# Axiomatisation of fully probabilistic design revisited

**Article** *in* Systems & Control Letters · July 2020

DOI: 10.1016/j.sysconle.2020.104719



1 author:

Miroslav Karny
The Czech Academy of Sciences

# Axiomatisation of Fully Probabilistic Design Revisited

Miroslav Kárný

*The Czech Academy of Sciences, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, 182 08 Prague, Czech Republic, http://www.utia.cas.cz/AS, school@utia.cas.cz*

**Abstract**

Fully probabilistic design (FPD) of control strategies models both the closed control loop and control objectives by joint probabilities of involved variables. It selects the optimal strategy as the minimiser of Kullback-Leibler (KL) divergence of the closed-loop model to its ideal counterpart expressing the control objectives. Since its proposal [1] and general algorithmisation [2], FPD has been axiomatised [3] and successfully applied both theoretically [4] and practically [5, 6]. This paper refines the FPD axiomatisation and bridges FPD to standard stochastic control theory, which it encompasses, in a better way. This enhances applicability of both as well as of its popular, independently proposed, special case known as KL control [7].

*Keywords:* Closed-loop control, Control theory, Stochastic control, Stochastic modelling, Performance indices

## 1. Introduction

The paper inspects an extension of the standard stochastic control [8, 9, 10]. The standard expresses the control aims by a performance index and it takes the minimiser of its expectation as the optimal control strategy. The studied fully probabilistic design [2] of control strategies specifies the control aims via an ideal (desired) probability distribution of variables in the closed control loop and minimizes Kullback-Leibler divergence [11] of their probability distribution to the chosen ideal probability distribution.

The optimal strategy design is studied under the presentation-simplifying relaxable assumptions that the system input[1] $u_t \in \boldsymbol{u}$ is selected at discrete time $t \in \boldsymbol{t} = \{1, \ldots, |\boldsymbol{t}|\}$, $|\boldsymbol{t}| < \infty$, and the closed-loop state $x_t \in \boldsymbol{x}$ is observed. The state $x_t$ and input $u_t$ pairs form the closed-loop behaviour

$$b = (x_{|\boldsymbol{t}|}, u_{|\boldsymbol{t}|}, \ldots, x_1, u_1) \in \boldsymbol{b} = (\boldsymbol{x} \times \boldsymbol{u})^{|\boldsymbol{t}|}.$$

An uncertain response of the controlled system and randomised system inputs $(u_t)_{t\in\boldsymbol{t}}$ make the behaviour $b \in \boldsymbol{b}$ random. The joint probability density $\mathsf{c}^{\mathsf{s}}(b)$ (pd[2]) is thus the most general model of the closed loop [13][3]. The joint pd $\mathsf{c}^{\mathsf{s}}(b)$ depends on the used, generally randomised, control strategy $\mathsf{s} \in \boldsymbol{\mathsf{s}}$. The chain rule for pds [15] and the fact that $x_t$ is the state imply

$$\mathsf{c}^{\mathsf{s}}(b) = \overbrace{\prod_{t\in\boldsymbol{t}} \mathsf{m}(x_t|u_t, x_{t-1})}^{\mathsf{m}(b)} \overbrace{\prod_{t\in\boldsymbol{t}} \mathsf{s}(u_t|x_{t-1})}^{\mathsf{s}(b)} = \mathsf{m}(b)\mathsf{s}(b). \tag{1}$$

The conditional pds $\mathsf{m}(x_t|u_t, x_{t-1})$ at all time instances $t \in \boldsymbol{t}$ model the controlled system. They describe probabilities of the transition from the state $x_{t-1}$ to the states $x_t \in \boldsymbol{x}$ for the system input $u_t$. The conditional pds $\mathsf{s}(u_t|x_{t-1})$ at all time instances $t \in \boldsymbol{t}$ model the strategy. They give probabilities of using the inputs $u_t \in \boldsymbol{u}$ at the state $x_{t-1}$.

Any design chooses a strategy $\mathsf{s}^{\circ} \in \boldsymbol{\mathsf{s}}$ and takes it as optimal under the design circumstances. Stochastic control theory arrives to it as follows. It specifies a loss $\mathsf{L}(b)$ assigning a real value to each behaviour $b \in \boldsymbol{b}$. The loss is bounded from below by $\mathsf{L}(b^{\circ}) > -\infty$, where $b^{\circ} \in \boldsymbol{b}$ is the most desired behaviour. The ex post accessible value $\mathsf{L}(b) \geq \mathsf{L}(b^{\circ})$ expresses the loss attributed to the deviation of the realised behaviour $b$ from the most desired behaviour $b^{\circ}$. By definition, stochastic nature of the closed-loop makes behaviour realisations $b \in$

[1] Throughout $\boldsymbol{z}$ denotes a set of $z$'s and $|\boldsymbol{z}|$ is cardinality of $\boldsymbol{z}$. If unspecified, $\boldsymbol{z}$ is a subset of finite-dimensional real space. Mappings are distinguished by san serif fonts.

[2] Existence of this Radon-Nikodým derivative with respect to Lebesgue's or counting measure is assumed, [12]. All statements on the behaviour $b \in \boldsymbol{b}$ are valid almost everywhere.

[3] Works analysing it mathematically call it strategic measure, e.g. [14].

$\boldsymbol{b}$ dependent on the used strategy $\mathsf{s} \in \mathbf{s}$ and "something else" [16], which is inaccessible by the strategy designer. Irrespectively of the cause – randomness, uncertainty, incomplete knowledge, vagueness, etc. – the loss does not a priori order quality of strategies. Bayesian methodology [17, 18] provides the cause-indifferent counteracting of this obstacle. It selects the optimal strategy $\mathsf{s}^{\mathsf{o}}$ as the minimiser of a suitable functional $\mathsf{T}^{\mathsf{s}}$ acting on uncertain losses.

Theorem 9.3-5 in [12] represents the functional $\mathsf{T}^{\mathsf{s}}$ as an expected performance index $\mathsf{T}^{\mathsf{s}} = \mathsf{E}^{\mathsf{s}}[\mathsf{I}^{\mathsf{s}}]$. The design of the optimal control strategy facing any uncertainty then reads

$$\begin{aligned} \mathsf{s}^{\mathsf{o}}_{\mathsf{I}} &\in \operatorname{Arg}\min_{\mathsf{s}\in\mathbf{s}} \int_{\boldsymbol{b}} \mathsf{U}^{\mathsf{s}}(\mathsf{L}(b), b)\mathsf{c}^{s}(b)\ \mathrm{d}b = \operatorname{Arg}\min_{\mathsf{s}\in\mathbf{s}} \mathsf{E}^{\mathsf{s}}[\mathsf{I}^{\mathsf{s}}] \\ \mathsf{I}^{\mathsf{s}}(b) &= \mathsf{U}^{\mathsf{s}}(\mathsf{L}(b), b), \quad \mathsf{E}^{\mathsf{s}}[\mathsf{I}^{\mathsf{s}}] = \int_{\boldsymbol{b}} \mathsf{I}^{\mathsf{s}}(b)\mathsf{c}^{s}(b)\ \mathrm{d}b. \end{aligned} \tag{2}$$

$\mathsf{U}^{\mathsf{s}}$ is a non-decreasing, real-valued utility function fulfilling $\mathsf{U}^{\mathsf{s}}(0, b) = 0$. For the assumed behaviours $b \in \boldsymbol{b}$, the minimised functional in (2) guarantees that $\mathsf{s}^{\mathsf{o}}$ is Pareto optimal. The representation (2) of functionals $\mathsf{T}^{\mathsf{s}}$ ordering control strategies is universal[4] whenever

A1 $\mathsf{T}^{\mathsf{s}}$ *is* locally *linear:* $\mathsf{T}^{\mathsf{s}}\left[\mathsf{L}^{\alpha} + \mathsf{L}^{\beta}\right] = \mathsf{T}^{\mathsf{s}}\left[\mathsf{L}^{\alpha}\right] + \mathsf{T}^{\mathsf{s}}\left[\mathsf{L}^{\beta}\right]$ *for losses* $\mathsf{L}^{\alpha}, \mathsf{L}^{\beta} \in \mathbf{L}$ *meeting* $\mathsf{L}^{\alpha}\mathsf{L}^{\beta} = 0$.

A2 $\mathsf{T}^{\mathsf{s}}$ *is sequentially and boundedly continuous.*

Local linearity is significantly weaker than the usual linearity required by the standard Bayesian expected utility theory [18]. It only requires linearity on loss functions $\mathsf{L}(b)$, which are non-zero on disjunct behaviour sets $\boldsymbol{b}^{\alpha}$, $\boldsymbol{b}^{\beta}$. The continuity requirement represents no practical constraint.

Bayesian, cause-independent, handling of stochasticity and knowledge elicitation based on minimum cross-entropy principle [4, 19] provide a systematic deductive methodology [20, 21, 15] giving the controlled-system model $\mathsf{m}(b)$ (1)

[4] It means that it serves to all control tasks dealing with the same behaviour set $\boldsymbol{b}$, facing the same uncertainty but possibly differing in control objectives or sets $\mathbf{s}$ of inspected strategies.

needed for the optimal design (2). The construction of the loss $\mathsf{L}$ and utility $\mathsf{U}$, determining the performance index $\mathsf{I}^{\mathsf{s}}$ (2), is still a methodological problem. There is no universal deductive way of combining multiple behaviour attributes [22, 23] expressing desirability of behaviours $b \in \boldsymbol{b}$ by a scalar-valued performance index $\mathsf{I}^{\mathsf{s}}(b)$. Its bad choice may make the optimal strategy $\mathsf{s}^{\mathsf{o}}_{\mathsf{I}}$ (2) quite poor. The probabilistic quantification of control objectives offers deductive rules of probability theory for their combinations. Thus, the revised fully probabilistic design conceptually overcomes the lack of rules for a deductive quantification of control objectives, for the choice of the performance index $\mathsf{I}^{\mathsf{s}}$ (2).

## 2. FPD Axiomatisation

The performance indices giving the optimal strategies (2), which result into the same closed-loop model are design-equivalent. Thus, a choice of a single *ideal closed-loop model* $\mathsf{c}^{\mathsf{i}}(b)$, $b \in \boldsymbol{b}$, meeting, cf. (1), (2),

$$\mathsf{c}^{\mathsf{i}} = \mathsf{c}^{\mathsf{s}^{\mathsf{o}}_{\mathsf{I}}} \tag{3}$$

replaces the choice of equivalent performance indices. It should assign high values to desired closed-loop behaviours and small values to undesired ones.

For a chosen ideal closed-loop model $\mathsf{c}^{\mathsf{i}}$ (3), it suffices to specify any representant $\mathsf{I}^{\mathsf{s}}$ of the equivalent performance indices. The paper [3] formulated several axioms (assumptions) under which such a representant is found. Their modified, less restrictive and more intuitive, version is now presented.

A3 *Let behaviours* $b^{\alpha}, b^{\beta} \in \boldsymbol{b}$ *have equal values of the closed-loop model,* $\mathsf{c}^{\mathsf{s}}(b^{\alpha}) = \mathsf{c}^{\mathsf{s}}(b^{\beta})$*, and also the values of the loss equal,* $\mathsf{L}(b^{\alpha}) = \mathsf{L}(b^{\beta})$*. Then, the corresponding values of the performance-index equal,* $\mathsf{I}^{\mathsf{s}}(b^{\beta}) = \mathsf{I}^{\mathsf{s}}(b^{\beta})$.

A3 demands equal contributions of equally probable behaviours $b^{\alpha}, b^{\beta}$ with the equal losses $\mathsf{L}(b^{\alpha}), \mathsf{L}(b^{\beta})$ to the value of the optimised functional $\mathsf{T}^{\mathsf{s}} = \mathsf{E}^{\mathsf{s}}[\mathsf{I}^{\mathsf{s}}]$ (2). This "natural" wish is met iff the utility $\mathsf{U}^{\mathsf{s}}$ depends on the behaviour $b \in \boldsymbol{b}$ and the strategy $\mathsf{s}$ only via the values of the loss $\mathsf{L}(b)$ and the values of the joint pd describing the closed loop $\mathsf{c}^{\mathsf{s}}(b)$

$$\mathsf{I}^{\mathsf{s}}(b) = \mathsf{U}^{\mathsf{s}}(\mathsf{L}(b), b) = \tilde{\mathsf{U}}(\mathsf{L}(b), \mathsf{c}^{\mathsf{s}}(b)).$$

There, the newly-introduced utility function $\tilde{\mathsf{U}}$ preserves monotonicity of utility $\mathsf{U}$ in the values of the loss $\mathsf{L}$ and its zero value for the zero loss.

A4 *No bijective mapping* $\boldsymbol{b} \leftrightarrow \tilde{\boldsymbol{b}}$ *of behaviours changes the value* $\mathsf{T}^{\mathsf{s}}$ *assigned to a strategy* $\mathsf{s} \in \mathbf{s}$.

A4 attributes a fixed quality to each strategy $\mathsf{s} \in \mathbf{s}$ irrespectively of the coordinate system of the behaviour $b \in \boldsymbol{b}$. It is simply met when (temporarily) assuming a strictly positive ideal closed-loop model

$$\mathsf{c}^{\mathsf{i}}(b) > 0 \quad \forall b \in \boldsymbol{b}. \tag{4}$$

Under (4), the substitution formula for multivariate integrals implies that A4 is met iff the performance index

$$\mathsf{I}^{\mathsf{s}}(b) = \tilde{\mathsf{U}}^{\mathsf{s}}(\mathsf{L}(b), \mathsf{c}^{\mathsf{s}}) = \mathsf{V}(\mathsf{L}(b), \rho^{\mathsf{s}}(b)), \quad \rho^{\mathsf{s}}(b) = \frac{\mathsf{c}^{\mathsf{s}}(b)}{\mathsf{c}^{\mathsf{i}}(b)}. \tag{5}$$

There, the utility function $\mathsf{V}$ preserves monotonicity of the utility $\tilde{\mathsf{U}}$ in the values of the loss $\mathsf{L}$ and its zero value for the zero loss.

A5 *Representant* $\mathsf{I}^{\mathsf{s}}(b)$ *is in the inspected equivalence class.*

A5 is the elementary property of any class representant. Operationally, it means that the optimal strategy $\mathsf{s}^{\mathsf{o}}_{\mathsf{I}}$ (2) computed for this representant $\mathsf{I}^{\mathsf{s}}$ is to guarantee (3) for the given ideal closed-loop model $\mathsf{c}^{\mathsf{i}}$ determining the equivalence class.

**Proposition 1 (Jensen's Representant).** *Let the closed-loop ideal model* $\mathsf{c}^{\mathsf{i}}$ *meet (4) and the utility function* $\mathsf{V}$ *(5) be a function* $\mathsf{W}$ *of the ratio* $\rho^{\mathsf{s}} = \frac{\mathsf{c}^{\mathsf{s}}}{\mathsf{c}^{\mathsf{i}}}$

$$\mathsf{I}^{\mathsf{s}}(b) = \mathsf{V}(\mathsf{L}(b), \rho^{\mathsf{s}}) = \mathsf{W}(\rho^{\mathsf{s}}(b)), \tag{6}$$

*while the function* $\rho\mathsf{W}(\rho)$ *is strictly convex for* $\rho > 0$ *and the value* $\mathsf{W}(1)$ *is finite. Then, the performance index* $\mathsf{I}^{\mathsf{s}}$ *(6) meets A1–A5.*

**Proof** A5, which remains to be proved, directly follows from Jensen's inequality [12], which can be seen as the definition of convexity. Indeed, for any $\mathsf{s} \in \mathbf{s}$,

$$\begin{aligned} \mathsf{E}^{\mathsf{s}}[\mathsf{I}^{\mathsf{s}}] &= \int_{\boldsymbol{b}} \rho^{\mathsf{s}}(b)\mathsf{W}(\rho^{\mathsf{s}}(b))\mathsf{c}^{\mathsf{i}}(b)\ \mathrm{d}b \\ &\geq \underbrace{\int_{\boldsymbol{b}} \rho^{\mathsf{s}}(b)\mathsf{c}^{\mathsf{i}}(b)\ \mathrm{d}b}_{=1} \times \mathsf{W}\Bigg(\underbrace{\int_{\boldsymbol{b}} \rho^{\mathsf{s}}(b)\mathsf{c}^{\mathsf{i}}(b)\ \mathrm{d}b}_{=1}\Bigg) = \mathsf{W}(1). \end{aligned} \tag{7}$$

The left-hand side of (7) reaches the minimum iff the strategy $\mathsf{s}_{\mathsf{I}}^{\mathsf{o}}$ guarantees $\rho^{\mathsf{s}_{\mathsf{I}}^{\mathsf{o}}}(b) = 1 \Leftrightarrow \mathsf{c}^{\mathsf{s}_{\mathsf{I}}^{\mathsf{o}}} = \mathsf{c}^{\mathsf{i}}$ on $\boldsymbol{b}$. □

A6 *The optimal strategy of concatenated but independent control tasks consists of the optimal strategies obtained for the individual control tasks.*

A6 prevents the design methodology to enforce dependence into the solution of independent control problems. It singles out Kullback-Leibler divergence [11] among $I$-divergences [24] given by (6).

**Proposition 2 (FPD).** *The utility* $\mathsf{W}(\rho) = \ln(\rho)$ *(6) meets A1-A6. It defines the optimal strategy as the minimiser of KL divergence* $\mathsf{D}(\mathsf{c}^{\mathsf{s}}||\mathsf{c}^{\mathsf{i}})$

$$\mathsf{s}^{\mathsf{o}} \in \mathrm{Arg}\min_{\mathsf{s}\in\mathbf{s}} \int_{\boldsymbol{b}} \mathsf{c}^{\mathsf{s}}(b) \ln\left(\frac{\mathsf{c}^{\mathsf{s}}(b)}{\mathsf{c}^{\mathsf{i}}(b)}\right)\ \mathrm{d}b = \mathrm{Arg}\min_{\mathsf{s}\in\mathbf{s}} \mathsf{D}(\mathsf{c}^{\mathsf{s}}||\mathsf{c}^{\mathsf{i}}). \tag{8}$$

*The optimisation (8) is dubbed* fully probabilistic design of decision strategies*.*

**Proof** It remains to inspect A6. A pair of independent control problems deals with the behaviour $b = (b^{\alpha}, b^{\beta}) \in \boldsymbol{b} = \boldsymbol{b}^{\alpha} \times \boldsymbol{b}^{\beta}$, $\boldsymbol{b}^{\alpha} \cap \boldsymbol{b}^{\beta} = \emptyset$. It uses the ideal closed-loop model $\mathsf{c}^{\mathsf{i}}(b) = \mathsf{c}^{\mathsf{i}\alpha}(b^{\alpha})\mathsf{c}^{\mathsf{i}\beta}(b^{\beta})$. The optimised functional (6) on the pair $b = (b^{\alpha}, b^{\beta})$ equals the sum of individual functionals

$$\begin{aligned} 0 = &\int_{\boldsymbol{b}^{\alpha}} \int_{\boldsymbol{b}^{s}} \mathsf{c}^{\mathsf{s}\alpha}(b^{\alpha})\mathsf{c}^{\mathsf{s}\beta}(b^{\beta}) \times \\ &\big[\mathsf{W}(\rho^{\mathsf{s}\alpha}(b^{\alpha})\rho^{\mathsf{s}\beta}(b^{\beta})) - \mathsf{W}(\rho^{\mathsf{s}\alpha}(b^{\alpha})) - \mathsf{W}(\rho^{\mathsf{s}\beta}(b^{\beta}))\big]\ \mathrm{d}b^{\alpha}\mathrm{d}b^{\beta}. \end{aligned}$$

This gives the functional equation for the utility function $\mathsf{W}$ (6), which has to be met for arbitrary ratios $\rho^{\alpha}, \rho^{\beta} > 0$

$$\mathsf{W}(\rho^{\alpha}\rho^{\beta}) = \mathsf{W}(\rho^{\alpha}) + \mathsf{W}(\rho^{\alpha}).$$

It has $\mathsf{W}(\rho) = \ln(\rho)$ for $\rho > 0$ as its only smooth solution, [25]. □

## 3. Relation to Standard Stochastic Control

The dependence of the performance index $\mathsf{I}^{\mathsf{s}}$ (2) on the optimised strategy $\mathsf{s} \in \boldsymbol{\mathsf{s}}$ makes the optimised functional $\mathsf{E}^{\mathsf{s}}[\mathsf{I}^{\mathsf{s}}]$ non-linear in the opted strategy $\mathsf{s}$. The standard stochastic control design deals with $\mathsf{s}$-independent performance indices $\mathsf{I}$. For them, the minimised $\mathsf{E}^{\mathsf{s}}[\mathsf{I}] = \int_{\boldsymbol{b}} \mathsf{I}(b)\mathsf{m}(b)\mathsf{s}(b)\ \mathrm{d}b$ is linear in the opted strategy $\mathsf{s}$. Consequently, the optimal strategy $\mathsf{s}^{\mathsf{o}}_{\mathsf{I}}$ is deterministic feedback [8]. The ideal closed-loop model specified by (3) then violates (4). The following proposition addresses this discrepancy and relates FPD to the standard stochastic control. It uses the closed-loop neg-entropy $\mathsf{H}^{\mathsf{s}} = \int_{\boldsymbol{b}} \mathsf{c}^{\mathsf{s}}(b) \ln(\mathsf{c}^{\mathsf{s}}(b))\ \mathrm{d}b$.

**Proposition 3 (Stochastic Control as FPD Limit).** *Let the performance* $\mathsf{I}$ *be strategy-independent. Then, the optimal strategy minimising* $\mathsf{E}^{\mathsf{s}}[\mathsf{I}]$ *over strategies with the closed-loop neg-entropy separated from its supremum,*

$$\mathsf{H}^{\mathsf{s}} \leq h < \sup_{\mathsf{s} \in \boldsymbol{\mathsf{s}}}(\mathsf{H}^{\mathsf{s}}) = \overline{h} \leq \infty, \tag{9}$$

*coincides with the FPD-optimal strategy* $\mathsf{s}^{\mathsf{o}}_{\lambda}$ *given by the ideal closed-loop model*

$$\mathsf{c}^{\mathsf{i}}(b) = \frac{\exp[-\mathsf{I}(b)/\lambda]}{\int_{\boldsymbol{b}} \exp[-\mathsf{I}(b)/\lambda]\ \mathrm{d}b}. \tag{10}$$

*The positive scalar* $\lambda = \lambda(h)$ *converges to zero if the separating parameter* $h$ *in (9) converges to* $\overline{h}$*. The FPD-optimal strategy* $\mathsf{s}^{\mathsf{o}}_{\lambda}$ *then converges to minimiser of* $\mathsf{E}^{\mathsf{s}}[\mathsf{I}]$*, i.e. to the strategy optimal in the standard stochastic-control sense.*

**Proof** The deterministic unconstrained optimal strategy reaches the supremum of the neg-entropy. Thus, the constraint (9) is active. The corresponding unconstrained minimisation of the Kuhn-Tucker functional, given by the multiplier $\lambda = \lambda(h) > 0$, reads

$$\begin{aligned} & \mathrm{Arg}\min_{\mathsf{s} \in \boldsymbol{\mathsf{s}}} \mathsf{E}^{\mathsf{s}}[\mathsf{I} + \lambda \ln(\mathsf{c}^{\mathsf{s}})] = \mathrm{Arg}\min_{\mathsf{s} \in \boldsymbol{\mathsf{s}}} \mathsf{E}^{\mathsf{s}}[\mathsf{I}/\lambda + \ln(\mathsf{c}^{\mathsf{s}})] \\ = \quad & \mathrm{Arg}\min_{\mathsf{s} \in \boldsymbol{\mathsf{s}}} \mathsf{E}^{\mathsf{s}}[\ln(\mathsf{c}^{\mathsf{s}}/\exp(-\mathsf{I}(b)/\lambda))] = \mathrm{Arg}\min_{\mathsf{s} \in \boldsymbol{\mathsf{s}}} \mathsf{D}(\mathsf{c}^{\mathsf{s}}||\mathsf{c}^{\mathsf{s}^{\mathsf{o}}_{\lambda}}). \end{aligned}$$

The claimed convergence is then an obvious consequence of the relaxation of the constraint (9). $\square$

Proposition 3 generalises its analogy in [3]. It origins in [26], where dual-control features of FPD are studied. Their discussion, connections with the theory of rational inattention [27], simulated annealing [28], Boltzmann machine, and etc., are out of scope of this brief paper. It is important to notice that by focusing on FPD no stochastic control problem is omitted. Formula (10) relates them constructively.

## 4. Concluding Remarks

Technically, the paper refines the axiomatisation [3]. Propositions 1, 2 lead to FPD under weaker assumptions than the former version based on variational arguments mimicking [29]. Proposition 3 provides a simpler connection of FPD with the standard stochastic control than that presented in [3].

The existence of the axiomatisation allowed us to squeeze the FPD theory into a short paper without entering subtleties of preference and strategy orderings. The interested reader is referred to it [3]. Even control experts who are uninterested in subtleties of this type could care about FPD, which:

- provides a unified theory properly extending the standard stochastic control, Proposition 3;
- unifies otherwise disparate languages describing controlled systems and control objectives;
- finds minimising strategy explicitly even in general setting [2, 30], which makes approximate dynamic programming [31] simpler as only the expectation is to be approximated instead of the operation pair (expectation,minimisation) of the standard optimal stochastic control;
- allows to address hard non-Gaussian control problems [32, 33, 34];
- has approximation [29] and generalisation of minimum KL principles [19] as simple consequences [4, 35];
- puts KL control [7, 36, 37] into a wider perspective;

- feeds a proper exploration into an implementable adaptive control [26];
- transforms quantitative description of control objectives into the choice of the ideal closed-loop model: this allows to employ estimation [38] and approximation [39] to this purpose;
- reveals that any control-objectives quantification is to respect the model of the controlled system [1, 6, 40] and in adaptive context it adapts performance index [41];
- converts cooperation of simple filters or controllers (agents) into the pooling problem [42] of mutually understandable shared pds [43, 44, 45, 46];
- offers unifying framework to probabilistic control design [47], etc.

*Acknowledgements*

This research has been supported by MŠMT ČR LTC18075 and EU-COST Action CA16228.

## References

[1] M. Kárný, Towards fully probabilistic control design, Automatica 32 (12) (1996) 1719–1722.

[2] M. Kárný, T. V. Guy, Fully probabilistic control design, Syst. & Con. Lett. 55 (4) (2006) 259–265.

[3] M. Kárný, T. Kroupa, Axiomatisation of fully probabilistic design, Inf. Sci. 186 (1) (2012) 105–113.

[4] M. Kárný, T. Guy, On support of imperfect Bayesian participants, in: T. Guy, et al (Eds.), Decision Making with Imperfect Decision Makers, Vol. 28, Springer, Int. Syst. Ref. Lib., 2012, pp. 29–56.

[5] A. Quinn, P. Ettler, L. Jirsa, I. Nagy, P. Nedoma, Probabilistic advisory systems for data-intensive applications, Int. J. of Adapt. Control & Signal Proc. 17 (2) (2003) 133–148.

[6] M. Kárný, J. Böhm, T. Guy, L. Jirsa, I. Nagy, P. Nedoma, L. Tesař, Optimized Bayesian Dynamic Advising: Theory and Algorithms, Springer, 2006.

[7] P. Guan, M. Raginsky, R. Willett, Online Markov decision processes with Kullback Leibler control cost, IEEE Trans. on AC 59 (6) (2014) 1423–1438.

[8] K. Åström, Introduction to Stochastic Control, Acad. Press, NY, 1970.

[9] D. Bertsekas, Dynamic Programming and Optimal Control, Athena Sci., 2001.

[10] P. R. Kumar, P. Varaiya, Stochastic Systems: Estimation, Identification, and Adaptive Control, Vol. 75 of Classics of Applied Math., SIAM, 2015.

[11] S. Kullback, R. Leibler, On information and sufficiency, An. of Math. Stat. 22 (1951) 79–87.

[12] M. Rao, Measure Theory and Integration, John Wiley, NY, 1987.

[13] M. Ullrich, Optimum control of some stochastic systems, in: Preprints of the VIII-th conference ETAN, Beograd, 1964.

[14] S. Yüksel, N. Saldi, Strategic measures approach to decentralized stochastic control, in: 2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2016, pp. 838–843.

[15] V. Peterka, Bayesian system identification, in: P. Eykhoff (Ed.), Trends and Progress in System Identification, Perg. Press, 1981, pp. 239–304.

[16] D. Lindley, Understanding Uncertainty, Wiley Interscience, 2006.

[17] A. Feldbaum, Theory of dual control, Autom. Remote Control 21-22 (9-2).

[18] L. Savage, Foundations of Statistics, Wiley, 1954.

[19] J. Shore, R. Johnson, Axiomatic derivation of the principle of maximum entropy & the principle of minimum cross-entropy, IEEE Tran. on Inf. Th. 26 (1) (1980) 26–37.

[20] J. Berger, Statistical Decision Theory and Bayesian Analysis, Springer, 1985.

[21] A. Jazwinski, Stochastic Processes and Filtering Theory, Ac. Press, 1970.

[22] L. Chen, P. Pearl, Survey of preference elicitation methods, Tech. Rep. IC/2004/67, HCI Group Ecole Politechnique Federale de Lausanne, Switzerland (2004).

[23] R. Keeny, H. Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs, J. Wiley, NY, 1978.

[24] I. Csiszar, I-divergence geometry of probability distribution and minimization problems, Annals of Probability 3 (1975) 146 – 158.

[25] J. Aczel, J. Dhombres, Functional equations in several variables, Cambridge University Press, 1989.

[26] M. Kárný, F. Hůla, Balancing exploitation and exploration via fully probabilistic design of decision policies, in: Proc. of the 11th Int. Conf. on Agents and Artificial Intelligence: ICAART, Vol. 2, 2019, pp. 857–864.

[27] C. Sims, Implications of rational inattention, J. of Monetary Economics 50 (3) (2003) 665 – 690.

[28] M. Tanner, Tools for statistical inference, Springer Verlag, NY, 1993.

[29] J. Bernardo, Expected information as expected utility, The An. of Stat. 7 (3) (1979) 686–690.

[30] J. Šindelář, I. Vajda, M. Kárný, Stochastic control optimal in the Kullback sense, Kybernetika 44 (1) (2008) 53–60.

[31] J. Si, A. Barto, W. Powell, D. Wunsch (Eds.), Handbook of Learning and Approximate Dynamic Programming, Wiley-IEEE Press, 2004.

[32] A. Bušič, S. Meyn, Action-constrained Markov decision processes with Kullback-Leibler cost, in: S. Bubeck, V. Perchet, P. Rigollet (Eds.), Proc. of Machine Learning Research, Vol. 75, 2018, pp. 1–14.

[33] A. Lebeda, J. Leth, Control synthesis using the fully probabilistic design, in: IEEE Conference on CACSD, 2016. `doi:10.1109/CACSD.2016.7602561`.

[34] L. Yao, J. Qin, H. Wang, B. Jiang, Design of new fault diagnosis and fault tolerant control scheme for non-Gaussian singular stochastic distribution systems, Automatica 48 (9) (2012) 2305–2313.

[35] A. Quinn, M. Kárný, T. Guy, Fully probabilistic design of hierarchical Bayesian models, Inf. Sci. 369 (2016) 532–547.

[36] H. Kappen, Linear theory for control of nonlinear stochastic systems, Physical review letters 95 (20) (2005) 200201.

[37] E. Todorov, Linearly-solvable Markov decision problems, in: B. Schölkopf, et al (Eds.), Adv. in Neur. Inf. Proc., MIT Press, 2006, pp. 1369 – 1376.

[38] M. Kárný, T. Guy, Preference elicitation in fully probabilistic design of decision strategies, in: Proc. of the 49th IEEE CDC, 2010.

[39] M. Kárný, Automated preference elicitation for decision making, in: T. Guy, et al (Eds.), Decision Making and Imperfection, Vol. 474, Springer, Berlin, 2013, pp. 65–99.

[40] M. Kárný, T. Guy, Preference elicitation within framework of fully probabilistic design of decision strategies, in: IFAC Int. Workshop on Adaptive and Learning Control Systems - ALCOS 2019, Vol. 52, 2019, pp. 239–244.

[41] K. Belda, Probabilistic design and tuning of LQ control, Archives of Control Sciences 19 (4 (2009)) (2009) 355–368.

[42] C. Genest, J. Zidek, Combining probability distributions: A critique and annotated bibliography, Stat. Sci. 1 (1) (1986) 114–148.

[43] S. Azizi, A. Quinn, Hierarchical fully probabilistic design for deliberator-based merging in multiple participant systems, IEEE Tran. on SMC PP (99) (2016) 1–9. doi:10.1109/TSMC.2016.2608662.

[44] C. Foley, A. Quinn, Fully probabilistic design for knowledge transfer in a pair of Kalman filters, IEEE Signal Proc. Letters 25 (4) (2018) 487–490.

[45] M. Kárný, R. Herzallah, Scalable harmonization of complex networks with local adaptive controllers, IEEE Trans. on SMC: Systems 47 (3) (2017) 394–404.

[46] A. Quinn, M. Kárný, T. Guy, Optimal design of priors constrained by external predictors, J. of Appr. Reason. 84 (2017) 150–158.

[47] G. Calafiore, F. Dabbeneb, R. Tempo, Research on probabilistic methods for control system design, Automatica 47 (2011) 1279 – 1293.