Information Value in Nonparametric Dirichlet-Process Gaussian-Process (DPGP) Mixture Models

Hongchuan Wei, Wenjie Lu, Pingping Zhu, Silvia Ferrari,
Miao Liu, Robert H. Klein, Shayegan Omidshafiei, Jonathan P. How

Abstract

This paper presents tractable information value functions for Dirichlet-process Gaussian-process (DPGP) mixture models obtained via collocation methods and Monte Carlo integration. Quantifying information value in tractable closed form is key to solving control and estimation problems for autonomous information-gathering systems. The properties of the proposed value functions are analyzed and then demonstrated by planning sensor measurements so as to minimize the uncertainty in DPGP target models that are learned incrementally over time. Simulation results show that sensor planning based on expected KL divergence outperforms algorithms based on mutual information, particle filters, and randomized methods.

Key words: Information theory; Bayesian nonparametric models; Gaussian process; Dirichlet process; information gain.

1 Introduction

Bayesian nonparametric models, such as the Dirichlet-process Gaussian-process (DPGP) mixtures, have been recently developed for modeling multiple dynamic processes adaptively from data. When data becomes available over time, DPGP clusters and parameters are expanded or compacted incrementally, as needed, to avoid growing the model dimensionality indefinitely as the size of the database increases [17]. Because of these characteristics, DPGP mixtures are particularly useful in lifelong learning and sensing problems, and present the opportunity for planning the measurement sequence so as to optimize the value of future data.

To date, DPGPs have been shown effective at modeling traffic patterns [9], clinical identification [19], and gene expression time series analysis [8]. Although information theoretic functions have become a common approach for representing information value in sensing and control problems [12], they are not directly applicable to DPGPs because they are defined in terms of finite-dimensional probability distributions [17]. Previous approaches to quantifying information value in GPs consist of inferring the process state and, then, determining the differential entropy or mutual information between the state and the measurements [11]. Similar work in [14] considered the mutual information between the state of a moving target and available measurements using a GP model. Previous work by the authors proposed using the Kullback-Leibler (KL) divergence obtained by approximating the GPs by probability distributions about a set of collocation points for a given DP prior [22, 24]. However, these GP-based methods are not applicable to the DPGP mixture models considered in this paper.

Building on the collocation point method in [22], this paper presents a systematic approach for obtaining tractable information value functions, such as the expected KL divergence, for DPGP models of multiple dynamic targets. New conditional and expected information theoretic functions for DPGPs are derived in Sections 4.2-4.3 to quantify information value in non-myopic sequential observations. New theoretical results (Section 4.4) are used to obtain a computationally efficient approximation of DPGP information value via Monte Carlo integration. The new analysis in Section 4.4 proves that this approximation is an unbiased estimator of the original DPGP value function, and is characterized by an error covariance that decreases linearly with the number of samples. As a result, it can be used to control moving sensors in real time while observing dynamic targets. Simulation results show that using this approximate DPGP KL divergence for sensor planning

⋆ This paper was not presented at any IFAC meeting. This work was supported by ONR MURI Grant N000141110688. Corresponding author: S. Ferrari (ferrari@cornell.edu).

Email addresses: {hw88, wl72}@duke.edu (Hongchuan Wei, Wenjie Lu), {pz224, ferrari}@cornell.edu (Pingping Zhu, Silvia Ferrari), {miaoliu, rhklein, shayegan, jhow}@mit.edu (Miao Liu, Robert H. Klein, Shayegan Omidshafiei, Jonathan P. How).

is far more effective than using DPGP mutual information, particle filtering [21], or randomized methods [7].

2 Problem Formulation and Assumptions

DPGP mixtures have been recently shown to be a new and useful paradigm for modeling multiple dynamic processes from data, as exemplified by traffic patterns modeled from measurements obtained by a helicopter flying over the greater Boston area [9]. The purpose of the information functions derived and analyzed in this paper is to provide design objectives for controlling moving sensors such that their measurement value is optimized over time. Consider the problem of modeling N unknown and independent targets in a workspace $\mathcal{W} \subset \mathbb{R}^2$, using multiple sensor measurements obtained over time. Assume each target can be described by a time-invariant nonlinear ordinary differential equation (ODE),

$$\dot{\mathbf{x}}_j(t) = \mathbf{f}_i[\mathbf{x}_j(t)] \triangleq \mathbf{v}_j(t), \quad j = 1, \ldots, N(t) \tag{1}$$

where the vector function $\mathbf{f}_i : \mathbb{R}^2 \to \mathbb{R}^2$, referred to as a *velocity field* (VF), is drawn from a set of continuously differentiable velocity fields, $\mathcal{F} = \{\mathbf{f}_1, \ldots, \mathbf{f}_M\}$, $\mathbf{x}_j \in \mathcal{W}$ denotes the jth target position, $\mathbf{v}_j \in \mathbb{R}^2$ denotes the jth target velocity, and $t \in [t_0,\ t_f]$. The number of targets N in $\mathcal{W}$ is unknown, time varying, and may not equal M.

At the initial time, t_0, no prior target information is available and, thus, M and $\mathcal{F}$ are to be learned from data. The data consists of position and velocity measurements,

$$\mathbf{m}_j(t) \triangleq \begin{bmatrix} \mathbf{y}_j(t) \\ \mathbf{z}_j(t) \end{bmatrix} = \begin{bmatrix} \mathbf{x}_j(t) \\ \mathbf{v}_j(t) \end{bmatrix} + \mathbf{n}(t),\ j = 1, \ldots, N(t) \tag{2}$$

for $\mathbf{x}_j(t) \in \mathcal{S}(t)$, $t > t_0$, where $\mathcal{S}(t) \subset \mathcal{W}$ is the moving sensor field-of-view (FoV), and $\mathbf{n} \in \mathbb{R}^4$ is an additive Gaussian measurement error with zero mean and known standard deviations σ_x and σ_v [25]. If $\mathbf{x}_j(t) \notin \mathcal{S}(t)$, $\mathbf{m}_j(t)$ belongs to an empty set. The measurement-target association is carried out by established data association algorithms [1, 2, 15], and any errors are assimilated in the DPGP model uncertainty representation.

From (1), it can be seen that a VF, $\mathbf{f}_i$, projects a target position, $\mathbf{x}_j$, to a velocity, $\mathbf{v}_j$, and, thus, can be viewed as a two-dimensional spatial phenomenon. Because $\mathbf{f}_i$ is unknown and is not necessarily drawn uniformly at random from $\mathcal{F}$, the association of a target with a VF in $\mathcal{F}$ can be represented by a discrete random variable, G_j, with a range $\mathcal{I} = \{1, \cdots, M\}$ that is possibly infinite. Then, the event $\{G_j = i\}$ represents the association of target j with the velocity field $\mathbf{f}_i \in \mathcal{F}$, as shown in (1), and the N targets can be modeled as a multioutput GP via nonlinear regression, as follows.

A multioutput GP defines a multivariate distribution over functions, $P(\mathbf{f}_i)$, where $\mathbf{f}_i : \mathcal{W} \to \mathbb{R}^n$, and here $n = 2$ and $i = 1, \ldots, M$ [9]. Let $F = \{\mathbf{f}_i(\mathbf{x}_1), \ldots, \mathbf{f}_i(\mathbf{x}_N) \mid \mathbf{x}_N \in \mathcal{W}\}$ be a set of vector function values evaluated at N points in $\mathcal{W}$. Then, $P(\mathbf{f}_i)$ is a multioutput GP if for any finite subset $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ the marginal distribution $P(F)$ is a joint multivariate Gaussian distribution [18]. A multioutput GP is specified by its mean vector function,

$$\boldsymbol{\theta}_i(\mathbf{x}_j) = \mathbb{E}_{\mathbf{v}_j}[\mathbf{f}_i(\mathbf{x}_j)], \quad \forall\ \mathbf{x}_j \in \mathcal{W} \tag{3}$$

and its covariance matrix function,

$$\boldsymbol{\Psi}_i(\mathbf{x}_j, \mathbf{x}'_j) = \mathbb{E}_{\mathbf{v}_j}\left\{[\mathbf{f}_i(\mathbf{x}_j) - \boldsymbol{\theta}_i(\mathbf{x}_j)][\mathbf{f}_i(\mathbf{x}'_j) - \boldsymbol{\theta}_i(\mathbf{x}'_j)]^T\right\} \tag{4}$$

for any $\mathbf{x}_j, \mathbf{x}'_j \in \mathcal{W}$, where $\mathbb{E}_{\mathbf{v}_j}[\cdot]$ denotes the expectation operator with respect to $\mathbf{v}_j$ [18]. Assuming stationary covariance functions, the notation,

$$\mathbf{f}_i(\mathbf{x}_j) \sim \mathrm{GP}_i\left(\boldsymbol{\theta}_i(\mathbf{x}_j), \boldsymbol{\Psi}_i(\mathbf{x}_j, \mathbf{x}'_j)\right), \quad \forall \mathbf{x}_j, \mathbf{x}_j \in \mathcal{W} \tag{5}$$

or in short $\mathbf{f}_i(\mathbf{x}_j) \sim \mathrm{GP}_i$, can be used to indicate that $\mathbf{f}_i$ is 'distributed as" the Gaussian process GP_i, for $i = 1, \ldots, M$. For simplicity, in this paper it is assumed that the elements of $\mathbf{v}_j$ are independent, such that $\boldsymbol{\Psi}_i$ is a diagonal and positive-definite matrix. Also, it is assumed that the M Gaussian processes, $\mathrm{GP}_1, \ldots, \mathrm{GP}_M$, share the same covariance matrix $\boldsymbol{\Psi}_i = \boldsymbol{\Psi}$ and, therefore, $\mathrm{GP}_i = \mathrm{GP}(\boldsymbol{\theta}_i, \boldsymbol{\Psi})$, where $\boldsymbol{\Psi}$ is known *a priori*.

For N independent targets, the target-VF association variables, $G_1, \ldots, G_N$, can be assumed to be independent and identically distributed (i.i.d.), such that the probability of event $\{G_j = i\}$ is

$$P(G_j = i) = \pi_i, \quad \forall j \tag{6}$$

Letting $\boldsymbol{\pi} \triangleq [\pi_1\ \cdots\ \pi_M]^T$ denote a prior probability vector that satisfies the properties, $\sum_{i=1}^{M} \pi_i = 1$, and $\pi_i \in [0,\ 1]$, $\forall i \in \mathcal{I}$, the prior distribution of G_j can be represented by an M-dimensional categorical distribution $\mathrm{Cat}(\boldsymbol{\pi})$, with probability mass function (PMF) $\boldsymbol{\pi}$ [4]. Then, the tuple $\{\mathcal{F}, \boldsymbol{\pi}\}$ provides sufficient statistics for modeling all N targets from data.

3 DPGP Target Modeling and Planning

When multiple targets obey the same VF, a prior distribution on $\boldsymbol{\pi}$ is required to provide the ability of learning the clustering of target behaviors from sensor measurements. Dirichlet processes (DPs) have been successfully applied to the clustering of data into an unknown number of clusters because they allow for the creation and deletion of clusters, as necessary, while new data is obtained over time. Let (A, B) be a measurable space, where B is a σ-algebra on a set A [4]. A finite measurable partition $\{B_i\}_{i=1}^{n}$ of A is a collection of sets $B_i \in B$, such

Preprint submitted to Automatica
Received February 27, 2016 13:19:09 PST

that $B_i \cap B_j = \emptyset$, if $i \neq j$; and $\cup_{i=1}^{n} B_i = A$. Let H be a finite non-zero measure on the measurable space (A, B), and let α be a positive real number. A DP with parameters H and α, denoted by $\mathrm{DP}(\alpha, H(A))$, is the distribution of a random probability measure P if, for any finite measurable partition $\{B_i\}_{i=1}^{n}$ of A, the following holds,

$$[P(B_1) \ \cdots \ P(B_n)]^T \sim \mathrm{Dir}\big(\alpha H(B_1), \ldots, \alpha H(B_n)\big) \quad (7)$$

where "Dir" denotes the Dirichlet distribution [20].

From Section 2, the base distribution H is a GP with zero mean, or $\mathrm{GP}_0 = \mathrm{GP}(\mathbf{0}, \boldsymbol{\Psi})$, and A is the space of continuously differentiable functions $C^1(\mathcal{W})$. Then, the following DPGP mixture [9],

$$\begin{aligned} \{\boldsymbol{\theta}_i, \boldsymbol{\pi}\} &\sim \mathrm{DP}(\alpha, \mathrm{GP}_0), \ i = 1, \ldots, \infty \\ G_j &\sim \mathrm{Cat}(\boldsymbol{\pi}), \ j = 1, \ldots, N \\ \mathbf{f}_{G_j}(\mathbf{x}) &\sim \mathrm{GP}(\boldsymbol{\theta}_{G_j}, \boldsymbol{\Psi}), \ \mathbf{x} \in \mathcal{W}, \ j = 1, \ldots, N, \end{aligned} \quad (8)$$

can be used to model the dynamic targets (1) from sensor measurements (2) that are obtained incrementally over a period of time $[t_0, \ t_f]$.

When new sensor measurements become available, (8) can be updated to make use of the latest target information for tracking, pursuit, or estimation. Let Δt denote a small but finite constant time interval required by the sensor to obtain and process new target measurements using a finite sampling rate. Then, $\mathbf{m}_j(k)$ denotes sensor measurements obtained from the jth target during the kth sampling interval, $[t_k, \ t_k + \Delta t)$. Since the sensor FoV is stationary during sampling, $\mathbf{m}_j(k)$ is known provided $\mathbf{x}_j(t) \in \mathcal{S}(k)$ for some $t \in [t_k, \ t_k + \Delta t)$. Because DPGP cluster learning depends on all targets observed over time, all measurements obtained up to time k, denoted by $M(k) \triangleq M_1(k) \cup \cdots \cup M_N(k)$, where $M_j(k) \triangleq \{\mathbf{m}_j(\ell) \mid 1 \leq \ell \leq k\}$, are used for the DPGP update at time k. Assume the DP prior is fixed during Δt. Then, the VF clusters can remain unchanged and the VF-target associations learned at time k, $\mathcal{G}(k) \triangleq \{G_j(k) \mid 1 \leq j \leq N\}$, are also part of the measurement database, $Q(k) \triangleq \{M(k), \mathcal{G}(k)\}$, used to update (8) at time k. Furthermore, from $Q(k)$, new sensor measurements can be planned at t_{k+1}, t_{k+2}, ..., so as to minimize the DPGP model uncertainty in future updates.

4 Information Value in DPGP Mixture Models

This section presents a novel approach for evaluating information value in DPGP mixture models so that the reduction in uncertainty brought about by new sensor measurements can be estimated from past data and a prior DPGP model. Evaluating information value in closed form is important both for assessing the benefit of a given database and for planning sensor motions and parameters. In order to derive information functions for the DPGP in (8), the M VFs in (1) are evaluated at a finite number of collocation points in the domain of integration $\mathcal{W}$ of the ODEs (1). The approach is illustrated for the KL divergence because it is typically most effective for solving sensor planning problems [10,22,26], but can easily be extended to other information theoretic functions, such as mutual information (Section 5).

4.1 DPGP KL Divergence

Let $p(x)$ and $q(x)$ denote two known probability density functions (PDFs) of a continuous random variable $x \in \mathbb{R}$. Then, the KL divergence or relative entropy,

$$D\big(p(x) \parallel q(x)\big) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} dx \quad (9)$$

can be used to represent the "distance" between $p(x)$ and $q(x)$. Although it does not constitute a true distance metric because it is nonadditive, nonsymmetric, and does not obey the triangle inequality, the KL divergence can represent the change in a PDF, e.g., due to a new measurement or observation [6].

Let $\boldsymbol{\xi}_l \in \mathcal{W}$ denote the lth collocation point chosen from a uniform grid of L points in $\mathcal{W}$ grouped in a $2L \times 1$ vector $\boldsymbol{\xi} = [\boldsymbol{\xi}_1^T \ \ldots \ \boldsymbol{\xi}_L^T]^T$, as in basic collocation methods [13]. Each ODE in (1) is discretized about every collocation point by evaluating $\mathbf{f}_i$ at $\boldsymbol{\xi}_l$ for $l = 1, \ldots, L$, such that,

$$\boldsymbol{\upsilon}_i \triangleq [\mathbf{f}_i(\boldsymbol{\xi}_1)^T \quad \cdots \quad \mathbf{f}_i(\boldsymbol{\xi}_L)^T]^T = \boldsymbol{\upsilon}_i(\boldsymbol{\xi}) \quad (10)$$

approximates $\mathbf{f}_i$ in $\mathcal{W}$. Similarly, all M velocity fields in $\mathcal{F}$ are discretized and grouped into a $2LM \times 1$ vector of random velocity variables $\boldsymbol{\upsilon} \triangleq [\boldsymbol{\upsilon}_1(\boldsymbol{\xi})^T \quad \cdots \quad \boldsymbol{\upsilon}_M(\boldsymbol{\xi})^T]^T$.

Now, let $p(\boldsymbol{\upsilon})$ and $q(\boldsymbol{\upsilon})$ denote two joint PDFs of the $2LM$ elements of the random velocity vector $\boldsymbol{\upsilon}$, where each PDF is obtained from the DPGP model (8). Then, the "distance" between the two DPGP parameterizations can be represented by the DPGP KL divergence

$$D\big(p(\boldsymbol{\upsilon}) \parallel q(\boldsymbol{\upsilon})\big) \triangleq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\boldsymbol{\upsilon}) \ln \frac{p(\boldsymbol{\upsilon})}{q(\boldsymbol{\upsilon})} \, d\boldsymbol{\upsilon} \quad (11)$$

In order to evaluate (11), the joint PDF of $p(\boldsymbol{\upsilon})$ needs to be expressed in terms of the DPGP parameters in (8). Given $\{\boldsymbol{\pi}, \boldsymbol{\theta}_i, \boldsymbol{\Phi}\}$, the random vector $\boldsymbol{\upsilon}_i$ has a mixed multivariate Gaussian distribution with mean and covariance matrix calculated from $\boldsymbol{\theta}_i$ and $\boldsymbol{\Phi}$. However, the number of components in the mixture model is infinite. Therefore, computing (11) is computationally very expensive due to the multiple integrals of the joint PDFs. Thus, the next subsections present several steps by which the KL divergence can be simplified and, then, approximated by an information function that is computation-

Preprint submitted to Automatica
Received February 27, 2016 13:19:09 PST

ally efficient, using conditional independence assumptions and Monte Carlo integration.

4.2 Conditional DPGP KL Divergence

For a non-myopic process, the information value is to be conditioned on all prior measurements or observations. Thus, as a first step, the KL divergence is obtained in terms of conditional probability densities that take into account the measurement database $Q(k)$ available or *given* at time t_k. The reduction in DPGP model uncertainty brought about by a measurement $\mathbf{m}_j(k+1)$ can be represented by the conditional KL divergence,

$$\begin{aligned} &D\big(\boldsymbol{\upsilon};\mathbf{m}_j(k+1)|Q(k)\big) \triangleq \\ &\quad D\left(p\big(\boldsymbol{\upsilon}|\mathbf{m}_j(k+1),Q(k)\big)\parallel p\big(\boldsymbol{\upsilon}|Q(k)\big)\right) \end{aligned} \tag{12}$$

in terms of the joint PDFs of $\boldsymbol{\upsilon}$ obtained from (8). Because the VF vectors of two targets are conditionally independent given $Q(k)$, the joint PDF is factorized as,

$$p(\boldsymbol{\upsilon}|Q(k))=p(\boldsymbol{\upsilon}_1,\ldots,\boldsymbol{\upsilon}_M|Q(k))=\prod_{i=1}^{M}p(\boldsymbol{\upsilon}_i|Q(k)) \tag{13}$$

where, each joint PDF $p(\boldsymbol{\upsilon}_i|Q(k))$ can be obtained from the GP covariance and mean as follows.

From GP regression [18], given a database $Q(k)$, $\boldsymbol{\upsilon}_i$ is characterized by the multivariate joint Gaussian PDF,

$$p(\boldsymbol{\upsilon}_i|Q(k)) = \frac{-\frac{1}{2}[\boldsymbol{\upsilon}_i-\boldsymbol{\mu}_i(k)]^T[\boldsymbol{\Sigma}_i(k)]^{-1}[\boldsymbol{\upsilon}_i-\boldsymbol{\mu}_i(k)]}{\sqrt{(2\pi)^L|\boldsymbol{\Sigma}_i(k)|}}$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ are the mean vector and covariance matrix of the measurements associated with the ith VF, respectively, and $|\cdot|$ denotes the matrix determinant. Both $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ can be computed from the cross-covariance matrix of the GP regression,

$$\begin{aligned} \boldsymbol{\Phi}(\mathbf{X},\mathbf{X}') &\triangleq \mathbb{E}_{\boldsymbol{\upsilon}_i,\boldsymbol{\upsilon}_i'}\left[\big(\boldsymbol{\upsilon}_i-\mathbb{E}[\boldsymbol{\upsilon}_i]\big)\big(\boldsymbol{\upsilon}_i'-\mathbb{E}[\boldsymbol{\upsilon}_i']\big)^T\right] \\ &= \begin{bmatrix} \boldsymbol{\Psi}(\mathbf{x}_1,\mathbf{x}_1') & \cdots & \boldsymbol{\Psi}(\mathbf{x}_1,\mathbf{x}_n') \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}(\mathbf{x}_m,\mathbf{x}_1') & \cdots & \boldsymbol{\Psi}(\mathbf{x}_m,\mathbf{x}_n') \end{bmatrix} \end{aligned} \tag{14}$$

where $\mathbf{X}=[\mathbf{x}_1^T\ \cdots\ \mathbf{x}_m^T]^T$ and $\mathbf{X}'=[(\mathbf{x}_1')^T\ \ldots\ (\mathbf{x}_n')^T]^T$ for any $\mathbf{x}_j,\mathbf{x}_l'\in\mathbb{R}^2$. From (10), $\boldsymbol{\upsilon}_i\triangleq[\mathbf{f}_i(\mathbf{x}_1)^T\ \cdots\ \mathbf{f}_i(\mathbf{x}_m)^T]^T$ and $\boldsymbol{\upsilon}_i'\triangleq[\mathbf{f}_i(\mathbf{x}_1')^T\ \cdots\ \mathbf{f}_i(\mathbf{x}_n')^T]^T$ are obtained by evaluating $\mathbf{f}_i$ at every point in $\mathbf{X}$ and $\mathbf{X}'$, respectively.

Now, let $\mathbf{Y}_i(k)=[\mathbf{y}_1^T(\cdot)\ \ \mathbf{y}_2^T(\cdot)\cdots]^T$ and $\mathbf{Z}_i(k)=[\mathbf{z}_1^T(\cdot)\ \ \mathbf{z}_2^T(\cdot)\cdots]^T$ denote two vectors containing all position and velocity measurements associated with $\mathbf{f}_i$ up to time k, respectively. Measurements $\mathbf{y}_j(l),\mathbf{z}_j(l)$ are associated with a velocity field $\mathbf{f}_i\in\mathcal{F}$ by the index set $\mathcal{G}(k)$ learned using (8). Let $\mathbf{A}\triangleq\boldsymbol{\Phi}[\mathbf{Y}_i(k),\mathbf{Y}_i(k)]+\sigma_v^2\mathbf{I}_{2k}$ denote the cross-covariance matrix obtained by GP regression [18]. Then, the measurement mean vector and covariance matrix for a collocation point vector $\boldsymbol{\xi}$ are

$$\begin{aligned} \boldsymbol{\mu}_i(k) &= \boldsymbol{\Phi}\left[\boldsymbol{\xi},\mathbf{Y}_i(k)\right]\mathbf{A}^{-1}\mathbf{Z}_i(k)\triangleq\boldsymbol{\mu}_{i,k} \\ \boldsymbol{\Sigma}_i(k) &= \boldsymbol{\Phi}(\boldsymbol{\xi},\boldsymbol{\xi})-\boldsymbol{\Phi}[\boldsymbol{\xi},\mathbf{Y}_i(k)]\mathbf{A}^{-1}\boldsymbol{\Phi}[\mathbf{Y}_i(k),\boldsymbol{\xi}]\triangleq\boldsymbol{\Sigma}_{i,k} \end{aligned} \tag{15}$$

Without loss of generality, consider $G_j(k+1)=i$. Then, since measurements obtained from one VF do not contain information about other VFs, it follows that,

$$D\big(\boldsymbol{\upsilon};\mathbf{m}_j(k+1)|Q(k)\big)=D\big(\boldsymbol{\upsilon}_i;\mathbf{m}_j(k+1)|Q(k)\big) \tag{16}$$

as proven in Appendix A. Now, if $\mathbf{m}_j(k+1)$ and $G_j(k+1)$ are both given, $\boldsymbol{\upsilon}_i$ is characterized by the multivariate joint Gaussian PDF in (14), and the conditional DPGP KL divergence (12) reduces to the closed form,

$$\begin{aligned} D\big(\boldsymbol{\upsilon};\mathbf{m}_j(k+1)\big) &= \frac{1}{2}\left[\mathrm{tr}\big(\boldsymbol{\Sigma}_{i,k}^{-1}\boldsymbol{\Sigma}_{i,k+1}\big)-\ln\left(\frac{|\boldsymbol{\Sigma}_{i,k+1}|}{|\boldsymbol{\Sigma}_{i,k}|}\right)\right] \\ &+\frac{1}{2}(\boldsymbol{\mu}_{i,k+1}-\boldsymbol{\mu}_{i,k})^T\boldsymbol{\Sigma}_{i,k}^{-1}(\boldsymbol{\mu}_{i,k+1}-\boldsymbol{\mu}_{i,k})-L \end{aligned} \tag{17}$$

where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix, and for brevity conditioning on $Q(k)$ is omitted here and in the remainder of the paper. When $\mathbf{m}_j(k+1)$ is unknown, as in sensor planning problems, the expected DPGP KL divergence can be derived as shown in the next subsection.

4.3 Expected DPGP KL Divergence

When planning the motion of a sensor, such as a camera mounted on a helicopter [9], the value of $\mathbf{m}_j(k+1)$ is unknown, and $G_j(k+1)$ cannot be learned from data. In this case, (17) cannot be evaluated because at time t_{k+1} the measurement mean and covariance in (15) are unknown. Then, the DPGP information value can be estimated by taking the expectation of (12) with respect to $\mathbf{m}_j(k+1)$ and $G_j(k+1)$ as follows

$$\hat{D}\left(\boldsymbol{\upsilon};\mathbf{m}_j(k+1)\right)\triangleq\mathbb{E}_{\mathbf{m}_j}\left\{\mathbb{E}_{G_j}\left\{D\left(\boldsymbol{\upsilon};\mathbf{m}_j(k+1)\right)\right\}\right\} \tag{18}$$

Let L_{ij} denote the likelihood of event $\{G_j(k+1)=i\}$, such that, $L_{ij}\triangleq p\left(M_j(k)|M_j^c(k),G_j(k+1)=i\right)$, where $M_j^c=Q\backslash M_j$ denotes the complement set of M_j in Q. Then, the posterior probability of $\{G_j(k+1)=i\}$ is,

$$w_{ij}\triangleq p\left(G_j(k+1)=i|Q(k)\right)=\frac{\pi_i L_{ij}}{\sum_{i=1}^{M}\pi_i L_{ij}} \tag{19}$$

where $\pi_i=p(G_j(k+1)=i)$, defined in (6), is the prior probability that the jth target obeys $\mathbf{f}_i$.

From (16) and (18), the DPGP expected KL divergence

Preprint submitted to Automatica
Received February 27, 2016 13:19:09 PST

(EKLD) can be written as,

$$\hat{D}\left(\boldsymbol{\upsilon};\mathbf{m}_j(k+1)\right)=\sum_{i=1}^{M} w_{ij}\hat{D}\left(\boldsymbol{\upsilon}_i;\mathbf{m}_j(k+1)\right) \quad (20)$$

and evaluated using the likelihood,

$$L_{ij}=\prod_{\ell=1}^{k}\frac{\exp\left\{-\frac{1}{2}[\mathbf{z}_{j,\ell}-\boldsymbol{\eta}_{j,\ell}]^T[\boldsymbol{\Lambda}_{j,\ell}]^{-1}[\mathbf{z}_{j,\ell}-\boldsymbol{\eta}_{j,\ell}]\right\}}{2\pi|\boldsymbol{\Lambda}_{j,\ell}|}$$

obtained from the ith GP component of the latest DPGP model, where $\mathbf{z}_{j,\ell}=\mathbf{z}_j(\ell)$, and the target velocity mean, $\boldsymbol{\eta}_{j,\ell}=\boldsymbol{\eta}_j(\ell)$, and covariance, $\boldsymbol{\Lambda}_{j,\ell}=\boldsymbol{\Lambda}_j(\ell)$, at $\mathbf{y}_j(\ell)$ are calculated by replacing $\boldsymbol{\xi}$ with $\mathbf{y}_j(\ell)$ in (15).

Then, the expected KL divergence of $\boldsymbol{\upsilon}_i$ in (20) is obtained by marginalizing the original KL divergence function over all possible values of $\mathbf{m}_j(k+1)$,

$$p\left(\mathbf{m}_j(k+1)|G_j(k+1)=i\right)= \quad (21)$$
$$p(\mathbf{z}_j(k+1)|\mathbf{x}_j(k+1))p\left(\mathbf{x}_j(k+1)|G_j(k+1)=i\right)$$

where $p(\mathbf{z}_j(k+1)|\mathbf{x}_j(k+1))$ can be calculated from (2). Using Euler integration and the ODE (1), it follows that,

$$p\left(\mathbf{x}_j(k+1)|G_j(k+1)=i\right)= \quad (22)$$
$$\int_{\mathbb{R}^2} p\left(\mathbf{v}_j(k)|\mathbf{x}_j(k)\right) f_X\left(\mathbf{x}_j(k+1)-\mathbf{v}_j(k)\Delta t\right) d\mathbf{v}_j(k)$$

where $f_X(\cdot)$ is the PDF of $\mathbf{x}_j(k)$ obtained via filtering techniques [23]. The conditional joint probability distribution of the target speed, $p\left(\mathbf{v}_j(k)|\mathbf{x}_j(k)\right)$, is a multivariate Gaussian distribution with mean and covariance calculated from (15) by replacing $\boldsymbol{\xi}$ with $\mathbf{x}_j(k)$. Thus, the information value of a future measurement $\mathbf{m}_j(k+1)$ can be estimated using equations (19)-(22).

When multiple target measurements can be obtained during one sampling interval $[t_k,\ t_k+\Delta t)$, say $\mathbf{m}(k)=[\mathbf{m}_j^T(k)\ \mathbf{m}_l^T(k)\cdots]^T$, the DPGP-EKLD is given by,

$$\hat{D}\left(\boldsymbol{\upsilon};\mathbf{m}(k+1)\right)=\sum_{j,\ \mathbf{x}_j\in\mathcal{S}(k+1)}\hat{D}\left(\boldsymbol{\upsilon};\mathbf{m}_j(k+1)\right)= \quad (23)$$
$$\sum_{j,i} w_{ij}\int_{\mathcal{S}(k+1)}\int_{\mathbb{R}^2} D\left(\boldsymbol{\upsilon}_i;\mathbf{m}_j(k+1)\right)p\left(\mathbf{z}_j(k+1)|\mathbf{x}_j(k+1)\right)$$
$$\times\, p\left(\mathbf{x}_j(k+1)|G_j(k+1)=i\right)d\mathbf{z}_j(k+1)d\mathbf{x}_j(k+1)$$

where all quantities are defined and calculated as shown in (19)-(22). Because the DPGP-EKLD function in (23) involves a 6th-order integral, its computation is typically very burdensome. Using the methodology presented in the next subsection, it is possible to simplify the DPGP-EKLD to a double integral, a form that can be implemented in real-time for sensor planning (Section 5).

4.4 DPGP Expected KL Divergence Approximation

The theoretical results presented in this subsection allow for the analytical approximation of the inner double integral in the DPGP-EKLD function (23), and provide performance guarantees for the subsequent Monte Carlo (MC) integration required to evaluate the DPGP-EKLD approximation. Since the probability $p\left(\mathbf{x}_j(k+1)|G_j(k+1)=i\right)$ is constant with respect to $\mathbf{z}_j(k+1)$, (23) can be simplified to,

$$\hat{D}\left(\boldsymbol{\upsilon};\mathbf{m}(k+1)\right)=\sum_{j,i} w_{ij}\int_{\mathcal{S}(k+1)} h_i[\mathbf{x}_j(k+1)] \quad (24)$$
$$\times\, p\left(\mathbf{x}_j(k+1)|G_j(k+1)=i\right)d\mathbf{x}_j(k+1)$$

where

$$h_i[\mathbf{x}_j(k+1)]\triangleq\int_{\mathbb{R}^2}\Big\{D\left(\boldsymbol{\upsilon}_i;\mathbf{m}_j(k+1)\right) \quad (25)$$
$$\times\, p\left(\mathbf{z}_j(k+1)|\mathbf{x}_j(k+1)\right)\Big\}d\mathbf{z}_j(k+1)$$

Then, the above second-order integral can be solved analytically according to the following proposition.

Proposition 1 *Consider a Gaussian process GP_i with known covariance matrix function $\boldsymbol{\Psi}_i$, and assume the target position is $\mathbf{x}_j(k+1)$. Then, the second-order integral in* (25) *affords the analytical solution,*

$$h_i[\mathbf{x}_j(k+1)]=\frac{1}{2}\Big[\mathrm{tr}\left(\boldsymbol{\Sigma}_{i,k}^{-1}\boldsymbol{\Sigma}_{i,k+1}\right)-\ln\left(|\boldsymbol{\Sigma}_{i,k+1}\boldsymbol{\Sigma}_{i,k}^{-1}|\right) \quad (26)$$
$$-2L+\mathrm{tr}(\mathbf{Q}^{-1}\mathbf{R}^T\boldsymbol{\Sigma}_{i,k}^{-1}\mathbf{R}\mathbf{Q}^{-1})\sigma_v^2\Big]$$

where $\boldsymbol{\Sigma}_{i,k}$ and $\boldsymbol{\Sigma}_{i,k+1}$ are the velocity covariance matrices at times k and $k+1$, respectively, and,

$$\mathbf{A}\triangleq\boldsymbol{\Phi}[\mathbf{Y}_i(k),\mathbf{Y}_i(k)]+\sigma_v^2\mathbf{I}_{2k} \quad (27)$$
$$\mathbf{B}\triangleq\boldsymbol{\Phi}[\mathbf{Y}_i(k),\mathbf{x}_j(k+1)] \quad (28)$$
$$\mathbf{D}\triangleq\boldsymbol{\Phi}[\mathbf{x}_j(k+1),\mathbf{x}_j(k+1)]+\sigma_v^2\mathbf{I}_2 \quad (29)$$
$$\mathbf{R}\triangleq\boldsymbol{\Phi}[\boldsymbol{\xi},\mathbf{x}_j(k+1)]-\boldsymbol{\Phi}[\boldsymbol{\xi},\mathbf{Y}_i(k)]\mathbf{A}^{-1}\mathbf{B} \quad (30)$$
$$\mathbf{Q}\triangleq\mathbf{D}-\mathbf{B}^T\mathbf{A}^{-1}\mathbf{B} \quad (31)$$

for any $\mathbf{x}_j(k+1)\in\mathcal{S}(k+1)$, where $\boldsymbol{\xi}$ is a vector of collocation points, $\mathbf{Y}_i$ is a vector of all past position measurements, and $\boldsymbol{\Phi}$ is the cross-covariance matrix of GP_i.

PROOF. Using the matrix inversion lemma, $[\boldsymbol{\Sigma}_{i,k+1}]^{-1}$ can be written as,

$$\begin{bmatrix}\mathbf{A}^{-1}(\mathbf{I}_{2k}+\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}^T\mathbf{A}^{-1}) & -\mathbf{A}^{-1}\mathbf{B}\mathbf{Q}^{-1}\\ \mathbf{Q}^{-1}\mathbf{B}^T\mathbf{A}^{-1} & \mathbf{Q}^{-1}\end{bmatrix}$$

Preprint submitted to Automatica
Received February 27, 2016 13:19:09 PST

where $\mathbf{Q}$, $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{A}$ are defined in (31). Substituting $\boldsymbol{\Sigma}_{i,k+1}^{-1}$ in the GP mean vector expression (15), the change in the mean during two consecutive time steps can be written as,

$$\boldsymbol{\mu}_i(k+1) - \boldsymbol{\mu}_i(k) = \mathbf{R}\mathbf{Q}^{-1}\mathbf{z}_j'(k+1) \tag{32}$$

where $\mathbf{z}_j'(k+1) \triangleq \mathbf{z}_j(k+1) - \mathbb{E}[\mathbf{z}_j(k+1)]$ is a Gaussian random vector with zero mean and covariance matrix $\sigma_v^2\mathbf{I}_2$. Substituting (32) into (25), the integral (26) can be simplified to

$$h_i[\mathbf{x}_j(k+1)] = \frac{1}{2}\Bigg[\operatorname{tr}\left(\boldsymbol{\Sigma}_{i,k}^{-1}\boldsymbol{\Sigma}_{i,k+1}\right) - \ln\left(|\boldsymbol{\Sigma}_{i,k+1}\boldsymbol{\Sigma}_{i,k}^{-1}|\right) - 2L + \operatorname{tr}(\mathbf{Q}^{-1}\mathbf{R}^T\boldsymbol{\Sigma}_{i,k}^{-1}\mathbf{R}\mathbf{Q}^{-1})\sigma_v^2\Bigg] \qquad \square$$

Even with the above simplification, the computational complexity of evaluating $h_i[\mathbf{x}_j(k+1)]$ is $O(L^3 + k^3)$, where L is the number of collocation points and k is the time index. Thus, it may be infeasible to evaluate (24) for all possible measurements. Because the integrand of (24) goes to zero when $p\left(\mathbf{x}_j(k+1)|G_j(k+1) = i\right)$ goes to zero, and $h_i[\mathbf{x}_j(k+1)]$ is finite, this computation can be significantly reduced by the approach known as Monte Carlo integration [16]. Let $\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(S)}$ denote S values of $\mathbf{x}_j(k+1)$ drawn identically and independently from $p\left(\mathbf{x}_j(k+1)|G_j(k+1) = i\right)$ in (22). Then, the integral in (24) can be computed numerically by evaluating its integrand at each sample with nonzero probability, i.e.,

$$\hat{D}\left(\boldsymbol{\upsilon}; \mathbf{m}(k+1)\right) \approx \sum_{j,i} \frac{w_{ij}}{S} \sum_{l=1}^{S} h_i(\boldsymbol{\xi}^{(l)}) \mathbf{1}_{\mathcal{S}(k+1)}(\boldsymbol{\xi}^{(l)}) \tag{33}$$

where,

$$\mathbf{1}_{\mathcal{S}(k+1)}(\boldsymbol{\xi}^{(l)}) \triangleq \begin{cases} 1, & \boldsymbol{\xi}^{(l)} \in \mathcal{S}(k+1) \\ 0, & \boldsymbol{\xi}^{(l)} \notin \mathcal{S}(k+1) \end{cases} \tag{34}$$

is the indicator function. Similarly to the collocation points used to discretize the DPGP information value, these samples represent points in $\mathcal{W}$. But, unlike collocation points, which are chosen from a uniform grid, the Monte Carlo integration samples are chosen by sampling a known distribution.

The remainder of this subsection proves that, based on the following three lemmas, the approximation in (33) is an unbiased estimator of the DPGP-EKLD in (24), and its error variance decreases linearly with S.

Lemma 2 *The velocity covariance matrix* $\mathbf{Q}$*, defined in* (31)*, obeys the element-wise inequality* $\boldsymbol{\Psi}_0 + \left\{\sigma_v^2/[k\,\operatorname{tr}(\boldsymbol{\Psi}_0) + \sigma_v^2]\right\}\mathbf{I}_2 \preceq \mathbf{Q} \preceq \boldsymbol{\Psi}_0 + \sigma_v^2\mathbf{I}_2$*, where* $\boldsymbol{\Psi}_0 = \boldsymbol{\Psi}(\mathbf{0}, \mathbf{0})$ *is a constant matrix obtained by evaluating the stationary covariance matrix function at zero,* $\sigma_v \in \mathbb{R}$ *is the standard deviation of the measurement noise of velocity, and* $k \in \mathbb{Z}_+$ *is the time index.*

PROOF.

From (27), the matrix $\mathbf{A}$ is symmetric and positive definite ($\mathbf{A} > 0$) because it is the sum of a real, symmetric, positive semi-definite matrix, $\boldsymbol{\Phi}[\mathbf{Y}_i(k), \mathbf{Y}_i(k)]$, and a real, symmetric, positive definite matrix, $\sigma_v^2\mathbf{I}_{2k}$. Then, $\mathbf{A}^{-1} > 0$ and $\mathbf{B}^T\mathbf{A}^{-1}\mathbf{B} \succ \mathbf{0}_{2\times 2}$, where $\mathbf{B}$ is defined in (28), and $\succ$ denotes element-wise inequality. From (31), the following also holds,

$$\mathbf{Q} \preceq \boldsymbol{\Psi}[\mathbf{x}_j(k+1), \mathbf{x}_j(k+1)] + \sigma_v^2\mathbf{I}_2 \tag{35}$$

and since the covariance matrix function is stationary, $\boldsymbol{\Psi}[\mathbf{x}_j(k+1), \mathbf{x}_j(k+1)] = \boldsymbol{\Psi}(\mathbf{0}, \mathbf{0}) = \boldsymbol{\Psi}_0$ for any $\mathbf{x}_j \in \mathcal{W}$, and thus $\mathbf{Q} \preceq \boldsymbol{\Psi}_0 + \sigma_v^2\mathbf{I}_2$.

Since $\boldsymbol{\Phi}[\mathbf{Y}_i(k), \mathbf{Y}_i(k)]$ is real, symmetric, and positive semi-definite, there exists an eigenvalue decomposition, $\boldsymbol{\Phi}[\mathbf{Y}_i(k), \mathbf{Y}_i(k)] = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{-1}$, with orthogonal eigenvectors, where $\boldsymbol{\Lambda}$ is a diagonal matrix obtained by placing the k eigenvalues of $\boldsymbol{\Phi}[\cdot]$ on the diagonal, and $\mathbf{U}$ is a $k \times k$ matrix whose columns are the eigenvectors of $\boldsymbol{\Phi}[\cdot]$, i.e.,

$$\boldsymbol{\Lambda} \triangleq \operatorname{diag}[\lambda_1 \ \cdots \ \lambda_k] \quad \text{and} \quad \mathbf{U} \triangleq [\mathbf{u}_1 \ \cdots \ \mathbf{u}_k]^T \tag{36}$$

and, thus, $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. Since the kth column of $\mathbf{Y}_i(k)$ is equal to $\mathbf{x}_j(k)$, the matrix $\mathbf{B}$ defined in (28) can be written as, $\mathbf{B} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{u}_k$. By substituting $\mathbf{B}$ into in (31), the matrix $\mathbf{Q}$ can be written as,

$$\mathbf{Q} = \boldsymbol{\Psi}_0 + \left[\sigma_v^2 - \sum_{\ell=1}^{k} \frac{\lambda_\ell}{\lambda_\ell + \sigma_v^2}\lambda_\ell(\mathbf{U}_{(\ell,k)})^2\right]\mathbf{I}_2 \tag{37}$$

where λ_ℓ is the ℓth eigenvalue of $\boldsymbol{\Phi}[\mathbf{Y}_i(k), \mathbf{Y}_i(k)]$, and $\mathbf{U}_{(\ell,k)}$ denotes the element in the ℓth row and kth column of $\mathbf{U}$. Because $\boldsymbol{\Phi}[\cdot]$ is symmetric and positive semi-definite, $\lambda_\ell \geq 0$ for all ℓ, and, since $\mathbf{u}_k^T\boldsymbol{\Lambda}\mathbf{u}_k = 1$, it follows that $\sum_{\ell=1}^{k}\lambda_\ell(\mathbf{U}_{(\ell,k)})^2 = 1$, which can be substituted into (37) to show that, $\mathbf{Q} \succeq \boldsymbol{\Psi}_0 + \left(\sigma_v^2 - \frac{\max_\ell\{\lambda_\ell\}}{\max_\ell\{\lambda_\ell\}+\sigma_v^2}\right)\mathbf{I}_2$. Because the trace of a real, symmetric matrix equals the sum of its eigenvalues, and that the diagonal blocks of $\boldsymbol{\Phi}[\mathbf{Y}_i(k), \mathbf{Y}_i(k)]$ are equal to $\boldsymbol{\Psi}(\mathbf{y}_l, \mathbf{y}_l) = \boldsymbol{\Psi}_0$, then $\sum_{\ell=1}^{k}\lambda_\ell = k\,\operatorname{tr}(\boldsymbol{\Psi}_0)$. Furthermore, $\max_\ell\{\lambda_\ell\} \leq k\,\operatorname{tr}(\boldsymbol{\Psi}_0)$, and thus,

$$\mathbf{Q} \succeq \boldsymbol{\Psi}_0 + \left[\frac{\sigma_v^2}{k\,\operatorname{tr}(\boldsymbol{\Psi}_0) + \sigma_v^2}\right]\mathbf{I}_2 \succeq \boldsymbol{\Psi}_0 \succ \mathbf{0} \tag{38}$$

which completes the proof. $\square$

Preprint submitted to Automatica
Received February 27, 2016 13:19:09 PST

The above result is used in the following lemma to establish a matrix inequality on consecutive covariance matrices $\mathbf{\Sigma}_{i,k}$ and $\mathbf{\Sigma}_{i,k+1}$.

Lemma 3 *Under the assumptions in Proposition* 1*, two consecutive covariance matrices* $\mathbf{\Sigma}_{i,k}$ *and* $\mathbf{\Sigma}_{i,k+1}$*, defined according to* (15)*, obey* $\mathbf{0} \prec \mathbf{\Sigma}_{i,k+1} \preceq \mathbf{\Sigma}_{i,k}$.

PROOF. Under the assumptions in Proposition 1, $\mathbf{\Sigma}_{i,k}, \mathbf{\Sigma}_{i,k+1} > 0$ because they are Gaussian covariance matrices. From the block matrix inversion of $\mathbf{\Sigma}_{i,k+1}$ the difference between two consecutive covariance matrices can be written as

$$\mathbf{\Sigma}_{i,k} - \mathbf{\Sigma}_{i,k+1} = \mathbf{R}\mathbf{Q}^{-1}\mathbf{R}^T \tag{39}$$

From Lemma 2, $\mathbf{\Psi}_0 \preceq \mathbf{Q}$ and, since $\mathbf{Q}$ is a diagonal positive-definite matrix, it follows that $\mathbf{Q}^{-1}$ is diagonal and positive definite. Then, there exists a diagonal positive-definite matrix, $\mathbf{V}$, such that $\mathbf{Q}^{-1} = \mathbf{V}\mathbf{V}^T$ and

$$\mathbf{\Sigma}_{i,k} - \mathbf{\Sigma}_{i,k+1} = \mathbf{R}\mathbf{V}\mathbf{V}^T\mathbf{R}^T = (\mathbf{R}\mathbf{V})(\mathbf{R}\mathbf{V})^T \succeq \mathbf{0} \qquad \square$$

The third and final lemma provides a bound on the trace of the quadratic form $\mathbf{R}^T\mathbf{R}$ that is later used to show that the Monte Carlo integration (33) is an unbiased estimator of the DPGP-EKLD in (24).

Lemma 4 *Under the assumptions in Proposition* 1*, the cross-covariance matrix* $\mathbf{R}$*, obtained from the target velocity and position estimates at the collocation points, and defined in* (30)*, obeys the inequality* $0 \leq \mathrm{tr}(\mathbf{R}\mathbf{R}^T) \leq 4k[(\mathbf{\Psi}_0)+2\sigma_v^2]$*, where* $\mathbf{\Psi}_0 = \mathbf{\Psi}(\mathbf{0},\mathbf{0})$ *is a constant matrix obtained by evaluating the stationary covariance matrix function at zero,* $\sigma_v \in \mathbb{R}$ *is the standard deviation of the measurement noise, and* $k \in \mathbb{Z}_+$ *is the time index.*

PROOF. Since $\mathbf{R}^T\mathbf{R}$ is a positive semi-definite matrix, it has positive or zero eigenvalues, and $\mathrm{tr}(\mathbf{R}\mathbf{R}^T) \geq 0$. From GP regression, the joint probability distribution of a vector $[\boldsymbol{\upsilon}_i^T \quad \mathbf{v}_j(k+1)^T \quad \mathbf{f}_i(\mathbf{y}_1)^T \quad \mathbf{f}_i(\mathbf{y}_2)^T \quad \cdots]^T$, comprised of target velocities at the collocation points (10) and at the measured target positions, is a multivariate Gaussian distribution with covariance matrix,

$$\begin{bmatrix} \mathbf{\Phi}(\boldsymbol{\xi},\boldsymbol{\xi}) + \sigma_v^2\mathbf{I}_{2L} & \mathbf{\Phi}[\boldsymbol{\xi}, \mathbf{x}_j(k+1)] & \mathbf{\Phi}[\boldsymbol{\xi}, \mathbf{Y}_i(k)] \\ \mathbf{\Phi}[\mathbf{x}_j(k+1), \boldsymbol{\xi}] & \mathbf{D} & \mathbf{B}^T \\ \mathbf{\Phi}[\mathbf{Y}_i(k), \boldsymbol{\xi}] & \mathbf{B} & \mathbf{A} \end{bmatrix}$$

where $\mathbf{A}, \mathbf{B}, \mathbf{D}$ are defined in (27)-(29). Then, the conditional marginal distribution of the vector $[\boldsymbol{\upsilon}_i^T \quad \mathbf{v}_j(k+1)^T]^T$, given the vector of target position measurements $\mathbf{Y}_i(k)$, is a multivariate Gaussian distribution with covariance matrix $\begin{bmatrix} \mathbf{\Sigma}_{i,k} & \mathbf{R} \\ \mathbf{R}^T & \mathbf{Q} \end{bmatrix}$, where $\mathbf{R}$ and $\mathbf{Q}$ are defined in (30)-(31). Because the covariance matrix is symmetric positive definite, the off-diagonal elements of $\mathbf{R}$ are smaller than the corresponding diagonal elements of $\mathbf{Q}$, or $\mathbf{R}_{(i,j)} < \mathbf{Q}_{(j,j)} < \mathrm{tr}(\mathbf{Q})$, for all i, j. From Lemma 2, $\mathbf{Q} \preceq \mathbf{\Phi}_0 + \sigma_v^2\mathbf{I}_2$, therefore it follows that $\mathbf{R}_{(i,j)} < \mathrm{tr}(\mathbf{\Psi}_0) + 2\sigma_v^2$ and, from the properties of quadratic forms, $\mathrm{tr}(\mathbf{R}\mathbf{R}^T) = \sum_{ij}[\mathbf{R}_{(i,j)}]^2 < 4k[\mathrm{tr}(\mathbf{\Psi}_0) + 2\sigma_v^2]$, completing the proof. $\square$

We are now ready to prove the following theorem on the properties of the DPGP-EKLD approximation:

Theorem 5 *Under the assumptions in Proposition* 1*, the Monte Carlo (MC) integration* (33) *is an unbiased estimator of the DPGP-EKLD* (24)*, and the variance of the difference between* (24) *and* (33) *decreases linearly with the number of samples, S, that are drawn independently and identically from the target state distribution* (22)*.*

PROOF. From the linearity of the expectation operation, the expected value of the DPGP-EKLD obtained via Monte Carlo integration in (33), is given by

$$\begin{aligned} \bar{D} &= \mathbb{E}_{\boldsymbol{\xi}^{(l)}}\{\hat{D}\left(\boldsymbol{\upsilon}; \mathbf{m}(k+1)\right)\} \\ &= \sum_{j=1}^{N}\sum_{i=1}^{M}\frac{w_{ij}}{S}\sum_{s=1}^{S}\mathbb{E}_{\boldsymbol{\xi}^{(l)}}\left\{h_i(\boldsymbol{\xi}^{(l)})\mathbf{1}_{\mathcal{S}(k+1)}(\boldsymbol{\xi}^{(l)})\right\} \end{aligned} \tag{40}$$

Since MC samples are iid, the following holds,

$$\begin{aligned} &\mathbb{E}_{\boldsymbol{\xi}^{(l)}}\left\{h_i(\boldsymbol{\xi}^{(l)})\mathbf{1}_{\mathcal{S}(k+1)}(\boldsymbol{\xi}^{(l)})\right\} \\ &\quad = \mathbb{E}_{\boldsymbol{\xi}^{(l)}}\left\{h_i\big[\mathbf{x}_j(k+1)\big]\mathbf{1}_{\mathcal{S}(k+1)}[\mathbf{x}_j(k+1)]\right\} \end{aligned} \tag{41}$$

for all $l = 1, \ldots, S$, where $h_i\big[\mathbf{x}_j(k+1)\big]$ is defined in (26), and thus the estimator (33) is unbiased. Also, the variance of the DPGP-EKLD in (40) can be written as,

$$\begin{aligned} \mathrm{var}(\hat{D}) &= \mathbb{E}\{(\hat{D} - \bar{D})^2\} \\ &= \sum_{j=1}^{N}\sum_{i=1}^{M} w_{ij}\frac{1}{S^2}\sum_{l=1}^{S}\mathbb{E}\Bigg\{\Bigg[h_i\big(\boldsymbol{\xi}^{(l)}\big)\mathbf{1}_{\mathcal{S}(k+1)}\big(\boldsymbol{\xi}^{(l)}\big) \\ &\quad - \mathbb{E}\Big\{h_i\big[\mathbf{x}_j(k+1)\big]\mathbf{1}_{\mathcal{S}(k+1)}\big[\mathbf{x}_j(k+1)\big]\Big\}\Bigg]^2\Bigg\} \\ &= \frac{1}{S}\mathrm{var}\Big\{h_i\big[\mathbf{x}_j(k+1)\big]\mathbf{1}_{\mathcal{S}(k+1)}\big[\mathbf{x}_j(k+1)\big]\Big\} \end{aligned} \tag{42}$$

where it can be proven that $\mathrm{var}\left\{h_i\big[\mathbf{x}_j(k+1)\big]\mathbf{1}_{\mathcal{S}(k+1)}\big[\mathbf{x}_j(k+1)\big]\right\}$ is a finite constant because $h_i[\cdot]$ is finite-valued as follows.

Preprint submitted to Automatica
Received February 27, 2016 13:19:09 PST

As a first step, we prove that $h_i[\cdot]$ is greater than or equal to zero. From (25), $h_i[\cdot]$ is the integral of the KL divergence weighted by a Gaussian distribution. Since the KL divergence is always greater than or equal to zero, and $p(\mathbf{z}_j(k+1)|\mathbf{x}_j(k+1)) \geq 0$, it also follows that $h_i[\cdot] \geq 0$. As a second step, we prove that $h_i[\cdot]$ is finite by showing that every term of $h_i[\cdot]$ is finite. The first term in (26) is finite-valued because from Lemma 3,

$$\mathrm{tr}(\boldsymbol{\Sigma}_{i,k}^{-1}\boldsymbol{\Sigma}_{i,k+1}) = \mathrm{tr}(\mathbf{I}_{2L} - \boldsymbol{\Sigma}_{i,k}\mathbf{R}\mathbf{Q}^{-1}\mathbf{R}^T) \leq 2L \quad (43)$$

where L is the number of collocation points in $\mathcal{W}$. The second term in (26) is also finite because from Lemma 3 $|\boldsymbol{\Sigma}_{i,k}| > 0$ and $|\boldsymbol{\Sigma}_{i,k+1}| > 0$, since the two covariance matrices are positive definite. Since $\boldsymbol{\Sigma}_{i,k}$ and $\boldsymbol{\Sigma}_{i,k+1}$ are Hermitian and $\boldsymbol{\Sigma}_{i,k} \succeq \boldsymbol{\Sigma}_{i,k+1}$, then $|\boldsymbol{\Sigma}_{i,k}| > |\boldsymbol{\Sigma}_{i,k+1}|$ and $0 < |\boldsymbol{\Sigma}_{i,k+1}|/|\boldsymbol{\Sigma}_{i,k}| < 1$, such that $0 < -\ln(|\boldsymbol{\Sigma}_{i,k+1}|/|\boldsymbol{\Sigma}_{i,k}|) < \infty$. Since the trace of a product of matrices is invariant under cyclic permutations, and the trace of the product of positive semi-definite matrices is less than or equal to the product of the individual traces [3], it follows that,

$$\begin{aligned}\mathrm{tr}(\mathbf{Q}^{-1}\mathbf{R}^T\boldsymbol{\Sigma}_{i,k}^{-1}\mathbf{R}\mathbf{Q}^{-1})\sigma_v^2 = \mathrm{tr}(\boldsymbol{\Sigma}_{i,k}^{-1})[\mathrm{tr}(\mathbf{Q}^{-1})]^2 \quad (44)\\ \times\mathrm{tr}(\mathbf{R}\mathbf{R}^T)\sigma_v^2 \leq 4\mathrm{tr}(\boldsymbol{\Sigma}_{i,k}^{-1})\left[4k\mathrm{tr}(\boldsymbol{\Phi}_0) + 2\sigma_v^2\right]\sigma_v^2\end{aligned}$$

where the last inequality is the result of Lemmas 2-4. Now that every term in (26) has been proven finite, $h_i[\cdot]$ is finite-valued for all $\mathbf{x}_j(k+1) \in \mathbb{R}^2$. Also from (34), the variance in (42) is a definite integral of $h_i[\cdot]$ over $\mathcal{S}(k+1)$, and thus a finite constant. Then, the error variance of the MC estimator in (33) is proportional to $1/S$ [5]. □

5 DPGP Sensor Planning

The DPGP-EKLD approximation in (33) is used to plan the sensor measurements in the target modeling problem formulated in Section 2. The performance obtained by planning the motion of the sensor FoV, $\mathcal{S}(k+1)$, using a DPGP-EKLD maximizing algorithm is compared to an algorithm based on DPGP mutual information (MI), derived by the approach in Section 4, a particle filter (PF) algorithm [21], and a randomized method [7]. The workspace is $\mathcal{W} = [0, 10] \times [0, 10]$ m^2, and the camera FoV is a square of side $s = 0.5$ m^2 and variable position in $\mathcal{W}$. The sensor measurements are characterized by $\sigma_x = \sigma_v = 0.1$ (m/s) and $\Delta t = 0.3$ sec. The unknown target ODEs are,

$$\begin{aligned}\mathbf{f}_1[\cdot] &= \left[-\sin\left([0.1 \;\; \pi/20]\mathbf{x}_j\right) \quad \sin\left([\pi/12 \;\; 1]\mathbf{x}_j + \pi/4\right)\right]^T\\ \mathbf{f}_2[\cdot] &= \left[-\cos\left([0 \;\; \pi/8]\mathbf{x}_j\right) \quad \sin\left([\pi/4 \;\; 0]\mathbf{x}_j\right)\right]^T\\ \mathbf{f}_3[\cdot] &= [-0.5 \quad \sin\left([\pi/4 \;\; 0.3]\mathbf{x}_j\right)]^T\\ \mathbf{f}_4[\cdot] &= [-\cos\left([\pi/8 \quad -0.5]\mathbf{x}_j\right) \quad 1]^T \end{aligned} \quad (45)$$

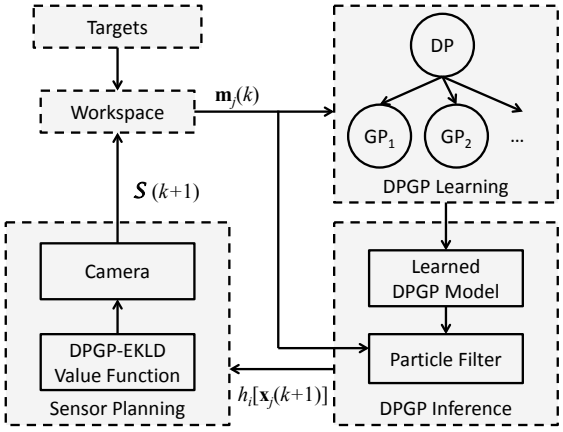

Fig. 1. Diagram of DPGP-EKLD sensor planning algorithms.

such that $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$, and $\boldsymbol{\pi} = [0.25\, 0.25\, 0.25\, 0.25]^T$.

The performance of the DPGP-EKLD algorithm is evaluated for $t_0 = 0$, $t_f = 300$ s. During this time, $N = 500$ targets enter and exit $\mathcal{W}$ following a VF in $\mathcal{F}$ selected randomly from $\mathrm{Cat}(\boldsymbol{\pi})$. The target VFs and statistics $\{\mathcal{F}, \boldsymbol{\pi}\}$ are learned from data, using $\boldsymbol{\Psi}(\mathbf{x}_j, \mathbf{x}'_j) = \exp\left(-\|\mathbf{x}_j - \mathbf{x}'_j\|_2^2/2\ell^2\right)\mathbf{I}_2$ and $\ell = \sqrt{10}$ based on the size of $\mathcal{W}$. An example of target trajectories used to learn the DPGP model is shown in Fig. 2, superimposed on a contour plot of $\mathrm{var}(\mathbf{v}_j)$. At the onset of the simulation, the DP concentration parameter is chosen as $\alpha = 1$. Subsequently, the DPGP model is updated at every time t_k, based on the complete database $Q(k)$. As schematized in Fig. 1, the target-VF associations $\mathcal{G}(k)$ are obtained from the DPGP using a particle filter [23]. Based on the latest DPGP, the position of the camera FoV in $\mathcal{W}$ is computed so as to maximize EKLD and obtain the target measurements $\mathbf{m}_j$ with the highest information value.

From (3), the DPGP target model can be evaluated by comparing its estimate of target velocity, $\hat{\mathbf{v}}_j(k) = \boldsymbol{\theta}_i(\mathbf{x}_j) = \mathbb{E}[\mathbf{f}_i(\mathbf{x}_j)]$, to the actual velocity $\mathbf{v}_j(k)$ such that, at t_k, the average root mean squared error (RMSE) is,

$$\varepsilon_k = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{M} w_{ij}\sqrt{\frac{1}{K_j}\sum_{k=1}^{K_j}\|\mathbf{v}_j(k) - \hat{\mathbf{v}}_j(k)\|_2^2} \quad (46)$$

where K_j is the number of time intervals spent in $\mathcal{W}$ by target j, and w_{ij} is the target-VF association probability in (19). The EKLD algorithm is compared to an MI algorithm that maximizes DPGP mutual information, a PF algorithm that places $\mathcal{S}(k+1)$ to track the nearest target, and a random algorithm that places $\mathcal{S}(k+1)$ by sampling a uniform multivariate distribution with support $\mathcal{W}$. Three cases with low, medium, and high prior information are considered, obtaining the results in Table 1. It can be seen that the EKLD information function is the most effective at improving RMSE, or reducing model error, over time.

Preprint submitted to Automatica
Received February 27, 2016 13:19:09 PST

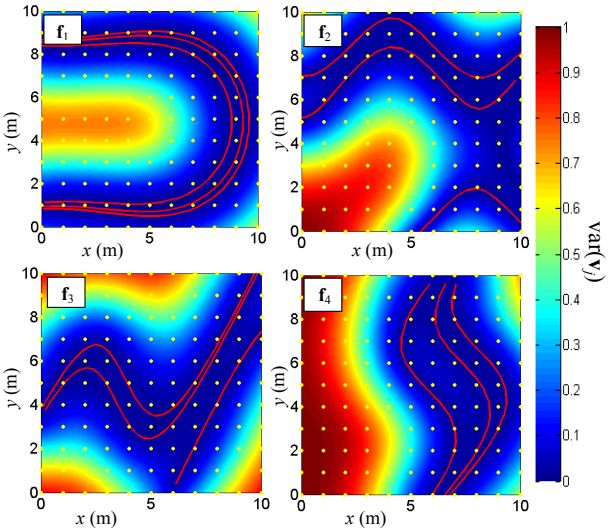

Fig. 2. Example target trajectories (red), collocation points (yellow) and velocity variance (contour plot) for the four velocity fields in (45).

Table 1
Average RMSE of DPGP model.

Algorithms:	EKLD	MI	PF	Random
Low prior information	0.35%	0.79%	3.31%	5.16%
Medium prior information	0.11%	0.25%	2.46%	3.32%
High prior information	0.09%	0.21%	0.56%	1.34%

The time history of RMSE and its variance for all four planning algorithms is plotted in Fig. 3, for the case of medium prior information. These results, obtained for a sample of fifty simulations performed under the same DPGP statistics, illustrate the performance robustness of the EKLD algorithm. The MI algorithm also displays good performance at the onset of learning, because the medium prior is approximately equally informative about all four VFs. However, as more target measurements of targets are obtained over time, the EKLD function is more effective because it prioritizes targets characterized by VFs with higher uncertainty. Finally, the results in Fig. 4 show that the posterior probabilities of target-VF association defined in (19), and obtained from the updated DPGP via PF, converge to their true values (dash-dotted lines) over time. The snapshots in Fig. 4.(a)-(c) show that, by planning the camera movements via EKLD, the FoV is able to intersect PF clusters with high information value and, thus, rapidly improve their posterior probability distributions.

6 Conclusion

This paper presents an approach for deriving tractable information value functions for nonparametric DPGP

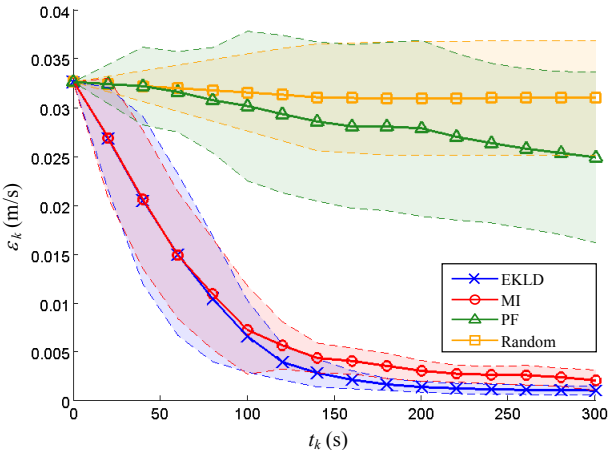

Fig. 3. Mean and (±) standard deviation of the VF RMSE in (46) resulting from EKLD, MI, PF, and random camera planning, for the case of medium prior information.

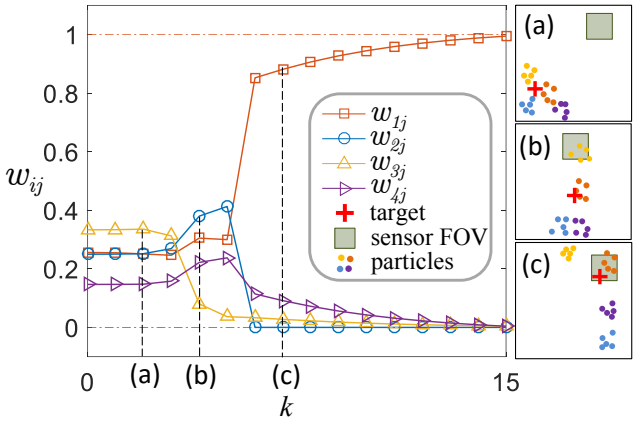

Fig. 4. Time history of target-VF posterior probabilities updated over time by the DPGP and PF algorithms in Fig. 1.

mixture models. Although information theoretic functions are a common approach for representing information value in sensing and control problems, they are not directly applicable to DPGP models of stochastic processes because they are typically defined in terms of finite-dimensional probability distributions. The proposed approach is demonstrated by deriving efficient representations of KL divergence in DPGP models via collocation methods and MC integration. In particular, the EKLD approximation, useful for planning and control of mobile sensors, is proven to be an unbiased estimator of the original information function, with an error variance that is guaranteed to decrease linearly with the number of MC samples. The effectiveness of the approach is demonstrated through a sensor planning problem in which the motion of the camera FoV is planned by optimizing the DPGP-EKLD value function over time. The results show that the EKLD sensor planning algorithm outperforms algorithms based on DPGP mutual information, particle filter, and randomized methods.

Preprint submitted to Automatica
Received February 27, 2016 13:19:09 PST

A Proof of Equation (16)

PROOF. When $G_j(k+1) = i$,

$$p(\boldsymbol{\upsilon}|Q(k+1)) = p(\boldsymbol{\upsilon}_i|Q(k+1)) \prod_{1\le l\le M,\ l\neq i} p(\boldsymbol{\upsilon}_l|Q(k)) \quad \text{(A.1)}$$

since $\boldsymbol{\upsilon}_i = \boldsymbol{\upsilon}_i(\boldsymbol{\xi})$ and $\boldsymbol{\upsilon}_l = \boldsymbol{\upsilon}_l(\boldsymbol{\xi})$ are conditional independent given $Q(k+1)$. Substituting (A.1) into (12), the conditional KL divergence can be written as

$$\begin{aligned} & D\left(\boldsymbol{\upsilon}; \mathbf{m}_j(k+1)|Q(k)\right) \qquad \text{(A.2)} \\ &= \int_{\mathbb{R}^{2LM}} \ln\left\{ \frac{p(\boldsymbol{\upsilon}_i|Q(k+1)) \prod_{1\le l\le M,\ l\neq i} p(\boldsymbol{\upsilon}_l(\boldsymbol{\xi})|Q(k))}{\prod_{\ell=1}^{2M} p(\boldsymbol{\upsilon}_\ell(\boldsymbol{\xi})|Q(k))} \right\} \\ &\quad \times p\big(\boldsymbol{\upsilon}|Q(k+1)\big)\, d\boldsymbol{\upsilon} \\ &= \int_{\mathbb{R}^{2M}} \ln\left\{ \frac{p\big(\boldsymbol{\upsilon}_i|Q(k+1)\big)}{p\big(\boldsymbol{\upsilon}_i|Q(k)\big)} \right\} p\big(\boldsymbol{\upsilon}_i|Q(k+1)\big) d\boldsymbol{\upsilon}_i \\ &= D\left(\boldsymbol{\upsilon}_i; \mathbf{m}_j(k+1)|Q(k)\right) \qquad \square \end{aligned}$$

References

[1] Y. Bar-Shalom and W.D. Blair. *Multitarget-Multisensor Tracking: Applications and Advances, Vol. III*. Artech House, 2000.

[2] Y. Bar-Shalom and E. Tse. Tracking in a cluttered environment with probabilistic data association. *Automatica*, 11(5):451–460, 1975.

[3] D. S. Bernstein. *Matrix Mathematics*. Princeton University Press, Princeton, NJ, 2005.

[4] V. I. Bogachev and M. A. S. Ruas. *Measure Theory*, volume 1. Springer, 2007.

[5] R. E. Caflisch. Monte Carlo and Quasi-Monte Carlo methods. *Acta numerica*, 7:1–49, 1998.

[6] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, NY, 1991.

[7] C. Fulgenzi, C. Tay, A. Spalanzani, and C. Laugier. Probabilistic navigation in dynamic environment using rapidly-exploring random trees and Gaussian processes. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 1056–1062. IEEE, 2008.

[8] J. Hensman, M. Rattray, and N. D. Lawrence. Fast nonparametric clustering of structured time-series. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 37(2):383–393, 2015.

[9] J. Joseph, F. Doshi-Velez, A. S. Huang, and N. Roy. A Bayesian nonparametric approach to modeling motion patterns. *Autonomous Robots*, 31(4):383–400, 2011.

[10] K. Kastella. Discrimination gain to optimize detection and classification. *IEEE Transactions on Systems, Man, and Cybernetics-Part A*, 27(1):112–116, 1997.

[11] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *The Journal of Machine Learning Research*, 9:235–284, 2008.

[12] W. Lu, G. Zhang, and S. Ferrari. An information potential approach to integrated sensor path planning and control. *Robotics, IEEE Transactions on*, 30(4):919–934, 2014.

[13] F. Nobile, R. Tempone, and C. G. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.

[14] J. L. Ny and G. J. Pappas. On trajectory optimization for active sensing in Gaussian process models. In *Proc. of the IEEE Conference on Decision and Control*, pages 6286 –6292, Shanghai, China, Dec 2009.

[15] K. R. Pattipati, S. Deb, Y. Bar-Shalom, and R. Washburn Jr. A new relaxation algorithm and passive sensor data association. *Automatic Control, IEEE Transactions on*, 37(2):198–213, 1992.

[16] W. H. Press. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge university press, 2007.

[17] C. E. Rasmussen. The infinite Gaussian mixture model. In *Proc. of Advances in Neural Information Processing Systems (NIPS)*, volume 12, pages 554–560, Denver, CO, USA, Dec 1999.

[18] C. E. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[19] J. Ross and J. Dy. Nonparametric mixture of Gaussian processes with constraints. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, pages 1346–1354, 2013.

[20] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Hierarchical Dirichlet processes. *Journal of the american statistical association*, 101(476), 2006.

[21] J. Wang, Y. Yin, and M. Hong. Multiple human tracking using particle filter with Gaussian process dynamical model. *EURASIP Journal on Image and Video Processing*, 2008.

[22] H. Wei, W. Lu, and S. Ferrari. An information value function for nonparametric Gaussian processes. In *Proc. Neural Information Processing Systems Conference*, Lake Tahoe, NV, 2012.

[23] H. Wei, W. Lu, P. Zhu, S. Ferrari, R. H. Klein, S. Omidshafiei, and J. P. How. Camera control for learning nonlinear target dynamics via Bayesian nonparametric Dirichlet-Process Gaussian-Process (DP-GP) models. In *Intelligent Robots and Systems, IEEE/RSJ International Conference On*, pages 95–102. IEEE, 2014.

[24] H. Wei, W. Lu, P. Zhu, G. Huang, J. Leonard, and S. Ferrari. Optimized visibility motion planning for target tracking and localization. In *Intelligent Robots and Systems, 2014 IEEE/RSJ International Conference On*, pages 76–82. IEEE, 2014.

[25] H. Wei, W. Ross, S. Varisco, P. Krief, and S. Ferrari. Modeling of human driver behavior via receding horizon and artificial neural network controllers. In *Decision and Control, IEEE Annual Conference On*, pages 6778–6785, Florence, Italy, Dec 2013.

[26] G. Zhang, S. Ferrari, and C. Cai. A comparison of information functions and search strategies for sensor planning in target classification. *IEEE Transactions on Systems, Man, and Cybernetics - Part B*, 42(1):2–16, 2012.

Preprint submitted to Automatica
Received February 27, 2016 13:19:09 PST