

# Near-Optimal Belief Space Planning via T-LQG\*

Mohammadhussein Rafieisakhaei<sup>1</sup>, Suman Chakravorty<sup>2</sup> and P. R. Kumar<sup>1</sup>

**Abstract**—We consider the problem of planning under observation and motion uncertainty for nonlinear robotics systems. Determining the optimal solution to this problem, generally formulated as a Partially Observed Markov Decision Process (POMDP), is computationally intractable. We propose a Trajectory-optimized Linear Quadratic Gaussian (T-LQG) approach that leads to quantifiably near-optimal solutions for the POMDP problem. We provide a novel “separation principle” for the design of an optimal nominal open-loop trajectory followed by an optimal feedback control law, which provides a near-optimal feedback control policy for belief space planning problems involving a polynomial order of calculations of minimum order.

## I. INTRODUCTION

Planning for systems with observation and motion uncertainty is generally formulated in the framework of a Partially Observed Markov Decision Process (POMDP), the general solution of which is provided by the Hamilton-Jacobi-Bellman equations [1]. Attempts to utilize this framework run into the intractability of the computations, referred to as the curse of dimensionality.

In this paper, we provide a structure under which the stochastic optimal control problem can be solved quantifiably near-optimal for moderate levels of noise. We utilize the Wentzell-Freidlin theory of large deviations for analyzing the asymptotics under small noise [2]. In particular, we consider a general nonlinear process and measurement models with additive white noise, and compensate the system with feedback. We show that the first-order stochastic error of the stochastic cost function for the feedback-compensated system is distributed according to a Gaussian distribution with zero expected value.

As a result of the independence of the first-order expected error from the feedback law, the optimal zeroth-order (nominal open-loop) control sequence can be designed separately from the optimal closed-loop feedback law; a result which we term as a “separation of the open-loop and closed-loop designs”. This leads to a novel design approach for partially-observed nonlinear stochastic systems whose characteristics we quantify. We also provide a tractable example of a robotic motion and path planning design based on this theory. Other than the HJB equations, this is the only structure to-date that provides quantifiably near-optimal solutions for a relatively

general stochastic optimal control problem. In addition, unlike the HJB, this approach does not run into the problem of curse of dimensionality, as the entire computation is of the order of  $O(Kn^3)$ , where  $K$  is the planning horizon, and  $n$  is the state dimension. Lastly, it is observed in simulations that the design is valid for a moderate-range of noise level due to the power of feedback compensation.

## II. GENERAL PROBLEM

The general belief space planning problem is formulated as a stochastic control problem in the space of feedback policies. In this section, we define the basic elements of the problem, including system equations and belief dynamics.

*SDE models:* We consider continuous-time Stochastic Differential Equation (SDE) models of the process and measurement as follows:

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t)dt + \epsilon\boldsymbol{\sigma}(t)d\boldsymbol{\omega}_t, \quad (1a)$$

$$d\mathbf{z}_t = \mathbf{h}(\mathbf{x}_t)dt + \epsilon d\boldsymbol{\nu}_t, \quad (1b)$$

where  $\{\boldsymbol{\omega}_t, \boldsymbol{\nu}_t, t \geq 0\}$  are two independent standard Wiener processes,  $\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^{n_x}$ ,  $\mathbf{u} \in \mathbb{U} \subset \mathbb{R}^{n_u}$ , and  $\mathbf{z} \in \mathbb{Z} \subset \mathbb{R}^{n_z}$ , denote the state, control and observation vectors, respectively, and  $\mathbf{f} : \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{X}$ ,  $\mathbf{h} : \mathbb{X} \rightarrow \mathbb{Z}$ ,  $\boldsymbol{\sigma}, \mathbf{a} : \mathbb{R} \rightarrow \mathbb{R}^{n_x \times n_x}$ ,  $\mathbf{f} = (f_i)_{0 \leq i \leq n_x}$ ,  $\mathbf{h} = (h_i)_{0 \leq i \leq n_z}$ , and  $\mathbf{a} := \boldsymbol{\sigma}\boldsymbol{\sigma}^T = (a_{i,j})_{0 \leq i,j \leq n_x}$ . We assume that the drift and diffusion coefficients,  $f_i, h_i, a_{i,j}$ , are bounded and uniformly Lipschitz continuous functions, and the diffusion matrix is uniformly positive-definite. Lastly,  $\mathbf{x}_0 \sim \mathcal{N}(\bar{\mathbf{x}}_0, \epsilon^2 \boldsymbol{\Sigma}_{\mathbf{x}_0})$ ,  $\epsilon > 0$ .

*Belief:* The conditional distribution of the state given the past observations, controls and the initial distribution is termed as “belief”. In the sequel, we denote the Gaussian belief by  $\mathbf{b}_t = (\hat{\mathbf{x}}_t^T, \text{vec}(\mathbf{P}_t)^T)^T \in \mathbb{B}$ , a vector of the mean and covariance of the estimation at time  $t$ .

*Problem 1: Stochastic Control Problem:* Given an initial belief state  $\mathbf{b}_0$ , the stochastic optimal control problem is:

$$\min_{\boldsymbol{\pi}} \mathbb{E} \left[ \sum_{t=0}^{K-1} c_t^{\boldsymbol{\pi}}(\mathbf{b}_t, \mathbf{u}_t) + c_K^{\boldsymbol{\pi}}(\mathbf{b}_K) \right] \\ \text{s.t. } \mathbf{b}_{t+1} = \boldsymbol{\tau}(\mathbf{b}_t, \mathbf{u}_t, \mathbf{z}_{t+1}), \quad (2)$$

where the optimization is over Markov policies,  $\boldsymbol{\pi}$ , and:

- $J^{\boldsymbol{\pi}} : \mathbb{B} \rightarrow \mathbb{R}$  is the cost function given the policy  $\boldsymbol{\pi} \in \boldsymbol{\Pi}$ , and  $J^{\boldsymbol{\pi}} := \sum_{t=0}^{K-1} c_t^{\boldsymbol{\pi}}(\mathbf{b}_t, \mathbf{u}_t) + c_K^{\boldsymbol{\pi}}(\mathbf{b}_K)$ ;
- $\boldsymbol{\pi} := \{\boldsymbol{\pi}_0, \dots, \boldsymbol{\pi}_t\}$ ,  $\boldsymbol{\pi}_t : \mathbb{B} \rightarrow \mathbb{U}$  and  $\mathbf{u}_t = \boldsymbol{\pi}_t(\mathbf{b}_t)$ ;
- $c_t^{\boldsymbol{\pi}}(\cdot, \cdot) : \mathbb{B} \times \mathbb{U} \rightarrow \mathbb{R}$  is the one-step cost function;
- $c_K^{\boldsymbol{\pi}}(\cdot) : \mathbb{B} \rightarrow \mathbb{R}$  denotes the terminal cost; and
- $K > 0$  is planning horizon, and  $\boldsymbol{\tau}$  defines belief evolution.

## III. METHOD AND MAIN RESULTS

*Feedback law:* We assume a Lipschitz continuous, bounded and smooth feedback law:

$$\mathbf{u}_t = \boldsymbol{\pi}_t(\hat{\mathbf{x}}_t). \quad (3)$$

\*This material is based upon work partially supported by NSF under Contract Nos. CNS-1646449 and Science & Technology Center Grant CCF-0939370, the U.S. Army Research Office under Contract No. W911NF-15-1-0279, and NPRP grant NPRP 8-1531-2-651 from the Qatar National Research Fund, a member of Qatar Foundation.

<sup>1</sup>M. Rafieisakhaei and P. R. Kumar are with the Department of Electrical and Computer Engineering, and <sup>2</sup>S. Chakravorty is with the Department of Aerospace Engineering, Texas A&M University, College Station, Texas, 77840 USA. {mrafieis, schakrav, prk@tamu.edu}

*Nominal ODEs:* Nominal (unperturbed) trajectories of the system can be obtained using a nominal control sequence (which is calculated using the separation result of this paper). The following Ordinary Differential Equations (ODEs) describe the nominal trajectories:

$$\dot{\mathbf{x}}_t^p = \mathbf{f}(\mathbf{x}_t^p, \mathbf{u}_t^p), \quad \dot{\mathbf{z}}_t^p = \mathbf{h}(\mathbf{x}_t^p), \quad \mathbf{u}_t^p = \boldsymbol{\pi}_t(\hat{\mathbf{x}}_t^p), \quad (4)$$

where  $\hat{\mathbf{x}}_0^p := \mathbf{x}_0^p := \mathbb{E}[\mathbf{b}_0]$ , and  $\mathbf{x}_t^p$  is the mean of nominal belief.

*Linearized equations:* We linearize the SDEs of (1) around nominal trajectories. Thus, if  $\|\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_t^p\| \leq \delta$  and  $\|\mathbf{x}_t - \mathbf{x}_t^p\| \leq \delta$ ,

$$\mathbf{u}_t = \mathbf{u}_t^p - \mathbf{L}_t(\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_t^p) + o(\delta), \quad (5a)$$

$$\begin{aligned} \dot{\mathbf{x}}_t &= \dot{\mathbf{x}}_t^p + \mathbf{A}_t(\mathbf{x}_t - \mathbf{x}_t^p) + \mathbf{B}_t(\mathbf{u}_t - \mathbf{u}_t^p) + \epsilon \mathbf{G}_t \frac{d\boldsymbol{\omega}_t}{dt} + o(\delta) \quad (5b) \\ &= \dot{\mathbf{x}}_t^p + \mathbf{A}_t(\mathbf{x}_t - \mathbf{x}_t^p) - \mathbf{B}_t \mathbf{L}_t(\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_t^p) + \epsilon \mathbf{G}_t \frac{d\boldsymbol{\omega}_t}{dt} + o(\delta), \end{aligned}$$

$$\dot{\mathbf{z}}_t = \dot{\mathbf{z}}_t^p + \mathbf{H}_t(\mathbf{x}_t - \mathbf{x}_t^p) + \epsilon \frac{d\boldsymbol{\nu}_t}{dt} + o(\delta). \quad (5c)$$

with Jacobians (the superscript  $p$  was dropped for simplicity):

$$\mathbf{A}_t^p := \nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u})|_{\mathbf{x}_t^p, \mathbf{u}_t^p}, \quad \mathbf{B}_t^p := \nabla_{\mathbf{u}} \mathbf{f}(\mathbf{x}, \mathbf{u})|_{\mathbf{x}_t^p, \mathbf{u}_t^p}, \quad \mathbf{G}_t := \boldsymbol{\sigma}(t),$$

$$\mathbf{L}_t^p := -\nabla_{\mathbf{x}} \boldsymbol{\pi}_t(\mathbf{x})|_{\hat{\mathbf{x}}_t^p}, \quad \mathbf{H}_t^p := \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x})|_{\mathbf{x}_t^p}.$$

*Kalman-Bucy Filter (KBF):* The linearized system's estimates can be obtained using the KBF equations:

$$\dot{\hat{\mathbf{x}}}_t = \hat{\mathbf{x}}_t^p + \mathbf{A}_t(\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_t^p) + \mathbf{B}_t(\mathbf{u}_t - \mathbf{u}_t^p) + \mathbf{K}_t(\dot{\mathbf{z}}_t - \dot{\mathbf{z}}_t^p - \mathbf{H}_t(\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_t^p)), \quad (6a)$$

$$\dot{\mathbf{P}}_t = \mathbf{A}_t \mathbf{P}_t + \mathbf{P}_t \mathbf{A}_t^T + \epsilon^2 \mathbf{G}_t \boldsymbol{\Sigma}_{\boldsymbol{\omega}} \mathbf{G}_t^T - \epsilon^2 \mathbf{K}_t \boldsymbol{\Sigma}_{\boldsymbol{\nu}} \mathbf{K}_t^T, \quad (6b)$$

$$\mathbf{K}_t = \epsilon^{-2} \mathbf{P}_t \mathbf{H}_t^T \boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1}. \quad (6c)$$

with  $\mathbf{P}_0 := \epsilon^2 \boldsymbol{\Sigma}_{\mathbf{x}_0}$  and  $\hat{\mathbf{x}}_0^p = \mathbf{x}_0^p$ , which implies  $\hat{\mathbf{x}}_t^p \equiv \mathbf{x}_t^p, t \geq 0$ .

*Stochastic differential equation governing the evolution of the augmented state:* Since the evolution of the covariance is deterministic, we define  $\mathbf{y}_t := (\mathbf{x}_t^T, \hat{\mathbf{x}}_t^T)^T$  (also denoted by  $\mathbf{y}_t^\epsilon$ ), which is the concatenation of the two vectors of state and mean of the belief, and define  $\boldsymbol{\zeta}_t := (\boldsymbol{\omega}_t^T, \boldsymbol{\nu}_t^T)^T$ . Then, the evolution of this augmented state random variable is:

$$d\mathbf{y}_t = \mathbf{g}(t, \mathbf{y}_t) dt + \epsilon \boldsymbol{\sigma}^{\mathbf{y}}(t) d\boldsymbol{\zeta}_t, \quad (7)$$

with  $\mathbf{y}_0 = (\mathbf{x}_0^T, (\mathbf{x}_0^p)^T)^T$ , where functions  $\mathbf{g}: \mathbb{R} \times \mathbb{R}^{2n_x} \rightarrow \mathbb{R}^{2n_x}$  and  $\boldsymbol{\sigma}^{\mathbf{y}}: \mathbb{R} \rightarrow \mathbb{R}^{2n_x \times 2n_x}$  are defined (with some abuse of notation) as:

$$\mathbf{g}(t, \mathbf{y}_t) := \begin{pmatrix} \mathbf{f}(\mathbf{x}_t, \boldsymbol{\pi}_t(\hat{\mathbf{x}}_t)) \\ \mathbf{f}(\mathbf{x}_t^p, \boldsymbol{\pi}_t(\hat{\mathbf{x}}_t^p)) + \mathbf{A}_t(\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_t^p) + \mathbf{B}_t(\boldsymbol{\pi}_t(\hat{\mathbf{x}}_t) - \boldsymbol{\pi}_t(\hat{\mathbf{x}}_t^p)) \\ + \mathbf{K}_t(\mathbf{h}(\mathbf{x}_t) - \mathbf{h}(\mathbf{x}_t^p) - \mathbf{H}_t(\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_t^p)) \end{pmatrix},$$

$$\boldsymbol{\sigma}^{\mathbf{y}}(t) := \begin{pmatrix} \boldsymbol{\sigma}(t) & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_t \end{pmatrix}.$$

*Lemma 1 (Initial State):* Let  $\{\mathbf{y}_t^p, t \geq 0\}$ ,  $\{\mathbf{y}_t^n, t \geq 0\}$ , and

$$\dot{\mathbf{y}}_t^p = \mathbf{g}(t, \mathbf{y}_t^p), \quad \mathbf{y}_0^p = ((\mathbf{x}_0^p)^T, (\mathbf{x}_0^p)^T)^T, \quad (8a)$$

$$\dot{\mathbf{y}}_t^n = \mathbf{g}(t, \mathbf{y}_t^n), \quad \mathbf{y}_0^n = (\mathbf{x}_0^T, \mathbf{x}_0^T)^T. \quad (8b)$$

Also, let  $K^{\mathbf{f}}$  and  $K^{\boldsymbol{\pi}_t}$  be the Lipschitz constants of  $\mathbf{f}$  and  $\boldsymbol{\pi}_t$ ,

$$c^\delta := \frac{\delta}{2} \exp\left(-\int_0^K K^{\mathbf{f}}(1 + K^{\boldsymbol{\pi}_t}) dr\right), \quad P_{\delta, \epsilon} := \int_{\|\mathbf{x}\| \leq c^\delta} \exp(\epsilon^2 \mathbf{x}^T \boldsymbol{\Sigma}_{\mathbf{x}_0} \mathbf{x}) d\mathbf{x},$$

and  $\delta > 0$ . Then,

$$P\{\|\mathbf{x}_K^n - \mathbf{x}_K^p\| \leq \delta/2\} \geq P_{\delta, \epsilon}. \quad (9)$$

*Linearization of the SDE:* Given  $\mathbf{F}_t^{\mathbf{g}} = \nabla_{\mathbf{y}} \mathbf{g}(t, \mathbf{y})|_{t, \mathbf{y}_t^p}$ , we linearize the SDE (7) around ODE (8a):

$$d\mathbf{y}_t = \mathbf{g}(t, \mathbf{y}_t^p) dt + \mathbf{F}_t^{\mathbf{g}}(\mathbf{y}_t - \mathbf{y}_t^p) dt + \epsilon \boldsymbol{\sigma}^{\mathbf{y}}(t) d\boldsymbol{\omega}_t + o(\|\mathbf{y}_t - \mathbf{y}_t^p\| dt).$$

If  $\|\mathbf{y}_t - \mathbf{y}_t^p\| \leq 2\delta$  (whose asymptotics are calculated using the Wentzell-Freidlin theory, next and Lemma 1),

$$d\mathbf{y}_t = \mathbf{g}(t, \mathbf{y}_t^p) dt + \mathbf{F}_t^{\mathbf{g}}(\mathbf{y}_t - \mathbf{y}_t^p) dt + \epsilon \boldsymbol{\sigma}^{\mathbf{y}}(t) d\boldsymbol{\omega}_t + o(\delta dt). \quad (10)$$

*Action functional [2]:* For  $[T_1, T_2] \subseteq [0, K]$ , the normalized action functional for the family of  $\epsilon$ -dependent stochastic processes of (7) is defined as:

$$S_{T_1, T_2}(\phi) := \frac{1}{2\epsilon^2} \int_{T_1}^{T_2} L(s, \phi_s, \dot{\phi}_s) ds, \quad (11)$$

for absolutely continuous  $\phi$ , and is set to  $+\infty$  for other  $\phi \in \mathbb{C}_{0K}(\mathbb{R}^{2n_x})$  (the space of continuous functions over  $[0, K]$ ), where  $L: \mathbb{R} \times \mathbb{R}^{2n_x} \times \mathbb{R}^{2n_x} \rightarrow \mathbb{R}$  is the Legendre transform of the cumulant of stochastic process of (7) (assuming  $\mathbf{K}_t \mathbf{K}_t^T \succ 0$ ):

$$L(t, \mathbf{x}, \boldsymbol{\beta}) = \frac{1}{2} (\boldsymbol{\beta} - \mathbf{b}(t, \mathbf{x}))^T \mathbf{a}(t, \mathbf{x})^{-1} (\boldsymbol{\beta} - \mathbf{b}(t, \mathbf{x})). \quad (12)$$

*Theorem 1 (Exponential Rate of Convergence):* Let:

- $\mathbb{D}$  be a domain in  $\mathbb{R}^{2n_x}$ , and denote its closure by  $\text{cl}(\mathbb{D})$ ;
- $\partial\mathbb{D}$  denote the boundary of  $\mathbb{D}$ ;
- $\mathbb{H}_{\mathbb{D}}(t, \mathbf{y}_0^n) = \{\phi \in \mathbb{C}_{0K}(\mathbb{R}^{2n_x}) : \phi_0 = \mathbf{y}_0^n, \phi_t \in \mathbb{D} \cup \partial\mathbb{D}\}$ .

Assume  $\partial\mathbb{D} = \partial\text{cl}(\mathbb{D})$ . Then, we have the following:

$$\lim_{\epsilon \rightarrow 0} \epsilon^2 \ln P_{\mathbf{y}_0^n} \{\mathbf{y}_t^\epsilon \in \mathbb{D}\} = - \inf_{\phi \in \mathbb{H}_{\mathbb{D}}(t, \mathbf{y}_0^n)} S_{0t}(\phi), \quad (13)$$

*Theorem 2 (Asymptotics of the Diffusion Process):* Let:

- $\mathbb{D}_t = \text{cl}(\mathbb{B}_{\delta/2}^c(\mathbf{y}_t^n))$ , the closure of the complement of a ball with radius  $\delta/2 > 0$  around the point  $\mathbf{y}_t^n$ ; and
- $\tau^\epsilon = \text{Min}\{t : \mathbf{y}_t^\epsilon \in \mathbb{D}_t\}$ .

Then,

$$\lim_{\epsilon \rightarrow 0} \epsilon^2 \ln P_{\mathbf{y}_0^n} \{\tau^\epsilon \leq t\} = - \inf_{\{\phi: \phi_0 = \mathbf{y}_0^n, \|\phi_t - \mathbf{y}_t^n\| > \delta/2\}} S_{0t}(\phi). \quad (14)$$

Proofs of Theorems 1 and 2 can be found in [2].

*Nominal belief:* Starting from  $\mathbf{b}_0^p = \mathbf{b}_0$ , the nominal belief evolution is given by  $\dot{\mathbf{b}}_t^p = \boldsymbol{\tau}(\mathbf{b}_t^p, \mathbf{u}_t^p, \dot{\mathbf{z}}_t^p)$ . Given equations (6),  $\mathbf{b}_t^p = ((\hat{\mathbf{x}}_t^p)^T, \text{vec}(\mathbf{P}_t^p)^T)^T = ((\mathbf{x}_t^p)^T, \text{vec}(\mathbf{P}_t^p)^T)^T$ , and linearizing  $\boldsymbol{\tau}$  only involves linearization of mean evolution:

$$\dot{\mathbf{b}}_t = \dot{\mathbf{b}}_t^p + \mathbf{T}_t^{\mathbf{b}}(\mathbf{b}_t - \mathbf{b}_t^p) + \mathbf{T}_t^{\mathbf{u}}(\mathbf{u}_t - \mathbf{u}_t^p) + \mathbf{T}_t^{\mathbf{z}}(\dot{\mathbf{z}}_t - \dot{\mathbf{z}}_t^p) + o(\delta),$$

with the Jacobians defined as usual.

*Linearization of belief and cost:* To address problem (1), we discretize the equations (5) in time with the discretization interval of  $dt \equiv 1$ . Let  $J^p := \sum_{t=0}^{K-1} c_t(\mathbf{b}_t^p, \mathbf{u}_t^p) + c_K(\mathbf{b}_K^p)$ , and linearize the cost function  $J^\pi$  around the nominal trajectories:

$$J^\pi = J^p + \tilde{J}_1 + o(\delta), \quad (15)$$

with  $\tilde{J}_1 = \sum_{t=0}^{K-1} (\mathbf{C}_t^{\mathbf{b}}(\mathbf{b}_t - \mathbf{b}_t^p) + \mathbf{C}_t^{\mathbf{u}}(\mathbf{u}_t - \mathbf{u}_t^p)) + \mathbf{C}_K^{\mathbf{b}}(\mathbf{b}_K - \mathbf{b}_K^p)$ , where the Jacobians are defined as usual.

If  $\|\mathbf{x}_K - \mathbf{x}_K^n\| \leq \delta/2$  and  $\|\mathbf{x}_K^n - \mathbf{x}_K^p\| \leq \delta/2$ , using the triangle inequality (note: as  $\epsilon \downarrow 0$ , using Theorems 1, 2, and Lemma 1, the probability of the first and second events tend exponentially to one, respectively; similarly for  $\|\hat{\mathbf{x}}_K - \mathbf{x}_K^p\|$ ):

$$\|\mathbf{x}_K - \mathbf{x}_K^p\| \leq \|\mathbf{x}_K - \mathbf{x}_K^n\| + \|\mathbf{x}_K^n - \mathbf{x}_K^p\| \leq \delta, \quad (16)$$

which means that all the linearizations are valid with a probability that tends to one as  $\epsilon \downarrow 0$ .

*Theorem 3 (First Order Cost Function Error):* For a time-discrete system, under a first-order approximation for the small noise paradigm, the stochastic cost function is dominated by the nominal part of the cost function, and the expected first-order error is zero:

$$\mathbb{E}[\tilde{J}_1] = 0.$$

Moreover, if the initial, process, and observation noises at each time are distributed according to zero mean Gaussian distributions, then  $\tilde{J}_1$  also has a zero mean Gaussian distribution.

*Corollary 1: Separation of the Open-Loop and Closed-Loop Designs Under Small Noise:* Based on Theorem 3, under the small noise paradigm, as  $\epsilon \downarrow 0$ , the design of the feedback law can be conducted separately from the design of the open loop optimized trajectory. Furthermore, this result holds with a probability that exponentially tends to one as  $\epsilon \downarrow 0$ .

Our separation principle combined with the usual separation principle provides a design structure where the optimal designs of the control law, nominal trajectory and estimator can be separated from each other. Thus, we couple the latter two, and design a nominal trajectory that aims for the best nominal estimation performance, which coincides with the Trajectory-optimized Linear Quadratic Gaussian (T-LQG) design [3].

**Problem 2: Trajectory Planning Problem:** Given an initial belief  $\mathbf{b}_0$ , a goal region of a ball with radius  $r_g$  around a goal state  $\mathbf{x}_g \in \mathbb{X}$ , horizon  $K > 0$ , and  $\mathbf{W}_t^u \succeq 0$ , solve:

$$\min_{\mathbf{u}_{0:K-1}^p} \sum_{t=1}^K [\text{tr}(\mathbf{P}_{\mathbf{b}_t^p}^+) + (\mathbf{u}_{t-1}^p)^T \mathbf{W}_t^u \mathbf{u}_{t-1}^p] \quad (17a)$$

$$s.t. \quad \mathbf{P}_t^- = \mathbf{A}_{t-1} \mathbf{P}_{t-1}^+ \mathbf{A}_{t-1}^T + \epsilon^2 \mathbf{G}_{t-1} \Sigma_\omega \mathbf{G}_{t-1}^T, \quad (17a)$$

$$\mathbf{S}_t = \mathbf{H}_t \mathbf{P}_t^- \mathbf{H}_t^T + \epsilon^2 \Sigma_\nu, \quad (17b)$$

$$\mathbf{P}_t^+ = (\mathbf{I} - \mathbf{P}_t^- \mathbf{H}_t^T \mathbf{S}_t^{-1} \mathbf{H}_t) \mathbf{P}_t^-, \quad (17c)$$

$$\mathbf{P}_0^+ = \epsilon^2 \Sigma_{\mathbf{x}_0}, \quad (17d)$$

$$\mathbf{x}_0^p = \mathbb{E}[\mathbf{b}_0], \quad (17e)$$

$$\mathbf{x}_{t+1}^p = \mathbf{f}(\mathbf{x}_t^p, \mathbf{u}_t^p), \quad 0 \leq t \leq K-1, \quad (17f)$$

$$\|\mathbf{x}_K^p - \mathbf{x}_g\|_2 < r_g, \quad (17g)$$

$$\|\mathbf{u}_t^p\|_2 \leq r_u, \quad 1 \leq t \leq K. \quad (17h)$$

*Control policy:* After linearizing the equations around the optimized nominal trajectory, the resulting control policy is a linear feedback policy [1],  $\mathbf{u}_t = \mathbf{u}_t^p - \mathbf{L}_t^p (\hat{\mathbf{x}}_t - \mathbf{x}_t^p)$ , where the feedback gain  $\mathbf{L}_t^p$  is:

$$\mathbf{L}_t^p = (\mathbf{W}_t^u + (\mathbf{B}_t^p)^T \mathbf{P}_{t+1}^f \mathbf{B}_t^p)^{-1} (\mathbf{B}_t^p)^T \mathbf{P}_{t+1}^f \mathbf{A}_t^p,$$

and the matrix  $\mathbf{P}_t^f$  is the result of backward iteration of the dynamic Riccati equation

$$\mathbf{P}_{t-1}^f = (\mathbf{A}_t^p)^T \mathbf{P}_t^f \mathbf{A}_t^p - (\mathbf{A}_t^p)^T \mathbf{P}_t^f \mathbf{B}_t^p (\mathbf{W}_t^u + (\mathbf{B}_t^p)^T \mathbf{P}_t^f \mathbf{B}_t^p)^{-1} (\mathbf{B}_t^p)^T \mathbf{P}_t^f \mathbf{A}_t^p + \mathbf{W}_t^x,$$

which is solvable with a terminal condition  $\mathbf{P}_K^f = \mathbf{W}_K^x \succeq 0$ .

#### IV. SIMULATION RESULTS

We consider a non-holonomic car-like robot in an environment with road-blocks equipped with landmark-based range and bearing measurement model. We use the MATLAB `fmincon` optimizer with no initial trajectory to obtain the solution of problem (2). For collision-avoidance, we utilize the Obstacle Barrier Function (OBF) method of [3]. Figures 1, 2, and 3 show the optimized planned, execution, and estimate trajectories, respectively.

#### V. CONCLUSION

We considered the general problem of controlling a stochastic nonlinear system with process and measurement uncertain-

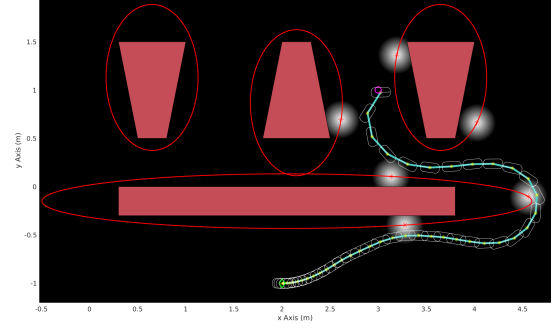


Fig. 1. Optimized planned trajectory for a car-like robot. Landmarks are shown with light areas. The planning horizon is 40 steps. Ellipsoids show the safety margin of the collision-avoidance function. Initial state is  $(x, y, \theta) = (2, -1, 0)$ , goal state is  $(3, 1, 0)$ ,  $\Sigma_{\mathbf{x}_0} = \Sigma_\omega = 0.01\mathbf{I}_{n_x}$ , and  $\Sigma_\nu = 0.01\mathbf{I}_{n_z}$ .

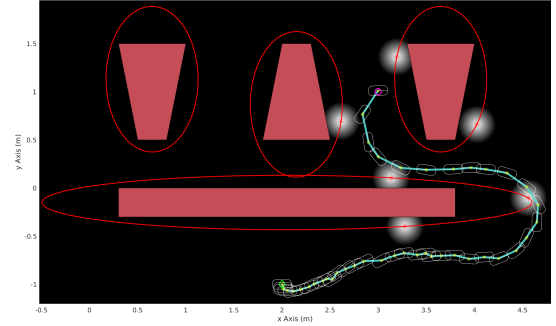


Fig. 2. A typical execution trajectory. Since, no significant deviation occurred, planning was only performed once.

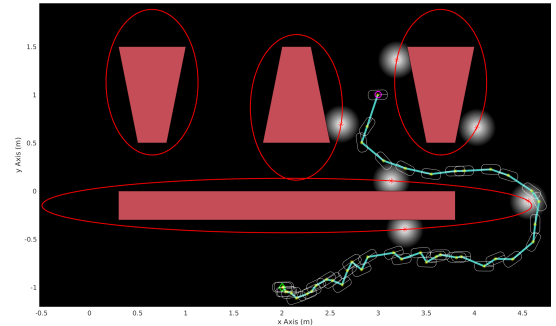


Fig. 3. Estimate trajectory. Although KF is used for planning, Extended KF is used during execution for better performance.

ties. We used the Wentzell-Freidlin theory of large deviations and provided a novel result of a “separation of the open-loop and closed-loop designs”. This result, combined with the usual separation principle (of the estimator and controller designs) leads to an asymptotically-optimal design approach of the Trajectory-optimized Linear Quadratic Gaussian (T-LQG) under small noise, and a near-optimal design for moderate noise levels involving a polynomial order of calculations of minimum order.

#### BIBLIOGRAPHY

- [1] P. R. Kumar and P. P. Varaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Englewood Cliffs: Prentice-Hall, 1986.
- [2] A. D. Wentzell, *Limit theorems on large deviations for Markov stochastic processes*. Springer Science & Business Media, 2012, vol. 38.
- [3] M. Rafieisakhaei, S. Chakravorty, and P. Kumar, “Belief Space Planning Simplified: Trajectory-Optimized LQG (T-LQG),” *arXiv preprint arXiv:1608.03013*, 2016.