MULTIOBJECTIVE PROBABILISTIC MIXTURE CONTROL

Josef Böhm, Tatiana V. Guy and Miroslav Kárný [1]

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
P. O. Box 18, 182 08 Prague, Czech Republic
e-mail:guy@ieee.org

Abstract: Paper formulates the problem of multiobjective probabilistic mixture control design and proposes its general solution with both system model and target represented by finite probabilistic mixtures. A complete feasible algorithmic solution for mixtures with components formed by normal auto-regression models with external variable is provided.

Keywords: Multiobjective control, probabilistic models, decision making

1. INTRODUCTION

The growing complexity of control problems, accompanied by increasing multiple performance criteria that should be reached simultaneously makes a *multiobjective control problem*, being particular case of *multiobjective decision making*, important. Despite a lot of approaches to multiobjective control developed (Stadler 1988), (Toivonen 1989), (Liao & Li 2002), there is still a lack of systematic methodology guaranteeing satisfactory solution of the problem. Creating efficient numerical algorithms of general structure is much complicated by the high dimensionality and uncertainty of the modern processes to be controlled.

The approach advocated here belongs to multi-model framework with the controlled system described by dynamic probabilistic mixture model (Kárný, et al. 2003). Historical process data, fully describing the closed-loop system behaviour are processed by quasi-Bayes algorithm (Kárný, et al. 2005) to build the mixture model. The model obtained reflects all significant operational modes of the system, each associated with so-called *mixture component*, while their weights indicate the probability of occurrence of a particular mode component. The mixture description of a system is especially useful in complex, large-scale systems, when the system behaviour exhibits several different modes and cannot be described by a linear model with fixed parameters. The underlying probabilistic extraction of information from process data does not require a detailed knowledge of system dynamics and thus supports the approach's generality.

The adopted fully probabilistic design methodology gives interesting insight into the *probabilistic decision making problem*. The approach designs a strategy that minimises the distance of the joint probability density function (pdf) describing closed-loop system behaviour to a joint pdf describing desired closed-loop behaviour. Kullback-Liebler divergence serves as a measure of the closeness and thus as universal quality criterion.

Algorithmic solution proposed has been obtained for the *system model* and the *user ideal* in the form of *finite mixtures* of uni-modal pdfs. This

[1] Partially supported by AVČR 1ET100750401, GAČR 102/02/0204 and GAČR 102/03/P010.

corresponds to the situations frequently met in practice, when there exist: i) several requirements imposed on single data item; ii) different ideals given by different users.

2. PRELIMINARIES

Throughout the text, $x(t)$ stands for $(x_1, \ldots, x_t)$, x^* denotes set of possible values of a variable x and $\mathring{x}$ is the number of entries in x^*. The symbol $f(x_{i;t})$, $t \in t^*$, denotes probability (density) function (p(d)f) of the ith entry of x at discrete time $t \in t^* \equiv \{1, \ldots, \mathring{t}\}$, $\mathring{t} \leq \infty$. Time index follows the semicolon in the subscript when multiple indices are used. The symbol x' means transposition of x.

Kullback-Leibler (KL) divergence is used as a measure of proximity of a pair of pdfs f, g acting on a common set x^*. Their KL divergence $\mathcal{D}(f||g)$, with the basic property $\mathcal{D}(f||g) \geq 0$, $\mathcal{D}(f||g) = 0$ iff $f = g$ almost everywhere, is defined by:

$$\mathcal{D}(f||g) \equiv \int f(x) \ln\left(\frac{f(x)}{g(x)}\right) dx. \qquad (1)$$

2.1 Probabilistic mixture modelling

The used probabilistic modelling operates with joint pdf on uncertain system quantities considered. They are supposed to consist of: the observed controlled system output y_t; the directly manipulated system input u_t; unobserved unknown time-invariant parameter $\Theta \in \Theta^*$ and a random pointer to the active mixture component $c_t \in c^* \equiv \{1, \ldots, \mathring{c}\}$, $\mathring{c} < \infty$.

The relation between the quantities involved is assumed to be (approximately) described by the mixture in the *component form*

$$f(d(\mathring{t}), c(\mathring{t}), \Theta) \equiv Di_\alpha(\kappa_0) \prod_{t \in t^*} f(d_t|\phi_{c_t;t-1}, \Theta_{c_t}, c_t) \times \alpha_{c_t} \prod_{c \in c^*} f(\Theta_c), \quad \text{where} \qquad (2)$$

$f(d_t|\phi_{c_t;t-1}, \Theta_{c_t}, c_t)$ is parameterised component, usually represented by uni-modal pdf of *data item* $d_t = [y_t, u_t]'$; the number of components $\mathring{c}$ is assumed to be finite and fixed;

$\phi_{c_t;t-1}$ is observable *state of* c_t*-th component*; the phase form $\phi_{c;t-1} \equiv [d'_{t-1}, \ldots, d'_{t-\partial_c}, 1]'$ of the state is considered; finite fixed orders ∂_c are assumed;

α_{c_t} is constant probability of the pointer c_t, called *component weight*; their collection forms probabilistic vector $\alpha \in \alpha^* \equiv \{\alpha_c \geq 0, \sum \alpha_c = 1\}$;

Θ is *mixture parameter* parameterising the model; it is formed of the component parameters and weights, i.e. $\Theta = \{\Theta_c, \alpha_c\}_{c \in c^*}$;

$f(\Theta_c)$ are prior pdfs of unknown $\Theta_c \in \Theta^*_c$ parameterising individual components $c \in c^*$;

$Di_\alpha(\kappa_0) \propto \prod_{c \in c^*} \alpha_c^{\kappa_{c;0}-1}$ is prior Dirichlet pdf (Fergusson 1973) of the component weights, determined by the vector statistic κ_0 with non-negative entries $\kappa_{c;0}$; α^* is its support; the symbol $\propto$ means that right hand side has to be normalised to the unit integral to get equality.

In the closed-loop, the data items $d_t \equiv (d_{1;t}, \ldots, d_{\mathring{d};t})$ are multivariate, i.e. $\mathring{d} > 1$. Using the chain rule, the individual components can be decomposed into parameterised *factors*

$$\begin{aligned} f(d_t \mid \phi_{c_t;t-1}, \Theta_{c_t}, c_t) &= \qquad (3) \\ &= \prod_{i=1}^{\mathring{d}} f(d_{i;t}|d_{i+1;t}, \ldots, d_{\mathring{d};t}, \phi_{c_t;t-1}, \Theta_{c_t}, c_t) \\ &\equiv \prod_{i=1}^{\mathring{d}} f(d_{i;t}|\psi_{ic_t;t}, \Theta_{ic_t}, c_t) \text{ with} \end{aligned}$$

regression vectors $\psi_{ic_t;t} = [d_{i+1;t}, \ldots, d_{\mathring{d};t}, \phi'_{c_t;t-1}]'$. The decomposition allows to model jointly scalar entries $d_{i;t}$, called *factor outputs*, and to parameterise factors individually. Moreover, only some parameters Θ_{ic} from Θ_c may be needed to describe the i-th factor in (3). The factorised description allows to treat factor outputs of a different nature, for instance, this enables joint modelling of mixed continuous and discrete data. To simplify the presentation, the model description at component level (2) is used below.

The model introduced is the *mixture model*. This becomes obvious when the unobserved pointers to components are excluded by marginalisation

$$\begin{aligned} f(d_t|d(t-1), \Theta) &= f(d_t|\phi_{1;t-1}, \ldots, \phi_{\mathring{c};t-1}, \Theta) \quad (4) \\ &= \sum_{c_t \in c^*} \alpha_{c_t} f(d_t|\phi_{c_t;t-1}, \Theta_{c_t}, c_t). \end{aligned}$$

The expression (4) extends usual finite mixtures as it consists of *dynamic* components. It is, however, still restricted to have constant component weights, i.e. $f(c_t = c|d(t-1), \Theta) = \alpha_c$. The assumption is used to get a feasible recursive estimation and can be weakened so that slow changes are admitted by employing stabilised forgetting. The approximate recursive quasi-Bayes mixture estimation (Kárný et al. 2005) is used. Subsequent control design adopts certainty-equivalence strategy and is based on recursively estimated model. Thus, parameter Θ can be dropped in pdfs from here onwards.

2.2 Mixture probabilistic control

The joint pdf $f(d(\mathring{t}), c(\mathring{t})) \equiv f(y(\mathring{t}), u(\mathring{t}), c(\mathring{t}))$ characterising possible *closed-loop behaviours* of the system and an input generator can be factorised in the following way:

$$f(d(\mathring{t}),\ c(\mathring{t})) = \prod_{t\in t^*} f(y_t|u_t, d(t-1), c_t) \qquad (5)$$
$$\times \prod_{t\in t^*} f(c_t|d(t-1), u_t) \prod_{t\in t^*} f(u_t|d(t-1)).$$

The factors $\{f(y_t|u_t, d(t-1), c_t)\}_{t\in t^*}$ are learned during the estimation and describe observable system reactions on the control actions u_t under the available experience reflected in the data $d(t-1)$ and for the fixed c_t. Pdfs $\{f(u_t|d(t-1))\}_{t\in t^*}$ describe models of the considered randomised control strategy. The factors $\{f(c_t|u_t, d_{t-1})\}_{t\in t^*}$ represent probabilities of pointers c_t to the particular components that determine component weights of the mixture model (2).

Control strategy designed is intended to make the closed-loop behaviour as close as possible to a desired one, when respecting given restrictions. Under the adopted probabilistic modelling, a control strategy can be searched as minimiser of the KL divergence of the joint pdf (5) from its prespecified ideal counterpart ${}^I f(d(\mathring{t}), c(\mathring{t}))$. The last pdf is called *user's ideal* or *ideal* and describes the desired closed-loop behaviour and given restrictions. The methodology providing the solution of the problem is known as *fully probabilistic design* (Kárný et al. 2005) and is recalled in the following Agreement.

Agreement 1. (Fully probabilistic design). The fully probabilistic design specifies its target through an *ideal pdf*

$${}^I f(d(\mathring{t}),\ c(\mathring{t})) = \qquad (6)$$
$$\prod_{t\in t^*} {}^I f(y_t|u_t, c_t, d(t-1))\, {}^I f(u_t, c_t|d(t-1)).$$

The optimal strategy is selected among causal, randomised, strategies $\{f(u_t, c_t|d(t-1)\}_{t\in t^*}$. It is defined as a minimiser of the KL divergence (1) of $f(d(\mathring{t}), c(\mathring{t}))$ from ${}^I f(d(\mathring{t}), c(\mathring{t}))$

$$\mathcal{D}\ \left(f||\,{}^I f\right) \equiv \qquad (7)$$
$$\int f(d(\mathring{t}), c(\mathring{t})) \ln\left(\frac{f(d(\mathring{t}), c(\mathring{t}))}{{}^I f(d(\mathring{t}), c(\mathring{t}))}\right) d(d(\mathring{t}), c(\mathring{t})).$$

The general formulation of fully probabilistic design guarantees its applicability to a wide class of problems.

While components, $f(y_t|u_t, d(t-1), c_t)$ (5), are obtained from the estimation and should be considered as given, the remaining two factors in (5) can be optimised and, by this, influence the closed-loop behaviour. Dependently on which factor is chosen for optimisation, there exist three types of the design: *academic*, *industrial* and *simultaneous*.
Academic design optimises probabilities of pointers to particular components c_t.
Industrial design optimises randomised control strategy $f(u_t|d(t-1))$ without changing the probabilities of components. This type of design has to be used whenever the component weights have an objective meaning that cannot be influenced by the optional controllers used.
Simultaneous design (Kárný et al. 2003) combines features both academic and industrial designs. This type of the design optimises the joint pdf $f(u_t, c_t|d(t-1)) \equiv f(c_t|u_t, d(t-1))f(u_t|d(t-1))$. It takes the model (4) as an approximation of a non-linear dynamic model and searches for the proper system inputs while respecting possible changes of operation mode.

3. MULTIOBJECTIVE CONTROL DESIGN

Probabilistic control for the described types of design was elaborated only for the case of *uni-modal user's ideal* (Kárný et al. 2003). The paper extends the technique to the case when *desired closed-loop behaviour is described by a mixture.* The mixture form of user's ideal allows to respect different user's demands expressed via particular components of the ideal mixture. The designed optimal strategy provides the compromise between these demands.
The *simultaneous* design selects the optimal pdf within the set of causal randomised strategies:

$$\{f(c_t|u_t, d(t-1))f(u_t|d(t-1))\}_{t\in t^*}. \qquad (8)$$

The optimal pdf defines such probabilities of particular components that make the resulting mixture closest to the user's ideal. Unlike uni-modal ideal case, the suggested *mixture form* of user's ideal needs evaluation the KL divergence between two mixtures. As this evaluation is difficult, the divergence between two *joint* pdfs describing *optimised* and *desired closed-loop behaviour* is used in minimisation.
The behaviour of the optimised closed-loop is described by the joint pdf

$$f(d(\mathring{t}), c(\mathring{t})) = \prod_{t\in t^*} f(d_t, c_t|d(t-1)) = \qquad (9)$$
$$\prod_{t\in t^*} f(y_t|u_t, d(t-1), c_t) f(u_t, c_t|d(t-1)),$$

where $f(y_t|u_t, d(t-1), c_t)$ are learned components of the system model and $f(c_t|u_t, d(t-1))$ is a strategy from (8).
The *ideal pdf*, describing the desired behaviour of the closed-loop, is expressed through the *finite mixture* of $\mathring{\tilde{c}}$ components:

$${}^I f(d(\mathring{t}), \tilde{c}(\mathring{t})) = \prod_{t\in t^*} {}^I f(y_t|d(t-1), \tilde{c}_t)\, {}^I f(u_t, \tilde{c}_t|d(t-1)) \qquad (10)$$

with their *supports* nested in the following way:

$$\text{supp}\ [f(d_t|d(t-1), c_t)] \subseteq \text{supp}\ \left[{}^I f(d_t|d(t-1), \tilde{c}_t)\right]. \tag{11}$$

The first factor in (10) describes user's wishes and restrictions on particular components $\tilde{c} \in \tilde{c}^*$ with respect to the system data. The number of components $\mathring{\tilde{c}}$ in the ideal mixture can generally be different from the number of components $\mathring{c}$ learned (2). The minimisation of KL divergence of two joint pdfs $\mathcal{D}(f(d(\mathring{t}), c(\mathring{t}))||\, {}^I f(d(\mathring{t}), \tilde{c}(\mathring{t})))$ is simplified when the numbers of components equal, i.e. $\mathring{c} = \mathring{\tilde{c}}$. To get this, Agreement 2 is used.

Agreement 2. (Extension of components number). Let mixtures f_1 and f_2 have different number of the components $\mathring{c}_2 > \mathring{c}_1$ and $m = \mathring{c}_2 - \mathring{c}_1$ is missing number of components of the mixture f_1.

Then, the amount of components $\mathring{c}_1$ of the mixture f_1 can be extended on missing number m by:

(1) "virtual" duplication of any m components of f_1, such that the probability of each original component is equally distributed among two "virtually" created components;
(2) "artificial" creation of a m new components of mixture f_1, such that each of new components is described by a flat prior pdf with a small weight.

A possible variant of the first way can be m-times duplication only one of the original components. The second choice guarantees that the inclusion (11) holds also for the joint pdfs on $(d_t, c_t)^*$. The first way cannot be applied to the extension of ideal mixtures as it causes decreasing the weights of those m components chosen for duplication.
Besides, the components of both ideal and learned mixtures can be ordered arbitrarily, so the design (Agreement 1) should take into account all possible permutations of the components. It means an optimal strategy from (8) for a permutation of components ${}^{opt}\pi_{\mathring{t}}(c(\mathring{t}))$ minimising the KL divergence between optimised and desired closed-loop behaviour, will be searched for:

$$ {}^{opt} f({}^{opt}\pi_{\mathring{t}}(c(\mathring{t}))) = \min_{\{\pi_{\mathring{t}}(c(\mathring{t}))\}} \tag{12}$$

$$\arg\min_{\{f(\pi_t(c_t|d(t-1))\}} \mathcal{D}(f(d(\mathring{t}), \pi_{\mathring{t}}(c(\mathring{t}))||\, {}^I f(d(\mathring{t}), c(\mathring{t}))).$$

Assuming a fixed, time-invariant permutation of the components $\pi_{t+1}(c_t) = \pi_t(c_t) = \pi(c_t)$, the solution of the addressed fully probabilistic design (Agreement 1) is described by the proposition:

Proposition 1. **Solution fully probabilistic design with multimodal target and fixed $\pi(c_t)$**
The optimal strategy minimising the KL divergence (1) has the following form ($\gamma(d(\mathring{t})) = 1$):

$$f(\pi(c_t), u_t|(t-1)) = {}^I f(c_t|u_t, d(t-1))$$
$$\times\, {}^I f(u_t|d(t-1)) \frac{\exp[-\omega_\gamma(\pi(c_t), u_t, d(t-1))]}{\gamma(d(t-1))},$$
$$\gamma(d(t-1)) \equiv \sum_{c_t \in c^*} {}^I f(c_t|u_t, d(t-1))\, {}^I f(u_t|d(t-1))$$
$$\times \exp[-\omega_\gamma(\pi(c_t), u_t, d(t-1))]$$
$$\omega_\gamma(\pi(c_t), u_t, d(t-1)) \equiv \int f(y_t|u_t, \pi(c_t), d(t-1))$$
$$\times \ln\left(\frac{f(y_t|u_t, \pi(c_t), d(t-1))}{\gamma(d(t))\, {}^I f(y_t|c_t, d(t-1))}\right) dy_t$$

The solution is performed against the time course, starting at $t = \mathring{t}$.
Proof: is similar to that for the uni-modal target, see (Kárný et al. 2005).◇

Time-invariant permutation of components corresponds to a frequently met real situation, when particular components describe known operation modes or physical states of the system considered. Then the permutation is defined by physical properties of the system and should be taken as given. If the permutation varies with time, an additional minimisation over possible component permutations is to be performed, see (12).
The exact evaluation of the KL divergence and thus, direct practical application of Proposition 1, is problematic as the reached minimum has the mixture form. The feasible variant of the design is then obtained when Jensen upper bound of it is used (Rao 1987) and the upper bound KL divergence is minimised.

4. CONTROL DESIGN FOR MIXTURES OF NORMAL ARX MODELS

The solution is applied to the mixtures of normal auto-regression models with exogenous variables (ARX). Models from this class can be easily identified (Peterka 1981), (Kárný et al. 2005) and they represent one of a few classes suitable for a numerical solution of large dimensional problems. The system is modelled by a mixture with normal components

$$f(d_t|d(t-1), \Theta_c, c) = \mathcal{N}_{d_t}(\theta_c' \phi_{c;t-1}, r_c), \tag{13}$$

where $\mathcal{N}_x(\bar{x}, r) \equiv |2\pi r|^{-0.5} \exp\left[-0.5(x-\bar{x})' r^{-1} (x-\bar{x})\right]$; parameters $\Theta_c = [\theta_c, r_c]'$ consist of matrix regression coefficients θ_c and covariance matrix r_c. The state vector $\phi_{c;t-1}$ contains the delayed data items $\phi_{c;t-1} = [d'_{(t-1)\ldots(t-\partial_c)}, 1]'$, $\partial_c \geq 0$. The regression coefficients are complemented by zeros so that all factors within a single component have a common state vector. The regression vectors of individual factors are nested in the following way:

$$\psi_{i;t} \equiv [d'_{i+1\ldots\mathring{d};t}, \phi'_{t-1}]' \equiv [d_{i+1;t}; \psi'_{i+1;t}]'$$

$$\psi_{\mathring{d};t} \equiv \phi_{t-1} \equiv [d'_{(t-1)\text{-}(t-\partial)}, 1]',\ \partial \geq 0, \qquad (14)$$

$$\psi_{0;t} \equiv \Psi_t \equiv [d'_{t\text{-}(t-\partial)}, 1]',\ i = 1, \ldots, \mathring{d} - 1.$$

The closed-loop model in the factor form is

$$f(d(\mathring{t}), c(\mathring{t})) = \prod_{t \in t^*} \prod_{i=1}^{\mathring{d}} \mathcal{N}_{d_{ic;t}}(\theta'_{ic}\psi_{ic;t}, r_{ic}) f(c_t|u_t, d(t-1)).$$

The parameterised factors of the ideal mixture are also supposed to be normal with parameters ${}^{I}\Theta_c = [{}^{I}\theta_c, {}^{I}r_c]'$ consisting of matrix regression coefficients ${}^{I}\theta_c$ and covariance matrix ${}^{I}r_c$:

$${}^{I}f(d_t|d(t-1), {}^{I}\Theta_c, c) = \mathcal{N}_{d_t}({}^{I}\theta'_c\phi_{c;t-1}, {}^{I}r_c), \qquad (15)$$

then the desired closed-loop behaviour reads:

$${}^{I}f(d(\mathring{t}), c(\mathring{t})) = \prod_{t \in t^*} \prod_{1}^{\mathring{d}} \mathcal{N}_{d_{ic;t}}({}^{I}\theta'_{ic}\psi_{ic;t}, {}^{I}r_{ic})\, {}^{I}f(c_t|u_t, d(t-1)).$$

The algorithmic realisation of the solution is based on factorised form of the components. Detailed description and proofs for the case of uni-modal target are given in (Kárný et al. 2005). An algorithmic solution of the simultaneous design for the normal mixtures and for a fixed permutation of the components is summarised in the following algorithm.

Algorithm 1. **Simultaneous fully probabilistic design with multimodal user's ideal applied to ARX mixtures.**
The optimal strategy minimising upper bound of KL divergence with models (13) and (15) is described by the pdfs

$$\begin{aligned}
{}^{I}f(c_t, u_t|\phi_{t-1}) &\propto {}^{U}f(c_t)\exp\left[-0.5\omega_\gamma(c_t, \phi_{t-1})\right] \\
&\quad \times {}^{I}f(u_t|\phi_{t-1}, c_t) \\
\omega_\gamma(c_t, \phi_{t-1}) &\equiv k_{c_t;t-1} + \\
&\quad \phi'_{t-1}L_{c_t;t-1}D_{c_t;t-1}L'_{c_t;t-1}\phi_{t-1} \\
{}^{I}f(u_t|\phi_{t-1}, c_t) &\equiv \prod_{i=\mathring{y}+1}^{\mathring{d}} \mathcal{N}_{u_{(i-\mathring{y});t}}\left({}^{I}\theta'_{ic_t;t-1}\psi_{i;t},\right. \\
&\quad \left.{}^{I}r_{ic_t;t-1}\right).
\end{aligned}$$

The solution is performed against time course, starting at $L_{\gamma;\mathring{t}} = I_{\mathring{\phi}}$, $D_{\gamma;\mathring{t}} = 0$, $k_{\gamma;\mathring{t}} = 0$.

For $t = \mathring{t}, \ldots, 1$

$L_{\mathring{d}} = I_{\mathring{d}},\ D_{\mathring{d}} = 0$

For $c = 1, \ldots, \mathring{c}$

$k_{0c} = -\mathring{d}_c + k_{\gamma;t}$,

$$L_{0c} = \begin{bmatrix} {}^{\psi_0}L_{\gamma;t} & 0 & 0 \\ 0 & L_{\mathring{d};t} & 0 \\ {}^{\psi_k}L_{\gamma;t} & 0 & 1 \end{bmatrix}$$

$D_{0c} = \mathrm{diag}\left[{}^{\psi_0}D_{\gamma;t}, D_{\mathring{d};t}, {}^{\psi_k}D_{\gamma;t}\right]$

$nor = 0$

For $i = 1, \ldots, \mathring{y}$

$$L_{ic}D_{ic}L'_{ic} = {}^{\psi_{i-1}}L_{(i-1)c}\, {}^{\psi_{i-1}}D_{(i-1)c}\, {}^{\psi_{i-1}}L'_{(i-1)c} + \left(\theta_{ic} + {}^{d\psi}L_{(i-1)c}\right) {}^{d}D_{(i-1)c} \left(\theta_{ic} + {}^{d\psi}L_{(i-1)c}\right)' + \frac{\left(\theta_{ic} - {}^{I}\theta_{ic}\right)\left(\theta_{ic} - {}^{I}\theta_{ic}\right)'}{{}^{I}r_{ic}}$$

$$k_{ic} = k_{(i-1)c} + {}^{d}D_{(i-1)c}r_{ic} + \left[\ln\left(\frac{{}^{I}r_{ic}}{r_{ic}}\right) + \frac{r_{ic}}{{}^{I}r_{ic}}\right]$$

end of the cycle over i

For $i = \mathring{y}+1, \ldots, \mathring{d}$

$$L_{ic}D_{ic}L'_{ic} = {}^{\psi_{i-1}}L_{(i-1)c}\, {}^{\psi_{i-1}}D_{(i-1)c}\, {}^{\psi_{i-1}}L'_{(i-1)c} + \left(\theta_{ic} + {}^{d\psi}L_{(i-1)c}\right) {}^{d}D_{(i-1)c} \left(\theta_{ic} + {}^{d\psi}L_{(i-1)c}\right)'$$

$$L_{ic} = \begin{bmatrix} 1 & 0 \\ -{}^{I}\theta_{ic;t-1} & L_{(i+1)c} \end{bmatrix}$$

$\tilde{D}_{ic} = \mathrm{diag}\left[{}^{I}r^{-1}_{ic;t-1}, D_{ic}\right]$, ${}^{I}r_{ic;t-1}$ is scalar,

$$k_{ic} = k_{(i-1)c} + {}^{d}D_{(i-1)c}r_{ic} + \left[\ln\left(\frac{{}^{I}r_{ic}}{r_{ic}}\right) + \frac{r_{ic}}{{}^{I}r_{ic}}\right]$$

end of the cycle over i

$k_{c;t-1} = k_{\mathring{d}c}$

$L_{c;t-1}D_{c;t-1}L'_{c;t-1} = L_{\mathring{d}c}D_{\mathring{d}c}L'_{\mathring{d}c}$

$\beta_c = {}^{I}f(c)\exp[-0.5k_{c;t-1}],\ nor = nor + \beta_c$

end of the cycle over c

$L_{\gamma;t-1} = I_{\mathring{\phi}}, D_{\gamma;t-1} = 0, k_{\gamma;t-1} = 0$

For $c = 1, \ldots, \mathring{c}$

$$L_{\gamma;t-1}D_{\gamma;t-1}L'_{\gamma;t-1} = L_{\gamma;t-1}D_{\gamma;t-1}L'_{\gamma;t-1} + \frac{\beta_c}{nor}L_{c;t-1}D_{c;t-1}L'_{c;t-1}$$

$$k_{\gamma;t-1} = k_{\gamma;t-1} + \frac{\beta_c}{nor}k_{c;t-1}$$

end of the cycle over c

end of the cycle over t.

Remarks 1:

- For a known state vector ϕ_{t-1}, it is possible to evaluate the achieved minimum as $\mathcal{D}(f(c)||\,{}^{I}f(c)) + \exp\left[-0.5\left(k_{\gamma;t} + \phi'_{t-1}L_{\gamma;t-1}D_{\gamma;t-1}L'_{\gamma;t-1}\phi_{t-1}\right)\right]$ and compare quality of various permutations. It is always possible for static systems with $\phi_{t-1} = 1$. Otherwise approximation of the future unknown states by the current known one is possible.
- The total number of such permutations is $\mathring{c}!$.
- Usually, the number of components is not very high. In rare cases when $\mathring{c}$ is big enough, the incomplete search for the maximum should be performed.

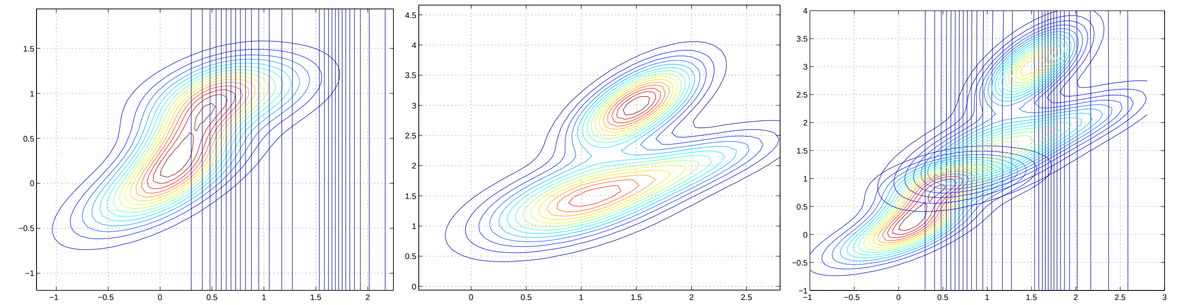

Fig. 1. Projections of mixtures.

5. ILLUSTRATIVE EXAMPLE

The data describing system with two data items $d_t = (d_{1;t}, d_{2;t})$ have been simulated. Mixture with three normal components (13) served as system model. The factors f_i, $i = \{1, \ldots, 6\}$ entering components and corresponding to the particular data items are also normal $\mathcal{N}_{d_i}(\bar{x}, r)$ with mean values $\bar{x}$ and variances r, see Table 1.

Table 1. System mixture

c_i	first factor	second factor
1	$f_1 \propto \mathcal{N}(1.0 * d_{2;t}, 0.12)$	$f_4 \propto \mathcal{N}(0.0, 0.12)$
2	$f_2 \propto \mathcal{N}(0.5 * d_{2;t}, 0.07)$	$f_5 \propto \mathcal{N}(0.0, 0.15)$
3	$f_3 \propto \mathcal{N}(0.7 * d_{2;t}, 0.20)$	$f_6 \propto \mathcal{N}(1.0, 0.07)$

Probabilities of pointers to the particular components are set to be equal $f(c) = [1/3, 1/3, 1/3]$, for the components ordering $c = \{1\ 2\ 3\}$. The user's ideal has been chosen also in a mixture form but with different components ordering, see Table 2.

Table 2. User's ideal mixture

c_i	first factor	second factor
1	${}^I f_3 \propto \mathcal{N}(2.0, 0.15)$	${}^I f_6 \propto \mathcal{N}(0, 1000.0)$
2	${}^I f_1 \propto \mathcal{N}(1.0, 0.10)$	${}^I f_4 \propto \mathcal{N}(0, 1000.0)$
3	${}^I f_2 \propto \mathcal{N}(1.5, 0.05)$	${}^I f_5 \propto \mathcal{N}(0, 1000.0)$

Probabilities of pointers to the particular components of the user's ideal (Table 2) are set to ${}^I f(\tilde{c}) = [1/6, 1/2, 1/3]$, with components ordering $\tilde{c} = \{1\ 2\ 3\}$. The equiprobability curves of the system model (Table 1) and ideal mixture (Table 2) are shown on the first plot, Fig. 1. The user's ideal is extremely prolonged at the direction of the second data item (note vertical lines on the plot), which means that no value of d_2 is preferred. Thus, the designed optimised mixture (Table 3) respects wishes and restrictions on d_1, mainly.

Table 3. Optimised mixture

c_i	first factor	second factor
1	${}^o f_1 \propto \mathcal{N}(1.9997, .12)$	${}^o f_4 \propto \mathcal{N}(1.9997, .1999)$
2	${}^o f_2 \propto \mathcal{N}(1.4990, .07)$	${}^o f_5 \propto \mathcal{N}(2.9990, .1999)$
3	${}^o f_3 \propto \mathcal{N}(0.9996, .20)$	${}^o f_6 \propto \mathcal{N}(1.4280, .2040)$

Resulting optimised weights are: [0.1588 0.3585 0.4827], which is close to the user ideal ones with the permutation of component [1 3 2]; The strategy designed gives the mixture, shown on the second plot, Fig. 1, which describes the optimised closed-loop behaviour. The last plot contains projections of all three mixtures.

6. CONCLUDING REMARKS

The paper proposes an efficient solution of multi-objective control design problem. The probabilistic mixture modelling complemented by the fully probabilistic control adopting certainty equivalence strategy is used. To express the multiple control objectives, the mixture modelling is employed, where particular objectives are represented by uni-modal mixture components.

The solution of simultaneous design that optimises joint pdfs of the system inputs and pointers to particular components is presented. The obtained solution has been transformed into computationally feasible algorithm for mixtures with components formed by normal auto-regression models with external variable. Comparing to the case when user ideal is described by uni-modal pdf, this solution became more complicated both theoretically and computationally, however, it substantially broads the applicability of the whole approach.

REFERENCES

T. Fergusson (1973). 'A Bayesian analysis of some nonparametric problems'. *The Annals of Statistics* **1**:209–230.

M. Kárný, et al. (2003). 'Mixture-based adaptive probabilistic control'. *International Journal of Adaptive Control and Signal Processing* **17**(2):119–132.

M. Kárný, et al. (2005). *Optimised Bayesian dynamic advising*. Springer Wien. In press

L.-Z. Liao & D. Li (2002). 'Adaptive differential dynamic programming for multiobjective optimal control'. *Automatica* **38**.

V. Peterka (1981). *Trends and Progress in System Identification*, pp. 239–304. Pergamon Press.

M. Rao (1987). *Measure Theory and Integration*. Pure and Applied Mathematics, Wiley.

W. Stadler (ed.) (1988). *Multicriteria Optimisation in Engineering and in the Sciences*, vol. 38. Plenum Press, NY.

H. Toivonen (1989). 'Computer-aided design procedure for multiobjective LQG control problems'. *Int. J.Control* **49**.