On-line MPC-based Stochastic Planning in the Non-Gaussian Belief Space with Non-convex Constraints*

Mohammadhussein Rafieisakhaei, Amirhossein Tamjidi, Suman Chakravorty
mrafieis@tamu.edu

Abstract

A convex program to design a feedback policy in non-Gaussian belief space is proposed for a class of systems with differentiable non-linear observation model. A Particle Filter is adopted for belief representation. The size of the optimization vector grows linearly with time horizon and dimension of the state. A new approach is proposed to deal with obstacles as a part of optimization problem. Included results show the efficacy of our method.

Introduction

Partially Observed Markov Decision Processes (POMDP) is a general framework for solving planning under uncertainty problems. POMDP is known to be PSPACE-complete in the general setting with continuous underlying spaces. Most of the POMDP solvers approach the problem in the discretized state, action and observation spaces. It becomes even harder in the existence of non-convex constraints over the state space. There have been successful efforts to solve the problem in continuous spaces with Gaussian assumptions (Agha-Mohammadi, Chakravorty, and Amato 2013), however, the results in the literature in the non-Gaussian belief space are sparse. Moreover, the existing algorithms are usually expensive in terms of time and memory. In this paper, we provide a feedback policy in non-Gaussian belief space that involves a convex program even for common non-linear observation models and is therefore very fast and globally optimal in terms of the considered cost. The solution involves a Model Predictive Control (MPC) solved in a Receding Horizon fashion using particle filters for the estimation of the non-Gaussian belief representation. Our solution avoids the maximum likelihood observation heuristics and avoids the filtering in the planning in a smart way using the observation covariance of the Maximum A-Posteriori (MAP) estimate of the state and the Jacobian matrix of the observation model. Then, we adjust the MPC to deal with obstacles and show the results.

*This material is based upon work partially supported by AFOSR Contract FA9550-13-1-0008, NSF under Science Technology Center Grant CCF-0939370, and NPRP grant 6-784-2-329 from the Qatar National Research Fund, a member of Qatar Foundation.

Obstacle-free Convex Environment

Here, we propose our solution for consisting of a linear process model and continuous observation model. This solution can be extended to nonlinear holonomic case.

State space definition: Throughout this paper we denote the random variables with upper case. Vectors $\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^{n_x}$, $\mathbf{u} \in \mathbb{U} \subset \mathbb{R}^{n_u}$, and $\mathbf{z} \in \mathbb{Z} \subset \mathbb{R}^{n_z}$ denote the state of the system, the action, and observation vector, respectively. We assume partial observability which means that we have incomplete and/or noisy information about the current state.

System equations: The dynamics of the system and the observations are as follows:

$$\mathbf{x}_{t+1}= \mathbf{A}_t\mathbf{x}_t + \mathbf{B}_t\mathbf{u}_t + \mathbf{G}_t\boldsymbol{\omega}_t, \quad \mathbf{z}_t= h(\mathbf{x}_t, \boldsymbol{\nu}_t).$$

where $\{\boldsymbol{\omega}_t\}$ and $\{\boldsymbol{\nu}_t\}$ are two zero mean independent, identically distributed (iid) random sequences, and are mutually independent. We also assume that $\mathbb{E}[\mathbf{z}_t\mathbf{z}_t^T|\mathbf{x}_t] = \mathbf{R}(\mathbf{x}_t) \succeq 0$ to be the state-dependent covariance of the measurement noise. In addition, at each time step, $\mathbf{A}_t, \mathbf{B}_t$ and $\mathbf{G}_t$ are appropriate constant matrices, and $h : \mathbb{X} \times \mathbb{R}^m \to \mathbb{Z}$ denotes the observation dynamics.

RHC policy: Suppose that at time step t' knowledge about the state is summarized in belief state $b_{t'}$. Let us call our RHC policy $\tilde{\pi}$. The belief state $b_{t'}$ is approximated by a set of particles $\{\mathbf{x}_{t'}^i\}_{i=1}^N$ whose approximation gets better as the number, N, of particles increases. Given this set of particles and a goal state $\mathbf{x}_g$, we perform an optimization as follows to obtain the appropriate action to be performed at time t':

$$\begin{aligned}\tilde{\pi}(b_{t'}):= \mathbf{e}_1^T\cdot \underset{\mathbf{u}_{t':t'+K-1}}{\arg\min} \sum_{t=t'+1}^{t'+K}[\mathbf{c}_t^T\overline{\mathbf{R}}_t(\mathbf{x}_t^{ML})\mathbf{c}_t+\mathbf{u}_{t-1}^T\mathbf{V}_t^u\mathbf{u}_{t-1}]\\ s.t.\ \mathbf{x}_{t+1}^{ML} = \mathbf{A}_t\mathbf{x}_t^{ML} + \mathbf{B}_t\mathbf{u}_t,\ t \in [t', t'+K]\\ \mathbf{x}_g = \mathbf{x}_{t'+K}^{ML} \qquad (1)\end{aligned}$$

where $\mathbf{x}_t^{ML} = \arg\max_{\mathbf{x}\in\mathbb{R}^{n_x}} b_t(\mathbf{x})$ is the MAP estimate of the state at time t which is obtained through the particle representation as the Most Likely (ML) state, K is the time horizon, $\mathbf{V}_t^u \succeq 0$ is an appropriate weight matrix, $\mathbf{e}_1$ is the K-dimensional unit vector where only the first element is non-zero. Note that we assume h to be differentiable and $\mathbf{H}_t = \partial h(\mathbf{x}, 0)/\partial \mathbf{x}|_{\mathbf{x}_t^{ML}}$, whence, $\mathbf{H}_t$ can be a function of the optimization variables if the observation model is nonlinear. Moreover, $\mathbf{c}_t := (\mathbf{c}_t^1, \mathbf{c}_t^2, \cdots, \mathbf{c}_t^N)^T \in \mathbb{R}^{n_z\times N}$ consists

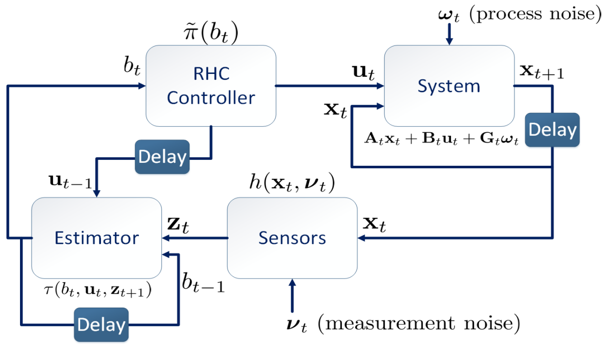

Figure 1: The RHC control loop.

of $\mathbf{c}_t^i := \mathbf{H}_t(\prod_{\tau=t'}^{t'+t} \mathbf{A}_\tau)(\mathbf{x}_{t'}^i - \mathbf{x}_{t'}^{ML}), \mathbf{c}_t^i \in \mathbb{R}^{n_z},\ \ 1 \leq i \leq N$, where $\overline{\mathbf{R}}_t(\mathbf{x}_t^{ML}) = BlockDiag(\frac{1}{N}\mathbf{R}(\mathbf{x}_t^{ML}))$ is a matrix with N equal block diagonal elements and is a diagonal matrix itself. Thus, the action $\mathbf{u}_{t'} = \tilde{\pi}(b_{t'})$ is produced by the function above. The overall control loop is shown in Fig. 1.

Obstacles as the Non-convex Constraints

Barrier function: We introduce a new method of incorporating the obstacles by adding a barrier function of the following form to the minimization objective: $\max_{1\leq i\leq n_b}\{Me^{\sum_{j=1}^{n_x} -\alpha_{ij}(\mathbf{x}_t^{ML}(j) - \mathbf{c}'_i(j))^2}\}$, where n_b is the number of the obstacles, $M > 0$, $\alpha_{ij} > 0$ are constant numbers, and $\mathbf{c}'_i \in \mathbb{R}^{n_x}$ is a constant vector. Adding this cost, ensures that the ML particle remains outside the infeasible regions along the whole time horizon. Thus, we inflate the size of the barriers to increase safety margin. We design the barrier function such that there is only one function per obstacle acting as a soft constraint.

Analysis and Results

In the main optimization problem (1), number of the decision variables is Kn_u, and the planning problem can be shown to be convex for some commonly used observation models in robotics. We have used a landmark-based observation model with observation covariance proportional to the range. Thus, the solution is a global minimizer in terms of the control effort in a trade-off with the uncertainty reduction along the trajectory. The weight matrix $\mathbf{V}_t^u$ determines the trade-off for the exploration vs. exploitation.

In the second case where we consider the obstacles by adding the barrier function to the optimization objective, the problem becomes non-convex, but still numerically well-behaved. The key to handle this problem is to increase the horizon increase to allow for more exploration that is needed to bypass the obstacles. By adding obstacle function the memory and computational requirements to solve the problem do not increase drastically as in the chance constrained problems (Zhang et al. 2013), since the size of the optimization vector remains the same and only the objective function changes. Moreover, there is no need to introduce any additional integer valued variables, and the optimization problem is only solved once (cf. (Platt 2013)). The simulation results are shown in the Fig. 2.

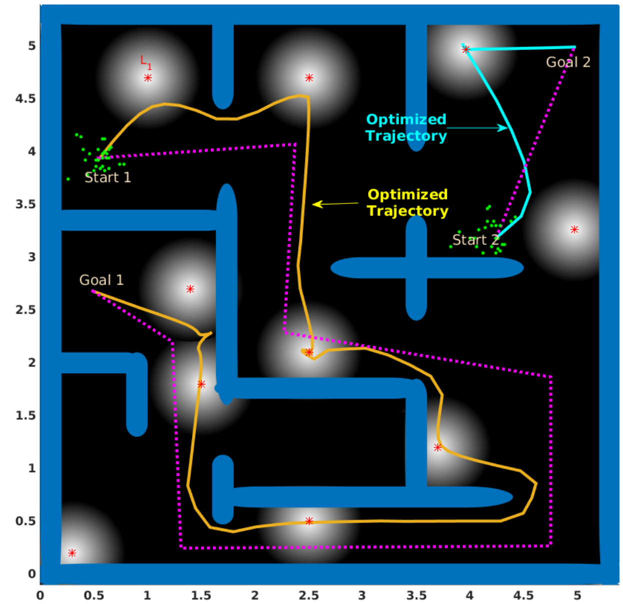

Figure 2: Solid lines show the optimal trajectories, dotted show the initial, for two different scenarios. The longer one includes obstacles, and the other, no obstacles. The dots around the start points show the initial particles. Landmarks are marked as stars and information is coded with color (lighter means more information, adapted from (Platt 2013)) The horizon for the longer one is 100 and 30 for the other.

Conclusion

In this paper, we proposed a new MPC strategy to solve the on-line planning problem in the cases where the state is only partially observed and the process model is stochastic. We proposed a new way of handling the problem that arises from the lack of observations at the planning stage avoiding the maximum likelihood heuristic. Our solution considers a non-Gaussian belief space and several landmarks in the field of view of the agent. We also proposed a different method of handling the obstacles and have shown the results. Note that the number of optimization variables is equal to the number of control inputs. In addition, in the presence of the obstacle, there is no increase in the number of variables. Therefore, the solution can be obtained fast enough for on-line applications. Our future work extends and proves our results.

References

Agha-Mohammadi, A.; Chakravorty, S.; and Amato, N. 2013. Firm: Sampling-based feedback motion planning under motion uncertainty and imperfect measurements. *The International Journal of Robotics Research* .

Platt, R. 2013. Convex receding horizon control in non-gaussian belief space. In *Algorithmic Foundations of Robotics X*. Springer. 443–458.

Zhang, X.; Margellos, K.; Goulart, P.; and Lygeros, J. 2013. Stochastic Model Predictive Control Using a Combination of Randomized and Robust Optimization. In *IEEE Conference on Decision and Control*, 7740 – 7745.