International Journal of Control

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tcon20

Exponential cone approach to joint chance constraints in stochastic model predictive control

Filipe Marques Barbosa & Johan Löfberg

To cite this article: Filipe Marques Barbosa & Johan Löfberg (23 Apr 2025): Exponential cone approach to joint chance constraints in stochastic model predictive control, International Journal of Control, DOI: 10.1080/00207179.2025.2492305

To link to this article: https://doi.org/10.1080/00207179.2025.2492305

INTERNATIONAL JOURNAL OF CONTROL
https://doi.org/10.1080/00207179.2025.2492305

Taylor & Francis
Taylor & Francis Group

Exponential cone approach to joint chance constraints in stochastic model predictive control

Filipe Marques Barbosa and Johan Löfberg

Department of Electrical Engineering, Division of Automatic Control, Linköping University, Linköping, Sweden

ABSTRACT

Stochastic model predictive control addresses uncertainties by incorporating the probabilistic description of the disturbances into joint chance constraints. Yet, the classic methods for handling this class of constraints are often computationally inefficient and overly conservative. To overcome this, we propose to replace the nonconvex inverse cumulative distribution function of the standard normal distribution in the deterministic counterpart of these constraints with a highly accurate, exponential cone-representable approximation. This allows the constraints to be formulated as exponential cone functions, and the problem is solved as an exponential cone optimization with risk allocation as decision variables. The main advantage of the proposed approach is that the optimization problem is efficiently solved with off-the-shelf software, and with reduced conservativeness. Moreover, it applies to any problem with linear joint chance constraints subject to normally distributed disturbances. We validate our method with numerical examples of stochastic model predictive control applications.

ARTICLE HISTORY
Received 30 September 2024
Accepted 21 March 2025

KEYWORDS
Joint chance constraint; exponential cone; risk allocation; stochastic systems

1. Introduction

With various engineering applications, model predictive control (MPC) is an advanced technique for controlling multivariable dynamic systems under constraints. MPC repeatedly solves an *open-loop* optimal control problem (OCP) in a receding horizon manner. The solution of the OCP gives a sequence of *control inputs* at every sampling instant at which the system's state is measured. The first control input of this sequence is applied to the system, and the OCP is solved again at the next sampling time. This provides a powerful 'implicit' feedback action to handle system uncertainties and disturbances.

Although the classic formulation of MPC problems provides a certain level of robustness, it is inadequate for systematically addressing uncertainties due to its deterministic formulation. With this in mind, robust MPC (RMPC) approaches incorporated descriptions of uncertainties in their formulation. Early works on RMPC used min-max formulations, accounting for worst-case uncertainty and computing control inputs that guarantee constraint satisfaction for all possible disturbances (Bemporad & Morari, 1999). Min-max approaches are, however, overly conservative and possibly infeasible (Paulson et al., 2017). Attempting to mitigate these problems, closed-loop approximations (Goulart et al., 2006; Löfberg, 2003) and tube-based MPC approaches were introduced (Langson et al., 2004), where constraints are tightened to ensure the system evolves within a 'tube' of trajectories around the nominal trajectory. RMPC formulations can nevertheless be overly conservative since they use bounded deterministic descriptions of uncertainties, accounting for the maximum possible disturbance and often neglecting their statistical properties. Moreover, disturbances have a stochastic nature in many practical cases.

A natural approach to RMPC with stochastic descriptions of uncertainties is to account for the probability of disturbance occurrences explicitly. This leads to stochastic MPC (SMPC), where constraint satisfaction is treated in a probabilistic sense. In other words, SMPC exploits the probabilistic description of the disturbances by replacing hard worst case constraints with soft probabilistic *chance constraints*, which must be satisfied with at least a predefined probability level. This allows for a systematic trade-off between performance and robustness, resulting in less conservative solutions (Farina et al., 2016).

In SMPC, chance constraints can be formulated as *individual* or *joint chance constraints*. Individual chance constraints restrict the probability of violating each constraint independently and have analytic deterministic counterparts for some instances. On the other hand, joint chance constraints limit the probability of violating any constraint over the entire prediction horizon. Although joint chance constraints are a more natural and intuitive choice for SMPC applications, they are generally nonconvex and computationally inefficient (Bertsimas et al., 2011). To overcome this, joint chance constraints are approximated to obtain convex and tractable[1] surrogates, i.e. solvable in polynomial time (Nemirovski & Shapiro, 2007). These approximations can be derived through sampling methods, confidence ellipsoids, or by decomposing them into individual chance constraints (Farina et al., 2016).

Also known as the scenario approach, sampling allows the stochastic chance constraint to be approximated by a standard

CONTACT Filipe Marques Barbosa filipe.barbosa@liu.se Department of Electrical Engineering, Division of Automatic Control, Linköping University, Linköping, Sweden

Supplemental data for this article can be accessed online at https://doi.org/10.1080/00207179.2025.2492305.

convex optimization problem, producing an approximately feasible solution for the original problem. It is, however, computationally expensive due to the large number of samples required for accuracy. We refer to Campi et al. (2009) for more details. Alternatively, confidence ellipsoids approximate the feasible region by a guaranteed convex ellipsoid and can be efficiently solved (Bertsimas & Sim, 2005; Calafiore & Ghaoui, 2006; Van Hessem, 2004). The main drawback of this approach is its over-conservativeness, which leads to sub-optimal solutions (Pilipovsky & Tsiotras, 2021; Vitus & Tomlin, 2011).

Decomposing the joint constraints into individual chance constraints is less conservative than confidence ellipsoids and may result in tractable surrogates. In this method, the individual violation probabilities are bounded using Boole's inequality, resulting in a tighter approximation. The main challenge lies in determining how the risk of constraint violation is spread along the prediction horizon, that is, the *risk allocation*. Thus, the risk allocation can either be fixed a priori, e.g. Blackmore et al. (2011), Knaup et al. (2023), and Li et al. (2022), or treated as a decision variable.

Even though fixing the risk allocation simplifies the optimisation problem, it yields conservative solutions. This is because the optimal risk allocation may change as the system's dynamics evolve. Conversely, treating the individual risks as decision variables may be an alternative to obtain less conservative solutions.

A common issue with having risk allocation as decision variables is that it may result in bilinear terms in the probabilistic constraints. This is usually due to the inclusion of feedback laws or inherited multiplicative disturbances from the system description. The typical approach to deal with this problem is to solve a two-stage (bilevel) optimisation, where an upper-stage optimises the risk allocation and a lower-stage optimises over the control inputs. Pioneering work on this approach include (Blackmore & Ono, 2009; Ono & Williams, 2008; Vitus & Tomlin, 2011) and further improvements were made in (Paulson et al., 2017; Pilipovsky & Tsiotras, 2021). However, the two-stage optimisation is not guaranteed to converge even to a local optimum, and the lower-stage essentially solves a problem with fixed risk allocation.

Note that, we address linear systems with additive disturbances and no feedback law is included, i.e. open-loop SMPC. Hence, bilinearity is not a problem here, but we believe that the upper-stage optimisation can benefit from the results presented in this paper. Notwithstanding, another issue is the handling of the quantile (inverse cumulative distribution) function, which is obtained when deriving a deterministic expression for the probabilistic constraints. Some authors have used the Cantelli-Chebyshev inequality under the assumption that only the first and second moments are known, but not the distribution itself (Li et al., 2022; Paulson et al., 2017; Wang et al., 2022). However, assuming normally distributed disturbance is nonetheless desired in many applications. For these cases, although tractable, replacing the quantile function with the Cantelli-Chebyshev inequality leads to a high degree of conservativeness. See Heirung et al. (2018) for a discussion. On the other hand, directly using the quantile function of a normal distribution (probit function) in the OCP formulation may be inefficient or even intractable as it leads to a general nonlinear, nonconvex optimisation problem involving a non-elementary function which is expensive to compute. Therefore, we argue that a better alternative to achieve tractability and reduce conservativeness is to replace the probit function with a tractable approximation already in the modelling phase.

With this in mind, exponential cone optimisation is a relatively new class of conic optimisation problems that generalises linear optimisation to incorporate inequalities defined by an exponential cone (Chares, 2009). It expands the general framework of conic optimisation and extensions such as mixed-integer conic optimisation. Exponential cones allow to model a variety of constraints involving exponentials and logarithms in a structured fashion. Most importantly, specialised solvers can exploit particular features of the model class and achieve good numerical performance and stability. This has been shown in early implementations in ECOS (Serrano, 2015) and SCS (O'Donoghue et al., 2016). Following this, Dahl and Andersen (2021) uses primal-dual scalings proposed by Tunçel (2001) and generalises the algorithm proposed by Nesterov and Todd (1997, 1998) to handle non-symmetric exponential cones. This results in a practical implementation with good numerical performance and tractable solutions for large-scale exponential cone optimisation problems. This implementation is available in the MOSEK ApS[2] and is the one used here.

In this paper, we propose to replace the probit function in each individual constraint with a highly accurate and exponential cone-representable approximation. This approach offers the advantage of reformulating the OCP as an exponential cone optimisation, allowing it to be efficiently solved. This also results in a significantly less conservative solution. To achieve this, we treat the risk allocation as decision variables and consider probabilistic joint chance constraints with additive normally distributed disturbances in a linear SMPC problem. The joint constraints are decomposed into individual chance constraints and then bounded using Boole's inequality. Finally, we derive the exponential cone representation using a product logarithm (Lambert W-function) as the core of the approximation.

2. Stochastic optimisation

This section presents the stochastic model predictive control problem and how incorporating joint chance constraints into it may lead to conservative solutions or intractable optimisation problems.

2.1 Stochastic model predictive control

Consider a stable discrete-time linear time-invariant system

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{G}\boldsymbol{\omega}_k \tag{1}$$

with the state vector $\boldsymbol{x}_k \in \mathbb{R}^n$, the control input vector $\boldsymbol{u}_k \in \mathbb{R}^m$, the additive disturbances $\boldsymbol{\omega}_k \in \mathbb{R}^p$, and known system matrices $\boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{G}$. Furthermore, the disturbances are assumed to be independent, identically distributed (i.i.d.) random normal variables defined as

$$\boldsymbol{\omega}_k \sim \mathcal{N}(\bar{\boldsymbol{\omega}}_k, \boldsymbol{\Sigma}_k).$$

In an SMPC problem, a sequence of control inputs that minimises the expected value of a cost $E[J(\cdot)]$ over a finite horizon

is computed at each time step k. This is done by solving a constrained optimisation problem that relies on the model (1) and depends on the system's current state. The first computed control input is then applied to the system, and following this, the process is repeated with a shifted horizon.

The states, control inputs, and disturbances over the prediction horizon $N \in \mathbb{N}$, at each time step k, can be expressed as stacked vectors

$$\begin{aligned}\boldsymbol{X} &= [\boldsymbol{x}_k^\top, \boldsymbol{x}_{k+1}^\top, \ldots, \boldsymbol{x}_{k+N}^\top]^\top \\ \boldsymbol{U} &= [\boldsymbol{u}_k^\top, \boldsymbol{u}_{k+1}^\top, \ldots, \boldsymbol{u}_{k+N-1}^\top]^\top \\ \boldsymbol{W} &= [\boldsymbol{\omega}_k^\top, \boldsymbol{\omega}_{k+1}^\top, \ldots, \boldsymbol{\omega}_{k+N-1}^\top]^\top.\end{aligned}$$

The system dynamics can then be written as

$$\boldsymbol{X} = \mathcal{A}\boldsymbol{x}_k + \mathcal{B}\boldsymbol{U} + \mathcal{G}\boldsymbol{W}, \tag{2}$$

where $\mathcal{A} \in \mathbb{R}^{(N+1)n\times n}$, $\mathcal{B} \in \mathbb{R}^{(N+1)n\times Nm}$ and $\mathcal{G} \in \mathbb{R}^{(N+1)n\times Np}$ are defined as

$$\mathcal{A} = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{A} \\ \boldsymbol{A}^2 \\ \vdots \\ \boldsymbol{A}^N \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{B} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{AB} & \boldsymbol{B} & \ldots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}^{N-1}\boldsymbol{B} & \boldsymbol{A}^{N-2}\boldsymbol{B} & \ldots & \boldsymbol{B} \end{bmatrix},$$

$$\mathcal{G} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{G} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{AG} & \boldsymbol{G} & \ldots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}^{N-1}\boldsymbol{G} & \boldsymbol{A}^{N-2}\boldsymbol{G} & \ldots & \boldsymbol{G} \end{bmatrix}. \tag{3}$$

2.1.1 Joint chance constraints

To account for the stochastic nature of (1), probabilistic chance constraints are incorporated into the optimization problem, along with hard constraints. Thus, the inequality constraints are allowed a certain level of violations in a probabilistic sense as

$$P(\boldsymbol{X} \in \mathcal{X}) \geq 1 - \xi, \tag{4}$$

where $\mathcal{X}$ is a convex polytope and ξ the risk of constraint violation. This way, the control inputs are computed so that the constraints containing stochastic parameters remain within the probabilistic bounds. The SMPC problem is formulated as

$$\begin{aligned} \min_{\boldsymbol{U}} \quad & E[J(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{W})] \\ \text{subject to} \quad & \boldsymbol{X} = \mathcal{A}\boldsymbol{x}_k + \mathcal{B}\boldsymbol{U} + \mathcal{G}\boldsymbol{W} \\ & P(\boldsymbol{X} \in \mathcal{X}) \geq 1 - \xi \\ & \boldsymbol{U} \in \mathcal{U} \end{aligned} \tag{5}$$

where $E[J(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{W})]$ is assumed convex and the control inputs are constrained to a convex region $\mathcal{U}$.

The incorporation of joint chance constraints is a natural choice for SMPC problems. However, they are, in general, nonconvex, and an exact tractable representation may not exist. This is because the evaluation of these constraints in principle requires the computation of a multivariate integral, which escalates in difficulty for higher dimensions. Thus, we decompose the joint chance constraints into individual chance constraints and use Boole's inequality to bound the probability of violation.

2.1.2 Boole's inequality

Boole's inequality states that the probability of at least one or several events occurring is less than the sum of the probabilities for the individual events to occur or

$$P(\boldsymbol{X} \notin \mathcal{X}) \leq \sum_{i=1}^{N} P(\boldsymbol{x}_{k+i} \notin \mathcal{X}). \tag{6}$$

This means that if an individual probability of constraint violation is bounded by $P(\boldsymbol{x}_{k+i} \notin \mathcal{X}) \leq \gamma_i$, then the probability of at least one constraint violation is bounded by $\sum_{i=1}^{N} \gamma_i$. As a result, the joint constraint (4) can be replaced by N individual chance constraints as

$$P(\boldsymbol{x}_{k+i} \in \mathcal{X}) \geq 1 - \gamma_i, \quad i = 1, \ldots, N, \tag{7}$$

and satisfy (6) when the *risk allocation* $\gamma_i \in [0, \xi]$ is bounded by

$$\sum_{i=1}^{N} \gamma_i \leq \xi. \tag{8}$$

Following this, the individual probabilistic constraints must be represented analytically in order to use them in an optimisation framework. With this in mind, we formulate the chance constraints as deterministic expressions using the stochastic characterisation of the disturbance.

2.1.3 Analytical chance constraints

The case of individual linear chance constraints with additive normally distributed disturbances is one of the few cases in which chance constraints can be derived into an *exact* analytic (deterministic) representation. Thus, we express the individual chance constraints in (7) explicitly as

$$P(f_j(\boldsymbol{X}, \boldsymbol{U}) + \boldsymbol{q}_j^\top \boldsymbol{W} \leq d_j) \geq 1 - \gamma_j, \quad j = 1, \ldots, M, \tag{9}$$

where

$$f_j(\boldsymbol{X}, \boldsymbol{U}) = \boldsymbol{a}_j^\top \boldsymbol{x}_k + \boldsymbol{b}_j^\top \boldsymbol{U} \tag{10}$$

is a scalar function, $d_j \in \mathcal{X}$ and $\boldsymbol{a}_j^\top, \boldsymbol{b}_j^\top$ and $\boldsymbol{q}_j^\top$ are vectors corresponding to the jth row of (3) as

$$\mathcal{A}_{j,*} \equiv \boldsymbol{a}_j^\top, \quad \mathcal{B}_{j,*} \equiv \boldsymbol{b}_j^\top \quad \text{and} \quad \mathcal{G}_{j,*} \equiv \boldsymbol{q}_j^\top. \tag{11}$$

Note: The individual chance constraints are now indexed by j in order to distinct them from the prediction steps. The constraints in (7) are not necessarily imposed on every state variable.

Then, assuming normally distributed $\boldsymbol{\omega}_k$, an exact analytic representation of (9) is obtained

$$\Phi^{-1}(1 - \gamma_j)\|\boldsymbol{F}^\top \boldsymbol{q}_j\|_2 + \boldsymbol{a}_j^\top \boldsymbol{x}_k + \boldsymbol{b}_j^\top \boldsymbol{U} + \boldsymbol{q}_j^\top \overline{\boldsymbol{W}} - d_j \leq 0, \tag{12}$$

where Φ^{-1} is the probit function[3], $\overline{\boldsymbol{W}}$ the means $\bar{\boldsymbol{\omega}}_i$ in the stacked form

$$\overline{\boldsymbol{W}} = [\bar{\boldsymbol{\omega}}_k^\top, \bar{\boldsymbol{\omega}}_{k+1}^\top, \ldots, \bar{\boldsymbol{\omega}}_{k+N-1}^\top]^\top, \tag{13}$$

and $\boldsymbol{F}$ is the lower triangular matrix obtained from the Cholesky factorisation

$$\boldsymbol{H} = \boldsymbol{F}\boldsymbol{F}^\top \tag{14}$$

where $\boldsymbol{H}$ is a block diagonal with the covariance along prediction horizon defined as

$$\boldsymbol{H} = \oplus_{i=1}^{N} \boldsymbol{\Sigma}_i. \tag{15}$$

Recall that $\boldsymbol{\omega}$ is assumed i.i.d. in this paper, that is, $\boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}$ and $\bar{\boldsymbol{\omega}}_k = \bar{\boldsymbol{\omega}}\ \forall\ k$. The result in (12) is well-known in the literature, and we refer to Appendix 1 for more details.

Remark 2.1: Assuming normally distributed disturbances is reasonable since, i.i.d. random variables can be well approximated to a normal distribution as a result of the central limit theorem (Casella & Berger, 2024).

The constraint *back-off* , which is how much $f_j(\boldsymbol{X}, \boldsymbol{U})$ must back off from d_j so that (9) is satisfied, is determined by the choice of values for γ_j i.e. the risk allocation.

2.1.4 Conservative risk allocation

One approach to handling the risk allocation is to fix the values of γ_j a priori. Following this, a typical choice for determining how the risk is allocated is to use the Bonferroni correction and spread the risk uniformly as $\gamma_j = \xi/M$ (Blackmore & Ono, 2009).

Though this approach simplifies the optimisation problem, it yields very conservative results. Consider, for instance, that the M probabilistic constraints in $\boldsymbol{X} \in \mathcal{X}$ happen to be the same. We have then unnecessarily tightened the risk to $\gamma_j = \xi/M$, instead of simply requiring $\gamma_j = \xi$. On the other hand, consider a trajectory following application with obstacles to be avoided. In this case, the risk allocation would be the same, regardless if the obstacles are far away or too close. In such situations, it is desirable to trade off the risk allocation with another performance criterion.

For minimal conservativeness we must include γ_j as decision variables. However, as $\Phi^{-1}(\cdot)$ is nonconvex and non-elementary, using it directly in the optimisation problem leads to a general nonlinear intractable optimisation problem. To overcome this, we propose an exponential cone representable approximation to the probit function, which is accurate within the interval of interest, i.e. the low risk region. This approach significantly reduces conservativeness, and allows (5) to be formulated as a convex optimisation problem and solved efficiently.

Remark 2.2: Although our focus is the case when risk is a decision variable, it is worth mentioning that (12) becomes a simple linear constraint for fixed risk allocation and a second-order cone when the disturbances are multiplicative.

Remark 2.3: The probit function has been extensively studied, and many highly accurate approximations are available in the literature. However, in addition to accuracy, it is imperative for us that the approximation is also conic-representable for the final problem to be conic representable.

3. Exponential cone representation of chance constraints

Stochastic model predictive control hinges on the computational efficiency of probabilistic constraints, and as seen in the previous section, incorporating joint chance constraints may result in nonconvex, intractable, or conservative optimisation problems. With this in mind, we propose a guaranteed outer exponential cone approximation to the probit function in (12).

3.1 Exponential cone programming

Exponential cone optimisation is an extension of the classical framework of cone optimisation, where the constraints involve the exponential cone. The exponential cone is the conic extension of the exponential function and, as the name suggests, involves an exponential in its definition. It is thus a convex subset of $\mathbb{R}^3$ defined as the set

$$\mathcal{K}_{\text{exp}} = \{(z, y, x) : z \geq ye^{x/y},\ y > 0\} \cup \{(z, 0, x) : z \geq 0,\ x \leq 0\}. \tag{16}$$

In addition to being convex by construction, the exponential cone is a non-symmetric proper cone with numerically stable self-concordant barrier functions. Thus, it has most of the duality, facial reduction, and interior-point properties required for conic optimisation (Permenter et al., 2017). Moreover, Dahl and Andersen (2021) introduced a primal-dual interior-point algorithm specifically designed to handle the non-symmetric exponential cone. This is an extension of the algorithm for symmetric conic optimisation proposed by Nesterov and Todd (1998), allowing the exponential cone optimisation to be solved efficiently. Additionally, MOSEK implementations have shown good numerical performances for large-scale problems, matching standard symmetric cone algorithms.

From the modelling perspective, the exponential cone can be used to model a variety of constraints involving exponentials and logarithms. Both the exponential epigraph and the logarithm hypograph sets can be represented using the exponential cone. In a wider context, it can be used to represent convex compositions of functions such as exponentials, logarithms, entropy, product logarithms, and softplus (commonly used in neural networks). Therefore, replacing the probit function with an exponential cone approximation allows solving the SMPC problem through exponential cone optimisation and efficiently. We refer to MOSEKApS (2024) for a comprehensive list of functions that can be represented using exponential cones.

3.2 Exponential cone approximation of the probit function

With the exponential cone and general conic representable functions, there is a large degree of freedom to create an upper approximation of the probit function. For example, one could use a convex combination of sufficiently many linear, quadratic and exponential functions to achieve a desired level of precision. However, our goal is to construct a compact, highly accurate approximation of the problem. We argue that the product logarithm is a strong candidate to achieve this, and we will use it as the core component of our approximation.

3.2.1 Product logarithm function

Also known as the *Lambert W-function*, the product logarithm has many engineering applications for modelling and solving characteristic, exponential, and inverse equations. It is defined as the solution $W(z)$ that satisfies

$$z = W(z)e^{W(z)}, \tag{17}$$

which is a multivalued function, typically defined for complex z and $W(z)$.

The product logarithm function has complex solutions for real-valued $z < -1/e$, one real negative solution for $z = -1/e$, two real negative solutions for $-1/e < z < 0$, and a *unique* non-negative solution for $z \geq 0$. We say that $W(z)$ has multiple *branches* to distinguish among the possible multiple solutions. The two branches that give real solutions are referred to as the *real branches*: $W_{-1}(z)$ for $-1/e \leq z < 0$ and $W_0(z)$ for $z \geq 0$. See Figure 1. We are interested only in the so-called *principal branch* $W_0 : \mathbb{R}_+ \to \mathbb{R}_+$, which is injective and concave.

Though there is no explicit analytic formula for $W_0(z)$, the hypograph $\{(z, y) : 0 \leq z,\ 0 \leq y \leq W_0(z)\}$ can be described equivalently as

$$z \geq ye^y = ye^{y^2/y}. \tag{18}$$

Thus, it can be modelled as a combination of exponential and quadratic cones as

$$(z, y, x) \in \mathcal{K}_{\text{exp}}, \quad (z \geq ye^{x/y}) \tag{19}$$

$$(1/2, x, y) \in \mathcal{Q}_r, \quad (x \geq y^2), \tag{20}$$

where $\mathcal{Q}_r$ denotes the rotated quadratic cone.

Thus, the principal branch W_0 is used to find an approximation to $\Phi^{-1}(1-\gamma)$ and express (7) as an exponential cone.

3.2.2 Constructing the approximation

A function $\Psi(\gamma)$ is constructed using W_0 and adding compensation terms that do not destroy convexity or tractability as

$$\Psi(\gamma) = -\alpha W_0(\beta\gamma) + \lambda\gamma + \varphi. \tag{21}$$

It is worth noting that terms that tend to infinity, such as $-\log(\gamma)$ or $1/\gamma$, could be included to address the asymptotic behaviour of the probit function ($\Phi^{-1}(1-\gamma) \to \infty$ as $\gamma \to 0$). However, this is not needed in practice, since the growth of Φ^{-1} is remarkably slow. Thus, the approximation is made within an interval of interest $\gamma \in [\gamma_0, 0.5]$, for small values of γ_0.

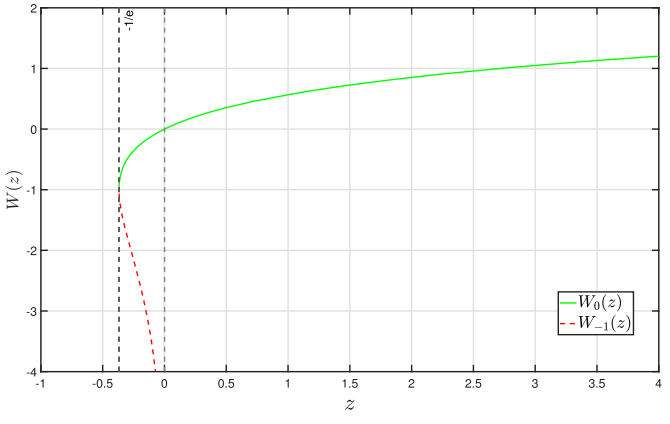


Figure 1. The two main branches of the Lambert W-function.

The curve fitting problem is formulated as a non-negative least-squares regression, optimising over α, β, λ and φ, and minimising the error $e(\gamma)$ in a set of sampling points γ_q.

$$\begin{aligned} \min_{\alpha,\beta,\lambda,\varphi} \quad & \sum_{q=0}^{P} \|\rho_q e(\gamma_q)\|_2^2 \\ \text{subject to} \quad & e(\gamma_q) = \Psi(\gamma_q) - \Phi^{-1}(1-\gamma_q) \\ & e(\gamma_q) \geq 0, \quad \alpha \geq 0 \end{aligned} \tag{22}$$

where the accuracy of the approximation in different points along the region of interest can be addressed by appropriate choice of ρ_q.

Remark 3.1: The constraint $e(\gamma_q) \geq 0$ ensures an upper approximation and is needed since it guarantees $\Psi \geq \Phi^{-1}$, which is necessary for (12) when replacing Φ^{-1} with Ψ.

Finally, the probit function is replaced by $\Psi(\gamma)$ and (12) written as an exponential cone constraint

$$\Psi(\gamma_j)\|\boldsymbol{F}^\top \boldsymbol{q}_j\|_2 + \boldsymbol{a}_j^\top \boldsymbol{x}_k + \boldsymbol{b}_j^\top \boldsymbol{U} + \boldsymbol{q}_j^\top \overline{\boldsymbol{W}} - d_j \leq 0, \tag{23}$$

allowing the risk allocation to be accounted for in the cost function.

4. Results and examples

In this section, we present the parameters of the obtained approximation $\Psi(\gamma)$. This is followed by two examples, with their descriptions and, relevant tables and plots. Lastly, we provide a discussion of the results.

4.1 Approximation results

Considering $\gamma_0 = 10^{-4}$ and using (22) to find the parameters in (21) we obtained

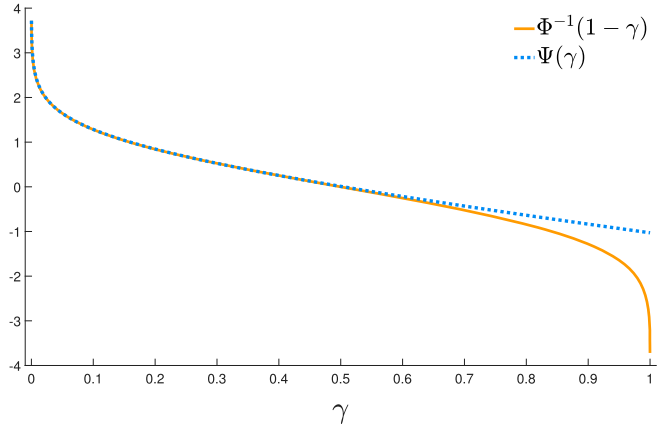
$$\alpha = 0.499492956059166 \quad \lambda = -1.475743096725997$$

$$\beta = 8.082867432374761 \times 10^3 \quad \varphi = 3.965651977413067$$

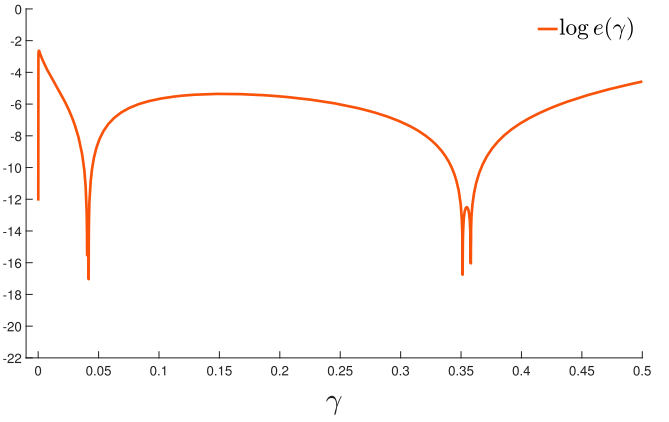
The difference between the obtained approximation $\Psi(\gamma)$ and the exact probit function $\Phi^{-1}(1-\gamma)$ is indiscernible in the region of interest ($\gamma \in [\gamma_0, 0.5]$). See the comparison between in Figure 2(a) and the error as $\log(e(\gamma))$ in Figure 2(b).

4.2 Examples

To validate our approach, we present two examples incorporating both upper and lower bound chance constraints, with γ^u and γ^l the respective associated risks. We use superscript u for upper-bounds associated variables and l for lower-bounds associated variables. The first example is a generic reference-following problem, while the second one is an overhead crane



(a) The probit function $\Phi^{-1}(1-\gamma)$ compared to the approximation $\Psi(\gamma)$.



(b) The logarithm error of the approximation.

Figure 2. The probit function compared to the exponential cone representable approximation for $\gamma \in [0, 1]$, and the logarithm function of relative error for $\gamma \in [0, 0.5]$. (a) The probit function $\Phi^{-1}(1-\gamma)$ compared to the approximation $\Psi(\gamma)$ (b) The logarithm error of the approximation.

application, where a time-optimal planner computed the reference trajectory without accounting for disturbances. The risk allocation was added to the cost function for illustrative purposes. Both examples have quadratic cost functions

$$J = \sum_{i=0}^{N} \|\boldsymbol{Q}^{1/2}(\boldsymbol{z}_{k+i} - \boldsymbol{r}_{k+i})\|_2^2 + \sum_{i=1}^{N} \|\gamma_{k+i}^{u}\|_2^2 + \|\gamma_{k+i}^{l}\|_2^2. \quad (24)$$

with $\boldsymbol{z}$ defined according to the problem and $\boldsymbol{r}$ the reference. For consistency, the predicted risks are also compactly written in the form

$$\boldsymbol{\Gamma} = [\gamma_{k+1}, \gamma_{k+2}, \ldots, \gamma_{k+N}]^\top. \quad (25)$$

Note that risks γ are scalars since the joint constraints were decomposed into individual ones. To avoid confusion, we opt for a more explicit description of the constraints when formulating the SMPC problem used in the examples, thus

$$\begin{aligned} \min_{\boldsymbol{u}_k,\ldots,\boldsymbol{u}_{k+N}} \quad & E\left[J(\boldsymbol{X}, \boldsymbol{\Gamma}^u, \boldsymbol{\Gamma}^l)\right] \\ \text{subject to} \quad & \boldsymbol{x}_{k+i+1} = \boldsymbol{A}\boldsymbol{x}_{k+i} + \boldsymbol{B}u_{k+i} \\ & u_{\min} \le u_{k+i} \le u_{\max} \end{aligned}$$

Table 1. Description and values of the parameters used in the examples.

Description	Parameter	Example 1	Example 2
Prediction horizon	N	6	10
Mean	μ	0	0.2
Covariance	Σ	0.2	0.1
Upper bound	d^u	3.4	0.15
Lower bound	d^l	−3.4	−0.15
Minimum allowed risk	$\xi_{\min}$	10^{-4}	10^{-4}
Maximum allowed risk	$\xi_{\max}$	0.5	0.15
State variable with probabilistic constraint	n_s	1	3

$$\begin{aligned} & \Psi(\gamma_{k+i}^{u})\|\boldsymbol{F}^\top q_j\|_2 + \boldsymbol{a}_j^\top \boldsymbol{x}_k + \boldsymbol{b}_j^\top \boldsymbol{U} \\ & \quad + \boldsymbol{q}_j^\top \overline{\boldsymbol{W}} - d^u \le 0 \\ & \Psi(\gamma_{k+i}^{l})\|\boldsymbol{F}^\top q_j\|_2 - \boldsymbol{a}_j^\top \boldsymbol{x}_k - \boldsymbol{b}_j^\top \boldsymbol{U} \\ & \quad - \boldsymbol{q}_j^\top \overline{\boldsymbol{W}} + d^l \le 0 \\ & \xi_{\min} \le \gamma_{k+i}^{u}, \quad \xi_{\min} \le \gamma_{k+i}^{l} \\ & \sum_{i=1}^{N} (\gamma_{k+i}^{u} + \gamma_{k+i}^{l}) \le \xi_{\max}. \end{aligned} \quad (26)$$

The lower bounds were obtained in the same fashion as the upper bound (see Appendix 1), with the appropriate modifications in the inequality. Moreover, for the sake of simplicity, in both examples, only one state variable is considered to have probabilistic constraints (upper and lower) along the prediction horizon. Thus, j is indexed as $j = in - n + n_s$, where n is the system's dimension and n_s indicates which state has probabilistic constraints.

Table 1 shows some parameter values used in the examples.

Both examples use the YALMIP toolbox to model the OCP (Löfberg, 2004), and MOSEK as the solver.

4.2.1 Example 1

The model parameters used in the first example are

$$\boldsymbol{A} = \begin{bmatrix} -0.1 & -0.3 \\ 1 & -0.5 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and}$$
$$\boldsymbol{C} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^\top, \quad (27)$$

with $z = \boldsymbol{C}\boldsymbol{x}$, the reference given by $r = 3\sin(t)$, $Q = 1$, $u \in [-2, 2]$ and the sampling time $T_s = 0.025$ s. The first state variable x_1 was constrained to not violate the upper and lower constraints with a probability of at least $1 - \gamma_i$ at each prediction step, or $1 - \xi$ for the entire prediction horizon.

Figure 3 compares the feasible set obtained when using the proposed approach to the one obtained when using fixed risk allocation with Bonferroni correction at $i = 1$. The Bonferonni correction was obtained by equally spreading the risk of both the upper and lower bound constraints along the prediction horizon with $\gamma_i^u = 0.5/(2N)$ and $\gamma_i^l = 0.5/(2N)$.

Figure 4 shows the reference to be followed $r(t)$ and the output $y(t)$ after the disturbance entered the system. The horizontal thick lines are the bounds that must not be violated with a probability of at least $1 - \xi$. Furthermore, Figure 5 shows how the risk at each prediction step γ_i evolves along the trajectory.

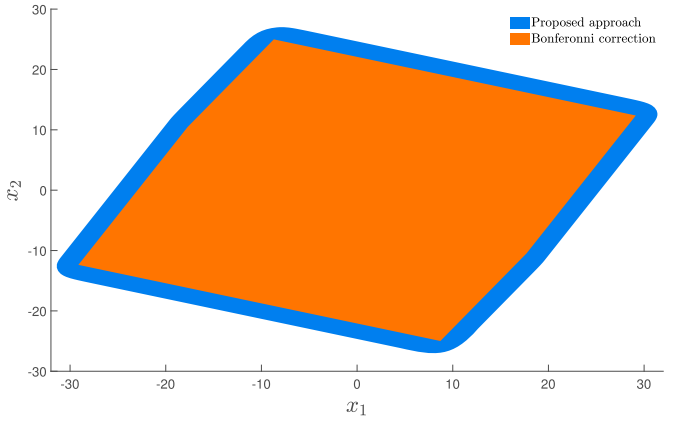


Figure 3. The feasible set at the initial state for the proposed approach (blue) and using the Bonferroni correction (orange) in example 1.

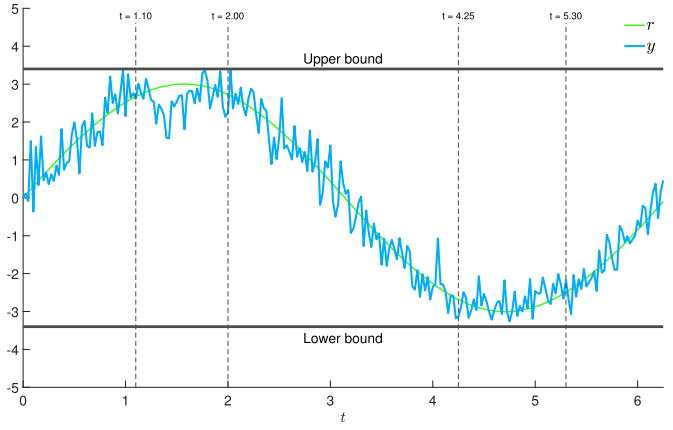


Figure 4. The obtained output trajectory and its reference in example 1.

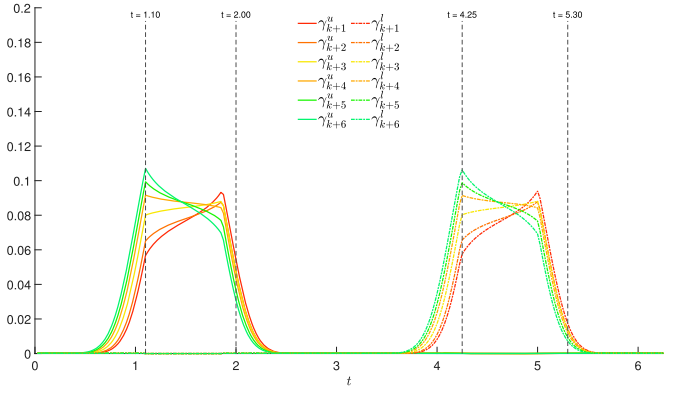


Figure 5. The risk allocation at each time step in example 1. The risk γ^u and γ^l correspond to the upper and lower bound constraints, respectively.

To verify the reliability of the approach in predicting the risk γ_i, Monte Carlo simulations were performed at different points along the trajectory. These are indicated by the dashed vertical lines in Figures 4 and 5. At each of these points, 10^7 realisations of $\boldsymbol{X}$ were generated. The violations in each prediction step i were counted, and the corresponding empirical probabilities $P_e(x_1 \notin \mathcal{X})$ (individual chance of violation) were obtained. The predicted risk of violation γ_i, obtained from the solution of (26), was then compared to $P_e(x_1 \notin \mathcal{X})$ at each prediction step. Additionally, the empirical joint probabilities $P_e(\boldsymbol{X} \notin \mathcal{X})$ were obtained and compared to $\sum \gamma_i$. Table 2 shows the results obtained from the Monte Carlo simulations.

A comparison between the computation costs associated with the proposed approach versus directly using $\Phi^{-1}(\cdot)$ was made. The former uses MOSEK as the solver, while the latter uses IPOPT. The problem was solved at each discrete time step k, and the mean computation time required by the solver to obtain a solution was obtained, along with the standard deviation. This was made for different prediction horizons, and the results are presented in Table 3.

4.2.2 Example 2

The second example illustrates a small-scale overhead crane application with probabilistic constraints on the sway angle θ. See Figure 6. The aim is to move the payload from an initial to a final position by applying a force $u = F_c$ while ensuring that the sway angle remains within the range $\theta \in [\theta_{\min}, \theta_{\max}]$ with a probability of at least $1 - \xi$ in the entire prediction horizon. This is done by following a reference trajectory obtained using the time-optimal control approach presented in (Barbosa & Löfberg, 2022). The reference trajectory includes both the cart position and its velocity, denoted as $\boldsymbol{r} = (x_c^r, \dot{x}_c^r)^\top$. Furthermore, the crane was modelled as a cart-pendulum system, where m_1 is the cart mass, m_2 the payload mass, and l is the rope length.

To apply the optimisation approach described here, a linearised state-space representation of the crane was obtained

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{m_2 g}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{(m_1+m_2)g}{m_1 l} & 0 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 0 \\ \frac{1}{m_1} \\ 0 \\ -\frac{1}{m_1 l} \end{bmatrix},$$

$$\boldsymbol{G} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}^\top, \quad \boldsymbol{C} = \boldsymbol{I},$$

$$m_1 = 1.2\,\text{kg}, \quad m_2 = 0.6\,\text{kg}, \quad l = 1.5\,\text{m} \quad \text{and}$$

$$g = 9.8\,\text{m/s}^2.$$

The state vector is $\boldsymbol{x} = (x_c, \dot{x}_c, \theta, \dot{\theta})^\top$, the control input is constrained to $u \in [-2, 2]$ and $\boldsymbol{Q} = \text{diag}(1, 1)$. Note that $\boldsymbol{G}$ implies that the disturbance affects the sway dynamics directly, and its physical meaning could be interpreted as wind disturbances. Additionally, the system is discretised using a sampling time[4] of $T_s = 0.0371$ s. Refer to Appendix 2 for more details on how the reference trajectory was obtained.

The two plots on top in Figure 7 show the trajectories of the cart position and velocity $(x_c, \dot{x}_c)$, together with their corresponding reference values $(x_c^r, \dot{x}_c^r)$. The bottom plot shows the performed trajectory of the sway velocity $\dot{\theta}$, which is the variable directly affected by the disturbance ω.

The first plot in Figure 8 shows the applied control actions and the control actions obtained from the trajectory planning algorithm u^r. The second plot shows the trajectory performed by θ, the variable with probabilistic constraint, together with the bounds $d^u = \theta_{\max}$ and $d^l = \theta_{\min}$. The colour map in the middle

Table 2. Monte Carlo simulations and predicted probabilities in the first example.

		$\sum \gamma_i$						
		$k+1$	$k+2$	$k+3$	$k+4$	$k+5$	$k+6$	$P_e(\boldsymbol{X} \notin \mathcal{X})$
$t = 1.1$	γ^u	0.0567	0.0651	0.0801	0.0916	0.0994	0.1071	0.5
	γ^l	0	0	0	0	0	0	
	$P_e(x_1 \geq d^u)$	0.0566	0.0650	0.0798	0.0911	0.0987	0.1063	0.4172
	$P_e(x_1 \leq d^l)$	0	0	0	0	0	0	
$t = 2$	γ^u	0.0543	0.0479	0.0470	0.0431	0.0366	0.0308	0.2603
	γ^l	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	
	$P_e(x_1 \geq d^u)$	0.0543	0.0479	0.0470	0.0431	0.0365	0.0308	0.2364
	$P_e(x_1 \leq d^l)$	0	0	0	0	0	0	
$t = 4.25$	γ^u	0	0	0	0	0	0	0.5
	γ^l	0.0576	0.0657	0.0803	0.0914	0.0988	0.1062	
	$P_e(x_1 \geq d^u)$	0	0	0	0	0	0	0.4173
	$P_e(x_1 \leq d^l)$	0.0574	0.0654	0.0798	0.0912	0.0982	0.1055	
$t = 5.3$	γ^u	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0746
	γ^l	0.0173	0.0141	0.0137	0.0120	0.0093	0.0072	
	$P_e(x_1 \geq d^u)$	0	0	0	0	0	0	0.0696
	$P_e(x_1 \leq d^l)$	0.0171	0.0137	0.0134	0.0116	0.0089	0.0068	

Note: Take $t = 1.1$ for example: At the first prediction step ($k+1$), the two first two rows show that the predicted risk of violating the upper bound is 5.67% and the lower bound is 0%. The third and fourth rows show that the upper and lower bounds were violated in 5.66% and 0% of the Monte Carlo experiments. The first row of the last column show that the sum of the predicted risks of violation in the prediction horizon is 50%, according to (8) and Table 1. The second row show that there was at least one constraint violation in 41.72% of the realisations of $\boldsymbol{X}$.

Table 3. A comparison between the computation cost of solving the problem using the proposed approach (with MOSEK) versus directly using Φ^{-1} (with IPOPT). The comparison was made for various prediction horizons.

	N	4	6	8	10	12	14
MOSEK	mean time	0.0021	0.0038	0.0044	0.0058	0.0063	0.0079
	standard deviation	0.0003	0.0005	0.0004	0.0004	0.0006	0.0008
IPOPT	mean time	0.0616	0.0909	0.1259	0.2101	0.4331	0.5406
	standard deviation	0.0194	0.0362	0.2979	0.7400	0.9096	1.3642

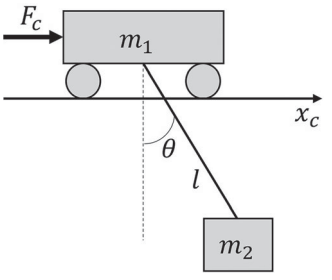


Figure 6. Cart-pendulum system representing a crane in example 2.

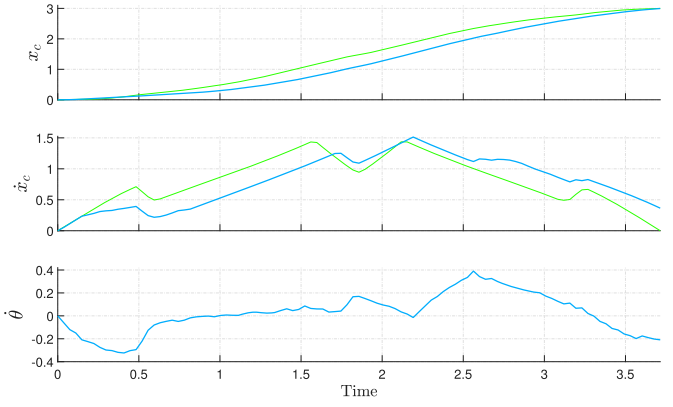


Figure 7. The trajectories performed by the cart position x_c and $\dot{x}_c$ velocity, and the sway angle speed $\dot{\theta}$ (blue). The reference for the cart position and velocity is shown in the first two plots (green).

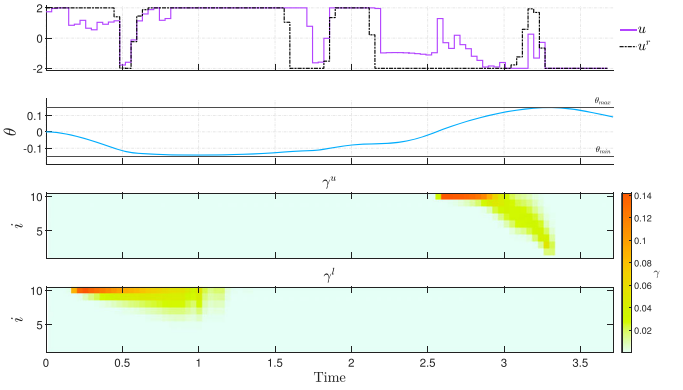


Figure 8. The applied and planned control inputs and the trajectory performed by the sway angle θ, the variable that must stay within the bounds with a certain probability. The colour maps show the risk allocation corresponding to the upper and lower bounds, i corresponds to the step in the prediction horizon, and the colours correspond to the risk, according to the colour scale on the right side.

concerns the risks of violating $\theta_{\max}$ while the colour map in the bottom the risk of violating $\theta_{\min}$. Each prediction step is associated with a specific risk allocation, indicated by the colour scale on the right-hand side.

4.3 Discussion

An exponential cone representable approximation to the probit function was constructed based on the product logarithm function. The parameters and the resulting approximation are shown in the subsection 4.1. By comparing the approximation

with the exact probit function in Figure 2(a) and assessing the error shown in Figure 2(b), one observes that $\Psi(\gamma)$ provides an indiscernible approximation of $\Phi^{-1}(1-\gamma)$ within the interval of interest. Moreover, the fact that Ψ is both exponential cone representable and an accurate approximation to Φ^{-1} implies that the optimal control problem can be solved efficiently, and the solution is significantly less conservative than the existing methods in the literature. The efficiency can be inferred from the comparison shown in Table 3, between solving the optimization problem with the proposed approach versus as a general nonlinear problem. This conservativeness is reflected in the feasible set: when risk is distributed along the prediction horizon with fixed γ_i, the resulting feasible set is smaller than when γ_i is treated as a decision variable. Figure 3 compares the feasible sets obtained for $i = 1$ in the first example, for the proposed approach and Bonferroni correction.

Incorporating the control performance and risk allocation into the cost function allows for a trade-off between them. From Figures 4 and 5, we observe that γ_i remains minimal when z is distant from the bounds. When z is closer to the bounds, γ_i increases, bounded by $\sum \gamma_i^u + \gamma_i^l \leq \xi$, relaxing the problem and thus allowing higher risks. This is because, when z approaches the bounds, the risks corresponding to the end of the prediction horizon are higher than those at its beginning, meaning that the predicted outputs that are closer to the bounds are allowed higher risk. Then, there is a switch in the risk allocation when z begins to move away from the bounds. This happens due to the end of the prediction horizon being farther from the bounds than its beginning, where more risk can now be allocated.

The reliability of the computed γ_i was verified by performing Monte Carlo simulations at different points along the trajectory. The results are shown in Table 2. The simulations indicated that γ_i tend to be slightly more conservative than the observed probabilities and consequently $P_e(\boldsymbol{X} \notin \mathcal{X}) \leq \sum \gamma_i$, which is consistent with the Boole's Inequality in (6).

The reference for the overhead crane (example 2) was computed using a nonlinear cart-pendulum model without considering disturbances. Since the disturbance in this example is additive to $\dot{\theta}$, the trajectory planned for the sway dynamics was discarded, resulting in a reference trajectory that only includes the cart position and velocity. The incorporation of probabilistic constraints on θ is the reason why the reference is not followed with precision, as shown in Figure 7. Now, the control inputs are computed accounting for the probability of violating the bounds due to disturbances, resulting in a more cautious controller than in a purely time-optimal setup.

The applied control inputs were especially less aggressive than those computed by the planning algorithm when the θ approaches the bounds. Furthermore, the colour maps show that the risks corresponding to the end of the prediction horizon were higher as θ approached the bounds. Conversely, when the end of the prediction horizon was farther from the bounds, more risk was allocated at its beginning. These observations align with the discussion made about risk allocation in the first example.

It is worth mentioning that an unstable model produces increasingly cautious inputs due to the disturbance propagation through the prediction horizon. When the system is unstable, the term $\boldsymbol{AG}$ will grow indefinitely, which in turn increases the effect of ω in the system. The standard approach to circumvent this pitfall is to use closed-loop control. However, as mentioned above, it produces a bilinear term in the probabilistic constraints, resulting in a nonconvex problem.

Lastly, we want to emphasise that this result is not limited to SMPC applications but to all classes of linear joint chance constraints subject to normally distributed additive disturbances. Moreover, we believe that systems with normally distributed bilinear chance constraints can also benefit from this approach. This can be done by incorporating this approach to the upper-stage optimisation and making this part less conservative and tractable.

5. Conclusion

The joint chance constraints in an SMPC problem with normally distributed disturbances were reformulated as exponential cone constraints. This was accomplished by first decomposing the joint constraints into individual chance constraints and deriving exact deterministic expressions to represent the probabilistic constraints. Then, the probit function in these expressions was then replaced with a highly accurate, exponential cone-representable approximation. This approach offers a significant advantage since the optimisation problem can be solved efficiently (tractable), and the solution becomes significantly less conservative than existing methods in the literature.

This result can benefit all applications with linear joint chance constraints subject to normally distributed additive disturbances, extending beyond SMPC problems. Future work will focus on developing convex and tractable formulations for the bilinear deterministic counterparts of probabilistic chance constraints.

Notes

1. We use the term tractability to refer to computational tractability throughout this paper. When we say a solution is tractable or computationally efficient, we mean that it can be solved in polynomial time.
2. Available at https://www.mosek.com/.
3. The probit is the inverse of the cumulative distribution function of the standard normal distribution.
4. T_s depends on both the number of control intervals used for constructing the reference and the total time T spent by the crane. See Appendix 2.

Acknowledgments

We would like to thank Stefanie Zimmermann and Anton Kullberg for their valuable discussions and insights.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by VINNOVA Competence Center Link-SIC.

References

Barbosa, F. M., & Löfberg, J. (2022). Time-optimal control of cranes subject to container height constraints. In *2022 American Control Conference (ACC)* (pp. 3558–3563). IEEE.

Bemporad, A., & Morari, M. (1999). Robust model predictive control: A survey. In *Robustness in identification and control* (pp. 207–226). Springer London.

Bertsimas, D., Brown, D. B., & Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM Review*, *53*(3), 464–501. https://doi.org/10.1137/080734510

Bertsimas, D., & Sim, M. (2005). Tractable approximations to robust conic optimization problems. *Mathematical Programming*, *107*(1–2), 5–36.

Blackmore, L., & Ono, M. (2009). Convex chance constrained predictive control without sampling. In *AIAA Guidance, Navigation, and Control Conference*. American Institute of Aeronautics and Astronautics.

Blackmore, L., Ono, M., & Williams, B. C. (2011). Chance-constrained optimal path planning with obstacles. *IEEE Transactions on Robotics*, *27*(6), 1080–1094. https://doi.org/10.1109/TRO.2011.2161160

Calafiore, G. C., & Ghaoui, L. E. (2006). On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications*, *130*(1), 1–22. https://doi.org/10.1007/s10957-006-9084-x

Campi, M. C., Garatti, S., & Prandini, M. (2009). The scenario approach for systems and control design. *Annual Reviews in Control*, *33*(2), 149–157. https://doi.org/10.1016/j.arcontrol.2009.07.001

Casella, G., & Berger, R. (2024). *Statistical inference*. Chapman and Hall/CRC.

Chares, R. (2009). *Cones and interior-point algorithms for structured convex optimization involving powers and exponentials*. UCL-Université Catholique de Louvain. http://hdl.handle.net/2078.1/28538

Dahl, J., & Andersen, E. D. (2021). A primal-dual interior-point algorithm for nonsymmetric exponential-cone optimization. *Mathematical Programming*, *194*(1-2), 341–370. https://doi.org/10.1007/s10107-021-01631-4

Farina, M., Giulioni, L., & Scattolini, R. (2016). Stochastic linear model predictive control with chance constraints – a review. *Journal of Process Control*, *44*, 53–67. https://doi.org/10.1016/j.jprocont.2016.03.005

Goulart, P. J., Kerrigan, E. C., & Maciejowski, J. M. (2006). Optimization over state feedback policies for robust control with constraints. *Automatica*, *42*(4), 523–533. https://doi.org/10.1016/j.automatica.2005.08.023

Heirung, T. A. N., Paulson, J. A., O'Leary, J., & Mesbah, A. (2018). Stochastic model predictive control – how does it work? *Computers and Chemical Engineering*, *114*, 158–170. https://doi.org/10.1016/j.compchemeng.2017.10.026

Knaup, J., Okamoto, K., & Tsiotras, P. (2023). Safe high-performance autonomous off-road driving using covariance steering stochastic model predictive control. *IEEE Transactions on Control Systems Technology*, *31*(5), 2066–2081. https://doi.org/10.1109/TCST.2023.3291570

Langson, W., Chryssochoos, I., Raković, S., & Mayne, D. (2004). Robust model predictive control using tubes. *Automatica*, *40*(1), 125–133. https://doi.org/10.1016/j.automatica.2003.08.009

Li, B., Tan, Y., Wu, A. G., & Duan, G. R. (2022). A distributionally robust optimization based method for stochastic model predictive control. *IEEE Transactions on Automatic Control*, *67*(11), 5762–5776. https://doi.org/10.1109/TAC.2021.3124750

Löfberg, J. (2003). Approximations of closed-loop minimax MPC. In *Proceedings of the 42nd IEEE Conference on Decision and Control* (Vol. 2, pp. 1438–1442). IEEE.

Löfberg, J. (2004). YALMIP a toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*. IEEE.

MOSEKApS (2024). *MOSEK modeling cookbook*.

Nemirovski, A., & Shapiro, A. (2007). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, *17*(4), 969–996. https://doi.org/10.1137/050622328

Nesterov, Y. E., & Todd, M. J. (1997). Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, *22*(1), 1–42. https://doi.org/10.1287/moor.22.1.1

Nesterov, Y. E., & Todd, M. J. (1998). Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, *8*(2), 324–364. https://doi.org/10.1137/S1052623495290209

O'Donoghue, B., Chu, E., Parikh, N., & Boyd, S. (2016). Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and Applications*, *169*(3), 1042–1068. https://doi.org/10.1007/s10957-016-0892-3

Ono, M., & Williams, B. C. (2008). Iterative risk allocation: A new approach to robust model predictive control with a joint chance constraint. In *2008 47th IEEE Conference on Decision and Control*. IEEE.

Paulson, J. A., Buehler, E. A., Braatz, R. D., & Mesbah, A. (2017). Stochastic model predictive control with joint chance constraints. *International Journal of Control*, *93*(1), 126–139. https://doi.org/10.1080/00207179.2017.1323351

Permenter, F., Friberg, H. A., & Andersen, E. D. (2017). Solving conic optimization problems via self-dual embedding and facial reduction: A unified approach. *SIAM Journal on Optimization*, *27*(3), 1257–1282. https://doi.org/10.1137/15M1049415

Pilipovsky, J., & Tsiotras, P. (2021). Covariance steering with optimal risk allocation. *IEEE Transactions on Aerospace and Electronic Systems*, *57*(6), 3719–3733. https://doi.org/10.1109/TAES.2021.3086956

Serrano, S. A. (2015). *Algorithms for unsymmetric cone optimization and an implementation for problems with the exponential cone*. Stanford University.

Tunçel, L. (2001). Generalization of primal–dual interior-point methods to convex optimization problems in conic form. *Foundations of Computational Mathematics*, *1*(3), 229–254. https://doi.org/10.1007/s002080010009

Van Hessem, D. H. (2004). *Stochastic inequality constrained closed-loop model predictive control with application to chemical process operation*. [Dissertation (TU Delft), Delft University of Technology]. DUP Science.

Vitus, M. P., & Tomlin, C. J. (2011). On feedback design and risk allocation in chance constrained control. In *IEEE Conference on Decision and Control and European Control Conference*. IEEE.

Wang, H., Wang, J., Xu, H., & Zhao, S. (2022). A self-triggered stochastic model predictive control for uncertain networked control system. *International Journal of Control*, *96*(8), 2113–2123. https://doi.org/10.1080/00207179.2022.2084163

Appendices

Appendix 1. Exact analytical representation of individual chance constraints

Consider a linear chance constraint with additive disturbance in the form

$$P(f(x) + \boldsymbol{c}^\top \boldsymbol{\omega} \leq b) \geq 1 - \xi, \tag{A1}$$

with constants b and $\boldsymbol{c}^\top$, convex $f : \mathbb{R}^n \to \mathbb{R}$, and normally distributed $\boldsymbol{\omega}$ with mean $\bar{\boldsymbol{\omega}}$ and covariance $\boldsymbol{\Sigma}$. An analytic representation can be derived using the stochastic characterization of $\boldsymbol{\omega}$, i.e. the fact that it is normally distributed.

Since the distribution is symmetric, arrange the terms in (A1) and write the constraint as

$$P\left(\frac{\boldsymbol{c}^\top \boldsymbol{\omega} - \boldsymbol{c}^\top \bar{\boldsymbol{\omega}}}{\sqrt{\boldsymbol{c}^\top \boldsymbol{\Sigma} \boldsymbol{c}}} \leq \frac{b - f(x) - \boldsymbol{c}^\top \bar{\boldsymbol{\omega}}}{\sqrt{\boldsymbol{c}^\top \boldsymbol{\Sigma} \boldsymbol{c}}}\right) \geq 1 - \xi. \tag{A2}$$

As $(\boldsymbol{c}^\top \boldsymbol{\omega} - \boldsymbol{c}^\top \bar{\boldsymbol{\omega}})/\sqrt{\boldsymbol{c}^\top \boldsymbol{\Sigma} \boldsymbol{c}}$ is a zero mean unit variance normal variable, the probability above is simply the cumulative distribution function of the standard normal distribution

$$\Phi\left(\frac{b - f(x) - \boldsymbol{c}^\top \bar{\boldsymbol{\omega}}}{\sqrt{\boldsymbol{c}^\top \boldsymbol{\Sigma} \boldsymbol{c}}}\right) \geq 1 - \xi. \tag{A3}$$

Finally, we arrive at the following expression, which is guaranteed convex for non-negative $\Phi^{-1}(1 - \xi)$, that is, $\xi \leq 0.5$

$$\Phi^{-1}(1 - \xi)\sqrt{\boldsymbol{c}^\top \boldsymbol{\Sigma} \boldsymbol{c}} + f(x) + \boldsymbol{c}^\top \bar{\boldsymbol{\omega}} - b \leq 0. \tag{A4}$$

Appendix 2. Time-optimal reference trajectory

The dynamics of an overhead crane with fixed rope length can be described by nonlinear state-space equations as

$$\dot{x}_1 = x_2$$

$$\begin{aligned}
\dot{x}_2 &= \frac{lm_2 \sin(x_3)x_4^2 + u + m_2 g \cos(x_3)\sin(x_3)}{m_1 + m_2(1-\cos^2(x_3))} \\
\dot{x}_3 &= x_4 \\
\dot{x}_4 &= -\frac{lm_2 \cos(x_3)\sin(x_3)x_4^2 + \cos(x_3)u + (m_1+m_2)g\sin(x_3)}{lm_1 + lm_2(1-\cos^2(x_3))}
\end{aligned} \tag{A5}$$

where the state variables are $\boldsymbol{x} = [x_c, \dot{x}_c, \theta, \dot{\theta}]^\top$.

To formulate the time-optimal control problem, reparametrize (A5) using spatial derivatives instead of temporal derivatives

$$\dot{x}_1 = \frac{\mathrm{d}x_1}{\mathrm{d}t} = x_2 \implies \frac{\mathrm{d}t}{\mathrm{d}x_1} = \frac{1}{x_2}, \tag{A6}$$

and the state vector becomes $\boldsymbol{x} = [t, \dot{x}_c, \theta, \dot{\theta}]^T$.

Note that now x_c is the free variable, and to avoid confusion, we use $x' = \mathrm{d}x/\mathrm{d}x_c$ for derivatives with respect to x_c, and $\dot{x} = \mathrm{d}x/\mathrm{d}t$ for derivatives with respect to time t. The time-optimal control problem is formulated with $x_1(x_c) = t(x_c)$ as the cost function

$$\begin{aligned}
\min \quad & T = t(x_{c_f}) \\
\text{subject to} \quad & x_2 \boldsymbol{x}'(x_c) = f(x_c, \boldsymbol{x}(x_c), u(x_c)) \\
& \boldsymbol{x}(0) = \boldsymbol{x}_0, \quad \boldsymbol{x}(x_{c_f}) = \boldsymbol{x}_f \\
& 0 \le t(x_c), \quad 0 \le x_2(x_c) \\
& x_3 \in [-0.15, 0.15], \quad u \in [-2, 2]
\end{aligned} \tag{A7}$$

with initial position of the cart $x_{c_0} = 0$ and its final position $x_{c_f} = 3$ and the initial and final conditions

$$\begin{aligned}
\boldsymbol{x}(0) &= \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^\top \\
\boldsymbol{x}(x_{p_f}) &= \begin{bmatrix} \quad 0 & 0 & 0 & 0 \end{bmatrix}^\top .
\end{aligned}$$

We refer to Barbosa and Löfberg (2022) for more details.

A trajectory plan with spatial derivatives is obtained from the solution of (A7), with 100 control intervals. Finally, the reference used for controlling the crane in Section 4.2.2 is obtained by resampling the cart position and velocity and expressing them with temporal derivatives as $\boldsymbol{r} = (x_c, \dot{x}_c)^\top$. The sampling time is $T_s = T/100$.

The primary motivation for not including the sway dynamics θ and $\dot{\theta}$ in the reference trajectory is that it is directly affected by the disturbance.