Geometry-Aware Probabilistic Shared Autonomy with Riemannian Motion Policies

Kay Pompetzki[1], Cristiana de Farias[1], Joao Carvalho[2], Georgia Chalvatzaki[1,3,4], Jan Peters[1−5]

***Abstract*—Shared autonomy in teleoperation remains challenging when human intent is ambiguous and environmental geometry restricts feasible motion. Existing policy-blending and observer-based methods often struggle to reconcile operator input, autonomy, and constraints in cluttered environments. To address this issue, we propose a geometry-aware probabilistic framework that combines Riemannian Motion Policies with entropic-regularized Bayesian inference for smooth multimodal arbitration and reliable control. Across two navigation tasks, our approach improves task success, alignment with user intent, and operator effort relative to established baselines.**

I. Introduction

In assisted teleoperation, shared autonomy combines human input with autonomous assistance to improve performance, enforce constraints, and reduce workload [1]. A central challenge is providing assistance that remains consistent with both human intent and task requirements, especially when these are uncertain or conflicting. Existing approaches typically follow two paradigms. Policy-blending methods combine human and autonomous commands through arbitration, often via weighted averaging [2]–[9]. Observer-based methods instead treat human input as an observation and generate motion autonomously [10]–[13]. While effective in structured settings, both methods struggle in cluttered or changing environments. Geometric variations, such as unfamiliar configurations or structural modifications, introduce ambiguity, restrict feasible motions, and may conflict with operator intent. This challenge comes from the designed interaction between humans and autonomy. In such scenarios, policy blending allows direct human influence but ignores geometric constraints, while observer-based methods limit the operator's ability to correct behavior when autonomy fails.

We instead treat the operator as an adaptive, geometry-aware policy capable of resolving ambiguities, escaping local minima, and exploring feasible motions. This requires shared autonomy to (i) integrate heterogeneous behaviors and constraints in a geometric manner, (ii) consistently fuse them with operator input, and (iii) maintain multiple hypotheses under uncertainty, including ambiguous human intent. Thus, we propose a shared-autonomy architecture combining Riemannian motion policy (RMP) [14]–[16] with entropic-regularized Bayesian filtering. RMP provides a metric-aware framework for composing behaviors into a stable controller, with the operator input being geometrically projected alongside autonomous behaviors. A multimodal Bayesian belief maintains hypotheses over behaviors, while entropic regularization ensures smooth arbitration, avoiding overcommitment or oscillations. This belief acts as a high-level arbitrator, dynamically weighting task-specific RMPs.

Our contributions are: (i) A geometry-aware shared-autonomy framework based on RMP combining operator input, autonomous skills, and online constraints into a single stable motion command. (ii) An entropic-regularized Bayesian filtering layer or multimodal arbitration under uncertainty. (iii) A reformulation of the operator as an active exploratory component in shared autonomy. (iv) A unified geometric–probabilistic framework bridging policy blending and observer methods.

II. Preliminaries

Geometric control as Inference: For reactive planning and control, many works compose local policies into a global behavior. Early approaches use superposition of attractive and repulsive fields in task space [17] and were later extended to operational space [18]–[20]. Later works generalize to non-Euclidean spaces employing Riemannian or Finsler geometry [14]–[16], [21], [22]. These approaches typically use differentiable task maps to project the policies' behaviors into a common space where they combine them. Most approaches formulate the composition of the local policies as a geometric least-squares problem

$$\begin{aligned}\boldsymbol{a}^r &= \arg\min_{\boldsymbol{a}\in\mathcal{A}} \mathcal{J}_{\mathrm{C}}(\boldsymbol{o}_t,\boldsymbol{a}) = \arg\min_{\boldsymbol{a}\in\mathcal{A}} \textstyle\sum_i w_i \mathcal{J}_{\mathrm{C}_i}(\boldsymbol{o}_t,\boldsymbol{a})\\ &= \arg\min_{\boldsymbol{a}\in\mathcal{A}} \textstyle\sum_i w_i(\boldsymbol{a}-\boldsymbol{a}_i^{\mathrm{des}}(\boldsymbol{o}_t))^\top \boldsymbol{M}_i(\boldsymbol{o}_t)(\boldsymbol{a}-\boldsymbol{a}_i^{\mathrm{des}}(\boldsymbol{o}_t)),\end{aligned}$$

where $\boldsymbol{a}^r$ denote the optimal action. $\boldsymbol{a}_i^{\mathrm{des}}$ being the desired acceleration from local policy i. A positive-semidefinite matrix $\boldsymbol{M}_i$ represents the metric or priority of each local policy. Both $\boldsymbol{a}_i^{\mathrm{des}}$ and $\boldsymbol{W}_i$ depend on observations $\boldsymbol{o}_t$ containing states and other contextual information. For clarity, we omit all dependencies on joint states, task maps, and task-space representations. This composition can also be addressed from a probabilistic perspective using energy based models (EBMs) [13], [20], [23]–[25]. and highlights a duality between geometric control and probabilistic inference.

III. Methodology

In goal-centric shared autonomy methods, the autonomy must infer two factors to assist the human effectively: (i)

[1]Technical University of Darmstadt [2]German Research Center for Artificial Intelligence (DFKI) [3]Robotics Institute Germany (RIG) [4] Hessian center for artificial intelligence (hessian.AI) [5]Centre for Cognitive Science; Corresponding author: kay@robot-learning.de

This work was supported by the German Federal Ministry of Research, Technology and Space (BMFTR) under the Robotics Institute Germany (RIG) and the DEMETER project (Grant no.: 01DR25003), the EU's Horizon Europe project ARISE (Grant no.: 101135959)

the human's intention and (ii) whether assistance is required. Let $s_t^r \in \mathcal{S}$ and $a_t^r \in \mathcal{A}^r$ be the robot's state and action, respectively. An observation at time t is given by $o_t = (s_t^r, a_t^h, c_t) \in \mathcal{S} \times \mathcal{A}^h \times \mathcal{C}$, comprising of the robot's state, the human's action, and additional context encountered at t. Here, we follow the traditional wait-and-move strategy [2], [5], [26], where assistance is conditioned on an inferred goal g drawn from the space of possible goals, which is a discrete or continuous subset of the state space, i.e., $g \in \mathcal{G} \subset \mathcal{S}$. In our work, we assume a discrete set of goals and want to estimate a belief b_t over the latent variable $z = [g, c]^\intercal$. The first latent variable, $c \in [0,1]^{n_\mathcal{G}}$, corresponds to selecting the correct goal given the observation. The second $c \in [0,1]^2$ represents whether assistance is required. We keep this notation to later extend uncertainty beyond the target distribution to assistance levels. We factorize the joint distribution

$$p(a_{0:T}^r, o_{0:T}, z_{0:T}, \mathcal{O}_{0:T}) = \tfrac{1}{Z} p(o_0) p(z_0) p(a_0^r) \quad (1)$$
$$\prod_{t=1}^{T} p(\mathcal{O}_t \mid o_t, z_t, a_t^r) p(o_t \mid o_{t-1}, a_{t-1}^r) p(z_t \mid z_{t-1}) p(a_t^r)$$

This model treats the robot's action a^r as a random variable. A binary variable $\mathcal{O}_t \in 0,1$ encodes task optimality at each time step [27], [28]. It depends on the observation o_t, belief b_t, and action a_t^r. We model it using a Gibbs likelihood $p(\mathcal{O}_t = 1 \mid o_t, z_t, a_t^r) \propto \exp(-\mathcal{J}_\text{C}(o_t, z_t, a_t^r))$, with a task-specific objective $\mathcal{J}_\text{C}(\cdot)$. For brevity, we omit $\mathcal{O}_t = 1$ in equation 1. The objective may be an aggregated cost $\mathcal{J}_\text{C}(\cdot) = \sum_i w_i \mathcal{J}_{\text{C}_\text{i}}(\cdot)$ where each weight $w_i \in \mathbb{R}_+$ represents a temperature parameter. These cost components may relate separately to belief update and action selection.

Sequential Belief Update with Entropic Mirror Descent: First, we aim to infer a belief $p(z_t \mid a_{:T}^r, o_{:t}, \mathcal{O}_{:t})$ over the latent state z_t from the history seen so far up to time step t. Assuming the Markov property and the joint distribution in equation 1, the posterior factorizes as

$$p(z_t \mid o_{:t}, \mathcal{O}_{:t}) \propto \int_\mathcal{A} p(\mathcal{O}_t \mid o_t, z_t, \hat{a}^r) p(\hat{a}^r)\, \mathrm{d}\hat{a}^r \quad (2)$$
$$\int_\mathcal{Z} p(z_t \mid \hat{z}_{t-1}) b_{t-1}(\hat{z}_{t-1})\, \mathrm{d}\hat{z}_{t-1}$$

where $b_{t-1}(\hat{z}_{t-1}) \approx p(\hat{z}_{t-1} \mid o_{:t}, \mathcal{O}_{:t})$ represents the previous belief. The first term represents prediction step $\hat{b}_t(\hat{z}_t)$ on the transition model $p(z_t \mid \hat{z}_{t-1})$. The second corresponds to the likelihood step $p(\mathcal{O}_t \mid o_t, z_t)$ based on a new observation [29]. Given a previous belief b_{t-1}, we are minimizing the KL-divergence subject to a relative-entropy constraint

$$b_t^\star(z) = \arg\min_{b \in \mathcal{P}_b} \mathbb{D}_\text{KL}(b(z), p(z \mid o_{:t}, \mathcal{O}_{:t})) \quad (3)$$
$$\text{s.t. } \mathbb{D}_\text{KL}(b(z), b_{t-1}(z)) \leq \epsilon, \int b(z)\, \mathrm{d}z = 1,$$

where $\mathcal{P}_b$ denotes the space of admissible belief distributions. This formulation corresponds to an entropic mirror descent [30], [31] update combined with a sequential importance sampling scheme [29]. The constraint on relative entropy prevents abrupt changes in belief, leading to a smooth evolution of the estimated posterior. The resulting closed-form update is

$$b_t(z) \propto b_{t-1}(z) \left[\frac{\hat{b}_t(z)}{b_{t-1}(z)}\right]^\eta p(\mathcal{O}_t \mid o_t, z)^\eta$$

where $\eta \in \mathbb{R}_+$ acts as an inverse temperature parameter that interpolates between the prior and likelihood contributions. The likelihood term incorporates an importance ratio between the predictive model $\hat{b}(\hat{z})$ and the prior $b_{t-1}(\hat{z})$. Thus, it predicts the change in the belief based on this ratio and corrects the estimate, taking into account the evidence from the observations.

We consider the belief $b_t(\cdot)$ over users intention and arbitration at time t represented hierarchically as

$$b_t(z) = w_1^c \delta(c_1) + w_2^c \delta(c_2) \sum_{i=1}^{n_\mathcal{G}} w_i^g \delta(g_i),$$
$$\text{s.t. } w_1^c + w_2^c \sum_{i=1}^{n_\mathcal{G}} w_i^g = 1.$$

At the inner level, $w^g \in \Delta^{n_\mathcal{G}-1}$ encodes the relevance of each assistant policy. At the outer level, $w^c \in \Delta^1$ represents the arbitration between user and autonomy. Although equivalent to a flat mixture over $n_\mathcal{G} + 1$ hypotheses, the hierarchical parameterization improves interpretability and separates human intent inference from policy blending.

Action selection through approximate Inference: We assume that the robot follows a policy that corresponds to

$$\pi_t(a_t^r \mid o_t) \propto \mathbb{E}_{z_t \sim b_t(\cdot)} [p(\mathcal{O}_t \mid o_t, z_t, a_t^r)]\, p(a_t^r).$$

given a curret estimate of $b_t(\cdot)$ at time t. Thus, we approximate it using a point estimate via maximum a posteriori (MAP) inference at equation 2. For our choice of RMP, this assumption is valid and computationally efficient as the MAP estimate aligns with the optimal action. Under a geometrical control-as-inference view, the posterior can be expressed as the EBM

$$\pi_t(a_t^r \mid o_t) \propto p(a_t^r) p(\mathcal{O}_t \mid o_t, a_t^r)$$
$$\sum_{i=1} w^i p(\mathcal{O}_t \mid o_t, z_t^i, a_t^r).$$

Like in prior work [13], we split the posterior into geometry parts, such as obstacle avoidance, and z_t^i-dependent parts, such as reaching-target policies or a velocity-tracking policy. The optimal action a_t^r can be computed in closed form and represents a softmax-weighted mixture over the different RMP policies. It resembles expectation-maximization like structure. We first update the belief over goals (e-step). Then we optimize the action given that belief (m-step).

The operator is modeled as a time-varying RMP $(M_{A_i}^\mathcal{H}, f_{A_i}^\mathcal{H})$ representing a velocity tracking policy. We model it as a quadratic cost in the tangent space of $SE(3)$ that minimizes the deviation between the robot's current twist and a desired twist provided by the operator. This operator-induced policy is incorporated as an additional expert in the energy-based action posterior.

IV. Experiments and Results

As a proof of concept, we investigate: (i) how does task performance scale with increasing task complexity; (ii) how well does the assisting output align with operator intent; and (ii) how does this affect intervention effort. Thus, we evaluate these questions in $2D$ navigation tasks with varying levels of complexity and goal ambiguity. The point mass must be steered from randomly sampled start locations to one of multiple goals. Two scenarios are used: *Toy Task*, where the primary

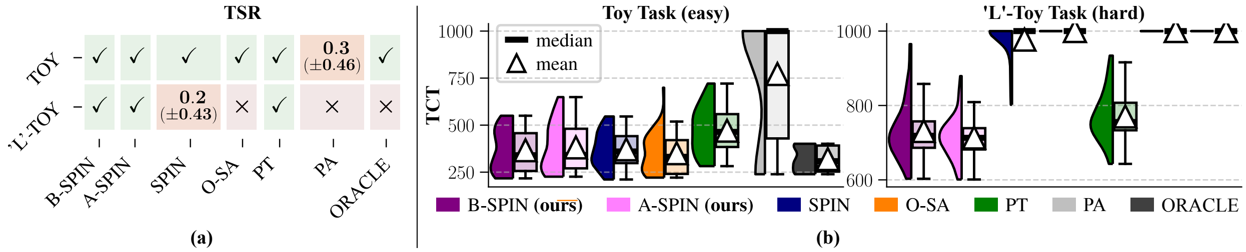

Fig. 1. The task success rate (TSR) and task completion time (TCT) for all methods across all evaluation stages. (a) TSR: A checkmark indicates full task success, numerical values show partial success rates (mean ± std), and a cross indicates complete failure. (b) TCT: Aggregated results across $\mathbb{R}^2$ and $SE(2)$ and all trials from the six operators. Columns represent the two complexity levels, from *No-Obstacle* to *L-Obstacle*. Each violin shows the full data distribution, with the boxplot indicating the mean, median, and interquartile range. B-SPIN and A-SPIN consistently solve the task and achieve lower TCT than the other assisted methods across all levels of complexity.

challenge is resolving goal ambiguity, and *L-Toy Task*, which adds an L-shaped obstacle inducing a strict local minimum for myopic controllers. Task difficulty is further varied by expanding the action space from 2D translations, i.e., $\mathbb{R}^2$, to $SE(2)$ control with an additional orientation component. To get an estimate of the algorithmic properties across different operator, the study is conducted with six expert participants who have prior experience with the system.

We compare our approach, B-SPIN, against structured ablations that isolate key design choices. A-SPIN replaces the belief-weighted goal distribution with a MAP estimate to evaluate hard goal selection. SC-SPIN removes adaptive arbitration by enforcing constant blending, representing shared-control approaches. O-SA removes direct operator control entirely, treating the operator as an observation for belief updates, corresponding to observer-based shared autonomy. We further include PT as a no-assistance baseline, as well as fully autonomous variants PA and ORACLE, the latter assuming access to the true goal as an oracle. All methods share identical RMP leaves, including goal attraction, obstacle avoidance, and smoothness, to ensure fair comparison. We evaluate performance in terms of task success rate (TSR), task completion time (TCT), task-space smoothness (TS), alignment (AL), operator effort (OE). These measures jointly capture effectiveness, feasibility, and usability.

In figure 1 and figure 2, results are aggregated across action-spaces and expert participants. The autonomous PA and ORACLE are included as reference points for task-space smoothness (TS) and task completion time (TCT) under fully autonomous control. The results indicate that across both action-spaces increasing complexity in form of obstacles causes a performance drop for several assisted teleoperation methods. Weighted SPIN (B-SPIN), argmax SPIN (A-SPIN), and pure teleoperation (PT) maintain a high task success rate, with B-SPIN, A-SPIN at substantially lower operator effort. In contrast, observer-based shared autonomy (O-SA) and shared-control SPIN (SC-SPIN) show higher TS and operator effort (OE) alongside a lower alignment (AL). This indicates a disagreement between the resulting action command and operator intent that likely increases mental demand of the participant. As the primary uncertainty in the navigation task is inferring the correct goal, B-SPIN and A-SPIN perform comparably, suggesting that hard MAP goal selection is sufficient when goal ambiguity is the dominant difficulty.

V. Conclusion

This work presented a geometry-aware shared autonomy framework that combines goal inference, belief-based arbitration, and policy blending. The proposed framework allows the operator to remain in control when autonomy fails, while still benefiting from assistance when it is reliable. In future work, we will evaluate our method on high-dimensional manipulation. We will further conduct a user study examining how operators develop trust in the system and whether reduced operator effort translates to lower perceived mental demand.

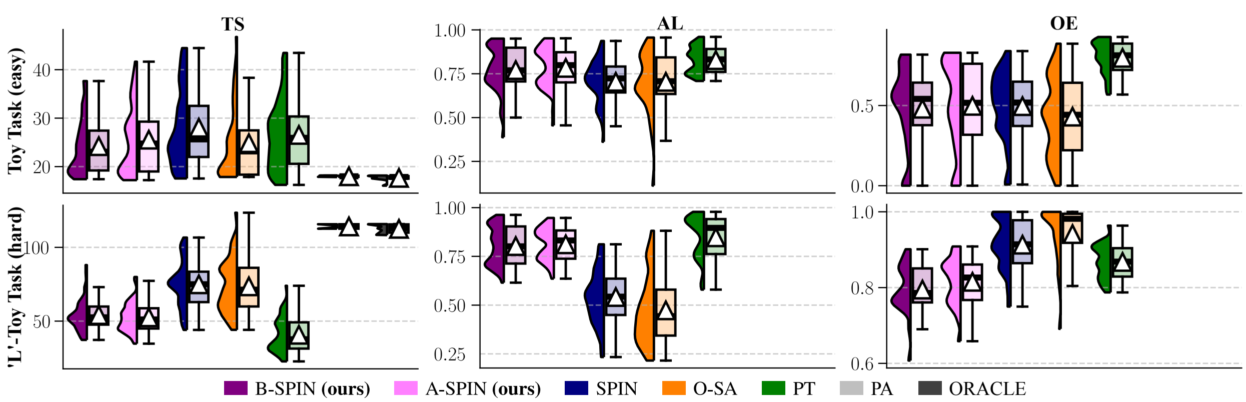

Fig. 2. Aggregated results across $\mathbb{R}^2$ and $SE(2)$ and all trials from the six operators. Rows represent the two complexity levels, from *No-Obstacle* and *L-Obstacle*. Columns show task-space smoothness (TS), alignment (AL), and operator effort (OE). Each violin shows the full data distribution, with the boxplot indicating the mean, median, and interquartile range. B-SPIN and A-SPIN consistently achieve lower OE as well as higher AL than the other assisted methods across all levels of complexity.

References

[1] M. Selvaggio, M. Cognetti, S. Nikolaidis, S. Ivaldi, and B. Siciliano, "Autonomy in physical human-robot interaction: A brief survey," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7989–7996, 2021.

[2] A. D. Dragan and S. S. Srinivasa, "A policy-blending formalism for shared control," *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 790–805, 2013.

[3] K. Muelling, A. Venkatraman, J.-S. Valois, J. E. Downey, J. Weiss, S. Javdani, M. Hebert, A. B. Schwartz, J. L. Collinger, and J. A. Bagnell, "Autonomy infused teleoperation with application to brain computer interface controlled manipulation," *Autonomous Robots*, vol. 41, pp. 1401–1422, 2017.

[4] M. Khoramshahi and A. Billard, "A dynamical system approach to task-adaptation in physical human-robot interaction," *Autonomous Robots*, vol. 43, pp. 927–946, 2019.

[5] S. Jain and B. Argall, "Probabilistic human intent recognition for shared autonomy in assistive robotics," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 9, no. 1, pp. 1–23, 2019.

[6] H. J. Jeon, D. P. Losey, and D. Sadigh, "Shared autonomy with learned latent actions," *arXiv preprint arXiv:2005.03210*, 2020.

[7] G. Maeda, "Blending primitive policies in shared control for assisted teleoperation," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 9332–9338.

[8] A. Gottardi, S. Tortora, E. Tosello, and E. Menegatti, "Shared control in robot teleoperation with improved potential fields," *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 3, pp. 410–422, 2022.

[9] P. Song, P. Li, E. Aertbeliën, and R. Detry, "Robot trajectron: Trajectory prediction-based shared control for robot manipulation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 5585–5591.

[10] S. Javdani, S. S. Srinivasa, and J. A. Bagnell, "Shared autonomy via hindsight optimization," *Robotics science and systems: online proceedings*, vol. 2015, pp. 10–15 607, 2015.

[11] S. Javdani, H. Admoni, S. Pellegrinelli, S. S. Srinivasa, and J. A. Bagnell, "Shared autonomy via hindsight optimization for teleoperation and teaming," *The International Journal of Robotics Research*, vol. 37, no. 7, pp. 717–742, 2018.

[12] S. Nikolaidis, Y. X. Zhu, D. Hsu, and S. Srinivasa, "Human-robot mutual adaptation in shared autonomy," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 2017, pp. 294–302.

[13] M. Allenspach, M. Pantic, R. Girod, L. Ott, and R. Siegwart, "Task adaptation in industrial human-robot interaction: Leveraging riemannian motion policies," *Robotics: Science and System XX*, p. 026, 2024.

[14] N. D. Ratliff, J. Issac, D. Kappler, S. Birchfield, and D. Fox, "Riemannian motion policies," *arXiv preprint arXiv:1801.02854*, 2018.

[15] C.-A. Cheng, M. Mukadam, J. Issac, S. Birchfield, D. Fox, B. Boots, and N. Ratliff, "Rmpflow: A geometric framework for generation of multitask motion policies," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 3, pp. 968–987, 2021.

[16] A. Li, C.-A. Cheng, M. A. Rana, M. Xie, K. Van Wyk, N. Ratliff, and B. Boots, "Rmp2: A structured composable policy class for robot learning," *arXiv preprint arXiv:2103.05922*, 2021.

[17] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *The international journal of robotics research*, vol. 5, no. 1, pp. 90–98, 1986.

[18] ——, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *IEEE Journal on Robotics and Automation*, vol. 3, no. 1, pp. 43–53, 1987.

[19] J. Nakanishi, R. Cory, M. Mistry, J. Peters, and S. Schaal, "Operational space control: A theoretical and empirical comparison," *The International Journal of Robotics Research*, vol. 27, no. 6, pp. 737–757, 2008.

[20] M. Toussaint and C. Goerick, "A bayesian view on motor control and planning," in *From Motor Learning to Interaction Learning in Robots*. Springer, 2010, pp. 227–252.

[21] A. Bylard, R. Bonalli, and M. Pavone, "Composable geometric motion policies using multi-task pullback bundle dynamical systems," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 7464–7470.

[22] K. Van Wyk, M. Xie, A. Li, M. A. Rana, B. Babich, B. Peele, Q. Wan, I. Akinola, B. Sundaralingam, D. Fox, *et al.*, "Geometric fabrics: Generalizing classical mechanics to capture the physics of behavior," *IEEE Robotics and Automation Letters*, 2022.

[23] K. Hansel, J. Urain, J. Peters, and G. Chalvatzaki, "Hierarchical policy blending as inference for reactive robot control," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 10 181–10 188.

[24] A. T. Le, K. Hansel, J. Peters, and G. Chalvatzaki, "Hierarchical policy blending as optimal transport," in *Learning for Dynamics and Control Conference*. PMLR, 2023, pp. 797–812.

[25] J. Urain, A. Li, P. Liu, C. D'Eramo, and J. Peters, "Composable energy policies for reactive motion generation and reinforcement learning," *The International Journal of Robotics Research*, vol. 42, no. 10, pp. 827–858, 2023.

[26] C. Ezeh, P. Trautman, L. Devigne, V. Bureau, M. Babel, and T. Carlson, "Probabilistic vs linear blending approaches to shared control for wheelchair driving," in *2017 International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 2017, pp. 835–840.

[27] M. Toussaint, "Robot trajectory optimization using approximate inference," in *Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 1049–1056.

[28] S. Levine, "Reinforcement learning and control as probabilistic inference: Tutorial and review," *arXiv preprint arXiv:1805.00909*, 2018.

[29] S. Särkkä and L. Svensson, *Bayesian filtering and smoothing*. Cambridge university press, 2023, vol. 17.

[30] A. Beck and M. Teboulle, "Mirror descent and nonlinear projected subgradient methods for convex optimization," *Operations Research Letters*, vol. 31, no. 3, pp. 167–175, 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167637702002316

[31] J. Peters, K. Mulling, and Y. Altun, "Relative entropy policy search," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 24, no. 1, 2010, pp. 1607–1612.