{a_t})}{f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)} \propto f(\Delta_{a_t}|\Theta_t, \mathcal{P}_{a_t}, a_t) f(\Theta_t|\mathcal{P}_{a_t}) \qquad (2.36)$$

  that incorporates the innovation Δ_{a_t} *and the decision* a_t, *and*

- Time updating

$$f(\Theta_{t+1}|\mathcal{P}_{a_{t+1}}) = \int_{\Theta_t^*} f(\Theta_{t+1}|\Theta_t, \mathcal{P}_{a_{t+1}}, a_{t+1}) f(\Theta_t|\mathcal{P}_{a_{t+1}})\, d\Theta_t \qquad (2.37)$$

  that reflects the time evolution $\Theta_t \to \Theta_{t+1}$.

Proof: It is a simple but instructive exercise in calculus with pdfs (marginalisation, Bayes rule) and natural conditions of decision-making (2.33). □

Agreement 2.6.1 (Filtering; predictive pdf) *The process of generating filtration is called (generalised Bayesian)* filtering.

The outer model of the system obtained by filtering is called predictive pdf.

Remarks 2.6.2

1. *The term* predictive pdf *reflects the way how the outer model of the system has been obtained. It uses the observed experience and extrapolates it into ignorance* assuming that the mechanism of generating Θ_t does not change.

   This accumulation of experience and its extrapolation represent a good formal model of learning.

2. *It has to be stressed that the accumulation of experience can take place only when the rules governing the behaviour are not changed during it, i.e. when we can rely on the validity of the underlying models.*

3. *The construction of the predictive pdf is our key motivation for filtering. Its results are often of an independent interest.*

4. *Under the natural conditions of decision-making, the filtering relies on the knowledge of decisions and not on the knowledge of rules* $\mathcal{R} : \mathcal{P}_a^* \to a^*$ *generating them. It is practically important when we learn under closed loop conditions, especially with a human decision maker.*

5. *The time evolution model* $f(\Theta_t|\Theta_{t-1}, \mathcal{P}_{a_t}, a_t)$ *as well as the observation model* $f(\Delta_{a_t}|\Theta_t, \mathcal{P}_{a_t}, a_t)$ *have to result from a theoretical modelling of the system in question. Such a modelling uses both field knowledge, like laws of conservation, and approximation capabilities of the selected family of the models involved. Often, deterministic relationships are modelled and then the "deviations" from an "expected" trajectory are described. For a complementary way see Section 4.1.3.*

6. *The prior pdf* $f(\Theta_1)$ *allows us to introduce an information based on expert knowledge or analogy to situations observed previously.*

7. *The observed data, the only bridge to reality, enter the evaluations in the data updating step only when the newest innovation-decision pair is processed. This simple observation is important for approximation of the time evolution model, see Section 5.2.2.*

8. *In summary,* the described Bayesian filtering combines prior information $(f(\Theta_1))$, theoretical knowledge of the specific fields $(f(\Delta_{a_t}|\Theta_t, \mathcal{P}_{a_t}, a_t)$, $f(\Theta_t|\Theta_{t-1}, \mathcal{P}_{a_t}, a_t))$ and observed data $(\Delta(\bar{t}), a(\bar{t}))$ by using coherent deductive rules of the calculus with pdfs.

   This combination of information sources is a powerful internally consistent framework describing the essence of learning. Due to its deductive structure, the incorrect modelling not an incorrect information processing can be blamed for a failure of the specific learning process.

Problem 2.6.1 (How to approximate filtering?) *The general filtering equations can rarely be solved analytically [22] or numerically [23]. In our opinion, this area is still waiting for efficient general approximate evaluation techniques. We claim it in spite of the progress observed and represented, for instance, by the family of bootstrap methods [24], see Section 5.3.2.*

2.6.2 Bayesian estimation

This section deals with a special version of filtering called *estimation*. It arises when the internal quantities Θ_t are time invariant

$$\Theta_t = \Theta, \ \forall t \in t^*. \tag{2.38}$$

The common value Θ is called *parameter*. In this case, the time evolution model is $f(\Theta_t|\Theta_{t-1}, \mathcal{P}_{a_t}, a_t) = \delta(\Theta_t - \Theta_{t-1})$. The employed Dirac delta function $\delta(\cdot)$ is a formal pdf of the measure fully concentrated on the zero argument.

Proposition 2.6.2 (Generalised Bayesian estimation) *Let Requirement 2.6.1 be met with time invariant $\Theta_t = \Theta \in \Theta^* \subset \mathcal{F}_{a_\tau}$, $\forall \tau \in t^*$. Then, the outer model of the system (2.20) is given by the formula*

$$f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t) = \int_{\Theta^*} f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) f(\Theta|\mathcal{P}_{a_t})\, d\Theta. \tag{2.39}$$

The evolution of the pdf $f(\Theta|\mathcal{P}_{a_t})$, called (generalised Bayesian) parameter estimate *or* posterior pdf of unknown parameter, *is described by the recursion identical with data updating (2.36)*

$$f(\Theta|\mathcal{P}_{a_{t+1}}) = \frac{f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) f(\Theta|\mathcal{P}_{a_t})}{f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)} \propto f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) f(\Theta|\mathcal{P}_{a_t}). \tag{2.40}$$

It starts from the prior pdf $f(\Theta) \equiv f(\Theta|\mathcal{P}_{a_1}, a_1) = f(\Theta|\mathcal{P}_{a_1})$.

The simplicity of the estimation formula allows us to write down its (non-recursive) batch version

$$f(\Theta|\mathcal{P}_{a_{t+1}}) = \frac{\prod_{\tau \le t} f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau) f(\Theta)}{\int_{\Theta^*} \prod_{\tau \le t} f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau) f(\Theta)\, d\Theta} \equiv \frac{\mathcal{L}(\Theta, \mathcal{P}_{a_{t+1}}) f(\Theta)}{\mathcal{I}(\mathcal{P}_{a_{t+1}})}. \tag{2.41}$$

The introduced likelihood function

$$\mathcal{L}(\Theta, \mathcal{P}_{a_{t+1}}) \equiv \prod_{\tau \le t} f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau) \tag{2.42}$$

evolves according to the recursion identical with that for the posterior pdf (2.40) but it starts from the $\mathcal{L}(\Theta, \mathcal{P}_{a_1})$ identically equal to 1.

The normalisation factor $\mathcal{I}(\cdot)$ is defined by the formula

$$\mathcal{I}(\mathcal{P}_{a_{t+1}}) = \int \mathcal{L}(\Theta, \mathcal{P}_{a_{t+1}}) f(\Theta)\, d\Theta. \propto f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t) \tag{2.43}$$

With it, the outer model of the system (2.20) can be alternatively expressed

$$f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t) = \frac{\mathcal{I}(\mathcal{P}_{a_{t+1}})}{\mathcal{I}(\mathcal{P}_{a_t})}. \tag{2.44}$$

Proof: It is again a simple exercise in calculus with pdfs (marginalisation, Bayes rule) and natural conditions of decision-making (2.33). □

Remarks 2.6.3

1. *The observation model* $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t)$ *is called* parameterised model *whenever estimation problem is considered. We respect this tradition.*

2. *Note that the recursive evolution of the pdf* $f(\Theta|\mathcal{P}_{a_t})$ *allows us to interpret the posterior pdf as the prior one before processing new observations.*

3. *The data inserted into the "objective" parameterised (observation) model correct gradually the subjectively chosen prior pdf* $f(\Theta)$. *The posterior pdf* $f(\Theta|\mathcal{P}_{a_t})$ *reflects always both ingredients. If the data are informative enough, the relative contribution of the single subjective factor* $f(\Theta)$ *to the posterior pdf is decreasing with increasing* t *as the likelihood function* $\mathcal{L}(\Theta, \mathcal{P}_{a_{t+1}})$ *contains* t *"objective" factors, see Proposition 4.1.3.*

4. *Zero values are preserved by multiplication. Thus, the posterior pdf redistributes probability mass only within the support of the prior pdf, i.e. within the set*
$$\mathrm{supp}\,[\,f(\Theta)] \equiv \{\Theta \in \Theta^* : \, f(\Theta) > 0\}.$$
This fact allows us to introduce hard bounds on possible parameter values but prevents us to "learn" about parameters Θ *out of the support* $\mathrm{supp}\,[\,f(\Theta)]$, *see also Proposition 4.1.3.*

5. *Remarks 2.6.2 related to filtering apply mostly to estimation, too: the parameter estimation is a task on its own; unknown parameters are always in the ignorance of the decision to be chosen; under the natural conditions of decision-making (2.33), decisions are needed and not the rules* $\mathcal{R}_t : \mathcal{P}^*_{a_t} \to a^*_t$ *generating them.*

6. *The internal quantities* Θ_t *are usually assumed to be finite-dimensional in order to avoid technicalities related to measure theory. In exceptional cases, like in description of so called equivalence approach, see Section 5.2.1, we deal with* potentially infinite-dimensional parameter. *It means that the number of unknown quantities is finite but increases without limitations. This case is often called* non-parametric estimation. *A useful and relatively general view on it is presented in Chapter 6.*

Problem 2.6.2 (How to approximate estimation?) *Estimation as a special case of filtering also waits for efficient evaluation techniques. In a sense, it is simpler as time updating is missing. At the same time, it is harder as the (usual) stabilising effect of the time updating cannot be exploited. An advanced and promising solution as well as review of the current state of the art can be found in [25], see also Section 5.2.1.*

Chapter 3

Asymptotics

The analysis outlined here serves us mainly for interpretation purposes. Thus, all technicalities are suppressed as much as possible.

3.1 Estimation

The "objective" pdf $f(\mathcal{Q})$, see Section 2.3.1, describing the system behaviour is denoted here ${}^{[o]}f(\mathcal{Q})$. The coresponding outer model of the system $f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)$ is denoted ${}^{[o]}f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)$. Its relationship to the predictive pdf $f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)$ obtained through the parameter estimation, Proposition 2.6.2, is inspected.

For the analysis, the notion of *entropy rate* $\mathcal{H}_\infty({}^{[o]}f||\Theta)$ is needed. For each $\Theta \in \Theta^*$, it is defined by the formula

$$\mathcal{H}_\infty({}^{[o]}f||\Theta) \equiv \lim_{t\to\infty} \sup \mathcal{H}_t({}^{[o]}f||\Theta) \tag{3.1}$$
$$\equiv \lim_{t\to\infty} \sup \frac{1}{t} \sum_{\tau\le t} \int_{\Delta^*_{a_\tau}} {}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau) \ln\left(\frac{{}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau)}{f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau)}\right) d\Delta_{a_\tau}.$$

Proposition 3.1.1 (Asymptotics of estimation) *Let natural conditions of decision-making (2.33) hold and $C_\Theta \le c < \infty$, $\bar{t}_\Theta \in \{1, 2, \ldots\}$ exist, for almost all $\Theta \in \Theta^*$, such that*

$$ {}^{[o]}f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t) \le C_\Theta f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t),\ \forall t > \bar{t}_\Theta,\ \forall \mathcal{P}_{a_{t+1}} \in \mathcal{P}^*_{a_{t+1}}. \tag{3.2}$$

Then, the posterior pdf $f(\Theta|\mathcal{P}_{a_t})$ (2.40) converges almost surely to a pdf $f(\Theta|\mathcal{P}_{a_\infty})$. It has the support coinciding with the set of minimising arguments in

$$\mathrm{supp}\,[\, f(\Theta|\mathcal{P}_{a_\infty})] = \mathrm{Arg} \min_{\Theta\in \mathrm{supp}[\, f(\Theta)]\cap\Theta^*} \mathcal{H}_\infty({}^{[o]}f||\Theta). \tag{3.3}$$

Proof: Under the natural conditions of decision-making (2.33), the posterior pdf (2.40) can be written in the form

$$f(\Theta|\mathcal{P}_{a_{t+1}}) \propto f(\Theta) \exp\left[-t\mathcal{H}(\mathcal{P}_{a_{t+1}}, \Theta)\right] \quad \text{with} \tag{3.4}$$
$$\mathcal{H}(\mathcal{P}_{a_{t+1}}, \Theta) = \frac{1}{t} \sum_{\tau\le t} \ln\left[\frac{{}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau)}{f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau)}\right]. \tag{3.5}$$

This form exploits the fact that un-normalised posterior pdf can be multiplied by any factor independent of Θ.

Let us fix the argument $\Theta \in \Theta^*$ and define

$$e_{\Theta;\tau} \equiv \ln\left[\frac{{}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau)}{f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau)}\right] - {}^{[o]}\mathcal{E}\left[\ln\left[\frac{{}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau)}{f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau)}\right] |\mathcal{P}_{a_\tau}, a_\tau\right]$$

$$\equiv \ln\left[\frac{{}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau)}{f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau)}\right] - \int {}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau) \ln\left[\frac{{}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau)}{f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau)}\right] d\Delta_{a_\tau}.$$

A direct check reveals that the introduced deviations $e_{\Theta;\tau}$ of values

$$\ln\left[\frac{{}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau)}{f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau)}\right]$$

from their conditional expectations ${}^{[o]}\mathcal{E}[\cdot|\mathcal{P}_{a_\tau}, a_\tau]$, given by ${}^{[o]}f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau)$, are zero mean and mutually uncorrelated. With them,

$$\mathcal{H}(\mathcal{P}_{a_{t+1}}, \Theta) = \mathcal{H}_t({}^{[o]}f||\Theta) + \frac{1}{t}\sum_{\tau \le t} e_{\Theta;\tau}.$$

The assumption (3.2) implies boundedness of the variance of $e_{\Theta;\tau}$. Consequently, the last term in the above expression converges to zero almost surely (a.s.), see [26], page 417. The first term on the right hand side of the last equality is non-negative as it can be viewed as a sum of Kullback-Leibler distances, see Proposition 2.5.3. Due to (3.2), it is also finite. Thus, (3.5) converges a.s. to the non-negative value $\mathcal{H}({}^{[o]}f||\Theta)$. The posterior pdf remains unchanged if we subtract $\min_{\Theta \in \mathrm{supp}[f(\Theta)] \cap \Theta^*} \mathcal{H}_\infty({}^{[o]}f||\Theta)$ from the exponent of its un-normalised version (3.4). Then, the exponent contains ($-t\times$ an_asymptotically_non-negative_factor). Thus, the posterior pdf $f(\Theta|\mathcal{P}_{a_\infty})$ may be asymptotically non-zero on minimising arguments (3.3) only. □

Remarks 3.1.1

1. *The entropy rate can be seen as an extension of Kullback-Leibler distance (2.26) that covers well asymptotic and controlled cases. It coincides with Kullback-Leibler distance in a range of particular cases.*

2. *The chosen assumption (3.2) can be weakened. It is, however, intuitively well acceptable. It excludes parameterised models, which assign zero belief to data generated by the system with a non-zero probability.*

3. *The Bayesian estimator minimises asymptotically the entropy rate between the objective pdf ${}^{[o]}f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)$ and a model chosen among candidates $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t)$, $\Theta \in \Theta^*$. In other words, a* best projection *of the objective pdf to the considered parameterised models is asymptotically found. The prior pdf can be interpreted as a prior belief assigned to the individual parameters $\Theta \in \Theta^*$ that the corresponding parameterised model is the best projection of the objective pdf [27]: not knowing the reality we do not know the best projection we finally arrive at.*

4. *The posterior pdf concentrates on a point if there is a unique minimiser of the entropy rate. In this case, the model is called* identifiable. *The identifiability can be influenced by:*

- *the considered class of the parameterised models and*
- *the decisions chosen, for instance, by the controller used (e.g. controller generating constant inputs does not allow us to learn their dynamic influence on outputs).*

5. *If the objective pdf* ${}^{[o]}f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)$ *coincides with* $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t)$ *for some* $\Theta = {}^{[o]}\Theta$ *with* $f({}^{[o]}\Theta) > 0$ *then* ${}^{[o]}\Theta$ *is in the support of the asymptotic posterior pdf* $f(\Theta|\mathcal{P}_{a_\infty})$. *If, moreover, the model is identifiable the objective pdf is asymptotically identified by the adopted Bayesian approach. This fact can be expressed in a more suggestive form:*

   Bayesian estimate is consistent whenever there is a consistent estimator.

6. *Often, a similar analysis is performed by measuring distance of parameterised models to empirical pdf of data [28]. It gives similar answers if the empirical pdf converges to the objective pdf. Moreover, it provides hints how to approximate the posterior pdf [25], see also Section 5.2.1. On the other hand, the known conditions of such convergence are more restrictive. For instance, an analysis of the controlled case is much harder.*

Problem 3.1.1 (How to unify statistics?) *Asymptotic analysis and finite-data oriented Bayesian approach are often perceived in an antagonistic way. Their simultaneous and harmonised use waits still for its full exploitation.*

3.2 Design

The asymptotics of the dynamic programming is analysed for horizon $\bar{t} \to \infty$ within this section. The data driven case with additive loss function (2.22) is considered. Note that the general loss function can always be converted in it by defining

$$z(\Delta(t), a(t))) = \begin{cases} \mathcal{Z}(\Delta(\bar{t}), a(\bar{t})) & \text{if } t = \bar{t} \\ 0 & \text{otherwise} \end{cases}. \tag{3.6}$$

We deal, however, with a simpler but still useful case by assuming:

- there is a finite-dimensional information state, i.e. $\mathcal{P}_{a_t} \equiv x_{t-1} \equiv$ a finite-dimensional vector,
- the partial loss depends on x_t and a_t only $z(\Delta(t), a(t)) \equiv z(x_t, a_t)$, i.e. the considered loss is
$$\mathcal{Z}(\Delta(\bar{t}), a(\bar{t})) = \sum_{t \in t^*} z(x_t, a_t).$$

Agreement 3.2.1 (Stabilising strategy) *Let us consider sequence of decision-making problems with the growing horizon* $\bar{t} \to \infty$. *A strategy* $\{\mathcal{R}_t : \mathcal{P}^*_{a_t} \to a^*_t\}_{t \in t^* \equiv \{1,\ldots,\bar{t}\}}$ $\forall \bar{t} \in \{1, 2, \ldots\}$ *is called* stabilising strategy *if there is a finite constant* c *such that*

$$\mathcal{E}[z(x_t, a_t)|\mathcal{P}_{a_t}, a_t] \le c < \infty, \; t \in t^*. \tag{3.7}$$

Proposition 3.2.1 (Asymptotics of design) *Let a stabilising strategy exist. Then there is an optimal strategy for $\bar{t} \to \infty$ whose decisions are minimising arguments in the formal analogy of (2.23)*

$$ {}^{[\infty]}\mathcal{V}(x_{t-1}) + C = \min_{a_t \in a_t^*} \mathcal{E}[z(x_t, a_t) + {}^{[\infty]}\mathcal{V}(x_t)|x_{t-1}, a_t]. \tag{3.8} $$

with a constant $C \le c$ and a time invariant Bellman function ${}^{[\infty]}\mathcal{V}(x)$.

Proof: (Outline) Let us take any finite horizon $\bar{t}$. We denote ${}^{[\bar{t}]}\mathcal{E}$ the expectation taken for the strategy $\{{}^{[\bar{t}]}\mathcal{R}_t\}_{t \in t^*}$ optimal for the additive loss specified on this horizon. Similarly, ${}^{[\bar{t}]}\tilde{\mathcal{V}}(\mathcal{P}_{a_t}) \equiv {}^{[\bar{t}]}\tilde{\mathcal{V}}(x_{t-1})$ denotes the optimal cost-to-go, see Agreement 2.5.3 and ${}^{[\bar{t}]}a_t = {}^{[\bar{t}]}\mathcal{R}_t(x_{t-1})$.

The definition of optimality, the additive form of the loss (2.22) and the assumption (3.7) imply that

$$ c \ge {}^{[\bar{t}]}C \equiv {}^{[\bar{t}]}\mathcal{E}[\mathcal{Z}(\Delta(\bar{t}), {}^{[\bar{t}]}a(\bar{t}))] - {}^{[\bar{t}-1]}\mathcal{E}[\mathcal{Z}(\Delta(\bar{t}-1), {}^{[\bar{t}-1]}a(\bar{t}-1))] \ge 0. $$

Consequently,

$$ \limsup_{\bar{t} \to \infty} \left\{ {}^{[\bar{t}]}\mathcal{E}[\mathcal{Z}(\Delta(\bar{t}), {}^{[\bar{t}]}a(\bar{t}))] - \sum_{t=1}^{\bar{t}} {}^{[t]}C \right\} \equiv {}^{[\infty]}\mathcal{V} < \infty. $$

If we redefine the partial loss to $z(x_t, a_t) - {}^{[t]}C$ then the optimal strategies for all $\bar{t}$ do not change and ${}^{[\bar{t}]}\mathcal{E}[\mathcal{Z}(\Delta(\bar{t}), {}^{[\bar{t}]}a(\bar{t}))] = {}^{[\infty]}\mathcal{V} + o(\bar{t})$ where the symbols $o(\bar{t})$ means, as usual, a term converging to zero for $\bar{t} \to \infty$.

Similar considerations apply to the modified optimal cost-to-go ${}^{[\bar{t}]}\mathcal{V}(x_t)$, Agreement 2.5.3, imply that (with the modified partial loss!) it converges for all $x_t \in x_t^*$ to some ${}^{[\infty]}\mathcal{V}(x_t)$.

With the original partial loss and modified cost-to-go, the dynamic programming (2.23) gets the form

$$ {}^{[\bar{t}]}\mathcal{V}(x_{t-1}) + {}^{[t]}C = \min_{a_t \in a_t^*} {}^{[\bar{t}]}\mathcal{E}[z(x_t, a_t) + {}^{[\bar{t}]}\mathcal{V}(x_t)|x_{t-1}, a_t] $$

which converges for $\bar{t} \to \infty$. □

Remarks 3.2.1

1. *The same proof is directly applicable to the fully probabilistic design as it can be seen as an instance of the additive loss function.*

2. *Solutions of the Bellman equation obtained for a growing finite horizon $\bar{t}$ can be interpreted as successive approximations for solving its stationary counterpart (3.8).*

3. *So called* iterations in strategy space *[29] is an alternative and efficient way of finding the asymptotic solution. Essentially, a stabilising stationary ($\equiv$ repeating a single rule) strategy $\{\mathcal{R}\}$ is selected and the linear equation*

$$ \mathcal{V}(x) + C = \mathcal{E}[z(\tilde{x}, \mathcal{R}(x)) + \mathcal{V}(\tilde{x})|x, \mathcal{R}(x)] $$

is solved for the function $\mathcal{V}(\cdot)$ *and constant* C. *Then, a new approximating strategy is found value-wise* $\mathcal{R}(x) = \mathrm{Arg}\min_{a \in a^*} \mathcal{E}[z(\tilde{x}, a) + \mathcal{V}(\tilde{x})|x, a]$ *with such a* $\mathcal{V}(\cdot)$. *Under general conditions, the newly found strategy is stabilising and iterations may be repeated till the guaranteed convergence. Details of this procedure are out of scope of this work but it should be considered when searching for efficient numerical procedures.*

Chapter 4

Decision making tasks

Here, we provide practical comments to elements occurring in all decision tasks. Then we list usual decision tasks met. Their specific cases are is the second part of this text. Their presentation follows more or less the structure of this chapter.

4.1 Construction elements

The discussion of "atoms" creating any decision task should help us to learn a good practice and to avoid common mistakes. Their choice, modification and use represent typically specific decision subtasks that have to be harmonised with the final aim considered. It is often very hard task but the following *design golden rule* should be respected as much as possible:

$$\text{Do not split the overall decision task into subtasks unless necessary.} \tag{4.1}$$

4.1.1 Data

Data connect the artificial world of evaluations with reality. Their information content is crucial for the success of the decision making that use them.

Experimental design

If the designer has an opportunity to influence the information content of data he/she should do that by a proper *experimental design*. At abstract level, it means that the chosen working conditions should suppress ambiguity of the best projections caused by quality of data, see Proposition 3.1.1. More practically, the optional data (inputs, set-points) have to be chosen so that they "excite" the system in a sufficient way. For instance, we cannot learn the dependence of outputs on inputs when inputs are not varying during the data acquisition. Obviously, data properties influence also speed of the learning.

Thus, if possible, it is reasonable to optimise information content of data. We outline such optimisation in the simpler case of estimation.

Mutual information, i.e. Kullback-Leibler distance (2.26) of the joint pdf of data and parameters to the product of their marginal pdfs, is known to be an adequate quantitative expression of the informational relationship inspected [14]. Under natural conditions of decision making (2.33), it takes the form

$$\mathcal{I}(\Delta(\bar{t}), a(\bar{t})||\Theta) \equiv \tag{4.2}$$
$$\int f(\Delta(\bar{t}), a(\bar{t}), \Theta) \ln\left(\frac{f(\Delta(\bar{t}), a(\bar{t}), \Theta)}{f(\Delta(\bar{t}), a(\bar{t}))f(\Theta)}\right) d(\Delta(\bar{t}), a(\bar{t}), \Theta) =$$
$$= \sum_{t\in t^*} \int f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) f(a_t|\mathcal{P}_{a_t}) f(\Theta|\mathcal{P}_{a_t}) f(\mathcal{P}_{a_t}) \ln\left(\frac{f(\Theta|\mathcal{P}_{a_t})}{f(\Theta)}\right) d(\Delta(t), a(t), \Theta).$$

This form demonstrates that the mutual information measures an average distance of the posterior and prior pdfs.

In order to learn as much as possible about unknown parameters, the mutual information should be maximised with respect to the optional admissible strategy described by its model $\{f(a_t|\mathcal{P}_{a_t})\}_{t\in t^*}$. Formally, dynamic programming, Proposition 2.5.1, could be used to this purpose.

Remarks 4.1.1

1. *The functional (4.2) depends linearly on pdfs $\{f(a_t|\mathcal{P}_{a_t})\}_{t\in t^*}$ describing the optional strategy. Thus, a practically meaningful solution can be obtained only when some restrictions are added. This property is not related to this specific distance but manifests the fact that an infinite stimulating energy, if allowable, may provide the best information about the system inspected.*

2. *The formulated maximisation task cannot be generally solved because of its complexity. It is as complex as dual control. It provides us at least the ideal solution to be approximated. For instance, it becomes a useful optimising guide when only a finite amount of competitive strategies is a priori allowed.*

   Mutual information can also be used for analysing data when results of parameter estimation are unsatisfactory.

3. *Specific feasible solutions for specific classes of models are well elaborated see, for instance, [30].*

Problem 4.1.1 (How to approximate experimental design?) *General, numerically feasible approximations of the optimal but infeasible solution of the experimental design are still waiting for their inventors. The problem is even more hard and challenging when the predictive pdfs result from filtering, see Section 2.6.1. It might be reasonable to consider the asymptotic version of this optimisation like in Section 3.2.*

Data pre-processing

The objective pdf describing reality is (practically) always out of the set considered parameterised models. The estimation searches for the best projection of all aspects of reality which are reflected in measured data, cf. Remarks 3.1.1. Some aspects may be a priori uninteresting for the problem at hands. The estimation, that always deals with a finite data set, has to handle them, too. Consequently, it might fail to provide a solid information on aspects of interest.

Thus, a wise and commonly used practice is to suppress uninteresting details in data before estimation or filtering, to pre-process data. A similar logic justifies pre-processing in control and other decision making tasks.

Pre-processing steps could be and ideally should be done using the framework described in this text. Some of them form essence of the presented examples. At the same time, there is a bunch of ready technique [31, 32] that can be *carefully* used.

Typical pre-processing steps are: data transformation linearising their relationships, data scaling, outliers removal, high frequency noise removal, filling of missing data, data aggregation etc. Each of them makes a pre-projection to a simplified world. It obviously influences quality of the resulting projection. Damages made in the pre-processing phase can hardly be removed in later design phases. Typical errors in pre-processing are:

- a premature reduction of data so that some informative attributes are lost,
- wasting of information in data by sampling them with too low frequency,
- a significant change of the modelled dynamics by underestimating the fact that the the final model describes coupling of the reality and the used pre-processing mechanism,
- a distortion of the inspected relationships by inadequate substitutions for missing data.

Problem 4.1.2 (How to harmonise pre-processing with ultimate goal?) *Similarly, as in other places, the optimal pre-processing demands to solve the decision problem in its entirety. It is mostly impossible and splitting the overall task into adequate and harmonised subtasks is left to a "sound" reasoning. It is pleasant as it requires creativity. It is unpleasant as the overall results of the decision making might be spoiled by an improper choice. The problem is severe especially in* dynamic *decision making problems in which there is a restricted freedom for an iterative trial-and-error treatment.*

4.1.2 Decisions

The whole design we deal with aims at selecting proper decisions from a set a^*. The choice of this set is discussed briefly here.

Elements

The set a^* of possible decisions contains often elements that can "substitute" each other. This incomplete pre-determination calls for selecting the most suitable elements. For instance, moisture of a produced paper can be influenced both by machine speed and energy used for drying [33]. The choice is first of all based on physical and economic priorities. Then structure estimation task, see Agreement 4.2.1 and Proposition 4.2.1, should follow when no other technical priorities can be recognised.

Restrictions

Causality restrictions, see Section 2.2, are always considered in our design. Their impact can be influenced at a pre-design stage. Typically, new sensors can be installed in order to enrich accessible experience. Their efficient choice is far from being trivial and respecting of the design golden rule (4.1) cannot be over-stressed in this instance.

Technological restrictions, i.e. final specification of the decision set $a^*(\bar{t})$, follows mostly from technological, economic or safety considerations. Generally, they increase complexity of the design of the optimal strategy and have to be modified in order to make the design and application feasible. When they are relaxed for complexity reasons, they have to be reflected in other elements of the design. Often, the loss function is modified. In a sense, a sort of penalty-based minimisation under restrictions is implicitly used.

Moreover, a good and justified optimisation practice has to be respected. For instance, when a particular unrestricted optimising decision is out of the target set a^* a simple clipping at the boundary of a^* should not be used: a proper (near) optimal projection on a^* is needed.

Also, unnecessary restrictions should be avoided. For instance, the restriction to unbiased estimators [34] is well justified in many statistical decision tasks. If it is, however, mechanically required for other tasks, say adaptive control design, it can make the final strategy (much) less efficient than possible.

Problem 4.1.3 (How to treat technical priorities?) *A systematic analysis of priorities and technical consideration mentioned in previous paragraphs and met on many places should be supported by formal tools. Those available now, like* qualitative modelling and simulation *[35], do not fit well the methodology described here.*

4.1.3 Parameterised model

Again, a non-formalised search for a compromise between the model quality and complexity of its treatment makes the model choice more art than science. The need for reaching this compromise is responsible for a prominent role of so called *exponential family* [36] and *conjugate prior*. They help us in struggle with the curse of dimensionality [37] and at the same time they are able to cover a range of practical problems..

Exponential family

Agreement 4.1.1 (Exponential family of parameterised models) *The parameterised model belongs to the (dynamic) exponential family iff it can be written in the form*

$$f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) = A(\Theta)\exp[B'(\Psi_t)C(\Theta)], \quad \textit{where} \tag{4.3}$$

$\Psi_t' \equiv [\Delta_{a_t}', \psi_t']$ *is a finite dimensional* data vector *determined by experience* $\mathcal{P}_{a_{t+1}}$ *and its values can be recursively updated according to a known formula* $(\Psi_{t-1}, a_t, \Delta_{a_t})^* \rightarrow \Psi_t^*$,

$'$ *means transposition,*

$A(\cdot)$ *is a non-negative scalar function defined on* Θ^*,

$B(\cdot)$, $C(\cdot)$ *are vector functions of the same finite fixed dimension. They are defined on respective arguments in* Ψ_t^* *and* Θ^*.

Remarks 4.1.2

1. *The* ψ *part of the data vector* Ψ *is often called* regression vector.
2. *Our definition of the exponential family contains the non-standard additional requirement on the recursive updating of the data vector* Ψ *(4.3). It is practically important for the dynamic situations we deal with.*
3. *Notice that equality is used in (4.3). The normalisation of this pdf must not spoil it. Due to it, the allowed form is rather restrictive. Let us indicate it.*

   Often, it is possible to assume that $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) = \Phi(\Psi_t, \Theta) > 0$. *Then, a first order Taylor-type expansion of logarithm of this function implies that* $\ln(\Phi(\Psi_t, \Theta)) \approx \tilde{B}'(\Psi_t)\tilde{C}(\Theta)$. *We exponentiate the right hand of this approximate equality and try to normalise it so that integral over* Δ_{a_t} *equals to one. Mostly, we find that the normalising factor* $A(\cdot)$ *will depend both on entries of* ψ_t *and* Θ. *Thus, the resulting pdf lies out of the exponential family.*

Majority of the textbook pdfs belong to the exponential family. The following table provides examples of such pdfs in the *static case* when the regression vector is empty, i.e. $\Psi \equiv \Delta_a \equiv y$ and y is either scalar or its entries are independent (recall Agreement 2.3.4).

Name	**Parameterised model**	**Innovation**	**Parameter**
Exponential	$\frac{1}{\lambda}\exp\left(-\frac{y}{\lambda}\right)$	$y \geq 0$	$\lambda > 0$
Poisson	$\frac{\mu^y}{\Gamma(y+1)}\exp(-\mu)$	$y \in \{0, 1, \ldots\}$	$\mu > 0$
Multinomial	$\prod_{i \in y^*} \Theta_i^{\delta(i,y)}$	$y \in \{1, \ldots, \bar{y}\}$	$\{\Theta_i \geq 0, \sum_{i \in y^*} \Theta_{i\|j} = 1\}$
Normal	$\frac{1}{\sqrt{2\pi r}}\exp\left[-\frac{(y-\mu)^2}{2r}\right]$	$y \in (-\infty, \infty)$	$\mu \in (-\infty, \infty)$, $r > 0$
Log-normal	$\frac{1}{y\sqrt{2\pi r}}\exp\left[-\frac{\ln^2\left(\frac{y}{\mu}\right)}{2r}\right]$	$y > 0$	$\mu > 0$, $r > 0$

The used *Euler gamma function*

$$\Gamma(x) \equiv \int_0^\infty z^{x-1}\exp(-z)\,dz \tag{4.4}$$

is finite for $x > 0$.

The introduced *Kronecker delta* is defined

$$\delta(i, y) = \begin{cases} 1 & \text{if } i = y \\ 0 & \text{if } i \neq y \end{cases}. \tag{4.5}$$

The *dynamic case* with a nonempty ψ is much more narrow (cf. Remarks above). Essentially, normal (Gaussian) linear-in-parameters model and Markov chains dominate this case.

Name	**Parameterised model**	**Data vector**	**Parameter**
Normal (Gaussian)	$\frac{1}{\sqrt{2\pi r}} \exp\left[-\frac{(y-\theta'\psi)^2}{2r}\right]$	$y \in (-\infty, \infty)$ $\psi \in (-\infty, \infty)^{\bar{\psi}}$	$\theta \in \theta^* \subset (-\infty, \infty)^{\bar{\theta}}$ $r > 0$
Markov chain	$\prod_{i \in y^*} \prod_{j \in \psi^*} \Theta_{i\mid j}^{\delta(i,y)\delta(j,\psi)}$	$y \in y^* \equiv \{1, \ldots, \bar{y}\}$ $\psi \in \psi^* \equiv \{1, \ldots, \bar{\psi}\}$	$\forall j \in \psi^*$, $\{\Theta_{i\mid j} \geq 0,$ $\sum_{i \in y^*} \Theta_i = 1\}$

Practical significance of the exponential family becomes obvious when we specialise to it Proposition 2.6.2 describing estimation and prediction.

Proposition 4.1.1 (Estimation and prediction in exponential family) *Let natural conditions of decision making, Requirement 2.6.1, be met with time-invariant $\Theta_t = \Theta \in \Theta^*$. Let the parameterised model belong to exponential family (4.3). Then, the* predictive pdf *(the outer model of the system) is given by the formula*

$$f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t) = \frac{\mathcal{I}(V_{t-1} + B(\Psi_t), \nu_{t-1} + 1)}{\mathcal{I}(V_{t-1}, \nu_{t-1})}, \ \textit{with} \tag{4.6}$$

$$V_t = V_{t-1} + B(\Psi_t),\ V_0 = 0;\quad \nu_t = \nu_{t-1} + 1,\ \nu_0 = 0, \quad \textit{and} \tag{4.7}$$

$$\mathcal{I}(V, \nu) = \int_{\Theta \in \Theta^*} A^{\nu}(\Theta) \exp[V'C(\Theta)] f(\Theta)\, d\Theta \tag{4.8}$$

where $f(\Theta)$ is a prior pdf. The Bayesian parameter estimate (posterior pdf) is

$$f(\Theta|\mathcal{P}_{a_{t+1}}) = \frac{A^{\nu_t}(\Theta) \exp[V_t'C(\Theta)] f(\Theta)}{\mathcal{I}(V_t, \nu_t)}, \tag{4.9}$$

i.e. the likelihood function is $\mathcal{L}(\Theta, \mathcal{P}_{a_{t+1}}) \equiv \mathcal{L}(\Theta, V_t, \nu_t) = A^{\nu_t}(\Theta) \exp[V_t'C(\Theta)]$.

Remarks 4.1.3

1. *Estimation and prediction within exponential family is extremely simple. The problem of updating of functions (pdfs) converts into the algebraic recursive updating of the finite-dimensional sufficient statistics V_t and of the* sample counter *ν_t. The wish to have this recursion complete explains the requirement for possibility to update Ψ_t recursively, see Agreement 4.1.1. Moreover, a single type of the normalisation integral $\mathcal{I}(V, \nu)$ has to be evaluated.*

2. *An inspection whether there is a wider set of parameterised models with advantageous properties of the exponential family opens just narrow space [38, 39]. Essentially, the exponential family coincides with all parameterised models that are sufficiently smooth functions of Θ and with supports independent of Θ. Uniform distribution with unknown constant boundaries represents one of a few feasible examples of pdfs out of exponential family, see Chapter dealing with uniform noise.*

3. *The class of models that lead to a finite-dimensional characterisation of pdfs occurring in filtering is even more restrictive. Its discussion can be found in [22].*

Completion of probabilistic models

The presented design relies (with a few exceptions) on a complete specification of the parameterised model. At the same time, the designer has always at disposal a finite amount of facts for its construction. Their necessary *completion to pdf* (an infinite-dimensional object) represents a non-standard and weakly supported interpolation task. We sketch its solution proposed in [40].

Let the designer know the form of the conditional expectation of a given vector function $H(\Delta_{a_t})$, i.e.

$$\mathcal{E}[G(\Delta_{a_t}, \mathcal{P}_{a_t}, a_t)|\Theta, \mathcal{P}_{a_t}, a_t] = g(\Theta, \mathcal{P}_{a_t}, a_t) \tag{4.10}$$

with a given vector function g. The constructed pdf $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t)$ should respect just this knowledge. It has to fulfill (4.10). It should be as uncertain as possible in order to leave as much freedom as possible for learning from data. We quantify these requirements by minimising Kullback-Leibler distance (2.26) of the constructed $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t)$ to the uniform pdf under the restriction (4.10). Note that the minimisation is equivalent to entropy maximisation [14].

Proposition 4.1.2 (Completion of parameterised models) *The minimiser $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t)$ of the functional*

$$\mathcal{D}(f||\text{uniform pdf}) = \int f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) \ln[f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t)]\, d\Delta_{a_t}$$

within the set of pdfs

$$\left\{ f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) : \int G(\Delta_{a_t}, \mathcal{P}_{a_t}, a_t) f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) dDt = g(\Theta, \mathcal{P}_{a_t}, a_t) \right\}$$

has the form

$$f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) = \gamma(\Theta, \mathcal{P}_{a_t}, a_t) \exp[\lambda'(\Theta, \mathcal{P}_{a_t}, a_t) G(\Delta_{a_t}, \mathcal{P}_{a_t}, a_t)]. \tag{4.11}$$

The scalar function $\gamma(\cdot)$ is determined by the normalisation, see Proposition 2.3.4. The vector function $\lambda(\cdot)$ has to fulfill the identity

$$\int G(\Delta_{a_t}, \mathcal{P}_{a_t}, a_t) \gamma(\Theta, \mathcal{P}_{a_t}, a_t) \exp[\lambda'(\Theta, \mathcal{P}_{a_t}, a_t) G(\Delta_{a_t})]\, d\Delta_{a_t} = g(\Theta, \mathcal{P}_{a_t}, a_t). \tag{4.12}$$

Proof: It is implied by an elementary calculus of variations. □

Remarks 4.1.4

1. *The result is more of a conceptual than of a practical interest as the solution of (4.12) can rarely be found. Moreover, the model found is mostly out of the favourite exponential family.*
2. *Gaussian parameterised model is obtained for $G(\Delta_{a_t}, \mathcal{P}_{a_t}, a_t) = [\Delta_{a_t}, \Delta_{a_t}^2]'$. This helps us in judging suitability of this popular option.*

Problem 4.1.4 (How to make the completion practicable?) *The weakness of the presented results indicates how neglected the art of the model completion is. The completion has to be either developed much more or avoided completely. The latter desirable direction would require to built the corresponding decision theory relying only on partial descriptions like (4.10). It is possible that it is impossible.*

Mixtures of models

The model constructed in previous Section exploits available physical knowledge expressed by the expectation (4.10). It is often obtained through modelling valid locally around some working point. Then, an extension to models covering a wider range of working conditions is desirable. *Mixture of models*

$$f(\Delta_{a_t}|\Theta,\mathcal{P}_{a_t},a_t) = \sum_{c\in c^*} \alpha_c f_c(\Delta_{a_t}|\Theta_c,\mathcal{P}_{a_t},a_t), \quad \Theta \in \Theta^* \equiv \tag{4.13}$$
$$\left\{\Theta = (\alpha_c,\Theta_c),\ \Theta_c \in \Theta_c^*\ c \in c^* \equiv \{1,\ldots,\bar{c}<\infty\},\ \alpha \in \alpha^* \equiv \left\{\alpha_c \geq 0,\ \sum_{c\in c^*} \alpha_c = 1\right\}\right\}$$

represents a good universal option for meeting this need as the mixtures of models have often so called *universal approximation property.* It means that the set (4.13) is dense in the set of outer models of the system. This property takes place whenever the number of *components* $\bar{c}$ is not limited and individual components, i.e. pdfs $f_c(\Delta_{a_t}|\Theta_c,\mathcal{P}_{a_t},a_t)$, can be made arbitrarily close to an arbitrarily positioned Dirac delta function [41].

The *exact estimation and prediction* with this model is practically *impossible* as the number of terms in likelihood function $\mathcal{L}(\Theta,\mathcal{P}_{a_{t+1}})$ – the product of parameterised models – increases exponentially. Luckily enough, good algorithms for approximate estimation of the mixture of models are available [42]. Even a good approximate *recursive* estimation is possible if the *components* $\{f_c(\Delta_{a_t}|\Theta_c,\mathcal{P}_{a_t},a_t)\}_{c\in c^*}$ belong to the exponential family or describe uniform distributions on rectangular boxes. This quasi-Bayes estimation is described in Section 5.2.3.

The quasi-Bayes estimation as well as other algorithms exploit the following alternative interpretation of the mixture of models (4.13). Let $c_t \in c^*$ be the unobservable random selector of the "active" component that takes the value c with probability α_c. Then,

$$f(\Delta_{a_t},c_t|\Theta,\mathcal{P}_{a_t},a_t) = f_{c_t}(\Delta_{a_t}|\Theta_{c_t},\mathcal{P}_{a_t},a_t)\alpha_{c_t}. \tag{4.14}$$

The marginalisation rule, see Proposition 2.3.4, implies that the mixture (4.13) is marginal pdf of the joint pdf (4.14).

4.1.4 Prior pdf

The need for selecting prior pdf is often regarded as the main disadvantage of the adopted Bayesian approach []. We contribute positively to the never-ending discussion on its pros and cons by showing that its use opens a regular gate through which prior "expert" information can be introduced into learning in a systematic way.

The posterior pdf (2.41) is a product of the likelihood function consisting of t factors $f(\Delta_{a_\tau}|\Theta,\mathcal{P}_{a_\tau},a_\tau)$ and of a single prior pdf $f(\Theta)$. If t is high enough and data bring a sufficient information on Θ then the posterior pdfs obtained for various prior pdfs (with a common support) resemble each other: the role of prior pdf is weak.

The posterior pdf is significantly influenced by the prior pdf when some of the above conditions is not fulfilled. Let us formulate simple propositions that formalise this.

Proposition 4.1.3 (Role of the prior pdf)

1. *Parameter values* $\Theta \notin \mathrm{supp}\,[\,f(\Theta)]$ *for which the prior pdf is zero get the zero posterior pdf, too. Formally,*

$$\mathrm{supp}\left[\,f(\Theta|\mathcal{P}_{a_{t+1}})\right] = \mathrm{supp}\left[\,\mathcal{L}(\Theta,\mathcal{P}_{a_{t+1}})\right] \cap \mathrm{supp}\,[\,f(\Theta)]\,.$$

2. *The recursive evolution of the likelihood function*

$$\begin{aligned}\mathcal{L}(\Theta,\mathcal{P}_{a_{t+1}}) \quad &\equiv \quad \prod_{\tau\le t} f(\Delta_{a_\tau}|\Theta,\mathcal{P}_{a_\tau},a_\tau) = f(\Delta_{a_t}|\Theta,\mathcal{P}_{a_t},a_t)\mathcal{L}(\Theta,\mathcal{P}_{a_t})\\ & \quad\quad t\in t^*,\ \mathcal{L}(\Theta,\mathcal{P}_{a_1}) = 1, \Theta\in\Theta^*_{\mathcal{L}}\end{aligned}$$

   does not depend on the prior pdf chosen.

3. *The posterior pdf exists iff the product* $\mathcal{L}(\Theta,\mathcal{P}_{a_{t+1}})f(\Theta)$ *is integrable.*

Proof: It is a direct consequence of the formula for posterior pdf (2.41). □

Remarks 4.1.5

1. *The prior pdf offers a simple and clear way for introducing hard restrictions on parameters.*

2. *The recursion (4.15) is valid even if* $\Theta^*_{\mathcal{L}} \neq \Theta^* \equiv \mathrm{supp}\,[\,f(\Theta)]$. *This trivial statement implies that hard bounds on parameter values must not influence likelihood function, i.e. the "objective" link to data reflecting reality. This fact is easily and repeatedly overlooked in recursive estimation. Instead of restricting the posterior pdf the statistics determining the likelihood are deformed with an adverse effect on the estimation quality.*

3. *Often, a flat prior pdf is chosen in order to model the lack of prior knowledge. Even integrability of the prior pdf is relaxed and the* improper prior pdfs $f(\cdot)\ge 0,\ \int f(\Theta)\,d\Theta = \infty$ *are used. For instance, the posterior pdf is proportional to the likelihood function if we allow the prior pdf be improper and equal 1. Then, the posterior pdf might be improper, too. Thus, it is more reasonable to select a flat but proper prior pdf as it serves as a regularising factor.*

Agreement 4.1.2 (Conjugate prior pdf) *Let the parameterised model belong to exponential family, see Agreement 4.1.1,*

$$f(\Delta_{a_t}|\Theta,\mathcal{P}_{a_t},a_t) = A(\Theta)\exp[B'(\Psi_t)C(\Theta)].$$

The prior pdf is called conjugate prior pdf *if it has the form*

$$f(\Theta) \propto A^{\nu_0}(\Theta)\exp[V_0'C(\Theta)]\chi(\Theta). \tag{4.15}$$

It is determined by the indicator χ *of the set* Θ^* *and "prior statistics"* V_0, ν_0.

Proposition 4.1.4 (Estimation with conjugate prior pdf) *Let the parameterised system belong to the exponential family and the conjugate prior pdf is chosen. Then, the posterior pdf has the same functional form as the prior one.*

Its statistics V_t, ν_t evolve according to (4.7) with zero initial conditions replaced by V_0, ν_0:

$$\begin{aligned} f(\Theta|\mathcal{P}_{a_t}) &\propto A^{\nu_t}(\Theta)\exp[V_t' C(\Theta)]\chi(\Theta),\ t \in t^* \\ V_t &= V_{t-1} + B(\Psi_t),\ \nu_t = \nu_{t-1} + 1 \\ V_0, \nu_0 &\equiv \textit{prior statistics.} \end{aligned} \tag{4.16}$$

Proof: It is a direct consequence of the formula for posterior pdf (2.41) applied to the exponential family. □

Agreement 4.1.3 (Fictitious data) *Let the parameterised model belong to the exponential family and a conjugate prior pdf $f(\Theta) \equiv f(\Theta|K_{t_f})$ respects a piece of knowledge K_{t_f}. The* fictitious data vector *Ψ_{t_f} is said to express this knowledge piece if*

$$f(\Theta|K_{t_f}) \propto A^{\bar{\nu}+1}(\Theta)\exp\left\{[B(\Psi_{t_f}) + \bar{V}]'C(\Theta)\right\} \propto A(\Theta)\exp[B'(\Psi_{t_f})C(\Theta)]\bar{f}(\Theta). \tag{4.17}$$

The used pdf $\bar{f}(\Theta) \propto A^{\bar{\nu}}(\Theta)\exp[\bar{V}'C(\Theta)]$ is a flat pre-prior pdf.

In other words, the inclusion of the individual knowledge piece K_{t_f} is expressed as an application of the Bayes rule with fictitious data vector inserted into the parameterised model.

Remarks 4.1.6

1. *Instructive examples of fictitious data are in in [43]. The construction is also demonstrated in examples of this text. The common procedure is as follows:*

   Algorithm 4.1.1 (Construction of fictitious data)

   (a) *Ask yourselves what response of the system you expect to some experience (often, past inputs or initial values) when your information piece is correct.*
   (b) *Think about uncertainty of the system response caused by your uncertainty about the treated knowledge piece.*
   (c) *Scale these data so that the response uncertainty resembles the uncertainty caused by random noises within the system.*

2. *It is possible and sometimes reasonable to group fictitious data into mutually consistent blocks that are treated as a block of real data vectors, i.e. modify the flat prior pdf by applying the Bayes rule.*

A collection of knowledge pieces $\{K_{t_f}\}_{t_f \in t_f^*}$ contains generally items of heterogeneous and uncertain nature. The auxiliary notion of fictitious data provides us their unified expression and give us a possibility to reflect their uncertainty. With fictitious data vectors, the knowledge items $K(\bar{t}_f)$, $\bar{t}_f < \infty$ can be incorporated into the prior pdf individually. Without knowing their mutual dependencies, like their complementarity, degree of compatibility, repetitions etc., we cannot construct the desirable prior pdf $f(\Theta|K(\bar{t_f}))$ reflecting the whole collection of available knowledge pieces. Thus, some heuristics is needed. The following simple proposition prepares one of possible and successful formalisations.

Proposition 4.1.5 (Geometric representant of a pair of pdfs) *Let an unknown pdf $f \in f^* \equiv \{f_1, f_2\}$ be equal to $f_1(x)$ with a probability $\alpha \in \alpha^* \equiv [0,1]$ and equal to $f_2(x)$ with the complementary probability $1-\alpha$. The pdfs f_1, f_2 are supposed to have a common support x^*. Then, the pdf*

$$^{[o]}\hat{f}(x) \propto f_1^{\alpha}(x) f_2^{1-\alpha}(x), \; x \in x^* \tag{4.18}$$

minimises the expected Kullback-Leibler distance (2.26) of $\hat{f}(\cdot)$ to $f(\cdot)$, i.e. the functional

$$\mathcal{E}[\mathcal{D}(\hat{f}||f)] \equiv \alpha \mathcal{D}(\hat{f}||f_1) + (1-\alpha)\mathcal{D}(\hat{f}||f_2). \tag{4.19}$$

The minimum reached is

$$\omega(\alpha) \equiv \mathcal{E}[\mathcal{D}(\,^{[o]}\hat{f}||f)] = -\ln \int_{x^*} f_1^{\alpha}(x) f_2^{1-\alpha}(x)\, dx. \tag{4.20}$$

If $f_1 \not\equiv f_2$, the function $\omega(\alpha)$ reaches its maximum on $(0,1)$.

Proof: (4.18) is implied by an elementary calculus of variations. (4.20) follows from it and (4.19). The last statement is implied by the assumption $f_1 \not\equiv f_2$, closeness of α^* and continuity of $\omega(\alpha)$ guaranteed by the common support of f_1, f_2. □

The combination or merging of a pair of knowledge pieces may be directly based on this proposition by taking $f_{t_f} \equiv f(\Theta|K_{t_f})$, $t_f = 1,2$ and accepting $^{[o]}\hat{f}$ as $f(\Theta|K(\bar{t_f}))$. The constructed prior pdf used should not claim misleading confidence thus it makes sense to choose the "worst" (conservative) α in (4.18), i.e. α that maximises the minimum reached.

It is straightforward to extend Proposition 4.1.5 for several pooled pdfs. However, the search for the worst probabilistic vector α becomes computationally intensive even for a medium $\bar{t_f}$.

There is an obvious approximation based on a sequential application of Proposition 4.1.5 with $f_1 \equiv f(\Theta|K(\bar{t}_f - 1))$, $f_2 \equiv f(\Theta|K_{\bar{t}_f})$ and $^{[o]}\hat{f} \equiv f(\Theta|K(\bar{t}_f))$. The result may, however, strongly depend on the evaluation order. For this reason, an alternative formulation has been proposed in [43] that allows us not only to add a knowledge piece but also to remove superfluous ones.

Let us describe this procedure assuming that $f(\Theta|K(t_f - 1))$ has been gained in some way and that there are two possibilities to built-in a new piece of knowledge K_{t_f}:

Optimistic variant: Bayes rule includes correctly the new knowledge piece

$$f_1(\Theta|K(t_f)) \propto f(\Delta_{a_{t_f}}|\Theta, \mathcal{P}_{a_{t_f}}, a_{t_f}) f(\Theta|K(t_f - 1)).$$

This variant is assigned the probability $\alpha \in [0,1]$

Cautious variant: Bayes rule can only be applied after removing the new factor $f(\Delta_{a_{t_f}}|\Theta, \mathcal{P}_{a_{t_f}}, a_{t_f})$ β-times

$$f_2(\Theta|K(t_f)) \propto [f(\Delta_{a_{t_f}}|\Theta, \mathcal{P}_{a_{t_f}}, a_{t_f})]^{1-\beta} f(\Theta|K(t_f - 1)), \; \beta \geq 0.$$

This variant is taken with the probability $1-\alpha$.

Proposition 4.1.6 (Recursive processing of prior knowledge) *Let $f(\Theta|K(t_f))$ have only the optimistic or cautious variant. Then, the minimiser $\hat{f}(\Theta|K(t_f))$ of the expected value of the Kullback-Leibler distance (2.26) to $f(\Theta|K(t_f))$ has the form*

$$\hat{f}(\Theta|K(t_f)) \propto [f(\Delta_{t_f}|\Theta, \mathcal{P}_{a_{t_f}}, a_{t_f})]^w f(\Theta|K(t_f-1)), \text{ with } w = \alpha+(1-\alpha)(1-\beta).$$

The unique "worst" values of $\alpha \in [0,1]$ maximising the reached minimum and

$$\beta = \left\{\inf \tilde{\beta} : \int [f(\Delta_{a_{t_f}}|\Theta, \mathcal{P}_{a_{t_f}}, a_{t_f})]^{-\tilde{\beta}} f(\Theta|K(t_f-1))\, d\Theta = \infty\right\}$$

exist.

Proof: Essentially, Proposition 4.1.5 is extended. For details see [43]. □

Remarks 4.1.7

1. *The upper bound on the set of possible β values is chosen so that out of it the removal of the factor $f(\Delta_{t_f}|\Theta, \mathcal{P}_{a_{t_f}}, a_{t_f})]^{-\beta}$ shifts the product with $f(\Theta|K(t_f-1))$ out of normalisable pdfs: nothing more can be removed.*
2. *The combination of pdfs with the worst α is tightly related to so called* barycenter *of a pdf group, see [44] and its approximation called pooling [45].*

In batch processing, when we can combine prior knowledge with the measured data at one shot, it is possible to estimate weights α by using the standard Bayes rule. This straightforward procedure provides directly estimate of the posterior pdf with the available prior knowledge included. We present the result for exponential family.

Agreement 4.1.4 (Batch processing of prior knowledge) *Let the parameterised model belong to exponential family (4.3) and prior knowledge pieces $K(\bar{t}_f)$ are expressed through fictitious data, see Agreement 4.1.3. We search for such an approximation $\hat{f}(\Theta|\mathcal{P}_{a_t}, K(\bar{t}_f))$ of the unknown posterior pdf $f(\Theta|\mathcal{P}_{a_t}, K(\bar{t}_f))$ that*

- *has experience $\mathcal{P}_{a_t}, K(\bar{t}_f)$,*
- *minimizes the expected Kullback-Leibler distance to projections of $f(\Theta|\mathcal{P}_{a_t}, K(\bar{t}_f))$ within the set*
$$\{f(\Theta|\mathcal{P}_{a_t}, K_{t_f})\}_{t_f \in t_f^*} \tag{4.21}$$
- *assigns to the elements of (4.21) the posterior probabilities $\alpha(K_{t_f}|\mathcal{P}_{a_t})$ corresponding to uniform prior probabilities on $K(\bar{t_f})$.*

A straightforward use of Proposition 4.1.5, definition of fictitious data, Agreement 4.1.3, estimation in exponential family, Proposition 4.1.4, and Bayes rule give:

Algorithm 4.1.2 (Batch processing of prior knowledge)

1. *Compute the statistics V, ν determining the posterior pdf corresponding to the pre-prior pdf only*

$$\bar{V}_t = \sum_{\tau \le t} B(\Psi_\tau) + \bar{V}, \quad \bar{\nu}_t = \bar{\nu} + t \tag{4.22}$$

where the data vectors Ψ_τ are constructed from the (real) data available.

2. *Evaluate the posterior probabilities*

$$\alpha(K_{t_f}|\mathcal{P}_{a_t}) \propto \frac{\mathcal{I}(\bar{V}_t + B(\Psi_{t_f}), \bar{\nu}_t + 1)}{\mathcal{I}(\bar{V} + B(\Psi_{t_f}), \bar{\nu} + 1)}, \; t_f \in t_f^* \tag{4.23}$$

where the data vectors Ψ_{t_f} are fictitious.

3. *Determine statistics of the final pdf $\hat{f}(\Theta|\mathcal{P}_{a_t}, K(\bar{t}_f)) \propto A^{\nu_t}(\Theta) \exp[V_t' C(\Theta)]$*

$$V_t = \bar{V}_t + \sum_{t_f \in t_f^*} \alpha(K_{t_f}|\mathcal{P}_{a_t}) B(\Psi_{t_f}), \quad \nu_t = \bar{\nu}_t. \tag{4.24}$$

Remarks 4.1.8

1. *The results of the batch version do not depend on the order of prior knowledge processing. It is obviously desirable property that favorise use of this version whenever possible.*

Problem 4.1.5 (Is the knowledge quantification good enough?) *There is a significant progress in this area but the results are not stabilised yet. For instance, there is little experience with a direct use of Proposition 4.1.6 and Algorithm 4.1.2. It is worth of being inspected more closely.*

Problem 4.1.6 (How to choose version of Kullback-Leibler distance?) *Kullback-Leibler distance is not symmetric in its arguments. Optimisation of the alternative version leads to the arithmetic mean of pdfs instead of the geometric one. There are indications [46] that the chosen version should be preferred but a definite answer does not exist.*

4.1.5 Loss function

The *loss function should express our wishes as precisely as possible.* Its choice is strongly influenced by the need to get feasible design. It restrict substantially the spectrum of options.

Popular options

A dominating use of a few loss functions reflects an indirect fight with the design complexity. Undoubtedly, the *quadratic loss function* is a prominent example of a compromise between wish to quantify aims of the design and the need to get feasible solution. Its general form, $||\mathcal{Q} - {}^{[I]}\mathcal{Q}||^2$, where $||\cdot||$ is a quadratic *weighted* norm and ${}^{[I]}$ denotes a given ideal (reference, set-point), has numerous variants. They differ in the target ideal, in the variables involved and weighting used. The quadratic loss:

+ is numerically feasible in conjuction with linear models,

+ is flexible as it allows us to express priorities by a proper choice of weights,

+ approximates majority of smooth loss functions through a second order Taylor expansion.

It is fair to say that:

− the quadratic loss over-stresses significance of large deviations from the ideal,

− the choice of weights and ideal reflecting true aims is not easy,

− usefulness of the quadratic loss as a "suppressor" of the design complexity decreases sharply whenever the underlying model is not linear or hard restrictions are involved.

When hard bounds on admissible decisions have to be considered, loss functions based on l-one norm ($\sum_i |\mathcal{Q}_i - {}^{[I]}\mathcal{Q}_i|$, $\mathcal{Q}_i \equiv$scalar entry of $\mathcal{Q}$) might be much better. Also, the design based on l-infinity norm ($\max_i |\mathcal{Q}_i - {}^{[I]}\mathcal{Q}_i|$) is nowadays widely inspected [47]. It enhances the attention paid to requirements on *robustness* of the designed strategy. Note that robustness is an engineering name to a smooth dependence of the strategy on various elements (initial conditions, model quality, etc.) on which it depends.

Ideal in fully probabilistic design

The choice of the loss function should be harmonised with the model adopted. For instance, if large random deviations are expected then the use of a quadratic loss might be doubtful.

The need for the harmonisation is an additional argument for our preference to the fully probabilistic design, see Section 2.5.2. In this case, the loss is determined by the ideal pdf ${}^{[I]}f(\Delta(\bar{t}), a(\bar{t}))$ we would like to approximate. Generally, it is wise to derive the ideal pdf from the outer description of closed loop $f(\Delta(\bar{t}), a(\bar{t}))$ by replacing those characteristics we hope to influence by their desirable values. In this way, we get, for instance, an equivalent of a well tailored quadratic performance index for linear systems with Gaussian distribution [20]. Then, the choice of adequate weights in the loss is much simplified. We get also a realistic loss if the well modelled system noise is non-normal.

Problem 4.1.7 (How to express multiple objectives?) *There are extensive theories how to handle multi-valued loss functions, what properties utility functions should have [48] etc. In spite of this, operational guidelines how to construct the loss function in specific cases are far from being cook-book.*

Common mistakes

The loss function is claimed to quantify aims we want to reach. This obvious demand is often violated. Then, the design provides non-acceptable decision strategies. Here we warn against the common mistakes made in this respect.

A wider than the practically admissible set of decisions a^* is often used in order to decrease complexity of the design. Then, the loss function has to penalise values out of the desirable set. Otherwise, the optimal strategy found may generate useless decisions.

Use of loss functions that do not respect decision making as a dynamic process is another common error. Generally, the decision a_t influences whole future behaviour of the coupling system with the adopted strategy. This behaviour becomes often far from the desired one if this fact is not respected and sequence of decisions is chosen by optimising a sequence of unrelated rules, see Section 5.4.1.

Instability of the closed loop with one-step-ahead controllers applied to non-minimum phase systems is the most known demonstration of this property. However, even stable behaviour is achieved with energy wasted. This is a less known manifestation of the same "mistake" [9].

4.2 Static tasks

The *static decision-making* deals with the design and use of a single decision rule. Here, we formulate and comment variants of the static design. Examples are postponed to the second part of this work.

Specific instances of the static design differ in:

- Decomposition of the behaviour $\mathcal{Q} = (\mathcal{P}_a, a, \mathcal{F}_a) \equiv$ (experience, decision, ignorance).
- Admissible rules $\mathcal{R} \in \mathcal{R}^*$ determined by the domain $\mathcal{P}_a^*$ and range a^*.
- Loss function $\mathcal{Z} : \mathcal{Q}^* \equiv (\mathcal{P}_a, a, \mathcal{F}_a)^* \rightarrow [0, \infty]$.

These elements determine the outer model $f(\mathcal{F}_a|\mathcal{P}_a, a)$ needed for a direct application of the basic decision-making Proposition 2.4.1.

4.2.1 Point estimation

Point estimation, i.e. a selection of the estimate $\hat{\Theta}$ of an unknown parameter Θ, can be cast in the considered framework as follows.

- Decomposition of the behaviour $\mathcal{Q} \equiv (\mathcal{P}_a, a, \mathcal{F}_a) \equiv (D, \hat{\Theta}, \Theta) \equiv$ (data at disposal, point estimate, unknown parameter).
- Admissible rules are of the form $\mathcal{R} : D^* \rightarrow \hat{\Theta}^*$, $\Theta^* \subset \hat{\Theta}^*$
- Loss function $\mathcal{Z}$ measures a distance of $\hat{\Theta}$ and Θ. This distance may depend on data D, too.

The ignorance $\mathcal{F}_{\hat{\Theta}}$ coincides with the unknown parameter Θ. Thus, the outer model of the system needed for decision making is $f(\Theta|D, \hat{\Theta})$. The adopted natural conditions of decision making (2.33) imply that the decision $\hat{\Theta}$ is superfluous in the conditioning. Thus, the needed model coincides with the (generalised) Bayesian estimate $f(\Theta|D)$ (posterior pdf) determined in Proposition 2.6.2. For the given D, the optimal point estimate is, Proposition 2.4.1,

$$\hat{\Theta}(D) \in \operatorname{Arg} \min_{\hat{\Theta} \in \hat{\Theta}^*} \int \mathcal{Z}(D, \hat{\Theta}, \Theta) f(\Theta|D) \, d\Theta.$$

Remarks 4.2.1

1. *Note that the needed outer model depends on the specific structuring of realisations $\mathcal{Q}$ not on the loss function chosen.*

   The specific result of the optimisation depends of course on the loss function chosen.

2. *The quadratic loss function* $\mathcal{Z}(D, \hat{\Theta}, \Theta) = ||\Theta - \hat{\Theta}||^2_{Q(D)} \equiv ||\Theta||^2_{Q(D)} + ||\hat{\Theta}||^2_{Q(D)} + 2(\Theta, \hat{\Theta})_{Q(D)}$ *is the most popular option. It is determined by a quadratic norm* $||\cdot||_{Q(D)}$ *with a positive definite, data-dependent, kernel* $Q(D)$ *and the related scalar product* $(\cdot,\cdot)_{Q(D)}$*. For it, the minimised quantity, see Proposition 2.4.1, is quadratic form in* $\hat{\Theta}$*, namely,*

$$\mathcal{E}[\mathcal{Z}|D, \hat{\Theta}] = \mathcal{E}[||\Theta||^2_{Q(D)}|D] + ||\hat{\Theta}||^2_{Q(D)} + 2(\mathcal{E}[\Theta|D], \hat{\Theta})_{Q(D)}.$$

   It is minimised by $\hat{\Theta} \equiv \mathcal{E}[\Theta|D]$ *irrespectively of the kernel* $Q(\cdot)$.

   This general result looks very promising as just the first moment (expectation) of the posterior pdf is needed. However, its evaluation requires mostly a complete knowledge of $f(\Theta|D)$.

Problem 4.2.1 (How to select the ideal pdf for estimation?) *It would be desirable to formulate point estimation in the fully probabilistic setting. There is, howeveer, little experience how to specify the ideal of the fully probabilistic design 2.5.2 in this case. The general guideline, see Section 4.1.5 should obviously be followed. A systematic inspection is needed.*

4.2.2 Set estimation

Set estimation, i.e. construction of a set $\theta^* \subset \Theta^*$ that contains with a high probability an unknown paramater Θ, is tightly related to the point estimation. It can be cast in the considered framework as follows.

- Decomposition of the behaviour $\mathcal{Q} \equiv (\mathcal{P}_a, a, \mathcal{F}_a) \equiv (D, \theta^*, \Theta) \equiv$ (data, set estimate, unknown parameter).
- Admissible rules are of the form

$$\mathcal{R} : D^* \to (\theta^*)^* \subset \left\{\theta^* \subset \Theta^*, \int_{\theta^*} f(\Theta|D)\, d\Theta \geq \alpha \in (0,1)\right\}$$

  where α is a pre-specified probability called *credibility level*.
- Loss function $\mathcal{Z}$ measures volume of the θ^*. The volume specification may depend on data D, too.

Structuring of realisations $\mathcal{Q}$ is similar to the point estimation. This, see Remark 1 in the previous section, implies that the outer system model needed is $f(\Theta|D)$. For the given D, the optimal interval estimate is (Proposition 2.4.1)

$$\theta^*(D) \in \text{Arg} \min_{(\theta^*) \in (\theta^*)^*} \int \mathcal{Z}(D, \theta^*, \Theta) f(\Theta|D)\, d\Theta.$$

Remarks 4.2.2

1. *Essentially, α-fractiles of $f(\Theta|D)$ are searched for.*
2. *Mostly, intervals are taken as the sets considered. Their length is then taken as a measure of their volume. Other choices might be more appropriate when a multivariate Θ is considered. For instance, hyper-ellipsoids fit well to normally distributed parameters.*
3. *The sets searched for are called* credibility sets *and they are quite close to so called* confidence intervals *studied in classical statistics.*

Problem 4.2.2 (Is there a cook-book giving the ideal pdf?) *Again, the experience how to specify the ideal of the fully probabilistic design 2.5.2 is missing. The general guideline, see Section 4.1.5 should be followed but a systematic inspection is needed.*

4.2.3 Testing of hypothesis

Hypothesis testing, a selection of the best variant among several alternative hypothesis H_h, can be put within our framework as follows.

- Decomposition of the behaviour $\mathcal{Q} \equiv (\mathcal{P}_a, a, \mathcal{F}_a) \equiv (D, \hat{h}, \{H_h\}_{h\in h^*}) \equiv$ (data at disposal, estimate $\hat{h} \in h^* \equiv \{1, \ldots, \bar{h} < \infty\}$ of a pointer h to H_h, $\{$hypothesis list$\}$).
- Admissible rules are of the form $\mathcal{R} : D^* \rightarrow h^*$.
- Loss function $\mathcal{Z}$ is a $(\bar{h}, \bar{h})$-table with non-negative entries $\mathcal{Z}(D, \hat{h}, h)$ penalising discrepancy between the pointer h and its estimate $\hat{h}$. Usually, the table has the zero diagonal as no penalty is paid when $\hat{h} = h$.

The outer system model (2.20) needed is $f(h|D, \hat{h})$. The adopted natural conditions of decision making (2.33) imply that the decision $\hat{h}$ is superfluous in conditioning. Thus, we need $f(h|D)$. This probability function (pf, $\bar{h} < \infty$) is obtained according to Proposition 2.6.2 as the generalised Bayesian estimate

$$f(h|D) \propto f(D|h)f(h), \;\; h \in h^* \tag{4.25}$$

where $f(h)$ is the prior pf of h's hypothesis. The pdf $f(D|h)$ relates observed data to individual hypotheses $h \in h^*$. For a given D, the optimal decision in hypothesis testing is, Proposition 2.4.1,

$$\hat{h}(D) \in \mathrm{Arg} \min_{\hat{h}\in\hat{h}^*} \sum_{h\in h^*} \mathcal{Z}(D, \hat{h}, h) f(h|D)\,. \tag{4.26}$$

Remarks 4.2.3

1. *Unlike in classical hypothesis testing [34], the testing is performed within a completely specified set of alternatives.*
2. *If a pair of hypothesis is compared, $\bar{h} = 2$, the decision rule constructed according to the described methodology coincides with a celebrated Neymann-Pearson lemma [1].*

   In this case, off diagonal elements of $\mathcal{Z}(D, \hat{h}, h)$ penalise the classical errors of the 1st and 2nd kind [34].

3. *The needed pdfs are rarely obtained directly. Instead, they are predictive pdfs obtained through filtering or parameter estimation and prediction, see Remarks 2.6.2. If other decisions like inputs influence the observed data then factors describing their generators cancels in (4.25). It follows from natural conditions of decision-making (2.33) we accept.*

4. *In spite of its formal simplicity, the testing of hypothesis is extremely powerful technique. It is especially true when dealing with predictive pdfs needed when so called compound hypothesis are treated, [34]. It, for instance, brought a whole set of novel and efficient solutions of so called* structure estimation problem *[49, 50, 51] (see below).*

5. *When dealing with predictive pdfs, a special care has to be devoted to the choice of the prior pdf of unknown parameters within the respective hypothesis. Their practical influence may be unexpectedly high.*

The importance of the structure estimation task calls for its formal specification:

Agreement 4.2.1 (Structure estimation) *Let us consider $\bar{h}$ groups of parameterised models $\{f(\Delta_{a_t}|\Theta_h, \mathcal{P}_{a_t}, a_t, h),\ \Theta_h \in \Theta_h^*\}_{h\in h^*;t\in t^*}$ as candidates for describing a single system. Let the hypothesis H_h state that the best projection of the "objective" pdf belongs to $\{f(\Delta_{a_t}|\Theta_h, \mathcal{P}_{a_t}, a_t, h),\ \Theta_h \in \Theta_h^*\}_{t\in t^*}$. Then, the corresponding hypothesis testing is called* structure estimation. *The generator of the additional decisions a_t (like system inputs) is supposed to fulfill natural conditions of decision making (2.33).*

Proposition 4.2.1 (Solution of structure estimation) *Let data $D \equiv (\mathcal{P}_{a_{t+1}}, a_{t+1})$ be available for structure estimation. Then, the outer model needed for selecting the best structure estimate according to (4.26) is*

$$\begin{aligned} f(h|\mathcal{P}_{a_{t+1}}) &\propto f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t, h) f(h|\mathcal{P}_{a_t}) \equiv \\ &\equiv \int f(\Delta_{a_t}|\Theta_h, \mathcal{P}_{a_t}, a_t, h) f(\Theta_h|\mathcal{P}_{a_t})\, d\Theta_h f(h|\mathcal{P}_{a_t}). \end{aligned} \tag{4.27}$$

The pf $f(h|\mathcal{P}_{a_t}, a_t)$ can be obtained recursively according to the formula (4.27) starting from a given prior pf $f(h) \equiv f(h|\mathcal{P}_{a_1})$. The needed generalised Bayesian parameter estimate evolves according to Bayes rule, $h \in h^*$,

$$f(\Theta_h|\mathcal{P}_{a_t}) \propto f(\Delta_{a_t}|\Theta_h, \mathcal{P}_{a_t}, a_t, h) f(\Theta_h|\mathcal{P}_{a_{t-1}}) \tag{4.28}$$

starting from a given prior pdf of parameters $f(\Theta_h|h) \equiv f(\Theta_h|h, \mathcal{P}_{a_1})$.

The equivalent batch version is, $h \in h^*$,

$$\begin{aligned} f(h|\mathcal{P}_{a_{t+1}}) &\propto \frac{\mathcal{I}_h(\mathcal{P}_{a_{t+1}})}{\mathcal{I}_h(\mathcal{P}_{a_1})} f(h|\mathcal{P}_{a_1}) \\ \mathcal{I}_h(\mathcal{P}_a) &\equiv \int \mathcal{L}(\Theta_h, \mathcal{P}_a) f(\Theta_h|h)\, d\Theta_h \\ \mathcal{L}(\Theta_h, \mathcal{P}_{a_{t+1}}) &= \prod_{\tau=1}^{t} f(\Delta_{a_\tau}|\Theta_h, \mathcal{P}_{a_\tau}, a_\tau, h),\ \mathcal{L}(\Theta_h, \mathcal{P}_{a_1}) \equiv 1. \end{aligned} \tag{4.29}$$

Proof: It uses basic algebra with pdfs, Proposition 2.3.4, respecting natural conditions of acting adopted for all involved decisions. □

Problem 4.2.3 (How to treat select the list of hypotheses?) *Mechanical ways of generating list of hypotheses make $\bar{h}$ extremely large and consequently their testing infeasible. GUHA methodology [] represents an example of a systematic way how to proceed. Nothing, however, has been elaborated which would fit the methodology presented here.*

Problem 4.2.4 (How to treat extending list of hypotheses?) *Hypotheses are usually created gradually. It opens a question, how to extend the existing set of hypothesis and how to exploit former data so that the new hypothesis is compared in a fair way. A lot of partial steps have been done in this respect but a systematic design and analysis are missing.*

4.2.4 One-step-ahead prediction

One-step-ahead prediction, an extrapolation of experience to the potentially observable data that are part of the ignorance set, is a very basic task that fits in our framework as follows.

- Decomposition of the behaviour $\mathcal{Q} \equiv (\mathcal{P}_a, a, \mathcal{F}_a) \equiv (\mathcal{P}_{\hat{\Delta}}, \hat{\Delta}, \Delta) \equiv$ (data at disposal, prediction, potentially observable data).
- Admissible rules are $\mathcal{R} : \mathcal{P}^*_{\hat{\Delta}} \rightarrow \hat{\Delta}^*$.
- Loss function $\mathcal{Z}$ measures a distance of $\hat{\Delta}$ and Δ and may depend on $\mathcal{P}_{\hat{\Delta}}$.

The outer system model (2.20) needed is $f(\Delta|\mathcal{P}_{\hat{\Delta}})$ as the natural conditions of decision making (2.33) imply that the decision $\hat{\Delta}$ is superfluous in conditioning. This pdf is obtained according to Proposition 2.6.2 as the generalised Bayesian prediction. For a given $\mathcal{P}_{\hat{\Delta}}$, the optimal point prediction is, Proposition 2.4.1,

$$\hat{\Delta}(\mathcal{P}_{\hat{\Delta}}) \in \text{Arg} \min_{\hat{\Delta} \in \hat{\Delta}^*} \int \mathcal{Z}(\mathcal{P}_{\hat{\Delta}}, \hat{\Delta}, \Delta) f(\Delta|\mathcal{P}_{\hat{\Delta}})\, d\Delta.$$

Remarks 4.2.4

1. *The prediction experience may contain results of other decisions, like system inputs. In this case $\mathcal{P}_{\hat{\Delta}} \equiv (\mathcal{P}_u, u)$, inputs are integral part of experience and they influence both value of the innovation and its prediction.*
2. *A popular option is the quadratic loss $\mathcal{Z}(\mathcal{P}_{\hat{\Delta}}, \hat{\Delta}, \Delta) = ||\Delta - \hat{\Delta}||^2_{Q(\mathcal{P}_{\hat{\Delta}})}$ with quadratic norm $||\cdot||_{Q(\mathcal{P}_{\hat{\Delta}})}$ weighted by a positive definite kernel $Q(\cdot)$. Similarly as in estimation case, see Remarks 4.2.1, the conditional expectation $\mathcal{E}[\Delta|\mathcal{P}_{\hat{\Delta}}]$ is the optimal prediction for any weight $Q(\cdot)$.*
3. *The discussed prediction should be called* point prediction. *As in the estimation, set estimates (credibility intervals) might be searched for. Their discussion is identical with that presented in Section 4.2.2.*

4.2.5 One-step-ahead control

One-step-ahead control tries to influence the response of the controlled system in the nearest future by choosing an appropriate input. This problem is formulated within our framework as follows.

- Decomposition of the behaviour $\mathcal{Q} \equiv [\mathcal{P}_a, a, \mathcal{F}_a] \equiv [\mathcal{P}_u, u, (y, x)] \equiv$ [(data at disposal and reference values, system input, (system output, system state)].
- Admissible rules are of the form $\mathcal{R} : \mathcal{P}_u^* \to u^*$.
- Loss function $\mathcal{Z}$ measures distance of (y, x, u) to desirable known reference values (y_r, x_r, u_r) and may depend on $\mathcal{P}_u$.

The outer system model (2.20) needed is $f(y, x|\mathcal{P}_u, u) = f(x|\mathcal{P}_u, u, y)f(y|\mathcal{P}_u, u)$. The factors creating it are obtained by applying Proposition 2.6.1 if we

- set $x \equiv \Theta_t$,
- take the filtration result after data updating as the first factor,
- use the output predictor $f(y|\mathcal{P}_u, u)$ as the second factor.

For the given $\mathcal{P}_u$, the optimal inputs is (Proposition 2.4.1)

$$u(\mathcal{P}_u) \in \text{Arg} \min_{u \in u^*} \int \mathcal{Z}(\mathcal{P}_u, y, x, u, y_r, x_r, u_r) f(x|\mathcal{P}_u, u, y) f(y|\mathcal{P}_u, u)\, dx dy.$$

Remarks 4.2.5

1. *The input u influences directly the outer model of the system. This property makes the main difference of the control task from estimation.*
2. *This case is a special version of the multi-step-ahead control design, see Section 4.3.4, where additional details are presented.*
3. *Often, just output of the system is controlled. Then the output predictor is needed only.*
4. *Output and state reference values are often called* set points.
5. *Quadratic distance is a popular option used. Then, the conditional expected loss optimised depends on 1st and 2nd conditional moments of the outer model. Often, the conditional variance does not depend on the chosen input. Then, a quadratic form in the conditional expectation $\mathcal{E}\{[y, x]|\mathcal{P}_u, u\}$ is just needed for determining the optimal input.*
6. *Often, multi-step-ahead control design is approximated by solving it as a sequence of unrelated one-step-ahead control design (using an extreme version of* receding-horizon strategy*, see Section 5.4.1). The substantial loss of quality up to the closed-loop instability may be encountered in this case [9]. This extreme should be avoided as much as possible.*
7. *Fully probabilistic version, see Proposition 2.5.4, of this problem may be easily formulated and solved. It is the simplest application of Proposition 2.5.4 and as such is omitted here.*

4.3 Dynamic tasks

Dynamic decision-making discussed here forms the core of this work. This justified a bit more detailed discussion.

4.3.1 Sequential estimation

Mostly, we consider decision tasks with a fixed horizon $\bar{t} < \infty$. Sequential estimation is an important exception.

If technically possible, it is often advantageous to decide whether to perform a new measurement before making the final decision. This task makes sense if we respect the price paid in connection with the data acquisition. The optimal decision strategies, called *sequential*, then lead in average to a shorter observation process. It makes the overall decision making cheaper in comparison with strategies that rely on a pre-defined sufficiently long decision horizon. Sequential estimation serves well for describing the idea. Other variants, like sequential testing of hypothesis might be formulated and solved similarly.

Sequential point estimation can be cast in our framework as follows.

- Decomposition of the behaviour $\mathcal{Q} \equiv [\mathcal{P}_{a_t}, a_t, \mathcal{F}_{a_t}] \equiv [\mathcal{P}_{a_t}, (\hat{\Theta}_t, s_t), (\Theta, \Delta_{a_t})]$ [data at disposal, (estimate,stopping flag), (unknown parameter,innovations)].
- Admissible strategies consist of rules $\mathcal{R}_t : \mathcal{P}^*_{a_t} \to (\hat{\Theta}^*_t, s^*_t)$, $\Theta^* \subset \hat{\Theta}^*$, $s^*_t \equiv$ {stop measuring and estimate Θ, wait for a new measurement} $\equiv \{0, 1\}$,
- Loss function

$$\mathcal{Z} = \begin{cases} \sum_{\tau \le t} c(\mathcal{P}_{a_\tau}) + z(\mathcal{P}_{a_t}, \Theta, \hat{\Theta}_t) & \text{if} \quad s_t = 0 \text{ and } s_\tau = 1, \ \forall \tau < t \\ \sum_{\tau \le t} c(\mathcal{P}_{a_\tau}) & \text{if} \quad s_\tau = 1 \ \forall \tau \le t \end{cases} \tag{4.30}$$

where $z(\mathcal{P}_{a_t}, \Theta, \hat{\Theta}_t)$ measures a distance of Θ and its estimate $\hat{\Theta}$. $c(\mathcal{P}_{a_\tau})$ denotes a positive price of τth observation.

Proposition 4.3.1 (Sequential estimation) *Let us consider the sequential estimation and assume that there is an admissible strategy for which the expected loss is finite. The following inequalities express the sufficient condition for an index t to be the time moment at which observation should be stopped*

$$\mathcal{E}\left[\left(z(\mathcal{P}_{a_t}, \Theta, \hat{\Theta}_t) - z(\mathcal{P}_{a_{t+k}}, \Theta, \hat{\Theta}_{t+k}) - \sum_{\tau > t}^{t+k} c(\mathcal{P}_{a_\tau})\right) |\mathcal{P}_{a_t}\right] \le 0, \ \forall k = 1, 2, \ldots. \tag{4.31}$$

In (4.31), $\hat{\Theta}_{t+k}$, $k = 0, 1, 2, \ldots$ denote parameter estimate based on $\mathcal{P}_{a_{t+k}}$ minimising $\mathcal{E}[z(\mathcal{P}_{a_{t+k}}, \Theta, \hat{\Theta})]$.

Proof: Let (4.31) be fulfilled. Then, combining the form of the loss function (4.30), the fact that the optimal stopping time has to be determined using its experience and finiteness of the loss for the optimal solution we get, $\forall k = 1, 2, \ldots$,

$$\mathcal{E}\left[\left(z(\mathcal{P}_{a_t}, \Theta, \hat{\Theta}_t) + \sum_{\tau=1}^{t} c(\mathcal{P}_{a_\tau})\right) |\mathcal{P}_{a_t}\right] \le \mathcal{E}\left[\left(z(\mathcal{P}_{a_{t+k}}, \Theta, \hat{\Theta}_{t+k}) + \sum_{\tau=1}^{t+k} c(\mathcal{P}_{a_\tau})\right) |\mathcal{P}_{a_t}\right].$$

Using isotonicity of the expectation (taken over $\mathcal{P}_{a_t}$), see Proposition 2.3.6, we find that the chosen decision cannot be improved by any estimate that uses more measurements than the inspected one. □

The outer system model (2.20) needed for this decision making coincides with the generalised Bayesian estimate (posterior pdf) given in Proposition 2.6.2.

Remarks 4.3.1

1. *The ability to evaluate* $\mathcal{E}\left[z(\mathcal{P}_{a_{t+k}},\Theta,\hat{\Theta}_{t+k})|\mathcal{P}_{a_t}\right]$ *is decisive for a practical solvability of the problem.*

2. *Stopping rules used for speeding up extensive simulations [52] based on a simple sequential estimation serve as an example of their, still underestimated, usefulness.*

3. *The dependence of the observation price can be effectively exploited when the sequential estimation is performed in an inner loop of some optimisation process: the closer we are to the optimum the lower this weight is.*

4.3.2 Multi-step-ahead prediction

Multi-step-ahead prediction, an extrapolation of experience to data in a more distant part of the ignorance set, is an extension of one-step-ahead prediction, Section 4.2.4, and fits in our framework as follows.

- Decomposition of the behaviour $\mathcal{Q} \equiv (\mathcal{P}_{a_t}, a_t, \mathcal{F}_{a_t}) \equiv (\mathcal{P}_{\hat{\Delta}_{t+j|t}}, \hat{\Delta}_{t+j|t}, \Delta_{t+j}) \equiv$ (data at disposal at time t, prediction of innovations at time $t+j$ based on data at disposal, innovation potentially observable at time $t+j$, $j>1$).
- Admissible strategies are formed by rules $\mathcal{R}_t : \mathcal{P}^*_{\hat{\Delta}_{t+j|t}} \to \hat{\Delta}^*_{t+j|t}$.
- Loss function $\mathcal{Z}$ measures $\mathcal{P}_{\hat{\Delta}_{t+j|t}}$-dependent distance of $\hat{\Delta}_{t+j|t}$ and Δ_{t+j}.

The outer system model (2.20) needed is $f(\Delta_{t+j}|\mathcal{P}_{\hat{\Delta}_{t+j|t}})$ and the optimal point prediction is (Proposition 2.4.1)

$$\hat{\Delta}_{t+j|t} \in \mathrm{Arg}\min_{\hat{\Delta}\in\hat{\Delta}^*} \int \mathcal{Z}(\mathcal{P}_{\hat{\Delta}_{t+j|t}}, \hat{\Delta}, \Delta_{t+j|j}) f(\Delta_{t+j}|\mathcal{P}_{\hat{\Delta}_{t+j|t}})\, d\Delta_{t+j}.$$

Let us illustrate the construction of the needed outer model in a special case of the controlled system with $\Delta_t = y_t =$ system output. It is controlled by a randomised controller described by its outer model $f(u_t|\mathcal{P}_{u_t})$, Agreement 2.4.1. The outer model of the system $f(y_t|\mathcal{P}_{u_t}, u_t)$ is either given or obtained as the predictive pdf by filtering, Proposition 2.6.1 or through parameter estimation, Proposition 2.6.2. The experience of the multi-step-ahead predictor is assumed to coincide with $\mathcal{P}_{u_t}, u_t$.

Using the calculus with pdfs, Proposition 2.3.4, the needed outer model can be expressed with the help of the available elements as follows

$$\begin{aligned}
&f(y_{t+j}|u_t, \mathcal{P}_{u_t}) = \\
&= \int f(y(t_(t+j)), u((t+1)_(t+j))|\mathcal{P}_{u_t}, u_t)\, d\,(y(t_(t+j-1)), u((t+1)_(t+j))) = \\
&= \int \prod_{\tau=t}^{t+j} f(y_\tau|\mathcal{P}_{u_\tau}, u_\tau,) f(u_\tau|\mathcal{P}_{u_\tau}) d\,(y(t_(t+j-1)), u((t+1)_(t+j)))\,.
\end{aligned}$$

Remarks 4.3.2

1. *The need to know the future* control strategy *makes the main difference of this task from one-step-ahead prediction where just knowledge of inputs is sufficient.*

2. *As a rule, the computed marginal pdf* $f(y_{t+j}|u_t, \mathcal{P}_{u_t})$, $j > 1$ *is much flatter than the one-step ahead predictor. Formally, it results from integration (averaging) over intermediate predicted values. It reflects the fact that it is much harder (less reliable) to make a long term prediction. The situation is the more pronounced the longer the prediction horizon* j *is.*

3. *Notice that integrations over the intermediate quantities is done also over their values in conditioning. This makes multi-step prediction highly non-linear task.*

4. *Sometimes, the parameterised model can be constructed directly with the gap* j, *i.e. in the form* $f(\Delta_{t+j}|\Theta, \mathcal{P}_{\hat{\Delta}_{t+j|t}})$. *Then, there is no formal difference to the one-step-ahead prediction. If there is a gap in system-output sequence and no gap in system-input sequence the prediction of future inputs is necessary.*

   Predictors of this type are used in connection with so called MUSMAR controllers *[53].*

5. *The point prediction is discussed here. Of course, the set estimation may be also formulated and solved using our general tools.*

4.3.3 Filtering derivates

Filtering, described in Proposition 2.6.1, provides filtration and one-step-ahead prediction of an unknown internal variable Θ_t and can be extended to multi-step prediction in the same manner as the innovation prediction discussed in Section 4.3.2. The fact that we never observe directly time varying Θ_t calls for a novel task called *smoothing*. The strategy to be designed is called *smoother*.

The smoother estimates the internal quantities Θ_{t-j}, $j > 1$ using also the measured data reflecting newer variables $\Theta((t-j+1)\_t)$. Its construction is cast in our framework as follows.

- Decomposition of the behaviour $\mathcal{Q} \equiv (\mathcal{P}_{a_t}, a_t, \mathcal{F}_{a_t}) \equiv (\mathcal{P}_{\hat{\Theta}_{t-j|t}}, \hat{\Theta}_{t-j|t}, \Theta_{t-j}) \equiv$ (data at disposal at time t, smoothed estimate of Θ_{t-j} based on data at disposal, unknown internal variable at time $t-j$, $j > 1$).

- Admissible rules are of the form $\mathcal{R} : \mathcal{P}^*_{\hat{\Theta}_{t-j|t}} \rightarrow \hat{\Theta}^*_{t-j|t}$.

- Loss $\mathcal{Z}$ measures distance of $\hat{\Theta}_{t-j|t}$ and Θ_{t-j} and may depend on $\mathcal{P}_{\hat{\Theta}_{t-j|t}}$.

Assuming controlled system and $\mathcal{P}_{\hat{\Theta}_{t-j|t}} \equiv \mathcal{P}_{u_{t+1}}$, basic decision making lemma, Proposition 2.4.1, reads

$$\hat{\Theta}_{t-j|t} \in \text{Arg} \min_{\hat{\Theta} \in \hat{\Theta}^*_{t-j}} \int \mathcal{Z}(\mathcal{P}_{u_{t+1}}, \hat{\Theta}_{t-j}, \Theta_{t-j}) f(\Theta_{t-j}|\mathcal{P}_{u_{t+1}})\, d\Theta_{t-j}$$

Let us assume that we have at disposal both observation $f(y_t|\Theta_t, \mathcal{P}_{u_t}, u_t)$ and time evolution $f(\Theta_t|\Theta_{t-1}, \mathcal{P}_{u_t}, u_t)$, $t \in t^*$ models used in Proposition 2.6.1. The applied control strategy meets natural conditions of decision making (2.33). Using the calculus with pdfs, Proposition 2.3.4, and (2.33), the model needed in the above optimisation can be expressed with the help of the available elements as follows

$$f(\Theta_{t-j}|\mathcal{P}_{u_{t+1}}) \propto f(\Theta_{t-j}, y((t-j)\_t), u((t-j)\_t)|\mathcal{P}_{u_{t-j}}) =$$
$$= \int \prod_{\tau=t-j}^{t} f(y_\tau|\Theta_\tau, \mathcal{P}_{u_\tau}, u_\tau) f(\Theta_\tau|\Theta_{\tau-1}, \mathcal{P}_{u_\tau}, u_\tau) f(\Theta_{t-j}|\mathcal{P}_{u_{t-j}}), d\Theta((t-j+1)\_t)$$

The integrand contains the available models with available data substituted. The factor $f(\Theta_{t-j}|\mathcal{P}_{u_{t-j}})$ results from filtering described by Proposition 2.6.1. Smoothing is simply derived from filtering: this justifies title of this section.

4.3.4 Multi-step-ahead control

Multi-step-ahead control (an extension of one-step-ahead control, see Section 4.2.5) tries to push quantities characterizing the controlled system, namely (system output, system input, system state)= (y, u, x), to the desired reference values (y_r, u_r, x_r). Their proximity should be enforced on a whole time-interval up to the horizon $\bar{t}$ by choosing an appropriate sequence of inputs $u(\bar{t})$.

We consider here the most general combination of tracking and regulation to uncertain reference values. The specialisation to the case of (partially) known reference values is straightforward. This problem is formulated in our framework as follows.

- Decomposition of the behaviour $\mathcal{Q} \equiv (\mathcal{P}_{a_t}, a_t, \mathcal{F}_{a_t}) \equiv [\mathcal{P}_{u_t}, u_t, \mathcal{F}_{u_t}] \equiv$ [(observed data $\equiv y(t-1), u(t-1), y_r(t-1), u_r(t), x_r(t-1)$), system input u_t, (future system outputs $y(t\_\bar{t})$, states $x(\bar{t})$ and references $y_r(t\_\bar{t})$, $u_r((t+1)\_\bar{t})$, $x_r(t\_\bar{t})$)].
- Admissible strategy consists of a sequence *control laws* (decision rules) of the form $\{\mathcal{R}_t : \mathcal{P}^*_{u_t} \to u^*_t\}_{t\in t^*}$.
- Loss $\mathcal{Z}$ measures distance of $\mathcal{Q}$ to reference values $y_r(\bar{t}), u_r(\bar{t}), x_r(\bar{t})$.

The optimal control strategy can be designed using directly dynamic programming, see the general version in Proposition 2.5.1, the version with additive loss in Proposition 2.5.2 or the fully probabilistic version in Proposition 2.5.4.

The needed outer model is $f(y_t, x_t, y_{r;t}, x_{r;t}|\mathcal{P}_{u_t}, u_t, u_{r;t})$. It is obtained via filtering with unknown internal variables $\Theta_t = (x_t, y_{r;t}, x_{r;t})$ using Proposition 2.6.1.

Remarks 4.3.3

1. *The optimal controller consists of the combination of filtering and controller acting on its results.*
2. *Control of the system with unknown time invariant parameters Θ, the central topic of main stream adaptive control [3], is a special case of the described design with the trivial time evolution $\Theta_t = \Theta_{t-1}$. The parameter estimation, see Proposition 2.6.1, provides the necessary outer system model in this case.*

3. *It is fair to say that the optimal multi-step-ahead control has the widest gap between the optimal design and practically optimal design. It is rarely analytically or numerically feasible. For this reason, a lot of heuristic approximation techniques have been developed. Some of them are discussed in Chapter 5.*

4. *The time evolution model* $f(x_t, y_{r;t}, x_{r;t}|x_{t-1}, y_{r;t-1}, x_{r:t-1}, \mathcal{P}_{u_t}, u_t, u_{r;t})$ *has to describe the evolution of both system state and uncertain reference values.*

5. *The typical control design optimising closed-loop behaviour with respect to randomly occurring step changes can be obtained by modelling the output and/or state references as random walks with time-varying dispersion [54].*

6. *The described general cases cover the situation when we try to follow evolution of another uncertain object. Rescue/military interpretations of this case are straightforward. The key message is that the dynamics of the uncertain target has to be modelled for finding the optimal strategy.*

   The situation is simplified whenever target values are (partially) known. Then the corresponding outer models reduce formally to Dirac delta functions on a given known support.

Chapter 5

Evaluation techniques

Here, we review basic evaluation techniques and tools that create *bridge between optimal and practically optimal design.* All of them try to cope with the complexity restrictions. The complexity of the optimal design stems from the fact that the optimal "plan" of dynamic decisions is essentially searched for by comparing all possible behaviours resulting from the interaction of the system and judged strategy. The design complexity is significantly influenced by the richness of the inspected space. Its reduction is behind the majority of available approximation schemes including those discussed here.

5.1 Adaptive systems

The ideal solution of the decision-making under uncertainty is described by the combination of Bayesian filtering, see Proposition 2.6.1, and dynamic programming, see Propositions 2.5.2, 2.5.4. The functional equations describing them are mostly computationally infeasible. Consequently, their use cannot live without approximations. Even the simpler fully probabilistic design, see Proposition 2.5.4, operates on infinite-dimensional objects.

The involved multivariate functions should be represented in a computer, i.e. in the device that can operate on a high but finite amount of values. Thus, a sort of approximation is needed. Global approximation of functions of many variables we are dealing with is known to be computationally hard. On the other hand, the application of the strategies resulting from the design requires knowledge of the discussed solutions only for the recorded experience. Thus, it is sufficient to know them *locally around the actual experience.* This can often be approximately achieved. Such *local approximators* can be identified with *adaptive systems* [55]. They exploit majority of approximation techniques described below.

Remarks 5.1.1

1. *Note that there is no formal definition of adaptive systems. Their operational description is, for instance, in [3]. Essentially, any combination of learning and decision making module is called adaptive.*

   The toolset described there has a significant overlap with techniques presented here.

2. *We found the coined understanding of adaptive systems as local approximators very useful. It helps us to have a unified view on existing practical strategies and opens a way for designing novel ones [55]. Moreover, it shows that the adaptive systems will be inevitably used in future due to the theoretically provable need for local approximations.*

5.2 Suboptimal estimation

Parameter estimation, described by Proposition 2.6.2, is a special case of filtration solved by Proposition 2.6.1. Recall that the estimation relies on a "trivial" time-evolution model $f(\Theta_{t+1}|\Theta_t, \mathcal{P}_{a_{t+1}}, a_{t+1}) = \delta(\Theta_{t+1} - \Theta_t)$. Majority of techniques suitable for approximate filtering, see Section 5.3, can be used in this case, too. The discussion of this section is restricted to evaluation techniques specific to estimation only.

5.2.1 Equivalence approach

Under natural conditions of decision-making (2.33), the generalised Bayesian estimation (Proposition 2.6.2) updates the posterior pdfs according to the Bayes rule (2.40)

$$f(\Theta|\mathcal{P}_{a_{t+1}}) = \frac{f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) f(\Theta|\mathcal{P}_{a_t})}{f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)}, \; t \in t^*. \tag{5.1}$$

Complexity of these pdfs increases quickly with an increasing amount of data, with increasing t. Exponential family (4.3) is essentially the only exception from this rule. This section tries to cope with *recursive estimation applicable out of exponential family.*

Recursively feasible representation

Always limited capabilities of a computer call for a reduced representation of the propagated posterior pdfs. It is a peculiar task as the posterior pdfs concentrate quickly on a very narrow support at a priori unknown position in Θ^*, see Proposition 3.1.1. Thus, a representation on a sufficiently fine grid that does not miss the final position becomes soon computationally prohibitive. The way out has been elaborated in a sequence of papers and summarised in [25]. Here, we just outline the essence of this *equivalence approach.*

Proposition 5.2.1 (Equivalence-preserving mapping) *Let the posterior pdf $f(\Theta|\mathcal{P}_{a_t}) \in f^*(\Theta|\mathcal{P}_{a_t}) \equiv$ a set of pdf with a common, time, data and parameter invariant support. Let the mapping*

$$\mathcal{G}_t : f^*(\Theta|\mathcal{P}_{a_t}) \rightarrow g_t^* \tag{5.2}$$

assign to each element of $f^*(\Theta|\mathcal{P}_{a_t})$ a finite-dimensional vector statistics $g_t \equiv g(\mathcal{P}_{a_t})$ representing it.

Then, the values of g_t can be exactly recursively updated *using only its previous value and the current parameterised model $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t)$ iff $\mathcal{G}_t$ is a* time-invariant linear mapping $\mathcal{G} \equiv \mathcal{G}_t$, $t \in t^*$, *acting on logarithms of the pdfs involved. The logarithmic pdfs are treated as functions of Θ.*

$\mathcal{G}_t$ has to map Θ-independent elements to zero.

Proof: To demonstrate necessity is rather hard and the interested reader is referred to [56, 57]. To show that the conditions on $\mathcal{G}_t \equiv \mathcal{G}$, $t \in t^*$, are sufficient is simple and instructive. They become obvious if we apply $\mathcal{G}$ to the logarithmic version of the Bayes rule (5.1) and use both its time-invariance and linearity. The normalising term $\ln(f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t))$ is independent of Θ and as such mapped to zero. The recursion for g_t is then simply

$$g_t = \mathcal{G}[\ln(f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t))] + g_{t-1}, \text{ with } g_0 = \mathcal{G}(f(\Theta)) \equiv \mathcal{G}(\text{prior pdf}). \quad (5.3)$$

□

Note that (5.3) becomes the true recursion if we need not store whole past observed data for evaluating $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t)$. Thus, similarly as in the case of exponential family, see Agreement 4.1.1, we restrict ourselves to such models. Together with their formal description we introduce other notions useful in the subsequent discussion.

Agreement 5.2.1 (Finite memory; Empirical pdf; Riezs representation) *The parameterised model is said to have* finite memory *iff*

$$f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) \equiv M(\Theta, \Psi_t) \quad (5.4)$$

where $M(\cdot,\cdot)$ is a known function of Θ and of a finite dimensional data vector Ψ_t that can be updated recursively

$$(\Psi_{t-1}, a_t, \Delta_{a_t})^* \rightarrow \Psi_t^*.$$

Let $\{\Psi_\tau\}_{\tau=1}^t$ be measured data vectors. Then,

$$f_t(\Psi) \equiv \frac{1}{t}\sum_{\tau=1}^{t} \delta(\Psi - \Psi_\tau), \;\; \Psi \in \Psi^* \equiv \bigcup_{t \in t^*} \Psi_t^* \quad (5.5)$$

is the (formal) empirical pdf *of Ψ. The used Dirac delta function $\delta(\cdot)$ is the linear functional assigning to (reasonable) functions $B(\Psi)$ their values at zero argument. It has Riezs integral representation, [58],*

$$\int_{\Psi^*} B(\Psi)\delta(\Psi)\, d\Psi = B(0) \quad (5.6)$$

where $\delta(\Psi)$ has to be taken as generalised function [59].

We assume that the same representation exists for the linear mapping $\mathcal{G}$ introduced in Proposition 5.2.1, i.e. there is a, possibly generalised, vector function $G(\Theta)$ such that

$$\mathcal{G}(C) \equiv \int_{\Theta^*} C(\Theta)G(\Theta)\, d\Theta \quad (5.7)$$

for any considered function $C(\Theta)$, $\Theta \in \Theta^*$.

Assuming (5.4), using the empirical pdf (5.5) and Riezs representation of G (5.7), we see that the representation of the posterior pdf can be written in the form

$$g_t = t\int_{\Psi^*}\left[\int_{\Theta^*} \ln[M(\Theta,\Psi)]G(\Theta)\, d\Theta\right] f_t(\Psi)\, d\Psi + g_0. \quad (5.8)$$

The integral in brackets defines the vector function

$$h(\Psi) \equiv \int_{\Theta^*} \ln[M(\Theta,\Psi)]G(\Theta)\,d\Theta. \tag{5.9}$$

With it, we get the equivalent form of (5.8)

$$g_t = t\int_{\Psi^*} h(\Psi)f_t(\Psi)\,d\Psi + g_0. \tag{5.10}$$

Recall, that left hand side of (5.10) is known as it can be update recursively according to (5.3) which has the equivalent "Riezs" form

$$g_t \underbrace{=}_{(5.3)} \int_{\Theta^*} \ln[M(\Theta,\Psi_t)]G(\Theta)\,d\Theta + g_{t-1} \underbrace{\equiv}_{(5.9)} h(\Psi_t) + g_{t-1},\ g_0 = \mathcal{G}(f(\Theta)). \tag{5.11}$$

Also, for the chosen parameterised model (5.4) and functions $G(\Theta)$ representing the admissible projections to g_t^*, the vector function $h(\Psi)$ is known as well as its values in measured data vectors Ψ_t. The *empirical pdf* $f_t(\Psi)$ (5.5) is the only incompletely known object in (5.10) as we are not able (or willing) to store all measured $\Psi_\tau,\ \tau \le t$ that determine it.

Approximation as point estimation

The posterior pdf, we are interested in, can be expressed in terms of empirical pdf as follows, cf. similar transformation in connection with Proposition 3.1.1,

$$f(\Theta|\mathcal{P}_{a_{t+1}}) \equiv f(\Theta|f_t(\cdot)) \propto f(\Theta)\exp\left[t\int_{\Psi^*} \ln[M(\Theta,\Psi)]f_t(\Psi)\,d\Psi\right]. \tag{5.12}$$

The empirical pdf is unknown to us, thus this posterior pdf is unknown, too. We want estimate it, we are facing the decision-making problem.

The selection of a suitable estimate $\hat{f}(\Theta|g_t)(\equiv$ a pdf acting on $\Theta^*)$ fits to our general decision-making framework as follows.

- $\mathcal{Q} \equiv (\mathcal{P}_a, a, \mathcal{F}_a) \equiv (g_t, \hat{f}(\Theta|g_t), f(\Theta|f_t(\cdot)) \equiv$ (stored statistic, estimate of the posterior pdf, the posterior pdf determined by the unknown empirical pdf $f_t(\Psi),\ \Psi \in \Psi^*$).
- Admissible decision rule are of the form

  $$\mathcal{R} : g_t^* \to \hat{f}^* \equiv \left\{\hat{f}(\Theta) : \int \hat{f}(\Theta)\,d\Theta = 1,\ \hat{f}(\Theta) \ge 0,\ \forall\Theta \in \Theta^*\right\}.$$

  Note that g_0 is a known fixed representation of the known prior pdf and as such is not explicitly mentioned as a part of the domain of $\mathcal{R}$.
- Loss function is Kulback-Leibler distance (2.26) $\mathcal{D}(\hat{f}(\Theta)||f(\Theta|f_t))$.

Formally, the optimal point estimate of the posterior pdf is, Proposition 2.4.1,

$$\begin{aligned} \hat{f}(\Theta|g_t) &\in \operatorname{Arg}\min_{\hat{f}\in\hat{f}^*} \mathcal{E}\left[\mathcal{D}(\hat{f}(\Theta)||f(\Theta||f_t))|g_t\right] = \qquad (5.13)\\ &= \operatorname{Arg}\min_{\hat{f}\in\hat{f}^*}\left[\mathcal{D}(\hat{f}(\Theta)||f(\Theta)) - t\int_{\Psi^*} \hat{f}(\Theta)\ln[M(\Theta,\Psi)]\mathcal{E}[f_t(\Psi)|g_t]\,d\Psi\right]. \end{aligned}$$

The conditional expectation $\mathcal{E}[\cdot|g_t]$ is taken with respect to the uncertain empirical pdf $f_t(\Psi),\ \Psi \in \Psi^*$.

The minimiser can be found explicitly as

$$\hat{f}(\Theta|g_t) \propto f(\Theta) \exp\left\{ t \int_{\Psi^*} \ln[M(\Theta,\Psi)] \mathcal{E}[f_t(\Psi)|g_t]\, d\Psi \right\}. \tag{5.14}$$

Specification of $\mathcal{E}[f_t(\Psi)|g_t]$

The estimate (5.14) depends only on the conditional expectation

$$E(\Psi) \equiv \mathcal{E}[f_t(\Psi)|g_t],\ \ \Psi \in \Psi^*. \tag{5.15}$$

Its choice and the resulting estimation algorithm are described in this subsection.

First we notice that the empirical pdf obeys the identity (5.10) that is linear in it. Consequently, its conditional expectation $\mathcal{E}[f_t(\Psi)|g_t]$ has to, see Proposition 2.3.6, fulfill it, too,

$$g_t = \int_{\Psi^*} h(\Psi) \mathcal{E}[f_t(\Psi)|g_t]\, d\Psi + g_0 \underbrace{\Leftrightarrow}_{(5.15)} g_t - g_0 = \int_{\Psi^*} h(\Psi) E(\Psi)\, d\Psi \tag{5.16}$$

We need the following agreement in order to explain the choice of $E(\Psi)$ we finally make.

Agreement 5.2.2 (Linear span over h-entries; Orthogonality) *The set of real functions defined on Ψ^* and spanned linearly over entries of the vector function $h(\Psi)$, see (5.9),*

$$e^* \equiv \left\{ \sum_{i \in i^* \equiv \{1,\ldots,\bar{h}\}} \alpha_i h_i(\Psi) \text{ with arbitrary real weights } \alpha_i \right\} \tag{5.17}$$

is embedded into a linear space E^* with the scalar product $< \cdot, \cdot >$. For $A, B \in E^*$ it is defined $< A, B > \equiv \int_{\Psi^*} A(\Psi) B(\Psi)\, d\Psi$.

The orthogonal complement $e^{*\perp}$ of e^* to E^* is the set of functions $A(\Psi)$

$$e^{*\perp} \equiv \{ A \in E^* \setminus 0 : < A, B >= 0\ \forall B \in e^* \}. \tag{5.18}$$

With this agreement, the restriction (5.16) on the constructed expectation $E(\Psi)$ (5.15) can be written in the equivalent form

$$g_{i;t} - g_{i;0} =< h_i, E >,\ \forall i \in i^* \equiv \{1, \ldots, \bar{h}\}. \tag{5.19}$$

This form and definition (5.18) imply that if some $E(\cdot) \in E^*$ fulfills (5.19) then also $E(\cdot) + A(\cdot)$ does whenever $A(\cdot) \in e^{*\perp}$. All such options are equivalent. At the same time, we have exploited all relevant information sources: the prior pdf $f(\Theta)$, the parameterised model $M(\Theta, \Psi)$ and data compressed into the recursively feasible statistics g_t. We have no reason to prefer any $A(\cdot) \in e^{*\perp}$. Consequently, we consider that $E(\cdot)$ (5.15)

$$E(\Psi) \in e^* \Leftrightarrow \mathcal{E}[f_t(\Psi)|g_t] = h'(\Psi)\alpha_t, \text{ for an } \bar{h}\text{-vector of weights } \alpha_t \equiv \alpha(g_t). \tag{5.20}$$

This assumption and the condition (5.19) determines the weights α_t as the solution of the set of linear algebraic equations

$$g_t - g_0 = H\alpha_t,\ H_{ij} \equiv < h_i, hj >,\ i,j \in i^*. \tag{5.21}$$

The constant square matrix H is positive semidefinite. Its regularity can simply be guaranteed by a proper choice of $h(\cdot)$ that can be influenced through the optional $G(\Theta)$ (5.7) determining the optional mapping (5.2). Then, the weights α_t and consequently of $E(\Psi) \equiv \mathcal{E}[f_t(\Psi)|g_t]$ (5.20) are uniquely determined and we have got the overall algorithm.

Algorithm 5.2.1 (Recursive estimation based on equivalence approach)
Off-line phase

1. *Select the parameterised model* $f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) \equiv M(\Theta, \Psi_t)$.
2. *Select the prior pdf* $f(\Theta)$.
3. *Select the (generalised) vector function* $G(\Theta)$ *such that* $\int G(\Theta)\, d\Theta = 0$.
4. *Prepare evaluation of the function*
$$h(\Psi) \equiv \int \ln(M(\Theta, \Psi))G(\Theta)\, d\Theta. \tag{5.22}$$
5. *Compute the positive definite matrix* $A \equiv \left[\int h(\Psi)h'(\Psi), d\Psi\right]^{-1}$.
6. *Prepare evaluation of the function*
$$H(\Theta) \equiv \int \ln(M(\Theta, \Psi))h(\Psi)\, d\Psi. \tag{5.23}$$
7. *Set* $t = 0$ *and* $\alpha_0 = 0$.

On-line use phase *for* $t \in t^*$

1. *Measure data vector* Ψ_t.
2. *Evaluate* $h(\Psi_t) \equiv \int \ln(M(\Theta, \Psi_t))G(\Theta)\, d\Theta$.
3. *Update weights* $\alpha_t = Ah_t + \alpha_{t-1}$.
4. *Exploit the obtained approximation of the posterior pdf*
$$\hat{f}(\Theta|g_t) \propto f(\Theta)\exp\left[H'(\Theta)\alpha_t\right]. \tag{5.24}$$

Remarks 5.2.1

1. *Zero value of initial weights* α_0 *reflects the fact that the weights are always determined by the increment* $g_t - g_0$*, see (5.21).*
2. *The name "equivalence approach" stresses the fact that the set of posterior pdfs is reduced to equivalence classes: pdfs with the same representation* g *cannot be distinguished.*

3. *The chosen form of the expectation (5.20) determining the final solution is definitely not unique. It has, however, the smallest quadratic norm among possible options and as such it has tendency to give the most flat estimate of the posterior pdf. This seems to be wise as the approximation involved adds uncertainty to the estimation results.*

4. *The required* commutativity of the updating and projecting *the posterior pdfs is crucial. The recursion for gs is exact and the approximation errors caused by the use of $\hat{f}(\Theta|g_t)$ instead of $f(\Theta|\mathcal{P}_{a_{t+1}}) \equiv f(\Theta|f_t(\cdot))$ do not accumulate! Use of a non-commutative projection $\mathcal{G}_t : f^*(\Theta|\mathcal{P}_{a_t}) \to g_t^*$ is always endangered by divergence as the estimation described by the Bayes rule can be viewed as a dynamic system evolving $f(\Theta|\mathcal{P}_{a_{t+1}})$ at stability boundary.*

5. *The indicated integrations represent computationally the most demanding part of the algorithm. They can be performed in off line mode if their results can be efficiently stored (the resulting functions interpolated).*

Problem 5.2.1 (How to choose the mapping $\mathcal{G}$?) *The (generalised) functions G represent the key tuning knobs of the approach. Options leading to discretisation of the function and/or its derivatives, or $M(\Theta, \Psi_i)$ on a grid of Ψ_i have been tried with a success but a deeper insight is needed in order to arrive to a cook-book.*

Problem 5.2.2 (Does exist an equivalent for filtration?) *There are indications that answer is negative. If it is true, other approaches relying on usual stabilising effect of time updating (2.37) have to be used. They, however, should have a smooth transition to the estimation as its special case.*

5.2.2 Estimation with forgetting

The case of *slowly varying parameters* is discussed here as a widely met intermediate case bridging estimation and filtering. It admits time variations of Θ_t but it assumes that $\Theta_{t+1} \approx \Theta_t$. At the same time, a more precise model of changes is not used.

We summarise an approach called *stabilised forgetting*, [60]. It is based on a flexible problem formulation and a simple application of Proposition 4.1.5. It proved to be a powerful tool for solving the considered ill-posed problem.

Let $f(\Theta_{t+1} = \Theta_t = \Theta|\mathcal{P}_{a_{t+1}})$ be the pdf that assumes no parameter changes and $f_A(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}})$ an *alternative pdf* that describes parameters after some changes. Let $\phi \in (0, 1]$ be the probability that the "correct" unknown pdf ${}^{[o]}f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}})$ has the best projection, see Remarks 3.1.1, equal to the former pdf and $1 - \phi$ the probability that the later one is relevant. Then, the best compromise $f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}})$ minimising its expected Kullback-Leibler distance to the unknown best projection is, see Proposition 4.1.5,

$$f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}}) \propto \left[f(\Theta_{t+1} = \Theta_t = \Theta|\mathcal{P}_{a_{t+1}})\right]^{\phi} \left[f_A(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}})\right]^{1-\phi}. \tag{5.25}$$

By using this formula, we approximate the time updating step (2.37) in filtering without explicitly specifying the model of time evolution (2.32).

Algorithm 5.2.2 (Stabilised forgetting)
Off-line phase

1. *Specify prior pdf* $f(\Theta_1 = \Theta) \equiv f(\Theta_1 = \Theta|\mathcal{P}_{a_1})$ *corresponding to the treated parameterised model* $f(\Delta_{a_t}|\Theta_t, \mathcal{P}_{a_t}, a_t)$.

2. *Select the probability* $\phi \in (0,1]$ *that parameters do not change.*

On-line phase, *do for* $t \in t^*$

1. *Collect the newest data* $\mathcal{P}_{a_{t+1}}$, a_{t+1}

2. *Perform data updating*

$$f(\Theta_t = \Theta|\mathcal{P}_{a_{t+1}}) \propto f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) f(\Theta_t = \Theta|\mathcal{P}_{a_t}).$$

3. *Select or update the alternative pdf* $f_A(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}})$.

4. *Approximate time updating (*forget*)*

$$f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}}) \propto \left[f(\Theta_{t+1} = \Theta_t = \Theta|\mathcal{P}_{a_{t+1}})\right]^{\phi} \left[f_A(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}})\right]^{1-\phi}.$$

Remarks 5.2.2

1. *The pdf (5.25) after the time updating is a compromise between the posterior pdf obtained under the hypothesis that* Θ_t *is time invariant and an externally supplied alternative* f_A. *The closer* ϕ *is to unity, the slower changes are expected, i.e. the higher weight the posterior pdf corresponding to the time invariant case gets.*

2. *Note that the* forgetting operation *(5.25) preserves the basic property of time updating: the posterior pdf on parameters propagates without obtaining any new measured information.*

3. *Let us assume that* $f_A \propto 1$. *Then* $f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}}) \propto \left[f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}})\right]^{\phi}$, *i.e. the pdf after time updating is a flattened version of the pdf obtained after data-updating (if* $\phi < 1$*). It is intuitively appealing as our uncertainty about parameters can hardly decrease without knowing a good time evolution model (2.32) and with no new information processed.*

4. *It is instructive to inspect the influence of forgetting on old data built in through the observation model. The older data are, the stronger flattening is applied to the corresponding model. Consequently, the older data influence the estimation results less than new ones. Data are gradually "forgotten". This explains why the probability* ϕ *is called* forgetting factor,

5. *The alternative pdf assigned to* Θ^* *expresses our belief where the parameters might move within the time interval* $[t, t+1)$ *while we have no new observable information.*

   Often, the pessimistic uniform alternative pdf ($\propto 1$*) has been used. This special case of stabilised forgetting is called* exponential forgetting. *It allows us to follow relatively fast parameter changes but it forgets the accumulated information with, often too high, exponential rate. For this reason, it is worth to preserve what we feel as a guaranteed information. The prior pdf* $f(\Theta_{t+1} = \Theta)$ *is a typical, reasonably conservative choice of the alternative pdf* $f_A(\Theta_{t+1} = \Theta|\mathcal{P}_{a_t})$.

6. *The non-trivial alternative pdf prevents us to forget the "guaranteed" information as it is always incorporated after flattening (exponential forgetting). This* stabilises *whole learning and reflects very positively in its numerical implementations. Without this, the posterior pdf may become too flat whenever the information brought by new data is not sufficient.*

   Note that insufficient informativeness of new data is more rule than exception. It is true especially in regulation problems in which the controller tries to make closed-loop as quiet as possible, i.e. it tries to suppress any new information in data.

7. *The forgetting factor ϕ can be either taken as tuning knob or estimated. The predictive pdf parameterised by it, however, depends on it in a very complex way so that a partitioned estimation has to be applied when its posterior pdf is estimated on a prespecified grid [61], see Section 5.3.1.*

   Also, the forgetting factor can be chosen in a pessimistic way as the maximiser of the reached minima, see Proposition 4.1.5.

8. *Practical importance of this particular case of slowly varying parameters cannot be over-stressed: vast majority of adaptive systems, see Section 5.1, rely on a version of forgetting combined with a version of receding horizon strategy, see Section 5.4.1. With it, the estimation and prediction are approximated locally within the time domain. They rely on data measured nearby the actual time.*

Problem 5.2.3 (How to make forgetting as good as possible?) *The stabilised forgetting is a well established and reliable technique that, however, still nee ds significant effort to be put in. For instance,*

- *the use of a fixed priror pdf as an alternative pdf seems to be over-conservative;*
- *permanent adaptation of the forgetting factor has not been inspected enough;*
- *use of multirate forgetting seems to be desirable.*

5.2.3 Quasi-Bayes estimation

The quasi-Bayes estimation described here is a special approximation techniques designed for mixtures of models introduced in Section 4.1.3. It can be seen as a very specialised case of filtration as it tries to estimate an internal time varying quantity (see below).

The standard Bayesian estimation, see Section 2.6.2, is infeasible with the mixture of models

$$f(\Delta_{a_t}|\Theta, \mathcal{P}_{a_t}, a_t) = \sum_{c\in c^*} \alpha_c f_c(\Delta_{a_t}|\Theta_c, \mathcal{P}_{a_t}, a_t), \quad \Theta \equiv \{\alpha_c, \Theta_c\}_{c\in c^*} \in \Theta^* \equiv$$

$$\left\{\Theta : \alpha \in \alpha^* \equiv \left\{\alpha : \alpha_c \ge 0, \sum_{c\in c^*} \alpha_c = 1\right\}, \ \Theta_c \in \Theta_c^*, \ c \in c^* \equiv \{1, \ldots, \bar{c} < \infty\}\right\}$$

as the number of terms in the likelihood function (2.42) (a product of parameterised models) blows up exponentially. The quasi-Bayes estimation is justified as follows.

Let us have the generalised Bayesian estimate in the form

$$f(\Theta|\mathcal{P}_{a_t}) \propto \prod_{c\in c^*} f(\Theta_c|\mathcal{P}_{a_t})\alpha_c^{v_{c;t-1}-1} \tag{5.26}$$

with some pdfs $f(\Theta_c|\mathcal{P}_{a_t})$ and scalars statistics $v_{c;t-1} > 0$.

In order to approximate Bayesian estimation, we exploit the following interpretation of the mixture of models (4.13). Let $c_t \in c^*$ be the unobservable random selector of the "active" component that takes the value c with probability α_c. Then,

$$f(\Delta_{a_t}, c_t|\Theta, \mathcal{P}_{a_t}, a_t) = f_{c_t}(\Delta_{a_t}|\Theta_{c_t}, \mathcal{P}_{a_t}, a_t)\alpha_{c_t}. \tag{5.27}$$

The marginalisation rule, see Proposition 2.3.4, implies that the mixture (4.13) is the marginal pdf of the joint pdf (5.27).

Using the joint pdf (5.27), the assumed form (5.26) of the parameter estimate to be updated, Bayes rule (2.8) and definition of Kronecker delta $\delta(c,\tilde{c})$ (4.5), we get

$$\begin{aligned} f(\Theta, c_t|\mathcal{P}_{a_{t+1}}) &\propto f_{c_t}(\Delta_{a_t}|\Theta_{c_t}, \mathcal{P}_{a_t}, a_t)\alpha_{c_t} \prod_{c\in c^*} f(\Theta_c|\mathcal{P}_{a_t})\alpha_c^{v_{c;t-1}-1} = \\ &= \prod_{c\in c^*} [f_c(\Delta_{a_t}|\Theta_c, \mathcal{P}_{a_t}, a_t)]^{\delta(c,c_t)} f(\Theta_c|\mathcal{P}_{a_t})\, \alpha_c^{v_{c;t-1}+\delta(c,c_t)-1}. \end{aligned} \tag{5.28}$$

The result of the data updating giving the posterior pdf of Θ could be gained from this pdf using marginalisation. It would help us to get rid off the unobservable selector c_t, see Proposition 2.3.4. It would, however, destroy the product form (5.26) of the posterior pdf. In order to avoid this trap, we simply replace the unknown $\delta(c, c_t)$ by its point estimate. We use its conditional expectation as the estimate optimal (at least) with respect to the quadratic loss, see Remarks 4.2.1. It can be computed explicitly:

$$\begin{aligned} \delta(c, c_t) &\approx w_{c;t} \equiv \mathcal{E}[c_t = c|\mathcal{P}_{a_{t+1}}] = f(c_t = c|\mathcal{P}_{a_{t+1}}) \propto \\ &\propto (v_{c;t-1} + 1) \int f_c(\Delta_{a_t}|\Theta_c, \mathcal{P}_{a_t}, a_t) f(\Theta_c|\mathcal{P}_{a_t})\, d\Theta_{c_t}. \end{aligned} \tag{5.29}$$

Note that the expectation over the unknown α is evaluated using formulae related to Dirrichlet distribution, see Chapter 6.

The substitution of the point estimate $w_{c;t}$ instead of $\delta(c, c_t)$ in (5.29) preserves the product form (5.26). Thus, the whole procedure can be prolonged. We summarise the algorithm for exponential family (4.3) for which it is of a practical interest. The results of Propositions 4.1.1, 4.1.4 are used.

Note that out of the exponential family an additional approximation has to be employed. It may be based, for instance, on the equivalence approach, see Section 5.2.1.

Algorithm 5.2.3 (Quasi-Bayes estimation in exponential family)
Off-line phase

1. *Select the number of components and their structure, i.e. the forms of pdfs $f_c(\Delta_{a_t}|\Theta_c, \mathcal{P}_{a_t}, a_t)$, $c \in c^*$. Exponential family (4.3) is at disposal for their choice.*

2. *Select prior statistics $V_{c,0}$, $\nu_{c,0} > 0$, $c \in c^*$ of conjugate prior pdfs (4.15) for estimation of parameters of individual components .*

3. *Select prior statistics $v_{c,0} > 0$, $c \in c^*$ for estimation of the unknown weights α_c of components.*

On-line phase, *do for $t \in t^*$*

1. *Collect the newest data Δ_{a_t}, a_t.*

2. *do for $c \in c^*$*

   (a) *Construct data vectors $\Psi_{c,t}$,*

   (b) *Compute probabilistic weights (5.29)*

$$w_{c;t} \propto (v_{c;t-1}+1)\frac{\mathcal{I}(V_{c;t-1}+B(\Psi_{c;t}),\nu_{c;t-1}+1)}{\mathcal{I}(V_{c;t-1},\nu_{c;t-1})}, \tag{5.30}$$

   (c) *Update statistics*

$$\begin{aligned} V_{c;t} &= V_{c;t-1}+w_{c;t}B(\Psi_{c;t}),\ \nu_{c;t}=\nu_{c;t-1}+w_{c|t} \\ v_{c;t} &= v_{c;t-1}+w_{c;t}. \end{aligned} \tag{5.31}$$

3. *Evaluate desirable characteristics of the approximate posterior pdf*

$$f(\Theta|\mathcal{P}_{a_{t+1}}) \propto \prod_{c\in c^*} A^{\nu_{c;t}}(\Theta_c)\exp[V'_{c;t}C(\Theta_c)]\alpha^{v_{c;t}-1}. \tag{5.32}$$

   and/or use this parameter estimate for evaluation of the predictive pdf.

Remarks 5.2.3

1. *The adopted approximation is intuitively appealing. The more we expect that the innovations Δ_{a_t} for given a_t has been generated by some component, the higher weight to the increment $B(\Psi_{c;t})$ of the corresponding statistics we assign.*

2. *It is straightforward and computationally cheap to allow slow parameter changes by employing the stabilised forgetting, see Algorithm 5.2.2.*

3. *Unlike other procedures applied to mixtures, it fits well to* on-line adaptive learning *of dynamic systems (with non-trivial regressors).*

4. *All known procedures are sensitive to initiation, [42]. It applies to the quasi-Bayes estimation even in a wider extent. This is the, probbably inevitable, price for its advantages.*

5. *The components describing multivariate innovations can be always factorised into product of parameterised models describing predicting single entry (single output), see Section 4.1.3. Each of these* factor*s may have its structure and may be even of a different type. Thus, we can combine, for instance, regression models with Markov chains. In this way, we get mixture model describing interrelated continuous and discrete data. It is of an extreme practical importance, especially in connection with the universal approximation property of mixtures.*

Problem 5.2.4 (How to make quasi-Bayes estimation complete?) *The quasi-Bayes algorithm proved to be a powerful complement to universal and flexible mixtures. A lot of subproblems still await for their solutions: initiation, structure estimation, multi-step control design etc.*

5.3 Suboptimal filtering

Filtering, see Proposition 2.6.1, is the basic technique for generating the outer model of the system needed for (sub)optimal decision-making. There is a bunch of, mutually complementary, general techniques that convert optimal learning into practically applicable suboptimal learning. Significant representants are outlined here.

5.3.1 Partitioned approach

This approach [61] suits to cases when the problem formulation, determined by the observation $f(\Delta_{a_t}|\Theta_t, \mathcal{P}_{a_t}, a_t)$ (2.31) and time evolution $f(\Theta_t|\Theta_{t-1}, \mathcal{P}_{a_t}, a_t)$ (2.32) models, has the following structure.

- $\Theta_t = (x_t, \Theta) \equiv$ (state of the system, time invariant parameters)
- filtering of x_t is feasible for any fixed $\Theta \in \Theta^*$
- a finite low-dimensional grid $\{\Theta_i \in \Theta^*\}_{i \in i^*}$ "covers" sufficiently the space of possible parameter values Θ^*.

The partitioned filtering runs as follows.

Algorithm 5.3.1 (Partitioned filtering)
Off-line phase

1. *Select a grid* $\{\Theta_i\}_{i \in i^*}$
2. *Assign prior probabilities* $f(\Theta_i|\mathcal{P}_{a_1})$ *of individual parameter values* Θ_i.
3. *Select prior pdfs of the initial state* $f(x_0|\Theta_i, \mathcal{P}_{a_1})$ *for all grid points* Θ_i.

On-line phase, *do for* $t \in t^*$

1. *Collect the newest data* Δ_{a_t}, a_t.
2. *Repeat for* $i \in i^*$
   (a) *Perform the conditional* data updating, *see (2.36),*

$$\begin{aligned} f(x_t|\Theta_i, \mathcal{P}_{a_{t+1}}) &= \frac{f(\Delta_{a_t}|x_t, \Theta_i, \mathcal{P}_{a_t}, a_t) f(x_t|\Theta_i, \mathcal{P}_{a_t})}{f(\Delta_{a_t}|\mathcal{P}_{a_t}, a_t)} \\ f(\Delta_{a_t}|\Theta_i, \mathcal{P}_{a_t}, a_t) &= \int f(\Delta_{a_t}|x_t, \Theta_i, \mathcal{P}_{a_t}, a_t) f(x_t|\Theta_i, \mathcal{P}_{a_t})\, dx_t \quad (5.33) \end{aligned}$$

   (b) *Update* probabilities of individual grid points, *see (2.8),*

$$f(\Theta_i|\mathcal{P}_{a_{t+1}}) \propto f(\Delta_{a_t}|\Theta_i, \mathcal{P}_{a_t}, a_t) f(\Theta_i|\mathcal{P}_{a_t}).$$

(c) *Select the action* a_{t+1} *you intend to use.*

(d) *Perform the conditional* time updating, *see (2.37),*

$$f(x_{t+1}|\Theta_i, \mathcal{P}_{a_{t+1}}, a_{t+1}) = \int f(x_{t+1}|x_t, \Theta_i, \mathcal{P}_{a_t}, a_t) f(x_t|\Theta_i, \mathcal{P}_{a_t})\, dx_t.$$

(e) *Evaluate the conditional* prediction, *see (2.35)*

$$f(\Delta_{a_{t+1}}|\Theta_i, \mathcal{P}_{a_{t+1}}, a_{t+1}) = \int f(\Delta_{a_{t+1}}|x_t, \mathcal{P}_{a_t}, a_t, \Theta_i) f(x_t|\mathcal{P}_{a_t}, \Theta_i)\, dx_t.$$

3. *Compute the approximate outer model of the system, see Proposition 2.3.4,*

$$f(\Delta_{a_{t+1}}|\mathcal{P}_{a_{t+1}}, a_{t+1}) = \sum_{i \in i^*} f(\Delta_{a_{t+1}}|\Theta_i, \mathcal{P}_{a_{t+1}}, a_{t+1}) f(\Theta_i|\mathcal{P}_{a_{t+1}}).$$

Remarks 5.3.1

1. *Essentially, a* bank of filters *"indexed" by* Θ_i *is run and probabilities of individual predictors are updated in a standard Bayesian way.*

2. *The approach relies on feasibility of individual filters and on the low dimension of the grid. Bootstrap type estimator, [24] and see Section 5.3.2, seems to be an efficient way how to reach it. The time variations of the corresponding grid bring, however, additional troubles.*

3. *Notice the order of the performed operations. Long term experience proved that it is far too easy to make an error in it.*

Problem 5.3.1 (How to select the grid and exploit results?) *Acceptable dimensionality of the grid is of a vital importance and it is often reached by adopting additional techniques of a "branch-and-bound" type. It is straightforward to remove* Θ_i *with low probabilities* $f(\Theta_i|\mathcal{P}_{a_t})$ *but generating of new promising values* Θ_i *is not easy. Moreover, the exploitation of the filtering results obtained so far has no systematic solution in multi-step control tasks [62].*

5.3.2 Bootstrap type methods

Knowing the posterior pdf or pdfs describing filtration it would be relatively simple to select a low dimensional non-uniform grid sufficient for a good approximation of the whole pdf. This simple observation forms a basis of a whole set approximate filters that use random samples from the current approximate pdfs for generating new grid serving for, hopefully better, approximation that is updated by new data, see e.g. [63]. We describe a version of such *bootstrap filtration* that exploits interpretation of random samples as a variable non-uniform grid.

Under natural conditions of decision-making (2.33), we search for practically feasible filtration

$$f_{t|t}(\Theta) \equiv f(\Theta_t = \Theta|\mathcal{P}_{a_{t+1}}, a_{t+1}) \underbrace{=}_{(2.33)} f(\Theta_t = \Theta|\mathcal{P}_{a_{t+1}}).$$

The internal quantities Θ_t evolve according to the time evolution model (2.32) $f(\Theta_t|\Theta_{t-1}, \mathcal{P}_{a_t}, a_t)$ and manifest themselves through the observation model $f(\Delta_{a_t}|\Theta_t, \mathcal{P}_{a_t}, a_t)$ (2.31), see Requirement 2.6.1. During the filtering also the one-step-ahead prediction of internal quantities

$$f_{t+1|t}(\Theta) \equiv f(\Theta_{t+1} = \Theta|\mathcal{P}_{a_{t+1}}, a_{t+1})$$

is computed, see Proposition 2.6.1.

Let $\Theta_{i;t|t-1},\ i \in i^* \equiv \{1, \ldots, \bar{i} < \infty\}$ denote random independent samples of Θ taken from the approximate pdf $f_{t|t-1}(\Theta)$. We also assume that the values $f_{t|t-1}(\Theta_{i;t|t-1})$ are known. Then, data updating (2.36) provides the values

$$f_{t|t}(\Theta_{i;t|t-1}) \propto f(\Delta_{a_t}|\Theta_{i;t|t-1}, \mathcal{P}_{a_t}, a_t) f_{t|t-1}(\Theta_{i;t|t-1}).$$

The pairs of arguments and corresponding pdf values $\{\Theta_{i;t|t-1}, f_{t|t}(\Theta_{i;t|t-1})\}_{i\in i^*}$ are available for an approximation of $f_{t|t}(\Theta),\ \Theta \in \Theta^*$. The simplest approximation by the empirical pdf is used here

$$f_{t|t}(\Theta) \approx \sum_{i\in i^*} \delta(\Theta - \Theta_{i;t|t-1}) w_{i;t|t},\ w_{i;t|t} \propto f_{t|t}(\Theta_{i;t|t-1}) \tag{5.34}$$

where δ is Dirac delta function and $\sum_{i\in i^*} w_{i;t|t} = 1$. With the approximation (5.34), the time updating (2.37) reduces to the the formula

$$f_{t+1|t}(\Theta) = \sum_{i\in i^*} f(\Theta_{t+1}|\Theta_{i;t|t-1}, \mathcal{P}_{a_{t+1}}, a_{t+1}) w_{i;t|t}.$$

The filtering recursion is then completed by generating random independent samples $\Theta_{i;t+1|t},\ i \in i^*$ from this pdf.

The predictive pdf is obtained by approximating the pdf $f_{t+1|t}(\Theta)$ by empirical pdf on the latest samples and using the standard prediction formula (2.35). Let us summarise the overall algorithm.

Algorithm 5.3.2 (Bootstrap filtering)
Off-line phase

1. *Select prior pdf* $f_{1|0}(\Theta) \equiv f(\Theta_1|\mathcal{P}_{a_1})$
2. *Generate random independent samples* $\{\Theta_{i;1|0}\}_{i\in i^*}$ *from it.*
3. *Evaluate the probabilistic weights* $w_{i;1|0} \propto f_{1|0}(\Theta_{i;1|0}),\ i \in i^*$.

On-line phase, *do for* $t \in t^*$

1. *Collect the newest data* Δ_{a_t}, a_t.
2. *Perform the* data updating *on the avaliable samples* $\Theta_{i;t|t-1}$

$$\begin{aligned} w_{i;t|t} &\propto f_{t|t}(\Theta_{i;t|t-1}) \propto f(\Delta_{a_t}|\Theta_{i;t|t-1}, \mathcal{P}_{a_t}, a_t) f_{t|t-1}(\Theta_{i;t|t-1}) \propto \\ &\propto f(\Delta_{a_t}|\Theta_{i;t|t-1}, \mathcal{P}_{a_t}, a_t) w_{i;t|t-1},\ i \in i^*. \end{aligned}$$

3. *Perform the approximate* time updating

$$f_{t+1|t}(\Theta) = \sum_{i\in i^*} f(\Theta_{t+1} = \Theta|\Theta_{i;t|t-1}, \mathcal{P}_{a_{t+1}}, a_{t+1}) w_{i;t|t}. \tag{5.35}$$

4. *Generate independent samples* $\{\Theta_{i;t+1|t}\}_{i\in i^*}$ *from the pdf* $f_{t+1|t}(\Theta)$.

5. *Evaluate the probabilistic weights* $w_{i;t+1|t} \propto f_{t+1|t}(\Theta_{i;t+1|t}),\ i \in i^*$.

6. *Compute the (approximate) outer model of the system*

$$f(\Delta_{a_{t+1}}|\mathcal{P}_{a_{t+1}}, a_{t+1}) = \sum_{i\in i^*} f(\Delta_{a_{t+1}}|\Theta_{i;t+1|t}, \mathcal{P}_{a_{t+1}}, a_{t+1}) w_{i;t+1|t}.$$

Remarks 5.3.2

1. *The algorithm is conceptually simple and universal. It needs, however, a large number of* Θ *samples. Their generating from the pdf (5.35) is, except of simple cases discussed in Section 5.6.2, far from being trivial. Complexity of this step might make the algorithm infeasible. Often, the feasibility is achieved by sophisticated techniques represented by Gibbs sampler [64] or adaptive sampling [52].*

2. *A massive parallelism might be used for making the algorithm computationally efficient.*

3. *Less crude approximations of pdfs than the empirical pdfs are desirable and necessary when time evolution is degenerated and some entries of* Θ_t *are time invariant. Typically, smooth approximations of Dirac delta function are used. A narrow Gaussian pdf "sitting" on the sample may be used. This can be interpreted as addition of an external noise to the original samples. Spline-type or mixture approximations might be used in some cases.*

4. *This group of methods is very adaptive, see Section 5.1 It generates approximators locally around measured behaviour.*

Problem 5.3.2 (How to make bootstrap complete?) *A lot has been reached in development of methods of this type. The state is, however, far from being finished. Problems range from sensitivity to initial values, efficient generating of samples, choice of approximating pdfs up to the choice of the number of samples and numerical (parallel) implementation.*

Problem 5.3.3 (Relationships of bootstrap & equivalence methods?) *The raw bootstrap method works satisfactory when the time evolution is sufficiently random so that approximation errors are overriden by it. Then, filtering is inherently stable and robust with respect to approximation errors. The closer we are to the determistic time evolution, the closer we are to parameter estimation which is described by operations at the stability boundary. Under such a situation, the bootstrap method should be close to a version of equivalence approach, see Section 5.2.1. This aspect has not been inspected and both these method could benefit from a deeper study of their relationships.*

5.3.3 Extended Kalman filtering

Here we outline a very special but a very widely used approximate filtering.

Filtering can be exactly solved if both observation and time evolution models are Gaussian with mean values depending linearly on the unknown Θ_t and with known covariances. All involved pdfs are Gaussian if the prior pdf is. Then, the filtering reduces to an algebraic evolution of two conditional moments, [10, 65] and also examples. In this case, the time evolution model can be characterised by

$$\begin{aligned} \mathcal{E}[\Theta_{t+1}|\Theta_t, \mathcal{P}_{a_{t+1}}, a_{t+1}] &= A(\mathcal{P}_{a_{t+1}}, a_{t+1})\Theta_t + B(\mathcal{P}_{a_{t+1}}, a_{t+1}) \quad (5.36) \\ \mathrm{cov}[\Theta_{t+1}|\Theta_t, \mathcal{P}_{a_{t+1}}, a_{t+1}] &= R(\mathcal{P}_{a_{t+1}}, a_{t+1}). \end{aligned}$$

A is a square matrix, B is a vector and R is a positive semidefinite square matrix. Similarly, the observation model is fully described by

$$\begin{aligned} \mathcal{E}[\Delta_{a_{t+1}}|\Theta_{t+1}, \mathcal{P}_{a_{t+1}}, a_{t+1}] &= C(\mathcal{P}_{a_{t+1}}, a_{t+1})\Theta_{t+1} + D(\mathcal{P}_{a_{t+1}}, a_{t+1}) \\ \mathrm{cov}[\Delta_{a_{t+1}}|\Theta_{t+1}, \mathcal{P}_{a_{t+1}}, a_{t+1}] &= r(\mathcal{P}_{a_{t+1}}, a_{t+1}). \quad (5.37) \end{aligned}$$

C is a matrix, D is a vector, and r is a positive semidefinite square matrix.

All involved arrays can be *known functions of available data and time.*

The *extended Kalman filter* is obtained if we approximate the treated conditional expectation $\mathcal{E}[\Theta_{t+1}|\Theta_t, \mathcal{P}_{a_{t+1}}, a_{t+1}]$ by the affine term of Taylor expansion around the latest point estimate $\hat{\Theta}_t \equiv \hat{\Theta}(\mathcal{P}_{a_t})$ of Θ_t and $\mathrm{cov}[\Theta_{t+1}|\Theta_t, \mathcal{P}_{a_{t+1}}, a_{t+1}]$ by the absolute term at the same point. Through a similar expansion for the observation model we arrive to forms (5.36), (5.37) and use them for an approximate updating of two moments of the considered filtration by using ordinary Kalman filter, see examples. The updated estimates serve for updating the expansion point.

Remarks 5.3.3

1. *Obviously, the local approximation used in extended Kalman filter is a rather crude. Consequently, a careful start up and various modifications based on a detailed analysis are needed in order to prevent divergence of* $\{\hat{\Theta}_t\}_{t\in t^*}$ *from* $\{\Theta_t\}_{t\in t^*}$*, [66].*

2. *The smoother the model is the higher is chance for success.*

Problem 5.3.4 (How to combine techniques?) *The extended Kalman filtering can be combined with partitioning approach, see Section 5.3.1, bootstrap methods, see Section 5.3.2, or mixture estimation, see Section 5.2.3. A systematic way is (to our best knowledge) not available.*

5.4 Suboptimal design

The design complexity is the key issue addressed repeatedly at various places of this text. All elements, see Chapter 4, are recommended to account for it. At the computational stage we are now discussing the design complexity stems from

- richness of the space of the ignorance part of the behaviour that have to be inspected for the optimal selection of decisions;
- complexity of models describing relationships of the experience and optional decisions to ignorance part of the behaviour.

The suboptimal design tries to reduce influence of one or both these sources of complexity. Here we describe the most common ways used.

5.4.1 Strategies examining reduced space

Here, we mention common techniques oriented on simplification of the space searched for.

Receding horizon

The reduction of the design horizon is the most obvious way to a simplified (suboptimal) design. The reduction obtained by planning just one-step-ahead has been popular for a long time [67].*Dynamic* decision-making, however, means that consequences of a decision are encountered far behind the time moment of its application. Consequently, the decision that is optimal when judged from a short-sighted perspective might be quite bad one from the long-term viewpoint, [9], see examples.

This observation has stimulated search for a compromise between the ideal planning over whole horizon of interest and short-sighted, locally optimising strategies.

When we plan looking just few steps ahead we get just an approximation of the optimal design. Thus, it is reasonable to correct the planned decisions permanently. After applying the initial decision(s) implied by the temporary design and after acquiring and exploiting a new measured knowledge about the system state, the design should be repeated. This is essence of the design technique called *receding horizon strategy.* Let us describe its algorithm in the case of additive loss function (2.22), recall (3.6), and for a pre-specified value T of the *receding horizon* $T < \bar{t}$.

Algorithm 5.4.1 (Receding-horizon strategy) *Repeat for* $t \in t^* \equiv \{1, \ldots, \bar{t}\}$,

1. *Find the strategy* $\mathcal{R}(t\_t + T) = (\mathcal{R}_t, \ldots, \mathcal{R}_{t+T})$ *approximately minimising*

$$\mathcal{E}\left[\sum_{\tau=t}^{t+T} z(\Delta(\tau), a(\tau))|\mathcal{P}_{a_t}\right]$$

   using the available experience and outer model of the system.

2. *Apply* $a_{t|t-1} \equiv \mathcal{R}_t(\mathcal{P}_{a_t})$ *for the available* $\mathcal{P}_{a_t}$.

3. *Extend the experience by the new data* Δ_{a_t}, $a_{t|t-1}$ *and use it in learning (perform filtration or estimation) so that an improved outer model of the system is obtained.*

Remarks 5.4.1

1. *Notice that receding horizon strategy makes sense even with fixed model (without learning). Then the plan is locally adapted to the observed experience. This version is important when a good but non-linear model is available.Such a decision system is adaptive according to our understanding outlined in Section 5.1.*

Problem 5.4.1 (How to choose lenght of receding horizon?) *Ideally,* $T = \bar{t}$ *should be chosen.* $T = 1$ *is known to be often insufficient for reaching a good quality. The selection of the smallest reasonable* T *is an open problem in spite of a wide range of strong but particular results available.*

Iterations spread in time

It is intuitively clear and practically verified that the good receding horizon T can be impractically large if the decision horizon $\bar{t}$ is large. In the case of additive loss, the asymptotic analysis of the design, see Section 3.2 and Proposition 3.2.1, shows that for $\bar{t} \rightarrow \infty$ the optimal decision is the minimiser in

$$^{[\infty]}\mathcal{V}(x_{t-1}) + C = \min_{a_t \in a_t^*} \mathcal{E}[z(x_t, a_t) + {}^{[\infty]}\mathcal{V}(x_t)|x_{t-1}, a_t].$$

This can be interpreted as one-step-ahead design for the partial loss $z(x_t, a_t)$ increased by the stationary Bellman function $^{[\infty]}\mathcal{V}(x_t)$. Knowing it, we could use receding horizon $T = 1$. At the same time, dynamic programming has been interpreted as an iterative search for this stationary value. These observations led to:

Algorithm 5.4.2 (Strategy with iterations spread in time) *Select an itial guess of stationary Bellman function $\mathcal{V}_0(x)$ and repeat for $t \in t^* \equiv \{1, \ldots, \bar{t}\}$,*

1. *Find the strategy $\mathcal{R}(t\_t + T) = (\mathcal{R}_t, \ldots, \mathcal{R}_{t+T})$ approximately minimising*

$$\mathcal{E}\left[\sum_{\tau=t}^{t+T} z(x_t, a_\tau) + \mathcal{V}_{t-1}(x_{t+T})|x_{t-1}\right]$$

   using the available experience and outer model of the system.

   Evaluate it for all possible $x \in x^*$ in the role of x_{t-1}.

2. *Take the final Bellman function $\mathcal{V}_t(x)$ resulting from this design as an updated approximation of the stationary Bellman function.*

3. *Apply $a_{t|t-1} \equiv \mathcal{R}_t(\mathcal{P}_{a_t})$ for the measured x_{t-1}.*

4. *Extend the experience by the new data Δ_{a_t}, $a_{t|t-1}$ and use it in learning (perform filtration or estimation) so that an improved outer model of the system is obtained.*

Remarks 5.4.2

1. *The receding horizon in this strategy can be chosen very low. Even the choice $T = 1$ does not prevent to reach the optimal stationary solution. The necessary optimising iterations are spread in real time $t \in t^*$.*

2. *The need to evaluate the reached minimum as a function (for all x) determines computational complexity of this procedure. For linear system and quadratic criterion (*LQ design*), the Bellman function is quadratic form in x and iterations are performed on a matrix kernel of this form. Then, this strategy becomes perfectly feasible. In more complex cases, an additional approximation is needed. It may, however, pay well back.*

3. *For LQ design combined with learning, the strategy is described in [9]. It has been found quite useful but at the same time sensitive to time variations of the estimated parameters.*

Problem 5.4.2 (Is the potential of this strategy well exploited?) *It seems that potential of this strategy is far from being exploited. Since the time of its invention computer power has increased substantially so that evaluation of minima on a rich subset of x^* could be feasible. Moreover, it could be combined with the simpler fully probabilistic design, see Section 2.5.2.*

Predictive strategies

There is a whole set of strategies that drastically reduce the space of inspected behaviour by working only with point predictions of uncertain quantities to be influenced. Sometimes they deal with a sort of set predictions by considering credibility predicion sets of a very simplified form. Such strategies are usually (a bit unluckily) labelled as *predictive strategies* [54, 68].

Let us consider an additive loss function with receding horizon T. It serves us an example demonstrating essence of these strategies. For minimisation purposes, the predictive strategies use the approximation

$$\mathcal{E}\left[\sum_{\tau=t}^{t+T} z(\Delta(\tau), a(\tau))|\mathcal{P}_{a_t}\right] \approx \sum_{\tau=t}^{t+T} z(\mathcal{E}\left[\Delta(\tau)||\mathcal{P}_{a_t}\right], a(\tau)).$$

Obviously, such an approximation may be quite rough when there is a non-negligible uncertainty (lack of knowledge or influence of noise). On the other hand, the gained simplification allows the user to concentrate on respecting hard bounds on actions and on other quantities. The practical significance of technological restrictions justifies this simplification and explains the popularity of such a treatment.

5.4.2 Strategies simplifying models

Let us consider the receding horizon strategy applied at time t. Then, a substantial degree of the design complexity is caused by the use of predictive pdfs $\{f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau)\}_{\tau=t}^{t+T}$ obtained through Bayesian estimation. They have the form, see Proposition 2.6.2,

$$f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau) = \int_{\Theta^*} f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau) f(\Theta|\mathcal{P}_{a_\tau})\, d\Theta. \tag{5.38}$$

For a relatively short planning horizon T, the predictors (5.38) can be approximated by

$$f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau) \approx \int_{\Theta^*} f(\Delta_{a_\tau}|\Theta, \mathcal{P}_{a_\tau}, a_\tau) \hat{f}_\tau(\Theta)\, d\Theta,\ \tau = t, \ldots, t+T, \tag{5.39}$$

with some $\hat{f}_\tau(\Theta) \approx f(\Theta|\mathcal{P}_{a_\tau})$. Bellow we outline wide-spread choices of $\hat{f}_\tau(\Theta)$.

Super-cautious strategy

If the posterior pdf $f(\Theta|\mathcal{P}_{a_t})$ is reasonably concentrated then it makes sense to exploit the approximate equality

$$f(\Theta|\mathcal{P}_{a_\tau}) \approx f(\Theta|\mathcal{P}_{a_t}) \equiv \hat{f}_\tau(\Theta),\ \tau = t, \ldots, t+T. \tag{5.40}$$

The receding horizon strategy combined with the approximation (5.39), (5.40) is called *super-cautious* as it is based on assumption that we learn nothing about unknown parameters during the planning horizon. The uncertainty of the parameter estimates (5.40) projected into the predictors through the formula (5.39) inhibits excessive decision values. It makes the strategy (typically controller) (super)cautious.

Cautious strategy

The assumption (5.40) is often too strong. Then, it is reasonable to weaken it to

$$f(\Theta|\mathcal{P}_{a_\tau}) \approx F(\tau, f(\Theta|\mathcal{P}_{a_t})) \equiv \hat{f}_\tau(\Theta),\ \tau = t,\ldots,t+T,\ \Theta \in \Theta^*. \tag{5.41}$$

The mapping $F(\tau,\cdot)$ modifies the posterior pdf $f(\Theta|\mathcal{P}_{a_t})$ to other, more concentrated, pdfs. The mapping F is chosen beforehand and does not use data belonging to the ignorance $\mathcal{F}_{a_t}$. Typically, the generated pdf $\hat{f}_\tau(\Theta)$ has the same 1st moment (expectation) as $f(\Theta|\mathcal{P}_{a_t})$ but its covariance $C_{\tau|t}$ decreases with increasing distance $\tau - t$ in a pre-specified way. Often,

$$C_{\tau|t} = b(\tau - t)C_t \equiv b(\tau - t) \times \text{ covariance computed for } f(\Theta|\mathcal{P}_{a_t}).$$

The following non-negative scalar functions $b(\kappa)$ serve as representative examples

$$b(\kappa) = 1/\kappa \text{ or } b(\kappa) = \begin{cases} 1 & \text{if } \kappa = 1 \\ 0 & \text{otherwise} \end{cases}. \tag{5.42}$$

Certainty-equivalence strategy

This strategy is the most wide spread one. It replaces unknown parameter in the parameterised model by a current point estimate $\hat{\Theta}_t$ of unknown parameters

$$f(\Delta_{a_\tau}|\mathcal{P}_{a_\tau}, a_\tau) \approx f(\Delta_{a_\tau}|\hat{\Theta}_t, \mathcal{P}_{a_\tau}, a_\tau),\ \tau = t,\ldots,t+T. \tag{5.43}$$

which corresponds to the approximate parameter estimate

$$f(\Theta|\mathcal{P}_{a_\tau}) \approx \hat{f}_\tau(\Theta) \equiv \delta(\Theta - \hat{\Theta}_t),\ \tau = t,\ldots,t+T. \tag{5.44}$$

where $\delta(\cdot)$ is Dirac delta function, a formal pdf of the measure concentrated on zero value.

Active strategies

All outlined strategies are *passive*. The planned decisions do not care about learning as the used $\hat{f}_\tau(\Theta)$ depend on the experience $\mathcal{P}_{a_t}$ only for all $\tau \geq t$. At the same time, it is known, see [5, 6] and Remarks 4.3.3, that the optimal strategy is *dual strategy*. It cares both about immediate decisions and learning process in a balanced way. This motivates a use of measures that support learning. The resulting strategies are known as *active strategies*. Two basic suboptimal ways are used:

- an external stimulating signal is fed into the closed loop (added to optional quantities like inputs or set points),

- a term is added to the loss original function that reflects quality of learning even when a passive-type design is applied [69].

Remarks 5.4.3

1. *The essence of suboptimal design is demonstrated mostly in combination with estimation. Similar techniques are applicable in conjunction with filtering.*

2. *Quality of approximations depends heavily on quality of the parameterised model and on parameter estimates $f(\Theta|\mathcal{P}_{a_t}, a_t)$. Thus, the care devoted to the construction of the parameterised model and the prior pdf $f(\Theta)$, see Section 4.1.4, might decide on a success/failure of the decision-making. This observation has been at roots of designing of algorithmic and software tools for the prior design of adaptive controllers [70, 71, 72].*

3. *A substantial effort with significant results has been put into analysis of various approximations. Mostly, however, the results are of an asymptotic nature. A little is known about practically significant finite time properties of such approximators.*

4. *It is fair to say that the "dangerous" certainty equivalent strategies dominate applications. They are simple, understandable and they can be directly extended to the combination of the design with filtering.*

5. *Active strategies are rarely used [3] in spite of known, theoretically proved, danger of passive strategies.*

Problem 5.4.3 (Is the back of trick a final stage?) *Still, the approximate design can be called more back of tricks than a systematic framework. Naturally, a systematic solution like the equivalence approach is desirable. The outlined understanding of suboptimal design as a complexity reduction could be useful in this respect.*

Problem 5.4.4 (Is the fully probabilistic strategy active?) *The strategy generated by the fully probabilistic design, see Proposition 2.5.4, is randomised. It can be interpreted as a combination of a deterministic feedback with a dither noise. Thus, there is a chance that it is "naturally" active. This hypothesis, however, has to be inspected in detail.*

5.5 Decomposition of decision-making

Repeatedly, splitting of the decision-making task in a chain of subtasks is declared as the way of converting optimal design to an approximation of the practical optimal design. At the same time, it is known that any violation of the golden decision-making rule (4.1) makes the solution to depart from optimality. Lack of the formal tools for the decomposition leave us with empirical rules in this area.

Here give as an example the list describing such decomposition in the case of adaptive control with learning part based on parameter estimation. Each item in the list has been found as a relatively self-containing decision sub-problem.

Whenever possible, we give a reference to the relevant support within the text. The design, as any human activity is iterative. Naturally, the majority of iterations should be concentrated in the off-line phase in order to minimise expenses related to the commission of the controller.

5.5.1 Off-line phase

The following indicative list of subtasks is solved, often with internal iterations, until the potential user is satisfied.

- Formulate the addressed problem, see Chapter 2, get the specification of
  - the technical control aims,
  - the system,
  - the available data, system inputs and outputs,
  - the technological and complexity restrictions,
  - the knowledge available.
- Perform experimental design and collect data for off-line analysis, see Section 4.1.1.
- Make data pre-processing, see Section 4.1.1.
- Select class of parameterised models, see Section 4.1.3.
- Quantify prior knowledge, see Section 4.1.4.
- Estimate model structure and control period, see Proposition 4.1.6.
- Estimate forgetting factor, see Propositions 4.1.5, 4.1.6
- Perform generalised Bayesian estimation based on prior knowledge and available data.

  These result will be used as the prior and/or alternative pdf in on-line phase, see Algorithm 4.1.2.
- Select the loss function.

  Typically, weights of quadratic loss or the ideal in fully probabilistic design, see Sections 4.1.5, 2.5.2 have to be chosen. The iterative choice runs as follows:

  Do until the results cannot be improved

  - Select suboptimal control design and its parameters.
  - Perform prior design of controller and prediction of closed loop (coupling system $\leftrightarrow$ controller) behaviour by transforming all uncertain quantities into to user-defined quality indicators.

    The formal solution provides Proposition 2.3.5with Monte Carlo evaluation employed, see Section 5.6.2. It becomes feasible using stopping rules, see Section 4.3.1.
  - Compare results with user's specification.

    Set estimation is useful to it, see Section 4.2.2.

- Validate the results.

  Again Monte Carlo techniques are quite useful.

5.5.2 On-line phase

These decision-making subtasks are solved in real time for $t \in t^*$. Here, there is almost no freedom for iterative trial-and-error solutions. Of course, it is wise to store the data collected during the on-line phase and use them for an improved off-line design.

- Collect and pre-process data, see Section 4.1.1.
- Generate reference signals.
- Estimate with stabilised forgetting, see Section 5.2.2.
- Use receding horizon or iterations-spread-in time design, see Section 5.4.1, in an appropriate version, see Section 4.3.1.
- Generate input using the designed strategy and measured data.
- Check possible discrepancies and make a finer tuning of optional parameters of the design.

  This supervision can be based on theory of operator control [73], that exploits mixture estimation, see Sections 4.1.3, 5.2.3.

Problem 5.5.1 (Completion of the design support) *The presented outline shows the complexity of the overall chain. Obviously, some steps are weakly supported by the presented theory and algorithms. Some of them are not supported at all. In spite of this, our experience indicates that the presented theory provides the most complete unified support available. Still a development of this support represents significant research challenge.*

5.6 Some numerics

The presented theory should lead to algorithms implemented in computers. Thus, their numerical realisation form a substantial part of the decision-making as an applied science. Naturally, each decision task requires its specific approach. There is, however, common or wide spread problems that are worth to be mentioned independently. This is done here.

5.6.1 Sharp likelihoods

Probability density functions are the basic object we are dealing with. The main numerical problems with them are encountered in parameter estimation and in testing of hypothesis. In these cases, they have tendency to converge to very sharp functions close to Dirac delta function. This fact follows from Proposition 3.1.1. It means that we have to work with logarithms of likelihood functions, $\ln[\mathcal{L}(\Theta, \mathcal{P}_{a_t})]$ and before converting them into posterior pdfs subtract $\max_{\Theta \in \Theta^*} \ln[\mathcal{L}(\Theta, \mathcal{P}_{a_t})]$. It is possible because the likelihood is determined uniquely up to a Θ independent factor. It is advantageous as it prevents numerical overflows and the dynamical range of the posterior pdfs causes at most numerical underflows for values for which the posterior pdf is anyway negligible.

5.6.2 Monte Carlo techniques

Monte Carlo (MC) techniques evaluate complex formulae by generating samples with a specific distribution and by forming (typically point) estimators of the results of interest. They rely on laws of large numbersthat guarantee convergence of such estimators to the target value. There is a numerous variants and problems that have to be resolve in order to make this idea work efficiently. Here, we can only touch this rich, useful and interesting area.

Generating of samples

For MC techniques samples from a pre-specified distribution have to be generated. We need them, for instance, in bootstrap-based filtering, see Section 5.3.2.

We denote $x \sim f$ if the sample x is generated according to the pdf f.

Majority of software tools, like MATLAB, provide (pseudo)random number generators. They generate independent samples uniformly distributed on interval (0,1), its pdf we denote $\mathcal{U}(0,1)$, or normally distributed with zero mean and unit variance, with the pdf $\mathcal{N}(0,1)$. Thus, we assume further on that samples of this type are at disposal. The results summarised in the following proposition serve often for solution of the task addressed.

Proposition 5.6.1 (Samples obtained from uniform and normal pdfs)

1. *Let $f(x)$ be a pdf of a univariate real variable $x \in x^* \equiv (-\infty, \infty)$ and $r \sim \mathcal{U}(0,1)$. The $x_r \sim f$ if it solves the equation*

$$r = \int_{-\infty}^{x_r} f(x)\, dx. \tag{5.45}$$

2. *Let the entries e_i of a vector e be independent and each of them $e_i \sim \mathcal{N}(0,1) \Leftrightarrow e \sim \mathcal{N}(0,I)$ where $I =$ unit matrix. Then,*

$$g \equiv \mu + Ge \sim \mathcal{N}(\mu, GG') \tag{5.46}$$

for any vector μ and matrix G of appropriate dimensions.

Proof:
ad 1. The equation (5.45) has a unique solution as the integral as a function of the upper bound is continuous non-decreasing function covering interval (0,1) due to the non-negativity and normalisation of pdf f, see Proposition 2.3.4.

Using Proposition 2.3.5 with $\alpha = x_r$, $\beta = r$ $f_T(\beta) = \mathcal{U}(0,1) \equiv 1$ on $(0,1)$, we find $|J(\alpha)| = f(\alpha)$
ad 2. Using Proposition 2.3.5, it can be verified that the considered affine transformation maps preserves normality. Transformation of moments is implied either by the same Proposition or can be simply verified using basic properties of expectation, see Proposition 2.3.6. □

Usefulness of the transformation (5.45) depends heavily on our ability to solve efficiently this equation. A bunch of other techniques has been developed to cope with (predominant) cases to which it cannot be used.

Prediction of achievable quality

This section presents an example of application of Monte Carlo technique for the prior prediction of achievable quality of the optimal strategy [70, 71, 74].

The basic scenario is as follows. For a chosen parameterised model, prior pdf $f(\Theta)$ is suppossed to be constructed. Let us assume that for each Θ we are able to find the stationary optimal strategy ($\bar{t} \to \infty$) that is not admissible as it depends on unknown parameter. Without too much loss of generality, we can assume that strategy is determined by a repetitive use of a single rule $\mathcal{R}_\Theta$. With this strategy, the achieved values of the considered stationary loss function become a function, say $Q(\Theta)$ of Θ only. These values represent lower bound on achievable quality as the set strategies working on $\mathcal{P}^*_{a_t}, \Theta^*$ is wider than the set of admінissible strategies. The admissible strategy can achieve this bound only when it successfully learn the unknown parameters.

As we are uncertain about Θ we are uncertain about values of $Q(\Theta)$. This quality indicator, however, has been computed through a very deterministic optimisation. Thus, conceptually we can get pdf of Q by using Proposition 2.3.5. Practically, the proposition cannot be used as the mapping $T : \Theta^* \to Q^*$ is non-linear, many to one and its values are mostly evaluated numerically. Here, Monte Carlo technique can directly be used. Simply, independent samples $\Theta_i \sim f(\Theta)$, $i \in i^*$ are generated and independent samples $Q_i \equiv Q(\Theta_i)$ computed in the mentioned deterministic way. Then, the desired pdf $f(Q)$ is estimated.

The estimation of the pdf $f(Q)$ is simple as Q is scalar variable.

The estimation of the pdf $f(Q)$ is complex as each sample Q result from a demanding optimisation. For this reason, all measures have to be taken in order to get results within an acceptable computing time. For instance,

- A parameterised version of $f(Q)$ is estimated [75].
- Sequential stopping rules [52], see Section 4.3.1, are employed.
- Partial characteristics of $f(Q)$, like credibility intervals, see Section 4.2.2, are searched for only.

5.6.3 Quadratic forms

Optimal designs with *q*uadratic loss functions and *G*aussian models with expected value formed by a *l*inear function of parameters and observed data are often *practically* optimal. Both learning and optimisation reduce to manipulations with positive semidefinite quadratic forms. This fact is behind the predominant use of this so called *LQG design*. Moreover, other decision problems are often approximated by an LQG one. This explains why we present the summary of basic operations within this general chapter.

Agreement 5.6.1 (Quadratic forms) *Quadratic form $\psi' Q \psi$ is the function of a real column vector $\psi \in \psi^*$, specified by a symmetric matrix $Q = Q'$. A quadratic form is said positive semidefinite if it is non-negative for any $\psi \in \psi^*$. It is called positive definite if $\psi' Q \psi > 0$, $\forall \psi \in \psi^* \setminus 0$. The matrix Q is called kernel of the form and the introduced terminology is used also for it. It means that positive (semi)definite kernel specifies positive (semi)definite quadratic form.*

Proposition 5.6.2 (Factorised kernels) *The kernel Q is positive definite iff it can be uniquely expressed (factorised) as follows*

$$Q = \underline{L}\,\underline{L}' = \underline{U}\,\underline{U}' = LDL' = U\underline{D}U', \quad \textit{where} \tag{5.47}$$

$\underline{L}$, ($\underline{U}$) is lower (upper) triangular matrix with positive diagonal entries

L, (U) is lower (upper) triangular matrix with unit diagonal entries

$D, \underline{D}$ are diagonal matrices with positive entries.

Proof: Let the first decomposition exist. Then, $\psi' Q \psi = \sum_i z_i^2 \geq 0$ with $z \equiv L'\psi$. The matrix L is regular and consequently maps non-zero ψ on non-zero z. Thus, the last inequality is sharp for $\psi \neq 0$. The similar arguments apply for the alternative decompositions.

Let Q be positive definite matrix and let us search for the first decomposition. If we write the desirable identity entry-wise and exploit the definition of $\underline{L}$ as well as symmetry of Q, we get a formal algorithm for computing $\underline{L}$

Algorithm 5.6.1 (Choleski algorithm)

$$\begin{array}{lll}
i \in i^* & = & \{1,\ldots,\bar{i}\} \\
& & \underline{L}_{ii} = \sqrt{Q_{ii} - \sum_{k \leq i-1} \underline{L}_{ik}^2} \\
& j & = \{i+1,\ldots,\bar{i}\} \\
& & \underline{L}_{ji} = \dfrac{Q_{ij} - \sum_{k \leq i-1} \underline{L}_{ik}\underline{L}_{jk}}{\underline{L}_{ii}} \\
& \text{end } j & \\
\text{end } i & &
\end{array}$$

The direct inspection shows that all quantities on right-hand side are available in time if we understand sums over empty set as zero. It remains to judge whether square-root and division make sense. If we, however, assume that it is not true at an ith step then we get the contradiction with positive definiteness by trying the vector ψ which is zero except of the unity on the ith position.

The LD factorisation is obtained from the Choleski ones by extracting diagonal entries from $\underline{L}$. Similar considerations verify the factorisations with upper triangular matrices. □

Remarks 5.6.1

1. *The expressions $\underline{L}\,\underline{L}' = \underline{U}\,\underline{U}'$ are called* Choleski factorisations *as the name of the algorithm indicate.*
2. *Algorithms providing LD or UD factorisations can be derived in the same way. They are even computationally simpler as they do not need to compute square roots. It gives them alternative name* square-root free factorisations.
3. *LD (UD) factorisations are easily extended (not uniquely) to a positive definite Q: whenever $D_{ii} = 0$, ($\underline{D}_{ii} = 0$) the corresponding non-diagonal entries in ith column (row) of L, (U) are set to zero.*

4. *The factorisation $Q = AA'$ with a general A is not unique even for positive definite Q. To see it, it is sufficient to consider any orthogonal matrix T for which $TT' = I$ =unit matrix. Then , obviously, also AT is a factor of Q if A is.*

Proposition 5.6.3 (Minimisation with factorised kernels) *Let $\psi' = [u', x']$ and $Q = LDL'$ with*

$$L = \begin{bmatrix} L_u & 0 \\ L_{xu} & L_x \end{bmatrix}, \; D = \begin{bmatrix} D_u & 0 \\ & D_x \end{bmatrix}$$

where dimensions of of L_u, D_u are compatible with the dimension of u. Then,

$$u = -(L_u')^{-1} L_{xu}' x = \arg\min_{u \in u^*} \psi' LDL' \psi'. \qquad (5.48)$$

The reached minimum is

$$\min_{u \in u^*} \psi' LDL' \psi' = x' L_x D_x L_x' x' \qquad (5.49)$$

Proof: The minimised quadratic form can be written as as a sum of squares of two weighted quadratic norms

$$\psi' LDL' \psi = ||u' L_u + x' L_{xu}||^2_{D_u} + ||x' L_x||^2_{D_x}.$$

Only the first one depends on the optional u. It reaches minimum for the zero argument which gives the unique minimiser. The second one gives the minimum reached. □

Remarks 5.6.2

1. *The minimisation of a quadratic form is very simple when the kernel is a factorised version. More importantly, it is numerically safe as the work with factors guarantees that the minimised form is always at least positive semidefinite even if L is evaluated with errors. The original Q lacks this property.*

2. *The advantageous numerical properties of factorised versions are preserved when we do not need to created the complete kernel during manipulations encountered. We indicate the basic algorithm with such a property in case when two kernels should be summed.*

   Let $Q = LDL'$ and l_i, d_i be ith column of L and ith diagonal entry of D_i respectively. Then, $Q = \sum_{i \in i^*} l_i d_i l_i'$ where $l_i' = [\underbrace{0, \ldots, 0}_{(i-1)}, 1, x, \ldots, x]$ where x denotes entries with arbitrary values. If the similar notation is used for another kernel $\tilde{Q} = \tilde{L} Q \tilde{L}'$ we have

$$Q + \tilde{Q} = \sum_{i \in i^*} \left(l_i d_i l_i' + \tilde{l}_i \tilde{d}_i \tilde{l}_i' \right) \qquad (5.50)$$

   It is straightforward to derive an algorithm for dyadic reduction *re-computing a sum of two weighted dyads with a common leading unit into sum of two dyads where the second gets zero of the leading unity. Taking into account*

(5.50), it is sufficient to take the first dyad from both decompositions, to reduce one of them by the dyadic reduction. Then the reduction is called twice on the triple of dyads having unit on the second position, etc. In this way, LD decomposition of the sum of kernels is directly obtained from factors of summed terms.

The factorised version of a *precision matrix* Q, i.e. inversion of of covariance matrix of a vector Ψ having normal pdf allows us to evaluate simply conditional and marginal pds of its subvectors.

Proposition 5.6.4 (Conditioning and marginalisation of normal pdfs) *Let* $\Psi' = [y', \psi']$ *be normally distributed with zero mean and* $Q \equiv [\mathrm{cov}(\Psi)]^{-1} = LDL'$ *with*

$$L = \begin{bmatrix} L_y & 0 \\ L_{\psi y} & L_\psi \end{bmatrix}, \ D = \begin{bmatrix} D_y & 0 \\ & D_\psi \end{bmatrix}$$

where dimensions of of L_y, D_y *are compatible with the dimension of* y. *Then,* $f(y|\psi)$ *is the normal pdf with the conditional mean and precision matrix*

$$\mathcal{E}[y|\psi] = -(L'_y)^{-1} L'_{\psi y}\psi, [\mathrm{cov}(y|\psi)]^{-1} = L_y D_y L'_y. \tag{5.51}$$

The marginal pdf $f(\psi)$ *of* ψ *is zero mean with precision matrix*

$$[\mathrm{cov}(x)]^{-1} = L_x D_x L'_x \tag{5.52}$$

Proof: It is obvious by writing the quadratic form in exponent of multivariate normal form exactly as in proof of Proposition 5.6.3 and recall chain rule for pdfs $f(u, x) = f(u|x)f(x)$. □

Remarks 5.6.3

1. *Notice that the position of the variable with respect to which minimisation or conditioning is performed is important and dependent on the factorisation used. This simple observation is often overlooked.*

Bibliography

[1] M.H. DeGroot, *Optimal Statistical Decisions*, McGraw-Hill Company, New York, 1970.

[2] J.O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York, 1985.

[3] K.J. Astrom and B. Wittenmark, *Adaptive Control*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1989.

[4] P.E. Wellstead and M.B. Zarrop, *Self-tuning Systems*, John Wiley & Sons, Chichester, 1991.

[5] A.A. Feldbaum, "Theory of dual control", *Autom. Remote Control*, vol. 21, no. 9, 1960.

[6] A.A. Feldbaum, "Theory of dual control", *Autom. Remote Control*, vol. 22, no. 2, 1960.

[7] A. Wald, *Statistical Decision Functions*, John Wiley & Sons, New York, London, 1950.

[8] T. Bohlin, *Interactive System Identification: Prospects and Pitfalls*, Springer-Verlag, Berlin, Heidelberg, New York, 1991.

[9] M. Kárný, A. Halousková, J. Böhm, R. Kulhavý, and P. Nedoma, "Design of linear quadratic adaptive control: Theory and algorithms for practice", *Kybernetika*, vol. 21, 1985, Supplement to No. 3, 4 ,5, 6.

[10] V. Peterka, "Bayesian system identification", in *Trends and Progress in System Identification*, P. Eykhoff, Ed., pp. 239–304. Pergamon Press, Oxford, 1981.

[11] R.L. Keeny and H. Raiffa, *Decisions with multiple objectives: Preferences and value tradeoffs*, J. Wiley and Sons, New York, 1978.

[12] M.M. Rao, *Measure Theory and Integration*, John Wiley and Sons, New York, Chichester, Brisbane, Toronto, Singapore, 1987, Pure and Applied Mathematics, Wiley-Interscience Series of Texts, Monographs and Tracts.

[13] V. Jarník, *Integral calculus II*, Academia, Prague, 1984, in Czech.

[14] I. Vajda, *Information theory and stadistical decision making*, Alfa, Bratislava, 1982, in Slovak.

[15] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, Champan & Hall, London, New York, 1991, ISBN 0 412 286602 (HB).

[16] G. Shafer, "Belief functions and parametric models (with discussion)", *J. Roy, Statist. Soc.*, vol. Ser. B 44, pp. 322–352, 1982.

[17] R. Bellman, *Introduction to the Mathematical Theory of Control Processes*, Academic Press, New York, 1967.

[18] M. Hendrikx, J. van Nunen, and J. Wessels, "On iterative optimization of structured markov decision processes with discounted rewards", *Math. Operationsforsch. und Statistics, ser. Optimization*, vol. 15, no. 3, pp. 439–459, 1984.

[19] S. Kullback and R. Leibler, "On information and sufficiency", *Annals of Mathematical Statistics*, vol. 22, pp. 79–87, 1951.

[20] M. Kárný, "Towards fully probabilistic control design", *Automatica*, vol. 32, no. 12, pp. 1719–1722, 1996.

[21] J. Šindelář and M. Kárný, "Dynamic decision making under uncertainty allows explicit solution", Tech. Rep. 1916, ÚTIA AVČR, POB 18, 18208 Prague 8, CR, 1997.

[22] E.F. Daum, "New exact nonlinear filters", in *Bayesian Analysis of Time Series and Dynamic Models*, J.C. Spall, Ed. Marcel Dekker, New York, 1988.

[23] M. Tanner, *Tools for statistical inference*, Springer Verlag, New York, 1993.

[24] A.F.M. Smith and A.E.Gelfand, "Bayesian statistics without tears: a smapling-resampling perspective", *The American Statistician*, vol. 46, pp. 84–88, 1992.

[25] R. Kulhavý, *Recursive Nonlinear Estimation: A Geometric Approach*, vol. 216 of *Lecture Notes in Control and Information Sciences*, Springer-Verlag, London, 1996.

[26] M. Loeve, *Probability theory*, D. van Nostrand Company Inc., Princeton, New Jersey, 1962, Russian translation, Moscow 1962.

[27] L. Berec and M. Kárný, "Identification of reality in Bayesian context", in *Preprints of the 2nd European IEEE Workshop on Computer Intensive Methods in Control and Signal Processing, CMP'96*, L. Berec, J. Rojíček, M. Kárný, and K. Warwick, Eds., pp. 211–216. ÚTIA AV ČR, Prague, 1996.

[28] I.N. Sanov, "On probability of large deviations of random variables", *Matematiceskij Sbornik*, vol. 42, pp. 11–44, 1957, in Russian, translation in Selected Translations mathematical statistics and probability, I, 1961, 213–244.

[29] H. Kushner, *Introduction to stochastic control*, Holt, Rinehart and Winston, New York, San Francisco, London, 1971.

[30] M. Zarrop, *Experiment Design for Dynamic System Identification. Lecture Notes in Control and Information Sciences 21*, Springer Verlag, Berlin-Heidelberg-New York, 1979.

[31] A.V. Oppenheim and A.S. Wilsky, *Signals and systems*, Prentice Hall, New Jersye, 1983.

[32] R.Todd Ogden, *Essential wawelets for statistical applications and data analysis*, Birkhäuser, Boston-Basel-Berlin, 1997.

[33] A. Lízr, "Linear quadratic self-tuning regulators in paper-machine control systems", in *Preprints of the 2nd IFAC Workshop on Adaptive Systems in Control and Signal Processing*, pp. 109–112. 1986.

[34] C.R. Rao, *Linear method of statistical inference and their applications*, Academia, Prague, 1987, in Czech.

[35] S. B. Jorgensen and K. M. Hangos, "Qualitative models as unifying modelling tool for grey box modelling", *International Journal of Adaptive Control and Signal Processing*, 1995.

[36] A. R. Barron and C. H. Sheu, "Approximation of density functions by sequences of exponential families", *The Annals of Statistics*, vol. 19, no. 3, pp. 1347–1369, 1991.

[37] K. Warwick and M. Kárný, *Computer-Intensive Methods in Control and Signal Processing: Curse of Dimensionality*, Birkhauser, 1997.

[38] B.O. Koopman, "On distributions admitting a sufficient statistic", *Transactions of American Mathematical Society*, vol. 39, pp. 399, 1936.

[39] M. Kárný, "Probabilistic identification and self-reproducing pdfs with application to order estimation", Tech. Rep., UTIA CSAV, Prague, 1976, in Czech, CSc thesis.

[40] M. Kárný, P. Nedoma, and J. Böhm, "On completion of probabilistic models", in *Preprints of the 2nd European IEEE Workshop on Computer Intensive Methods in Control and Signal Processing, CMP'96*, L. Berec, J. Rojíček, M. Kárný, and K. Warwick, Eds., pp. 59–64. ÚTIA AV ČR, Prague, 1996.

[41] V. Kurková, "Incremental approximation by neural networks", in *Dealing with Complexity: A Neural networks Approach*, M. Kárný, K. Warwick, and V. Kurková, Eds., pp. 177–188. Springer, London, 1997, ISBN 3 540 76160.

[42] D.M. Titterington, A.F.M. Smith, and U.E. Makov, *Statistical Analysis of Finite Mixtures*, John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1985, ISBN 0 471 90763 4.

[43] M. Kárný and P. Nedoma, "Automatic processing of prior information with application to identification of regression model", *ACASP*, 1998, submitted.

[44] A. Perez and R. Jiroušek, "Constructing an intensional expert system (INES)", in *Medical Decision Making*, pp. 307–313. Elsevier, Amsterdam, 1986.

[45] M. Kárný, J. Šindelář, and M. Valečková, "Towards Bayesian pooling of probabilistic controllers", in *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*, J. Rojíček, M. Valečková, M. Kárný, and K. Warwick, Eds., Praha, September 1998, pp. 19–24, ÚTIA AV ČR.

[46] J.M. Bernardo and A.F.M. Smith, *Bayesian theory*, John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1997, 2nd edition.

[47] M.A. Kaashoek, J.H van Schuppen, and A.C.M Ran, *Robust Control of Linear Systems and Nonlinear Control*, vol. 2, Birkhäuser, Boston, Basel Berlin, 1990, Proceedings of the International Symposium MNTS-89.

[48] P.C. Fishburn, *Utility Theory for Decision Making*, J. Wiley and Sons, New York, London, Sydney, Toronto, 1970.

[49] M. Kárný, "Algorithms for determining the model structure of a controlled system", *Kybernetika*, vol. 19, no. 2, pp. 164–178, 1983.

[50] M. Kárný and R. Kulhavý, "Structure determination of regression-type models for adaptive prediction and control", in *Bayesian Analysis of Time Series and Dynamic Models*, J.C. Spall, Ed. Marcel Dekker, New York, 1988, chapter 12.

[51] L. Berec, "Model structure identification: Global and local views", Tech. Rep., ÚTIA AVČR and Faculty of Nuclear and Physical Engineering, Czech Technical Univeristy, Prague, 1997, PhD thesis.

[52] J. Rojíček and M. Kárný, "A sequential stopping rule for extensive simulations", in *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*, J. Rojíček, M. Valečková, M. Kárný, and K. Warwick, Eds., Praha, September 1998, pp. 145–150, ÚTIA AV ČR.

[53] E. Mosca, L. Chisci, and L. Giarré, "Sidestepping the certainty equivalence in 2-DOF adaptive control via multiple implicit identifiers", in *Mutual Impact of Computing Power and Control Theory*, M. Kárný and K. Warwick, Eds., pp. 41–64. Plenum Press, New York, London, 1993.

[54] V. Peterka, "Predictor-based self-tuning control", *Automatica*, vol. 20, no. 1, pp. 39–50, 1984.

[55] M. Kárný, "Adaptive systems: Local approximators?", in *Preprints of the IFAC Workshop on Adaptive Systems in Control and Signal Processing*, Glasgow, August 1998, pp. 129–134, IFAC.

[56] R. Kulhavý, "A Bayes-closed approximation of recursive non-linear estimation", *International Journal Adaptive Control and Signal Processing*, vol. 4, pp. 271–285, 1990.

[57] R. Kulhavý, "Recursive Bayesian estimation under memory limitations", *Kybernetika*, vol. 26, pp. 1–20, 1990.

[58] R. Taylor, *Introduction to functional analysis*, Academia, Prague, 1973, Czech translation.

[59] V.S. Vladimirov, *Generalized functions in mathematical physics*, Mir Publishers, Moscow, 1979.

[60] R. Kulhavý and M. B. Zarrop, "On general concept of forgetting", *International Journal of Control*, vol. 58, no. 4, pp. 905–924, 1993.

[61] R.K. Mehra and D.G. Lainiotis (Eds.), *System identification – Advances and case studies*, Pergamon Press, New York, 1976.

[62] R. Murray-Smith and T.A. Johansen, *Multiple Model Approaches to Modelling and Control*, Taylor & Francis, London, 1997.

[63] J.S. Liu and R. Chen, "Sequential monte carlo methods for dynamic systems", *Journal of the American Statistical Association*, vol. 93, 1998.

[64] C.K. Carter and R. Kohn, "On Gibbs sampling for state space models", *Biometrika*, vol. 81, pp. 541–553, 1994.

[65] J.S. Meditch, *Stochastic Optimal Linear Estimation and Control*, Mc. Graw Hill, 1969.

[66] A.M. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[67] V. Peterka, "Adaptive digital regulation of noisy systems", in *Preprints of the 2nd IFAC Symposium on Identification and Process Parameter Estimation*, p. paper 6.2. Prague, 1970.

[68] D.W. Clarke, Mohtadi, and P.S. Tuffs, "Generalized predictive control", *Automatica*, vol. 23, no. 2, pp. 137–160, 1987.

[69] O.L.R. Jacobs and J.W. Patchell, "Caution and probing in stochastic control", *International Journal of Control*, vol. 16, pp. 189–199, 1972.

[70] M. Kárný and A. Halousková, "Designer – package for preliminary tuning of LQG adaptive control. programmer's manual (with user's guide). version 4.0", Tech. Rep. 1665, UTIA AV ČR, POBox 18, 182 08 Prague 8, Czech Republic, 1990.

[71] M. Kárný and A. Halousková, "Preliminary tuning of self-tuners", in *Lecture Notes: Advanced Methods in Adaptive Control for Industrial Application (Joint UK-CS seminar)*, K. Warwick, M. Kárný, and A. Halousková, Eds., vol. 158. Springer Verlag, 1991, held in May 1990, Prague.

[72] P. Nedoma, M. Kárný, and J. Böhm, "Software tools for use of prior knowledge in design of LQG adaptive controllers", in *Preprints of the IFAC Workshop on Adaptive Systems in Control and Signal Processing*, Glasgow, August 1998, pp. 425–429, IFAC.

[73] M. Kárný, E. Sutanto, P. Ettler, J. Kadlec, A. Quinn, and L. Tesař, "Decision support tool for complex industrial processes based on probabilistic data clustering", Tech. Rep., 1998, deliverable of Esprit project ProDac Tools, No. 25729.

[74] M. Kárný and A. Halousková, "Pre-tuning of self-tuners", in *Advances in Model-Based Predictive Control*, D. Clarke, Ed., pp. 333–343. Oxford University Press, Oxford, 1994.

[75] M. Kárný, T. Jeníček, and W. Ottenheimer, "Contribution to prior tuning of LQG selftuners", *Kybernetika*, vol. 26, no. 2, pp. 107–121, 1990.

Index