Bayesian Filtering for States Uniformly Distributed on a Parallelotopic Support

Ladislav Jirsa
The Czech Academy of Sciences
Institute of Information Theory and Automation
Prague, Czech Republic
jirsa@utia.cas.cz

Lenka Kuklišová Pavelková
The Czech Academy of Sciences
Institute of Information Theory and Automation
Prague, Czech Republic
pavelkov@utia.cas.cz

Anthony Quinn
Czech Acad Sci
Inst Inform Th & Autom
Prague, Czech Republic
& Trinity College Dublin
Dublin, Ireland
aquinn@utia.cas.cz

***Abstract*—This paper contributes to the literature on Bayesian filtering in the case where the processes driving the states and observations are uniformly distributed on finite intervals. We introduce the class of uniform distributions on parallelotopic supports (UPS). We derive optimal local distributional projections (i.e. approximations) within this UPS class—in the sense of minimum Kullback-Leibler divergence—of the outputs of the data and time updates of filtering. We demonstrate that the UPS class provides a tighter approximation (and therefore more precise inferences) than a previously reported approximation on orthotopic supports. It does this, while still achieving bounded complexity in the resulting recursive filtering algorithm. The comparative performance of the UPS-closed filtering algorithm is explored—via both Bayesian and frequentist performance measures—as a function of signal-to-noise ratio and state dimension in a position-velocity system.**

***Index Terms*—Bayesian filtering, uniform distribution on a parallelotopic support (UPS), local approximation, Kullback-Leibler divergence**

I. Introduction

Stochastic filtering [1], [2]—i.e. the sequential estimation of hidden (system) states via noisy observations and known inputs/controls [1]—has many applications in contexts where either the states, or the observations, or both, are constrained to particular sets. There is significant current interest in these constrained filtering applications, particularly in problems of fault detection [3], [4] and robust model-predictive control [3], [5], and in applications involving constrained dynamics in physical processes [6]–[8] and distributed estimation in resource-constrained wireless sensor networks [9], for instance.

Deterministic approaches for processing the sequential knowledge constraints propose so-called unknown-but-bounded approaches [5], [9], [10], and investigate classes of constrained sets within which set membership of the states can be guaranteed [3], [11]–[14], notably convex polytopes and their specializations (to zonotopes and orthotopes), and ellipsoids (and their specialization to hyperspheres).

These geometric considerations are essential in Bayesian filtering involving constrained (typically uniformly distributed) stochastic state (i.e. process) and observation drivers (i.e. noise processes) [4], [7], [15]. The key advantages of Bayesian filtering are (i) the opportunity to take account of the distribution (often non-uniform, e.g. the truncated Gaussian pdf [8]) of the states within their constrained support sets, (ii) the quantification of uncertainty in the states [2], and (iii) formal procedures for optimal (in the minimum-Bayes'-risk sense) sequential estimation and decision-making, including control design [16].

In the authors' recent work, the relaxation of the Kalman filtering assumptions of linearity in the state and observation transition kernels, and known parameterization have been adopted, but, now, specifically in the case of uniformly distributed and white noise processes, yielding sequential state inferences optimally approximated—in the sense of minimum Kullback-Leibler divergence (KLD) [17]—within the class of uniform distributions on orthotopic supports [18]. The framework has been extended to yield the consistent observation predictor [19] and this has been used as the basis for probabilistic knowledge transfer between constrained and uniformly approximated Bayesian nodes [20]. The key advantage of distributional closure within the class of uniform distributions on orthotopic supports is that the number of degrees of freedom in this distribution—and, so, the number of statistics which must be sequentially computed—is quadratic in the dimension of the state variables, ensuring a tractable and computationally efficient recursive algorithm. Nevertheless, the adoption of orthotopic supports is very restrictive, leading to significant accumulation of distributional error, particularly when approximations are performed locally at each step of filtering.

In this paper, we relax the class into which we sequentially project the state inferences by proposing the uniform distribution on a parallelotopic support (UPS), arguing that the degrees-of-freedom remain polynomial in the number of state dimensions. This preserves the tractability of the resulting sequential Bayesian filtering algorithm, while attaining much tighter approximation to the true distributions.

In Section II, we introduce the UPS class of distributions,

This research has been supported by GAČR grant 18-15970S.

978-1-7281-5341-4/19/$31.00 ©2019 IEEE

based on the geometry of the parallelotope, emphasizing issues of degrees-of-freedom and measure that are important in the design and properties of the subsequent Bayesian filtering algorithm. We also derive the optimal projections (i.e. approximations, in the minimum-KLD sense) of more complicated distributions on constrained supports. The latter emerge from the data and time steps of Bayesian filtering, as shown in Section III, and so we deploy these optimal UPS projections there, summarizing the resulting recursive filtering algorithm. In Section IV, we present experiments on uniformly-driven position-velocity systems of increasing dimension, demonstrating the improved tracking accuracy attained by the UPS algorithm in comparison to the cruder orthotopically constrained variant previously reported in [19].

The following notation is used. Matrices are in capital letters (e.g. A), vectors and scalars are in lowercase letters (e.g. b). A_{ij} is the element of a matrix A on i-th row and j-th column. A_i denotes the i-th row of A. ℓ_z denotes the length of a (column) vector z and $\mathbb{Z}$ means a set of z. $\|z\|_2$ is the Euclidean norm of z. Also, $\|z\|_\infty = \max_i |z_i|$, $i = 1, \dots, \ell_z$, is the H-∞ norm of z, $\equiv$ means equality by definition. Note that no notational distinction is made between a random variable and its realisation. z_t is the value of a column vector z at a discrete time instant $t \in \mathbb{T} \equiv \{1, 2, \dots, \bar{t}\}$, being typically a random process realisation and $z_{t;i}$ is the i-th entry of z_t; $z(t) \equiv \{z_t, z_{t-1}, \dots, z_1\}$. The symbol $f(\cdot|\cdot)$ denotes a conditional probability density function of known but unspecified type (pdf); names of arguments distinguish respective pdfs.

II. UPS class of parametric models

Conventionally, the multivariate uniform distribution is considered on a rectangular (orthotopic) support. In Bayesian filtering, we need to move beyond this restriction as the orthotopic support is too conservative. Not least, it implies independence between the coordinates. We therefore introduce the multivariate uniform distribution on a parallelotopic support (UPS class), which is more flexible and yet tractable in Bayesian filtering.

In the paper, we aim to develop a sequential Bayesian filtering algorithm within this UPS class. Unfortunately, we are not able to propagate the support and functional form of the involved pdfs within exact Bayesian filtering, i.e. the UPS class is not closed under the data and time updates of Bayesian filtering.

Therefore, we propose optimal local projections (approximations) of non-uniform and non-parallelotopic distributions into the UPS class.

A. Polytopes and their specializations

We consider a finite-dimensional vector random variable z with realisations in a bounded subset of R^{ℓ_z}. We now define appropriate support sets in a ℓ_z-dimensional space.

A *polytope* is a bounded set defined (bounded) by a finite number of flat facets. In this paper, we specialise this to the following types of convex polytope (Fig. 1).

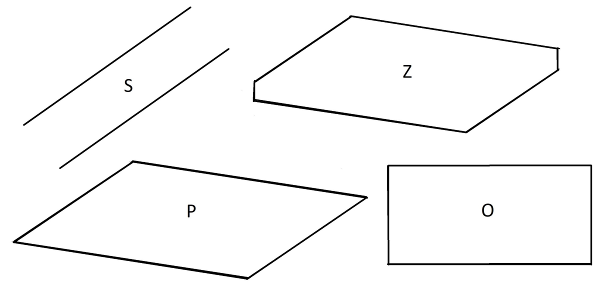

Fig. 1. A two-dimensional strip (S) and two-dimensional convex polytopes: zonotope (Z), parallelotope (P) and orthotope (O).

TABLE I
THE DEGREES OF FREEDOM (DOF) AND LEBESQUE MEASURE, I.E. VOLUME, $\mathcal{V}$ OF VARIOUS CONVEX POLYTOPE SPECIALISATIONS.

	dof	volume $\mathcal{V}$
orthotope	$2\ell_z$	$\prod_{i=1}^{\ell_z} (\bar{z}_i - \underline{z}_i)$
parallelotope	$\ell_z(\ell_z + 2)$	$\lvert \det V \rvert^{-1} \prod_{i=1}^{\ell_z} (b_i - a_i)$, see [18]
zonotope	$k(k+2)$, $k > \ell_z$	the sum of the $\mathcal{V}$s of its generating parallelotopes, see [21]

A *zonotope* $\mathbb{Z}_Z$ is a convex polytope formed by the intersection of k strips, $k \geq \ell_z$. It can be expressed as

$$\mathbb{Z}_Z = \{z : a \leq Vz \leq b\}, \tag{1}$$

where a and b are vectors of length k, of lower and upper bounds, respectively, which are meant entry-wise; V is a matrix of size $k \times \ell_z$ with rank ℓ_z. The i-th *strip* is therefore given by the inequality

$$\mathbb{Z}_{S_i} = \{z : a_i \leq V_i z \leq b_i\}. \tag{2}$$

A *parallelotope* $\mathbb{Z}_P$ is a special case of a zonotope (1) with $k = \ell_z$, so that V is a square invertible matrix.

An *orthotope* $\mathbb{Z}_O$ is a special case of the parallelotope with adjacent facets orthogonal. It then holds that $V = I$ in (1), where I denotes an identity matrix. Furthermore, $a = \underline{z}$ and $b = \bar{z}$, being the lower and upper bounds of z, respectively. Then, the orthotope is specified by

$$\mathbb{Z}_O = \{z : \underline{z} \leq z \leq \bar{z}\}, \tag{3}$$

Table I compares the above defined support sets from the point of view of their volumes — i.e. the Lebesque measure — and degrees-of-freedom (dof). The dof corresponds to the minimal number of geometric parameters that unambiguously defines the mentioned set.

Besides the standard form (1), we introduce another two equivalent descriptions of parallelotope $\mathbb{Z}_P$ that will be used further. The $[-\mathbf{1}, \mathbf{1}]$-form of parallelotope equivalent to standard form (1) is defined

$$\mathbb{Z}_P = \{z : -\mathbf{1}_{(\ell_z)} \leq Wz - c \leq \mathbf{1}_{(\ell_z)}\}, \tag{4}$$

where $\mathbf{1}_{(\ell_z)}$ is a unit vector of length ℓ_z and

$$W_{ij} = \frac{2V_{ij}}{b_i - a_i}, \qquad c_i = \frac{b_i + a_i}{b_i - a_i}, \tag{5}$$

$i, j = 1, \ldots, \ell_z$. We can transform it back to the standard form (1) using $a = c - \mathbf{1}_{(\ell_z)}$, $b = c + \mathbf{1}_{(\ell_z)}$, $V = W$.

An expression for the parallelotope (1) in terms of its centroid $\hat{z}$ is [24]

$$\mathbb{Z}_P = \{z : z = \hat{z} + T\xi\}, \tag{6}$$

where $T = W^{-1}$, $\hat{z} = Tc$, $\forall \xi$ s.t. $\|\xi\|_\infty \leq 1$.

B. UPS class

We define a uniform distribution of z on a parallelotopic support (1), i.e. the UPS distribution, as

$$\mathcal{U}_z(a, b, V) \equiv \mathcal{V}^{-1} \chi_z \left(a \leq Vz \leq b\right) \tag{7}$$

where $\mathcal{V}$ is given in the second row of Table I, and $\chi_z(\mathbb{Z})$ is the set indicator, which equals 1 if $z \in \mathbb{Z}$ and 0 otherwise.

In the case of orthotopic support (3), i.e. for $V = I$, we simplify the notation to

$$\mathcal{U}_z(a, b) \equiv \mathcal{U}_z(a, b, I). \tag{8}$$

The first moment (mean value) of the UPS distribution (7) is

$$\mathsf{E}[z|a, b, V] \equiv \hat{z} = V^{-1} \ \frac{b + a}{2}. \tag{9}$$

The second central moment (covariance) of the UPS is

$$\mathsf{cov}[z|a, b, V] = \frac{1}{3} \ V^{-1} G G' \big(V^{-1}\big)', \tag{10}$$

where $G_{ii} = \frac{b_i - a_i}{2}$ and $G_{ij} = 0$, $i \neq j$, $i, j = 1, \ldots, \ell_z$. For details on moments see [22].

C. Projection into the UPS class

1) Bayes-optimal projection/approximation: We aim to approximate an original pdf g by a simpler pdf f. The pdf f is to be a projection of g on a properly selected set $\mathbb{F}$ of feasible pdfs. According to [23], minimisation of Kullback-Leibler divergence (KLD) [17] gives, in a Bayesian sense, the optimal approximation of a pdf, E means an expectation,

$$f^O \in \mathrm{Arg}\min_{f \in \mathbb{F}} D(g||f) = \mathrm{Arg}\min_{f \in \mathbb{F}} E_g \ln\left(\frac{g}{f}\right). \tag{11}$$

Below, we propose (i) a functional approximation of a non-uniform pdf on a bounded support by the uniform pdf and (ii) approximation of the zonotopic support of a uniform pdf by a parallelotopic support.

2) Non-uniform to uniform distribution: Given pdf on a bounded support, $g(z) = k(z)\chi_z(\mathbb{Z}_g)$, $0 < k(z) < +\infty$ $\forall z \in \mathbb{Z}_g$, we search for the optimal approximation of $g(z)$ by a uniform pdf, $f(z) = \mathcal{V}^{-1}\chi_z(\mathbb{Z}_f)$, $\mathcal{V} = \mathrm{vol}(\mathbb{Z}_f)$, using (11). Then, function arguments omitted, $D(g||f) = \int k \ln \frac{\mathcal{V} k \chi_g}{\chi_f} \,\mathrm{d}z = \int_{\mathbb{Z}_g} k \ln k \,\mathrm{d}z + \int_{\mathbb{Z}_g} k \ln \frac{\chi_g}{\chi_f} \,\mathrm{d}z + \ln \mathcal{V} \int_{\mathbb{Z}_g} k \,\mathrm{d}z$. The first term is independent of f, the second term is finite (zero) if $\mathbb{Z}_g \subset \mathbb{Z}_f$. The third term depends on f through $\mathcal{V}$: the larger support of f, the higher $\mathcal{V}$. Hence, to minimise KLD, we minimise the measure of $\mathbb{Z}_f$ choosing $\mathbb{Z}_f = \mathbb{Z}_g$, i.e. $\hat{f}_f$ is to be the uniform pdf on the support of g.

3) Transformation of a parallelotope: Let us express a parallelotope in the direct form (6) as $\mathbb{Z} = \{z : z = \hat{z} + T_z\xi\}$, $\|\xi\|_\infty \leq 1$. Consider a linear transformation $w = Mz + m$ where M is an invertible matrix, m is a vector. Then, the parallelotope set $\mathbb{W}$ corresponding to w is

$$\mathbb{W} = \{w : w = \hat{w} + T_w\xi\}, \ \|\xi\|_\infty \leq 1 \tag{12}$$

where $\hat{w} = M\hat{z} + m$, $T_w = MT_z$.

4) Expansion of the parallelotope: Consider a random variable z defined on a parallelotopic set $\mathbb{Z} = \{z : a \leq Vz \leq b\}$ (1) and random variable w, $w = z + e$, where e is defined on an orthotopic set $\mathbb{E} = \{e : -\sigma \leq e \leq \sigma\}$.

Then, w is defined on a zonotope given by the Minkowski sum $\oplus$ of the parallelotope and the orthotope [3]

$$\mathbb{W} = \{z : a \leq Vz \leq b\} \oplus \{-\sigma \leq e \leq \sigma\} \tag{13}$$

5) Approximation of the zonotope by a parallelotope: We consider a polytope that corresponds to the intersection of $\ell_z + k$ strips (2), $a_i \leq V_i z \leq b_i$, $i = 1, \ldots, \ell_z + k$, ℓ_z corresponds to the dimension of z, $k \in$ N. We aim to obtain the smallest parallelotope that contains the above mentioned polytope. For this purpose, we use the adapted algorithm from [24] as follows: (i) Consider a parallelotope given by ℓ_z defined strips. (ii) Add another one strip to the parallelotope and *tighten* all these $\ell_x + 1$ strips to remove redundancy, i.e. all strips are narrowed and/or shifted so that their intersection is unchanged. (iii) Discard one strip of these $\ell_x + 1$ strips so that the intersection of the remaining ℓ_x strips has minimal volume. (iv) Repeat the procedure for all remaining $k - 1$ strips. For details, see [18].

III. LSU FILTERING WITHIN THE UPS CLASS

In this section, Bayesian filtering is summarised and the state space model with uniformly distributed noise terms (LSU) model is introduced. Then, state filtering within the UPS class is proposed.

A. Bayesian filtering and the LSU model

In the considered Bayesian set up [16], the system is described by the following pdfs:

$$\begin{aligned} \text{prior pdf:} \quad & f\left(x_1\right) \\ \text{observation model:} \quad & f\left(y_t | x_t\right) \\ \text{time evolution model:} \quad & f\left(x_{t+1} | x_t, u_t\right) \end{aligned} \tag{14}$$

where y_t is a scalar observable output, u_t is a known system input (optional, for generality), and x_t is an ℓ_x-dimensional unobservable system state, $t \in \mathbb{T}$.

We assume that (i) the hidden process x_t satisfies the Markov property, (ii) no direct relationship between input and output exists in the observation model, and (iii) the inputs consist of a known sequence $u_1, \ldots, u_{\bar{t}}$.

Bayesian filtering, i.e, state estimation, consists of the evolution of the posterior pdf $f(x_t|d(t))$ where $d(t)$ is a sequence of observed data records $d_t = (y_t, u_t)$, $t \in \mathbb{T}$. The evolution of $f(x_t|d(t))$ is described by a two-steps recursion that starts from the prior pdf $f(x_1)$) and ends with the data update at the final time $t = \bar{t}$:

- Data update (Bayes rule)

$$f(x_t|d(t)) = \frac{f(y_t|x_t)f(x_t|d(t-1))}{\int_{x_t^*} f(y_t|x_t)f(x_t|d(t-1))\mathrm{d}x_t}, \quad (15)$$

- Time update

$$f(x_{t+1}|d(t)) = \int_{x_t^*} f(x_{t+1}|u_t, x_t)f(x_t|d(t))\,\mathrm{d}x_t. \quad (16)$$

A linear state space model with a uniform noise (LSU model) is defined as

$$\begin{aligned} f(y_t|x_t) &= \mathcal{U}_y(\tilde{y}_t - r, \tilde{y}_t + r) \quad (17) \\ f(x_{t+1}|x_t, u_t) &= \mathcal{U}_x(\tilde{x}_{t+1} - \rho, \tilde{x}_{t+1} + \rho) \end{aligned}$$

where $\tilde{y}_t = Cx_t$, $\tilde{x}_{t+1} = Ax_t + Bu_t$, A, B, C are the known model matrices/vectors of appropriate dimensions, $\nu_t \in (-\rho, \rho)$ is the uniform state noise with known parameter ρ, $n_t \in (-r, r)$ is the uniform observation noise with known parameter r.

State estimation for the LSU model (17), according to (15) and (16), leads to a very complex form of posterior pdf. In [18], [19], an approximate Bayesian state estimation of this model is proposed. That algorithm is based on minimising the KLD (11) of two pdfs and provides the evolution of the approximate posterior pdf $f(x_t|d(t))$ by the time and data update steps. The pdf $f(x_t|d(t))$ is uniformly distributed on a parallelotopic support. Nevertheless, to close the recursion, the parallelotopic support is circumscribed by an orthotope before the next time update step. In this paper, we extend our previous results on approximated Bayesian filtering with uniform noise and closure on an orthotopic support to the case uniform distribution with closure on a parallelotopic support, i.e. we propose a UPS-closed recursion without the above mentioned circumscription.

B. Approximate Bayesian filtering within UPS class

We consider the system (14) with observation model and state evolution model (17) and with prior pdf $f(x_1)$. Performing (15) and (16) expels the posterior pdf from the UPS class. We use the results of Section II to re-admit it.

Approximate data update: The data update (15) processes $f(x_t|d(t-1))$ together with the $f(y_t|x_t)$ (17) according to the Bayes rule. It starts in $t = 1$ with $f(x_1) = \mathcal{U}_x(a_1^+, b_1^+, M_1^+)$. The exact pdf is uniformly distributed on a zonotopic support that results from the intersection of a parallelotope (7) obtained during previous time update—or $f(x_1)$ in the first step—and strips (8) given by new data

$$f(x_t|d(t)) \propto \mathcal{U}_x(a_t^+, b_t^+, M_t^+)\mathcal{U}_{y_t}(Cx_t - r, Cx_t + r) \propto$$

$$\propto \chi\left(\begin{bmatrix} a_t^+ \\ y_t - r \end{bmatrix} \le \begin{bmatrix} M_t \\ C \end{bmatrix} x_t \le \begin{bmatrix} b_t^+ \\ y_t + r \end{bmatrix}\right). \quad (18)$$

We approximate (18) by a pdf uniformly distributed on a parallelotopic support, see Sec. II-C5. Then, the approximate pdf takes the form

$$f(x_t|d(t)) \approx \mathcal{U}_x(a_t, b_t, M_t). \quad (19)$$

Approximate time update: The time update (16) processes $f(x_t|d(t))$ (19) together with $f(x_{t+1}|x_t, u_t)$ (17). The exact pdf $f(x_{t+1}|d(t))$ is non-uniformly distributed on a zonotopic support. It has a linear piecewise shape with shaping parameters ρ, a_t and b_t. We approximate it by the uniform pdf, see Sec. II-C2. The support of $f(x_{t+1}|d(t))$ is computed in two steps. Firstly, the support $\mathbb{X}_t$ of $f(x_t|d(t))$ (19) is transformed according to the deterministic part, i.e. $\tilde{x}_{t+1}$, of (17). For this, the parallelotope $\mathbb{X}_t$ of form (1) is converted into the form (6) and then the linear transformation $\tilde{x}_{t+1} = Ax_t + Bu_t$ is performed according to Sec. II-C3. The resulting support, $\tilde{\mathbb{X}}_{t+1}^+$, is then transformed back to the form (1). Secondly, the parallelotope $\tilde{\mathbb{X}}_{t+1}^+$ is expanded by the set $[-\rho, \rho]$ which corresponds to the stochastic part of (17). The resulting support, $\mathbb{X}_{t+1}^+$, corresponds to the Minkowski sum of $\tilde{\mathbb{X}}_{t+1}^+$ and the set $[-\rho, \rho]$, see II-C4 which is a zonotope (1).

We project the above mentioned uniform pdf with support $\mathbb{X}_{t+1}^+$ into UPS class, see Sec. II-C5. Then, a_{t+1}^+, b_{t+1}^+ and M_{t+1}^+ are derived and the approximate pdf has the form

$$f(x_{t+1}|d(t)) \approx \mathcal{U}_x(a_{t+1}^+, b_{t+1}^+, M_{t+1}^+). \quad (20)$$

IV. Experiments

In this section, we design simulations around a position-velocity system [2] of increasing dimension n and report some key Bayesian and frequentist performance measures to compare the performance of the proposed LSU filtering under UPS and orthotopically supported local approximations [18]. The system matrix A of size $2n \times 2n$ is

$$\begin{aligned} A_{ii} &= 1, \quad i = 1, \dots, 2n, \\ A_{i,i+n} &= \Delta t, \quad i = 1, \dots, n, \end{aligned} \quad (21)$$

otherwise 0. The sampling period $\Delta t = 1$ in our case. The matrix $B = 0$ (i.e. a system without input u_t) and C of size $n \times 2n$ is $C_{ii} = 1, \quad i = 1, \dots, n$, otherwise 0. The dimensions $\ell_y = n$ and $\ell_x = 2n$.

As a key operating parameter, the ratio ρ/r (17) was chosen as a proxy for signal-to-noise ratio. We use two relative performance measures (subscripts o and p mean orthotope and parallelotope, respectively):

- volume ratio $\mathcal{V}_\mathrm{o}/\mathcal{V}_\mathrm{p}$, where the volumes are specified in Table I (Bayesian measure),
- total norm-squared error (TNSE) ratio, $\mathrm{TNSE}_\mathrm{o}/\mathrm{TNSE}_\mathrm{p}$ (frequentist measure), where

$$\mathrm{TNSE} = \sum_{t=1}^{\bar{t}} \|\hat{x}_t - x_t\|_2^2.$$

As performance mesures, we also use these probabilities:

- the nesting probability, p_n, whether the parallelotopic support of $f(x_t|d(t))$ is a proper subset of the orthotopic support,

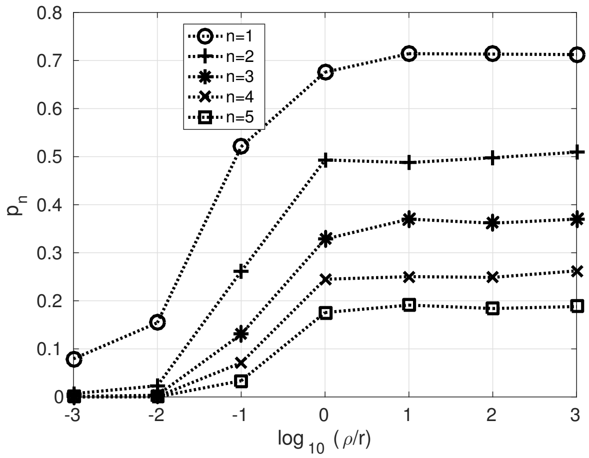

Fig. 2. Probability that the parallelotopic support (UPS class) is nested within the orthotopic support, i.e. whether the parallelotope is a proper subset of the orthotope.

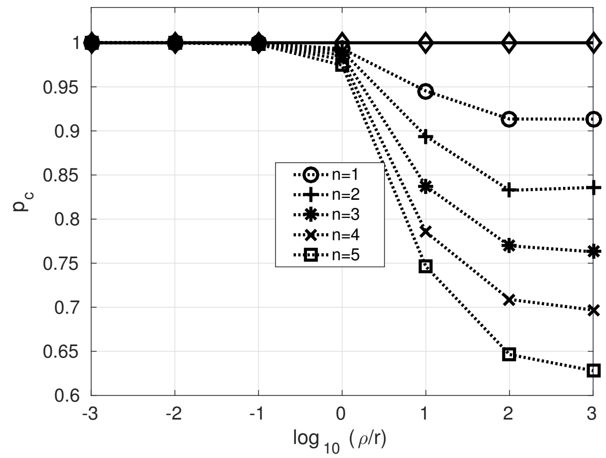

Fig. 3. Probability of containing the true state vector within the support of the posterior UPS (dotted lines). The solid line shows these probabilities for the orthotopically approximated filtering distribution, which are invariant with dimension n.

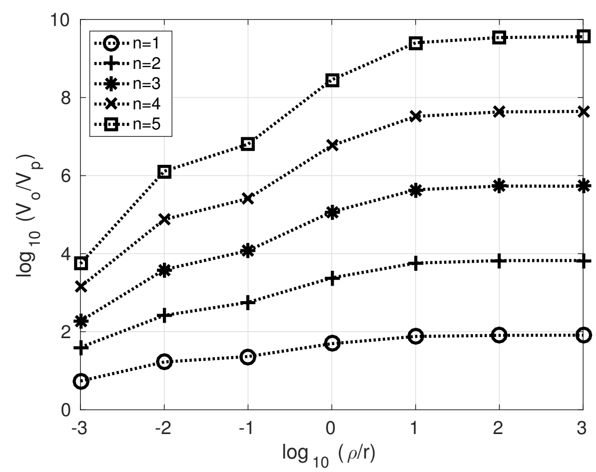

Fig. 4. Volume ratio, $\mathcal{V}_o/\mathcal{V}_p$, of the orthopic to parallelotopic supports.

- the probability, p_c, that the support of the UPS-closed filtering distribution, $f(x_t|d(t))$, contains the true state.

For every operating condition, we took 100 Monte Carlo runs for the UPS and orthotopic algorithms, and reported the average per operating condition for each algorithm.

A. *Performance comparison*

Firstly, we wanted to check whether—at a fixed ρ/r—the results depend on r itself. It was observed that the dependence is insignificant for a wide range of r and n, and therefore ρ/r can be used as the operating parameter.

In the following figures, we display the performance measures defined above. In all cases, the higher the performance measure, the better the relative performance of our new UPS-closed algorithm versus the formerly reported orthotopically closed variant [20].

After each filtering step, t, we computed the nesting probability, p_n. The situation is shown in Fig. 2. The p_n decreases with decreasing ρ/r and with increasing dimension n.

Fig. 3 shows the probability p_c. We show that this probability depends on n in a manner similar to Fig. 2. However, its dependence on ρ/r is in opposition to the findings in Fig. 2. Fig. 4 shows the ratio of the support volumes (Table I), and Fig. 5 shows the ratio TNSEs. The volume ratio increases with ρ/r. Note also that it increases exponentially with n (note the log scale). In Fig. 5, the ratio $\mathrm{TNSE_o}/\mathrm{TNSE_p}$ is almost invariant with dimension, n. While the strongly monotonically improving relative performance of of the UPS-closed filtering algorithm—compared to the orthotopically-closed variant—is again observed as a function of increasing ρ/r, we notice two anomalies: (i) the UPS-closed TNSE is the greater one for $\rho/r = 10^{-3}$; and (ii) the monotonicity is lost around $\rho/r = 10^{-1}$.

B. *Discussion*

The greatly reduced volume ratios for the UPS-closed algorithm versus the orthotopically closed variant point to far greater precision in the UPS state inferences (Fig. 4). This comes at the cost of slowly decreasing containment probability, particularly in high dimensions (Fig. 3). This dimension sensitivity is corroborated by the decreasing nesting probability with increasing dimension (Fig. 2). This may be caused by the mechanism of approximating zonotopes by parallelotopes, i.e. discarding strips with respect to minimal volume. This procedure can shift the resulting parallelotope.

Variable nesting probability can be caused by a similar effect. More research is required to understand these behaviours more formally, and, indeed, the fact that the effect of increasing ρ/r is contrary for p_n versus p_c.

Nevertheless, the frequentist performance of the posterior mean estimate is greatly enhanced in our novel UPS-closed filtering algorithm compared to the earlier variant, as seen in Fig. 5. It is interesting to note that this performance enhancement seems to be relatively robust to increasing

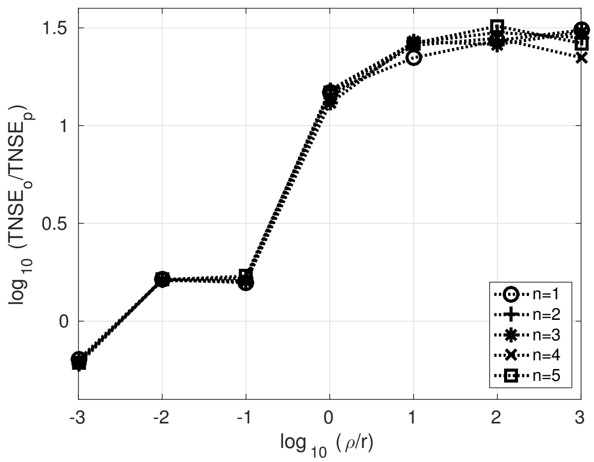

Fig. 5. Ratio of TNSEs for the orthopically-closed to the UPS-closed filtering algorithms.

state dimension. Again, more work is needed to understand the detailed performance comparison in stressful regime.

V. Concluding remarks

Recall that the projection of the zonotopic into the parallelotopic support in Section II-C was proposed in a strictly geometric sense; i.e. the optimality of the approximation in terms of distribution (11) has not been demonstrated. This may explain the effects notes in Section IV-B. In future work, we hope to explore these considerations.

Our algorithm enforces closure of both substeps of Bayesian filtering within the UPS class. This sequential application of local approximation brings with it the risk of unbounded propagation of distributional error with t. We can apply a forgetting operation in each step of approximate Bayesian filtering [25]. Another approach is to combine the two UPS projections into a single one at each step.

In all events, this paper clearly points to the benefits for LSU filtering of adopting UPS as the distributional invariant. As well as the Bayesian and frequentist performance enhancements reported here, UPS captures correlation between state variables, something that is lost in the orthotopically supported variant.

Our future work will focus on the derivation of the data predictor consistent with this UPS-closed filtering algorithm, for application in knowledge transfer between LSU filters. This will advance the preliminary solution—based on orthotopically supported distributions—reported in [20].

References

[1] A. M. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic Press, 1970.

[2] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

[3] J. K. Scott, D. M. Raimondo, G. R. Marseglia, and R. D. Braatz, "Constrained zonotopes: A new tool for set-based estimation and fault detection," *Automatica*, vol. 69, pp. 126 – 136, 2016.

[4] C. Combastel, "An extended zonotopic and Gaussian Kalman filter merging set-membership and stochastic paradigms: Toward non-linear filtering and fault detection," *Annual Reviews in Control*, vol. 42, pp. 232 – 243, 2016.

[5] U. Sharma, S. Thangavel, A. R. G. Mukkula, and R. Paulen, "Effective recursive parallelotopic bounding for robust output-feedback control," *IFAC Papersonline*, vol. 51, no. 15, pp. 1032–1037, 2018.

[6] D. Simon and D. L. Simon, "Constrained Kalman filtering via density function truncation for turbofan engine health estimation," *International Journal of Systems Science*, vol. 41, pp. 159–171, 2010.

[7] L. Lang, W. Chen, B. R. Bakshi, P. K. Goel, and S. Ungarala, "Bayesian estimation via sequential Monte Carlo sampling – Constrained dynamic systems," *Automatica*, vol. 43, no. 9, pp. 1615–1622, 2007.

[8] X. Shao, B. Huang, and J. M. Lee, "Constrained Bayesian state estimation - A comparative study and a new particle filter based approach," *Journal of Process Control*, vol. 20, no. 1, pp. 143–157, 2010.

[9] X. Ge, Q. Han, and Z. Wang, "A dynamic event-triggered transmission scheme for distributed set-membership estimation over wireless sensor networks," *IEEE Transactions on Cybernetics*, vol. 49, no. 1, pp. 171–183, 2019.

[10] M. Pourasghar, V. Puig, and C. Ocampo-Martinez, "Interval observer versus set-membership approaches for fault detection in uncertain systems using zonotopes," *International Journal of Robust and Nonlinear Control*, vol. 29, no. 10, pp. 2819–2843, 2019.

[11] Y. Wang, V. Puig, and G. Cembrano, "Set-membership approach and Kalman observer based on zonotopes for discrete-time descriptor systems," *Automatica*, vol. 93, pp. 435 – 443, 2018.

[12] Y. K. Foo, Y. C. Soh, and M. Moayedi, "Linear set-membership state estimation with unknown but bounded disturbances," *International Journal of Systems Science*, vol. 43, no. 4, pp. 715–730, 2012.

[13] L. Zhang, X. Liu, and X. Kong, "State estimators for uncertain linear systems with different disturbance/noise using quadratic boundedness," *Journal of Applied Mathematics*, 2012.

[14] L. Chisci, A. Garulli, and G. Zappa, "Recursive state bounding by parallelotopes," *Automatica*, vol. 32, no. 7, pp. 1049–1055, 1996.

[15] R. Fletcher, *Practical Methods of Optimization*. John Wiley & Sons, 2000.

[16] 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.

[17] S. Kullback and R. Leibler, "On information and sufficiency," *The Annals of Statistic*, vol. 22, pp. 79–87, 1951.

[18] L. Pavelková and L. Jirsa, "Approximate recursive Bayesian estimation of state space model with uniform noise," in *Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, Porto, Portugal, 2018, pp. 388–394.

[19] L. Jirsa, L. Pavelková, and A. Quinn, "Approximate Bayesian prediction using state space model with uniform noise," in *Informatics in Control Automation and Robotics*, ser. Lecture Notes in Electrical Engineering, O. Gusikhin and K. Madani, Eds. Springer, 2018, in press.

[20] ——, "Knowledge transfer in a pair of uniformly modelled Bayesian filters," in *Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, Prague, Czech republic, 2019, in press.

[21] E. Gover and N. Krikorian, "Determinants and the volumes of parallelotopes and zonotopes," *Linear Algebra and its Applications*, vol. 433, no. 1, pp. 28 – 40, 2010.

[22] L. Jirsa, "Linear ARX and state-space model with uniform noise: computation of first and second moments," Institute of Information Theory and Automation, Tech. Rep. 2357, 2016. [Online]. Available: http://library.utia.cas.cz/separaty/2016/AS/jirsa-0463816.pdf

[23] J. M. Bernardo, "Expected information as expected utility," *The Annals of Statistic*, vol. 7, no. 3, pp. 686–690, 1979.

[24] A. Vicino and G. Zappa, "Sequential approximation of feasible parameter sets for identification with set membership uncertainty," *IEEE Transactions on Automatic Control*, vol. 41, no. 6, pp. 774–785, 1996.

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