Variational Bayes in Distributed Fully Probabilistic Decision Making

Václav Šmídl and Ondřej Tichý,
Institute of Information Theory and Automation,
Pod vodárenskou věží 4, Prague 8, Czech republic,
{smidl,otichy}@utia.cas.cz

Abstract

We are concerned with design of decentralized control strategy for stochastic systems with global performance measure. It is possible to design optimal centralized control strategy, which often cannot be used in distributed way. The distributed strategy then has to be suboptimal (imperfect) in some sense. In this paper, we propose to optimize the centralized control strategy under the restriction of conditional independence of control inputs of distinct decision makers. Under this optimization, the main theorem for the Fully Probabilistic Design is closely related to that of the well known Variational Bayes estimation method. The resulting algorithm then requires communication between individual decision makers in the form of functions expressing moments of conditional probability densities. This contrasts to the classical Variational Bayes method where the moments are typically numerical. We apply the resulting methodology to distributed control of a linear Gaussian system with quadratic loss function. We show that performance of the proposed solution converges to that obtained using the centralized control.

1 Introduction

Complexity of large-scale uncertain systems, such as traffic light signalization in urban areas, prevents effective use of centralized design of control strategy. The technology of multi-agent systems [1] offers technical background how to build a distributed control system. The mainstream multi-agent theory is concerned with deterministic systems for which the majority of results on communication protocols and negotiation strategies are established. As a result, many stochastic problems are converted into deterministic formulation and solved as such. This is typical e.g. in design of distributed traffic light control, where the certainty equivalence assumption is used in all agents [2].

Design methodologies for optimal control strategies of large-scale decentralized stochastic systems are available, e.g. [3], however, the complexity of the decision maker is rather high. In this paper, we propose to design suboptimal (imperfect) decision makers by imposition of additional restrictions within an established centralized design methodology. Specifically, we focus on the theory of Fully Probabilistic Control Design (FPD) [4, 5] for centralized control strategies. This theory is based on minimization of Kullback-Leibler divergence (KLD) [6] and it has been extended to multiple participants using heuristic arguments [7, 8]. An independently developed variant of this approach was used in multi-agent setup in [9]. In this paper, we enforce distribution of control between decision makers via the constraint of conditional independence. Minimization of Kullback-Leibler divergence under this constraint is well known as the Variational Bayes approach [10, 11]. Generalization of these results yields a design methodology of approximate decision makers that are capable to design their own control strategy using probabilistic moments obtained from their neighbors.

We study two computation schemes in this contribution. The first scheme allows unlimited communication with small messages. The second scheme allows much lower number of messages, however, the messages contain much more information than in the first case. In both cases, the Variational Bayes approach is capable to compute approximate results in limited time depending on the number of iterations.

2 Review of Centralized Fully Probabilistic Design

Consider a probabilistic model of a stochastic system

$$y_t \backsim f(y_t|u_t, d_{t-m:t-1}), \tag{1}$$

where symbol $y \backsim f$ denotes that y is a realization from probability density f; vector y_t denotes system output at discrete time t; vector u_t is system input; $d_t = [y_t', u_t']'$ is an aggregation of output and input, where $(.)'$ denotes a transposition of vector or matrix; and $d_{t-m:t-1} = [d_{t-m}, \dots, d_{t-1}]$ is a matrix of the last m observation vectors. Our aim is to design a probabilistic control strategy $f\,(u_t|d_{1:t-1})$ such that the closed loop behavior is as close to the desired behavior as possible.

The Fully Probabilistic Design is based on probabilistic description of the desired behavior represented by the target (ideal) probability density, ${}^If(d_{1:t+h})$, which expresses its aim and constraints. Closeness of the real and the target behavior is measured by the Kullback-Leibler divergence. The optimal control strategy on a horizon of length h is then found recursively for $\tau = t+h, \dots, t+1$,

$$\begin{aligned} {}^of(u_\tau|d_{1:\tau-1}) &= \arg \min_{f(u_\tau|d_{1:\tau-1})} KLD\left[f(d_{t+1:t+h})||{}^If(d_{t+1:t+h})\right], && (2) \\ &= \arg \min_{f(u_\tau|d_{1:\tau-1})} E_{f(d_\tau|d_{1:\tau-1})}\left[\ln \frac{f(d_{t+1:t+h})}{{}^If(d_{t+1:t+h})}\right], && (3) \end{aligned}$$

where $E_{f(x)}(.)$ is the expected value of the argument with respect to probability density $f(x)$; it is abbreviated as $E_{f(x)}(x) \equiv E(x)$ when no confusion can arise. $KLD(.||.)$ is the Kullback-Leibler divergence between the first and the second argument. The optimal solution can be found in the following form, [12]:

$${}^of\,(u_\tau|d_{1:\tau-1}) = {}^If(u_\tau|d_{1:\tau-1}) \frac{\exp[-\omega(u_\tau, d_{1:\tau-1})]}{\gamma(d_{1:\tau-1})}. \tag{4}$$

Here, functions $\omega(.)$ and $\gamma(.)$ are recursively evaluated as

$$\begin{aligned} \omega(u_\tau, d_{1:\tau-1}) &= E_{f(y_\tau|u_\tau, d_{1:\tau-1})}\left(\ln \frac{f(y_\tau|u_\tau, d_{1:\tau-1})}{\gamma(d_{1:\tau}){}^If(y_\tau|u_\tau, d_{1:\tau-1})}\right), && (5) \\ \gamma(d_{1:\tau-1}) &= \int {}^If(u_\tau|d_{1:\tau-1}) \exp[-\omega(u_\tau, d_{1:\tau-1})]\, du_\tau, && (6) \end{aligned}$$

initialized at time $\tau = t+h$ as $\gamma(d_{1:t+h}) = 1$.

2.1 Special case of Linear Quadratic design

Linear Quadratic Gaussian (LQG) control arise as a special case of FPD (4)–(6), when both the model and the target probability densities are Gaussian with linear function of their mean value:

$$\begin{aligned} f(y_t|u_t, d_{1:t-1}) &= \mathcal{N}(\Theta\psi_t, R), && (7) \\ {}^If(y_t, u_t|d_{1:t-1}) &= \mathcal{N}\left(\begin{bmatrix} \overline{y}_t \\ \overline{u}_t \end{bmatrix}, \begin{bmatrix} Q_y & 0 \\ 0 & Q_u \end{bmatrix}\right). && (8) \end{aligned}$$

Here, $\mathcal{N}(\mu, \Sigma)$ denotes Gaussian probability density with mean value μ and covariance Σ; Θ is a matrix of known parameters; ψ_t is a vector composed from an arbitrary combination of elements of $y_{t-m:t-1}$ and $u_{t-m:t}$, and any deterministic transformation of these elements.

Substitution of (8) into (5) at $\tau = t + h$, i.e. $\gamma(d_{1:t+h}) = 1$, yields:

$$\omega(u_\tau, d_{1:\tau-1}) = \frac{1}{2}\left[\ln(Q_y R^{-1}) - n_y + tr(RQ_y^{-1}) + (\Theta\psi_\tau - \overline{y}_\tau)' Q_y^{-1}(\Theta\psi_\tau - \overline{y}_\tau)\right], \quad (9)$$

$$= [\psi'_\tau, 1]\Psi_\tau[\psi'_\tau, 1]' \quad (10)$$

where n_y denotes dimension of vector y_τ. Note that the first three terms in $\omega(.)$ are independent of u_τ and y_τ making them irrelevant to this time step. Evaluation of probability ${}^o f(u_\tau|\phi_\tau)$ from (4) is achieved by reordering the quadratic form in (10) into

$$[\psi'_\tau, 1]\Psi_\tau[\psi'_\tau, 1]' = [u_\tau, \phi'_\tau, 1]\Psi_{\omega,\tau}[u_\tau, \phi'_\tau, 1]', \quad (11)$$

where u_τ was extracted from ψ_τ (the rest of the elements from ψ_τ are in vector with time-delayed values, ϕ_τ, related to the time τ), and $\Psi_{\omega,\tau}$ is composed of the same elements as Ψ_τ in adapted order with respect to vector $[u_\tau, \phi'_\tau, 1]$. Since (8) is independent in y_τ and u_τ, the marginal on u_τ can be written as

$$f(u_\tau|d_{1:\tau-1}) \propto \exp\left(-\frac{1}{2}[u_\tau, \phi'_\tau, 1]\Psi_{u,\tau}[u_\tau, \phi'_\tau, 1]'\right), \quad \Psi_{u,\tau} = \begin{bmatrix} Q_u^{-1} & 0 & Q_u\overline{u}_\tau \\ 0 & 0 & 0 \\ \overline{u}'_\tau Q_u^{-1} & 0 & \overline{u}_\tau Q_u^{-1}\overline{u}_\tau \end{bmatrix}.$$

The joint probability density (4) is then a quadratic form (11) with kernel $\Psi_{f,\tau} = \Psi_{\omega,\tau} + \Psi_{u,\tau}$. The kernel can be decomposed using Cholesky factorization into $\Psi_{f,\tau} = L_\tau L'_\tau$ where lower triangular matrix L_τ is decomposed into $L_\tau = \begin{pmatrix} \Upsilon_\tau & 0 \\ \Omega_\tau & \Lambda_\tau \end{pmatrix}$, with Υ_τ being triangular matrix of the same dimension as u_τ. Probability density (4) has form

$${}^o f(u_\tau|\phi_\tau) = \mathcal{N}(-(\Upsilon'_\tau)^{-1}\Omega_\tau[\phi'_\tau, 1]', (\Upsilon_\tau\Upsilon'_\tau)^{-1}). \quad (12)$$

and the remainder

$$\gamma(d_{1:\tau-1}) = \exp\left(-\frac{1}{2}[\phi'_\tau, 1]\Lambda_\tau\Lambda'_\tau[\phi'_\tau, 1]'\right). \quad (13)$$

The recursion from $\tau = t + h$ to t reveals the same quadratic forms with the exception that there are additional element in $\Psi_{f,\tau}$ from function $\gamma(d_{1:\tau-1})$.

The mean value of (12), i.e. $\hat{u}_\tau = -(\Upsilon'_\tau)^{-1}\Omega_\tau[\phi'_\tau, 1]'$, is equivalent to LQG designed strategy with loss function given by the quadratic form from (9) in exp(.) [4].

3 Distributed FPD via Variational Bayes

Consider a case where (1) describes a complex system, with vector inputs $u_t = [u_{1,t}, \ldots, u_{n,t}]$, where vectors $u_{i,t}$, $i = 1, \ldots, n$ are logically separated so that they represent independent decision makers. Without any additional assumptions on the model (1), solution (2) would be a complex probability density with no guide how to implement it in a distributed way.

As a first step to decentralization of the control strategy, we impose the restriction of conditional independence of control inputs

$$f(u_t|.) = \prod_{i=1}^{n} f(u_{i,t}|.), \forall t. \quad (14)$$

If the solution is in this form, each decision maker can handle its own inputs via $f(u_{i,t}|\cdot)$. The task is to find a way how to design it.

We repeat minimization (3), under constraint (14)

$$\prod_{i=1}^{n} {}^o f(u_{i,\tau}|d_{1:\tau-1}) = \arg\min_{\prod_i f(u_{i,\tau}|\cdot)} E_{f(d_\tau|d_{1:\tau-1})}\left[\ln\frac{f(d_{t+1:t+h})}{{}^I f(d_{t+1:t+h})}\right]. \quad (15)$$

Using the chain rule of probability calculus and definitions (5)–(6) we obtain

$$\prod_{i=1}^{n} {}^o f(u_{i,\tau}|d_{1:\tau-1}) = \arg\min_{\prod_i f(u_{i,\tau}|\cdot)} KLD\left[f(u_\tau|d_{1:\tau-1})||{}^o f(u_\tau|d_{1:\tau-1})\right]. \quad (16)$$

Minimum of (16) is well known from the Variational Bayes method [11] to satisfy the following set of conditions:

$$ {}^{o}f(u_{i,\tau}|d_{1:\tau-1}) \propto \exp\left(E_{f(u_{/i,\tau}|d_{1:\tau-1})}\left[\ln {}^{o}f(u_\tau|d_{1:\tau-1})\right]\right), i = 1,\dots,n. \tag{17} $$

Here, $u_{/i,\tau}$ denotes a subset of elements of vector u_τ without the element $u_{i,\tau}$, i.e. $u_{/i,\tau} = [u_{1,\tau},\dots,u_{i-1,\tau},u_{i+1,\tau},\dots,u_{n,\tau}]$, and $\propto$ is equality up to normalizing constant.

Substitution of (4) into (17) at each step on the horizon, $\tau = t+h,\dots,t+1$, yields the following set of implicit equations for $i = 1,\dots,n$,

$$ {}^{o}f(u_{i,\tau}|d_{1:\tau-1}) \propto \exp\left(E_{f(u_{/i,\tau}|d_{1:\tau-1})}\left(\ln {}^{I}f(u_\tau|d_{1:\tau-1}) - \omega(u_\tau, d_{1:\tau-1})\right)\right), \tag{18} $$

The normalizing constant of (18) is

$$ \gamma_i(d_{1:\tau-1}) = \int \exp\Big(E_{f(u_{/i,\tau}|d_{1:\tau-1})}\big[\ln {}^{I}f(u_\tau|d_{1:\tau-1}) - \omega(u_\tau, d_{1:\tau-1})\big]\Big)\mathrm{d}u_{i,\tau}, \tag{19} $$

hence $\gamma(d_{1:\tau-1})$ required in (5) of the previous step factorizes into $\gamma(d_{1:\tau-1}) = \prod_{i=1}^{n}\gamma_i(d_{1:\tau-1})$.

Typically, set (18) does not have a closed form solution and must be solved iteratively using the iterative VB (IVB) algorithm. It has been shown that the IVB algorithm monotonically decrease the KLD in each iteration and thus converging to a local minimum [13].

Note that $d_{1:\tau-1}$ in $f(u_{i,\tau}|d_{1:\tau-1})$ are symbolic random variables. This contrasts to the typical application of the Variational Bayes where $d_{1:\tau-1}$ are measured data.

3.1 Special case of LQG

For the special case of linear Gaussian system discussed in Section 2.1, the Variational Bayes method [11] is to be applied to Gaussian probability density (12) with logarithm

$$ \ln f(u_\tau|d_{1:\tau-1}) = c - \frac{1}{2}(u_\tau - (\Upsilon'_\tau)^{-1}\Omega_\tau[\phi'_\tau,1]')'(\Upsilon_\tau\Upsilon'_\tau)^{-1}(u_\tau - (\Upsilon'_\tau)^{-1}\Omega_\tau[\phi'_\tau,1]') \tag{20} $$

$$ = c - \frac{1}{2}[u'_\tau,\phi'_\tau,1]\Phi_\tau[u'_\tau,\phi'_\tau,1]', \tag{21} $$

Here, we use the same notation as in the previous section for ϕ_τ, Υ_τ, and Ω_τ, $c = \ln|\Upsilon_\tau|$ which is independent of control action u_τ, and Φ_τ is the kernel of quadratic form (21). For simplicity, we consider partitioning $u_\tau = [u_{1,\tau}, u_{2,\tau}]'$, generalization to n partitions is straightforward. The expected value of (20) with respect to $f(u_{2,\tau}|d_{1:\tau-1})$ is again a quadratic form

$$ E_{f(u_{2,\tau}|d_{1:\tau-1})}(\ln f(u_\tau|d_{1:\tau-1})) = E_{f(u_{2,\tau}|d_{1:\tau-1})}\left(c - \frac{1}{2}[u_{2,\tau},\zeta_\tau]\begin{bmatrix}\Phi_{uu,\tau} & \Phi_{u\zeta,\tau}\\ \Phi_{\zeta u,\tau} & \Phi_{\zeta\zeta,\tau}\end{bmatrix}[u'_{2,\tau},\zeta_\tau]'\right) \tag{22} $$

$$ = c - \frac{1}{2}[E_{f(u_{2,\tau}|d_{1:\tau-1})}(u_{2,\tau}\Phi_{uu,\tau}u_{2,\tau}) + \zeta_\tau\Phi_{\zeta u,\tau}E(u_{2,\tau}) + E(u_{2,\tau})\Phi_{u\zeta,\tau}\zeta_\tau + \zeta_\tau\Phi_{\zeta\zeta,\tau}\zeta_\tau] \tag{23} $$

$$ = c - \frac{1}{2}\zeta_\tau\overline{\Phi}_{u_1,\tau}\zeta_\tau, \tag{24} $$

where $\Phi_{uu,\tau}, \Phi_{u\zeta,\tau}, \Phi_{\zeta\zeta,\tau}$ are composed of elements of Φ_τ restructured to match the new decomposition of the $[u_\tau,\phi'_\tau,1]$ to $u_{2,\tau}$, $\zeta_\tau = [u_{1,\tau},\phi'_\tau,1]$ and $\overline{\Phi}_{u_1,\tau}$ is given by reordering to match the quadratic form in ζ_τ.

Note that (24) is equivalent to (10) and the control law can be obtained using the same derivation that lead to (12). In this case

$$ f(u_{1,\tau}|d_{1:\tau-1}) = \mathcal{N}(Q_{1,\tau}[\phi'_\tau,1]',\sigma_{1,\tau}), \tag{25} $$

$$ f(u_{2,\tau}|d_{1:\tau-1}) = \mathcal{N}(Q_{2,\tau}[\phi'_\tau,1]',\sigma_{2,\tau}), \tag{26} $$

Algorithm 1 DP-VB variant of the distributed control design.

Off-line:

Choose control horizon h, target probability density ${}^I f\left(d_{1:t+h}\right)$, and initial value of $f^{(0)}(u_{i,\tau}|\cdot)$ for each decision maker $i = 1 \dots n$.

On-line:

At each time t, for each decision-maker i, do:

1. For each $\tau = t + h, t + h - 1, \dots, t$ do
   (a) Start negotiation with counter $j = 1$, and initial guess $f^{(0)}(u_{i,\tau}|\cdot)$,
   (b) Compute moments required by the neighbors and communicate them,
   (c) Compute jth value of control strategy $f^{(j)}(u_{i,\tau}|\cdot)$ using moments obtained from the neighbors,
   (d) If the strategy convergence is not reached and $j < j_{IVB}$, increase j and goto (a), stop otherwise.
2. Apply designed control action $u_{i,t}$ from the converged strategy.

where $\sigma_{i,\tau}$ is given using Cholesky decomposition of $\overline{\Phi}_{u_i,\tau}$ in the same form as in (12) and the second line follows from equivalent derivation for $u_{2,\tau}$. Now, we can formulate the necessary moments for substitution into (22):

$$E_{f(u_{i,\tau}|d_{1:\tau-1})}(u_{i,\tau}) = Q_{i,\tau}[\phi'_\tau, 1]', \tag{27}$$

$$E_{f(u_{i,\tau}|d_{1:\tau-1})}(u_{i,\tau}\Phi_{uu,\tau}u'_{i,\tau}) = Q_{i,\tau}[\phi'_\tau, 1]\Phi_{uu,\tau}[\phi'_\tau, 1]'Q'_{i,\tau} + \Phi_{uu,\tau}\sigma_{i,\tau}. \tag{28}$$

This finalizes the list of results that are necessary to run the IVB algorithm in each time step of the horizon $\tau = t + h, \dots, t + 1$. This variant will be denoted as **DP-VB** algorithm, Algorithm 1.

3.2 Alternative evaluations

Note that the set of conditions (18) has to be met for each time of the horizon, τ. Put together, we may interpret it as a set of $n \times (h + 1)$ conditions of optimality. If the control strategies $f(u_{i,\tau}|\cdot)$ were conditionally independent from $f(u_{i,\tau-1}|\cdot)$, then the iterations could be performed in any order and still guaranteed to converge to a local minimum. This would be a great property since it would allow asynchronous communication between the decision makers, and guarantee robustness against lost messages. However, this is not automatically guaranteed due to dependence $f(u_{i,\tau}|u_{i,\tau-1})$. Therefore, a change of order of the time index can lead to an increase of the KL divergence within one iteration due to inaccurate $\gamma(d_{1:\tau-1})$ from (19). However, similar difficulty arise in the case of on-line variational Bayes and the convergence is still guaranteed by means of stochastic approximations [14]. The only drawback is slower convergence in comparison to the standard IVB algorithm. We conjecture that it is also the case in our approach.

If our conjecture holds, then we may change the order of dynamic programming and IVB iterations. Specifically, each decision maker first exchange messages about expected values $[Q_{i,t}, \dots, Q_{i,t+h}, \sigma_{i,t}, \dots, \sigma_{i,t+h}]$ with its neighbors and then designs its strategy using backward evaluation (5), see Algorithm 2 for details. The new moments are send to the neighbors for the next iterations. This algorithm will be denoted as **VB-DP**.

4 Example

Consider the following 3-output 2-input system:

$$f(y_t|\psi_t, \Sigma) = \mathcal{N}(\Theta\psi_t, \Sigma_y), \tag{29}$$

where

$$y_t = [y_{1,t}, y_{2,t}, y_{3,t}]', \qquad \psi_t = [y_{1,t-1}, y_{2,t-1}, y_{3,t-1}, u_{1,t}, u_{2,t}, u_{1,t-1}, u_{2,t-1}]',$$

Algorithm 2 VB-DP variant of the distributed control design.

Off-line:

Choose control horizon h, target probability density ${}^{I}f\left(d_{1:t+h}\right)$, and initial value of $f^{(0)}(u_{i,\tau}|\cdot)$ for each decision maker $i = 1 \dots n$.

On-line:

At each time t, for each decision-maker i, do:

1. Start negotiation with counter $j = 1$, and $f^{(0)}(u_{i,\tau}|\cdot)$.
2. Compute jth value of control strategy $f^{(j)}(u_{i,\tau}|\cdot)$ on the whole horizon $\tau = t+h, \dots t$ using moments obtained from the neighbors, evaluate moments required by the neighbors and communicate them,
3. If the strategy convergence is not reached and $j < j_{IVB}$, increase j and goto 2, stop otherwise.
4. Apply designed control action $u_{i,t}$ from the converged strategy.

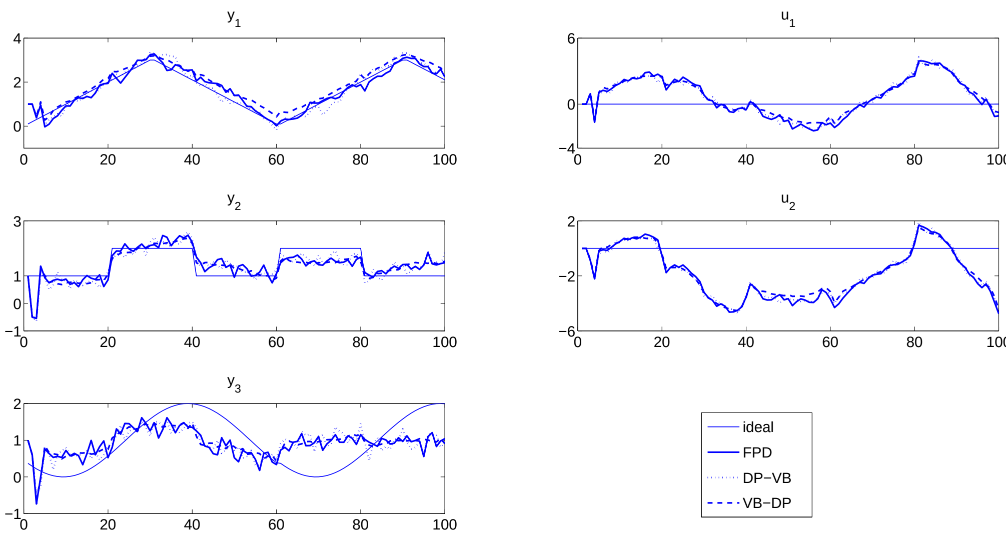

Figure 1: Example run of the controlled system. The target values for the inputs and the outputs are displayed in thin full line. The typical realization of outputs and inputs for all tested algorithms are also displayed for illustration.

$$\Theta = \begin{bmatrix} 0.8 & 0.2 & 0 & -0.3 & 0.4 & 0 & 0 & 0 \\ -0.2 & 0.5 & -0.8 & 0.2 & 0.5 & -0.2 & -0.5 & 0 \\ 0 & 1.1 & -0.5 & 0 & 0 & -0.2 & 0.3 & 0 \end{bmatrix}. \tag{30}$$

The target probability densities are

$${}^{I}f(y_t) = \mathcal{N}\left(\begin{bmatrix} \overline{y}_{1,t} \\ \overline{y}_{2,t} \\ \overline{y}_{3,t} \end{bmatrix}, \begin{bmatrix} 0.01 & & \\ & 0.01 & \\ & & 0.01 \end{bmatrix}\right), \quad {}^{I}f(u_t) = \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 100 & \\ & 100 \end{bmatrix}\right), \tag{31}$$

with values of $\overline{y}_{1,t}, \overline{y}_{2,t}, \overline{y}_{3,t}$ displayed in Fig 1 (solid lines). The choice of diagonal covariance matrices in (31) allows the convergence of algorithms from Section 3 to the centralized solution, Section 2.

Three control strategies were tested:

FPD: as the centralized strategy (Section 2.1),

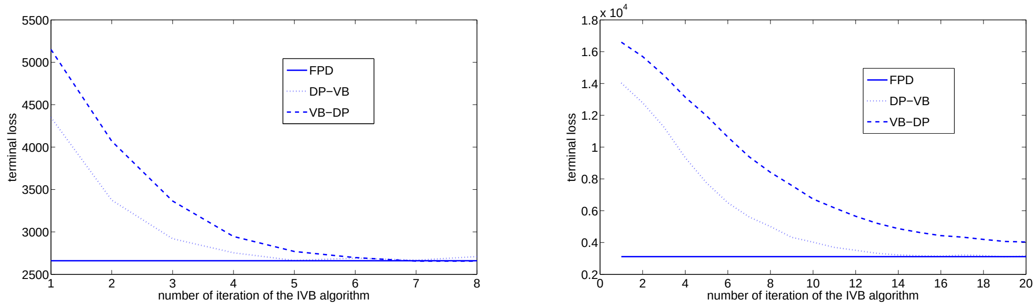

Figure 2: Convergence of the terminal loss (the sum of differences from target values) of the decentralized DP-VB and VB-DP algorithms to the terminal loss of the centralized FPD solution as a function of the number of IVB iterations for two variants of system parameters.

DP-VB: decentralized evaluation of the FPD control via multiple VB algorithms, one at each time τ on the horizon (Section 3.1),

VB-DP: decentralized evaluation of the FPD control via a single the VB algorithm on the whole horizon (Section 3.2).

A comparative Monte Carlo study with 15 runs of the system with parameter (30) was performed to establish convergence of the decentralized strategy design to the centralized one. An example run of the controlled system is shown in Fig. 1. Results of the study are displayed in Fig. 2 via dependence of the terminal loss on the number of iterations of the IVB algorithm, j_{IVB}. Note that the results converged to the centralized solution after a few iterations; the full convergence is allowed using diagonal covariance matrices in (31). As expected, the DP-VB variant converges faster than the VB-DP. Suitability of each strategy than depends on the quality of communication between agents. The VB-DP algorithm may be attractive especially for systems with higher latency in communication.

The difference is even more visible on a more demanding system with parameters

$$\Theta = \begin{bmatrix} 0.8 & 0.2 & 0.5 & -0.3 & 0 & 0.4 & 0 & 0 \\ -0.2 & 0.5 & -0.2 & 0.2 & -0.2 & 0.5 & -0.5 & 0 \\ 0.5 & 1.1 & -0.5 & 0 & -0.2 & 0 & 0.3 & 0 \end{bmatrix}. \tag{32}$$

The results of the same Monte Carlo experiment for the new value of parameter Θ are displayed in Fig. 2, right. While the DP-VB algorithm reaches performance of the centralized FPD after 14 iterations, the VB-DP algorithms requires more than 20 iterations to converge. The number of iterations required to reach the centralized solution is rather high, since the IVB algorithm was initialized with $f^{(0)}(u_{i,t}|\cdot) = {}^I f(u_{i,t})$ for both variants. The purpose of this choice was to verify if the algorithm converges to the correct solution even from poor initial conditions.

5 Conclusion

The presented methodology for design of approximate decision makers is based on fully probabilistic control and decentralization is achieved by imposing conditional independence between control inputs. The general method yields two principle outputs: (i) an iterative algorithm that is known to systematically decrease the loss function, and (ii) the moments that needs to be exchanged to achieve optimum performance. Under the condition of diagonal covariance matrices of target probability densities, the simulation results suggest that the decentralized control is able to reach the same performance as the centralized one. This was achieved at the price of all decision makers having full model of the system and intensive negotiation with high volume of communication. We have shown in simulation that the intensity of communication can be lowered by an alternative order of evaluation and communication. The Variational Bayes approach can cope with limited computational time, the quality of the solution depends on the number of iterations in the IVB algorithm.

Further simplifications can be achieved by imposing additional restrictions (e.g. in the form of conditional independence) on the solution.

Acknowledgment

Support of grants MŠMT 1M0572 and TAČR TA01030603 is gratefully acknowledged.

References

[1] G. Weiss, ed., *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. The MIT Press, 2000.

[2] R. Negenborn, B. De Schutter, and J. Hellendoorn, "Multi-agent model predictive control for transportation networks: Serial versus parallel schemes," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 3, pp. 353–366, 2008.

[3] M. Hutter, *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*. Berlin: Springer, 2005.

[4] M. Kárný, "Towards fully probabilistic control design," *Automatica*, vol. 32, no. 12, pp. 1719–1722, 1996.

[5] M. Kárný and T. Kroupa, "Axiomatisation of fully probabilistic design," *Information Sciences*, 2011. DOI=10.1016/j.ins.2011.09.018.

[6] S. Kullback and R. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, pp. 79–87, 1951.

[7] V. Šmídl and J. Andrýsek, "Merging of multistep predictors for decentralized adaptive control," in *Proc. of the American Control Conference*, (Seattle, USA), 2008.

[8] V. Šmídl, "On adaptation of loss functions in decentralized adaptive control," in *Proc. of the 12th IFAC symposium on Large Scale Systems*, (Villeneuve d'Ascq, France), 2010.

[9] B. van den Broek, W. Wiegerinck, and B. Kappen, "Graphical model inference in optimal control of stochastic multi-agent systems," *J. of Artificial Intelligence Research*, vol. 32, no. 1, pp. 95–122, 2008.

[10] Z. Ghahramani and M. Beal, "Propagation algorithms for variational Bayesian learning," *Advances in Neural Information Processing Systems*, vol. 13, pp. 507–513, 2001.

[11] V. Šmídl and A. Quinn, *The Variational Bayes Method in Signal Processing*. Springer, 2005.

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

[13] S. Amari, S. Ikeda, and H. Shimokawa, "Information geometry of α-projection in mean field approximation," in *Advanced Mean Field Methods* (M. Opper and D. Saad, eds.), (Cambridge, Massachusetts), The MIT Press, 2001.

[14] M. Sato, "Online model selection based on the variational Bayes," *Neural Computation*, vol. 13, pp. 1649–1681, 2001.