

# Applying Gaussian distributed constraints to Gaussian distributed variables

Andrew W. Palmer, Andrew J. Hill, Steven J. Scheduling\*

*Australian Centre for Field Robotics, The University of Sydney, NSW, Australia*

---

## Abstract

This paper develops an analytical method of truncating inequality constrained Gaussian distributed variables where the constraints are themselves described by Gaussian distributions. Existing truncation methods either assume hard constraints, or use numerical methods to handle uncertain constraints. The proposed approach introduces moment-based Gaussian approximations of the truncated distribution. This method can be applied to numerous problems, with the motivating problem being Kalman filtering with uncertain constraints. In a simulation example, the developed method is shown to outperform unconstrained Kalman filtering by over 40% and hard-constrained Kalman filtering by over 17%.<sup>1</sup>

*Keywords:* Constrained Kalman filter, uncertain constraints

---

## 1. Introduction

Gaussian distributed variables are widely used to represent the state of a system in many problems ranging from state estimation [1] to scheduling [2, 3]. In practice, the state vectors in many systems are known to satisfy inequality constraints. Examples of state-constrained systems include health monitoring [4], vision systems [5], robotics [6], binary sensor networks [7], and object tracking [8]. This paper deals specifically with systems that are subject to inequality constraints where the constraints themselves have uncertainty described by Gaussian distributions. Constraints described by

---

\*Email address: {a.palmer;a.hill;s.scheduling}@acfr.usyd.edu.au

<sup>1</sup>©2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

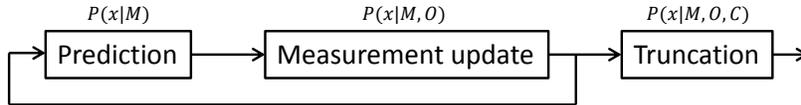


Figure 1: The Kalman filter is run independent of the truncation method, with the truncation being applied to the state estimate that is the output of the Kalman filter. The prediction step of the Kalman filter results in a probability distribution describing the state,  $x$ , conditioned on the system model,  $M$ . The measurement update step further conditions the state estimate on the observations,  $O$ . Finally, the truncation step conditions the estimate on the constraints acting on the state,  $C$ .

Gaussian distributions can arise from many sources in state estimation problems including discrete sensors, such as position or level switches, that have uncertainty on their activation point, obstacles whose positions are uncertain, and other physical and model-derived bounds such as maximum fuel levels based on historical fuel burn rates. Constrained Gaussian distributed variables also appear in scheduling applications where the distribution describing the time at which an event is predicted to occur is constrained by the time distributions of other events.

Hard inequality constraints are well studied [1], where the main approaches are estimate projection [4], gain projection [9], and Probability Density Function (PDF) truncation [10]. Estimate and gain projection approaches incorporate the constraints into the derivation of the Kalman filter, resulting in a constrained optimisation problem that can be solved using quadratic programming, least squares approaches, amongst others [1, 11]. Truncation methods, on the other hand, are applied directly to the PDF resulting from a Kalman filter, as outlined in Figure 1. This approach truncates the PDF at the constraints and calculates the mean and covariance of the truncated PDF, which become the constrained state estimate and its covariance. The PDF truncation approach was shown in [10] to, in general, outperform the estimate projection method. The truncation approach has been applied to probabilistic collision checking for robots [12], and has been extended to non-linear systems [13, 14].

Soft constraints correspond to uncertain or noisy constraints, and are less studied than hard constraints. Soft equality constraints are typically incorporated as noisy measurements [1, 15]. However, soft inequality constraints

are significantly more difficult to deal with, and numerical filters such as a Particle Filter (PF) are typically used for these problems [16]. Several numerical methods have been examined for incorporating soft constraints into the Kalman filter. A numerical PDF truncation method was used in [6] for robot localisation using Radio Frequency IDentification (RFID) tags, where the noise on the inequality constraints was highly non-Gaussian. Compared with a PF approach, the numerical PDF truncation method was 2 to 3 orders of magnitude faster while, in general, providing similar results. A similar RFID problem was examined in [7] where aspects of the Unscented Kalman Filter (UKF) and PF were combined—the prediction step used the standard UKF step, while the correction step was modified to weight the sigma-points of the UKF in a similar manner to the weighting process in a PF. It was shown to outperform a PF as well as the Quantised Extended Kalman Filter (QEKF) presented in [17].

The literature on soft inequality constraints has focused on constraints with non-Gaussian distributions, where the constrained state estimates are, by necessity, calculated using numerical methods. The main contribution of this paper is an analytical method for PDF truncation with soft constraints where the soft constraints are described by Gaussian distributions. This reduces the computational requirement compared to numerical methods, and it is shown to provide superior estimation performance compared to unconstrained and hard-constrained state estimation methods. The truncation approach presented in this paper is not limited to Kalman filters and can be applied to any constrained system using Gaussian distributions to represent the state and constraints.

The rest of this paper is structured as follows: Section 2 introduces the constrained Kalman filtering problem, Section 3 shows how the state and constraints can be transformed such that each state has only one constraint acting on it, Section 4 presents the truncation method for a one-sided constraint, and Section 5 extends this to an interval constraint. The performance of the methods are evaluated in Section 6, and the paper is concluded in Section 7. Appendix A and Appendix B provide in-depth derivations of the integrals used in this paper.

## 2. Problem definition

This paper adapts the notation used in [10]. A discrete linear time-invariant system is described by:

$$\begin{aligned}\mathbf{x}(k) &= \mathbf{F}\mathbf{x}(k-1) + \mathbf{G}\mathbf{u}(k) + \mathbf{w}(k) \\ \mathbf{y}(k) &= \mathbf{H}\mathbf{x}(k) + \mathbf{v}(k)\end{aligned}\tag{1}$$

where  $k$  is the time index,  $\mathbf{x}$  is the state vector with  $n$  states,  $\mathbf{u}$  is the vector of known control inputs, and  $\mathbf{y}$  is the vector of measurements. The vectors  $\mathbf{w}$  and  $\mathbf{v}$  contain the process and measurement noise respectively. The process noise,  $\mathbf{w}$ , is assumed to be zero mean Gaussian white noise with a covariance matrix of  $\mathbf{Q}$ . The measurement noise,  $\mathbf{v}$ , is similarly assumed to be zero mean Gaussian white noise with a covariance matrix of  $\mathbf{R}$ . The noises at each time-step are assumed to be independent.

For the given system, the Kalman filter prediction equations are [18]:

$$\begin{aligned}\hat{\mathbf{x}}(k|k-1) &= \mathbf{F}\hat{\mathbf{x}}(k-1|k-1) + \mathbf{G}\mathbf{u}(k-1) \\ \mathbf{P}(k|k-1) &= \mathbf{F}\mathbf{P}(k-1|k-1)\mathbf{F}^T + \mathbf{Q}\end{aligned}\tag{2}$$

and the measurement update equations are:

$$\begin{aligned}\mathbf{K} &= \mathbf{P}(k|k-1)\mathbf{H}^T (\mathbf{H}\mathbf{P}(k|k-1)\mathbf{H}^T + \mathbf{R})^{-1} \\ \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(\mathbf{y}(k) - \mathbf{H}\hat{\mathbf{x}}(k|k-1)) \\ \mathbf{P}(k|k) &= \mathbf{P}(k|k-1) - \mathbf{K}\mathbf{H}\mathbf{P}(k|k-1)\end{aligned}\tag{3}$$

where  $\hat{\mathbf{x}}(k|k)$  is the state estimate, and  $\mathbf{P}(k|k)$  is the covariance of the state estimate. The state estimate is initialised with  $\hat{\mathbf{x}}(0) = E[\mathbf{x}(0)]$ , where  $E[\cdot]$  is the expectation operator. The covariance matrix is initialised with  $\mathbf{P}(0) = E[(\mathbf{x}(0) - \hat{\mathbf{x}}(0))(\mathbf{x}(0) - \hat{\mathbf{x}}(0))^T]$ .

Now consider the following  $s$  linearly independent constraints on the system:

$$A_m(k) \leq \phi_m^T(k)\mathbf{x}(k) \leq B_m(k) \quad m = 1, \dots, s\tag{4}$$

where the constraints are uncertain and normally distributed:

$$A_m(k) \sim \mathcal{N}(\mu_{a,m}, \sigma_{a,m}^2) \quad B_m(k) \sim \mathcal{N}(\mu_{b,m}, \sigma_{b,m}^2)\tag{5}$$

Equation (4) describes a two-sided constraint on the linear function of the state described by  $\phi_m^T(k)\mathbf{x}(k)$ . One sided constraints can be represented by

setting  $\mu_{a,m} = -\infty$ , or  $\mu_{b,m} = \infty$ , and hard constraints can be implemented by setting  $\sigma_{a,m} \approx 0$  or  $\sigma_{b,m} \approx 0$  as required.

Given an estimate  $\hat{\mathbf{x}}(k)$  with covariance  $\mathbf{P}(k)$  at time  $k$ , the problem is to truncate the Gaussian PDF  $\mathcal{N}(\hat{\mathbf{x}}(k), \mathbf{P}(k))$  using the  $s$  constraints described above, and then find the mean  $\tilde{\mathbf{x}}(k)$  and covariance  $\tilde{\mathbf{P}}(k)$  of the truncated PDF. The calculated mean and covariance represent the constrained estimate of the state.

### 3. Transforming the state vector and constraints

To apply the constraints via the truncation method, the state vector must be transformed so that the constraints are decoupled. This will result in  $s$  transformed constraints that each involve only one element of the transformed state, allowing the constraints to be enforced individually on each element of the transformed state. It should be noted that the order in which constraints are applied can change the final state estimate. However, if the initial constraints are decoupled, the order of constraint application does not change the result [10].

The transformation process is outlined in [1] and [10], and is summarised here in equations (6)–(12) and (24)–(26). For ease of notation, the  $(k)$  after each variable will be dropped. Let the vector  $\tilde{\mathbf{x}}_i$  be the truncated state estimate, and the matrix  $\tilde{\mathbf{P}}_i$  be the covariance of  $\tilde{\mathbf{x}}_i$ , after the first  $i - 1$  constraints have been enforced. To initialise the process:

$$i = 1 \quad \tilde{\mathbf{x}}_i = \hat{\mathbf{x}} \quad \tilde{\mathbf{P}}_i = \mathbf{P} \quad (6)$$

The transformed state vector is given by:

$$\mathbf{z}_i = \boldsymbol{\rho}_i \mathbf{W}_i^{-1/2} \mathbf{T}_i^T (\mathbf{x} - \tilde{\mathbf{x}}_i) \quad (7)$$

where the matrices  $\mathbf{T}_i$  and  $\mathbf{W}_i$  are derived from the Jordan canonical decomposition of  $\tilde{\mathbf{P}}_i$ :

$$\mathbf{T}_i \mathbf{W}_i \mathbf{T}_i^T = \tilde{\mathbf{P}}_i \quad (8)$$

$\mathbf{T}_i$  is an orthogonal matrix, and  $\mathbf{W}_i$  is a diagonal matrix. The matrix  $\boldsymbol{\rho}_i$  is derived by the Gram–Schmidt orthogonalisation [19] which finds the orthogonal  $\boldsymbol{\rho}_i$  that satisfies:

$$\boldsymbol{\rho}_i \mathbf{W}_i^{1/2} \mathbf{T}_i^T \boldsymbol{\phi}_i = \left[ (\boldsymbol{\phi}_i^T \tilde{\mathbf{P}}_i \boldsymbol{\phi}_i)^{1/2} \quad 0 \quad \dots \quad 0 \right]^T \quad (9)$$

Now only one element of  $\mathbf{z}_i$  is constrained, and the states in the transformed state vector  $\mathbf{z}_i$  are independent standard normal distributions. Let  $\mathbf{e}_i$  be the  $i$ th column of an  $n \times n$  identity matrix. Transforming the constraints results in:

$$C_i \leq \mathbf{e}_i^T \mathbf{z}_i \leq D_i \quad (10)$$

where

$$\begin{aligned} C_i &\sim \mathcal{N}(\mu_{c,i}, \sigma_{c,i}^2) \\ \mu_{c,i} &= \frac{\mu_{a,i} - \boldsymbol{\phi}_i^T \tilde{\mathbf{x}}_i}{\sqrt{\boldsymbol{\phi}_i^T \tilde{\mathbf{P}}_i \boldsymbol{\phi}_i}} \quad \sigma_{c,i} = \frac{\sigma_{a,i}}{\sqrt{\boldsymbol{\phi}_i^T \tilde{\mathbf{P}}_i \boldsymbol{\phi}_i}} \end{aligned} \quad (11)$$

and

$$\begin{aligned} D_i &\sim \mathcal{N}(\mu_{d,i}, \sigma_{d,i}^2) \\ \mu_{d,i} &= \frac{\mu_{b,i} - \boldsymbol{\phi}_i^T \tilde{\mathbf{x}}_i}{\sqrt{\boldsymbol{\phi}_i^T \tilde{\mathbf{P}}_i \boldsymbol{\phi}_i}} \quad \sigma_{d,i} = \frac{\sigma_{b,i}}{\sqrt{\boldsymbol{\phi}_i^T \tilde{\mathbf{P}}_i \boldsymbol{\phi}_i}} \end{aligned} \quad (12)$$

The equations for calculating the standard deviation of each constraint are not present in [1, 10], but they are a trivial extension from the equations provided for calculating the mean.

#### 4. One-sided constraint

First, consider the case where there is only one constraint on the transformed state, in this case a lower constraint:

$$C_i \leq \mathbf{e}_i^T \mathbf{z}_i \quad (13)$$

Applying a lower constraint to the transformed state is equivalent to finding the conditional probability distribution of the transformed state given that it is higher than the constraint. Using Bayes' theorem, the conditional probability distribution,  $p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta | C_i \leq \zeta)$ , as a function of  $\zeta$  is given by:

$$p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta | C_i \leq \zeta) = \frac{p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta) \times P(C_i \leq \zeta)}{P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i)} \quad (14)$$

where  $p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta)$  is the PDF of  $\mathbf{e}_i^T \mathbf{z}_i$ ,  $P(C_i \leq \zeta)$  is the probability that a point  $\zeta$  is greater than the constraint, and  $P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i)$  is the probability that the transformed state is greater than the constraint.  $P(C_i \leq \zeta)$  is given by:

$$\begin{aligned} P(C_i \leq \zeta) &= \int_{-\infty}^{\zeta} \text{PDF}_{C_i}(c) \, dc \\ &= \text{CDF}_{C_i}(\zeta) \end{aligned} \quad (15)$$

where  $\text{PDF}_{C_i}(c)$  is the PDF of the constraint  $C_i$  evaluated at  $c$ , and  $\text{CDF}_{C_i}(\zeta)$  is the Cumulative Distribution Function (CDF) of  $C_i$  evaluated at  $\zeta$ .  $P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i)$  is given by:

$$\begin{aligned} P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i) &= P(C_i - \mathbf{e}_i^T \mathbf{z}_i \leq 0) \\ &= \int_{-\infty}^0 \text{PDF}_{C_i - \mathbf{e}_i^T \mathbf{z}_i}(\zeta) \, d\zeta \\ &= \text{CDF}_{C_i - \mathbf{e}_i^T \mathbf{z}_i}(0) \end{aligned} \quad (16)$$

where  $C_i - \mathbf{e}_i^T \mathbf{z}_i \sim \mathcal{N}(\mu_{c,i}, \sigma_{c,i}^2 + 1)$  since  $\mathbf{e}_i^T \mathbf{z}_i$  is a standard normal distribution. The conditional probability distribution of the transformed state given that it is higher than the constraint is then given by:

$$\begin{aligned} p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta | C_i \leq \zeta) &= \frac{p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta) \times P(C_i \leq \zeta)}{P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i)} \\ &= \frac{\text{PDF}_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta) \times \text{CDF}_{C_i}(\zeta)}{\text{CDF}_{C_i - \mathbf{e}_i^T \mathbf{z}_i}(0)} \end{aligned} \quad (17)$$

The denominator of (17) can be thought of as a normalising factor—it is the area of the numerator and ensures that the CDF of  $p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta | C_i \leq \zeta)$  is bound between 0 and 1. For states and constraints described by Gaussian distributions,  $p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta | C_i \leq \zeta)$  is given by:

$$p_{e_i^T z_i}(\zeta | C_i \leq \zeta) = \frac{\frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\zeta^2}{2}\right) \frac{1}{2} \left[ 1 + \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right]}{\frac{1}{2} \left[ 1 - \operatorname{erf}\left(\frac{\mu_{c,i}}{\sqrt{2}(\sigma_{c,i}^2 + 1)}\right) \right]} \quad (18)$$

where  $\operatorname{erf}(\cdot)$  is the error function, defined as:

$$\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t \exp(-\tau^2) d\tau \quad (19)$$

Let:

$$\alpha_i = \frac{1}{\sqrt{2\pi} \left[ 1 - \operatorname{erf}\left(\frac{\mu_{c,i}}{\sqrt{2}(\sigma_{c,i}^2 + 1)}\right) \right]} \quad (20)$$

then

$$p_{e_i^T z_i}(\zeta | C_i \leq \zeta) = \alpha_i \exp(-\zeta^2/2) \left[ 1 + \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right] \quad (21)$$

To approximate  $p_{e_i^T z_i}(\zeta | C_i \leq \zeta)$  with a Gaussian distribution, the mean and variance are calculated as follows:

$$\begin{aligned} \mu_i &= \alpha_i \int_{-\infty}^{\infty} \zeta \exp(-\zeta^2/2) \left[ 1 + \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right] d\zeta \\ &= \frac{2\alpha_i}{\sqrt{\sigma_{c,i}^2 + 1}} \exp\left(-\frac{\mu_{c,i}^2}{2(\sigma_{c,i}^2 + 1)}\right) \end{aligned} \quad (22)$$

$$\begin{aligned} \sigma_i^2 &= \alpha_i \int_{-\infty}^{\infty} (\zeta - \mu_i)^2 \exp(-\zeta^2/2) \left[ 1 + \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right] d\zeta \\ &= \alpha_i \left[ \sqrt{2\pi} \left( (1 + \mu_i^2) \left( 1 - \operatorname{erf}\left(\frac{\mu_{c,i}}{\sqrt{2}(\sigma_{c,i}^2 + 1)}\right) \right) \right) \right. \\ &\quad \left. + \frac{2}{\sqrt{\sigma_{c,i}^2 + 1}} \exp\left(-\frac{\mu_{c,i}^2}{2(\sigma_{c,i}^2 + 1)}\right) \left( \frac{\mu_{c,i}}{\sigma_{c,i}^2 + 1} - 2\mu_i \right) \right] \end{aligned} \quad (23)$$

The derivations of the mean and variance can be found in Appendix A and Appendix B respectively. The transformed state estimate, after the  $i$ th constraint has been applied, has the following mean and covariance:

$$\begin{aligned}\tilde{\mathbf{z}}_{i+1} &= \mu_i \mathbf{e}_i \\ \tilde{\mathbf{G}}_{i+1} &= \mathbf{I}_n + (\sigma_i^2 - 1) \mathbf{e}_i \mathbf{e}_i^T\end{aligned}\tag{24}$$

where  $\mathbf{I}_n$  is an  $n \times n$  identity matrix. Taking the inverse of the transformation in (7) gives the mean and variance of the state estimate after the truncation of the  $i$ th constraint:

$$\begin{aligned}\tilde{\mathbf{x}}_{i+1} &= \mathbf{T}_i \mathbf{W}_i^{1/2} \boldsymbol{\rho}_i^T \tilde{\mathbf{z}}_{i+1} + \tilde{\mathbf{x}}_i \\ \tilde{\mathbf{P}}_{i+1} &= \mathbf{T}_i \mathbf{W}_i^{1/2} \boldsymbol{\rho}_i^T \tilde{\mathbf{G}}_{i+1} \boldsymbol{\rho}_i \mathbf{W}_i^{1/2} \mathbf{T}_i^T\end{aligned}\tag{25}$$

This process (from (7) to (25)) is repeated for the  $s$  constraints, incrementing  $i$  each time and using the constrained state estimate after constraint  $i$  has been applied as the input state estimate for constraint  $i + 1$ . After the  $s$  constraints have been applied, the constrained state estimate is:

$$\begin{aligned}\tilde{\mathbf{x}} &= \tilde{\mathbf{x}}_{s+1} \\ \tilde{\mathbf{P}} &= \tilde{\mathbf{P}}_{s+1}\end{aligned}\tag{26}$$

The equations for applying an upper constraint of the form:

$$\mathbf{e}_i^T \mathbf{z}_i \leq D_i\tag{27}$$

are as follows:

$$\alpha_i = \frac{1}{\sqrt{2\pi} \left[ 1 + \operatorname{erf} \left( \frac{\mu_{d,i}}{\sqrt{2(\sigma_{d,i}^2 + 1)}} \right) \right]}\tag{28}$$

$$\mu_i = -\frac{2\alpha_i}{\sqrt{\sigma_{d,i}^2 + 1}} \exp \left( -\frac{\mu_{d,i}^2}{2(\sigma_{d,i}^2 + 1)} \right)\tag{29}$$

$$\sigma_i^2 = \alpha_i \left[ \sqrt{2\pi} \left( (1 + \mu_i^2) \left( 1 + \operatorname{erf} \left( \frac{\mu_{d,i}}{\sqrt{2(\sigma_{d,i}^2 + 1)}} \right) \right) \right) - \frac{2}{\sqrt{\sigma_{d,i}^2 + 1}} \exp \left( -\frac{\mu_{d,i}^2}{2(\sigma_{d,i}^2 + 1)} \right) \left( \frac{\mu_{d,i}}{\sigma_{d,i}^2 + 1} - 2\mu_i \right) \right] \quad (30)$$

Several examples of the proposed method with a lower constraint are shown in Figure 2. As can be seen, the Gaussian approximation is very close to the actual truncated distribution, with the approximation improving as  $\sigma_{c,i}$  increases. As  $\sigma_{c,i} \rightarrow \infty$ , the CDF ( $C_i$ ) approaches a uniform distribution, which means that the truncated PDF approaches the original PDF. In Figure 2c, the lower constraint is higher than the original distribution, resulting in the truncated distribution moving towards the constraint. In this case, the truncated distribution is actually still below the majority of the constraint distribution. Here, the uncertainty of the original and constraint distributions are balanced against one another—the more certain that one of the distributions is, the closer the truncated distribution will be to that distribution. For example, as the uncertainty of the constraint is decreased, the truncated distribution will move to the right. As the uncertainty of the constraint approaches 0, the constraint approaches a hard constraint and the majority of the truncated distribution will be above the constraint.

#### 4.1. Feedback of the truncated estimate

There is disagreement amongst authors as to whether or not the truncated state estimate should be fed back into the Kalman filter, with some suggesting using feedback [13, 14] as shown in Figure 3a and others stating that the truncation process should be kept independent of the unconstrained Kalman filter [10] as shown in Figure 3b. In [10], it is argued that this feedback can lead to overconfident estimates as the information provided by the constraints is used multiple times. In reality, there are two issues to consider when deciding whether or not to use feedback. The first issue concerns the uncertainty model of the constraints—if the constraints are noisy, and if that noise is independent from one time-step to another, then the truncated estimate can be fed back into the Kalman filter. Under these conditions, the constraints are similar to independent noisy measurements. However, many

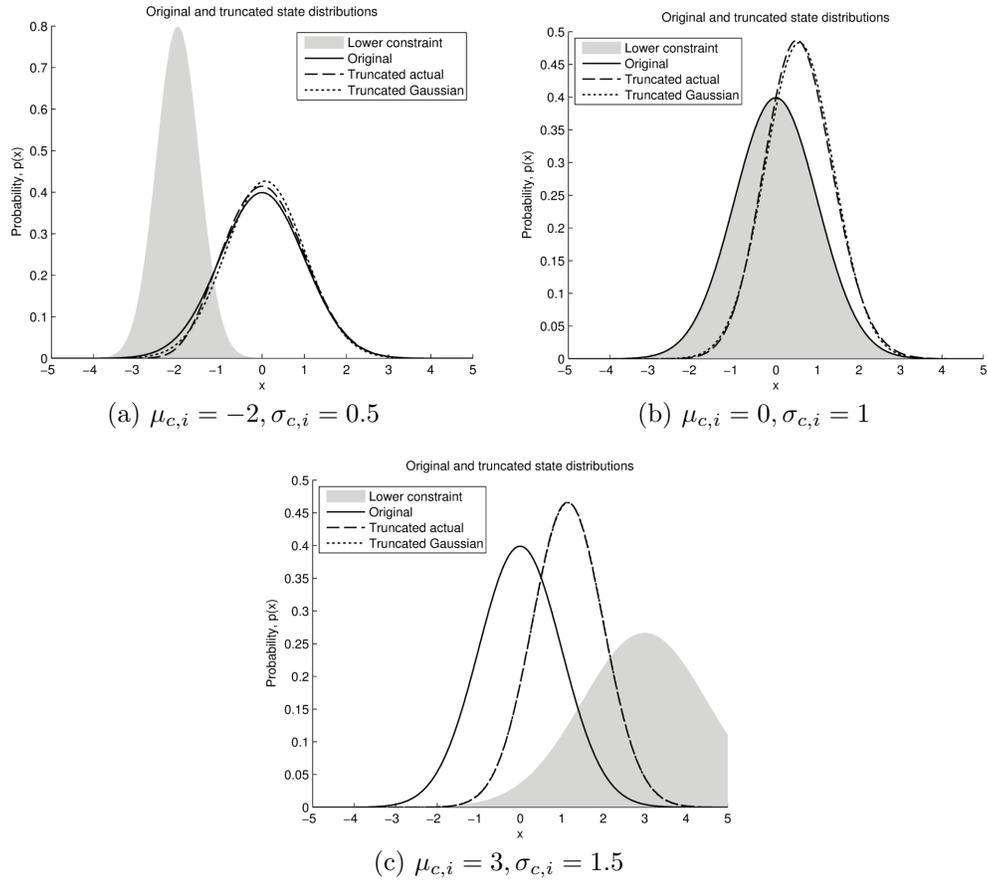
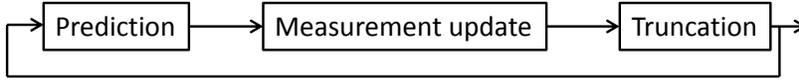


Figure 2: Comparison of the actual truncated distributions and Gaussian approximations of the truncated distributions for several lower constraints. In (c), the lower constraint is higher than the original distribution, resulting in the truncated distribution moving towards the constraint. In this case, the Gaussian approximation is an almost perfect approximation of the truncated distribution and the two lines overlap.



(a) The truncated state estimate is fed back into the Kalman filter.



(b) The Kalman filter is run independently of the truncation, with the feedback occurring after the measurement update and before the truncation.

Figure 3: Feedback of the state estimate into the Kalman filter can occur either before or after the truncation process. Deciding which feedback approach to use depends on the uncertainty model of the constraints and the shape of the truncated distribution.

physical constraints are uncertain rather than noisy—that is, the actual value of the constraint is constant and is not resampled at each time-step. Feeding the truncated estimate back into the Kalman filter in this case can result in overconfident estimates, as will be shown in Section 6.

The second issue arises when the truncated distribution is highly non-Gaussian, which is commonly the case when the uncertainty of the constraint is low in comparison to the uncertainty of the estimate. In these cases, the Gaussian approximation of the truncated distribution introduces error that can accumulate if it is fed back into the Kalman filter, leading to unrepresentative state estimates. An example of such a situation is given in [10]. Consider a scalar system with no process noise such that  $x_{k+1} = x_k$ , no measurements, and a hard constraint of  $x \geq 0$ . If the initial state estimate is a standard normal distribution, then the truncated distribution will have a Gaussian shape for  $x \geq 0$  and 0 otherwise. Approximating this as a Gaussian distribution changes the mean from 0 to  $\sqrt{2/\pi}$  and the variance from 1 to  $(\pi - 2)/\pi$  after the truncation has been applied once. If this were fed back into the Kalman filter, it would result in a monotonically increasing estimate mean and a monotonically decreasing estimate variance for successive applications of the truncation approach. The authors of [10] argue that this is a result of incorporating the information from the constraints multiple times.

However, it is actually the Gaussian approximation of the truncated distribution that causes this behaviour. If it were possible to feed the truncated distribution back into the filter without approximating it as a Gaussian distribution, then applying the constraint at the next time step would result in the exact same truncated distribution.

When deciding whether or not to use feedback, the above issues should be carefully considered. Not using feedback is the conservative option—the resultant estimates will have a higher uncertainty compared to methods using feedback, but will avoid most of these issues. Provided the constraints are noisy rather than uncertain, and have a high noise uncertainty, using feedback is valid and can yield a significant performance benefit over not using feedback. The difference between uncertain and noisy constraints is minimal when the uncertainty of the constraints is small in comparison to the uncertainty of the state estimate, and the main source of error in these cases is the approximation of the truncated distribution. One possible way of dealing with this, suggested by [10], is to only feed back the elements of the truncated estimate where the elements of the original estimate violate the constraint. This was suggested in the context of hard constraints, however, and determining the point at which a soft constraint is violated is ambiguous and is left to the reader.

## 5. Interval constraint

Now consider the case where there are two constraints. After transforming the state using Equations (6)–(12), the two constraints acting on the transformed state are:

$$C_i \leq \mathbf{e}_i^T \mathbf{z}_i \leq D_i \quad (31)$$

Using Bayes' theorem, the conditional probability distribution of the transformed state satisfying the constraints,  $p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta | (C_i \leq \zeta \leq D_i))$ , as a function of  $\zeta$  is given by:

$$p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta | C_i \leq \zeta \leq D_i) = \frac{p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta) \times P(C_i \leq \zeta \leq D_i)}{P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i \leq D_i)} \quad (32)$$

$$= \frac{p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta) \times P(C_i \leq \zeta) \times P(\zeta \leq D_i)}{P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i \leq D_i)} \quad (33)$$

Replacing  $P(C_i \leq \zeta \leq D_i)$  in (32) with  $P(C_i \leq \zeta) \times P(\zeta \leq D_i)$  in (33) is possible since  $C_i$  and  $D_i$  are independent random variables. Note, however, that  $P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i \leq D_i)$  cannot be split up in this way and requires the evaluation of a multi-variate CDF, which does not have an analytical solution. This probability is a normalising factor for the numerator.

This gives the following distribution:

$$\begin{aligned}
p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta | C_i \leq \zeta \leq D_i) &= \frac{p_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta) \times P(C_i \leq \zeta) \times P(\zeta \leq D_i)}{P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i \leq D_i)} \\
&= \frac{\text{PDF}_{\mathbf{e}_i^T \mathbf{z}_i}(\zeta) \times \text{CDF}_{C_i}(\zeta) \times \text{CDF}_{D_i}(\zeta)}{P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i \leq D_i)} \\
&= \frac{\frac{1}{\sqrt{2\pi}} \exp(-\frac{\zeta^2}{2}) \frac{1}{2} \left[ 1 + \text{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right] \frac{1}{2} \left[ 1 - \text{erf}\left(\frac{\zeta - \mu_{d,i}}{\sigma_{d,i}\sqrt{2}}\right) \right]}{P(C_i \leq \mathbf{e}_i^T \mathbf{z}_i \leq D_i)} \quad (34)
\end{aligned}$$

The integrals required to calculate the mean and variance of (34) contain integrals of the form:

$$\int_{-\infty}^{\infty} \exp(-x^2) \text{erf}(ax + b) \text{erf}(cx + d) dx \quad (35)$$

which does not have an analytical solution. The following approximation is proposed:

$$\begin{aligned}
&\left[ 1 + \text{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right] \left[ 1 - \text{erf}\left(\frac{\zeta - \mu_{d,i}}{\sigma_{d,i}\sqrt{2}}\right) \right] \\
&\approx 2 \left[ \text{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) - \text{erf}\left(\frac{\zeta - \mu_{d,i}}{\sigma_{d,i}\sqrt{2}}\right) \right] \quad (36)
\end{aligned}$$

This yields an approximation of the numerator in (34) of:

$$Z(\mathbf{e}_i^T \tilde{\mathbf{z}}_{i+1}) = \frac{1}{4\sqrt{2\pi}} \exp(-\zeta^2/2) \left[ 2 \left( \text{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) - \text{erf}\left(\frac{\zeta - \mu_{d,i}}{\sigma_{d,i}\sqrt{2}}\right) \right) \right] \quad (37)$$

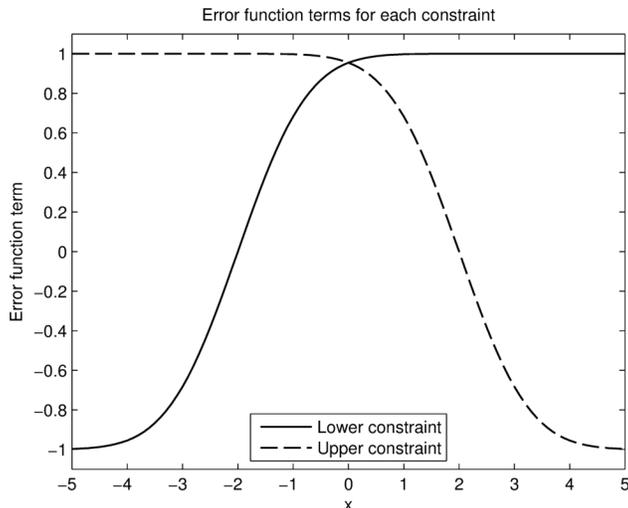


Figure 4: Error function terms for a lower constraint with  $\mu_{c,i} = -2$  and  $\sigma_{c,i} = 1$  and an upper constraint with  $\mu_{d,i} = 2$  and  $\sigma_{d,i} = 1$ . In this case, one of the error function terms is always close to 1.

where  $Z(\mathbf{e}_i^T \tilde{\mathbf{z}}_{i+1})$  is an unnormalised function describing the truncated distribution. This approximation relies on the condition that  $\mu_{c,i} < \mu_{d,i}$ , and the assumption that the constraint distributions do not significantly overlap. If these are satisfied, it is highly likely that one of the erf terms will be equal to 1 when the other is not, giving a good approximation of the actual distribution. This is illustrated in Figure 4.

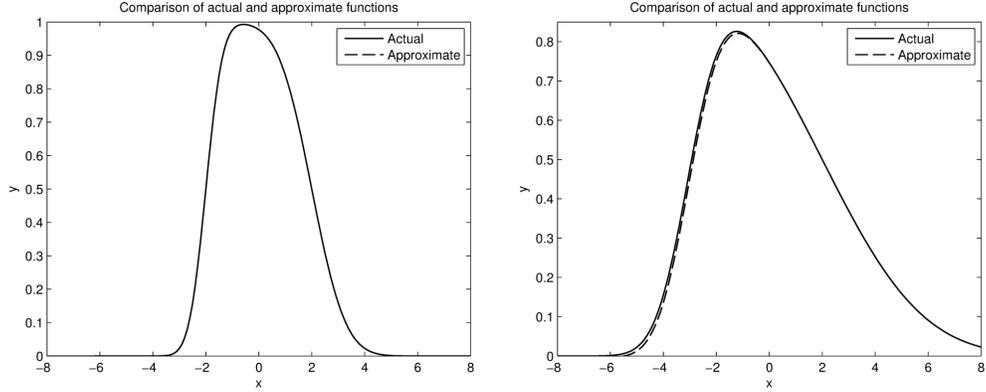
An overlap metric,  $\gamma$ , is defined as:

$$\gamma = \frac{\mu_{d,i} - \mu_{c,i}}{\sigma_{d,i} + \sigma_{c,i}} \quad (38)$$

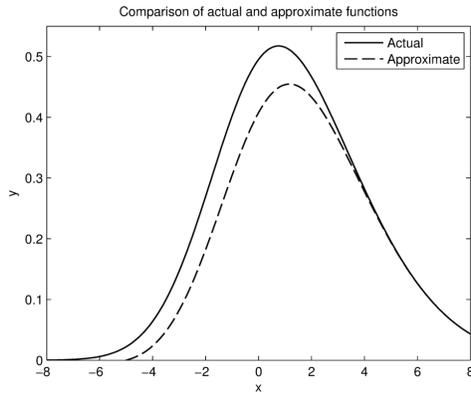
and a shape metric,  $\delta$ , is defined as:

$$\delta = \left| \log \left( \frac{\sigma_{c,i}}{\sigma_{d,i}} \right) \right| \quad (39)$$

$\gamma$  is a measure of how much the probability distributions of the two constraints overlap, and  $\delta$  is a measure of how different the shapes of the probability distributions of the constraints are. Figure 5 shows examples of the approximation applied to several different cases of  $\gamma$  and  $\delta$ . As can be seen, the approximation is an almost perfect approximation in Figure 5a, with the approximation degrading as  $\gamma$  decreases in the other examples.



(a)  $\mu_{c,i} = -2, \sigma_{c,i} = 0.5, \mu_{d,i} = 2, \sigma_{d,i} = 1$ , corresponding to  $\gamma = 2.67$  and  $\delta = 0.30$       (b)  $\mu_{c,i} = -3, \sigma_{c,i} = 1, \mu_{d,i} = 2, \sigma_{d,i} = 3$ , corresponding to  $\gamma = 1.25$  and  $\delta = 0.48$



(c)  $\mu_{c,i} = -1, \sigma_{c,i} = 2, \mu_{d,i} = 2, \sigma_{d,i} = 3.5$ , corresponding to  $\gamma = 0.55$  and  $\delta = 0.24$

Figure 5: Comparison of the actual and approximate functions from (36) for several values of  $\gamma$  and  $\delta$ . In (a), the approximation is an almost perfect approximation of the actual distribution and the two lines overlap.

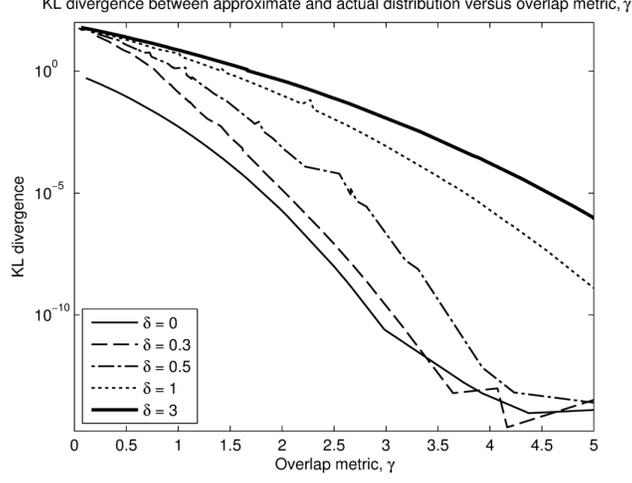


Figure 6: Comparison of the KL divergence between the actual and approximate probability distributions as a function of the overlap metric,  $\gamma$ , where the lines are constant  $\delta$ . For  $\delta > 3$ , the KL divergence is approximately the same as for  $\delta = 3$ .

Figure 6 shows the Kullback–Leibler (KL) divergence between the actual and approximate probability distributions for various  $\delta$  and  $\gamma$ . Increasing  $\gamma$  and decreasing  $\delta$  improves the approximation. For  $\gamma \geq 3$ , the approximation very closely matches the actual distribution, regardless of  $\delta$ .

To normalise  $Z(\mathbf{e}_i^T \tilde{\mathbf{z}}_{i+1})$  to a PDF, the area of the function is calculated as:

$$\begin{aligned}
 & \int_{-\infty}^{\infty} Z(\mathbf{e}_i^T \tilde{\mathbf{z}}_{i+1}) d\zeta \\
 &= \int_{-\infty}^{\infty} \frac{1}{4\sqrt{2\pi}} \exp(-\zeta^2/2) \left[ 2 \left( \operatorname{erf} \left( \frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}} \right) - \operatorname{erf} \left( \frac{\zeta - \mu_{d,i}}{\sigma_{d,i}\sqrt{2}} \right) \right) \right] d\zeta \\
 &= \frac{1}{2} \left[ \operatorname{erf} \left( \frac{\mu_{d,i}}{\sqrt{2}(\sigma_{d,i}^2 + 1)} \right) - \operatorname{erf} \left( \frac{\mu_{c,i}}{\sqrt{2}(\sigma_{c,i}^2 + 1)} \right) \right]
 \end{aligned} \tag{40}$$

The mean of the PDF is then given by:

$$\begin{aligned}
\mu_i &= \alpha_i \int_{-\infty}^{\infty} \zeta \exp(-\zeta^2/2) \left[ \left( \operatorname{erf} \left( \frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}} \right) - \operatorname{erf} \left( \frac{\zeta - \mu_{d,i}}{\sigma_{d,i}\sqrt{2}} \right) \right) \right] d\zeta \\
&= 2\alpha_i \left( \frac{1}{\sqrt{\sigma_{c,i}^2 + 1}} \exp \left( -\frac{\mu_{c,i}^2}{2(\sigma_{c,i}^2 + 1)} \right) \right. \\
&\quad \left. - \frac{1}{\sqrt{\sigma_{d,i}^2 + 1}} \exp \left( -\frac{\mu_{d,i}^2}{2(\sigma_{d,i}^2 + 1)} \right) \right)
\end{aligned} \tag{41}$$

where

$$\alpha_i = \frac{1}{\sqrt{2\pi} \left[ \operatorname{erf} \left( \frac{\mu_{d,i}}{\sqrt{2(\sigma_{d,i}^2 + 1)}} \right) - \operatorname{erf} \left( \frac{\mu_{c,i}}{\sqrt{2(\sigma_{c,i}^2 + 1)}} \right) \right]} \tag{42}$$

The variance is given by:

$$\begin{aligned}
\sigma_i^2 &= \alpha_i \int_{-\infty}^{\infty} (\zeta - \mu_i)^2 \exp(-\zeta^2/2) \left[ \left( \operatorname{erf} \left( \frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}} \right) - \operatorname{erf} \left( \frac{\zeta - \mu_{d,i}}{\sigma_{d,i}\sqrt{2}} \right) \right) \right] d\zeta \\
&= \alpha_i \left[ \sqrt{2\pi} \left( (1 + \mu_i^2) \left( \operatorname{erf} \left( \frac{\mu_{d,i}}{\sqrt{2(\sigma_{d,i}^2 + 1)}} \right) - \operatorname{erf} \left( \frac{\mu_{c,i}}{\sqrt{2(\sigma_{c,i}^2 + 1)}} \right) \right) \right) \right. \\
&\quad \left. + \frac{2}{\sqrt{\sigma_{c,i}^2 + 1}} \exp \left( -\frac{\mu_{c,i}^2}{2(\sigma_{c,i}^2 + 1)} \right) \left( \frac{\mu_{c,i}}{\sigma_{c,i}^2 + 1} - 2\mu_i \right) \right. \\
&\quad \left. - \frac{2}{\sqrt{\sigma_{d,i}^2 + 1}} \exp \left( -\frac{\mu_{d,i}^2}{2(\sigma_{d,i}^2 + 1)} \right) \left( \frac{\mu_{d,i}}{\sigma_{d,i}^2 + 1} - 2\mu_i \right) \right]
\end{aligned} \tag{43}$$

The derivations of these integrals are not provided, but can be easily derived using the solutions provided in Appendix A and Appendix B for the one-sided constraint. The truncated state estimate is then obtained by applying Equations (24) and (25). This process is repeated for the  $s$  constraints, incrementing  $i$  each time.

Several examples of the proposed method for interval constraints are shown in Figure 7. The Gaussian approximation method produces distributions that are very similar to the actual truncated distributions. Figure 7c shows an example where soft and hard constraints are combined—the hard constraint has been modelled as a soft constraint with a very small standard deviation.

## 6. Results

Consider a robot moving along a corridor, as shown in Figure 8. The corridor has a wall 10m in front of the initial position of the robot, and discrete position sensors placed at 1m intervals. The position sensors can detect which side of the set-point of the sensor the robot is on, and have uncertainty on the set-point. The robot has an initial velocity of 10cm/s and accelerates at  $1\text{cm/s}^2$  for 20s, then decelerates at  $1\text{cm/s}^2$  for 20s, before again accelerating at  $1\text{cm/s}^2$  until it reaches the far wall. Two types of robots were considered—one with standard deviations on the acceleration and initial velocity of  $\sigma_a = 1\text{cm/s}^2$  and  $\sigma_v = 3\text{cm/s}$  respectively (Robot A), and a less uncertain one with standard deviations on the acceleration and initial velocity of  $\sigma_a = 0.5\text{cm/s}^2$  and  $\sigma_v = 1.5\text{cm/s}$  respectively (Robot B). The standard deviation of the set-point of the sensors was also varied, with standard deviations (in cm) of  $\sigma_s \in \{0, 5, 10, 15, 20, 25, 30\}$  tested. The position of each robot was tracked using the following Kalman filter run at 10Hz:

$$\mathbf{x} = \begin{bmatrix} x \\ \dot{x} \end{bmatrix} \quad \mathbf{u} = [\ddot{x}] \quad \mathbf{F} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} \frac{\Delta t^2}{2} \\ \Delta t \end{bmatrix} \quad (44)$$

The covariance of the process noise was given by:

$$\mathbf{Q} = \mathbf{G}\mathbf{G}^T\sigma_a^2 = \begin{bmatrix} \frac{\Delta t^4}{4} & \frac{\Delta t^3}{2} \\ \frac{\Delta t^3}{2} & \Delta t^2 \end{bmatrix} \sigma_a^2 \quad (45)$$

The Kalman filter was initialised with:

$$\hat{\mathbf{x}} = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} \quad \mathbf{P} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_v^2 \end{bmatrix} \quad (46)$$

A position sensor changing its reading was incorporated as a noisy measurement of the position:

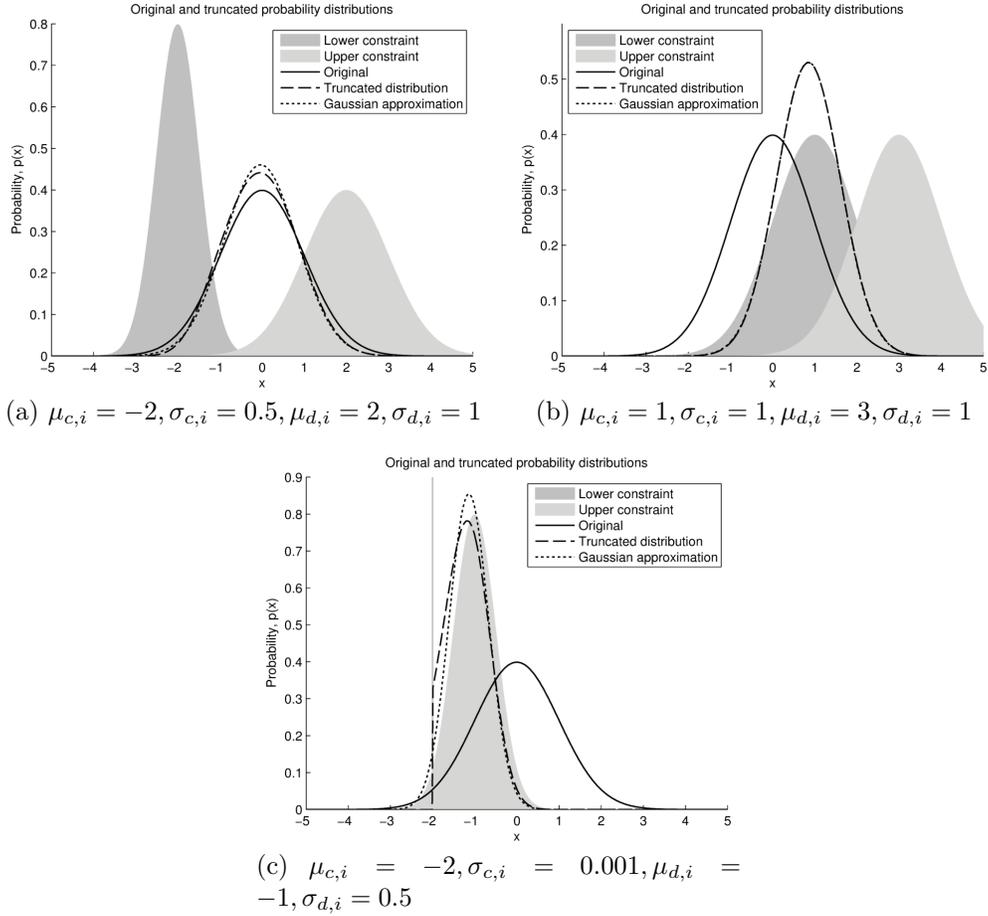


Figure 7: Comparison of the actual and approximate distributions for several values of  $\gamma$  and  $\delta$ . In (b), the Gaussian approximation is an almost perfect approximation of the truncated distribution and the two lines overlap.

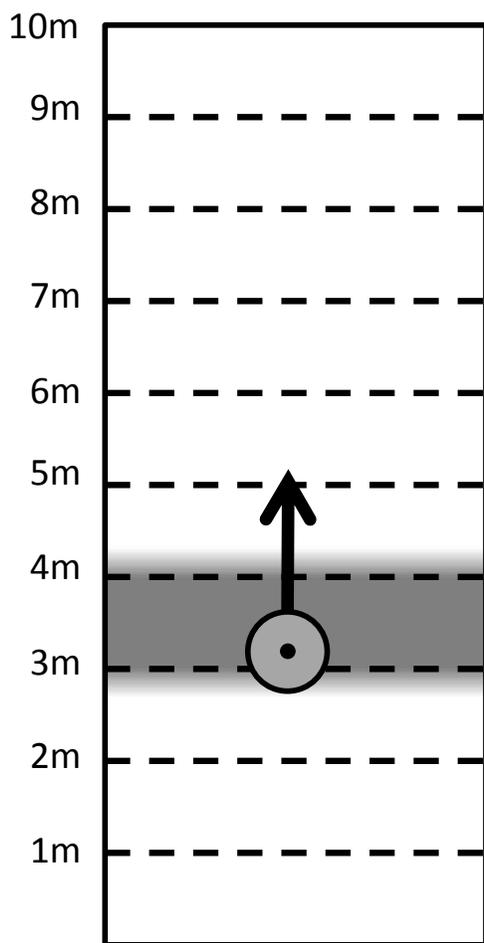


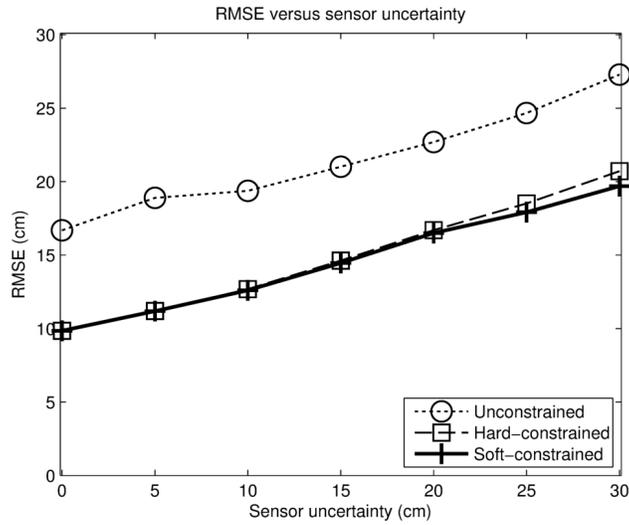
Figure 8: The robot (circle) is moving along the corridor. Positioned at 1m intervals are sensors with uncertain positions that can detect which side of the sensor the robot is on. These sensor readings are used as both measurements (when the sensor reading changes) and constraints. In the image, the position estimate of the robot would be constrained between the sensors at 3m and 4m (shaded area). The uncertainty of the constraints is shown by the shading.

$$\mathbf{H} = [1 \ 0] \quad \mathbf{R} = [\sigma_s^2] \quad (47)$$

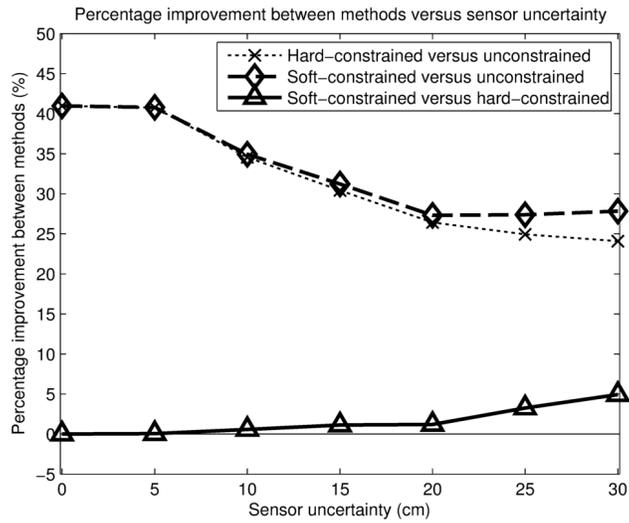
At each time-step, position sensors whose reading did not change were not incorporated into the Kalman filter. The aim of the constrained Kalman filtering approach here was to use the absence of measurements to improve the state estimate. While the robot was in between sensors, the sensors were treated as constraints on the state of the system. The truncation method proposed in this paper (which will be referred to as the soft-constrained Kalman filter) was compared with an unconstrained Kalman filter, and a constrained Kalman filter using the truncation method that ignored the uncertainty of the constraints and treated them as hard constraints (referred to as the hard-constrained Kalman filter). Each combination of robot, sensor uncertainty, and Kalman filter method was tested 1000 times.

The time-average Root Mean Square Error (RMSE) and percentage improvement between methods for Robot A are shown in Figure 9, and the results for Robot B are shown in Figure 10. For Robot A, both the hard-constrained and soft-constrained methods provided a significant benefit over the unconstrained Kalman filter, with an improvement of over 40% in tracking performance when the sensors have no uncertainty. As the uncertainty of the sensors was increased, the soft-constrained method slightly outperformed the hard-constrained method. The process noise for Robot B was significantly less than Robot A. As a result, the uncertainty of the sensors played a larger role in determining the performance of the methods. As can be seen in Figure 10, the hard-constrained Kalman filter was significantly outperformed by the unconstrained Kalman filter once the sensor uncertainty was above 10cm. In these cases, the estimate produced by the hard-constrained Kalman filter was overconfident, and the proposed method outperformed the hard-constrained Kalman filter by over 17%. An example of the overconfident estimates produced by the hard-constrained method is shown in Figure 11. The proposed method strikes a balance between the high uncertainty of the unconstrained Kalman filter and the overconfident estimates of the hard-constrained Kalman filter.

As discussed at the end of Section 4, under certain conditions the truncated state distribution can be fed back into the Kalman filter. For the scenario considered, the discrete position sensors are uncertain rather than noisy, and thus using feedback is not valid. Figure 12b shows the effects of using feedback in the example scenario, with the result that the estimate is

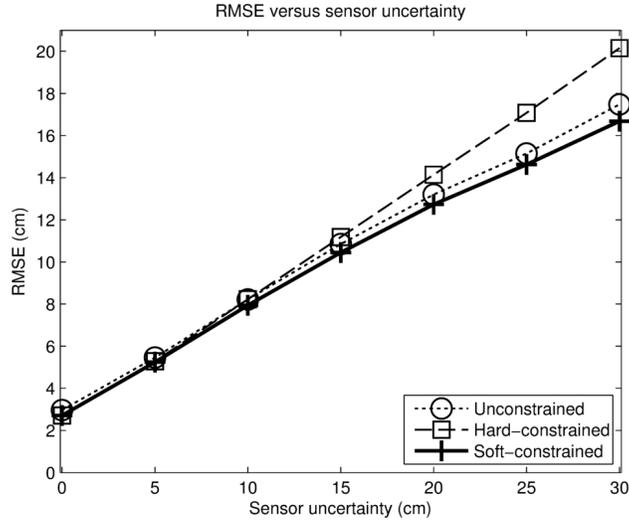


(a) RMSE

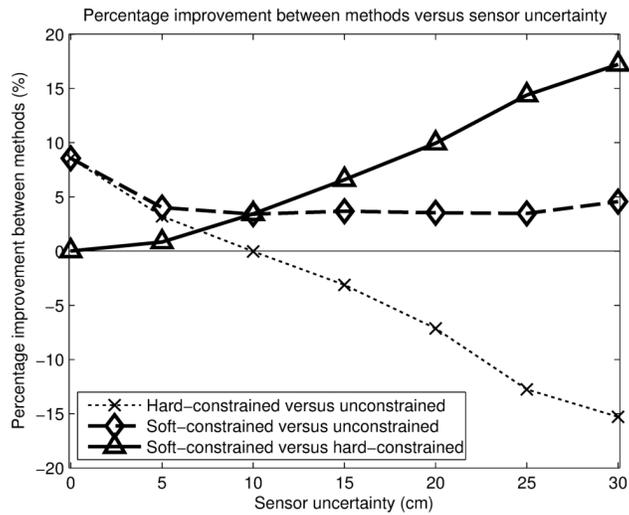


(b) Percentage improvement between methods

Figure 9: RMSE and percentage improvement between methods for Robot A as the sensor uncertainty is varied. The soft-constrained approach is equal to or better than the unconstrained and hard-constrained approaches in all cases.



(a) RMSE



(b) Percentage improvement between methods

Figure 10: RMSE and percentage improvement between methods for Robot B as the sensor uncertainty is varied. The soft-constrained approach is equal to or better than the unconstrained and hard-constrained approaches in all cases. The hard-constrained approach is outperformed by the unconstrained approach when the sensor uncertainty is large.

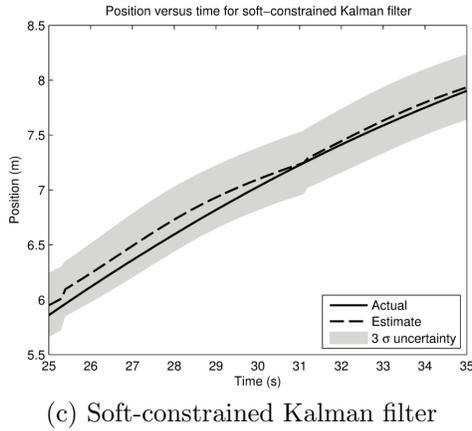
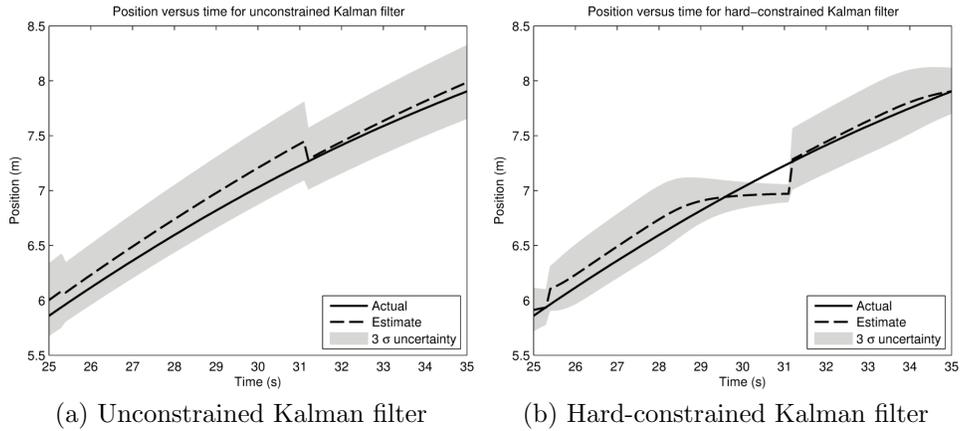


Figure 11: Comparison of the actual and estimated positions with uncertainty for Robot B with sensor uncertainty of 15cm. This illustrates a case where the hard-constrained approach is outperformed by the unconstrained and soft-constrained approaches. The unconstrained estimate has large uncertainty, while the hard-constrained estimate is overconfident. The soft-constrained approach has a lower uncertainty compared to the unconstrained approach without producing the overconfident estimates of the hard-constrained method.

more confident. However, in some cases this estimate can become overconfident and fail to accurately represent the actual state. The method without feedback (Figure 12a) has less confident estimates, but they are not overconfident.

## 7. Conclusion

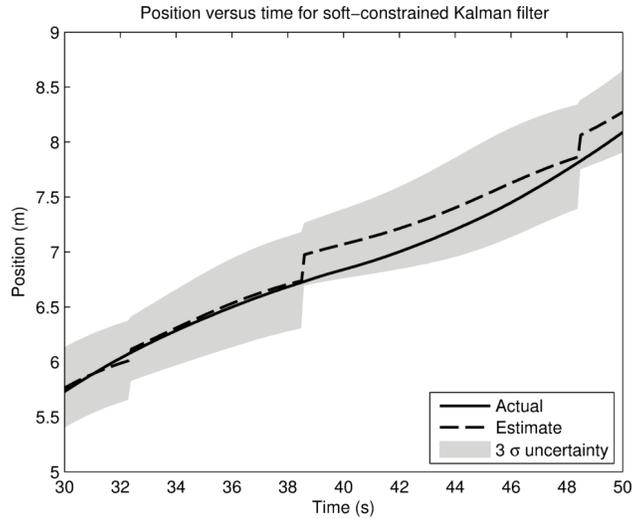
This paper developed an analytical method of truncating an inequality constrained Gaussian distributed variable where the constraints themselves are described by Gaussian distributions. A key aspect of the approach was the use of moment-based Gaussian approximations of the truncated distribution. This truncation method was applied to the constrained Kalman filtering problem where it was shown to outperform unconstrained Kalman filtering and the existing constrained Kalman filter using hard constraints in a simulation example. A key benefit of the developed method compared to hard-constrained Kalman filters is that it is not overconfident near the uncertain constraints. It is an analytical version of existing numerical integration methods, thus providing a computational benefit over the existing numerical methods.

## Acknowledgements

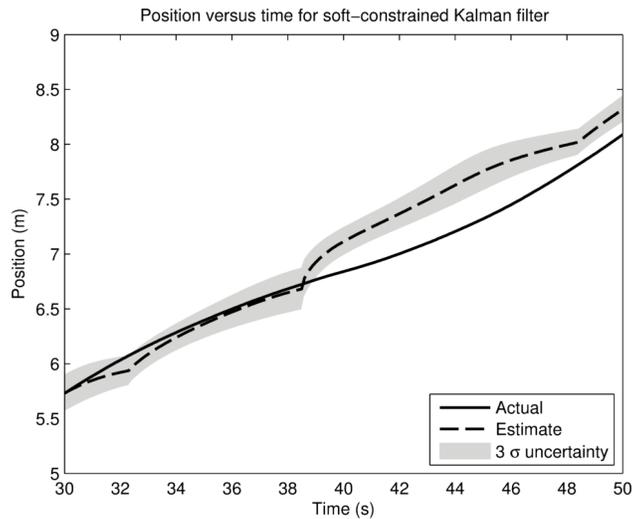
This work was supported by the Rio Tinto Centre for Mine Automation and the Australian Centre for Field Robotics, University of Sydney, Australia.

## Appendix A. Calculation of the mean

The integral in equation (22) is calculated as follows:



(a) Soft-constrained Kalman filter without feedback of the truncated estimate into the Kalman filter



(b) Soft-constrained Kalman filter with feedback of the truncated estimate into the Kalman filter

Figure 12: Comparison of the actual and estimated positions with uncertainty for Robot A with sensor uncertainty of 15cm. In (a), the truncated estimate is not fed back into the Kalman filter, while in (b), the truncated estimate is used by the Kalman filter and the resultant estimate is overconfident.

$$\begin{aligned} & \int_{-\infty}^{\infty} \zeta \exp(-\zeta^2/2) \left[ 1 + \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right] d\zeta \\ &= \int_{-\infty}^{\infty} \zeta \exp(-\zeta^2/2) d\zeta \end{aligned} \quad (\text{A.1})$$

$$+ \int_{-\infty}^{\infty} \zeta \exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) d\zeta \quad (\text{A.2})$$

The first integral, (A.1), equates to 0. For integral (A.2), performing integration by parts with:

$$u = \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \quad dv = \zeta \exp(-\zeta^2/2) d\zeta \quad (\text{A.3})$$

gives:

$$\begin{aligned} & \int_{-\infty}^{\infty} \zeta \exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) d\zeta \\ &= \left[ -\exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right]_{-\infty}^{\infty} \end{aligned} \quad (\text{A.4})$$

$$+ \int_{-\infty}^{\infty} \frac{\exp(-\zeta^2/2)}{\sigma_{c,i}} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{(\zeta - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right) d\zeta \quad (\text{A.5})$$

Equation (A.4) equals 0, and completing the square allows (A.5) to be calculated as:

$$\int_{-\infty}^{\infty} \frac{\exp(-\zeta^2/2)}{\sigma_{c,i}} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{(\zeta - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right) d\zeta = \frac{2 \exp\left(-\frac{\mu_{c,i}^2}{2(\sigma_{c,i}^2+1)}\right)}{\sqrt{\sigma_{c,i}^2+1}} \quad (\text{A.6})$$

To summarise:

$$\boxed{\int_{-\infty}^{\infty} \zeta \exp(-\zeta^2/2) \left[ 1 + \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right] d\zeta = \frac{2}{\sqrt{\sigma_{c,i}^2 + 1}} \exp\left(-\frac{\mu_{c,i}^2}{2(\sigma_{c,i}^2 + 1)}\right)}$$

(A.7)

## Appendix B. Calculation of the variance

The integral in equation (23) is calculated as follows:

$$\begin{aligned} & \int_{-\infty}^{\infty} (\zeta - \mu_i)^2 \exp(-\zeta^2/2) \left[ 1 + \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right] d\zeta \\ &= \int_{-\infty}^{\infty} \zeta^2 \exp(-\zeta^2/2) d\zeta \end{aligned} \quad (\text{B.1})$$

$$+ \int_{-\infty}^{\infty} \zeta^2 \exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) d\zeta \quad (\text{B.2})$$

$$- \int_{-\infty}^{\infty} 2\mu_i \zeta \exp(-\zeta^2/2) d\zeta \quad (\text{B.3})$$

$$- \int_{-\infty}^{\infty} 2\mu_i \zeta \exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) d\zeta \quad (\text{B.4})$$

$$+ \int_{-\infty}^{\infty} \mu_i^2 \exp(-\zeta^2/2) d\zeta \quad (\text{B.5})$$

$$+ \int_{-\infty}^{\infty} \mu_i^2 \exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) d\zeta \quad (\text{B.6})$$

Integral (B.1) equates to  $\sqrt{2\pi}$ . For integral (B.2), performing integration by parts with:

$$u = \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \quad dv = \zeta^2 \exp(-\zeta^2/2) d\zeta \quad (\text{B.7})$$

gives:

$$\begin{aligned} & \int_{-\infty}^{\infty} \zeta^2 \exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) d\zeta \\ &= \left[ \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \left( \sqrt{\frac{\pi}{2}} \operatorname{erf}\left(\frac{\zeta}{\sqrt{2}}\right) - \zeta \exp(-\zeta^2/2) \right) \right]_{-\infty}^{\infty} \end{aligned} \quad (\text{B.8})$$

$$- \int_{-\infty}^{\infty} \frac{1}{\sigma_{c,i}} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{(\zeta - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right) \left( \sqrt{\frac{\pi}{2}} \operatorname{erf}\left(\frac{\zeta}{\sqrt{2}}\right) - \zeta \exp(-\zeta^2/2) \right) d\zeta \quad (\text{B.9})$$

Equation (B.8) is equal to 0, and integral (B.9) can be split into the following integrals:

$$\begin{aligned} & \int_{-\infty}^{\infty} \frac{1}{\sigma_{c,i}} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{(\zeta - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right) \left( \sqrt{\frac{\pi}{2}} \operatorname{erf}\left(\frac{\zeta}{\sqrt{2}}\right) - \zeta \exp(-\zeta^2/2) \right) d\zeta \\ &= \int_{-\infty}^{\infty} \frac{\exp\left(-\frac{(\zeta - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right)}{\sigma_{c,i}} \operatorname{erf}\left(\frac{\zeta}{\sqrt{2}}\right) d\zeta \end{aligned} \quad (\text{B.10})$$

$$- \int_{-\infty}^{\infty} \sqrt{\frac{2}{\pi}} \frac{\zeta}{\sigma_{c,i}} \exp\left(-\frac{(\zeta - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right) \exp(-\zeta^2/2) d\zeta \quad (\text{B.11})$$

Integral (B.10) does not have an indefinite integral, but a definite integral is provided in [20]:

$$\int_{-\infty}^{\infty} \exp(-(\alpha t + \beta)^2) \operatorname{erf}(at + b) dt = \frac{\sqrt{\pi}}{\alpha} \operatorname{erf}\left[\frac{\alpha b - \beta a}{\sqrt{\alpha^2 + a^2}}\right] \quad (\text{B.12})$$

Using this formula, integral (B.10) equates to:

$$\int_{-\infty}^{\infty} \frac{\exp\left(-\frac{(\zeta - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right)}{\sigma_{c,i}} \operatorname{erf}\left(\frac{\zeta}{\sqrt{2}}\right) d\zeta = \sqrt{2\pi} \operatorname{erf}\left(\frac{\mu_{c,i}}{\sqrt{2}(\sigma_{c,i}^2 + 1)}\right) \quad (\text{B.13})$$

Integral (B.11) has the same exponent as in equation (A.6). Completing the square gives:

$$\int_{-\infty}^{\infty} \sqrt{\frac{2}{\pi}} \frac{\zeta}{\sigma_{c,i}} \exp\left(-\frac{(\zeta - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right) \exp(-\zeta^2/2) d\zeta = \frac{2\mu_{c,i} \exp\left(-\frac{\mu_{c,i}^2}{2(\sigma_{c,i}^2+1)}\right)}{(\sigma_{c,i}^2+1)^{3/2}} \quad (\text{B.14})$$

To summarise, integral (B.2) simplifies to:

$$\begin{aligned} \int_{-\infty}^{\infty} \zeta^2 \exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) d\zeta \\ = \frac{2\mu_{c,i} \exp\left(-\frac{\mu_{c,i}^2}{2(\sigma_{c,i}^2+1)}\right)}{(\sigma_{c,i}^2+1)^{3/2}} - \sqrt{2\pi} \operatorname{erf}\left(\frac{\mu_{c,i}}{\sqrt{2(\sigma_{c,i}^2+1)}}\right) \end{aligned} \quad (\text{B.15})$$

Integral (B.3) equates to 0. Integral (B.4) is similar to the integral evaluated in Appendix A and equates to:

$$\int_{-\infty}^{\infty} 2\mu_i \zeta \exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) d\zeta = \frac{4\mu_i}{\sqrt{\sigma_{c,i}^2+1}} \exp\left(-\frac{\mu_{c,i}^2}{\sigma_{c,i}^2+1}\right) \quad (\text{B.16})$$

Integral (B.5) is the integral of a Gaussian distribution and equates to  $\mu_i^2 \sqrt{2\pi}$ . Using the formula provided in [20], integral (B.6) equates to:

$$\int_{-\infty}^{\infty} \mu_i^2 \exp(-\zeta^2/2) \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) d\zeta = \mu_i^2 \sqrt{2\pi} \operatorname{erf}\left(\frac{-\mu_{c,i}}{\sqrt{2(\sigma_{c,i}^2+1)}}\right) \quad (\text{B.17})$$

Summarising:

$$\begin{aligned}
& \int_{-\infty}^{\infty} (\zeta - \mu_i)^2 \exp(-\zeta^2/2) \left[ 1 + \operatorname{erf}\left(\frac{\zeta - \mu_{c,i}}{\sigma_{c,i}\sqrt{2}}\right) \right] d\zeta \\
&= \left[ \sqrt{2\pi} \left( (1 + \mu_i^2) \left( 1 - \operatorname{erf}\left(\frac{\mu_{c,i}}{\sqrt{2(\sigma_{c,i}^2 + 1)}}\right) \right) \right) \right. \\
&\quad \left. + \frac{2}{\sqrt{\sigma_{c,i}^2 + 1}} \exp\left(-\frac{\mu_{c,i}^2}{2(\sigma_{c,i}^2 + 1)}\right) \left( \frac{\mu_{c,i}}{\sigma_{c,i}^2 + 1} - 2\mu_i \right) \right]
\end{aligned}
\tag{B.18}$$

## References

- [1] D. Simon, *Optimal State Estimation: Kalman, H infinity, and nonlinear approaches*, John Wiley & Sons, Hoboken, New Jersey, 2006.
- [2] A. W. Palmer, A. J. Hill, S. J. Scheduling, Stochastic collection and replenishment (SCAR): Objective functions, in: *Proceedings of the 2013 IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2013, pp. 3324–3331.
- [3] A. W. Palmer, A. J. Hill, S. J. Scheduling, Stochastic Collection and Replenishment (SCAR) Optimisation for Persistent Autonomy, in: *Proceedings of the 2014 IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2014, pp. 2943–2949.
- [4] D. Simon, D. L. Simon, Kalman filtering with inequality constraints for turbofan engine health estimation, *IEE Proceedings - Control Theory and Applications* 153 (3) (2006) 371–378.
- [5] N. Shimada, Y. Shirai, Hand gesture estimation and model refinement using monocular camera-ambiguity limitation by inequality constraints, in: *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, 1998.
- [6] M. Boccadoro, F. Martinelli, S. Pagnottelli, Constrained and quantized Kalman filtering for an RFID robot localization problem, *Autonomous Robots* 29 (3-4) (2010) 235–251.

- [7] C. Manes, F. Martinelli, State estimation under quantized measurements: A Sigma-Point Bayesian approach, in: Proceedings of the 52nd IEEE Conference on Decision and Control, 2013, pp. 5024–5029.
- [8] V. Romero-Cano, G. Agamennoni, J. Nieto, A variational approach to simultaneous multi-object tracking and classification, *The International Journal of Robotics Research* (2015) 1–18.
- [9] N. Gupta, R. Hauser, Kalman Filtering with Equality and Inequality State Constraints, Tech. Rep. 07, Oxford University Computing Laboratory (2007). [arXiv:0709.2791](https://arxiv.org/abs/0709.2791).
- [10] D. Simon, D. L. Simon, Constrained Kalman filtering via density function truncation for turbofan engine health estimation, *International Journal of Systems Science* 41 (2) (2010) 159–171.
- [11] D. Simon, Kalman filtering with state constraints: a survey of linear and nonlinear algorithms, *IET Control Theory & Applications* 4 (8) (2010) 1303–1318.
- [12] S. Patil, J. Van Den Berg, R. Alterovitz, Estimating probability of collision for safe motion planning under Gaussian motion and sensing uncertainty, in: Proceedings of the 2012 IEEE International Conference on Robotics and Automation, 2012, pp. 3238–3244.
- [13] B. O. S. Teixeira, L. a. B. Tôrres, L. a. Aguirre, D. S. Bernstein, On unscented Kalman filtering with state interval constraints, *Journal of Process Control* 20 (1) (2010) 45–57.
- [14] O. Straka, J. Duník, M. Šimandl, Truncation nonlinear filters for state estimation with nonlinear inequality constraints, *Automatica* 48 (2) (2012) 273–286.
- [15] Y. Hel-Or, A. Rappoport, M. Werman, Relaxed parametric design with probabilistic constraints, in: Proceedings on the second ACM symposium on Solid modeling and applications, 1993, pp. 261–270.
- [16] X. Shao, B. Huang, J. M. Lee, Constrained Bayesian state estimation - A comparative study and a new particle filter based approach, *Journal of Process Control* 20 (2) (2010) 143–157.

- [17] E. DiGiampaolo, F. Martinelli, A passive UHF-RFID system for the localization of an indoor autonomous vehicle, *IEEE Transactions on Industrial Electronics* 59 (10) (2012) 3961–3970.
- [18] R. Faragher, Understanding the basis of the kalman filter via a simple and intuitive derivation, *IEEE Signal Processing Magazine* 29 (5) (2012) 128–132.
- [19] T. K. Moon, W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, New Jersey: Prentice Hall, Upper Saddle River, 2000.
- [20] H. A. Fayed, A. F. Atiya, An evaluation of the integral of the product of the error function and the normal probability density with application to the bivariate normal integral, *Mathematics of Computation* 83 (285) (2014) 235–250.