Structured and Distributed Representations for Perception and Control

Alexander Sasha Lambert

In order to make learning and inference for robotics more data-efficient and scalable, we should integrate domain knowledge into our representations when possible [1], with careful consideration of inductive biases. Luckily, robotics problems often exhibit rich *structure* and known dependencies which can be exploited in different ways, particularly when learning predictive models for inference and control. Such examples include system appearance and geometry [2, 3], kinematics [4], and physics-based priors [5, 6]. A large part of my research has been motivated by this notion, where I have addressed the problem of embedding structure in perceptual models for both visual and tactile sensing modalities [3, 7]. This includes a framework for cross-modal compensation and efficient inference in environments with severe occlusion [8].

More recently, I have turned to the question of efficient representations for control and planning problems. Sampling-based approaches to model predictive control (MPC) have risen in popularity, largely due to their speed and ease of implementation in model-based reinforcement learning schemes, as well as their success in noisy real-world environments [9, 10]. These methods resort to open-loop, Monte Carlo (MC) sampling for estimating expected costs over finite-length, stochastic trajectories using simple uni-modal control distributions [11, 12]. This can make it challenging to resolve multi-modal, complex posteriors which might arise from non-convexity of the optimization problem (ex. due to obstacles) or state uncertainty (ex. from localization). My work has focused on leveraging non-parametric representations for optimal control distributions in MPC and planning problems. These are *distributed*, in that they consist of a collection of unique parameters which require local evaluation, but interact in a de-centralized way. By casting optimization as a Bayesian inference problem and leveraging recent developments in particle-based Variational Inference (ParVI), sample-efficient control schemes can be achieved by maintaining a system of interacting particles over a distribution of control parameters.

With the growth in availability of on-board GPUs, we should approach online control and inference with parallelization in mind, and extend ideas from the statistics community to efficiently handle high-dimensional uncertainty common in robotics. I advocate that applying non-parametric approaches to online learning and prediction will allow systems to become more robust and adaptive in noisy and dynamic environments.

I. Flow-Based Models for Visual Prediction

Predictive models for visual data, such as image frames or video, define a mapping from a latent space to pixel-level observations. They have been used for learning unsupervised visuo-motor policies [2, 13], visual task planning [14], and model-predictive control [15–17]. This has demonstrated the utility of defining desired *visual* states and trajectories directly in observation space, for both manipulation and navigation tasks. However, prediction using purely parameterized deep networks, such as de-convolutional networks, VAEs or GANs, often suffers in quality (ex. overly blurry images) or requires a large amount of data to train on. In [3], we devised a method for predicting photo-realistic observations in robot manipulation by leveraging a key fact: the geometry and kinematics of the system is effectively constant, and the configuration space is a well-defined closed set. By collecting key-frame data of different robot poses, a flow-based transformation can be learned to generate novel viewpoints from nearest-neighbour images. This can then be used for visual prediction of desired joint-space trajectories, with the added benefit of detecting occlusions in cluttered scenes.

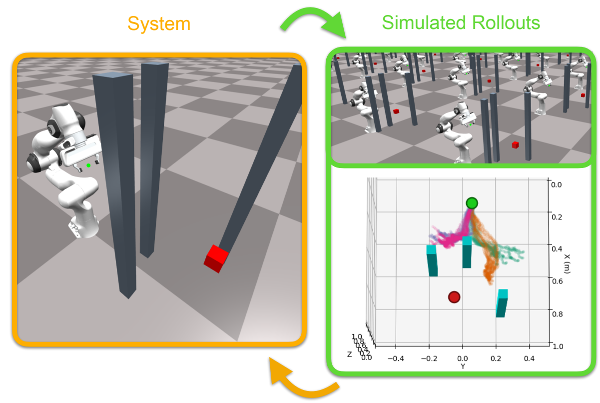

Fig. 1: Particle-based VI-MPC for a 7-dof reaching task. Expected behavior of each particle is depicted by a uniquely colored set of end-effector trajectories.

II. Structured Models for Tactile Sensing

The advent of sophisticated tactile sensors [18] has provided increased sensitivity to forces induced by contact dynamics, allowing for a diversity of applications in robotics ranging from object class and pose identification, surface texture reasoning, and slip detection [19–24]. Yet, the output signals of these devices are often noisy and difficult to interpret. There has been a lack of accurate, generalizable models which can correctly map raw sensory signals to useful force information across different tasks. Providing reliable, continuous measurements on force direction and magnitude could allow for improved robustness in controller design used in contact-rich manipulation. We addressed these shortcomings in [7] using a large-scale data-driven approach, where training examples were collected across different contact domains in order to learn a tactile sensor model. To improve prediction of directional force measurements, the spatial structure and surface geometry of the sensor was encoded directly into the architecture of the network used for the model. The proposed method was shown to outperform current state-of-the-art methods for force estimation using the same device, including both analytic [25] and learned [26] baselines.

III. Joint Inference for Multi-Modal Sensing

Having used different sensor modalities in my work, a natural extension was to tackle the question of how to combine tactile and visual measurements for efficient inference and state estimation. Occlusion is a common issue for visual object tracking in robot manipulation, particularly in cluttered scenes or during in-hand re-grasping. Tactile sensing offers an additional modality that can compensate for such partial observability in contact-rich tasks. Likewise, estimation of contact-point and force normals can be informed by visually-tracked poses and knowledge of system dynamics. Inspired by previous work [27, 28], we leveraged a factor-graph representation developed in the SLAM community for multi-modal sensor fusion [29]. In [8], we describe a framework for joint inference over visuo-tactile measurements which integrates geometric and physics-based priors (such as quasi-static mechanics) to minimize state uncertainty during task execution. This was demonstrated to improve both contact-force and pose estimation for non-prehensile and under-actuated object manipulation in heavily occluded scenes, combining observations from tactile/force-torque sensors and depth-based object tracking.

IV. Variational Inference for Control

Approximate inference has been widely explored for Stochastic Optimal Control and risk-sensitive, or entropy-regularized, MPC [9, 30–33]. Solutions have typically resorted to importance-sampling schemes using high-entropy, open-loop control distributions [11, 12], where dense Monte-Carlo sampling is generally required to mitigate noisy system behavior. Otherwise, using narrower, low-entropy distributions can give rise to greedy, highly optimistic action selection. This can become problematic when considering high-dimensional control inputs, where sample-efficiency becomes increasingly important in resolving expected costs over finite horizons. Instead, we can consider the full posterior distribution over control parameters, and attempt to resolve the multi-modal probabilities of value-weighted actions. In [34], we accomplish this by formulating MPC as a Variational Inference (VI) problem. The posterior is approximated as a distributed set of particles, where each particle constitutes a control or decision trajectory (Fig.1). We use a recent kernel-based ParVI algorithm, Stein Variational Gradient Descent, or SVGD [35], to adapt the distribution in an online fashion. Favorable performance in dealing with local minima is observed in common robot scenarios, including manipulation and navigation. The approach allows for gradient-based information to be derived from differentiable cost and dynamics models, and the algorithm can be trivially modified to solve motion planning tasks for deterministic systems. In more recent work, we combine this approach with a non-parametric filtering algorithm for online parameter adaptation to resolve model uncertainty [36].

V. Future Directions

Going forward, I plan to use particle-based MPC to accelerate reinforcement learning for continuous-control tasks. This will specifically target actor-critic frameworks which incorporate model-based control with value-function learning [37, 38]. Generating low-variance, sample-efficient approximations of expected rewards is crucial for efficient learning in this context. By defining a non-parametric i.e. *distributed* policy over action sequences, ParVI methods can be used for resolving multi-modal action distributions and improve value estimation. They are known to be more sample efficient, and converge faster, than classical Markov-Chain Monte Carlo methods [39, 40]. SVGD, as a particular case of ParVI, makes use of likelihood gradients, which has the potential to further reduce variance by back-propagating through the rolled-out state transitions (ex. via the re-parameterization trick for differentiable state transitions). However, we must ensure adequate scaling of particle dynamics to higher dimensions by using appropriately structured and factorized kernels [41, 42].

Additionally, incorporating parameter uncertainty into model-based reinforcement learning has been generally limited to the episodic setting [43, 44]. To realize safe and robust learning, we should consider how the agent manages uncertainty and adapts its belief during execution. This would mean including online parameter estimation and adaptive control for minimizing Bayesian regret at both the continual and episodic level [45, 46]. This could be addressed using particle-based methods for resolving complex posterior distributions

Further, by considering a Bayesian formulation of model-predictive control, we can incorporate priors over action spaces in a principled way. Obtaining meaningful priors is non-trivial. However these can be derived from expert or human demonstrations [47, 48], learned from experience [49, 50] or take the form of a trajectory or skill library [51]. Ideally, such informed priors may be conditioned on the context, such as the task and environment setting [52].

With the availability of fast, GPU-accelerated simulators [53] and increasingly sophisticated methods for bridging the sim-to-real gap [54–57], it is becoming conceivable to employ simulators within sampling-based control and state-estimation loops during real-time execution [58, 59]. However, to effectively utilize such parallelized computation, we need principled methods for resolving high-dimensional uncertainty over actions and model parameters. This could be achieved by considering the natural, non-Euclidean geometry induced by system kinematics and constraints of the system [60]. By ensuring that our sampling space lies on a known Riemannian manifold, for example, we can improve sample efficiency by implicitly accounting for system geometry [61].

References

[1] Niko Sünderhauf, Oliver Brock, Walter Scheirer, Raia Hadsell, Dieter Fox, Jürgen Leitner, Ben Upcroft, Pieter Abbeel, Wolfram Burgard, Michael Milford, et al. The limits and potentials of deep learning for robotics. *The International Journal of Robotics Research*, 37(4-5):405–420, 2018.

[2] Arunkumar Byravan, Felix Leeb, Franziska Meier, and Dieter Fox. Se3-pose-nets: Structured deep dynamics models for visuomotor control. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3339–3346. IEEE, 2018.

[3] Alexander Lambert, Amirreza Shaban, Amit Raj, Zhen Liu, and Byron Boots. Deep forward and inverse perceptual models for tracking and prediction. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 675–682. IEEE, 2018.

[4] Arunkumar Byravan and Dieter Fox. Se3-nets: Learning rigid body motion using deep neural networks. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 173–180. IEEE, 2017.

[5] Rico Jonschkowski, Roland Hafner, Jonathan Scholz, and Martin Riedmiller. Pves: Position-velocity encoders for unsupervised learning of structured state representations. *arXiv preprint arXiv:1705.09805*, 2017.

[6] Michael Lutter, Christian Ritter, and Jan Peters. Deep lagrangian networks: Using physics as model prior for deep learning. In *International Conference on Learning Representations*, 2018.

[7] Balakumar Sundaralingam, Alexander Sasha Lambert, Ankur Handa, Byron Boots, Tucker Hermans, Stan Birchfield, Nathan Ratliff, and Dieter Fox. Robust learning of tactile force estimation through robot interaction. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 9035–9042. IEEE, 2019.

[8] Alexander Sasha Lambert, Mustafa Mukadam, Balakumar Sundaralingam, Nathan Ratliff, Byron Boots, and Dieter Fox. Joint inference of kinematic and force trajectories with visuo-tactile sensing. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 3165–3171. IEEE, 2019.

[9] Grady Williams, Paul Drews, Brian Goldfain, James M Rehg, and Evangelos A Theodorou. Aggressive driving with model predictive path integral control. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1433–1440. IEEE, 2016.

[10] Nolan Wagener, Ching-An Cheng, Jacob Sacks, and Byron Boots. An online learning approach to model predictive control. In *Proceedings of Robotics: Science and Systems (RSS)*, 2019.

[11] Zdravko I Botev, Dirk P Kroese, Reuven Y Rubinstein, and Pierre L'Ecuyer. The cross-entropy method for optimization. In *Handbook of statistics*, volume 31, pages 35–59. Elsevier, 2013.

[12] Grady Williams, Nolan Wagener, Brian Goldfain, Paul Drews, James M Rehg, Byron Boots, and Evangelos A Theodorou. Information theoretic mpc for model-based reinforcement learning. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1714–1721. IEEE, 2017.

[13] Chelsea Finn, Ian Goodfellow, and Sergey Levine. Unsupervised learning for physical interaction through video prediction. In *Advances In Neural Information Processing Systems*, pages 64–72, 2016.

[14] Chris Paxton, Yotam Barnoy, Kapil Katyal, Raman Arora, and Gregory D Hager. Visual robot task planning. In *2019 international conference on robotics and automation (ICRA)*, pages 8832–8838. IEEE, 2019.

[15] Chelsea Finn and Sergey Levine. Deep visual foresight for planning robot motion. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2786–2793. IEEE, 2017.

[16] Frederik Ebert, Chelsea Finn, Alex X Lee, and Sergey Levine. Self-supervised visual planning with temporal skip connections. In *CoRL*, pages 344–356, 2017.

[17] Noriaki Hirose, Fei Xia, Roberto Martin-Martin, Amir Sadeghian, and Silvio Savarese. Deep visual mpc-policy learning for navigation. *IEEE Robotics and Automation Letters*, 4(4):3184–3191, 2019.

[18] Pedro Silva Girão, Pedro Miguel Pinto Ramos, Octavian Postolache, and José Miguel Dias Pereira. Tactile sensors for robotic applications. *Measurement*, 46(3):1257–1271, 2013.

[19] Danfei Xu, Gerald E Loeb, and Jeremy A Fishel. Tactile identification of objects using bayesian exploration. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 3056–3061, 2013.

[20] Rui Li, Robert Platt, Wenzhen Yuan, Andreas ten Pas, Nathan Roscup, Mandayam A Srinivasan, and Edward Adelson. Localization and manipulation of small parts using GelSight tactile sensing. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3988–3993. IEEE, 2014.

[21] Filipe Veiga, Herke van Hoof, Jan Peters, and Tucker Hermans. Stabilizing Novel Objects by Learning to Predict Tactile Slip. In *iros*, 2015.

[22] Nicholas Wettels, Avinash R Parnandi, Ji-Hyun Moon, Gerald E Loeb, and Gaurav S Sukhatme. Grip control using biomimetic tactile sensing systems. *IEEE/ASME Transactions On Mechatronics*, 14(6):718–723, 2009.

[23] Ioannis Agriomallos, Stefanos Doltsinis, Ioanna Mitsioni, and Zoe Doulgeri. Slippage detection generalizing to grasping of unknown objects using machine learning with novel features. *IEEE Robotics and Automation Letters*, 3(2):942–948, 2018.

[24] Janine Hoelscher, Jan Peters, and Tucker Hermans. Evaluation of Tactile Feature Extraction for Interactive Object Recognition. In *IEEE-RAS International Conference on Humanoid Robotics*, 2015.

[25] Lin Chia-Hsien, Jeramy A Fishel, and Gerald E Loeb. Estimating point of contact, force and torque in a

biomimetic tactile sensor with deformable skin. *Tech. Rep.*, 2013.

[26] Zhe Su, Karol Hausman, Yevgen Chebotar, Artem Molchanov, Gerald E Loeb, Gaurav S Sukhatme, and Stefan Schaal. Force estimation and slip detection/classification for grip control using a biomimetic tactile sensor. In *IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pages 297–303, 2015.

[27] Kuan-Ting Yu, John Leonard, and Alberto Rodriguez. Shape and pose recovery from planar pushing. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 1208–1215. IEEE, 2015.

[28] Kuan-Ting Yu and Alberto Rodriguez. Realtime state estimation with tactile and visual sensing. application to planar manipulation. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 7778–7785. IEEE, 2018.

[29] Frank Dellaert and Michael Kaess. Square root SAM: Simultaneous localization and mapping via square root information smoothing. *The International Journal of Robotics Research*, 25(12):1181–1203, 2006.

[30] Marc Toussaint and Amos Storkey. Probabilistic inference for solving discrete and continuous state markov decision processes. In *Proceedings of the 23rd international conference on Machine learning*, pages 945–952, 2006.

[31] Konrad Rawlik, Marc Toussaint, and Sethu Vijayakumar. Approximate inference and stochastic optimal control. *arXiv preprint arXiv:1009.3958*, 2010.

[32] Hilbert J Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of statistical mechanics: theory and experiment*, 2005(11):P11011, 2005.

[33] Evangelos A Theodorou and Emanuel Todorov. Relative entropy and free energy dualities: Connections to path integral and kl control. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 1466–1473. IEEE, 2012.

[34] Alexander Lambert, Adam Fishman, Dieter Fox, Byron Boots, and Fabio Ramos. Stein variational model predictive control. In *Conference on Robot Learning*, 2020.

[35] Qiang Liu and Dilin Wang. Stein variational gradient descent: A general purpose bayesian inference algorithm. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 2378–2386. Curran Associates, Inc., 2016.

[36] Lucas Barcelos, Alexander Lambert, Rafael Oliveira, Paulo Borges, Byron Boots, and Fabio Ramos. Dual online stein variational inference for control and dynamics. In *Robotics: Science and Systems*, 2021.

[37] Kendall Lowrey, Aravind Rajeswaran, Sham Kakade, Emanuel Todorov, and Igor Mordatch. Plan online, learn offline: Efficient learning and exploration via model-based control. In *International Conference on Learning Representations*, 2018.

[38] David Hoeller, Farbod Farshidian, and Marco Hutter. Deep value model predictive control. In *Conference on Robot Learning*, pages 990–1004. PMLR, 2020.

[39] Michael Betancourt. A conceptual introduction to hamiltonian monte carlo. *arXiv preprint arXiv:1701.02434*, 2017.

[40] Chang Liu, Jingwei Zhuo, Pengyu Cheng, Ruiyi Zhang, and Jun Zhu. Understanding and accelerating particle-based variational inference. In *International Conference on Machine Learning*, pages 4082–4092. PMLR, 2019.

[41] Dilin Wang, Zhe Zeng, and Qiang Liu. Stein variational message passing for continuous graphical models. In *International Conference on Machine Learning*, pages 5219–5227. PMLR, 2018.

[42] Jingwei Zhuo, Chang Liu, Jiaxin Shi, Jun Zhu, Ning Chen, and Bo Zhang. Message passing stein variational gradient descent. In *International Conference on Machine Learning*, pages 6018–6027. PMLR, 2018.

[43] Kurtland Chua, Roberto Calandra, Rowan McAllister, and Sergey Levine. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. In *NeurIPS*, 2018.

[44] Masashi Okada and Tadahiro Taniguchi. Variational inference mpc for bayesian model-based reinforcement learning. In *Conference on Robotics Learning (CoRL)*, 2019.

[45] Elena Arcari, Lukas Hewing, Max Schlichting, and Melanie Zeilinger. Dual stochastic mpc for systems with parametric and structural uncertainty. In *Learning for Dynamics and Control*, pages 894–903. PMLR, 2020.

[46] Kim Peter Wabersich and Melanie Zeilinger. Bayesian model predictive control: Efficient model exploration and regret bounds using posterior sampling. In *Learning for Dynamics and Control*, pages 455–464. PMLR, 2020.

[47] Rae Jeong, Jost Tobias Springenberg, Jackie Kay, Daniel Zheng, Yuxiang Zhou, Alexandre Galashov, Nicolas Heess, and Francesco Nori. Learning dexterous manipulation from suboptimal experts. In *Conference on Robot Learning*, 2020.

[48] Nicholas Rhinehart, Rowan McAllister, and Sergey Levine. Deep imitative models for flexible inference, planning, and control. In *International Conference on Learning Representations*, 2019.

[49] Benjamin Eysenbach, Abhishek Gupta, Julian Ibarz, and Sergey Levine. Diversity is all you need: Learning skills without a reward function. In *International Conference on Learning Representations*, 2018.

[50] Avi Singh, Huihan Liu, Gaoyue Zhou, Albert Yu, Nicholas Rhinehart, and Sergey Levine. Parrot: Data-driven behavioral priors for reinforcement learning. *arXiv preprint arXiv:2011.10024*, 2020.

[51] Tianyu Li, Nathan Lambert, Roberto Calandra, Franziska Meier, and Akshara Rai. Learning generalizable locomotion skills with hierarchical reinforcement learning. In *2020 IEEE International Conference on Robotics and*

Automation (ICRA), pages 413–419. IEEE, 2020.

[52] Karl Pertsch, Youngwoon Lee, and Joseph J Lim. Accelerating reinforcement learning with learned skill priors. In *Conference on Robot Learning*, 2020.

[53] Miles Macklin, Kenny Erleben, Matthias Müller, Nuttapong Chentanez, Stefan Jeschke, and Viktor Makoviychuk. Non-smooth newton methods for deformable multi-body dynamics. *ACM Transactions on Graphics (TOG)*, 38(5):1–20, 2019.

[54] Rika Antonova, Akshara Rai, Tianyu Li, and Danica Kragic. Bayesian optimization in variational latent spaces with dynamic compression. In *Conference on Robot Learning*, pages 456–465. PMLR, 2020.

[55] Rika Antonova, Akshara Rai, and Christopher G Atkeson. Deep kernels for optimizing locomotion controllers. In *Conference on Robot Learning*, pages 47–56. PMLR, 2017.

[56] Fabio Ramos, Rafael Possas, and Dieter Fox. BayesSim: Adaptive domain randomization via probabilistic inference for robotics simulators. In *Proceedings of Robotics: Science and Systems*. doi: 10.15607/RSS.2019.XV.029.

[57] Homanga Bharadhwaj, Zihan Wang, Yoshua Bengio, and Liam Paull. A data-efficient framework for training and sim-to-real transfer of navigation policies. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 782–788. IEEE, 2019.

[58] Ian Abraham, Ankur Handa, Nathan Ratliff, Kendall Lowrey, Todd D Murphey, and Dieter Fox. Model-based generalization under parameter uncertainty using path integral control. *IEEE Robotics and Automation Letters*, 5(2):2864–2871, 2020.

[59] Jacky Liang, Ankur Handa, Karl Van Wyk, Viktor Makoviychuk, Oliver Kroemer, and Dieter Fox. In-hand object pose tracking via contact feedback and gpu-accelerated robotic simulation. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6203–6209. IEEE, 2020.

[60] Sylvain Calinon. Gaussians on riemannian manifolds: Applications for robot learning and adaptive control. *IEEE Robotics & Automation Magazine*, 27(2):33–45, 2020.

[61] Chang Liu and Jun Zhu. Riemannian stein variational gradient descent for bayesian inference. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.