

Risk Aversion in Belief-space Planning under Measurement Acquisition Uncertainty

Stephen M. Chaves, Jeffrey M. Walls, Enric Galceran, and Ryan M. Eustice

Abstract—This paper reports on a Gaussian belief-space planning formulation for mobile robots that includes random measurement acquisition variables that model whether or not each measurement is actually acquired. We show that maintaining the stochasticity of these variables in the planning formulation leads to a *random* belief covariance matrix, allowing us to consider the risk associated with the acquisition in the objective function. Inspired by modern portfolio theory and utility optimization, we design objective functions that are *risk-averse*, and show that risk-averse planning leads to decisions made by the robot that are desirable when operating under uncertainty. We show the benefit of this approach using simulations of a planar robot traversing an uncertain environment and of an underwater robot searching for loop-closure actions while performing visual SLAM.

I. INTRODUCTION

Autonomous mobile robots operating in real-world environments often perform simultaneous localization and mapping (SLAM) to estimate their state and model their environment [1]. A robot must plan actions within this framework in order to perform tasks like exploration, inspection, target-tracking, reconnaissance, and others. This has led to research in *belief-space planning*, where the robot makes decisions about actions to execute based on the belief of its state and other variables of interest [2].

Recently, the Gaussian belief-space planning problem was extended to include stochastic measurements [3], leading to a robot belief that is a “distribution of distributions.” In this case, the belief is represented by a mean state vector that is stochastic and a covariance matrix that is deterministic, as the covariance update equations depend on the *a priori* known measurement noise covariance but not the random measurement value.

In this paper, we develop a belief-space planning formulation that also includes stochastic measurement acquisition variables. The acquisition variables are Bernoulli random variables that model whether or not each measurement is actually acquired. For example, an acquisition variable might describe whether a landmark is in the field of view of the robot, whether two camera images overlap and are registered, or whether a message is received over a lossy communication channel. As a motivating example, consider the underwater environment encountered in ship hull inspection. Visual features in this environment tend to be sparsely distributed

*This work was supported by the Office of Naval Research under award N00014-12-1-0092; S. Chaves was supported by The SMART Scholarship for Service Program by the Department of Defense.

S. Chaves, J. Walls, E. Galceran, and R. Eustice are with the University of Michigan, Ann Arbor, MI 48109, USA {schaves, jmwalls, egalceran, eustice}@umich.edu.

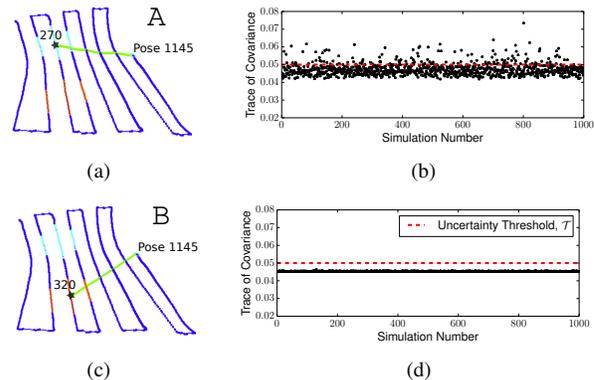


Fig. 1: Risk-averse planning for an underwater robot performing visual SLAM. The robot must decide between path A (a) and path B (c) for gathering loop-closure camera registrations. Previous methods select path A, which has a lower expected uncertainty and shorter path length than B, but the proposed risk-averse planning framework selects path B. Monte Carlo simulations in (b) and (d) show that path B is indeed preferable. Path A exceeds the final pose uncertainty threshold in 147 of 1000 trials.

such that much of the imagery collected by the robot is not useful for registration. The probability of a successful two-view registration (i.e., probability of acquisition) is directly related to a measure of visual saliency of the proposed camera images [4]. At execution time, the SLAM system does not explicitly model the acquisition variables, but simply incorporates successful measurements when they occur. However, for these systems, the nondeterministic nature of measurement acquisition should be considered at planning time when the set of received measurements is still unknown. We show that maintaining the stochasticity of the acquisition variables throughout the planning formulation leads to a *random* covariance matrix over the robot’s state, allowing us to consider the risk associated with measurement acquisition in the objective function.

The contributions of this work can be summarized as:

- We include measurement acquisition variables in the planning prediction and show that maintaining their stochasticity throughout the formulation leads to a random belief covariance matrix.
- We design *risk-averse* objective functions for selecting control actions that account for the stochasticity of the belief with respect to the acquisition variables.
- We show that risk-averse planning under uncertainty leads to decisions that result in more desirable outcomes than decisions made with traditional approaches.

Fig. 1 illustrates the benefit of the proposed risk-averse planning framework with an autonomous underwater robot

performing active visual SLAM. This example is discussed further in §IV.

A. Related Work

The integration of planning with SLAM has its roots in active exploration [5–7], where research focused on reducing uncertainty in the map representation. Combining traditional sampling-based planners with decision-theoretic formulations on uncertainty reduction led to sampling-based approaches that plan in the belief-space of the robot [8–10]. Lately, belief-space planning research has focused on trajectory smoothing frameworks for planning in the continuous domain of control actions [2, 3, 11]. These works find locally-optimal solutions to the planning problem and provide promising results for robotics applications, especially point-to-point planning queries. Work in path planning for information gathering [12] and active SLAM systems [13–15] focused more on the interaction between planning and SLAM, and how the performance and efficiency of SLAM is improved with intelligent decisions regarding which paths to travel.

Our proposed methodology is specifically interested in accurately modeling the stochasticity of variables within the planning formulation. Van den Berg et al. [3] first relaxed the assumption of maximum-likelihood measurements in planning [16] to consider the measurements as random variables in the prediction. In this paper, we include random acquisition variables as well. Our approach is similar to the formulation of Sinopoli et al. [17], who studied Kalman filtering given measurements over a lossy communication channel. They derived the Kalman filter equations as functions of the stochastic acquisition variables. Kim and Eustice [14] and Indelman et al. [2] included acquisition variables in belief-space planning for robotics, but their formulations removed the effect of the acquisition randomness in the resulting belief. Instead, we seek to model the variability in the outcomes with respect to uncertainty in order to design objective functions for planning that are sensitive to risk.

Extensive ongoing research in economics and finance continues to examine methods for making smart investment decisions given a number of risky assets [18]. Modern portfolio theory originates from the classical methods of Markowitz [19], who used the graphical concept of an efficient frontier to maximize expected return for a given amount of tolerance to risk. Another approach to portfolio optimization, which we examine in this paper, optimizes a von Neumann-Morgenstern utility function that defines rational investor behavior [20]. This approach has been frequently examined in the broader literature discussing risk aversion in Markov decision processes (MDPs), specifically related to exponential utility optimization [21, 22] and stemming from the original work of Howard and Matheson [23]. Our proposed method also closely resembles methods for analyzing the internal variance [24, 25] and parametric variance [26, 27] present in MDPs, although our focus is on developing a framework for planning in real-world mobile robotics applications.

II. METHOD

The derivation of our method closely follows that of Indelman et al. [2] in format and notation.

A. Optimal Planning Problem

We are interested in finding a set of control actions, $U_{0:K-1}$, over a horizon of K planning steps¹ that minimizes some function of the robot’s belief over this horizon, $\mathcal{B}_{0:K}$. The optimal planning problem is formulated as

$$U_{0:K-1}^* = \arg \min_{U_{0:K-1}} J(\mathcal{B}_{0:K}, U_{0:K-1}), \quad (1)$$

with an objective function comprised of stage costs and a final cost dependent on the predicted belief of the robot at each planning step:

$$J(\mathcal{B}_{0:K}, U_{0:K-1}) = \sum_{k=0}^{K-1} c_k(\mathcal{B}_k, \mathbf{u}_k) + c_K(\mathcal{B}_K). \quad (2)$$

Previous works solved for the control actions in a trajectory-smoothing optimization, such as gradient descent [2] or dynamic programming [3]. Alternatively, we can perform this optimization within a sampling-based planning framework by selecting the sampled set of control actions that minimizes the objective [9, 10]. We will return to the discussion of the optimal planning problem after examining the belief of the robot and its evolution.

B. Belief Inference

We define the belief of the robot at a given planning step $k \in [1, K]$ as

$$\mathcal{B}_k = p(X_k | \mathcal{Z}_0, \mathcal{U}_0, Z_{1:k}, \Gamma_{1:k}, U_{0:k-1}), \quad (3)$$

where X_k is the state vector of interest and \mathcal{Z}_0 and \mathcal{U}_0 are the prior measurements and controls, respectively, up to the planning event. $Z_{1:k}$ are the random update measurements and $\Gamma_{1:k}$ are the corresponding random acquisition variables. The belief vector is represented by a multivariate Gaussian with mean and covariance matrix (given in terms of the inverse information matrix):

$$\mathcal{B}_k \sim \mathcal{N}(X_k^*, \Lambda_k^{-1}), \quad (4)$$

found using the maximum *a posteriori* (MAP) estimate

$$X_k^* = \arg \min_{X_k} -\log \mathcal{B}_k. \quad (5)$$

The Gaussian motion model for transitioning to step $k+1$ from step k is

$$\begin{aligned} \mathbf{x}_{k+1} &= f(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k, \\ \mathbf{w}_k &\sim \mathcal{N}(\mathbf{0}, \Omega_w^{-1}). \end{aligned} \quad (6)$$

However, at planning time, the update measurements ($Z_{1:k}$) are unknown. In addition, it is also unknown whether or not each measurement will be acquired. Therefore, we introduce a Bernoulli random variable for each measurement that models its acquisition. The set of random acquisition variables at

¹Note that we do not compute a policy, but a set of actions to be executed in an open-loop or model predictive control scheme with replanning.

a given planning step k is $\Gamma_k = \{\gamma_{k,j}\}_{j=1}^{n_k}$, where n_k is the number of possible measurements at step k . Therefore, we use the following Gaussian observation model for the update measurements:

$$\begin{aligned} \mathbf{z}_{k,j} &= h(X_k^j) + \mathbf{v}_{k,j}(\gamma_{k,j}), \\ \mathbf{v}_{k,j}(\gamma_{k,j}) &\sim \mathcal{N}(\mathbf{0}, (\gamma_{k,j}\Omega_v^{k,j} + (1 - \gamma_{k,j})\Omega_0)^{-1}), \end{aligned} \quad (7)$$

where $\Omega_v^{k,j}$ is the information contributed by a successful measurement and Ω_0 is the information contributed by an unsuccessful measurement. It is easily identified that unsuccessful measurements add zero information; that is, the second term of (7) vanishes, resulting in the following observation model that is still Gaussian:

$$\begin{aligned} \mathbf{z}_{k,j} &= h(X_k^j) + \mathbf{v}_{k,j}(\gamma_{k,j}), \\ \mathbf{v}_{k,j}(\gamma_{k,j}) &\sim \mathcal{N}(\mathbf{0}, (\gamma_{k,j}\Omega_v^{k,j})^{-1}). \end{aligned} \quad (8)$$

Using the standard assumption of an uninformative prior on the measurements, we can write the distribution of the state from (3) as

$$\begin{aligned} p(X_k | \mathcal{Z}_0, \mathcal{U}_0, Z_{1:k}, \Gamma_{1:k}, U_{0:k-1}) &\propto \\ p(X_0 | \mathcal{Z}_0, \mathcal{U}_0) \prod_{i=1}^k p(\mathbf{x}_i | \mathbf{x}_{i-1}, \mathbf{u}_{i-1}) p(Z_i, \Gamma_i | X_i), \end{aligned} \quad (9)$$

where

$$p(Z_i, \Gamma_i | X_i) = \prod_{j=1}^{n_i} p(\mathbf{z}_{i,j} | \gamma_{i,j}, X_i^j) p(\gamma_{i,j} | X_i^j), \quad (10)$$

and X_i^j are the state variables associated with measurement j at planning step i .

Online, each $\gamma_{i,j}$ of interest is observed by the robot upon receiving the associated measurement $\mathbf{z}_{i,j}$. Acquisition variables corresponding to measurements not received do not inform the estimate of the state. But within the prediction, it is unknown which measurements will be received ahead of time. Previous work incorporated the random acquisition variables within an expectation-maximization (EM) framework [2], but this method results in a *deterministic* covariance matrix by evaluating $\gamma_{i,j}$ at its mean value of $p(\gamma_{i,j} = 1)$ within the MAP estimate. Instead, we want to maintain the randomness of the acquisition variables throughout the formulation such that the robot can be aware of the associated acquisition risk in the optimization. We approximate the acquisition variables as independent, and therefore uninformative to the estimate, such that $p(\gamma_{i,j} | X_i^j) \approx p(\gamma_{i,j})$, allowing (9) to take the form

$$p(X_0 | \mathcal{Z}_0, \mathcal{U}_0) \prod_{i=1}^k p(\mathbf{x}_i | \mathbf{x}_{i-1}, \mathbf{u}_{i-1}) \prod_{j=1}^{n_i} p(\mathbf{z}_{i,j} | \gamma_{i,j}, X_i^j). \quad (11)$$

This approach essentially borrows the rationale behind EM but delays taking the expectation over the acquisition variables until the evaluation of the objective function (described later in §II-C).

Inserting the Gaussian motion and observation models of (6) and (8) into the MAP estimate of (5) and (11), we

minimize the negative log likelihood to arrive at a nonlinear least-squares problem common to graph-based SLAM [1]:

$$\begin{aligned} X_k^* &= \arg \min_{X_k} \left[\|X_0 - X_0^*\|_{\Lambda_0}^2 + \right. \\ &\quad \sum_{i=1}^k \|f(\mathbf{x}_{i-1}, \mathbf{u}_{i-1}) - \mathbf{x}_i\|_{\Omega_w}^2 + \\ &\quad \left. \sum_{i=1}^k \sum_{j=1}^{n_i} \gamma_{i,j} \|h(X_i^j) - \mathbf{z}_{i,j}\|_{\Omega_v^{i,j}}^2 \right], \end{aligned} \quad (12)$$

where both $\gamma_{i,j}$ and $\mathbf{z}_{i,j}$ are random. We can compute a linearization point for the problem by compounding the given set of controls, yielding the nominal mean estimate $\bar{X}_k(U_{0:k-1}) = \{X_0^*, \bar{\mathbf{x}}_1, \dots, \bar{\mathbf{x}}_k\}$. Linearizing about this nominal mean estimate, the problem collapses into the following representation for the state update vector ΔX_k :

$$\|A_k(U_{0:k-1})\Delta X_k - \mathbf{b}_k(Z_{1:k}, U_{0:k-1})\|_{\mathcal{G}_k(\Gamma_{1:k})}^2, \quad (13)$$

where

$$A_k = \begin{bmatrix} \left[\begin{array}{cc} \Lambda_0^{\frac{1}{2}} & 0 \end{array} \right] \\ \mathcal{D}(\Omega_w)^{\frac{1}{2}} F_k \\ \mathcal{D}(\Omega_v^{i,j})^{\frac{1}{2}} H_k \end{bmatrix}, \quad \mathbf{b}_k = \begin{pmatrix} \mathbf{0} \\ \mathcal{D}(\Omega_w)^{\frac{1}{2}} \mathbf{b}_k^f \\ \mathcal{D}(\Omega_v^{i,j})^{\frac{1}{2}} \mathbf{b}_k^h \end{pmatrix}, \quad (14)$$

and

$$\mathcal{G}_k = \begin{bmatrix} I \\ I \\ \mathcal{D}(\gamma_{i,j}) \end{bmatrix}. \quad (15)$$

Here, $\mathcal{D}(\cdot)$ denotes a diagonal matrix with the specified elements, F_k and H_k are the sparse Jacobians from the motion and observation models, respectively, and \mathbf{b}_k^f and \mathbf{b}_k^h are the corresponding residual vectors with stacked elements

$$\mathbf{b}_i^f = \bar{\mathbf{x}}_i - f(\bar{\mathbf{x}}_{i-1}, \mathbf{u}_{i-1}), \quad (16)$$

$$\mathbf{b}_{i,j}^h = \mathbf{z}_{i,j} - h(\bar{X}_i^j). \quad (17)$$

Solving (13) around the linearization point \bar{X}_k , we find the update vector as a function of the random measurements and the random acquisition variables:

$$\Delta X_k(Z_{1:k}, \Gamma_{1:k}, U_{0:k-1}) = (A_k^\top \mathcal{G}_k A_k)^{-1} A_k^\top \mathcal{G}_k \mathbf{b}_k. \quad (18)$$

Thus, the belief at planning step k is represented by the mean vector $X_k^* = \bar{X}_k + \Delta X_k$ and the associated information matrix as a function of the acquisition variables,

$$\Lambda_k(\Gamma_{1:k}, U_{0:k-1}) = A_k^\top \mathcal{G}_k A_k. \quad (19)$$

This gives us a stochastic belief as a function of the random measurements and acquisition variables at each planning step, k .

C. Objective Functions

Here we examine costs to insert into the general objective function of (2). We recall that we derived the belief as a function of the random measurements and acquisition variables, meaning we must take the expectation of the objective function with respect to these variables. Following

the literature [2, 3], we consider costs at stage k that penalize control effort and the robot uncertainty:

$$c_k(\mathcal{B}(X_k), \mathbf{u}_k) = g_u(\mathbf{u}_k) + \mathbb{E}_{\Gamma_{1:k}} [g_\Lambda(\Lambda_k^{-1})], \quad (20)$$

where we leave the functions $g(\cdot)$ undefined for now. The expectation is taken on the uncertainty term because the information matrix Λ_k is stochastic due to the effect of the random acquisition variables. Similarly, we can write the final cost to penalize distance from a desired goal pose and the final robot uncertainty:

$$c_K(\mathcal{B}(X_K)) = \mathbb{E}_{Z_{1:K}, \Gamma_{1:K}} [g_{\mathbf{x}}(\mathbf{x}_K^* - \mathbf{x}_G)] + \mathbb{E}_{\Gamma_{1:K}} [g_\Lambda(\Lambda_K^{-1})]. \quad (21)$$

The expectation is taken for the goal pose term since the mean estimate of the belief is random, as it is a function of both the random measurements and the random acquisition variables, evidenced in (18).

III. RISK AVERSION

Given the random belief from the planning prediction, we seek to design objective functions that consider the stochasticity. Modern portfolio theory provides insight into how to design objective functions that are *risk-averse*. We can consider the robot as an investor with the goal of maximizing its wealth from a number of risky investments. However, rather than maximizing wealth, the robot seeks to maximize information from a number of uncertain sensor measurements.

Investor behavior can be encoded in a utility function of wealth, $\mathcal{U}(W)$, such that maximizing the expected utility, $\mathbb{E}[\mathcal{U}(W)]$, results in more desirable decisions than directly maximizing the expected wealth, $\mathbb{E}[W]$. A utility function that encodes rational investor behavior satisfies four axioms [18]:

- 1) Investors exhibit non-satiation, $\mathcal{U}'(W) > 0$.
- 2) Investors exhibit risk aversion, $\mathcal{U}''(W) < 0$.
- 3) Investors exhibit decreasing absolute risk aversion, $\mathcal{A}'(W) < 0$, where

$$\mathcal{A}(W) = -\frac{\mathcal{U}''(W)}{\mathcal{U}'(W)}. \quad (22)$$

- 4) Investors exhibit constant relative risk aversion, $\mathcal{R}'(W) = 0$, where

$$\mathcal{R}(W) = W \cdot \mathcal{A}(W). \quad (23)$$

The first two axioms equate to $\mathcal{U}(W)$ being monotonic and concave with respect to W . Absolute risk aversion is related to the absolute amount of wealth an investor puts toward risky assets. Similarly, relative risk aversion is related to the fraction of wealth invested in risky assets.

In the robotics planning literature, it is common to design costs that are quadratic with respect to the control effort and distance from a goal pose. However, the uncertainty costs within the planning objective are typically linear in the trace (or determinant) of the belief covariance [2, 3, 6, 15].

The linear objective function is monotonic but not concave, meaning it is *risk-neutral*. Without randomness in the acquisition, this cost is sensible because the belief covariance is deterministic. However, our formulation leads to a random belief covariance.

Instead, we prefer to design objective functions that are risk-averse. We could consider replacing the linear cost with a quadratic cost, but despite being risk-averse, the quadratic function does not satisfy the third and fourth axioms above [18]. Specifically, the quadratic function exhibits *increasing* absolute and *increasing* relative risk aversion. There is a similar drawback with the exponential utility function [22], which exhibits constant absolute risk aversion and increasing relative risk aversion.

A utility function that follows rational investor behavior defined by the four axioms is the power function [18], given by

$$\mathcal{U}(W) = \begin{cases} \frac{W^{(1-\eta)}}{(1-\eta)}, & \eta \neq 1 \\ \log W, & \eta = 1 \end{cases}. \quad (24)$$

Here, the relative risk aversion actually equals the tunable (user-defined) parameter η , such that $\eta = 0$ corresponds to a risk-neutral utility and risk aversion increases as η increases. Since we seek to minimize uncertainty rather than maximize wealth, we define

$$W = \mathcal{T} - \text{tr}(\Lambda_k^{-1}) = \mathcal{T} - \mathbf{m}^\top \text{vec}(\Lambda_k^{-1}), \quad (25)$$

where \mathcal{T} is a user-specified upper bound on the uncertainty and the trace is expanded using an element selection vector \mathbf{m} . This allows us to write an equivalent penalty function to the power utility function:

$$\mathcal{P}(W) = -\mathcal{U}(W) = \begin{cases} -\frac{W^{(1-\eta)}}{(1-\eta)}, & \eta \neq 1 \\ -\log W, & \eta = 1 \end{cases}. \quad (26)$$

For a random belief covariance and $\eta \neq 1$, the expected value of the power penalty function is approximated using a Taylor series expansion as

$$\mathbb{E}_{\Gamma_{1:k}} [\mathcal{P}(W)] \approx -\frac{\mathbb{E}[W]^{(1-\eta)}}{1-\eta} + \frac{\eta}{2} \text{Var}[W] \mathbb{E}[W]^{(-\eta-1)}, \quad (27)$$

with

$$\begin{aligned} \mathbb{E}_{\Gamma_{1:k}} [W] &= \mathcal{T} - \mathbf{m}^\top \mathbb{E}_{\Gamma_{1:k}} [\text{vec}(\Lambda_k^{-1})] \\ &\approx \mathcal{T} - \mathbf{m}^\top \text{vec} \left(\mathbb{E}_{\Gamma_{1:k}} [\Lambda_k]^{-1} \right), \end{aligned} \quad (28)$$

and

$$\text{Var}_{\Gamma_{1:k}} [W] = \mathbf{m}^\top \text{Var}_{\Gamma_{1:k}} [\text{vec}(\Lambda_k^{-1})] \mathbf{m}. \quad (29)$$

Finding the necessary terms in the variance equation (29) is easier with the helpful equation for first-order propagation of uncertainty:

$$\text{Var}[y(X)] \approx \frac{\partial y}{\partial X} \Big|_{\mathbb{E}[X]} \cdot \text{Var}[X] \cdot \frac{\partial y}{\partial X} \Big|_{\mathbb{E}[X]}^\top. \quad (30)$$

The variance of the belief covariance matrix is written in terms of the variance of the belief information matrix:

$$\text{Var}_{\Gamma_{1:k}} [\text{vec}(\Lambda_k^{-1})] \approx L_k \cdot \text{Var}_{\Gamma_{1:k}} [\text{vec}(\Lambda_k)] \cdot L_k^\top, \quad (31)$$

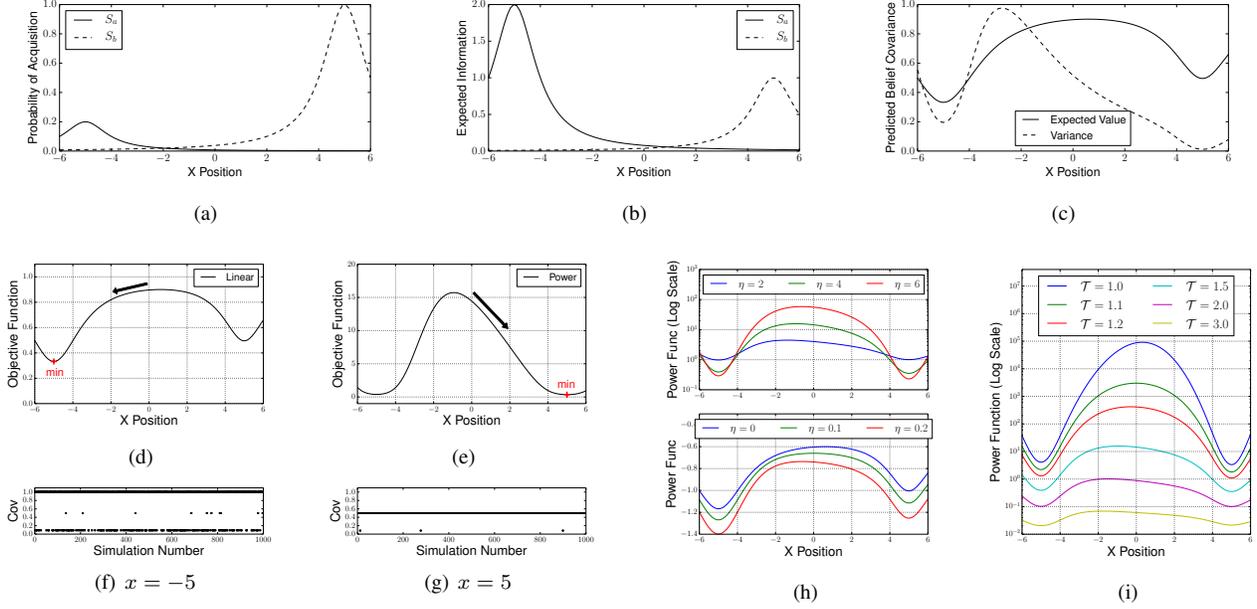


Fig. 2: One-dimensional intuition. (a) The functions describing the probability of acquisition for each measurement source in the one-dimensional example. (b) The expected value of information contributed to the belief by a measurement from each source. (c) The predictions for the expected value and variance of the belief covariance as a function of sensor placement. (d) A linear objective function of $J = \Lambda^{-1}$, which simply minimizes the expected value of the predicted belief covariance. (e) The power objective function of the form of (26), with $\eta = 4$ and $\mathcal{T} = 1.5$. (f)(g) The resulting covariance for 1000 trials of Monte Carlo simulations of the sensor belief, placed at $x = \{-5, 5\}$. (h) The power objective function graphed for varying values of the parameter η with $\mathcal{T} = 1.5$. (i) The power objective function graphed for varying values of the uncertainty bound \mathcal{T} with $\eta = 4$.

where the partial derivative L_k is

$$L_k = - \left(\mathbb{E} [\Lambda_k]^{-\top} \otimes \mathbb{E} [\Lambda_k]^{-1} \right), \quad (32)$$

and \otimes denotes the Kronecker product. The variance of the belief information matrix is

$$\text{Var} [\text{vec}(\Lambda_k)] = M_k \cdot \text{Var}[\Gamma_{1:k}] \cdot M_k^\top, \quad (33)$$

with each column of the partial derivative M_k (indexed by c) corresponding to an individual $\gamma_{i,j}$ given by

$$M_k^{(c)} = \text{vec} \left(H_k^{i,j \top} \Omega_v^{i,j} H_k^{i,j} \right). \quad (34)$$

L_k and M_k are the Jacobians used to propagate the uncertainty in the acquisition variables into the space of the belief covariance. It is worth noting that $H_k^{i,j}$ is sparse for many robotics applications, allowing us to efficiently compute (31) by leveraging sparsity patterns.

We propose the use of the risk-averse power penalty function within the planning objective function. Replacing the commonly-used linear uncertainty costs with the power function of (26) naturally encodes rational decision-making with respect to the belief uncertainty.

IV. RESULTS

We now present simulation results that show the effect of the random acquisition variables on the belief and the benefit of risk-averse planning.

A. One-dimensional Intuition

The following example shown in Fig. 2 illustrates the intuition behind the method for risk-averse planning. We are interested in placing a sensor in a one-dimensional environment given prior belief information $\Lambda_0 = 1.0$. The sensor is able to receive measurements from two sources with constant information. Measurement source S_a has information Ω_a and measurement source S_b has information $\Omega_b = 0.1\Omega_a$. Each measurement source also has a binary variable describing whether it is acquired with a parameter dependent on the placement of the sensor. As such, acquisition variable γ_a reaches a peak probability of success of $\max p(\gamma_a = 1) = 0.2$ at $x = -5$. Acquisition variable γ_b reaches a peak probability of success of $\max p(\gamma_b = 1) = 1.0$ at $x = 5$. The state-dependent parameter functions for the acquisition variables are shown in Fig. 2(a). Each measurement source contributes expected information $\mathbb{E}[\Lambda_i] = p(\gamma_i = 1)\Omega_i$, shown over the one-dimensional environment in Fig. 2(b). At their respective peak probabilities of acquisition, S_a contributes twice the expected information than source S_b , as S_a is 10 times more informative but S_b is 5 times more likely to be acquired. Using the method from §II, Fig. 2(c) shows the predictions of the expected value and variance of the belief covariance for placing the sensor along the environment.

Consider the simple risk-neutral, linear objective function of $\text{tr}(\Lambda^{-1})$, graphed in Fig. 2(d). With an initial sensor placement of $x = 0$ and a gradient descent update framework, the placement follows the gradient and converges to $x = -5$. Now consider minimizing the risk-averse power function of (26) with $\eta = 4$ and $\mathcal{T} = 1.5$. This objective function accounts for the uncertain measurement acquisition

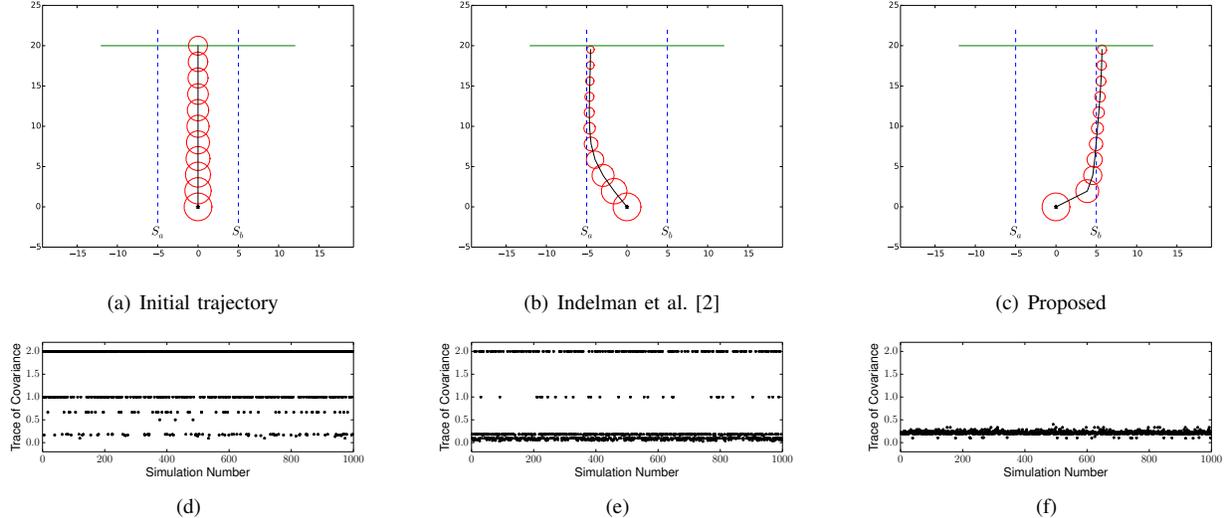


Fig. 3: Results from the planar robot example. (a) The initial trajectory. (b) The resulting trajectory from the gradient-based optimization method of Indelman et al. [2] with an objective function that is linear in the belief covariance. (c) The resulting trajectory from the gradient-based optimization with the proposed risk-averse formulation using an objective that includes a power function of the belief covariance. (d)(e)(f) The trace of the robot’s terminating covariance for 1000 trials of Monte Carlo simulations of the above trajectories. The proposed risk-averse method results in a path that consistently receives measurements for improved localization.

and is graphed in Fig. 2(e). In this case, the placement initially at $x = 0$ converges to the risk-averse location of $x = 5$. Monte Carlo simulations of the resulting covariance from each placement are shown in Fig. 2(f) and Fig. 2(g). While the placement at $x = -5$ often yields a very low uncertainty, it also often receives no measurements. The placement at $x = 5$ is guaranteed to receive the measurement from S_b . Fig. 2(h) and (i) plot the power penalty function with varying values of η and \mathcal{T} .

B. Planar Robot

We can apply the intuition from the previous example to a planar robot with control authority along the x and y directions. Rather than placing a sensor, we are interested in localizing the robot along a trajectory and reaching a goal region. Similar to the previous example, the robot receives absolute measurements in x and y from two different sources of constant information. The information and acquisition properties for these sources are the same as in the one-dimensional example (Fig. 2(a), (b)) but extend along the y -direction. The robot starts at pose $(x, y) = (0, 0)$ with the goal of reaching a final pose at planning step $K = 10$ with $y = 20$. We update the trajectory using a gradient descent optimization and use this example to show the applicability of our formulation to trajectory-smoothing planners.

In this example, we compare our proposed risk-averse planner to the method of Indelman et al. [2]. Their method uses the following forms of the costs in (20) and (21):

$$\begin{aligned} g_u(\mathbf{u}_k) &= \|\delta(\mathbf{u}_k)\|_{M_u}^2, \\ g_\Lambda(\Lambda_k^{-1}) &= \mathbf{m}^\top \text{vec}(\Lambda_k^{-1}), \\ g_x(\mathbf{x}_K^* - \mathbf{x}_G) &= \|\mathbf{x}_K^* - \mathbf{x}_G\|_{M_x}^2, \end{aligned} \quad (35)$$

where M represents a weight matrix and (in our implementation) $\delta(\cdot)$ penalizes deviation from the nominal step

length. The Indelman et al. approach does not model the variability in the belief covariance. With this method, the robot settles on a path traveling the peak acquisition zone of S_a , shown in Fig. 3(b). However, our proposed framework accounts for the randomness of the acquisition in the belief covariance matrix by replacing the uncertainty cost above with $g_\Lambda(\Lambda_k^{-1}) = \mathcal{P}(W)$ from (26). With the risk-averse optimization, the robot prefers a path traveling the peak acquisition zone of S_b , shown in Fig. 3(c).

The effect of the randomness of acquisition is clearly seen when we simulate 1000 runs of the robot traversing the selected paths. Fig. 3(d), (e), and (f) show the trace of the marginal covariance of the belief at the final planning step for the simulations. The uncertainty is often lower with the Indelman et al. path, but the proposed method path is preferable with respect to worst-case uncertainties. The risk-averse path consistently receives measurements for improved localization despite having larger expected uncertainty.

C. Underwater Visual Inspection Robot

Consider a Hovering Autonomous Underwater Vehicle (HAUV) performing visual SLAM to inspect a ship hull, as in Fig. 4. The robot executes a lawnmower-like trajectory over the underwater portion of the hull to collect camera images of the environment in a coverage-efficient manner. However, execution of this policy results in navigation drift, so the robot must perform loop-closing revisit actions throughout the mission to bound its uncertainty. Loop-closures in the visual SLAM formulation come from pairwise camera registrations between overlapping images. For this simulation, we design a synthetic environment where most of the environment is feature-less and has zero registrability. We use a measure of visual saliency [14] and a Gaussian process (GP) prediction [15] to model the distribution of

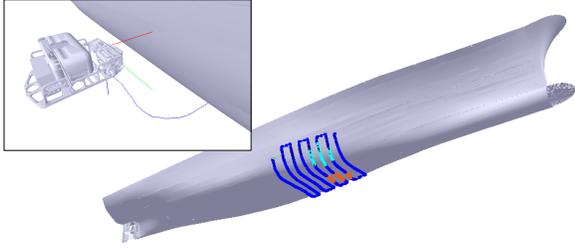


Fig. 4: Visualization of the HAUV performing visual ship hull inspection.

the acquisition variables throughout the environment. Given a path planning algorithm for finding possible revisit paths, we evaluate a candidate path based upon its distance traveled and its final uncertainty, which we frame as the risk-averse power penalty function. The objective function becomes

$$J = g_u(U_{0:K-1}) + \mathbb{E}_{\Gamma_{1:K}} [\mathcal{P}(\mathcal{T} - \mathbf{m}^\top \text{vec}(\Lambda_K^{-1}))], \quad (36)$$

where $g_u(U_{0:K-1})$ computes the path length (scaled by a weight), the threshold is set to $\mathcal{T} = 0.05$, and \mathbf{m} selects the diagonal elements of the final pose marginal covariance. The risk parameter is $\eta = 4$. We show the benefit of the proposed risk-averse framework within this type of sampling-based active SLAM system by comparing to the path evaluation method of our previous work [15].

Fig. 5 shows the HAUV deciding between two candidate loop-closure paths at pose number 770 of the mission. Candidate path A considers revisiting a moderately-salient portion of the environment centered at pose number 270 in the graph. Candidate path B considers revisiting a more salient area centered at pose 320 in the graph. Here we see the tradeoff illustrated throughout this paper: path A has a high risk-reward ratio. Registering to poses along path A provides greater information gain as the resulting loop-closures are larger than loops closed via path B. However, the higher visual saliency for images along path B means that registrations to these poses are more likely to occur than those along path A.

Both the previous and proposed methods select path A in this scenario. Table I presents statistics from each method related to the selection, including evaluation times. It also presents the number of proposed camera registrations for each path, the average probability of acquisition of these hypotheses, and the expected values and variances of the uncertainties predicted using the methodology from §II. We overlay these predictions on the penalty function contour plot of Fig. 6(a). Despite the higher variance and longer length, path A has a lower expected uncertainty than path B. We see why preferring path A is sensible given the Monte Carlo simulation results for traveling each path in Fig. 5(b) and (d). Only 2 trials in 1000 from path A result in uncertainties greater than the threshold and many trials outperform the resulting uncertainties of traveling path B.

We investigate a second scenario later in the mission. At pose number 1145, the robot again decides between revisiting the same locations along paths A and B (Fig. 1). This time,

TABLE I: Underwater Robot Path Predictions & Statistics

DECISION AT POSE 770	Path A	Path B
Distance [m]	21.23	20.45
Registration Hypotheses	39	37
Avg. $p(\gamma_{i,j} = 1)$	0.232	0.804
$\mathbb{E}[\mathbf{m}^\top \text{vec}(\Lambda_K^{-1})]$	0.02932	0.03047
$\text{Var}[\mathbf{m}^\top \text{vec}(\Lambda_K^{-1})]$	9.743E-07	3.791E-09
PREVIOUS METHOD[15]		
Evaluation Time [ms]	45.73	59.29
Selected Path	A	
PROPOSED METHOD		
Evaluation Time [ms]	117.75	138.66
Selected Path	A	
DECISION AT POSE 1145	Path A	Path B
Distance [m]	31.79	32.59
Registration Hypotheses	35	33
Avg. $p(\gamma_{i,j} = 1)$	0.276	0.893
$\mathbb{E}[\mathbf{m}^\top \text{vec}(\Lambda_K^{-1})]$	0.04464	0.04503
$\text{Var}[\mathbf{m}^\top \text{vec}(\Lambda_K^{-1})]$	2.400E-06	3.048E-09
PREVIOUS METHOD[15]		
Evaluation Time [ms]	90.03	100.46
Selected Path	A	
PROPOSED METHOD		
Evaluation Time [ms]	150.36	198.50
Selected Path		B

the robot is farther along in the mission and must travel farther to close loops, leading to higher predicted uncertainties than in the first scenario. Here, the previous method of [15] once again selects path A. In contrast, the proposed risk-averse method strongly prefers path B even though it predicts a higher expected uncertainty and longer traveling distance than path A. Fig. 6(b) shows how the penalty function contours change given the much closer predicted proximity to the threshold. The Monte Carlo simulations of Fig. 1(b) and (d) show why choosing path B is desirable in this case. Traveling path A results in 147 trials of 1000 that exceed the uncertainty threshold, but the robot can confidently travel path B without concern.

These results show how the power function naturally lends itself to desirable behavior in active SLAM. While the uncertainty is low, the robot is willing to make risky decisions for possible high rewards. But as the uncertainty approaches the threshold, the robot exhibits greater absolute risk aversion, making more conservative but safer decisions.

V. CONCLUSION

We proposed a risk-averse framework for Gaussian belief-space planning with stochastic measurement acquisition. We developed a planning formulation that maintains the randomness of the measurement acquisition variables and showed that this resulted in a random belief covariance matrix. We leveraged this randomness in the belief covariance to design objective functions for the planning problem that are risk-averse, inspired by utility optimization in modern portfolio theory. Our simulation results showed that risk-averse path planning for mobile robotics applications yields more desirable outcomes than paths found with previous

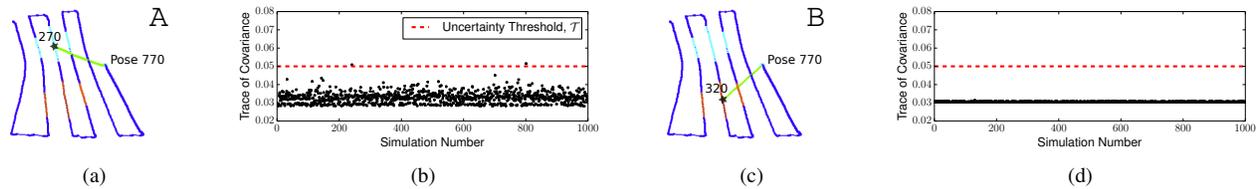


Fig. 5: Results from the first HAUV planning scenario. (a) The trajectory of revisit path A from pose 770 to pose 270 and back. (b) The trace of the final pose covariance for 1000 trials of a Monte Carlo simulation of traveling path A. (c) The trajectory of revisit path B from pose 770 to pose 320 and back. (d) The trace of the final pose covariance for 1000 trials of a Monte Carlo simulation of traveling path B.

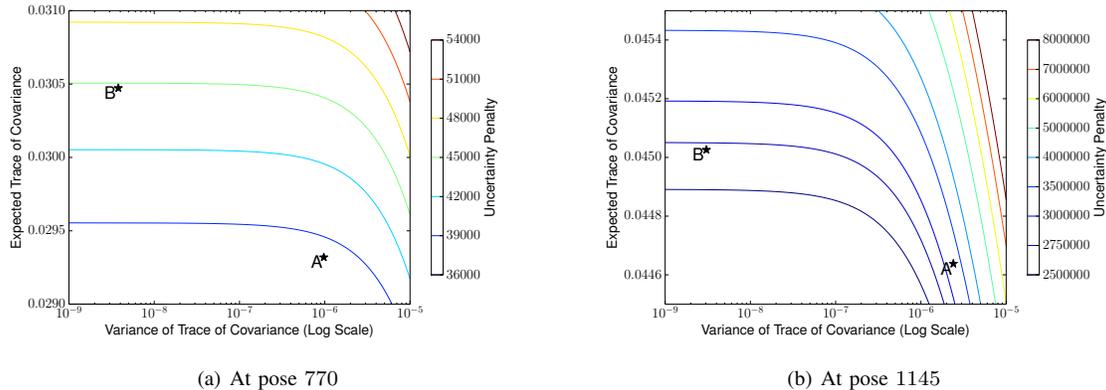


Fig. 6: The uncertainty penalty function contours for the HAUV planning scenarios. (a) At pose 770. (b) At pose 1145.

approaches, using both trajectory-smoothing and sampling-based frameworks.

REFERENCES

- [1] F. Dellaert and M. Kaess, "Square root SAM: Simultaneous localization and mapping via square root information smoothing," *Int. J. Robot. Res.*, vol. 25, no. 12, pp. 1181–1203, 2006.
- [2] V. Indelman, L. Carlone, and F. Dellaert, "Planning in the continuous domain: A generalized belief space approach for autonomous navigation in unknown environments," *Int. J. Robot. Res.*, vol. 34, no. 7, pp. 849–882, 2015.
- [3] J. van den Berg, S. Patil, and R. Alterovitz, "Motion planning under uncertainty using iterative local optimization in belief space," *Int. J. Robot. Res.*, vol. 31, no. 11, pp. 1263–1278, 2012.
- [4] A. Kim and R. M. Eustice, "Real-time visual SLAM for autonomous underwater hull inspection using visual saliency," *IEEE Trans. Robot.*, vol. 29, no. 3, pp. 719–733, 2013.
- [5] H. H. Gonzalez-Banos and J.-C. Latombe, "Navigation strategies for exploring indoor environments," *Int. J. Robot. Res.*, vol. 21, no. 10–11, pp. 829–848, 2002.
- [6] R. Sim and N. Roy, "Global A-optimal robot exploration in SLAM," in *Proc. IEEE Int. Conf. Robot. and Automation*, Barcelona, Spain, Apr. 2005, pp. 661–666.
- [7] C. Stachniss, G. Grisetti, and W. Burgard, "Information gain-based exploration using Rao-Blackwellized particle filters," in *Proc. Robot.: Sci. & Syst. Conf.*, Cambridge, MA, USA, Jun. 2005.
- [8] S. Prentice and N. Roy, "The belief roadmap: Efficient planning in belief space by factoring the covariance," *Int. J. Robot. Res.*, vol. 28, no. 11–12, pp. 1448–1465, 2009.
- [9] A. Bry and N. Roy, "Rapidly-exploring random belief trees for motion planning under uncertainty," in *Proc. IEEE Int. Conf. Robot. and Automation*, Shanghai, China, May 2011, pp. 723–730.
- [10] J. van den Berg, P. Abbeel, and K. Goldberg, "LQG-MP: Optimized path planning for robots with motion uncertainty and imperfect state information," *Int. J. Robot. Res.*, vol. 30, no. 7, pp. 895–913, 2011.
- [11] S. Patil, G. Kahn, M. Laskey, J. Schulman, K. Goldberg, and P. Abbeel, "Scaling up gaussian belief space planning through covariance-free trajectory optimization and automatic differentiation," in *Proc. Int. Work. Algorithmic Foundations of Robot.*, Istanbul, Turkey, Aug. 2014.
- [12] G. A. Hollinger and G. S. Sukhatme, "Sampling-based robotic information gathering algorithms," *Int. J. Robot. Res.*, vol. 33, no. 9, pp. 1271–1287, 2014.
- [13] R. Valencia, J. Miro, G. Dissanayake, and J. Andrade-Cetto, "Active pose SLAM," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, Vilamoura, Portugal, Oct. 2012, pp. 1885–1891.
- [14] A. Kim and R. M. Eustice, "Active visual SLAM for robotic area coverage: Theory and experiment," *Int. J. Robot. Res.*, vol. 34, no. 4–5, pp. 457–475, 2015.
- [15] S. M. Chaves, A. Kim, and R. M. Eustice, "Opportunistic sampling-based planning for active visual SLAM," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, Chicago, IL, USA, Sep. 2014, pp. 3073–3080.
- [16] R. Platt, R. Tedrake, L. Kaelbling, and T. Lozano-Perez, "Belief space planning assuming maximum likelihood observations," in *Proc. Robot.: Sci. & Syst. Conf.*, Zaragoza, Spain, Jun. 2010, pp. 587–593.
- [17] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, 2004.
- [18] J. C. Francis and D. Kim, *Modern Portfolio Theory: Foundations, Analysis, and New Developments*. Hoboken, NJ: John Wiley & Sons, 2013.
- [19] H. Markowitz, "Portfolio selection," *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.
- [20] R. C. Merton, *Continuous-Time Finance*. Cambridge, MA: Wiley-Blackwell, 1992.
- [21] S. Koenig and R. G. Simmons, "Risk-sensitive planning with probabilistic decision graphs," in *Proc. Int. Conf. Princ. Knowledge Representation and Reas.*, Bonn, Germany, May 1994, pp. 363–373.
- [22] T. M. Moldovan and P. Abbeel, "Risk aversion in Markov decision processes via near-optimal Chernoff bounds," in *Proc. Advances Neural Inform. Process. Syst. Conf.*, Lake Tahoe, NV, USA, Dec. 2012, pp. 3140–3148.
- [23] R. A. Howard and J. E. Matheson, "Risk-sensitive Markov decision processes," *Management Science*, vol. 18, no. 7, pp. 356–369, 1972.
- [24] M. J. Sobel, "The variance of discounted Markov decision processes," *J. Applied Prob.*, vol. 19, no. 4, pp. 794–802, 1982.
- [25] J. A. Filar, L. C. M. Kallenberg, and H.-M. Lee, "Variance-penalized Markov decision processes," *Mathematics of Operations Research*, vol. 14, no. 1, pp. 147–161, 1989.
- [26] S. Mannor, D. Simester, P. Sun, and J. N. Tsitsiklis, "Bias and variance approximation in value function estimates," *Management Science*, vol. 53, no. 2, pp. 308–322, 2007.
- [27] M. M. Ford, J. Pineau, and P. Sun, "A variance analysis for POMDP policy evaluation," in *Proc. AAAI Nat. Conf. Artif. Intell.*, Chicago, IL, USA, Jul. 2008, pp. 1056–1061.