

# Variational Inference with Mixture Model Approximation for Applications in Robotics

Emmanuel Pignat, Teguh Lembono and Sylvain Calinon

**Abstract**—We propose to formulate the problem of representing a distribution of robot configurations (e.g. joint angles) as that of approximating a product of experts. Our approach uses variational inference, a popular method in Bayesian computation, which has several practical advantages over sampling-based techniques. To be able to represent complex and multimodal distributions of configurations, mixture models are used as approximate distribution. We show that the problem of approximating a distribution of robot configurations while satisfying multiple objectives arises in a wide range of problems in robotics, for which the properties of the proposed approach have relevant consequences. Several applications are discussed, including learning objectives from demonstration, planning, and warm-starting inverse kinematics problems. Simulated experiments are presented with a 7-DoF Panda arm and a 28-DoF Talos humanoid.

## I. INTRODUCTION

Many robotic approaches rely on the capability to generate in a quick and efficient way robot configurations fulfilling multiple objectives. These configurations are typically computed using iterative methods, which can leverage good and diversified initial guesses to improve the convergence speed and to find the global optimum.

Our work addresses the problem of representing the distribution  $p(\mathbf{x})$  of robot configurations that satisfy multiple given objectives.<sup>1</sup> We first show that our goal of representing the distribution of *good and diversified* configurations is the same as the goal pursued by variational inference (VI), a popular and scalable method for Bayesian computation. VI methods approximate an unnormalized density by a tractable one.

Similarly as [1] for a 10-DoF planar robot, we propose to use a mixture model to represent the complex and often multimodal distribution of solutions. We extend the framework to conditional distributions using a mixture of experts (MoE) approximation. Conditional distributions provide greater online adaptation capabilities to the robot, as it allows us to represent solutions corresponding to the distributions of parametrized objectives. We discuss how several forms of objectives in robotics can be included in this framework, illustrated by various robot applications.

The authors are with the Idiap Research Institute, CH-1920 Martigny, Switzerland. E-mail: {name.surname}@idiap.ch

The research leading to these results has received funding from the European Commission’s Horizon 2020 Programme through the MEMMO Project (Memory of Motion, <http://www.memmo-project.eu/>, grant agreement 780684) and COLLABORATE project (<https://collaborate-project.eu/>, grant agreement 820767)

<sup>1</sup>Robot configurations are often denoted  $\mathbf{q}$ , but we will avoid this notation because  $\mathbf{q}$  are also used in variational inference to denote approximate distributions.

As a practitioner guide, we gather several transformations and experts models related to a set of standard robotics problems in an Appendix. The paper is accompanied by a full Python/Tensorflow library (the codes are included in the submission), allowing researchers to build upon the method and/or use it with other robots, by providing a URDF file.

## II. TWO INTERPRETATIONS OF THE INVESTIGATED PROBLEM

We start by presenting two interpretations of the problem, both leading to the optimization of the same objective.

### A. Product of experts

A product of experts [2] is a model multiplying several densities  $p_m$  (which are called *experts*) together and renormalizing the density. The renormalization makes sure that it is a proper density and introduces dependencies between the experts. Illustratively, each expert  $m$  gives its opinion on a different view or transformation  $\mathcal{T}_m(\mathbf{x})$  of the data  $\mathbf{x}$ . Their opinions are then fused together as the product

$$p(\mathbf{x}|\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M) = \frac{\prod_m p_m(\mathcal{T}_m(\mathbf{x})|\boldsymbol{\theta}_m)}{\int_z \prod_m p_m(\mathcal{T}_m(\mathbf{z})|\boldsymbol{\theta}_m)}. \quad (1)$$

For example, recalling that  $\mathbf{x}$  denotes the configuration of a robot (joint angles and floating base 6-DoF transformation for a humanoid),  $\mathcal{T}_m(\mathbf{x})$  can be the position of a foot computed using forward kinematics and  $p_m$  a Gaussian distribution defining where the foot should be. The product of experts becomes the distribution of configuration where the foot follows the distribution  $p_m$ . By adding multiple experts representing various objectives, the product becomes the distribution fulfilling the set of objectives. We provide various examples of representations and transformations in the Appendix, including forward and inverse kinematics, the center of mass, relative distance spaces, orientations and hierarchies (nullspace structures).

We point out that the experts do not need to be properly normalized themselves, as the normalization already occurs at the level of the product. For compactness, we will later refer to the unnormalized product as

$$\tilde{p}(\mathbf{x}) = \prod_m p_m(\mathcal{T}_m(\mathbf{x})|\boldsymbol{\theta}_m), \quad (2)$$

where we dropped the parameters of the experts  $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M$  in the notation. In the general case, the normalized product  $p(\mathbf{x})$  has no closed-form expression. Gaussian experts with linear transformations are a notable exception, where  $p(\mathbf{x})$  is Gaussian, but is of limited interest in our case, due to the

nonlinearity of the transformations we are considering. Approximation methods, as used in Bayesian computation, are required; the posterior distribution is indeed the renormalized product on  $\theta$  of a likelihood and a prior.

*Markov chain Monte Carlo:* Markov chain Monte Carlo (MCMC) is a class of methods to approximate  $p(\mathbf{x})$  with samples. If MCMC methods can represent arbitrary complex distributions, they suffer from some limitations, particularly constraining for our application. They are known not to scale well to a high dimensional space. Except for some methods [3], they require an exact evaluation of  $\tilde{p}(\mathbf{x})$  while stochastic variational inference only requires a stochastic estimate of its gradient. MCMC methods also struggle with multimodal distributions and require particular proposal steps to move from distant modes [4]. Designing a good proposal step is also algorithmically restrictive. Furthermore, it is difficult to obtain good acceptance rates in high dimension, especially with highly correlated  $\tilde{p}(\mathbf{x})$ . When using sampling-based techniques, it is also difficult to assess if the distribution  $\tilde{p}(\mathbf{x})$  is well covered.

*Variational Inference:* Variational inference (VI) [5] is another popular class of methods that recasts the approximation problem as an optimization. VI approximates the *target density*  $\tilde{p}(\mathbf{x})$  with a *tractable density*  $q(\mathbf{x}; \lambda)$ , where  $\lambda$  are the *variational parameters*. *Tractable density* means that drawing samples from  $q(\mathbf{x}; \lambda)$  should be easy and  $q(\mathbf{x}; \lambda)$  should be properly normalized. VI tries to minimize the intractable KL-divergence

$$D_{\text{KL}}(q||p) = \int_{\mathbf{x}} q(\mathbf{x}; \lambda) \log \frac{q(\mathbf{x}; \lambda)}{p(\mathbf{x})} d\mathbf{x} \quad (3)$$

$$= \int_{\mathbf{x}} q(\mathbf{x}; \lambda) \log \frac{q(\mathbf{x}; \lambda)}{\tilde{p}(\mathbf{x})} d\mathbf{x} + \log \mathcal{C}, \quad (4)$$

where  $\mathcal{C}$  is the normalizing constant. Instead, it minimizes the negative *evidence lower bound* (ELBO) which can be estimated by sampling as

$$\mathcal{L}(\lambda) = \int_{\mathbf{x}} q(\mathbf{x}; \lambda) \log \frac{q(\mathbf{x}; \lambda)}{\tilde{p}(\mathbf{x})} d\mathbf{x} \quad (5)$$

$$= \mathbb{E}_q[\log q(\mathbf{x}; \lambda) - \log \tilde{p}(\mathbf{x})] \quad (6)$$

$$\approx \sum_{n=1}^N (\log q(\mathbf{x}^{(n)}; \lambda) - \log \tilde{p}(\mathbf{x}^{(n)})) \quad (7)$$

$$\text{with } \mathbf{x}^{(n)} \sim q(\cdot | \lambda)$$

The reparametrization trick [6] [7] allows us to compute a noisy estimate of the gradient  $\mathcal{L}(\lambda)$ , which is compatible with stochastic gradient optimization like Adam [8]. For example, if  $q$  is Gaussian, this is done by sampling  $\boldsymbol{\eta}^{(n)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  and applying the continuous transformation  $\mathbf{x}^{(n)} = \boldsymbol{\mu} + \mathbf{L}\boldsymbol{\eta}^{(n)}$ , where  $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^\top$  is the covariance matrix.  $\mathbf{L}$  and  $\boldsymbol{\mu}$  are the *variational parameters*  $\lambda$ . More complex mappings as normalizing flows can be used [9].

*a) Zero forcing properties of minimizing  $D_{\text{KL}}(q||p)$ :*

It is important to note that, due to the objective  $D_{\text{KL}}(q||p)$ ,  $q$  is said to be zero forcing. If  $q$  is not expressive enough to approximate  $\tilde{p}$ , it would miss some mass of  $\tilde{p}$  rather than

giving probability to locations where there is no mass (see Fig. 1 for an illustration). In our applications, it means that we are more likely to miss some solutions than retrieve wrong ones.

## B. Maximum entropy

Another interpretation is that the robot configurations should minimize a sum of cost

$$c(\mathbf{x}) = \sum_m c_m(\mathbf{x}), \quad (8)$$

representing the different objectives. Computing  $\arg \min_{\mathbf{x}} c(\mathbf{x})$  is equivalent to *maximum a posteriori* in Bayesian statistics. Retrieving a distribution of configurations can be done by finding the distribution  $q$  under which the expectation of cost  $\mathbb{E}_q[c(\mathbf{x})]$  is minimized with the widest entropy  $\mathcal{H}(q)$

$$\mathcal{L}(\lambda) = \mathbb{E}_q[\log q(\mathbf{x}; \lambda) + c(\mathbf{x})] \quad (9)$$

$$= \mathbb{E}_q[c(\mathbf{x})] - \mathcal{H}(q). \quad (10)$$

It actually corresponds to (6), where  $\tilde{p}(\mathbf{x})$  was replaced with

$$\tilde{p}(\mathbf{x}) = \exp(-c(\mathbf{x})). \quad (11)$$

Intuitively, it also fits with our initial objective to generate *good* ( $\mathbb{E}_q[c(\mathbf{x})]$ ) and *diversified* ( $\mathcal{H}(q)$ ) samples.

## III. MIXTURE MODEL VARIATIONAL DISTRIBUTION

For computational efficiency, the *approximate distribution*  $q(\mathbf{x}; \lambda)$  is often chosen as a factorized distribution, using the mean-field approximation [5]. Correlated distribution can be approximated by a full-covariance Gaussian distribution [10]. These approaches fail to capture the multimodality and arbitrary complexity of  $\tilde{p}(\mathbf{x})$ . The idea to use a mixture for greater expressiveness as *approximate distribution* is not new [11], but the approach has regained popularity recently, see, e.g., [12], [13], [1].

A mixture model is built by summing the probability of  $k$  mixture components

$$q(\mathbf{x}|\lambda) = \sum_{k=1}^K \pi_k q_k(\mathbf{x}|\lambda_k), \quad \sum_{k=1}^K \pi_k = 1, \quad (12)$$

where  $\pi_k$  is the total mass of component  $k$ . The components  $q_k$  can be of any family accepting a continuous and invertible mapping between  $\lambda$  and the samples. The discrete sampling of the mixture components according to  $\pi_k$  has no such mapping. Instead, the variational objective can be rewritten as

$$\mathcal{L}(\lambda) = \mathbb{E}_q[\log q(\mathbf{x}; \lambda) - \log \tilde{p}(\mathbf{x})] \quad (13)$$

$$= \sum_{k=1}^K \pi_k \mathbb{E}_{q_k}[\log q(\mathbf{x}; \lambda) - \log \tilde{p}(\mathbf{x})], \quad (14)$$

meaning that we need to compute and get the derivatives of expectations only under each *component distribution*  $q_k(\mathbf{x}|\lambda_k)$ .

*Mixture components distributions:* Gaussian components with a full covariance matrix are a natural choice for robotics, due to the quadratic form of its log-likelihood. It can be exploited in standard robotics approaches like linear quadratic tracking (LQT) or inverse kinematics (IK), more details in Sec IV. For some situations, where  $\tilde{p}(\mathbf{x})$  is highly correlated, we propose to use “*banana-shaped*” distribution [14], which is done by applying the following differentiable mapping to Gaussian samples  $\boldsymbol{\eta}^{(n)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ ,

$$f_{\boldsymbol{\kappa}}(\boldsymbol{\mu} + \mathbf{L}\boldsymbol{\eta}^{(n)}) = f_{\boldsymbol{\kappa}}(\mathbf{z}^{(n)}) = \begin{bmatrix} z_0^{(n)} + \sum_{i=1}^N \kappa_i z_i^{(n)2} \\ z_1^{(n)} \\ \vdots \\ z_N^{(n)} \end{bmatrix}. \quad (15)$$

This mapping can be applied along different parametrized directions. As illustrated in Fig. 1 (b), we get a Gaussian with full covariance, where  $\boldsymbol{\kappa}$  is a supplementary *variational parameter* encoding curvature.

*Conditional distribution:* Let us suppose the following problem: we want to sample humanoid configurations in static equilibrium within the joint limits where the feet are close to given poses. As this distribution differs substantially given the poses of the feet, we would like to retrieve this distribution conditioned on the feet locations. That way, no new computation is required according to the current objective, providing a very fast online adaptability. More generally, we want to approximate  $\tilde{p}(\mathbf{x}|\mathbf{y})$  where  $\mathbf{y}$  is a *task parameter*, for a distribution of possible  $p(\mathbf{y})$ .

For example, if  $\mathbf{y}_m$  defines a forward kinematics target, the expert  $p_m$  can be

$$p_m(\mathbf{x}|\mathbf{y}_m) \propto \mathcal{N}(\mathbf{F}(\mathbf{x})|\mathbf{y}_m, \sigma\mathbf{I}). \quad (16)$$

We propose to use as *approximate distribution* a mixture of experts (ME)

$$q(\mathbf{x}|\mathbf{y}; \boldsymbol{\lambda}) = \sum_{k=1}^K h_k(\mathbf{y}; \boldsymbol{\lambda}_k) q_k(\mathbf{x}|\mathbf{y}; \boldsymbol{\lambda}_k), \quad \sum_{k=1}^K h_k(\mathbf{y}; \boldsymbol{\lambda}_k) = 1, \quad (17)$$

where  $h_k$  are the gate and  $q_k$  are the conditional mixture components as in (12).  $q_k$  are normally referred to as *experts*.<sup>2</sup> From (12), we change the mass  $\pi_k$  and the *mixture components*  $q_k$  to depend on the *task parameters*  $\mathbf{y}$ .

In this work, we will use  $q_k$  as Gaussian around a linear function, but more complex functions can be considered. Samples from  $q_k(\mathbf{x}|\mathbf{y}; \boldsymbol{\lambda}_k)$  can be generated by first drawing  $\boldsymbol{\eta}^{(n)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  and then applying the mapping  $\mathbf{x}^{(n)} = \mathbf{W}_k\boldsymbol{\eta} + \mathbf{c}_k + \mathbf{L}_k\boldsymbol{\eta}^{(n)}$ . The *variational parameters*  $\boldsymbol{\lambda}_k$  are  $\mathbf{W}_k$ ,  $\mathbf{c}_k$  and  $\mathbf{L}_k$ .

Given  $\mathbf{y}$ , the objective (6) becomes

$$\mathcal{L}(\boldsymbol{\lambda}, \mathbf{y}) = \mathbb{E}_q[\log q(\mathbf{x}|\mathbf{y}; \boldsymbol{\lambda}) - \log \tilde{p}(\mathbf{x}|\mathbf{y})], \quad (18)$$

<sup>2</sup>We will avoid this term because of possible confusion with the product of experts.

which we would like to minimize under the distribution of possible *task parameters*  $p(\mathbf{y})$ , namely

$$\begin{aligned} \int_{\mathbf{y}} p(\mathbf{y}) \mathcal{L}(\boldsymbol{\lambda}, \mathbf{y}) d\mathbf{y} &= \mathbb{E}_{p(\mathbf{y})} \left[ \mathbb{E}_q[\log q(\mathbf{x}|\mathbf{y}; \boldsymbol{\lambda}) - \log \tilde{p}(\mathbf{x}|\mathbf{y})] \right] \\ &= \mathbb{E}_{p(\mathbf{y})} \left[ \sum_{k=1}^K h_k(\mathbf{y}; \boldsymbol{\lambda}_k) \mathbb{E}_{q_k(\cdot|\mathbf{y}; \boldsymbol{\lambda}_k)}[\log q(\mathbf{x}|\mathbf{y}; \boldsymbol{\lambda}) - \log \tilde{p}(\mathbf{x}|\mathbf{y})] \right]. \end{aligned} \quad (19)$$

This objective can also be minimized by stochastic gradient optimization [6], [7], by sampling first  $\mathbf{y}^{(n)} \sim p(\mathbf{y})$  and then  $\mathbf{x}^{(l,n)} \sim q(\mathbf{x}|\mathbf{y}^{(n)}; \boldsymbol{\lambda})$  for each  $\mathbf{y}^{(n)}$ .

## IV. APPLICATIONS

### A. Learning objectives from demonstration

It might seem intuitive to choose the transformations needed to represent a desired joint distribution for a given task, but choosing the experts parameters  $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M$  can be more complex (if not infeasible). These parameters can be learned by providing samples of possible desired configurations, for example through kinesthetic teaching. From a given dataset of robot configurations  $\mathbf{X}$ , maximum likelihood (or maximum a posteriori) of the intractable distribution  $p(\mathbf{x}|\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)$  should be computed. It can be done using gradient descent [2] with

$$\begin{aligned} \frac{\partial \log p(\mathbf{x}|\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)}{\partial \boldsymbol{\theta}_m} &= \frac{\partial \log p_m(\mathbf{x}|\boldsymbol{\theta}_m)}{\partial \boldsymbol{\theta}_m} - \frac{\partial \log \mathcal{C}(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)}{\partial \boldsymbol{\theta}_m} \\ &= \frac{\partial \log p_m(\mathbf{x}|\boldsymbol{\theta}_m)}{\partial \boldsymbol{\theta}_m} - \int_{\mathbf{c}} p(\mathbf{c}|\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M) \frac{\partial \log p_m(\mathbf{c}|\boldsymbol{\theta}_m)}{\partial \boldsymbol{\theta}_m} d\mathbf{c}, \end{aligned} \quad (20)$$

where  $\mathcal{C}(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)$  is the intractable normalizing constant. Computing a step of gradient requires us to compute an expectation under the intractable distribution  $p(\mathbf{x}|\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)$ , making the process computationally expensive. In [15], it is proposed to use a few sampling steps initialized at the data distribution  $\mathbf{X}$ . Unfortunately this approach fails when  $p(\mathbf{x}|\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)$  has multiple modes. The few sampling steps never move between modes, resulting in the incapacity of estimating their relative mass. *Wormholes* have been proposed as a solution [16], but are algorithmically restrictive.

Instead, we propose to use VI with the proposed mixture distribution to approximate  $p(\mathbf{x}|\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_M)$ . The training process thus alternates between minimizing  $D_{\text{KL}}(q||p)$  with current  $p$  and using current  $q$  to compute the gradient (20).  $q$  can either be used as an importance sampling distribution (or directly if expressive enough to represent  $p$ ). The mass of the modes is directly encoded as  $\pi_k$  which makes it much more convenient than moving samples between distant modes.

This process seems overly complex compared to directly encode  $p(\mathbf{x})$ , for example as a mixture, a non-parametric density or with a variational autoencoder. However, it offers several advantages, especially when we have access only to small datasets, and/or when very good generalization capabilities are required, namely:

- Simple and compact explanation can be found for complex distributions.

- Domain specific a priori knowledge can be included in the form of particular transformations and experts, which would reduce the need for data. More complex transformations can still be learned, if more data are available and more complex objectives are considered.
- The model is more structured and interpretable for a human user. It could facilitate interactive or active learning procedure.

For example, to learn directly the distribution shown in Fig. 1, a lot of samples spanning the whole distribution would be required. Finding a simple explanation under a known transformation, following Occam’s razor principle, increase generalization capabilities. Very sharp and precise distributions could be obtained with a very limited number of parameters.

### B. Planning

Sampling-based robot motion planning [17] requires sampling of configurations in obstacle-free space, while possibly satisfying other constraints. One of the difficulty is to generate those samples, especially with precise constraints and in high dimension space, which cannot be done with standard rejection sampling approaches. Uniform sampling and projection into the constraint is proposed in [18], and dedicated approaches for closed chains are proposed in [19], [20]. Variational methods can be interesting when high dimension configuration spaces are involved, as they are known to scale better than sampling.

Another challenge of sampling-based planning is to make sure that the space is well covered, ensuring good connectivity. This is indeed a general problem of MCMC methods when approximating a distribution. Using a mixture model, it is easier to assess for connectivity as the existence of common mass between components. The overlap between two components  $k$  and  $l$  can be computed for example with

$$a_{kl} = \int_{\mathbf{x}} q_k(\mathbf{x}; \boldsymbol{\lambda}_k) q_l(\mathbf{x}; \boldsymbol{\lambda}_l) d\mathbf{x}, \quad (21)$$

which is closed-form in case of Gaussian components. Due to the zero forcing properties of VI, if the components have some common mass, it is very likely that they are connected.

### C. Warm-starting inverse kinematics

Inverse kinematics problems are typically solved with iterative techniques that are subject to local optimum and benefit from a good initialization. It often arises that the poses we find should satisfy general, task-specific and goal-specific constraints. For example, a humanoid can achieve a wide range of tasks with its hands while constantly keeping its feet on the ground and ensuring balance. To be able to warm-start the optimization of all the constraints together, it would be beneficial to pre-compute the distribution of poses, already satisfying the general constraints.

If Gaussian components are used, a nice property is to be able to compute the intersection of objectives without further optimization. For example, if we have  $q_A(\mathbf{x}; \boldsymbol{\lambda})$  approximating the distribution of poses satisfying the set of

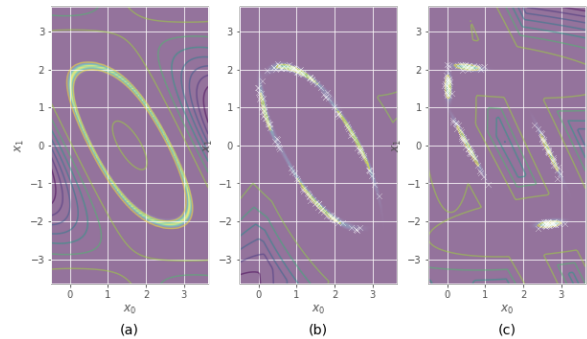


Fig. 1. Precise forward kinematics objectives often create a highly correlated distribution of joint angles, which is difficult to sample from or to estimate as a mixture of Gaussians. (a) Joint distribution of a 2-DoF planar robot whose end-effector should stay on a line. The colormap represents  $\exp(-c(\mathbf{x}))$  and the lines are isolines of  $c(\mathbf{x})$ . Probability mass is located in yellow (or lighter) areas. (b) A mixture of “banana-shaped” distributions can substantially reduce the number of mixture components required to represent the distribution, thanks to its capacity to adapt to the local curvature. (c) A mixture of Gaussians with the same number of components would not be able to cover efficiently the distribution. Here, for illustrating our point, we employed fewer components than required. The zero forcing properties of  $D_{\text{KL}}(q||p)$  are visible here, as no mass of  $q$  is assigned where  $\tilde{p}$  has none.

objectives  $A$ , defined by the intractable distribution  $\tilde{p}_A(\mathbf{x})$  and  $q_B(\mathbf{x}; \boldsymbol{\lambda})$  for the set  $B$ . We can directly compute  $q_{A \cap B}(\mathbf{x}; \boldsymbol{\lambda})$ , the distribution of configurations satisfying both sets as a product of mixtures of Gaussians, which is itself a mixture of Gaussians [21].

## V. EXPERIMENTS

In the experiments, we evaluate the quality of the approximation  $q$  with respect to the number of components and the choice of the component distribution. The model with one component is the baseline, which corresponds to the common Gaussian variational approximation (GVA). The quality of the approximation is evaluated using different similarity and divergence measures between  $\tilde{p}$  and  $q$ . These quantities are computed with integrals, which are evaluated by discretizing the configuration space within joint limits. We consider 3 scenarios:

- a 2-DoF planar robot whose end-effector should stay on a line;
- a 7-DoF Panda arm with a given position target for its gripper;
- a 28-DoF and floating base (6-DoF) Talos humanoid with a given position target of left gripper and feet, where the robot should also ensure static equilibrium on a flat ground.

In each one, the robot has more degrees of freedom than required by the task, creating a rich distribution of solutions. For the floating base, the rotation is parametrized using Euler angles. We make the assumption that the solution has a small variance. For a more rigorous treatment of the problem, a circular or higher-dimensional *tractable density* should be chosen. Due to the generally intractable normalizing constant and the difficulties to sample such directional distributions, it is still an open research question.

TABLE I

QUANTITATIVE EVALUATION OF THE APPROXIMATION QUALITY OF  $q$  FOR DIFFERENT NUMBERS OF COMPONENTS AND COMPONENT DISTRIBUTIONS.

THE DISTRIBUTION TO APPROXIMATE  $\tilde{p}$  IS SHOWN ON FIG. 1 AND CORRESPONDS TO CONFIGURATIONS OF A 2-DOF PLANAR ROBOT, WHOSE END-EFFECTOR IS ON A LINE. THE MEASURES ARE BHATTACHARYYA COEFFICIENT [22] AND OVERLAPPING COEFFICIENT [23] (HIGHER IS BETTER, 1 IS FOR PERFECT RECONSTRUCTION), ALPHA-DIVERGENCE (OR RÉNYI DIVERGENCE) WITH  $\alpha = 1/2$  [24] (LOWER IS BETTER). THESE METRICS WERE SELECTED BECAUSE OF THEIR SYMMETRY, UNLIKE THE KL DIVERGENCE AND THE REVERSE KL DIVERGENCE.

	Bhatt.	OVL	$D_{\alpha=1/2}$		Bhatt.	OVL	$D_{\alpha=1/2}$
<i>Banana-shaped</i>				<i>Gaussian</i>			
1 comp.	0.452	0.211	1.096	GVA - 1 comp.	0.313	0.110	1.374
5 comp.	0.793	0.561	0.414	5 comp.	0.680	0.439	0.640
10 comp.	0.963	0.826	0.073	10 comp.	0.927	0.738	0.146
15 comp.	0.977	0.854	0.046	15 comp.	0.948	0.783	0.104
20 comp.	0.995	0.936	0.009	20 comp.	0.989	0.895	0.022

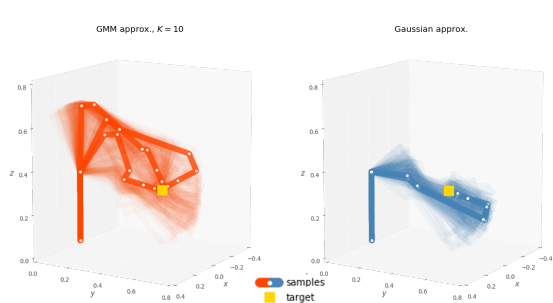


Fig. 2. Comparing approximations  $q$  of the distribution of configuration of the 7-DoF Panda robot whose end-effector position follows a Gaussian distribution. (left) Gaussian mixture model approximation with  $K = 10$  components. (right) Gaussian approximation. For each approximation, 3 samples drawn from the distribution are shown as well as 200 more samples in transparency. The GMM approximation covers a significantly higher proportion of the distribution and proposes much more diversified solutions.

Table I reports the results for the 2-DoF robot and Fig. 1 displays  $\tilde{p}$  and  $q$  as colormaps. As expected, one Gaussian is not expressive enough to represent the highly correlated distribution created by precise forward kinematic objectives. In a 2-dimensional space, a relatively low number of components is sufficient to approximate  $\tilde{p}$  almost exactly. Models with “banana-shaped” component reach quicker the distribution due to their ability to bend.

Table II reports the results for the 7-DoF Panda robot. Because of the 7 dimensions, the distribution is significantly richer and more complex to represent. If the results suggest that we are not representing the distribution as well as in the previous scenario, Fig. 2 shows that we already have a quite good variety of configurations with 10 Gaussian components. For a better approximation, an iterative training procedure can be used, such as proposed in [12], [13]. It incrementally builds a richer approximation by increasing the number of states. Compared to sampling, an interesting feature of the proposed method is that the insensitivity of the task with respect to the last joint of the robot (the wrist) can be directly encoded in the covariance matrices. This greatly reduces the number of units (samples or clusters) to represent the distribution, which can be crucial in high-dimension space.

For the humanoid robot, running evaluations using discretization is too expensive. We show in Fig. 3 samples of

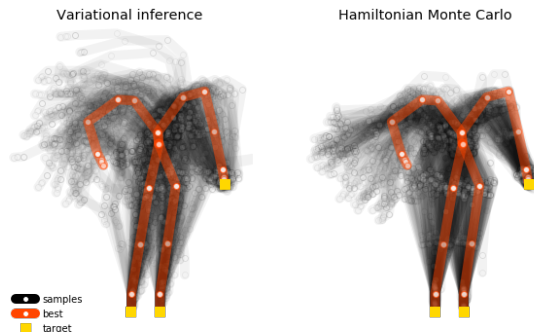


Fig. 3. Samples of configurations where the Talos robot should have its left hand at a given position and both feet at a given pose. A static equilibrium should also be ensured. (left) 100 samples from GMM variational distribution with  $K = 50$  components. (right) 100 randomly picked samples of a 2000 steps Hamiltonian Monte Carlo chain. Both learning procedures were prematurely stopped. It illustrates the advantages of the variational inference by its efficiency (in terms of time and memory) to represent irrelevant dimensions.

configuration from  $q$ , a Gaussian mixture model with  $K = 50$  components, using Hamiltonian Monte Carlo (HMC). We allowed for a limited 40 sec. of stochastic optimization for variational inference and HMC steps (equivalent to 2000 steps). For HMC, the kernel was chosen Gaussian with a standard deviation of 0.1. Fig. 3 shows a higher variance of the right arm for variational inference. The related degrees of freedom have little to no influence on the task (only for balancing). Thanks to the encoding of Gaussian covariances and the maximization of entropy in (10), the method can quickly localize unimportant dimensions. Compared to the sampling approach, it has benefits both in terms of computational and memory efficiency, as well as in generalization capabilities.

In [25], the authors also addressed a similar problem on the Talos robot using a GMM. They approximated the feasible positions of the CoM given the set of contact points, which is only a 3 dimension problem. Learning the GMM is done with a two-steps procedure implying sampling and then fitting the model. This procedure would be problematic with higher-dimensional problems. Our approach directly copes with full state of the robot but further developments are required to tackle the feasibility in case of a non-flat terrain.

TABLE II

RESULTS FOR 7-DOF PANDA ROBOT (SEE TABLE I FOR DESCRIPTION). THE POSITION OF THE GRIPPER SHOULD FOLLOW A GAUSSIAN DISTRIBUTION. SAMPLE CONFIGURATIONS OF THESE APPROXIMATIONS ARE SHOWN IN FIG. 2.

	Bhatt.	$D_{\alpha=1/2}$		Bhatt.	$D_{\alpha=1/2}$
“Banana-shaped”			<i>Gaussian</i>		
1 comp.	0.091	1.81	GVA - 1 comp.	0.085	1.89
10 comp.	0.362	1.27	10 comp.	0.225	1.55
100 comp.	0.587	0.82	100 comp.	0.505	0.98

## VI. CONCLUSION

We proposed the use of variational techniques with mixture models as a general tool to represent distributions of robot configurations satisfying several forms of objectives. We emphasized its advantages over sampling methods, showing that computational and memory requirements can be significantly reduced in high dimension space, which can be crucial in many applications. The developed approach, and associated source codes, can be exploited in a wide range of applications. It is particularly promising for training unnormalized density (e.g., energy-based model or PoE) as it can handle multimodality better than sampling approaches. This direction will be further developed and presented in future studies.

## REFERENCES

- [1] O. Arenz, G. Neumann, and M. Zhong, “Efficient gradient-free variational inference using policy search,” in *Proc. Intl Conf. on Machine Learning (ICML)*, ser. Proceedings of Machine Learning Research, J. Dy and A. Krause, Eds., vol. 80. Stockholm: PMLR, 10–15 Jul 2018, pp. 234–243.
- [2] G. E. Hinton, “Products of experts,” *Proc. Intl Conf. on Artificial Neural Networks (ICANN)*, pp. 1–6, 1999.
- [3] T. Chen, E. Fox, and C. Guestrin, “Stochastic gradient hamiltonian monte carlo,” in *Proc. Intl Conf. on Machine Learning (ICML)*, 2014, pp. 1683–1691.
- [4] C. Smichisescu, M. Welling, and G. Hinton, “A mode-hopping MCMC sampler,” Technical Report CSRG-478, University of Toronto, submitted to Machine , Tech. Rep., 2003.
- [5] M. J. Wainwright, M. I. Jordan *et al.*, “Graphical models, exponential families, and variational inference,” *Foundations and Trends® in Machine Learning*, vol. 1, no. 1–2, pp. 1–305, 2008.
- [6] T. Salimans, D. A. Knowles *et al.*, “Fixed-form variational posterior approximation through stochastic linear regression,” *Bayesian Analysis*, vol. 8, no. 4, pp. 837–882, 2013.
- [7] R. Ranganath, S. Gerrish, and D. Blei, “Black box variational inference,” in *Artificial Intelligence and Statistics*, 2014, pp. 814–822.
- [8] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *Proc. Intl Conf. on Learning Representation (ICLR)*, 2015.
- [9] D. J. Rezende and S. Mohamed, “Variational inference with normalizing flows,” in *Proceedings of the 32nd International Conference on International Conference on Machine Learning-Volume 37*. JMLR.org, 2015, pp. 1530–1538.
- [10] M. Opper and C. Archambeau, “The variational Gaussian approximation revisited,” *Neural computation*, vol. 21, no. 3, pp. 786–792, 2009.
- [11] C. M. Bishop, N. D. Lawrence, T. Jaakkola, and M. I. Jordan, “Approximating posterior distributions in belief networks using mixtures,” in *Advances in Neural Information Processing Systems (NIPS)*, 1998, pp. 416–422.
- [12] A. C. Miller, N. J. Foti, and R. P. Adams, “Variational boosting: Iteratively refining posterior approximations,” in *Proc. Intl Conf. on Machine Learning (ICML)*, 2017, pp. 2420–2429.
- [13] F. Guo, X. Wang, K. Fan, T. Broderick, and D. B. Dunson, “Boosting variational inference,” *Advances in Neural Information Processing Systems (NIPS)*, 2016.
- [14] H. Haario, E. Saksman, and J. Tamminen, “Adaptive proposal distribution for random walk Metropolis algorithm,” *Computational Statistics*, vol. 14, no. 3, pp. 375–396, 1999.

- [15] G. E. Hinton, “Training products of experts by minimizing contrastive divergence,” *Neural computation*, vol. 14, no. 8, pp. 1771–1800, 2002.
- [16] M. Welling, A. Mnih, and G. E. Hinton, “Wormholes improve contrastive divergence,” in *Advances in Neural Information Processing Systems (NIPS)*, 2004, pp. 417–424.
- [17] M. Elbanhawi and M. Simic, “Sampling-based robot motion planning: A review,” *IEEE access*, vol. 2, pp. 56–77, 2014.
- [18] D. Berenson, S. Srinivasa, and J. Kuffner, “Task space regions: A framework for pose-constrained manipulation planning,” *The International Journal of Robotics Research*, vol. 30, no. 12, pp. 1435–1460, 2011.
- [19] C. Voss, M. Moll, and L. E. Kavraki, “Atlas+ X: Sampling-based planners on constraint manifolds,” RICE computer science, Tech. Rep., 2017.
- [20] Y. Zhang, K. Hauser, and J. Luo, “Unbiased, scalable sampling of closed kinematic chains,” in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2013, pp. 2459–2464.
- [21] M. J. F. Gales and S. S. Airey, “Product of Gaussians for speech recognition,” *Computer Speech and Language*, vol. 20, no. 1, pp. 22–40, jan 2006.
- [22] A. Bhattacharyya, “On a measure of divergence between two statistical populations defined by their probability distributions,” *Bull. Calcutta Math. Soc.*, vol. 35, pp. 99–109, 1943.
- [23] H. F. Inman and E. L. Bradley Jr, “The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities,” *Communications in Statistics-Theory and Methods*, vol. 18, no. 10, pp. 3851–3874, 1989.
- [24] A. Rényi *et al.*, “On measures of entropy and information,” in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the University of California, 1961.
- [25] J. Carpentier, R. Budhiraja, and N. Mansard, “Learning feasibility constraints for multi-contact locomotion of legged robots,” in *Robotics: Science and Systems*, 2017, p. 9p.
- [26] S. Calinon, “A tutorial on task-parameterized movement learning and retrieval,” *Intelligent Service Robotics*, vol. 9, no. 1, pp. 1–29, 2016.
- [27] Y. Yang, V. Ivan, and S. Vijayakumar, “Real-time motion adaptation using relative distance space representation,” in *Proc. Intl Conf. on Advanced Robotics (ICAR)*. IEEE, 2015, pp. 21–27.
- [28] C. Khatri and K. Mardia, “The von Mises-Fisher matrix distribution in orientation statistics,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 95–106, 1977.
- [29] A. Kume, S. Preston, and A. T. Wood, “Saddlepoint approximations for the normalizing constant of Fisher–Bingham distributions on products of spheres and stiefel manifolds,” *Biometrika*, vol. 100, no. 4, pp. 971–984, 2013.
- [30] Y. Nakamura, H. Hanafusa, and T. Yoshikawa, “Task-priority based redundancy control of robot manipulators,” *The International Journal of Robotics Research*, vol. 6, no. 2, pp. 3–15, 1987.

## APPENDIX

In this document, several transformations  $\mathcal{T}^m$  and experts models  $p_m$  related to common robotic problems are presented as a cookbook.

a) *Forward kinematics (FK)*: One of the most common transformation used in robotics is forward kinematics, computing poses (position and orientation) of links given the robot configuration  $\mathbf{x}$ . Using variational inference with a mixture of Gaussians was already proposed in [1] to find all



the distribution of configurations of a 10 DoF planar robot where the end link is close to a target.

Forward kinematics can also be computed in several task spaces associated with objects of interest [26].

*b) Center of Mass (CoM):* From the forward kinematics of the center of mass of each link and their mass, it is possible to compute the center of mass (CoM) of the robot. To satisfy static equilibrium, the CoM should be on top of the support region. In the case of a flat terrain, this region is easy to compute and is the convex hull of the robot's feet. When it is not the case, the non-slipping condition of the contact points should be ensured. A second-order cone program should be solved to determine if there exist contact forces under this condition. It is thus not trivial to determine if the CoM is on top of the support region; the function and its gradient cannot be evaluated directly. A first solution is to approximate, in a preprocessing step, the support region by a differentiable cost function. The value of this function would be 0 when inside and very high outside, while being continuous and differentiable. The other solution is to add the contact forces into the state of the system and to approximate the joint distribution of configurations and contact forces ensuring a static equilibrium.

*c) Distance:* A *relative distance space* is proposed in [27]. It computes the distances from multiple virtual points on the robot to other objects of interests (targets, obstacles). It can for example be used in environments with obstacles, providing an alternative or complementing forward kinematics.

*d) Steps of Jacobian pseudo-inverse:* Precise kinematics constraints imply a highly correlated  $\tilde{p}(\mathbf{x})$  from which it is difficult to sample and represent as a mixture of Gaussians. In the extreme case of hard kinematics constraints, the solutions are on a low dimensional manifold embedded in configuration space. Dedicated methods address the problem of representing [19] or sampling this manifold [20]. In [18], a *projection* strategy is proposed. Configurations are sampled randomly and projected back using an iterative process. We propose a similar approach where the projection operator  $\mathcal{P}_N$  would be used as transformation  $\mathcal{T}^m$ . A variety of optimization techniques can be used to solve the inverse kinematics problem. A standard way is to apply the following relation recursively

$$\mathcal{P}(\mathbf{x}) = \mathbf{x} + J(\mathbf{x})^\dagger (\hat{\mathbf{p}} - \mathbf{F}(\mathbf{x})), \quad (22)$$

where  $\hat{\mathbf{p}}$  is the target and  $J(\mathbf{x})^\dagger$  is the Moore-Penrose pseudo-inverse of the Jacobian. This relation is derivable and can be applied recursively with

$$\mathcal{P}_0(\mathbf{x}) = \mathcal{P}(\mathbf{x}) \quad (23)$$

$$\mathcal{P}_{n+1}(\mathbf{x}) = \mathcal{P}(\mathcal{P}_{n-1}(\mathbf{x})). \quad (24)$$

Then, the distribution

$$p_m(\mathbf{x}) \propto \mathcal{N}(\mathcal{P}_N(\mathbf{x}) | \hat{\mathbf{p}}, \sigma \mathbf{I}), \quad (25)$$

is the distribution of configurations which converge in  $N$  steps to  $\mathcal{N}(\hat{\mathbf{p}}, \sigma \mathbf{I})$ , see Fig. 4. Thanks to the very good

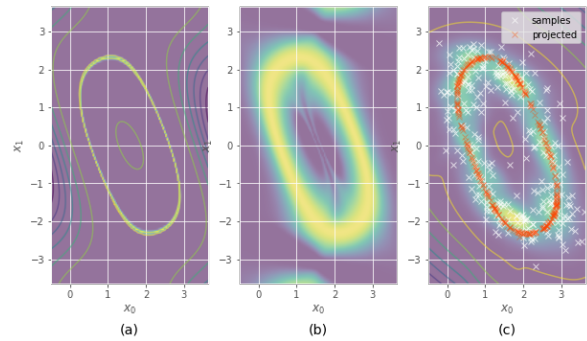


Fig. 4. Precise forward kinematics objectives often create a highly correlated distribution of joint angles, which is difficult to sample from or to estimate as a mixture of Gaussians. (a) Joint distribution of a 2-joint robot whose end-effector should stay on a line. The colormap represents  $\exp(-c(\mathbf{x}))$  and the lines are isolines of  $c(\mathbf{x})$ . Probability mass is located in yellow (or lighter) areas. (b) Distribution of points from which two iterations of IK leads to distribution (a). (c) Its approximation as a mixture of Gaussians with samples and projected samples with two iterations.

convergence of the iterative process (22),  $\sigma$  can be set very small. However, this approach has a similar (but less critical) problem as [18]. The *approximate distribution*  $q(\mathbf{x}; \lambda)$  will be slightly biased toward zones where the forward kinematics is close to linear (constant Jacobian), which are those where more mass converge to the manifold.

With high DoF robots, it might be computationally expensive to run iteration steps inside the *stochastic gradient optimization* and propagate the gradient. Another approach would be to define heuristically, or learn,  $\Sigma_h$  such that  $\mathcal{N}(\mathbf{x} | \hat{\mathbf{p}}, \sigma \mathbf{I} + \Sigma_h)$  is close to  $\mathcal{N}(\mathcal{P}_N(\mathbf{x}) | \hat{\mathbf{p}}, \sigma \mathbf{I})$ .

We present several distribution that can be used as cost on the transformations.

*e) Multivariate normal distribution (MVN):* An obvious choice for forward kinematics objective is the Gaussian or multivariate normal distribution (MVN). Its log-likelihood is quadratic, making it compatible with standard inverse kinematics and optimal control techniques,

$$p_{\text{MVN}}(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right), \quad (26)$$

where  $\boldsymbol{\mu}$  is the location parameter and  $\boldsymbol{\Sigma}$  is the covariance matrix.

*f) Matrix Bingham-von Mises-Fisher distribution (BMF):* To cope with orientation, for example represented as rotation matrix, Matrix Bingham-von Mises-Fisher distribution (BMF) [28] can be used. Its normalizing constant is intractable and requires approximation [29], which is not a problem in our case, as we integrate over robot configurations. Its density

$$p_{\text{BMF}}(\mathbf{X} | \mathbf{A}, \mathbf{B}, \mathbf{C}) \propto \exp\left(\text{tr}(\mathbf{C}^\top \mathbf{X} + \mathbf{B} \mathbf{X}^\top \mathbf{A} \mathbf{X})\right), \quad (27)$$

has a linear and a quadratic term and can be written as an MVN with a particular structure of the covariance matrix. Correlations between rotations can be represented with this distribution. Rewritten as an MVN, it is possible to create a joint distribution of position and rotation matrices that also encodes correlations between them.

g) *Cumulative distribution function (CDF)*: Inequality constraints, such as static equilibrium, obstacles or joint limits can be treated using cumulative distribution function

$$p(x \leq b), \quad \text{with } x \sim \mathcal{N}(\mathcal{T}(x), \sigma^2), \quad (28)$$

where  $\mathcal{T}(x)$  is a scalar. For example for half-plane constraints,  $\mathcal{T}(x)$  could be  $w^\top x$ , or for joint limits on first joint  $\mathcal{T}(x) = x_0$ .

The use of the CDF makes the objectives continuous and allows us to consider a safety margin determined by  $\sigma$ .

Obstacles constraints might be impossible to compute exactly and require collision checking techniques [17]. Due to the stochastic optimization, our approach is compatible with stochastic approximation of the collision related cost, which might speed up computation substantially.

h) *Uni-Gauss distribution*: To represent hierarchy between multiple tasks in our framework, we propose to use uni-Gauss experts [2]. It combines the distribution defining a non-primary objective  $p_m$  with a uniform distribution

$$p_{UG,m}(x) = \pi_m p_m(x) + \frac{1 - \pi_m}{c}, \quad (29)$$

which means that each objective has a probability  $p_m$  to be fulfilled. It can also be interpreted as a cost of  $\log(1 - \pi_m)$  penalizing the neglect of task  $m$ .

i) *Prioritization*: Classical prioritized approaches [30] exploit spatial or temporal redundancies of the robot to achieve multiple tasks simultaneously. They use a null-space projection matrix, such that commands required to solve a secondary task do not influence the primary task. As our tasks are defined as distributions, we do not necessarily need redundancies. As each non-primary tasks has a probability to be abandoned, approximating the PoE would evaluate if there is sufficient mass at the intersection of the objective.

There are two possible ways of estimating  $\tilde{p}(x)$  in case of Uni-Gauss experts. If the number of tasks is small, we can introduce, for each task  $m$ , a binary random variable indicating if the task is fulfilled or not. For each combination of these variable, we can then compute  $\tilde{p}(x)$ . The ELBO can be used to estimate the relative mass of each of these combinations, as done in model selection. For example, if the tasks are not compatible, their product would have a very small mass, as compared to the primary task. In the case of numerous objectives, this approach becomes computationally expensive because of the growing number of combinations. We can instead marginalize these variables and we fall back on (29). For practical reasons of flat gradients, the uniform distribution can be implemented as a distribution of the same family as  $p_m$  with a higher variance, that changes across optimization.

Fig. 5 (a) shows an example where forward kinematics objectives are given for the feet and a hand while static balance has to be ensured.

The proposed approach is also compatible with classical null-space formulations [30]. We can exploit the fact that stochastic variational inference does not need to evaluate the

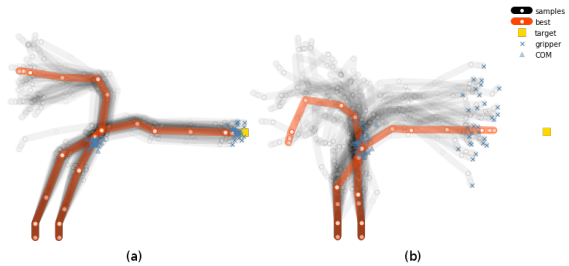


Fig. 5. Distribution of poses satisfying static equilibrium, joint limits, feet and left gripper IK objectives. Feet are projected to their exact locations using three Newton’s method steps. (a) No hierarchy between the different objectives. (b) Left gripper is set as a secondary objective using Uni-Gauss expert. Given that the secondary objective has some probability not to be fulfilled, there is more mass in the intersection of the other constraints than in the intersection of all.

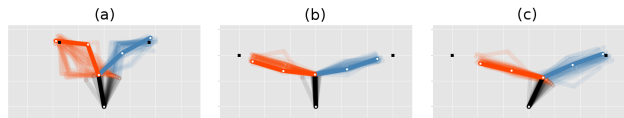


Fig. 6. 5 DoF bimanual planar robot with two forward kinematics objectives. (a) The two tasks are compatible and the distribution of solution is approximated. (b) No null-space, the two tasks are of the same importance. (c) The gradient of the objective of the orange arm is projected onto the null-space of the Jacobian of the forward kinematics of the blue arm. This is made possible because stochastic variational inference only requires to evaluate the gradient of the unnormalized density.

unnormalized density  $\tilde{p}$  but only its gradient. This characteristic is shared with only a very few Monte Carlo methods such as [3]. It allows us to filter the gradient of the different objectives (expert and transformation). The derivative of the expert log-probability with respect to the configuration  $x$  can be written as

$$\frac{\partial \log p(\mathcal{T}_m(x)|\theta_m)}{\partial x} = \frac{\partial \log p(y|\theta_m)}{\partial y} \frac{\partial \mathcal{T}_m(x)}{\partial x} \quad (30)$$

$$= \frac{\partial \log p(y|\theta_m)}{\partial y} \mathbf{J}_m(x), \quad (31)$$

where we make the Jacobian  $\mathbf{J}_m(x)$  of the transformation appear. When multiple tasks should be prioritized, a null-space filter  $\mathbf{N}_1(x)$  can be added such that the gradient of the secondary objective only acts perpendicularly to the main objective, namely

$$\frac{\partial \log p(\mathcal{T}_m(x)|\theta_1, \theta_2)}{\partial x} = \frac{\partial \log p(y|\theta_1)}{\partial y} \mathbf{J}_1(x) + \frac{\partial \log p(y|\theta_2)}{\partial y} \mathbf{J}_2(x) \mathbf{N}_1(x) \quad (32)$$

$$\text{where } \mathbf{N}_1(x) = \mathbf{I} - \mathbf{J}_1(x)^\dagger \mathbf{J}_1(x). \quad (33)$$

Note that when using automatic differentiation library, gradients can be easily redefined with this filter.

Fig. 6 shows a 5 DoF bimanual planar robot with two forward kinematics objectives. When the tasks are compatible, the filtering has no effect. The gradient of the objective of the orange arm can be projected onto the null-space of the Jacobian of the forward kinematics of the blue arm, resulting in a prioritization.