Learning deep observation models for factor graph inference

Paloma Sodhi
The Robotics Institute, Carnegie Mellon University

I. Introduction

We address the problem of learning observation models end-to-end for estimation. Robots operating in partially observable environments must infer latent states from multiple sensory inputs, e.g., image and touch measurements. Consider a robot hand manipulating an object with tactile feedback. It must reason over a *sequence* of touch measurements over time to collapse uncertainty about the latent object pose. A common way to solve this is as an inference over a factor graph which relies on having observation models that can map observations to states [6, 8, 9]. However, in many domains, the sensors that produce observations are complex and difficult to model. *Can we instead learn observation models from data?*

Given a batch of ground truth trajectories and corresponding measurements, how should we learn observations models? One option is directly learn a mapping from measurements to states, for example as a regression or classification problem [20, 25, 27]. However, we note that this only minimizes a surrogate loss independent of the graph optimizer – what we actually care about is the final task, i.e., how well we track the ground truth trajectory.

Instead, we aim to directly optimize end-to-end tracking performance. At a first glance, this appears to require the factor graph inference algorithm to be fully differentiable. However, many state-of-the-art factor graph optimizers e.g. iSAM2 [16] are not natively differentiable due to operations such as iterative re-linearizations. Hence, we are limited to black-box search for learning the parameters of these optimizers which can be very sample inefficient.

Instead of differentiating through the optimization process, we note that what we ultimately care about is the final solution from the optimizer, which depends *only on the shape* of the optimized cost function. We would like a cost function that has low cost around the observed ground truth trajectories and high cost elsewhere. This is precisely what energy-based models aim to do by shaping an "energy" function to be low around observed data and high elsewhere [2, 3, 11, 19, 23, 29, 31].

Our key technical insight is to cast our problem of learning observation models as energy-based learning. Our proposal is based on the central principle:

> *Learn observation models that directly minimize end-to-end tracking errors. We can achieve this by coupling learning and optimization so that the learner extracts the most salient information to aid optimization.*

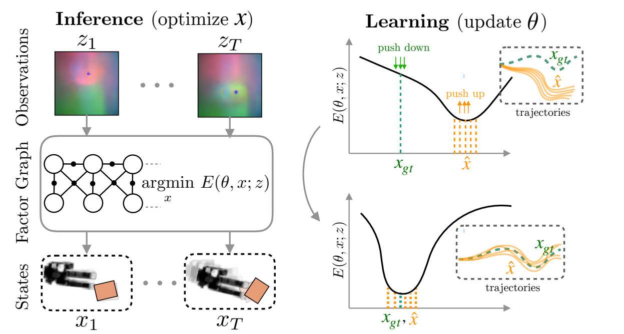

Fig. 1: We learn observation models in non-differentiable graph optimizers to minimize end-to-end tracking errors. **Inference** solves for the most likely states x given the model and input measurements z. **Learning** updates model parameters θ from training data.

II. Related Work

Observation models represent joint or conditional distributions between states and measurements. Early estimation methods utilized these models within filtering contexts such as Kalman filters and EKFs/UKFs [15, 17, 28]. Recent methods have also looked at making these filters differentiable for end-to-end learning [12, 14, 18]. However, filtering has difficulty scaling and can have inconsistencies due to linearization choices on past, marginalized states that cannot be undone. Localization and SLAM problems are now instead increasingly solved as smoothing or nonlinear optimization objectives over a graph that can leverage the inherent sparsity of the problem to give tractable and more accurate solutions [6, 8, 9].

Typically, observation models used in the factor graph are analytic models defined a-priori [10, 21, 24]. Recent work has looked at using learned models [25, 27] or learned measurement representations [5, 7] either to be used independently or to be plugged into a graph optimizer. However, these model parameters are learned on surrogate losses independent of the graph optimizer, and hence do not directly attempt to minimize final tracking errors. Another recent line of work has been on building fully differentiable optimizers [4, 13, 30] via unrolling where the optimization is represented as a series of gradient based updates. While such approaches allow end-to-end training, these are currently limited to optimizers that can be expressed as a composition of differentiable operations. Moreover, vanilla unrolling suffers from a few drawbacks notably that the learned cost function can be sensitive to the specific optimization procedure [1].

III. Problem Formulation

Inference. At inference time, given a sequence of measurements z, we wish to solve for the most likely sequence

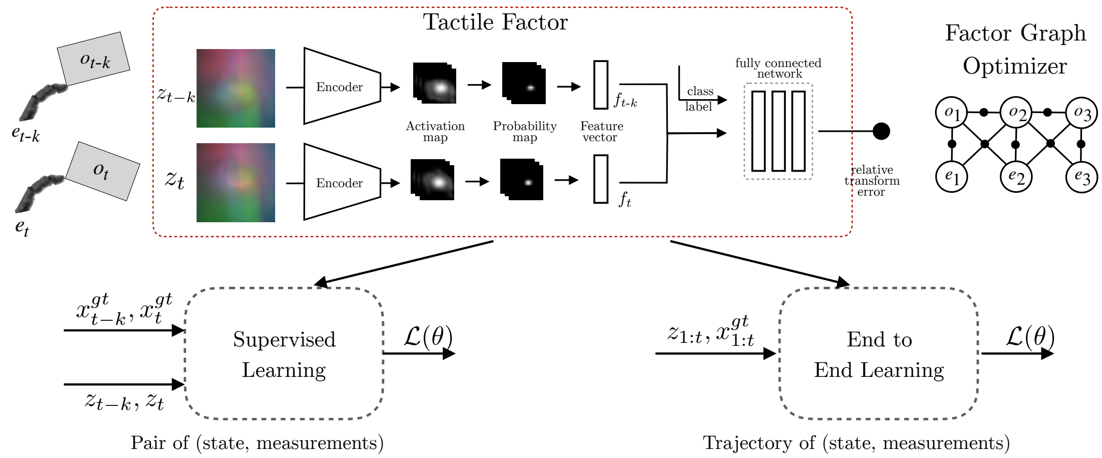

Fig. 2: Overall framework showing two approaches for learning tactile observation models **(a)** supervised learning on a surrogate loss **(b)** end-to-end learning on final tracking loss.

of latent states x. This can be formulated as maximum-a-posteriori objective over a factor graph, which under Gaussian noise models, is equivalent to solving a nonlinear least-squares problem of the form,

$$E(\theta, x; z) = \underset{x}{\operatorname{argmin}} \sum_t ||f_t(\theta, x; z)||^2 \tag{1}$$

where, $E(\cdot)$ is a scalar energy value. $||f_t(\theta, x; z)||^2$ are factor costs at time t with $f_t(\theta, x; z)$ being the local observation model mapping a *subset* of states x and measurements z to a cost. θ the the learnable model parameters that includes covariance Σ_t in the weighted norm $|| \cdot ||_{\Sigma_t}$.

Learning. At train time, our goal is to learn a model that explains training data of pairs (x_{gt}, z). We express this as minimizing a loss $\mathcal{L}(\theta)$ over the training data

$$\mathcal{L}(\theta) = \frac{1}{|\mathcal{D}|} \sum_{(x^i_{gt}, z^i) \in \mathcal{D}} \mathcal{L}(\theta; x^i_{gt}, z^i) \tag{2}$$

where, $\{x^i_{gt}, z^i\} \in \mathcal{D}$ is a training dataset of ground truth trajectories and measurements.

IV. Approaches

We propose two set of approaches to solve for the objective in Eq. 2: (a) supervised learning on a surrogate loss [25] and (b) end-to-end learning on final tracking loss [26].

A. Supervised learning on a surrogate loss

We train a tactile factor network that takes in pairs of tactile image measurements and predicts relative object poses (Fig. 2). The loss is a mean-squared-error against ground truth poses. While easy to train, given a direct supervised loss, this only minimizes a surrogate loss independent of the graph optimizer and is hence not guaranteed to minimize the final trajectory tracking errors that we care about. It is also not straightforward to design surrogate losses for learning parameters such as factor covariances that don't have direct ground truth supervision.

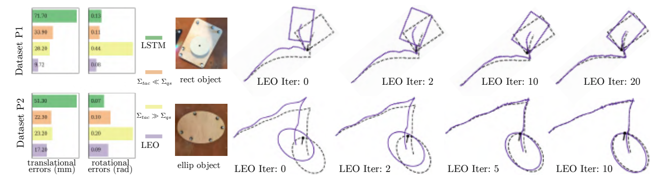

Fig. 3: Final tracking results for real-world planar pushing. We show convergence of optimizer trajectories to ground truth (dotted grey) over successive learner iterations on datasets with two different objects, a rectangle and ellipse.

B. End-to-end learning on final tracking loss

Instead of a pairwise loss, we minimize a final tracking loss against an entire trajectory of ground truth poses (Fig. 2). Directly minimizing such a loss is not possible, since the graph inference path from θ to $\hat{x}$ is not differentiable. We instead consider a loss that is a function of the energy, i.e. $\mathcal{L}(E(\theta, \cdot); x^i_{gt}, z^i)$. Intuitively, this assigns a low loss to *well-behaved* energy functions, i.e. functions that give the lowest energy to training data of ground truth trajectories (correct answers) and higher energy to unseen data (incorrect answers).

This energy-based loss is highly correlated with the final tracking loss as it shapes the energy or cost landscape so as to make the inference step return trajectories closer to the ground truth (Fig. 3). In practice, this requires matching gradients of the observation likelihoods on ground truth trajectories against samples generated from the factor graph optimizer [26].

C. Application to tactile manipulation

We apply both these approaches for tracking object poses during manipulation using tactile image measurements. We focus on the task of planar pushing where we use a DIGIT tactile sensor [22] mounted on an end-effector that is used to push objects of different shapes.

V. Future Work

As future work, we would like investigate,

1) Can we design a unified framework that combines the robustness of supervised learning with guarantees of end-to-end learning?
2) Can we extend these approaches to 3D manipulation tasks where sequence of tactile images may correspond to multiple possible object motions?

Acknowledgements

This work is in collaboration with Stuart Anderson, Mustafa Mukadam at Facebook AI Research and Eric Dexheimer, Michael Kaess at Carnegie Mellon University. The content of this work appears in [25, 26].

References

[1] Brandon Amos and Denis Yarats. The differentiable cross-entropy method. In *International Conference on Machine Learning*, pages 291–302. PMLR, 2020.

[2] Sarah Bechtle, Artem Molchanov, Yevgen Chebotar, Edward Grefenstette, Ludovic Righetti, Gaurav Sukhatme, and Franziska Meier. Meta learning via learned loss. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 4161–4168, 2021.

[3] David Belanger, Bishan Yang, and Andrew McCallum. End-to-end learning for structured prediction energy networks. In *International Conference on Machine Learning (ICML)*, pages 429–439, 2017.

[4] Mohak Bhardwaj, Byron Boots, and Mustafa Mukadam. Differentiable Gaussian process motion planning. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 10598–10604, 2020.

[5] Michael Bloesch, Jan Czarnowski, Ronald Clark, Stefan Leutenegger, and Andrew J Davison. CodeSLAM–learning a compact, optimisable representation for dense visual slam. In *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pages 2560–2568, 2018.

[6] Cesar Cadena, Luca Carlone, Henry Carrillo, Yasir Latif, Davide Scaramuzza, José Neira, Ian Reid, and John J Leonard. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Trans. Robotics*, 32(6): 1309–1332, 2016.

[7] Jan Czarnowski, Tristan Laidlow, Ronald Clark, and Andrew J Davison. DeepFactors: Real-time probabilistic dense monocular SLAM. *IEEE Robotics and Automation Letters*, 5(2):721–728, 2020.

[8] Frank Dellaert. Factor graphs: Exploiting structure in robotics. *Annual Review of Control, Robotics, and Autonomous Systems*, 4, 2020.

[9] Frank Dellaert and Michael Kaess. Factor graphs for robot perception. *Foundations and Trends in Robotics*, 6(1-2):1–139, 2017.

[10] Jakob Engel, Thomas Schöps, and Daniel Cremers. LSD-SLAM: Large-scale direct monocular SLAM. In *Eur. Conf. on Computer Vision (ECCV)*, pages 834–849. Springer, 2014.

[11] Fredrik K Gustafsson, Martin Danelljan, Radu Timofte, and Thomas B Schön. How to train your energy-based model for regression-supplementary material. 2020.

[12] Tuomas Haarnoja, Anurag Ajay, Sergey Levine, and Pieter Abbeel. Backprop kf: Learning discriminative deterministic state estimators. In *Advances in Neural Information Processing Systems 24 (NeurIPS)*, 2016.

[13] Krishna Murthy Jatavallabhula, Ganesh Iyer, and Liam Paull. GradSLAM: Dense SLAM meets automatic differentiation. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2020.

[14] Rico Jonschkowski, Divyam Rastogi, and Oliver Brock. Differentiable particle filters: End-to-end learning with algorithmic priors. *arXiv preprint arXiv:1805.11122*, 2018.

[15] Simon J Julier and Jeffrey K Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.

[16] Michael Kaess, Hordur Johannsson, Richard Roberts, Viorela Ila, John Leonard, and Frank Dellaert. iSAM2: Incremental smoothing and mapping using the Bayes tree. *Intl. J. of Robotics Research*, 31(2):216–235, February 2012.

[17] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. 1960.

[18] Alina Kloss, Georg Martius, and Jeannette Bohg. How to train your differentiable filter. *arXiv preprint arXiv:2012.14313*, 2020.

[19] Rithesh Kumar, Sherjil Ozair, Anirudh Goyal, Aaron Courville, and Yoshua Bengio. Maximum entropy generators for energy-based models. *arXiv preprint arXiv:1901.08508*, 2019.

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

[21] 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 *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 3165–3171, 2019.

[22] Mike Lambeta, Po-Wei Chou, Stephen Tian, Brian Yang, Benjamin Maloon, Victoria Rose Most, Dave Stroud, Raymond Santos, Ahmad Byagowi, Gregg Kammerer, et al. DIGIT: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation. *IEEE Robotics and Automation Letters*, 5(3):3838–3845, 2020.

[23] Yann LeCun, Sumit Chopra, Raia Hadsell, M Ranzato, and F Huang. A tutorial on energy-based learning. *Predicting structured data*, 1(0), 2006.

[24] Raul Mur-Artal and Juan D Tardós. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Trans. Robotics*, 33(5):1255–1262, 2017.

[25] Paloma Sodhi, Michael Kaess, Mustafa Mukadam, and Stuart Anderson. Learning tactile models for factor graph-based estimation. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2021.

[26] Paloma Sodhi, Eric Dexheimer, Mustafa Mukadam, Stuart Anderson, and Michael Kaess. LEO: Learning energy-based models in graph optimization. In *Conference on Robot Learning (CoRL)*, 2021 (submitted).

[27] 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 *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2019.

[28] Sebastian Thrun, Dieter Fox, Wolfram Burgard, and Frank Dellaert. Robust Monte Carlo localization for mobile robots. *Artificial intelligence*, 128(1-2):99–141, 2001.

[29] Markus Wulfmeier, Peter Ondruska, and Ingmar Posner. Maximum entropy deep inverse reinforcement learning. *arXiv preprint arXiv:1507.04888*, 2015.

[30] Brent Yi, Michelle Lee, Alina Kloss, Roberto Martín-Martín, and Jeannette Bohg. Differentiable factor graph optimization for learning smoothers. *arXiv preprint arXiv:2105.08257*, 2021.

[31] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *AAAI Conf. on Artificial Intelligence*, volume 8, pages 1433–1438, 2008.