2020 European Control Conference (ECC)
May 12-15, 2020. Saint Petersburg, Russia

Minimum Expected Relative Entropy Principle

Miroslav Kárný

***Abstract*— Stochastic filtering estimates a time-varying (multivariate) parameter (a hidden variable) from noisy observations. It needs both observation and parameter evolution models. The latter is often missing or makes the estimation too complex. Then, the axiomatic minimum relative entropy (MRE) principle completes the posterior probability density (pd) of the parameter. The MRE principle recommends to modify a prior guess of the constructed pd to the smallest extent enforced by new observations. The MRE principle does not deal with a generic uncertain prior guess. Such uncertainty arises, for instance, when the MRE principle is used recursively. The paper fills this gap. The proposed minimum expected relative entropy (MeRE) principle: (a) makes Bayesian estimation less sensitive to the choice of the prior pd; (b) provides a stabilised parameter tracking with a data-dependent forgetting that copes with abrupt parameter changes; (c) applies in all cases exploiting MRE, for instance, in stochastic modelling.**

Keywords minimum relative entropy principle; uncertain prior probability; forgetting; fully probabilistic design; abrupt parameter changes;

I. Introduction

Stochastic filtering [1] evaluates posterior probability, here given by probability density (pd). It is the sufficient statistic for estimation of an unobserved, time-varying parameter. Filtering exploits data while using observation and parameter evolution models. The former model is always needed but filtering complexity often discards the latter one. Then, the partially known posterior pd is to be completed. The axiomatic [2], often discussed [3], and used [4]–[6], minimum relative entropy (MRE) principle serves to this purpose. The completed pd is the minimiser of the relative entropy (RE, aka Kullback-Leibler divergence, [7], cross entropy [2], [8], etc.) to its prior guess. The minimisation runs over pds compatible with the acquired knowledge.

The result of the MRE principle use depends on the prior guess, which is often uncertain. For instance, if its outcome serves as the prior guess in the next step of recursive estimation. Such a guess can hardly be taken as certain [9]. This case has motivated the paper.

The paper extends the MRE principle to the minimum *expected* RE (MeRE[1]) principle. MeRE copes with the uncertainty of the prior guess. The extension mimics the derivation of the MRE principle in [10]. The derivation relies on the axiomatic fully probabilistic design (FPD) [11] and it is believed to be simpler than a surely possible extension of the way in [2].

Section II recalls FPD. Section III derives MeRE principle. Section IV illustrates its use on parameter estimation with uncertain prior pd and on tracking of a varying parameter. The former case makes the estimation robust with respect to the choice of the prior pd. It opens an efficient way to the estimation of the model structure. The latter case improves the stabilised forgetting [12], which robustly tracks the varying parameter of a parametric model[2]. The ability to cope with abrupt changes is an immediate contribution brought. Section V comments the results and open problems.

II. Static FPD

FPD is a prescriptive framework for decision making (DM) under uncertainty. FPD generalises the standard Bayesian DM [13], [14]. Its static version suffices to the solution of the treated problem.

The static DM uses the knowledge[3] $k \in \{k\}$ available for choosing of an action $a \in \{a\}$. The action influences the considered but unavailable ignorance $g \in \{g\}$. Practically, any part of the (closed-loop) behaviour $b = (g, a, k) \in \{b\} = \{g\} \times \{a\} \times \{k\}$ can be uncertain. The FPD axiomatisation [11] shows that, for such a DM, the behaviour $b \in \{b\}$ is to be modelled as the multivariate random variable. Its joint pd $\mathsf{j}(b)$ factorises [15]

$$\mathsf{j}(b) = \mathsf{j}(g, a, k) = \mathsf{m}(g|a, k)\mathsf{r}(a|k)\mathsf{k}(k).$$

The individual factors have specific meanings motivating their mnemonic names:

- m models the environment relevant to the solved DM task; it relates the action $a \in \{a\}$ and the knowledge $k \in \{k\}$ to the ignorance $g \in \{g\}$;
- r is the *optional* randomised DM rule; it describes the random choice of the action $a \in \{a\}$ under the given knowledge $k \in \{k\}$;
- k describes the knowledge $k \in \{k\}$; the static DM, which selects single (multivariate) action, does not needs this model.

The action choice is purposeful. FPD expresses DM aims by an *ideal joint pd* $\mathsf{j}_{\mathfrak{i}}(b)$ of behaviours $b \in \{b\}$. The ideal joint pd assigns high values to the desired

[1] The abbreviation stresses the added "expectation" and differentiates itself from MERE used for something else.

[2] It coincides with the observation model of the stochastic filtering.

[3] $\{k\}$ denotes a set of ks. The sets are defined only when needed. $\mathsf{San\ serif}$ fonts mark mappings, predominantly pds. $\mathcal{C}$aligraphic fonts mark functionals. Mathfrak indices $\mathfrak{o}$, $\mathfrak{i}$, and $\mathfrak{p}$ concern optimal, ideal and prior objects, respectively.

Copyright ©2020 EUCA

behaviours and small ones to undesired $b \in \{b\}$. FPD takes the minimiser of the relative entropy $\mathcal{R}(\mathsf{j}||\mathsf{j}_\mathfrak{i})$ as the optimal DM rule

$$\mathsf{r}_\mathfrak{o} \in \text{Arg} \min_{\mathsf{r}\in\{\mathsf{r}\}} \mathcal{R}(\mathsf{j}||\mathsf{j}_\mathfrak{i}) = \text{Arg} \min_{\mathsf{r}\in\{\mathsf{r}\}} \int_{\{b\}} \mathsf{j}(b) \ln\left(\frac{\mathsf{j}(b)}{\mathsf{j}_\mathfrak{i}(b)}\right) \mathrm{d}b.$$

The optimal DM rule $\mathsf{r}_\mathfrak{o}$ can be found explicitly, [16], [17], using: (a) Fubini theorem on multiple integration [18]; (b) the facts that $\mathcal{R}(\mathsf{f}||\mathsf{g}) \geq 0$ while $\mathcal{R}(\mathsf{f}||\mathsf{g}) = 0$ iff $\mathsf{f} = \mathsf{g}$; and (c) the factorisation of the ideal joint pd $\mathsf{j}_\mathfrak{i}(b) = \mathsf{m}_\mathfrak{i}(g|a,k)\mathsf{r}_\mathfrak{i}(a|k)\mathsf{k}_\mathfrak{i}(k)$. The pd $\mathsf{m}_\mathfrak{i}$ is called the ideal environment model and $\mathsf{r}_\mathfrak{i}$ is the ideal DM rule. It holds

$$\mathsf{r}_\mathfrak{o}(a|k) \propto \mathsf{r}_\mathfrak{i}(a|k) \exp[-\mathcal{R}(\mathsf{m}||\mathsf{m}_\mathfrak{i})]. \tag{1}$$

$\mathcal{R}(\mathsf{m}||\mathsf{m}_\mathfrak{i})$ in (1) is evaluated for the action a and the knowledge k in $\mathsf{r}_\mathfrak{o}(a|k)$ conditioning the pds $\mathsf{m}(g|a,k)$, $\mathsf{m}_\mathfrak{i}(g|a,k)$.

The solution of the inspected problem chooses its ideal DM rule in the leave-to-the-fate way (LTF, [17]). LTF suits whenever no demand is put on the optimal DM rule (except its domain). LTF identifies the ideal DM rule with the optimised DM rule, $\mathsf{r}_\mathfrak{i} = \mathsf{r}$. Under LTF, the FPD-optimal DM rule $\mathsf{r}_\mathfrak{o}$ selects the optimal action $a_\mathfrak{o}$ [17]

$$a_\mathfrak{o} \in \text{Arg} \min_{a\in\{a\}} \mathcal{R}(\mathsf{m}||\mathsf{m}_\mathfrak{i}). \tag{2}$$

Again, the conditional RE is used as in (1). The choice (2) reveals that FPD under LTF reduces to the Bayesian DM [13] with the loss function $\ln(\mathsf{m}/\mathsf{m}_\mathfrak{i})$.

III. Minimum Expected RE Principle

The addressed completion of a partially described pd is a meta-DM problem. It chooses a "best" completion among considered ones. This meta-DM problem is here formulated and solved as an FPD task. Its constituents are denoted by capital counterparts of the DM symbols introduced in Section II.

The knowledge K of the meta-DM task, used for the action A choice, is:

- an underlying DM task to which the completion serves; the underlying DM task operates on behaviours $b \in \{b\}$, modelled by a pd j;
- the known set $\{\mathsf{j}\}$ has the cardinality $|\{\mathsf{j}\}| > 1$, i.e. the joint pd $\mathsf{j} \in \{\mathsf{j}\}$ of behaviours $b \in \{b\}$ is incompletely known;
- the set $\{\mathsf{j}_\mathfrak{p}\}_{\mathfrak{p}\in\{\mathfrak{p}\}}$ of prior guesses $\mathsf{j}_\mathfrak{p}$ of the pd j is available: the set $\{\mathsf{j}_\mathfrak{p}\}_{\mathfrak{p}\in\{\mathfrak{p}\}}$ *contains more members*, $|\{\mathfrak{p}\}| > 1$; this assumption differentiates the solved problem from its predecessors [2], [10]; the set $\{\mathfrak{p}\}$ of pointers to pds in $\{\mathsf{j}_\mathfrak{p}\}$ is at most countable;
- a pd q on $\{\mathfrak{p}\}$ is given; $\mathsf{q}(\mathsf{j}_\mathfrak{p}) = \mathsf{q}(\mathfrak{p})$ quantifies the belief that $\mathsf{j}_\mathfrak{p} \in \{\mathsf{j}_\mathfrak{p}\}$ is the best prior model of $b \in \{b\}$.

The opted action A of the meta-DM is a pd in $\{\mathsf{j}\}$. The chosen join pd serves as a complete description of the uncertain behaviour $b \in \{b\}$. The action is generated by a DM rule selected among DM rules acting on the joint pds $\{\mathsf{j}\}$ under the given K. The FPD-optimal DM rule is searched for.

The ignorance G of the meta-DM reads

$$G = (b, \mathsf{j}, \mathsf{j}_\mathfrak{p}) \in \{G\} = \{b\} \times \{\mathsf{j}\} \times \{\mathsf{j}_\mathfrak{p}\} \tag{3}$$

as neither behaviour b, nor its complete description $\mathsf{j}(b)$ nor the prior guess $\mathsf{j}_\mathfrak{p}(b)$ are known when choosing $A \in \{\mathsf{j}\}$.

The LTF specification of the ideal DM rule is adopted. It leads to the deterministic FPD-optimal DM rule providing the optimal action, cf. (2),

$$j_\mathfrak{o} = A_\mathfrak{o} \in \text{Arg} \min_{A\in\{\mathsf{j}\}} \mathcal{R}(\mathsf{M}||\mathsf{M}_\mathfrak{i}). \tag{4}$$

The evaluation of the optimal completion (4) needs the model $\mathsf{M}(G|A,K)$ and its ideal counterpart $\mathsf{M}_\mathfrak{i}(G|A,K)$. The used description of M, $\mathsf{M}_\mathfrak{i}$ and the next formal operations assume that the cardinality of $\{b\}$ is finite. This:

- allows us to avoid the measure theory as the joint pds in $\{\mathsf{j}\}$ $\{\mathsf{j}_\mathfrak{p}\}_{\mathfrak{p}\in\{\mathfrak{p}\}}$ are probabilistic vectors;
- imposes no practical constraint as any DM with numerically expressed preferences is to operate on separable spaces [19];
- leads to the result valid without this assumption.

The next model M *unambiguously* respects the formulated meta-DM problem

$$\begin{aligned}
\mathsf{M}(G|A,K) &= \mathsf{M}(b,\mathsf{j},\mathsf{j}_\mathfrak{p}|A,K) \\
&= \mathsf{M}(b|\mathsf{j},\mathsf{j}_\mathfrak{p},A,K)\mathsf{M}(\mathsf{j}|\mathsf{j}_\mathfrak{p},A,K)\mathsf{M}(\mathsf{j}_\mathfrak{p}|A,K) \\
&= \mathsf{j}(b)\delta(\mathsf{j}-A)\mathsf{q}(\mathsf{j}_\mathfrak{p}).
\end{aligned} \tag{5}$$

In (5), the first equality uses the ignorance definition (3). The second one is just the chain rule for pds [15]. The last key equality respects that:

- j models behaviours $b \in \{b\}$;
- A uniquely selects j; it is formalised by Dirac's symbol δ, a pd of the measure concentrated on zero, and
- q quantifies the prior, A independent, belief in the prior guesses $\{\mathsf{j}_\mathfrak{p}\}_{\mathfrak{p}\in\{\mathfrak{p}\}}$ of the best model of $b \in \{b\}$.

The ideal model $\mathsf{M}_\mathfrak{i}$ expressing the meta-DM aim is similar

$$\mathsf{M}_\mathfrak{i}(G|A,K) = \mathsf{j}_\mathfrak{p}(b)\delta(\mathsf{j}-A)\mathsf{q}(\mathsf{j}_\mathfrak{p}). \tag{6}$$

The only difference is the use of $\mathsf{j}_\mathfrak{p}$ as the model of $b \in \{b\}$. $\mathsf{M}_\mathfrak{i}$ says that the selected $A_\mathfrak{o} = \mathsf{j}_\mathfrak{o} \in \{\mathsf{j}\}$ should deviate from the pd $\mathsf{j}_\mathfrak{p}$ only if the used knowledge demands it.

The minimised RE (4) with models (5), (6) reads

$$\begin{aligned}
&\mathcal{R}(\mathsf{M}||\mathsf{M}_\mathfrak{i}) \\
&= \int_{\{\mathsf{j}_\mathfrak{p}\}} \mathsf{q}(\mathsf{j}_\mathfrak{p}) \left\{ \int_{\{\mathsf{j}\}} \delta(\mathsf{j}-A) \left[\int_{\{b\}} \mathsf{j}(b) \ln\left(\frac{\mathsf{j}(b)}{\mathsf{j}_\mathfrak{p}(b)}\right) \mathrm{d}b \right] \mathrm{d}\mathsf{j} \right\} \mathrm{d}\mathsf{j}_\mathfrak{p} \\
&= \sum_{\mathfrak{p}\in\{\mathfrak{p}\}} \mathsf{q}(\mathfrak{p})\mathcal{R}(A||\mathsf{j}_\mathfrak{p}) = \mathcal{E}[\mathcal{R}(A||\mathsf{j}_\mathfrak{p})].
\end{aligned}$$

The expectation $\mathcal{E}$ is taken over the uncertain $\mathsf{j}_\mathfrak{p} \in \boldsymbol{\{}\mathsf{j}_\mathfrak{p}\boldsymbol{\}}$.

Thus, the optimal choice $A_\mathfrak{o} = \mathsf{j}_\mathfrak{o}$ (4) for (5) and (6) specialises to the *minimum expected relative entropy (MeRE) principle*

$$\mathsf{j}_\mathfrak{o} \in \mathrm{Arg}\min_{\mathsf{j}\in\boldsymbol{\{}\mathsf{j}\boldsymbol{\}}} \mathcal{E}[\mathcal{R}(\mathsf{j}||\mathsf{j}_\mathfrak{p})]. \tag{7}$$

A simple algebra maps (7) to the form[4]

$$\mathsf{j}_\mathfrak{o} \in \mathrm{Arg}\min_{\mathsf{j}\in\boldsymbol{\{}\mathsf{j}\boldsymbol{\}}} \mathcal{R}(\mathsf{j}||\mathsf{j}_{\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}}) \ \text{ with } \ \mathsf{j}_{\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}} \propto \prod_{\mathfrak{p}\in\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}} \mathsf{j}_\mathfrak{p}^{\mathsf{q}(\mathfrak{p})}. \tag{8}$$

Thus, the *MeRE principle is the MRE principle that uses the weighted geometric mean (aka logarithmic pool [20]) of the prior guesses as the prior guess of the constructed pd.*

This simple result could be expected. However, the use of the weighted arithmetic mean (aka linear pool [20]) as the representative of prior guesses also sound as a good solution. The fact that the result follows from the axiomatic FPD under explicit applicability conditions makes the recommendation (8) well grounded.

IV. Use of MeRE Principle in Estimation

This section shows the use of the MeRE principle (8).

The inspected cases support underlying DM tasks, which operate on the ignorance $g \in \boldsymbol{\{}g\boldsymbol{\}}$ made of an unused observation $o \in \boldsymbol{\{}o\boldsymbol{\}}$ and of an unknown parameter $\theta \in \boldsymbol{\{}\theta\boldsymbol{\}}$, $g = (o, \theta) \in \boldsymbol{\{}g\boldsymbol{\}} = \boldsymbol{\{}o\boldsymbol{\}} \times \boldsymbol{\{}\theta\boldsymbol{\}}$. Both o and θ can be multivariate. The knowledge k consists of a prior knowledge enriched by a collection $\{o, a\}$ of observations o made after applying actions a. The action $a \in \boldsymbol{\{}a\boldsymbol{\}}$ is generated by an informationally causal, randomised DM rule $\mathsf{r} \in \boldsymbol{\{}\mathsf{r}\boldsymbol{\}}$. The parameter $\theta \in \boldsymbol{\{}\theta\boldsymbol{\}}$ is unknown to it, cf. natural conditions of control in [15],

$$\mathsf{r}(a|\theta, k) = \mathsf{r}(a|k) \Leftrightarrow \mathsf{p}(\theta|a, k) = \mathsf{p}(\theta|k). \tag{9}$$

There, the posterior pd $\mathsf{p}(\theta|k)$ models $\theta \in \boldsymbol{\{}\theta\boldsymbol{\}}$.

The joint pd $\mathsf{j}(b)$ of $b \in \boldsymbol{\{}b\boldsymbol{\}} = \boldsymbol{\{}o\boldsymbol{\}} \times \boldsymbol{\{}\theta\boldsymbol{\}} \times \boldsymbol{\{}a\boldsymbol{\}} \times \boldsymbol{\{}k\boldsymbol{\}}$ is

$$\mathsf{j}(o, \theta, a, k) = \mathsf{m}(o|\theta, a, k)\mathsf{p}(\theta|k)\mathsf{r}(a|k)\mathsf{k}(k). \tag{10}$$

The known parametric model m relates observations $o \in \boldsymbol{\{}o\boldsymbol{\}}$ to an unknown parameter $\theta \in \boldsymbol{\{}\theta\boldsymbol{\}}$, to an action $a \in \boldsymbol{\{}a\boldsymbol{\}}$ and to the knowledge $k \in \boldsymbol{\{}k\boldsymbol{\}}$ modelled by $\mathsf{k}(k)$.

A. Uncertain Prior Pd

This case deals with an uncertain prior pd $\mathsf{p} = (\mathsf{p}(\theta))_{\theta\in\boldsymbol{\{}\theta\boldsymbol{\}}}$ of the unknown parameter $\theta \in \boldsymbol{\{}\theta\boldsymbol{\}}$. The a priori considered instances of $\mathsf{p} \in \{\mathsf{p}_\mathfrak{p}\}_{\mathfrak{p}\in\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}}$ with $|\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}| > 1$ are qualified by prior beliefs[5] q on $\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}$. Obviously, the posterior pd $\mathsf{p}(\theta|k)$ is uncertain, too, even when the pds m, r and k in (10) are assumed to be given.

This choice of the joint pd j represents frequent cases in which its multiple prior guesses $\mathsf{j}_\mathfrak{p} \in \boldsymbol{\{}\mathsf{j}_\mathfrak{p}\boldsymbol{\}}$ differ only in a marginal pd, here, operating on the parameter $\theta \in \boldsymbol{\{}\theta\boldsymbol{\}}$,

$$\boldsymbol{\{}\mathsf{j}_\mathfrak{p}\boldsymbol{\}} = \{\mathsf{j}_\mathfrak{p}(b) = \mathsf{m}(o|a, \theta, k)\mathsf{r}(a|k)\mathsf{p}_\mathfrak{p}(\theta|k)\mathsf{k}(k)\}_{\mathfrak{p}\in\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}}. \tag{11}$$

The product mrk of pds is the common factor of the prior guesses in (11). The MeRE principle (8) needs

$$\mathsf{p}_{\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}}(\theta|k) \propto \prod_{\mathfrak{p}\in\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}} \mathsf{p}_\mathfrak{p}^{\mathsf{q}(\mathfrak{p}|k)}(\theta|k). \tag{12}$$

Bayes' rule provides

$$\mathsf{p}_\mathfrak{p}(\theta|k) = \frac{\mathsf{l}(\theta|k)\mathsf{p}_\mathfrak{p}(\theta)}{\mathsf{f}_\mathfrak{p}}, \ \mathfrak{p} \in \boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}, \text{ where} \tag{13}$$

$\mathsf{l}(\theta|k)$ is the likelihood, i.e. the product of parametric models[6] with the inserted collection of the observed $\{o, a\}$ enriching the prior knowledge;

$\mathsf{f}_\mathfrak{p}$ is the normalising factor, called evidence; it is the likelihood $\mathsf{l}(\theta|k)$ multiplied by the prior pd $\mathsf{p}_\mathfrak{p}(\theta)$ with $\theta \in \boldsymbol{\{}\theta\boldsymbol{\}}$ integrated out.

Bayes' rule also updates prior beliefs $\mathsf{q}(\mathfrak{p})$ to the posterior beliefs $\mathsf{q}(\mathfrak{p}|k)$, $\mathfrak{p} \in \boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}$. It only uses the evidences

$$\mathsf{q}(\mathfrak{p}|k) \propto \mathsf{f}_\mathfrak{p}\mathsf{q}(\mathfrak{p}), \ \mathfrak{p} \in \boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}. \tag{14}$$

The final outcome of the MeRE principle simply reads

$$\mathsf{j}_\mathfrak{o}(b) \propto \mathsf{m}(o|\theta, a, k)\mathsf{r}(a|k) \prod_{\mathfrak{p}\in\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}} \mathsf{p}_\mathfrak{p}^{\mathsf{q}(\mathfrak{p}|k)}(\theta|k)\mathsf{k}(k). \tag{15}$$

It is (10) for $\mathsf{p}(\theta|k) = \mathsf{p}_{\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}}(\theta|k) \propto \mathsf{l}(\theta) \prod_{\mathfrak{p}\in\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}} \mathsf{p}_\mathfrak{p}^{\mathsf{q}(\mathfrak{p}|k)}(\theta)$. Formally, it assigns the posterior belief $\mathsf{q}(\mathfrak{p}|k)$ to prior pds $\mathsf{p}_\mathfrak{p}(\theta)$.

Commentary

- The posterior beliefs $\mathsf{q}(\mathfrak{p}|k)$ (14) into the uncertain posterior pds $\mathsf{p}_\mathfrak{p}(\theta|k)$ reflect the predictive abilities $\mathsf{f}_\mathfrak{p}$ induced by the respective prior pds $\mathsf{p}_\mathfrak{p}(\theta)$ on the observed collection $\{o, a\}$.
- It is practically important that the likelihood $\mathsf{l}(\theta|k)$ (13) is evaluated just once when computing $\mathsf{j}_\mathfrak{o}$ (15).
- The pd $\mathsf{q}(\mathfrak{p}|k)$ asymptotically points to a single prior $\mathsf{p}_\mathfrak{p}(\theta)$ (and thus posterior $\mathsf{p}_\mathfrak{p}(\theta|k)$) pd if the observed collection $\{o, a\}$ is informative and the model identifiable. This statement reflects Sanov's type analysis [21], [22] extended to closed DM loops in [23].
- Altogether, the use of multiple prior pds $\boldsymbol{\{}\mathsf{p}_\mathfrak{p}\boldsymbol{\}}$ is computationally cheap whenever evidences $\boldsymbol{\{}\mathsf{f}_\mathfrak{p}\boldsymbol{\}}$ are cheaply gained. This makes the Bayesian estimation robust with respect to a bad choice of a candidate prior pd. It cannot help if all prior pds are bad.
- The simple result (15) is extremely useful in structure estimation tasks, i.e. estimation of model order, time delay, importance of regressors etc. It suffices to exploit that the likelihood evaluated for the most

[4] $\propto$ denotes proportionality. The pd $\mathsf{j}_{\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}}$ represents $\boldsymbol{\{}\mathsf{j}_\mathfrak{p}\boldsymbol{\}}_{\mathfrak{p}\in\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}}$.

[5] The choice based on Laplace's insufficient reasons is often used. It recommends uniform $\mathsf{q}(\mathfrak{p}) = \mathsf{u}(\mathfrak{p}) = \mathit{constant}$ on $\boldsymbol{\{}\mathfrak{p}\boldsymbol{\}}$.

[6] The product of DM rules does not enter it due to (9).

rich structure often cheaply provides likelihoods of embedded models. Then, the structure estimation becomes a simple search for maxima [24].

- Let us stress that the combination of priors pds belongs to repeatedly addressed problem of the pd pooling [20]. Weighted geometric and arithmetic means (aka logarithmic and linear pools) are the main "pooling competitors". The geometric one has appeared as preferable in the MeRE scenario.
- The support of $\mathsf{p}_{\{\mathfrak{p}\}}$ (12) lies in the intersection of supports of pds $\{\mathsf{p}_\mathfrak{p}\}$. This property of logarithmic pooling makes no real harm in the discussed case. It suffices to replace the prior pds $\mathsf{p}_\mathfrak{p}$ by $\alpha\mathsf{p}_\mathfrak{p} + (1-\alpha)\underline{\mathsf{p}}$. There α approaches 1 from below and with an optional "background" $\underline{\mathsf{p}}$ has the support covering the union of supports of pds in $\{\mathsf{p}_\mathfrak{p}\}$. This avoids the troubles with the support of $\mathsf{p}_{\{\mathfrak{p}\}}$. It (practically) preserves the evidences $\mathsf{f}_\mathfrak{p}$, which decide on influence of the pds in $\{\mathsf{p}_\mathfrak{p}\}$ on the final outcome.

B. Robust Tracking of Varying Parameter

This case deals the unknown parameter θ of the parametric model $\mathsf{m}(o|\theta, a, k)$ (10), which varies during estimation while the model of variations is unavailable. This old problem is typically, and often successfully, solved by forgetting [12]. Multitude variants exist motivated by differing adopted methodologies as well as by requirements on the acceptable solution [25]. MeRE offers a prescriptive methodology. A specific version is obtained by the choice of the set of the prior guesses $\{\mathsf{p}_\mathfrak{p}\}_{\mathfrak{p}\in\{\mathfrak{p}\}}$ and their beliefs $\mathsf{q}(\mathfrak{p})$, $\mathfrak{p} \in \{\mathfrak{p}\}$. Our quite general version follows.

Requirements on and choices of $\{\mathsf{p}_\mathfrak{p}\}_{\mathfrak{p}\in\{\mathfrak{p}\}}$ and $\mathsf{q}(\mathfrak{p})$

a) Tracking must be robust with respect to long periods of processing non-informative data: The use of the very prior, guaranteed-knowledge quantifying, pd $\mathsf{p}_0 \in \{\mathsf{p}_\mathfrak{p}\}$ assures this. The stabilised forgetting, see [12] and the recent survey in [25], requires exactly this.

b) Tracking must cope both with slow and abrupt parameter changes: This is the requirement uncovered by the stabilised forgetting. Its use thus requires extra measures based on a change detection [26]. MeRE just needs to keep permanently $\mathsf{q}(0|k) > 0$.

c) Computational overhead has to be kept low: It means that a few pds $\{\mathsf{p}_\mathfrak{p}\}$ can be *stored*. The low computational effort is also required for gaining the set of pds $\{\mathsf{p}_\mathfrak{p}\}$ with their beliefs $\mathsf{q}(\mathfrak{p})$, $\mathfrak{p} \in \{\mathfrak{p}\}$. The simplest case with two stored posterior pds is considered; $\mathsf{p}_0(\theta)$, with the belief $\mathsf{q}(0)$, and a pd $\mathsf{p}_1(\theta)$, with the belief $\mathsf{q}(1) = 1 - \mathsf{q}(0)$.

d) The quality of the estimation, consisting of data updating and approximation, is uncertain: [7] The outcome of this updating must be built in $\mathsf{p}_{\{\mathfrak{p}\}}$ (12) according its, locally judged, performance.

For this, a cheaply gained, additional member of the set $\{\mathsf{p}_\mathfrak{p}\} = \{\mathsf{p}_0(\theta), \mathsf{p}_1(\theta), \mathsf{p}_2(\theta)\}$ is considered before processing the new data record. The updating, see Subsection IV-A, requires assigning them the relevant beliefs $\mathsf{q}(0), \mathsf{q}(1), \mathsf{q}(2)$. MRE principle offers to re-distribute the belief on the extended set $\mathfrak{p} \in \{0, \dots, 2\}$. This problem of an extending set of hypotheses is addressed [27]. Its version used here is presented.

A given pd $\mathsf{q}(\mathfrak{p})$ with $\mathfrak{p} \in \{0, 1, \dots, \chi-1\}$, $\chi \geq 2$ determines the set of pds $\mathsf{q}_\kappa(\mathfrak{p}) \in \{\mathsf{q}_\kappa(\mathfrak{p})\}$ with $\mathfrak{p} \in \{0, \dots, \chi\}$

$$\{\mathsf{q}_\kappa\} = \{\mathsf{q}_\kappa(\mathfrak{p}) = \kappa\mathsf{q}(\mathfrak{p}),\ \mathfrak{p} \in \{0, \dots, \chi-1\}, \mathsf{q}_\kappa(\chi) = 1-\kappa\}. \tag{16}$$

The optional $\kappa \in (0, 1]$ determines its members. The extensions in $\{\mathsf{q}_\kappa\}$ preserve mutual relations of the given beliefs $\mathsf{q}(\mathfrak{p})$ on $\mathfrak{p} \in \{0, 1, \dots, \chi-1\}$. MRE selects the optimal extension within the set $\{\mathsf{q}_\kappa\}$. The admission of abrupt changes, see the requirement b), implies inevitability to use Laplace's insufficient reasons for the uniform choice u of the prior guess of the best extension in $\{\mathsf{q}_\kappa\}$. Simple evaluations imply that the MRE-optimal $\mathsf{q}_\mathsf{o} = \mathsf{q}_{\kappa_\mathsf{o}}$ with

$$\kappa_\mathsf{o} = \frac{1}{1 + \frac{1}{\chi-1}\exp[\mathcal{R}(\mathsf{q}||\mathsf{u})]}. \tag{17}$$

The re-distributed prior belief $\mathsf{q}_\mathsf{o}(\mathfrak{p})$, $\mathfrak{p} \in \{0, 1, 2\}$ allows us to process a new data record $\{o, a\} = \{\underline{o}, \underline{a}\}$ and to use evidences $\mathsf{f}_\mathfrak{p}$ needed for their updating. The requirement a) insists on preservation of p_0. Thus, the new $\mathsf{p}_2(\theta|\underline{o}, \underline{a})$ may only replace $\mathsf{p}_1(\theta|\underline{o}, \underline{a})$ if it got higher belief $\mathsf{q}(2|\underline{o}, \underline{a})$ than $\mathsf{q}(1|\underline{o}, \underline{a})$. Otherwise, it is discarded. New beliefs into the new considered pair $\mathsf{p}_0(\theta), \mathsf{p}_1(\theta|\underline{o}, \underline{a})$ are simply re-normalised the corresponding beliefs in $\mathsf{q}(\mathfrak{p}|\underline{o}, \underline{a})$, $\mathfrak{p} \in \{0, 1, 2\}$.

Commentary

- The belief into the fixed safe prior pd $\mathsf{p}_0(\theta)$ may fall numerically to zero when no abrupt changes occur for a long time. This should be counteracted by adding the condition $\mathsf{q}_\kappa(0) \geq \nu > 0$, with ν bigger than numerical zero, to the definition of the set $\{\mathsf{q}_\kappa\}$ (16), see the requirement a). It is omitted for simplicity and respected in the numerical implementation.
- The Bayes' updated p_0 and/or storing more possible realisations $\mathsf{p}_\mathfrak{p}$ of p are surely possible whenever the additional computational load is acceptable.
- Stabilised forgetting, improved by the above solution, used a pair of pds, typically, the safe fixed p_0 and $\mathsf{p}_1 \propto (\mathsf{l}(\theta|k)\mathsf{p}_0)^\phi \mathsf{p}_0^{1-\phi}$, $\phi \in [0, 1]$. It is algorithmically quite close to our solution with ϕ corresponding to $\mathsf{q}(1)$. As said, it requires a detection of abrupt parameter changes [26].

  Even when the parameter vary slowly the critical choice of the forgetting factor ϕ is to be addressed

[7] Generally, the outcome of Bayes' rule (13) has to be projected into the class of feasible pds on $\{\theta\}$ [9]. Here, the need is left implicit for the presentation simplicity sake. It changes neither concept nor subsequent evaluations.

in classical solutions. Two options prevail.

- $\star$ The choice ϕ is left to the user of the tracking algorithm. She is given a chance introduce her intuition about the nature of changes at her deliberation cost. This usual "academic" solution is flexible but hard to apply.
- $\star\star$ Under slow changes of the parameter the appropriate forgetting factor is expected to vary slowly, too, and can be learnt in Bayesian way. This often works but the Bayes' rule should use a "second layer forgetting". Otherwise ϕ, sooner or later, converges [23], mostly to zero or unity when the tracking ability is lost.
  The choice of the second layer forgetting might be less critical and the option $\star$ can be used. The decreased deliberation costs and the danger of significant user's error together with a built-in detection of abrupt parameter changes are relatively strong arguments for the proposed solution.

- A single application of Bayes' rule updates the posterior pd on a subspace of $\{\theta\}$. Assuming that non-negligible parameter changes happen at most in this subspace, the forgetting should be applied to it only. This makes forgetting "directional" [28].

1) Illustrative Example: The example is similar to numerous illustrations of stabilised forgetting. Linear Gaussian second order model was simulated with abrupt parameter changes and a period of insufficiently exciting actions. The estimation reduces to the algorithmic equivalent of recursive least squares. It determines the self-reproducing Gauss-inverse-Wishart posterior pd of the unknown regression coefficient θ and of the unknown environment-noise variance[8] r [15], [17]. The fixed prior p_0 delimited the units as the range of $\theta, r^{0.5}$ around zero. It was modified by one data vector $o_t, o_{t-1}, o_{t-1}, a_t$. Results as well as the remaining specific details are in Figures 1 – 7. They just confirm theoretical expectations.

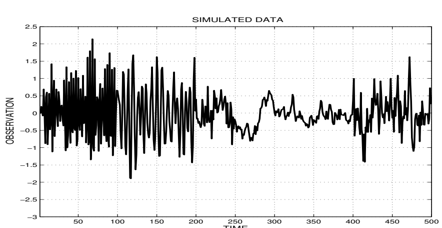

Fig. 1. Observations made on the environment $o_t = \theta_t \psi_t + r^{0.5} e_t$, for θ_t see caption of Figure 5, r Figure 4; $\psi_t = [o_{t-1}, o_{t-2}, a_t, 1]'$, e_t is white, zero-mean Gaussian sequence with unit variance.

V. Concluding Remarks

The paper derives the minimum *expected* relative entropy principle and indicates its application width.

[8]In this example, θ, r replaces θ of the general discussion.

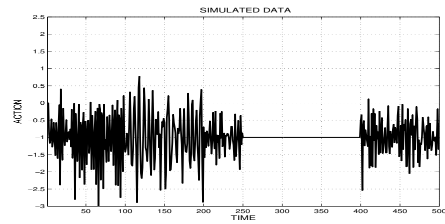

Fig. 2. Actions applied to the simulated environment, $a_t = -0.9 o_{t-1} - 1 + 0.1^{0.5} \varepsilon_t$ for $t \leq 250$, $a_t = -1$ for $250 < t \leq 400$ (poor excitation period) and $a_t = -0.6 o_{t-1} - 1 + 0.2^{0.5} \varepsilon_t$ for $400 < t$. ε_t is white, zero-mean Gaussian sequence with unit variance. It is independent of the noise e_t, see caption of Figure 1.

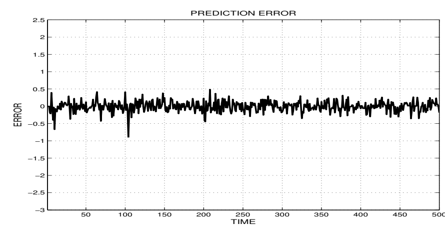

Fig. 3. Prediction errors $o_t - \hat{\theta}_{t-1} \psi_t$ with $\hat{\theta}_{t-1}$ being the newest point estimate of the regression coefficients θ_t, see Figures 1, 5.

A reader naturally expects an extensive experimental study. Our experience and space are limited but we feel that the presented simple idea is so useful that it makes sense to "sell" it as it is. We are aware that a future research should:

- make the said extensive experiments;
- elaborate details of the mentioned structure estimation, which is expected to bridge the approximate maximisation of the posterior likelihood over the space of competitive structures [24] with the popular automatic relevance determination [29];
- verify hypothesis that the proposed parameter tracking can practically cope well with slow parameter changes combined with more rare abrupt changes [30];
- apply the MeRE principle in areas relying now on the standard MRE principle as [4];
- apply the MeRE principle to the completion of the preference-describing ideal pd: this is a significant part of preference elicitation [31], [32].

We invite the interested readers to join or replace us.

References

[1] A. Jazwinski, *Stochastic Processes and Filtering Theory.* Ac. Press, 1970.
[2] J. Shore and R. Johnson, "Axiomatic derivation of the principle of maximum entropy & the principle of minimum cross-entropy," *IEEE Tran. on Inf. Th.*, vol. 26, no. 1, pp. 26–37, 1980.
[3] Z. Botev and D. Kroese, "The generalized cross entropy method, with applications to probability density estimation," *Methodol. Comput. Appl. Probab.*, vol. 13, pp. 1 – 27, 2011.

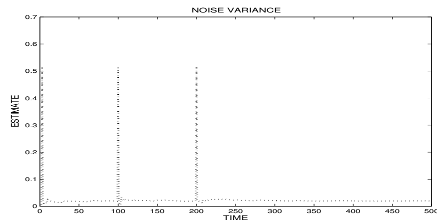

Fig. 4. Estimates of the environment-noise variance r. The simulated value was $r = 0.02$.

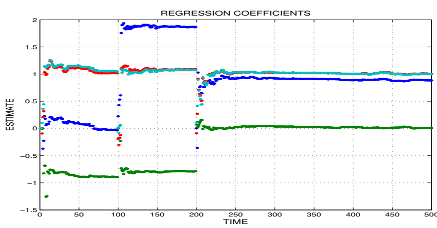

Fig. 5. Estimates of the regression coefficients θ_t, see Figure 1. Simulated values were $\theta_t = [0, -0.9, 1, 1]$ for $t \leq 100$ (a single root, delay 2), $\theta_t = [1.8, -0.81, 1, 1]$ for $100 < t \leq 200$ (a double root), and $\theta_t = [0.9, 0, 1, 1]$ for $200 < t$ (a single root, delay 1).

[4] S. Pare, A. Kumar, V. Bajaj, and G. Singh, "An efficient method for multilevel color image thresholding using cuckoo search algorithm based on minimum cross entropy," *Applied Soft Computing*, vol. 61, pp. 570 – 592, 2017.

[5] H. Shao, H. Jiang, Y. Lin, and X. Li, "A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders," *Mechanical Systems and Signal Processing*, vol. 102, pp. 278 – 297, 2018.

[6] M. Sheppard, J. Sullivan, P. Lipa, and A. Terrazas, "Methods and apparatus to utilize minimum cross entropy to calculate granular data of a region based on another region for media audience measurement," February 2019, patent US20190069024A1.

[7] S. Kullback and R. Leibler, "On information and sufficiency," *An. of Math. Stat.*, vol. 22, pp. 79–87, 1951.

[8] I. Good, "Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables," *Annals Math. Stat.*, vol. 34, pp. 911–934, 1963.

[9] M. Kárný, "Approximate Bayesian recursive estimation," *Inf. Sci.*, vol. 289, pp. 100–111, 2014.

[10] M. Kárný and T. Guy, "On support of imperfect Bayesian participants," in *Decision Making with Imperfect Decision Makers*, T. Guy and et al, Eds. Springer, Int. Syst. Ref. Lib., 2012, vol. 28, pp. 29–56.

[11] M. Kárný and T. Kroupa, "Axiomatisation of fully probabilistic design," *Inf. Sci.*, vol. 186, no. 1, pp. 105–113, 2012.

[12] R. Kulhavý and M. B. Zarrop, "On a general concept of forgetting," *Int. J. of Control*, vol. 58, no. 4, pp. 905–924, 1993.

[13] L. Savage, *Foundations of Statistics*. Wiley, 1954.

[14] D. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 2017, two volumes.

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

[16] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 1991.

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

[18] M. Rao, *Measure Theory and Integration*. NY: John Wiley, 1987.

[19] G. Debreu, "Representation of a preference ordering by a numerical function," in *Decision Processes*, R. Thrall, C. Coombs, and R. Davis, Eds. NY: Wiley, 1954.

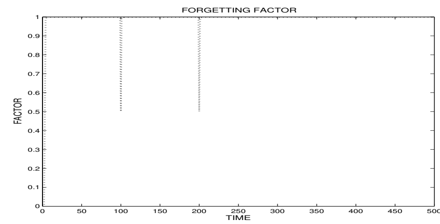

Fig. 6. Forgetting factor coincides with the belief $\mathsf{q}(1|k)$ into the permanently updated posterior pd $\mathsf{p}_1(\theta, r|k)$.

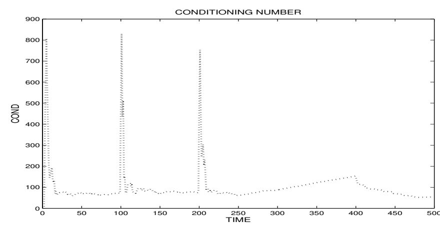

Fig. 7. The evolution of the conditioning number of the permanently updated covariance factor confirms the suppressed influence of the poor excitation.

[20] C. Genest and J. Zidek, "Combining probability distributions: A critique and annotated bibliography," *Stat. Sci.*, vol. 1, no. 1, pp. 114–148, 1986.

[21] I. Sanov, "On probability of large deviations of random variables," *Matematičeskij Sbornik*, vol. 42, pp. 11–44, 1957, (Russian); in Selected Translations Mathematical Statistics and Probability, I, 1961, 213–244.

[22] P. Algoet and T. Cover, "A sandwich proof of the Shannon-McMillan-Breiman theorem," *The Annals of Probability*, vol. 16, pp. 899–909, 1988.

[23] L. Berec and M. Kárný, "Identification of reality in Bayesian context," in *Computer-Intensive Methods in Control and Signal Processing*, K. Warwick and M. Kárný, Eds. Birkhäuser, 1997, pp. 181–193.

[24] 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. Spall, Ed. New York: Marcel Dekker, 1988.

[25] J. Dokoupil and P. Václavek, "Data-driven stabilized forgetting design using the geometric mean of normal probability densities," in *2018 IEEE Conference on Decision and Control (CDC)*, 2018, pp. 1403–1408.

[26] J. Holst and N. Poulsen, "Self tuning control of plants with abrupt changes," *IFAC Proceedings Volumes*, vol. 17, no. 2, pp. 923 – 928, 1984, 9th IFAC World Congress.

[27] M. Kárný, "On assigning probabilities to new hypotheses," *Signal Processing*, 2020, submitted.

[28] R. Kulhavý, "Directional tracking of regression-type model parameters," in *Preprints of the 2nd IFAC Workshop on Adaptive Systems in Control and Signal Processing*, Lund, Sweden, 1986, pp. 97–102.

[29] C. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[30] F. Gustafsson, *Adaptive Filtering and Change Detection*. John Wiley & Sons, 2000.

[31] J. Branke, S. Corrente, S. Greco, and W. Gutjahr, "Efficient pairwise preference elicitation allowing for indifference," *Computers & Operations Research*, vol. 88, no. Suppl. C, pp. 175 – 186, 2017.

[32] M. Kárný and T. Guy, "Preference elicitation within framework of fully probabilistic design of decision strategies," in *IFAC International Workshop on Adaptive and Learning Control Systems - ALCOS 2019*, vol. 52, no. 29, 2019, pp. 239–244. [Online]. Available: https://www.sciencedirect.com/journal/ifac-papersonline/vol/52/issue/29