Stochastic Optimal Control based on Value-Function Approximation using Sinc Interpolation

Florian Weissel [*] **Marco F. Huber** [*] **Dietrich Brunn** [*]
Uwe D. Hanebeck [*]

[*] *Intelligent Sensor-Actuator-Systems Laboratory,*
Institute of Computer Science and Engineering,
Universität Karlsruhe (TH), Germany;
e-mail: {weissel | marco.huber | uwe.hanebeck}@ieee.org,
brunn@ira.uka.de.

Abstract: An efficient approach for solving stochastic optimal control problems is to employ dynamic programming (DP). For continuous-valued nonlinear systems, the corresponding DP recursion generally cannot be solved in closed form. Thus, a typical approach is to discretize the DP value functions in order to be able to carry out the calculation. Especially for multidimensional systems, either a large number of discretization points is necessary or the quality of approximation degrades. This problem can be alleviated by interpolating the discretized value function. In this paper, we present an approach based on optimal low-pass interpolation employing sinc functions (sine cardinal). For the important case of systems with Gaussian mixture noise (including the special case of Gaussian noise), we show how the calculations required for this approach, especially the nontrivial calculation of an expected value of a Gaussian mixture random variable transformed by a sinc function, can be carried out analytically. We illustrate the effectiveness of the proposed interpolation scheme by an example from the field of Stochastic Nonlinear Model Predictive Control (SNMPC).

1. INTRODUCTION

For high-quality control of nonlinear heavily noise affected systems, the system and measurement noise has to be explicitly considered and incorporated in the control. Stochastic optimal control aims at finding a sequence of optimal control input policies under consideration of the noise influence for a given system. Here, the case of discrete-time systems with continuous state spaces and discrete control inputs is of special importance. A system of this kind may be a robot, whose state comprises its pose and the discrete controls are steering actions.

Stochastic optimal control of this kind forms the basis for the important field of Stochastic Nonlinear Model Predictive Control (Weissel et al. (2007a), Weissel et al. (2007b), Deisenroth et al. (2007)) and the closely related field of Markov decision processes with continuous-valued state spaces (Marecki et al. (2007)). Here, the future system behavior for a given control sequence is evaluated based on a set of reward functions. The optimal control sequence that gives the maximum expected reward is determined and then applied to the system.

The solution to a closed-loop feedback stochastic optimal control problem can be calculated efficiently for systems with discrete state spaces by employing dynamic programming (DP). Here, the necessary value function is calculated recursively employing the so-called Bellman equation (Bellman (1957)). Its recursive structure leads to a computational demand that increases just linearly with the length of the horizon.

In case of nonlinear systems with *continuous state space*, the recursive value function calculation cannot be carried out in closed form. Thus, some kind of value function approximation is inevitable (Bertsekas and Tsitsiklis (1996)). A common approach is to discretize the value function at a set of certain supporting points and then to use interpolation for obtaining an approximate solution of the value function in-between. By using interpolation, a significantly lower number of discretization points is necessary to gain high quality results.

For interpolation, numerous approaches are available. The most common one is by piecewise defined lower-degree polynomials, i.e., splines, where the most simple one is based on linear interpolation, which suffers from low accuracy (Philbrick, Jr. and Kitanidis (2001)). Using quadratic (Schumaker (1983)) or cubic (Johnson et al. (1993), Deisenroth et al. (2007)) splines is also very common. All these approaches are based on the arbitrary assumption that the value function behaves according to a polynomial of employed order. Another approach is based on the interpolation of the supporting points by Gaussians as described in (Nikovski and Brand (2003)), which is computational very demanding due to an inversion of a large matrix.

In this paper, we propose to use an ideal low-pass reconstruction (cardinal hold) for interpolation on a regular grid. This ideal low-pass reconstruction, which is well-known in signal processing theory, gives the interpolation with least oscillation possible and provides an upper bound for the signal frequency of the interpolated signal. Further-

more, it is the optimal reconstruction in case the sample distance is sufficiently small and no aliasing occurs, which is a necessary assumption for any discretization. For the important case of Gaussian mixture noise, which allows approximate modeling of any noise density shape, we present an analytic solution to the Bellman recursion employing discretization and ideal low-pass reconstruction.

The remainder of this paper is structured as follows: In the next section, the considered stochastic optimal control problem, which leads to the discretization, is described. In Section 3, the applied discretization and the subsequent ideal low-pass interpolation is introduced for a one-dimensional system. The main contribution of the paper, an analytical solution to the expected value of a Gaussian random variable transformed by a sinc function, is presented in Section 4. Then, in Section 5, the solution for one-dimensional systems is extended to the multidimensional case for a regular grid and Gaussian mixture noise that is comprised of axis-aligned components. In Section 6, an example from the field of Stochastic Nonlinear Model Predictive Control is used to illustrate the effectiveness of the proposed interpolation. The paper closes with conclusions.

2. CONSIDERED PROBLEM

In this paper, discrete-time dynamic systems with continuous state space $\underline{x}_k$ and discrete control inputs $\underline{u}_k$ of the form

$$\underline{x}_{k+1} = \underline{a}_k(\underline{x}_k, \underline{u}_k) + \underline{w}_k \qquad (1)$$

are considered, where $\underline{x}_k$ denotes the vector-valued random variable of the system state and $\underline{a}_k(\cdot)$ a nonlinear, possibly time-variant function. $\underline{w}_k$ denotes the additive white noise. In the following, the important case of Gaussian mixture noise of the form

$$\underline{w}_k \sim f_k^w(\underline{w}_k) = \sum_{h=1}^{H} \omega_k^{(h)} \cdot \mathcal{N}\left(\underline{x}_k - \underline{\mu}_k^{(h)}; \mathbf{C}_k^{(h)}\right) ,$$

is considered, where $\mathcal{N}\left(\underline{x} - \underline{\mu}; \mathbf{C}\right)$ denotes a Gaussian density with mean $\underline{\mu}$ and covariance matrix $\mathbf{C}$. Employing Gaussian mixtures to represent the system noise, arbitrary noise distributions can be realized due to the Gaussian mixtures' universal approximation property (Maz'ya and Schmidt (1996)), where Gaussian noise is a special case with $H = 1$ and $\omega^{(h)} = 1$.

For the system (1), an optimal control policy sequence $\pi_0^* = \{\underline{u}_0^*(\underline{x}_0), \underline{u}_1^*(\underline{x}_1), \ldots, \underline{u}_{N-1}^*(\underline{x}_{N-1})\}$ for a N-step horizon needs to be found, where the optimality is defined based on a cumulative reward

$$g_N(\underline{x}_N) + \sum_{k=0}^{N-1} g_k(\underline{x}_k, \underline{u}_k) ,$$

which comprises the step reward $g_k(\underline{x}_k, \underline{u}_k)$, depending on the predicted system states $\underline{x}_k$ and the corresponding control inputs $\underline{u}_k$, as well as a terminal reward $g_N(\underline{x}_N)$. Based on this reward function, the value function

$$J_k(\underline{x}_k) = \max_{\pi_k} \mathrm{E}_{\underline{x}_{k+1:N}} \left\{ g_N(\underline{x}_N) + \sum_{n=k}^{N-1} g_n(\underline{x}_n, \underline{u}_n) \right\} ,$$

representing the maximum expected future reward can be defined. Here, the expected value is employed, as only an uncertain prediction of the system state within the horizon is available.

If the value function is available, the optimal control sequence can then be found directly by calculating the maximizing argument according to

$$\underline{u}_k^*(\underline{x}_k) = \arg \max_{\underline{u}_k} J_{k+1}(\underline{x}_{k+1}) ,$$

with $\underline{x}_{k+1} \sim f_k^w(\underline{x}_{k+1} - \underline{a}(\underline{x}_k, \underline{u}_k))$ for a given $\underline{x}_k$.

The value function can be calculated efficiently by employing dynamic programming (DP), a recursive backward calculation that exploits Bellman's principle of optimality (Bellman (1957)) according to

$$J_N(\underline{x}_N) = g_N(\underline{x}_N) ,$$
$$J_k(\underline{x}_k) = \max_{\underline{u}_k} \left\{ g_k(\underline{x}_k, \underline{u}_k) + \mathrm{E}_{\underline{x}_{k+1}} \left\{ J_{k+1}(\underline{x}_{k+1}) | \underline{x}_k \right\} \right\} ,$$
$$k = N - 1, \ldots, 0 , \qquad (2)$$

which is possible since the value function only consists of additive terms (Bellman (1957), Bertsekas (2000)). Using DP, the computational demand increases just linearly with the length of the horizon due to the recursive structure of the calculation.

To employ DP, the recursive backward calculation in (2) needs to be evaluated, which is typically not possible for nonlinear systems with continuous state space. A common approach to overcome this problem is to discretize the value function $J_k(\boldsymbol{x}_k)$. Then, the maximization with respect to $\boldsymbol{u}_k$ as well as the calculation of the conditional expected value $\mathrm{E}_{\underline{x}_{k+1}} \left\{ J_{k+1}(\underline{x}_{k+1}) | \underline{x}_k \right\}$ only need to be performed at discrete points, which allows closed-form calculations as illustrated in the next section.

3. VALUE FUNCTION DISCRETIZATION AND INTERPOLATION

For clarity and brevity of presentation, the discretization and interpolation of a one-dimensional system is described in the following. The straightforward extension to M-dimensional systems is described in Section 5.

3.1 Discretization

An approximate closed-form solution to the recursion in (2) can be found if the value function $J_k(\boldsymbol{x}_k)$ is calculated for a fixed number of equally spaced discretization points $\mu_k^{(i)} = \mu_k^{(1)} + \alpha \cdot (i - 1)$ for $i = 1, ..., L$, where α is the grid spacing. Using this discretization leads to the approximate value function $\tilde{J}_k(\boldsymbol{x}_k)$

$$J_k(\boldsymbol{x}_k) \approx \tilde{J}_k(\boldsymbol{x}_k) = \sum_{i=1}^{L} \kappa_k^{(i)} \cdot \delta(x_k - \mu_k^{(i)}) ,$$

where $\delta(\cdot)$ denotes the Dirac delta distribution and $\kappa_k^{(i)}$ are weighting coefficients.

Employing the discretized value function $\tilde{J}_k(\boldsymbol{x}_k)$ respectively $\tilde{J}_{k+1}(\boldsymbol{x}_{k+1})$ in (2) allows closed-form calculation of

$$\kappa_k^{(i)} = \max_{u_k} \left\{ g_k(\mu_k^{(i)}, u_k) + \mathrm{E}_{x_{k+1}} \left\{ \tilde{J}_{k+1}(\boldsymbol{x}_{k+1}) | \mu_k^{(i)} \right\} \right\} ,$$
$$(3)$$

with

$$\mathrm{E}_{x_{k+1}}\left\{\tilde{J}_{k+1}(\boldsymbol{x}_{k+1})|\mu_k^{(i)}\right\}$$

$$= \int_{\mathbb{R}^M} f_k^w(x_{k+1} - a_k(\mu_k^{(i)}, u_k)) \cdot \tilde{J}_{k+1}(x_{k+1})\,\mathrm{d}x_{k+1}$$

$$= \int_{\mathbb{R}^M} \sum_{h=1}^{H} \omega_k^{(h)} \cdot \mathcal{N}\left(x_{k+1} - a_k(\mu_k^{(i)}, u_k) - \mu_k^{(h)}; \mathbf{C}_k^{(h)}\right)$$

$$\cdot \sum_{j=1}^{L} \kappa_{k+1}^{(j)} \cdot \delta(x_{k+1} - \mu_{k+1}^{(j)})\,\mathrm{d}x_{k+1} \ ,$$

where the maximization in (3) can be carried out element by element. Using a discretized value function gives a good approximation of the true value function if the grid is sufficiently fine, i.e., the grid spacing α is sufficiently small, for sampling the noise density $f_{k+1}^w(\cdot)$ with high precision. If this is not the case, the quality of the approximation degrades rapidly, which is illustrated exemplarily in the simulations in Section 6.

3.2 Interpolation

A common approach to overcome this problem is to use an interpolated version of the approximate value function $\tilde{J}_{k+1}(\cdot)$ instead of using it directly. This allows a substantially wider spacing of the sample points, which is highly desirable especially for multidimensional systems, as the number of necessary sample points depends exponentially on the dimension of the state space M, while the computational demand for interpolation just depends linearly on the dimension M.

In case of an equally spaced grid, the ideal low-pass reconstruction of a sampled signal can be obtained if the signal is reconstructed with sinc functions (sine cardinal), whose roots are located at the other sampling points as illustrated in Fig. 1. This interpolation is the well-known Whittaker-Shannon interpolation (Zayed (1993)). The sinc functions used for interpolation are of the form

$$\mathrm{sinc}(x/\alpha) = \begin{cases} 1 & \text{for } x = 0 \\ \dfrac{\sin(\pi x/\alpha)}{\pi x/\alpha} & \text{otherwise} . \end{cases} \quad (4)$$

Using the sinc function leads to an interpolation with least oscillation, where the signal frequency of the interpolated signal has an upper bound of

$$\frac{1}{2 \cdot \alpha} \ ,$$

which can be seen directly by calculating its Fourier transform (6). Using the sinc-interpolated function $\hat{J}_{k+1}$ for calculation leads to

$$\mathrm{E}_{x_{k+1}}\left\{\hat{J}_{k+1}(\boldsymbol{x}_{k+1})|\mu_k^{(i)}\right\} \quad (5)$$

$$= \int_{\mathbb{R}} f_k^w(x_{k+1} - a_k(\mu_k^{(i)}, u_k)) \cdot \hat{J}_{k+1}(x_{k+1})\,\mathrm{d}x_{k+1}$$

$$= \int_{\mathbb{R}} \sum_{h=1}^{H} \omega_k^{(h)} \cdot \mathcal{N}\left(x_{k+1} - a_k(\mu_k^{(i)}, u_k) - \mu_k^{(h)}; c_k^{(h)}\right)$$

$$\cdot \sum_{j=1}^{L} \kappa_{k+1}^{(j)} \cdot \mathrm{sinc}\left(\frac{x_{k+1} - \mu_{k+1}^{(j)}}{\alpha}\right)\,\mathrm{d}x_{k+1} \ .$$
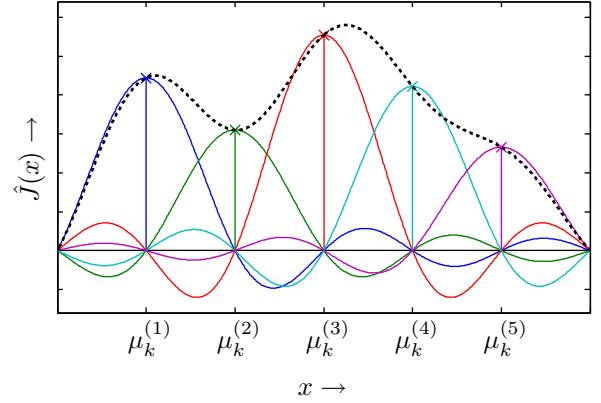
Employing the sum rule results in

Fig. 1. Interpolation for five sampling points using the ideal low-pass reconstruction. The dotted black line depicts the overall reconstruction, the colored solid lines represent the sinc functions with their maximum values at the sampling points.

$$\mathrm{E}_{x_{k+1}}\left\{\hat{J}_{k+1}(\boldsymbol{x}_{k+1})|\mu_k^{(i)}\right\}$$

$$= \sum_{h=1}^{H} \sum_{j=1}^{L} \int_{\mathbb{R}} \omega_k^{(h)} \cdot \mathcal{N}\left(x_{k+1} - a_k(\mu_k^{(i)}, u_k) - \mu_k^{(h)}; c_k^{(h)}\right)$$

$$\cdot \kappa_{k+1}^{(j)} \cdot \mathrm{sinc}\left(\frac{x_{k+1} - \mu_{k+1}^{(j)}}{\alpha}\right)\,\mathrm{d}x_{k+1} \ ,$$

which is the sum of expected values of Gaussian random variables that are transformed by sinc functions. Here, $c = \sigma^2$ denotes the variance of the Gaussians. If this expected values can be calculated in closed form, which is indeed the case as shown in the next section, efficient value function approximation using an ideal low-pass interpolation is possible.

4. CALCULATION OF THE EXPECTED VALUE

In this section, the main contribution of this paper, an analytic solution for calculating the expected value of a Gaussian random variable $\boldsymbol{x}$ with probability density function

$$f(x) = \mathcal{N}(x - \mu; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right)$$

transformed by a single sinc function (4), is presented for the one-dimensional case. In Section 5, this basic case is extended to the multidimensional case.

Generally, the expected value of a random variable transformed by a function $g(x)$ can be calculated according to

$$\mathrm{E}_{\boldsymbol{x}}\{g(\boldsymbol{x})\} = \int_{-\infty}^{\infty} f(x) \cdot g(x)\,\mathrm{d}x \ .$$

This can be solved employing Parseval's relation (Arfken and Weber (2005) p. 953)

$$\int_{-\infty}^{\infty} f(x) \cdot g^*(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} F(f) \cdot G^*(f)\,\mathrm{d}f \ ,$$

where $(\cdot)^*$ is used to denote the complex conjugate, $F(f)$ and $G(f)$ are the Fourier transforms of $f(x)$ and $g(x)$ respectively, which are given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

$$\circ\!\!-\!\!\bullet$$

$$F(f) = \mathrm{e}^{-2\pi^2\sigma^2 f^2} \cdot \mathrm{e}^{-\mathrm{j}\,2\pi f\mu} \ ,$$

(Bracewell (1999)), where $\mathrm{j} = \sqrt{-1}$ is the imaginary unit, and

$$g(x) = \mathrm{sinc}(x/\alpha)$$

$$\circ\!\!-\!\!\bullet$$

$$G(f) = \alpha \cdot \mathrm{rect}_{1/\alpha}(f) = \begin{cases} 1 & \text{for } |f| \leq \dfrac{1}{2\alpha} \\ 0 & \text{otherwise ,} \end{cases} \quad (6)$$

where $\mathrm{rect}_{1/\alpha}(\cdot)$ denotes the rectangular function with width $1/\alpha$. Employing that $g(x)$ and $G(f)$ have no imaginary part and thus $g^*(x) = g(x)$ and $G^*(f) = G(f)$ this allows calculating the desired expected value

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{x}}\{g(\boldsymbol{x})\} \\
&= \int_{-\infty}^{\infty} F(f) \cdot G(f)\,\mathrm{d}f \\
&= \alpha \int_{-\frac{1}{2\alpha}}^{\frac{1}{2\alpha}} \exp\left\{-2(\pi\sigma f)^2 - \mathrm{j}\,2\pi f\mu\right\}\mathrm{d}f \ .
\end{aligned}$$

Completing the square with $\left(\frac{\mathrm{j}\mu}{2\sigma}\right)^2$ results in

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{x}}\{g(\boldsymbol{x})\} \\
&= \alpha \int_{-\frac{1}{2\alpha}}^{\frac{1}{2\alpha}} \exp\left\{-2\left(\left(\pi\sigma f + \mathrm{j}\frac{\mu}{2\sigma}\right)^2 + \left(\frac{\mu}{2\sigma}\right)^2\right)\right\}\mathrm{d}f \ .
\end{aligned}$$

Substitution of

$$z = \sqrt{2}\pi\sigma f + \mathrm{j}\frac{\mu}{\sqrt{2}\sigma} \ ,$$

$$\mathrm{d}z = \sqrt{2}\pi\sigma\,\mathrm{d}f$$

leads to

$$\mathrm{E}_{\boldsymbol{x}}\{g(\boldsymbol{x})\} = \frac{\alpha}{\sqrt{2}\pi\sigma}\,\mathrm{e}^{-\frac{\mu^2}{2\sigma^2}} \cdot \int_{-\frac{\pi\sigma}{\sqrt{2}\alpha}+\mathrm{j}\frac{\mu}{\sqrt{2}\sigma}}^{\frac{\pi\sigma}{\sqrt{2}\alpha}+\mathrm{j}\frac{\mu}{\sqrt{2}\sigma}} \mathrm{e}^{-z^2}\,\mathrm{d}z \ . \quad (7)$$

To solve the resulting integral, which is an integral of e^{-z^2} with complex symmetric bounds $-a+\mathrm{j}b$ and $a+\mathrm{j}b$, the following lemma is necessary.

Lemma 1. The integral of e^{-z^2} with complex-valued but symmetric bounds $-a+\mathrm{j}b$ and $a+\mathrm{j}b$ is given by

$$\int_{-a+\mathrm{j}b}^{a+\mathrm{j}b} \mathrm{e}^{-z^2}\,\mathrm{d}z = \sqrt{\pi}\cdot\mathrm{Re}\left\{\mathrm{erf}(a+\mathrm{j}b)\right\} \ ,$$

where $\mathrm{Re}\{\cdot\}$ represents the real part of a complex number and $\mathrm{erf}(\cdot)$ denotes the complex Gaussian error function.

Proof.

$$\int_{-a+\mathrm{j}b}^{a+\mathrm{j}b} \mathrm{e}^{-z^2}\,\mathrm{d}z = \frac{\sqrt{\pi}}{2}\left(\mathrm{erf}(a+\mathrm{j}b) - \mathrm{erf}(-a+\mathrm{j}b)\right) \ .$$

Employing the symmetry relations $\mathrm{erf}(-z) = -\mathrm{erf}(z)$ (odd) and $\mathrm{erf}(z^*) = \mathrm{erf}^*(z)$ (Abramowitz and Stegun (1972) p. 297), leads to
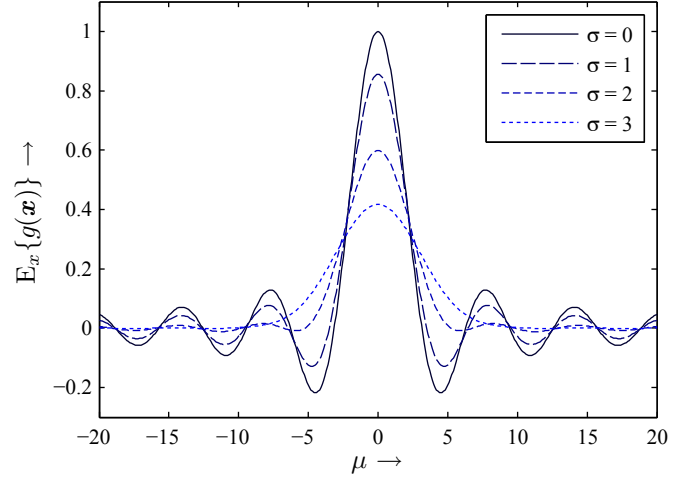

Fig. 2. Plot of the calculated expected value for different μ and σ, where $\alpha = \pi$.

$$\begin{aligned}
\int_{-a+\mathrm{j}b}^{a+\mathrm{j}b} \mathrm{e}^{-z^2}\,\mathrm{d}z &= \frac{\sqrt{\pi}}{2}\left(\mathrm{erf}(a+\mathrm{j}b) + \mathrm{erf}(a-\mathrm{j}b)\right) \\
&= \frac{\sqrt{\pi}}{2}\left(\underbrace{\mathrm{erf}(a+\mathrm{j}b)}_{=:c+\mathrm{j}\,d} + \mathrm{erf}^*(a+\mathrm{j}b)\right) \ ,
\end{aligned}$$

substituting $c + \mathrm{j}\,d := \mathrm{erf}(a+\mathrm{j}b)$ finally gives the result

$$\begin{aligned}
&= \frac{\sqrt{\pi}}{2}\left((c+\mathrm{j}\,d) + (c-\mathrm{j}\,d)\right) \\
&= \sqrt{\pi}\cdot c \\
&= \sqrt{\pi}\cdot\mathrm{Re}\left\{\mathrm{erf}(a+\mathrm{j}b)\right\} \ .
\end{aligned}$$

$\square$

The final result

$$\mathrm{E}_{\boldsymbol{x}}\{g(\boldsymbol{x})\} = \frac{\alpha}{\sqrt{2}\sigma}\,\mathrm{e}^{-\frac{\mu^2}{2\sigma^2}} \cdot \mathrm{Re}\left\{\mathrm{erf}\left(\frac{\pi\sigma}{\sqrt{2}\alpha} + \mathrm{j}\frac{\mu}{\sqrt{2}\sigma}\right)\right\}$$

is obtained using Lemma 1 in (7). In Fig. 2, a plot of the resulting function is shown for different μ and σ. It can be seen that for small σ the function approaches a sinc function, while for large σ it approaches a Gaussian. This is as expected, since for small σ the Gaussian approaches a Dirac delta distribution. Thus, the integral approaches the sinc function. For large σ, the effect is inverted, i.e., the sinc function approaches a Dirac delta distribution and thus, the integral approaches the Gaussian.

Remark 1. (Complex error function). The required calculation of the real part of the complex error function can be performed efficiently by employing the approximations given in Abramowitz and Stegun (1972), (7.1.23) and (7.1.29). In case the imaginary part $\mu/(\sqrt{2}\sigma)$ of the argument of the complex error function increases, the complex error function itself becomes increasingly large, which prohibits numerical calculation. If the Gaussian part $\exp\{-\mu^2/(2\sigma^2)\}$ is directly included in the approximation, this effect is avoided.

5. EXTENSION TO MULTIDIMENSIONAL SYSTEMS

In the previous sections, just one-dimensional systems have been considered. The extension to systems with M dimensions is straightforward in case of discretization with a regular grid and Gaussian mixture noise, which is comprised of components that are axis-aligned Gaussians. In the following, vector-valued variables $\underline{x}_k^{\mathrm{T}} = \left[x_k^{(1)}, \ldots, x_k^{(M)}\right]$ are employed. The discretization points of the regular grid

$$\kappa^{(i_{1:M})} \cdot \delta^M\left(\underline{x}_k - \underline{\mu}_k^{(i_{1:M})}\right)$$

$$= \kappa^{(i_{1:M})} \cdot \prod_{m=1}^{M} \delta\left(x_k^{(m)} - \mu_k^{(m,i_m)}\right)$$

are equally spaced in any dimension m according to $\mu_k^{(m,i_m)} = \mu_k^{(m,1)} + \alpha^m \cdot (i_m - 1)$ for $i_m = 1, ..., L^m$, where α^m is the grid spacing and L^m the number of sample points in dimension m. $\delta^M(\cdot)$ denotes the M-dimensional Dirac delta distribution.

The ideal low-pass interpolation can be obtained if M-dimensional sinc functions

$$\mathrm{sinc}(\underline{x}; \underline{\alpha}) = \prod_{m=1}^{M} \mathrm{sinc}(x^m/\alpha^m)$$

are used for reconstruction.

The probability density function of Gaussian mixture noise that is comprised of axis-aligned components is a weighted sum of products of Gaussians

$$f_k^w(\underline{w}_k) = \sum_{h=1}^{H} \omega_k^{(h)} \cdot \mathcal{N}\left(\underline{w}_k - \underline{\mu}_k^{(h)}; \mathbf{C}_k^{(h)}\right)$$

$$= \sum_{h=1}^{H} \omega_k^{(h)} \cdot \prod_{m=1}^{M} \mathcal{N}\left(w_k^{(m)} - \mu_k^{(m,h)}; c_k^{(m,h)}\right) ,$$

where the covariance matrix $\mathbf{C}_k^{(h)}$ is a diagonal matrix comprising the elements $c_k^{(m,h)}$. Using this representation, any kind of noise density, especially noise with correlated dimensions, can be realized due to the Gaussian mixtures' universal approximation property (Maz'ya and Schmidt (1996)).

Using the above definitions, the calculation of the desired expected value, which is the multidimensional extension of (5), can be carried out dimension by dimension

$$\mathrm{E}_{\underline{x}_{k+1}}\left\{\hat{J}_{k+1}(\underline{x}_{k+1})|\underline{\mu}_k^{(i_{1:M})}\right\}$$

$$= \sum_{h=1}^{H} \sum_{j=1}^{L} \omega_k^{(h)} \cdot \kappa_{k+1}^{(j)} \cdot \prod_{m=1}^{M} \mathrm{E}_{\boldsymbol{x}}\left\{g_{x_{k+1}}^{(m)}\left(x_{k+1}^{(m)}\right)\right\} ,$$

where $x_{k+1}^{(m)} \sim f_{x_{k+1}}^{(m)}\left(x_{k+1}^{(m)}\right)$.

Proof.

$$\mathrm{E}_{\underline{x}_{k+1}}\left\{\hat{J}_{k+1}(\underline{x}_{k+1})|\underline{\mu}_k^{(i_{1:M})}\right\}$$

$$= \sum_{h=1}^{H} \sum_{j=1}^{L} \int_{\mathbb{R}^M} \omega_k^{(h)} \cdot \mathcal{N}\left(\underline{x}_{k+1} - \underline{a}_k(\underline{\mu}_k^{(i_{1:M})}, \underline{u}_k) - \underline{\mu}_k^{(h)}; \mathbf{C}_k^{(h)}\right)$$

$$\cdot \kappa_{k+1}^{(j)} \cdot \mathrm{sinc}\left(\underline{x}_{k+1} - \underline{\mu}_{k+1}^{(j)}; \underline{\alpha}\right) \mathrm{d}\underline{x}_{k+1}$$
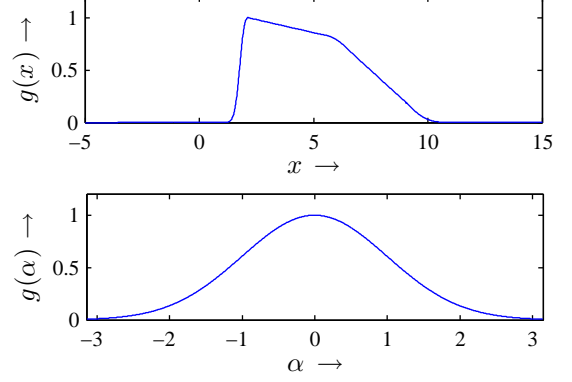

Fig. 3. x and α components of the reward function.

$$= \sum_{h=1}^{H} \sum_{j=1}^{L} \int_{\mathbb{R}^M} \omega_k^{(h)} \cdot \kappa_{k+1}^{(j)}$$

$$\cdot \prod_{m=1}^{M} \underbrace{\mathcal{N}\left(x_{k+1}^{(m)} - a_k^{(m)}(\underline{\mu}_k^{(i_{1:M})}, \underline{u}_k) - \mu_k^{(m,h)}; c_k^{(m,h)}\right)}_{f_{x_{k+1}}^{(m)}\left(x_{k+1}^{(m)}\right)}$$

$$\cdot \prod_{m=1}^{M} \underbrace{\mathrm{sinc}\left((x_{k+1}^m - \mu_{k+1}^{(m,j)})/\alpha^m\right)}_{g_{x_{k+1}}^{(m)}\left(x_{k+1}^{(m)}\right)} \mathrm{d}\underline{x}_{k+1}$$

$$= \sum_{h=1}^{H} \sum_{j=1}^{L} \omega_k^{(h)} \cdot \kappa_{k+1}^{(j)} \cdot \prod_{m=1}^{M} \mathrm{E}_{\boldsymbol{x}}\left\{g_{x_{k+1}}^{(m)}\left(x_{k+1}^{(m)}\right)\right\} .$$

$\square$

6. SIMULATIONS

To illustrate the effectiveness of the proposed interpolation scheme, we use an example from the field of Stochastic Nonlinear Model Predictive Control (SNMPC) that is inspired by the one presented in (Weissel et al. (2007b)). There, no interpolation is employed, which is, as shown in that paper, sufficient for the considered setting.

In the considered example system, a miniature walking robot is supposed to move with constant velocity along a given trajectory, e.g., along a wall. Additionally, constraints like not hitting the wall are modeled with the reward function. The robot is able to superimpose left and right turns onto the forward motion. The robot's motion can be modeled similar to the motion of a two-wheeled differential-drive robot, which leads to the nonlinear discrete-time system equation

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + s \cdot \sin(\boldsymbol{\alpha}_k) + \boldsymbol{w}_k^x ,$$
$$\boldsymbol{\alpha}_{k+1} = \boldsymbol{\alpha}_k + u_k + \boldsymbol{w}_k^\alpha ,$$

where $\underline{\boldsymbol{x}}_k = [\boldsymbol{x}_k, \boldsymbol{\alpha}_k]^{\mathrm{T}}$, with the distance to the wall $\boldsymbol{x}_k$ and its orientation relative to the wall $\boldsymbol{\alpha}_k$. The robot's constant step size is $s = 1$. The noise influence on the system is described by $\boldsymbol{w}_k^x$, $\boldsymbol{w}_k^\alpha$, where $\boldsymbol{w}_k^x$, $\boldsymbol{w}_k^\alpha$ is zero-mean white Gaussian noise with standard deviation $\sigma_w^x = 0.2$ and $\sigma_w^\alpha = 0.02 \approx 1°$, respectively. The input $u_k \in \{-0.2, \ 0, \ 0.2\}$ is the steering action, i.e., turn left, go straight, or turn right. Furthermore, the robot is equipped with sensors for measuring distance $\boldsymbol{y}_k^x = \boldsymbol{x}_k + \boldsymbol{v}_k^x$ and orientation $\boldsymbol{y}_k^\alpha = \boldsymbol{\alpha}_k + \boldsymbol{v}_k^\alpha$ with respect to the wall,

Table 1. Simulation results from
100 Monte Carlo runs each.

simulation	average reward		computation time	
sinc $L = 36$	34.49	(95.8%)	0.11 sec	(0.33%)
grid $L = 36$	5.23	(14.5%)	0.02 sec	(0.05%)
sinc $L = 100$	34.65	(96.28%)	0.54 sec	(1.54%)
grid $L = 100$	13.14	(36.51%)	0.10 sec	(0.29%)
grid $L = 1444$	34.64	(96.24%)	19.21 sec	(55.26%)
grid $L = 1936$	35.99	(100.00%)	34.77 sec	(100.00%)

where the measurements, on which the control is based, are distorted by zero-mean white Gaussian noise $\boldsymbol{v}_k^x$, $\boldsymbol{v}_k^\alpha$ with standard deviation $\sigma_v^x = 0.2$ and $\sigma_v^\alpha = 0.02 \approx 1°$, respectively.

All simulations are performed for an $N = 4$ step prediction horizon, with a reward function just comprising the terminal reward $g_N(\boldsymbol{x}_N, \boldsymbol{\alpha}_N)$ as depicted in Fig. 3.

In Table 1, the simulation results are given for six sets comprising 100 Monte Carlo runs each, where the uniformly distributed initial values are sampled from the intervals $x_0 \in [0, 6]$ and $\alpha_0 \in [-\pi/4, \pi/4]$. For all six simulation sets, the same noise realization has been employed. The given average reward is calculated based on the reward function according to Fig. 3 for a 40 step SNMPC run. The considered state space is restricted to $x \in [-5, 15], \alpha \in [-\pi, \pi]$. Computation times are given for the calculation of the DP recursion (2) implemented in MATLAB. The simulation employing a grid with $L = 44 \times 44 = 1936$ components is used as a ground truth, an increased number of components did not increase the quality of approximation noticeably.

The results show that the sinc interpolation with $L = 6 \times 6 = 36$ already leads to a good approximation (an average of 95.8% of the ground truth reward requireing just 0.33% of the computation time), where the grid approximation leads to far inferior results (just an average of 14.5% of the original reward). If the number of samples is increased to $L = 10 \times 10 = 100$, the computational demand employing the grid approximation is similar to the case employing the sinc interpolation with $L = 36$, but leads to a far inferior approximation quality (with an average reward of 36% compared to 95.8%). Even a grid with $L = 38 \times 38 = 1444$ leads to an approximation quality slightly better than using the sinc interpolation with $L = 6 \times 6 = 36$, but requires over 168 times the computation time.

7. CONCLUSIONS

In this paper, the analytic solution of the expected value of a Gaussian random variable that is transformed by a sinc function has been presented. This permits approximate closed-form calculation of the Bellman recursion, which is an efficient approach to solve stochastic optimal control problems. Additionally, a straightforward extension to multidimensional Gaussian mixture random variables has been derived and applied to an example from the field of Stochastic Nonlinear Model Predictive Control. Based on this example, the effectiveness of the proposed optimal low-pass interpolation scheme has been demonstrated by Monte Carlo simulations.

REFERENCES

Milton Abramowitz and Irene A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover Publications, 10th edition, December 1972.

George Arfken and Hand J. Weber. *Mathematical Methods for Physicists.* Elsevier Academic Press, 6th edition, 2005.

Richard E. Bellman. *Dynamic Programming.* Princeton University Press, Princeton, New Jersey, 1957.

Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, Belmont, Massachusetts, 2nd edition, 2000.

Dimitri P. Bertsekas and John N. Tsitsiklis. *Neuro-Dynamic Programming.* Athena Scientific, Belmont, Massachusetts, 1996.

Ronald N. Bracewell. *The Fourier Transform and its Applications.* McGraw-Hill, 3rd edition, 1999.

Marc P. Deisenroth, Florian Weissel, Toshiyuki Ohtsuka, and Uwe D. Hanebeck. Online-Computation Approach to Optimal Control of Noise-Affected Nonlinear Systems with Continuous State and Control Spaces. In *Proceedings of the European Control Conference (ECC 2007)*, Kos, Greece, July 2007.

Sharon A. Johnson, Jery R. Stedinger, Christine A. Shoemaker, Ying Li, and Jose Alberto Tejada-Guibert. Numerical Solution of Continuous-State Dynamic Programs Using Linear and Spline Interpolation. *Operations Research*, 41(3):484–500, May–June 1993.

Janusz Marecki, Sven Koenig, and Milind Tambe. A Fast Analytical Algorithm for Solving Markov Decision Processes with Real-Valued Resources. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI-07)*, Hyderabad, India, January 2007.

Vladimir Maz'ya and Gunther Schmidt. On approximate approximations using Gaussian kernels. *IMA Journal of Numerical Analysis*, 16(1):13–29, 1996.

Daniel Nikovski and Matthew Brand. Non-Linear Stochastic Control in Continuous State Spaces by Exact Integration in Bellman's Equations. In *Proceedings of the 2003 International Conference on Automated Planning and Scheduling*, pages 91–95, June 2003.

C. Russell Philbrick, Jr. and Peter K. Kitanidis. Improved Dynamic Programming Methods for Optimal Control of Lumped-Parameter Stochastic Systems. *Operations Research*, 49(3):398–412, May–June 2001.

Larry L. Schumaker. On Shape-Preserving Quadratic Spline Interpolation. *SIAM Journal on Numerical Analysis*, 20(4):854–864, August 1983.

Florian Weissel, Marco F. Huber, and Uwe D. Hanebeck. Efficient Control of Nonlinear Noise–Corrupted Systems Using a Novel Model Predictive Control Framework. In *Proceedings of the 2007 American Control Conference (ACC 2007)*, New York City, New York, July 2007a.

Florian Weissel, Marco F. Huber, and Uwe D. Hanebeck. A Nonlinear Model Predictive Control Framework Approximating Noise Corrupted Systems with Hybrid Transition Densities. In *IEEE Conference on Decision and Control (CDC 2007)*, New Orleans, Louisiana, December 2007b.

Ahmed I. Zayed. *Advances in Shannon's Sampling Theory.* CRC Press, Boca Raton, 1993.