Delft University of Technology

Delft Center for Systems and Control

Technical report 17-004

On optimization of stochastic max-min-plus-scaling systems – An approximation approach*

S.S. Farahani, T. van den Boom, and B. De Schutter

If you want to cite this report, please use the following reference instead:
S.S. Farahani, T. van den Boom, and B. De Schutter, "On optimization of stochastic max-min-plus-scaling systems – An approximation approach," *Automatica*, vol. 83, pp. 20–27, Sept. 2017. doi:10.1016/j.automatica.2017.05.001

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

* This report can also be downloaded via https://pub.bartdeschutter.org/abs/17_004.html

On Optimization of Stochastic Max-Min-Plus-Scaling Systems – An Approximation Approach

Samira S. Farahani ⋆, Ton van den Boom, Bart De Schutter

Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

Abstract

A large class of discrete-event and hybrid systems can be described by a max-min-plus-scaling (MMPS) model, i.e., a model in which the main operations are maximization, minimization, addition, and scalar multiplication. Accordingly, optimization of MMPS systems appears in different problems defined for discrete-event and hybrid systems. For a stochastic MMPS system, this optimization problem is computationally highly demanding as often numerical integration has to be used to compute the objective function. The aim of this paper is to decrease such computational complexity by applying an approximation method that is based on the moments of a random variable and that can be computed analytically.

Key words: Discrete event systems; optimization; max-min-plus-scaling systems; stochastic disturbance; moments.

1 Introduction

Stochastic max-min-plus-scaling (MMPS) systems construct a special class of stochastic discrete-event and hybrid systems, in which processing times and/or transportation times are stochastic quantities; in practice stochastic fluctuations of these times can, e.g., be caused by machine failure or depreciation [23]. The system dynamics of an MMPS system are defined by MMPS expressions, i.e., expressions constructed using the operations maximization, minimization, addition, and multiplication by a scalar. In [22] it was shown that the class of MMPS systems encompasses other classes of discrete-event systems such as max-plus linear systems. Furthermore, it has been shown in [15, 17, 24] that MMPS systems are equivalent to a particular class of hybrid systems, namely continuous piecewise affine (PWA) systems.

In optimization problems for stochastic MMPS or continuous PWA systems, the objective function is often defined as the expected value of an MMPS or a continuous PWA function. Since, in general, there are no analytic expressions for such an expected value, the computation of the objective function in principle involves numerical integration, which is computationally complex and very time consuming. The aim of this paper is to develop an approximation method to compute the expected value of a stochastic MMPS or continuous PWA function with focus on reducing the computational complexity and the computation time. This approximation method is an extension of the method presented in [13], which is inspired by the relation between different types of vector norms, namely the p-norm and the ∞-norm and which in [13] has been only applied to max-plus linear systems with normally distributed disturbances. In [11], the method proposed in [13] has been applied in the context of model predictive control for stochastic MMPS systems and in [12], the approximation method has been extended to a more general class of distributions and an upper bound for the error of this method has been introduced.

The main contributions of the current paper are as follows: 1) proposing a solution for the optimization problem of stochastic MMPS systems using an approximation method that is based on moment-generating functions and is applicable to any distribution with finite moments; 2) discussing the error of the proposed approximation method and presenting finite upper bounds for the error caused by this approximation method. In the discussion of the general optimization problem of stochastic MMPS systems, we introduce two main applications of such systems, namely, the filtering problem and the reference tracking problem. To solve the optimization problem, we use the approximation method proposed in [12], which provides an upper bound for the expected value of a stochastic MMPS function and which can be used as a replacement of the objective function in the optimization problem. In the error discussion, besides presenting an upper bound, we show how different parameters in the approximation function may influence the error bounds.

⋆ Corresponding author S. S. Farahani.

Email addresses: samifarahani@gmail.com (Samira S. Farahani), a.j.j.vandenboom@tudelft.nl (Ton van den Boom), b.deschutter@tudelft.nl (Bart De Schutter).

2 Max-Min-Plus-Scaling Systems

A large class of discrete-event and hybrid systems can be described by a max-min-plus-scaling (MMPS) model [1] . These models are described using MMPS functions.

Definition 1 ([8]) *A function $g : \mathbb{R}^n \to \mathbb{R}$ is a scalar-valued MMPS function of the variables $x_1, \ldots, x_n$ if there exist an index $i \in \{1, \ldots, n\}$ and scalars $\alpha, \beta \in \mathbb{R}$ such that*

$$g(x) = x_i | \alpha | \max(g_k(x), g_l(x)) | \min(g_k(x), g_l(x)) | g_k(x) + g_l(x) | \beta g_k(x),$$

where | stands for "or" and g_k and g_l are scalar-valued MMPS functions .

Accordingly, for a vector-valued MMPS function g, each component of g is an MMPS function of the above form.

A state space representation of a stochastic MMPS system, in which noise and modeling errors are present, can be described as

$$x(k) = \mathcal{M}_x(x(k-1), u(k), \omega(k)) \quad (1)$$
$$y(k) = \mathcal{M}_y(x(k), u(k), \omega(k)). \quad (2)$$

where $\mathcal{M}_x$, $\mathcal{M}_y$ are MMPS functions, $x(k) \in \mathbb{R}^n$ is the system state, $u(k) \in \mathbb{R}^m$ is the system input, and $y(k) \in \mathbb{R}^s$ is the system output at time or event step k. We present both noise and modeling errors in a single framework using a vector $\omega(k)$, which is a vector of independent random variables with a given probability distribution.

The class of MMPS systems is equivalent to a particular class of hybrid systems, namely the class of continuous PWA systems (see [4,7,18,20] for more details on PWA systems).

Proposition 2 ([15,24]) *Any MMPS function can be written as a continuous PWA function and vice versa.*

Moreover, any MMPS function can be written in a canonical form, as expressed in the following proposition.

Proposition 3 ([9]) *Any scalar-valued MMPS function g can be rewritten into the min-max canonical form $g(x) = \min_{i=1,\ldots,K} \max_{j \in n_i}(\alpha_{ij}^T x + \beta_{ij})$ or into the max-min canonical form $g(x) = \max_{i=1,\ldots,L} \min_{j \in m_i}(\gamma_{ij}^T x + \delta_{ij})$ for some integers K, L, non-empty subsets n_i and m_i of the index sets $\{1, 2, \ldots, K\}$ and $\{1, 2 \ldots, L\}$ respectively, real numbers β_{ij}, δ_{ij}, and vectors α_{ij}, γ_{ij}.*

Furthermore, the following proposition from [11, Corollary 5] shows that an MMPS function can be written as a difference of two convex functions.

Proposition 4 *The function $g(x) = \max_{i=1,\ldots,L} \min_{j=1,\ldots,m_i} l_{ij}(x)$ where $l_{ij}(x) = \gamma_{ij}^T x + \xi_{ij}$ is an affine function in x, can be written as $g(x) = s(x) - r(x)$ where $s(\cdot)$ and $r(\cdot)$ are both convex functions defined as follows*

$$r(x) = -\sum_{i=1}^{L} \min_{j=1,\ldots,m_i} l_{ij}(x) = \sum_{i=1}^{L} \max_{j=1,\ldots,m_i} (-l_{ij}(x)) \quad (3)$$

$$s(x) = r(x) + \max_{i=1,\ldots,L} \min_{j=1,\ldots,m_i} l_{ij}(x) = \max_{l=1,\ldots,L} \max_{(j_1,\ldots,j_{i-1},j_{i+1},\ldots,j_L) \in \mathcal{C}(m_1,\ldots,m_{i-1},m_{i+1},\ldots,m_L)} \sum_{\substack{i'=1 \\ i' \neq i}}^{L} (-l_{i'j_{i'}}(x)). \quad (4)$$

The last equality is obtained using the distributive property of addition w.r.t. maximization in which for some integers $L, m_1, \ldots, m_L$, the set $\mathcal{C}(m_1, \ldots, m_L)$ is defined as $\mathcal{C}(m_1, \ldots, m_L) = \{(q_1, \ldots, q_L) | q_k \in \{1, 2, \ldots, m_k\}, k = 1, \ldots, L\}$.

3 Optimization of stochastic MMPS systems

We consider minimization of a stochastic MMPS or continuous PWA function with a random vector ω that has a given probability density function. The class of minimization problems under consideration [2] can be formulated as

$$\min_{u \in \mathbb{R}^n} \ \mathbb{E}_\omega[F(u, \omega)] \quad \text{s.t.} \ \ G(u) \leq 0 \quad (5)$$

where $\mathbb{E}_\omega[\cdot]$ is the expected value operator with respect to ω, F is a scalar MMPS function of u and ω, and G is a convex function of u specifying the input constraints. In order to solve the optimization problem (5), we need to determine the expected value of an MMPS function. One possible approach is numerical integration using the available methods.However, numerical integration is in general both cumbersome and time-consuming, and it becomes even more complicated as the probability density function of ω becomes more complex. Therefore, it is desired to find an alternative approach that is more efficient than numerical integration.

First, we apply Proposition 4 to rewrite the objective function $\tilde{J}(u) = \mathbb{E}_\omega[F(u, \omega)]$ as a difference of two convex functions:

$$\tilde{J}(u) = \mathbb{E}_\omega[F(u, \omega)] = \mathbb{E}_\omega[s(u, \omega) - r(u, \omega)] = \mathbb{E}_\omega[s(u, \omega)] - \mathbb{E}_\omega[r(u, \omega)] \quad (6)$$

where $s(\cdot, \cdot)$ and $r(\cdot, \cdot)$ are defined as given in Proposition 4, and where the last equality stems from the fact that $\mathbb{E}[\cdot]$ is a linear operator. Note that $\tilde{J}(u)$ in (6) results in a non-convex optimization problem. To solve the optimization problem (5), it is only left to compute the expected values in (6). Note that $s(u, \omega)$ and $r(u, \omega)$ both consist of a maximization of set of affine terms. Therefore, our aim is to find an efficient way to compute the following general expression:

$$\mathbb{E}[\max_{j=1,\ldots,n} (\xi_j + \gamma_j^T \omega)] \quad (7)$$

[1] Note that generalized Lindley recursion models [6,30] are special case of MMPS systems.

[2] This class consists of one-stage horizon and receding horizon (model predictive control) optimization problems and the class of control problems with static open-loop inputs.

where $\xi_j \in \mathbb{R}$, $\gamma_j \in \mathbb{R}^{n_\omega}$ is a scaling factor, and $\omega \in \mathbb{R}^{n_\omega}$ is a vector of independent random variables with a given probability distribution. Note that by assumption $\xi_j = \alpha_j + \beta_j u$, for $\alpha_j \in \mathbb{R}$ and $\beta_j \in \mathbb{R}^m$ but that the dependence of ξ_j on u is dropped in the rest of the paper for brevity. Next, we present two popular cases in which the optimization of stochastic MMPS functions appears.

3.1 Filtering Problem

The first problem is a filtering problem [21,29] for which we consider a two-player setting. Assume, e.g., that player one tries to reach his final destination during the discrete time span $\{1, \ldots, \mathcal{T}\}$ and does not want his final location $x(\mathcal{T})$ to be detected by player 2. There exists a control action w that navigates player 1 towards the destination, and a cloaking action v by him that perturbs the measurements of player 2 and he tends to minimize the costs for taking these actions. Player 2 wants to determine an estimate $\hat{x}_\mathcal{T}$ of $x(\mathcal{T})$ as accurately as possible using the obtained measurements of location of player 1 at each time step k for $k = 1, \ldots, \mathcal{T}$.

Let $x(k) \in \mathbb{R}^n$ denote the state, $w(k) \in \mathbb{R}^n$ the control input, $y(k) \in \mathbb{R}^l$ the measurement, $v(k) \in \mathbb{R}^l$ the measurement noise, and $g(\cdot)$ the measurement model. Filtering begins at time step $k=0$ and we assume that player 2 has an estimate $\hat{x}_0$ of the initial state and the measurement time-history $\{y(k)\}_{k=1}^{\mathcal{T}}$ while the initial state $x(0)$ of player 1, the final state $x(\mathcal{T})$, and $w(\cdot)$ are unknown. The estimated final state $\hat{x}_\mathcal{T}$ is a decision variable for player 2. We define the state space model of such system as follows

$$x(k) = f(x(k-1)) + w(k) \tag{8}$$
$$y(k) = g(x(k)) + v(k) \tag{9}$$

where w and v are control variables and f and g are MMPS functions. Player 1 tends to minimize the control and cloaking cost and to maximize the difference between the final state $x(\mathcal{T})$ and estimation of his final state $\hat{x}_\mathcal{T}$ by player 2, while player 2 tends to minimize this difference and has to take the worst-case control and cloaking action of player 1 into account as well as the worst-case uncertainty about the initial state $x(0)$. Hence, the overall objective function is defined as $J(\hat{x}_\mathcal{T}, x(\mathcal{T}), \tilde{w}, \tilde{v}) = -\gamma \sum_{k=1}^{\mathcal{T}} \left[\|w(k)\|_\infty + \|v(k)\|_\infty\right] + \|x(\mathcal{T}) - \hat{x}_\mathcal{T}\|_\infty$ for some $\gamma > 0$, where $\tilde{w} = [w^T(1), \ldots, w^T(\mathcal{T})]^T$ ($\tilde{v}$ is defined similarly) and $x(\mathcal{T})$ can be defined as a function of $x(0)$ and $\tilde{w}$ in a recursive manner using (8), while $\tilde{v}$ can be eliminated through (8)-(9). We assume that w, v, and $x(0)$ are stochastic variables with a given probability distribution and that player 2 knows typical probability density functions of these variables based on previous experience or a priori knowledge. Having this assumption, the optimization problem can be defined as

$$\min_{\hat{x}_\mathcal{T} \in \mathbb{R}^n} \mathbb{E}_{x(0),\tilde{w}}[J(\hat{x}_\mathcal{T}, x(0), \tilde{w})]. \tag{10}$$

This is a problem of the form (5) in which $J(\hat{x}_\mathcal{T}, x(0), \tilde{w})$ is an MMPS function of its arguments.

3.2 Reference Tracking Problem

The second problem is a reference tracking problem [22,28] in which the aim is to minimize the difference between the output of the system and the given reference signal. The system is defined similarly as in (8)-(9), except that the state function contains the input variable $u(k)$. Here, $w(k)$ and $v(k)$ are external noise vectors that perturb the system at each time or event step k. Assume that the initial state $x(0)$ is known and that f and g are MMPS functions of their arguments. The objective function is defined as $J(\tilde{u}, \tilde{w}, \tilde{v}) = \sum_{k=1}^{\mathcal{T}} \lambda_k \|y(k) - r(k)\|_\infty$ where $r(k)$ is the reference signal and λ_k is a weighting factor at time or event step k, and $\tilde{u} = [u^T(1), \ldots, u^T(\mathcal{T})]^T$. Denote the constraints on $\tilde{u}$ by $G(\tilde{u}) \leq 0$, where $G(\cdot)$ is assumed to be a convex function. Note that $x(k)$ has the following form:

$$\begin{aligned} x(k) &= f(x(k-1), u(k)) + w(k) \\ &= f(f(x(k-2), u(k-1)) + w(k-1), u(k)) + w(k) \\ &\cdots \\ &= f(f(f(\ldots(f(x(0), u(1)) + w(1), u(2)) + \ldots) \\ &\qquad + w(k-1), u(k)) + w(k) \\ &= h_k(u(1), \ldots, u(k), w(1), \ldots, w(k)) \end{aligned}$$

where $h_k(\cdot)$ is an MMPS function of its arguments. Hence,

$$\begin{aligned} y(k) &= g(x(k)) + v(k) \\ &= g(h_k(u(1), \ldots, u(k), w(1), \ldots, w(k))) + v(k) \\ &= H_k(u(1), \ldots, u(k), w(1), \ldots, w(k), v(k)) \end{aligned} \tag{11}$$

where $H_k(\cdot)$ is also an MMPS function of its arguments. Accordingly, the objective function can be rewritten as

$$\begin{aligned} &J(\tilde{u}, \tilde{w}, \tilde{v}) \\ &= \sum_{k=1}^{\mathcal{T}} \lambda_k \|H_k(u(1), \ldots, u(k), w(1), \ldots, w(k), v(k))) - r(k)\|_\infty \end{aligned} \tag{12}$$

which is an MMPS function of its arguments. Assuming that both $w(k)$ and $v(k)$ are stochastic variables with a given probability distribution, the optimization problem can be defined as follows:

$$\begin{aligned} \min_{\tilde{u} \in \mathbb{R}^n} \quad & \mathbb{E}_{\tilde{w},\tilde{v}}[J(\tilde{u}, \tilde{w}, \tilde{v})] \\ \text{s.t.} \quad & G(\tilde{u}) \leq 0 \end{aligned} \tag{13}$$

which is again a problem of the form (5).

4 Approximation Method

In this section, an approximation method based on the higher-order moments of a random variable is proposed to compute the expected value of the maximum of several affine expressions. This approach is based on the method presented in [13] for max-plus linear systems.

Let $x = [x_1, \ldots, x_n]^T$ be a vector of random variables in $\mathbb{R}^n$; accordingly, for $p \geq 1$, $\|x\|_p = \left(|x_1|^p + \cdots + |x_n|^p\right)^{1/p}$

defines the p-norm and $\|x\|_\infty = \max(|x_1|, \ldots, |x_n|)$ defines the ∞-norm of x. These norms are related as follows [14]: $\|x\|_\infty \leq \|x\|_p \leq m^{1/p}\|x\|_\infty$, and the proposed approximation method is based on this relation between the vector norms. The following proposition [13, Proposition 2] shows how we can apply p-norms to find an upper bound for $\mathbb{E}\big[\max(x_1, \ldots, x_n)\big]$.

Proposition 5 *Consider random variables x_j for $j = 1, \ldots, n$ and let $p > 1$. Then*

$$\begin{aligned}\mathbb{E}\big[\max(x_1, \ldots, x_n)\big] &\overset{(i)}{\leq} \mathbb{E}\big[\max(|x_1|, \ldots, |x_n|)\big] \\ &\overset{(ii)}{\leq} \mathbb{E}\big[(|x_1|^p + \cdots + |x_n|^p)^{1/p}\big] \\ &\overset{(iii)}{\leq} \Big(\sum_{j=1}^{n} \mathbb{E}\big[|x_j|^p\big]\Big)^{1/p}. \end{aligned} \tag{14}$$

Remark 6 *For a positive even integer $p = 2q,\ q \in \mathbb{N}\setminus\{0\}$, we have $\mathbb{E}[x^p] = \mathbb{E}[|x|^p]$. Hence, if p is an even integer, we can use $\mathbb{E}[x^p]$ in* (14)*. So from now on, p is assumed to be an even integer larger than or equal to* 2*.*

Considering the above remark, we can approximate $\mathbb{E}[\max(x_1, \ldots, x_n)]$ by the following upper bound:

$$\mathfrak{U}\Big(\mathbb{E}[\max(x_1, \ldots, x_n)]\Big) = \Big(\sum_{j=1}^{n} \mathbb{E}\big[(x_j)^p\big]\Big)^{1/p}. \tag{15}$$

Recall that our aim is to compute $\mathbb{E}[\max_{j=1,\ldots,n}(\xi_j + \gamma_j^T\omega)]$ in (7) efficiently. Let $x_j = \xi_j + \gamma_j^T\omega,\ j = 1, \ldots, n$ where ξ_j is an affine expression in u and the elements of the stochastic vector ω, i.e., $\omega_1, \ldots, \omega_{n_\omega}$, are independent random variables, as mentioned before. Hence, in (15), we need to obtain the p-th moment of each random variable $x_j, j = 1, \ldots, n$, which is in fact a shifted, weighted sum of independent random variables $\omega_1, \ldots, \omega_{n_\omega}$. To this end, we use the following property of the moment generating function [26],

$$M_y(t) = E[e^{ty}] = E[e^{ty_1}] \times \cdots \times E[e^{ty_m}] \tag{16}$$

where $y = \sum_{i=1}^{m} y_i$ such that $y_1, \ldots, y_m$ are independent random variables. Now, by adopting (14) and applying (16) to calculate the p-th moment of y, which is the p-th derivative of the moment generating function of y at $t = 0$, an upper bound for (7) can be obtained as follows.

Theorem 7 *For $x_j = \xi_j + \gamma_j^T\omega,\ j = 1, \ldots, n$ in which the elements of the vector ω are independent random variables, an upper bound for $\mathbb{E}[\max(x_1, \ldots, x_n)]$ can be obtained as*

$$\mathbb{E}[\max_{j=1,\ldots,n}(\xi_j + \gamma_j^T\omega)] \leq \tag{17}$$
$$\Bigg(\sum_{j=1}^{n} \sum_{\substack{k_0+k_1+\cdots+k_{n_\omega}=p \\ k_0,k_1,\ldots,k_{n_\omega}\in\mathbb{N}}} \frac{p!}{k_0!\,k_1!\cdots k_{n_\omega}!}\xi_j^{k_0}\prod_{t=1}^{n_\omega}\gamma_{j,t}^{k_t}\mathbb{E}\big[\omega_t^{k_t}\big]\Bigg)^{1/p}.$$

Proof: The proof is straightforward by using the multinomial theorem [16, Section 2.3], and by considering the fact that the elements of the stochastic vector ω, i.e., $\omega_1, \ldots, \omega_{n_\omega}$ are independent and for independent random variables $Z_1, \ldots, Z_{n_\omega}$, we have $\mathbb{E}[\prod_{t=1}^{n_\omega} Z_t] = \prod_{t=1}^{n_\omega}\mathbb{E}[Z_t]$. $\square$

Consequently, we can rewrite (15) as follows

$$\mathfrak{U}\Big(\mathbb{E}[\max_{j=1,\ldots,n}(\xi_j + \gamma_j^T\omega)]\Big) = \tag{18}$$
$$\Bigg(\sum_{j=1}^{n} \sum_{\substack{k_0+k_1+\cdots+k_{n_\omega}=p \\ k_0,k_1,\ldots,k_{n_\omega}\in\mathbb{N}}} \frac{p!}{k_0!\,k_1!\cdots k_{n_\omega}!}\xi_j^{k_0}\prod_{t=1}^{n_\omega}\gamma_{j,t}^{k_t}\mathbb{E}\big[\omega_t^{k_t}\big]\Bigg)^{1/p}$$

where $\mathbb{E}[(\xi_j)^{k_0}] = \xi_j^{k_0}$ since ξ_j does not depend on the stochastic vector ω and hence is not a random variable. In the approximation function (18), we have to compute the k_t-th moment of each random variable $\omega_t,\ t = 0, \ldots, n_\omega$. In general, moments of a random variable can be finite or infinite. Hence, to be able to usefully apply $\mathfrak{U}\Big(\mathbb{E}[\max_{j=1,\ldots,n}(\xi_j + \gamma_j^T\omega)]\Big)$ as an approximation of $\mathbb{E}[\max_{j=1,\ldots,n}(\xi_j + \gamma_j^T\omega)]$, we need to consider random variables with *finite* moments for which a *closed-form* expression exists, such as variables with a uniform distribution, a normal distribution, a Beta distribution, etc. [19,25]. Note that if moments do not have a closed-form expression, one has to obtain them using numerical integration and hence, the approximation method will not be time-efficient anymore. Next, we present a theorem for the case that the independent elements of the stochastic vector w are normally distributed. This theorem allows a faster computation compared to the case using the upper bound in Theorem 7, since we will have less terms in the summation (cf. (19) and (17) for comparison).

Theorem 8 *Let $x_j = \xi_j + \gamma_j^T\omega,\ j = 1, \ldots, n$ in which ω is a stochastic vector and its elements ω_t are independent and normally distributed random variables with mean $\tilde{\mu}_t$ and standard deviation $\tilde{\sigma}_t, t = 1, \ldots, n_\omega$, i.e., $\omega_t \sim \mathcal{N}(\tilde{\mu}_t, \tilde{\sigma}_t)$. For an even integer p, an upper bound for $\mathbb{E}[\max(x_1, \ldots, x_n)]$ can be obtained as*

$$\mathbb{E}[\max(x_1, \ldots, x_n)] \leq \Big(\sum_{j=1}^{n} \sigma_j^p i^{-p} H_p(i\mu_j/\sigma_j)\Big)^{1/p} \tag{19}$$

where μ_j and σ_j are the mean and standard deviation of $x_j, j = 1, \ldots, n$, respectively, and

$$H_p(x) = p!\sum_{l=0}^{p/2} \frac{(-1)^l x^{p-2l}}{2^l l!(p-2l)!} \tag{20}$$

is the p-th Hermite polynomial [1, equations (26.2.51) and (22.3.11)] with $p/2 \in \mathbb{N}\setminus\{0\}$ since p is an even integer.

Proof: For the case of a normally distributed stochastic vector ω, the random variable $x_j = \xi_j + \gamma_j^T\omega$ is also normally distributed with appropriately defined mean μ_j and variance

σ_j^2, using the property of the normal distribution that propagates through linear transformation [10]. Hence, using the p-th Hermite polynomial, we can immediately compute the p-th moment in (15), defined as $\mathbb{E}\big[x_j^p\big] = \sigma_j^p i^{-p} H_p(i\mu_j/\sigma_j)$, which is a real number since p is an even integer. $\square$

Remark 9 *Theorem 8 is actually valid for all distributions that are either preserved under the summation, such as the Poisson and the Gamma distributions, or the distribution of the sum is known, such as the Irvin-Hall distribution, which is the sum of n i.i.d. uniformly distributed random variables [25].*

Recall that in the optimization problem (5), we minimize $\mathbb{E}_\omega[F(u,\omega)]$, which actually leads to the minimization of $\mathbb{E}[s(u,\omega)]$ and the maximization of $\mathbb{E}[r(u,\omega)]$ in (6). Hence, we need to have an upper bound for $\mathbb{E}[s(u,\omega)]$ and a lower bound for $\mathbb{E}[r(u,\omega)]$. Let us consider again the general function $\mathbb{E}[\max_{j=1,\dots,n}(\xi_j + \gamma_j^T\omega)]$ in (7). An upper bound for (7) can be obtained easily by using (18) or equivalently (19), depending on the distribution of ω. To compute a lower bound for (7), we can apply Jensen's inequality for convex functions since $\max(\cdot)$ is a convex function. Hence,

$$\max(\mathbb{E}[x_1],\dots,\mathbb{E}[x_n]) \le \mathbb{E}\big[\max(x_1,\dots,x_n)\big].$$

Therefore, a lower bound for $\mathbb{E}[\max_{j=1,\dots,n}(\xi_j+\gamma_j^T\omega)]$ can be defined as follows

$$\mathfrak{L}\Big(\mathbb{E}[\max_{j=1,\dots,n}(\xi_j+\gamma_j^T\omega)]\Big) = \max_{j=1,\dots,n}(\mathbb{E}[\xi_j+\gamma_j^T\omega]). \quad (21)$$

Consequently, instead of minimizing the objective function $\tilde{J}(u)$ in (6), we will minimize its upper bound $\tilde{J}_{\text{up}}(u) = \mathfrak{U}\Big(\mathbb{E}[s(u,\omega)]\Big) - \mathfrak{L}\Big(\mathbb{E}[r(u,\omega)]\Big)$.

5 On the Error of the Approximation Method

In this section, we show that $\tilde{J}_{\text{up}}(u) - \tilde{J}(u)$ is bounded from above. This means that the errors caused by approximating $\mathbb{E}[s(u,\omega)] = \mathbb{E}[\max(x_1,\dots,x_n)]$ by its upper bound $\mathfrak{U}\Big(\mathbb{E}[s(u,\omega)]]\Big) = \big(\sum_{j=1}^n \mathbb{E}[(x_j)^p]\big)^{1/p}$ (cf. (15)) and $\mathbb{E}[r(u,\omega)] = \mathbb{E}[\sum_{i=1}^M \max(x_1,\dots,x_{m_i})]$ by its lower bound $\mathfrak{L}\Big(\mathbb{E}[r(u,\omega)]\Big) = \sum_{i=1}^M \max(\mathbb{E}[x_1],\dots,\mathbb{E}[x_{m_i}])$ (cf. (21)) are bounded from above.

First the error of approximating $\mathbb{E}[s(u,\omega)]$ by its upper bound will be discussed. Note that $\mathbb{E}[\max(x_1,\dots,x_n)]$ is generally bounded from below and above by

$$\mathfrak{L}\Big(\mathbb{E}[\max(x_1,\dots,x_n)]\Big) \le \mathbb{E}[\max(x_1,\dots,x_n)] \le \mathfrak{U}\Big(\mathbb{E}[\max(x_1,\dots,x_n)]\Big). \quad (22)$$

Hence, the error of approximating $\mathbb{E}[\max(x_1,\dots,x_n)]$ by its upper bound is always bounded by

$$0 \le \mathfrak{U}\Big(\mathbb{E}[\max(x_1,\dots,x_n)]\Big) - \mathbb{E}[\max(x_1,\dots,x_n)] \quad (23)$$
$$\le \mathfrak{U}\Big(\mathbb{E}[\max(x_1,\dots,x_n)]\Big) - \mathfrak{L}\Big(\mathbb{E}[\max(x_1,\dots,x_n)]\Big)$$

and, since by assumption, $x_j,\ j=1,\dots,n$ have finite moments, this upper bound (cf. (15) and (21)) is finite.

Now, consider the error of approximating $\mathbb{E}[r(u,\omega)]$ by its lower bound. Due to linearity of $\mathbb{E}[\cdot]$, we may assume, without loss of generality, that $M = 1$. By Jensen's inequality for convex functions we have

$$\max(\mathbb{E}[x_1],\dots,\mathbb{E}[x_m]) \le \mathbb{E}[\max(x_1,\dots,x_m)] \Rightarrow$$
$$0 \le \mathbb{E}[\max(x_1,\dots,x_m)] - \max(\mathbb{E}[x_1],\dots,\mathbb{E}[x_m]) \quad (24)$$

and we want to show that the right-hand side of (24) is bounded from above. Note that $\max(\mathbb{E}[x_1],\dots,\mathbb{E}[x_m])$ is finite since we assume that the random variables $x_r = \xi_r + \gamma_r^T\omega,\ r=1,\dots,m$ have finite moments. To obtain an upper bound for (24), we show that $\mathbb{E}[\max(x_1,\dots,x_m)]$ is bounded from above. Let $x_1,\dots,x_m$ be random variables that are not necessarily independent or identically distributed and let $x_{(m)} = \max(x_1,\dots,x_m)$. Denote the mean and variance of each x_r by $\mathbb{E}[x_r] = \mu_r$ and $\sigma^2(x_r) = \sigma_r^2$ for $r = 1,\dots,m$. An upper bound for $\mathbb{E}[x_{(m)}]$ can be then defined as follows [2]

$$\mathbb{E}[x_{(m)}] \le \frac{\sum_{r=1}^m \mu_r}{m} + \sqrt{\frac{m-1}{m}\sum_{r=1}^m \Big[\sigma_r^2 + \Big(\mu_r - \frac{\sum_{k=1}^m \mu_k}{m}\Big)^2\Big]}$$

Hence, for a general case with $M \ge 1$, we have the following upper bound for $\sum_{i=1}^M \mathbb{E}[x_{(m)}]$:

$$\sum_{i=1}^M \left(\frac{\sum_{r=1}^{m_i}\mu_r}{m_i} + \sqrt{\frac{m_i-1}{m_i}\sum_{r=1}^{m_i}\Big[\sigma_r^2 + \Big(\mu_r - \frac{\sum_{k=1}^{m_i}\mu_k}{m_i}\Big)^2\Big]}\right)$$

Moreover, two other upper bounds for $\mathbb{E}[x_{(m_i)}]$ are given in [5]. Therefore, we can choose the smallest upper bound among these three, denote it by $\mathcal{U}\Big(\sum_{i=1}^M \mathbb{E}[x_{(m_i)}]\Big)$. Hence, an upper bound for (24) with $M \ge 1$, can be written as

$$\mathbb{E}[\sum_{i=1}^M \max(x_1,\dots,x_{m_i})] - \sum_{i=1}^M \max(\mathbb{E}[x_1],\dots,\mathbb{E}[x_{m_i}])$$
$$\le \mathcal{U}\Big(\mathbb{E}[x_{(m)}]\Big) - \sum_{i=1}^M \max(\mathbb{E}[x_1],\dots,\mathbb{E}[x_{m_i}]) \quad (25)$$

Hence, we have shown that the error of approximating $\tilde{J}$ by $\tilde{J}_{\text{up}}$, i.e., $\tilde{J}_{\text{up}}(u) - \tilde{J}(u)$, is bounded from above by (23)+(25).

5.1 Alternative upper bound for the approximation method

Since the proposed approximation method is also valid for distributions with a bounded domain, here we discuss this case separately to propose an alternative upper bound for the error caused by applying the upper bound approximation function (15). To this end, we consider the three inequalities in (14) and their corresponding error. For random variables $x_j,\ j=1,\dots,n$ with a *bounded* domain $\mathcal{X}_j$, Inequality (i) turns into an equality if all x_j are nonnegative. Hence, we introduce a constant $L = \min_{j=1,\dots,n}(\min \mathcal{X}_j, 0)$ and then

replace each x_j by $y_j = x_j - L,\ j = 1, \ldots, n$ and add L to the right-hand side of all inequalities in (14); in this way, the error due to (i) is zero. The error due to (ii) approaches zero if $p \to \infty$, since by definition $\|x\|_\infty = \lim_{p\to\infty} \|x\|_p$. However, a large p affects the error due to (iii), which is the error of Jensen's inequality, differently and we discuss it here in more detail.

For a differentiable, concave function f defined on an interval $[a, b]$, the absolute error of Jensen's inequality has the following upper bound [27]:

$$0 \leq f(\mathbb{E}[x]) - \mathbb{E}[f(x)] \leq \max_{\theta\in[0,1]} [f(\theta a + (1-\theta)b) - \theta f(a) - (1-\theta) f(b)] =: \bar{e}_{\text{abs}}(a,b) \tag{26}$$

and it has been shown in [27] that there exists a unique $\theta_0 \in (0, 1)$ for which $\bar{e}_{\text{abs}}(a, b)$ is maximal. In our case, the concave function f is given by $f(x) = x^{1/p}$ and we have $f'(x) = \frac{1}{p} x^{\frac{1}{p}-1}$. Since we assume that p is a positive even integer larger than or equal to 2, the argument x has to be larger than or equal to zero, which is the case since $x = \sum_{j=1}^{n} x_j^p$. Let us first consider the case where x is strictly positive and hence, $a, b > 0$. The case where $x = 0$ will be considered later on (see Propositions 10 and 11). By substituting f in (26) and by determining the optimal value of θ, the following expression for $\bar{e}_{\text{abs}}(a, b)$ is obtained:

$$\bar{e}_{\text{abs}}(a,b) = \Big(\frac{a-b}{p(a^{\frac{1}{p}} - b^{\frac{1}{p}})}\Big)^{\frac{1}{p-1}} - \left(\frac{1}{a-b}\left[(a^{\frac{1}{p}} - b^{\frac{1}{p}})\left(\frac{a-b}{p(a^{\frac{1}{p}} - b^{\frac{1}{p}})}\right)^{\frac{p}{p-1}} - a^{\frac{1}{p}} b + a b^{\frac{1}{p}}\right]\right) \tag{27}$$

Hence, we derive the following proposition based on the above formula. Since the proof of this proposition is straightforward, we skip it here.

Proposition 10 *Considering our assumptions that $a, b > 0$ and $p \geq 2$ is an even integer, we obtain the following result:*

$$\lim_{a\to 0^+} \bar{e}_{\text{abs}}(a,b) = b^{\frac{1}{p}} \left(\Big(\frac{1}{p}\Big)^{\frac{1}{p-1}} - \Big(\frac{1}{p}\Big)^{\frac{p}{p-1}} \right).$$

As we have assumed that $y_j = x_j - L = \xi_j + \gamma_j^T \omega - L$ has a probability distribution with a finite domain, a and b can be easily obtained. Indeed, assume that each independent element of the stochastic vector ω, i.e., $\omega_t,\ t = 1, \ldots, n_\omega$, belongs to the interval $[\underline{c}_t, \bar{c}_t]$ where $\underline{c}_t, \bar{c}_t \in \mathbb{R}$ and without loss of generality, we assume that $\underline{c}_t < \bar{c}_t$ for all t. Since γ_{jt} can be positive or negative, we have $\min(\gamma_{jt}\underline{c}_t, \gamma_{jt}\bar{c}_t) \leq \gamma_{jt}\omega_t \leq \max(\gamma_{jt}\underline{c}_t, \gamma_{jt}\bar{c}_t)$. Hence, we can show that each $y_j,\ j = 1, \ldots, n$, also belongs to the interval $[a_j, b_j]$ where $0 \leq a_j < b_j$ are defined as follows: since $x_j = \xi_j + \gamma_j^T \omega = \xi_j + \sum_{t=1}^{n_\omega} \gamma_{jt}\omega_t$, we have

$$\underbrace{\xi_j + \sum_{t=1}^{n_\omega} \min(\gamma_{jt}\underline{c}_t, \gamma_{jt}\bar{c}_t)}_{\bar{a}_j} \leq x_j \leq \underbrace{\xi_j + \sum_{t=1}^{n_\omega} \max(\gamma_{jt}\underline{c}_t, \gamma_{jt}\bar{c}_t)}_{\bar{b}_j}$$

Let $L = \min_j \bar{a}_j$ and define $a_j = \bar{a}_j - L$ and $b_j = \bar{b}_j - L$. Therefore, $a_j \leq y_j \leq b_j$. Note that by this choice of L, $0 \leq y_j, \forall j = 1, \ldots, n$, so $0 \leq a_j < b_j$ and $a_j^p < b_j^p$. Hence,

$$a_j^p \leq y_j^p \leq b_j^p \quad \Rightarrow \quad \underbrace{\sum_{j=1}^{n} a_j^p}_{a} \leq \underbrace{\sum_{j=1}^{n} y_j^p}_{x} \leq \underbrace{\sum_{j=1}^{n} b_j^p}_{b} \tag{28}$$

with $a < b$. Recall that the error caused by inequality (ii) in (14) approaches 0 as $p \to \infty$, which suggests that p should be selected very large. However, we need to investigate the effect of having a large p on $\bar{e}_{\text{abs}}(a, b)$, since both a and b depend on p in (28). To this end, let $\alpha = \max_{j=1,\ldots,n} a_j$ and $\beta = \max_{j=1,\ldots,n} b_j$. Denote the number of a_j values that are equal to α by A and the number of b_j values that are equal to β by B. Hence, $\beta > \alpha$ since $b > a$. Now, for a large p, we rewrite a and b as $a \approx A\alpha^p$ and $b \approx B\beta^p$. Using this notation, we have the following proposition.

Proposition 11 *Assuming that $a \approx A\alpha^p,\ b \approx B\beta^p$ for a large positive even integer p with A, B positive integers and $0 \leq \alpha < \beta$, we have $\lim_{p\to\infty} \bar{e}_{\text{abs}}(\alpha, \beta) = \beta$.*

The proof has been skipped as it is straightforward. This proposition shows that if $p \to \infty$, the absolute error converges to β, which indicates that for a large p, the error may become too large depending on the value of β. Consequently, there is a trade-off between having a small error in inequality (iii) by choosing a relatively small p and having a small error in inequality (ii) by choosing a very large p. So, the value of p has to be tuned accordingly.

6 Worked Example

In this example, we will illustrate that our approximation method works efficiently when it is applied to a reference tracking problem. As a specific case, we study control of the temperature of a room (see [22]) and we consider the model predictive control (MPC) optimization problem of an MMPS system. For more details on MPC for (stochastic) MMPS systems, the interested reader is referred to [11] and the references therein. The following continuous discrete-time PWA system is considered in [22]:

$$x(k+1) = \begin{cases} 1/2x(k) + u(k) + \omega_1(k) + 1 & \text{if } x(k) < 0 \\ u(k) + \omega_1(k) + 1 & \text{if } x(k) \geq 0 \end{cases}$$
$$y(k) = x(k) + \omega_2(k).$$

where x, y, and u denote the state (room temperature), measurement, and heat input, respectively, and ω_1 and ω_2 denote the disturbance. We also have the following input constraints for all k: $u(k) \geq 0$ and $-4 \leq \Delta u(k) = u(k+1) - u(k) \leq 4$. In [22], it is assumed that ω_1 and ω_2 belong to a bounded polyhedral set. However, to illustrate our approach, here we assume that the error components have a standard normal distribution, i.e., $\omega_1(k), \omega_2(k) \sim \mathcal{N}(0, 1)$. The equivalent MMPS representation of the above PWA system is

$$x(k+1) = \min(\frac{1}{2}x(k) + u(k) + \omega_1(k) + 1,\ u(k) + \omega_1(k) + 1)$$
$$y(k) = x(k) + \omega_2(k).$$

For the MPC setting (cf. [22]), the prediction horizon is $N_{\rm p} = 2$ and the control horizon is $N_{\rm c} = 2$ accordingly, $\tilde{y}(k) = [y(k+1), y(k+2)]^T$, $\tilde{r}(k) = [r(k+1), r(k+2)]^T$, and $\tilde{u}(k) = [u(k), u(k+1)]^T$. Let the uncertainty vector be $\omega(k) = [\omega_1(k), \omega_2(k+1)]^T$. Therefore, $\tilde{\omega}(k) = [\omega^T(k), \omega^T(k+1)]^T$. The objective function is defined as $\tilde{J}(k) = \mathbb{E}\big[\|\tilde{y}(k) - \tilde{r}(k)\|_\infty + \lambda \|\tilde{u}(k)\|_1\big]$, which is the expected value of an MMPS function. Here, we have the expected value due to the stochastic setting while in [22] the worst-case optimization problem was solved due to considering a bounded disturbance. We compute the closed-loop MPC control signal by minimizing the upper bound of $\tilde{J}(k)$ over the simulation period $[1, 20]$, with $\lambda = 0.01$, $x(0) = -6, u(-1) = 0$, and $p = 26$ in the approximation method (cf. (19)). The reference signal is given as $\{r(k)\}_{k=1}^{20} = \{-5, -5, -5, -5, -3, -3, 1, 3, 3, 8, 8, 8, 8, 10, 10, 7, 7, 7, 4, 3\}$. Figure 1 shows the results of the simulation in which we compare our proposed approach (cf. Section 4) with the "Exact solution" obtained by using numerical integration and with robust MPC (RMPC) in which the random variables have a truncated normal distribution using $3\sigma, 5\sigma$, and 7σ bounds, respectively. The simulations are done in MATLAB R2016b on a 2.6 GHz Intel Core i5 processor and the optimizations are solved using *fmincon* with the SQP solver.

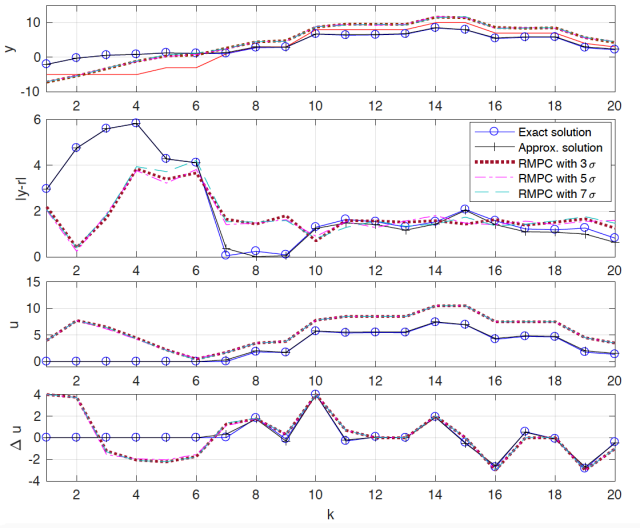

Fig. 1. Results of the stochastic and robust MPC-MMPS optimization problem. 'o'-line: exact solution; '+'-line: approximation approach; solid line without marker (first plot): reference signal; dotted line: robust MPC with 3σ bounded error; dash-dotted line: robust MPC with 5σ bounded error; dashed line: robust MPC with 7σ bounded error.

Since we have a stochastic system, we have repeated the simulations for each approach 100 times with different noise realizations and then report the mean and variance of the obtained trajectories. The top plot in Figure 1 shows the reference signal and the mean of the output trajectories of the system using the "Exact solution", using the approximation approach, and using the robust MPC approach with different error bounds; the second plot presents the mean of the tracking error using the mentioned approaches; the third plot shows the mean of the optimal input trajectories using the different mentioned approaches; the last plot illustrates the mean of $\Delta u(k)$ for each approach and the fact that the input constraint is satisfied, i.e., $-4 \leq \Delta u(k) \leq 4$. The maximum values for the variance of the output trajectories are as follows: "Exact solution": 2.5283, approximation approach: 2.4992, RMPC with 3σ: 5.2118, RMPC with 5σ: 4.9976, and RMPC with 7σ: 5.9361. The maximum values for the variance of the optimal input trajectories are: "Exact solution": 0.0742, approximation approach: 0.0687, RMPC with 3σ: 5.8396, RMPC with 5σ: 7.0336, and RMPC with 7σ: 6.7088. These values show that the variances in robust MPC are much larger compared to stochastic MPC, which results in more deviation in the obtained trajectories.

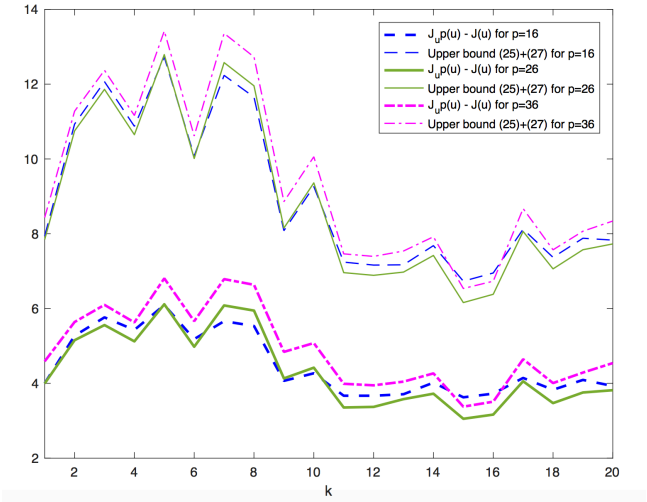

Fig. 2. The error $\tilde{J}_{\rm up}(u) - \tilde{J}(u)$ and its upper bound, i.e., sum of (23) and (25), for different values of p, applying (19).

As shown in Figure 1, the results of the approximation approach are very close to the ones from the "Exact solution". Moreover, the overall performance of robust MPC is worse than that of stochastic MPC; although the mean of the output trajectories in robust MPC in the first 4 time steps is closer to the reference signal, this changes in the next steps in favor of the stochastic approach using both the approximation method and the exact method. Furthermore, the third plot confirms the conservativeness of the robust MPC method as that approach results in much larger inputs, which results in more energy used (in terms of heat) compared to the stochastic approaches.

Furthermore, in the error analysis in Section 5, we have shown that the error of approximating $\tilde{J}$ by $\tilde{J}_{\rm up}$, i.e., $\tilde{J}_{\rm up}(u) - \tilde{J}(u)$, is bounded from above. Figure 2 shows this error and the obtained upper bound, which is the sum of the two upper bounds in (23) and (25), having different values for p in the approximation method (cf. (19)). As can be seen in this figure, $p = 26$ gives mostly a smaller error compared

to $p = 16$ and $p = 36$. This is due to Jensen's inequality used in obtaining the approximation function (cf. (14)), and hence, a trade-off has to be made in the choice of the p, as explained in Section 5.1. For more detail on the effect of p on the performance of the controller, the reader is referred to [13].

7 Conclusions

This paper has discussed an optimization problem of stochastic max-min-plus-scaling (MMPS) systems in which the objective function is defined as an expected value of stochastic MMPS functions. As the available numerical approaches for computing this expected value are both complex and time consuming, we have proposed an approximation method in which the objective function is replaced by its upper bound. We have shown that for distributions with closed-form moments, this upper bound has a closed-form expression and hence, can be computed analytically. This way, we have avoided the cumbersome numerical or analytic integrations needed for the calculation of the expected value. We have also shown that the error resulting from approximating the original objective function by its upper bound is bounded from above. In our ongoing and future research, we will investigate approaches to decrease the error of this approximation method by, e.g. considering truncated distributions or finding an appropriate L for distributions with an unbounded domain. Moreover, it is interesting to compare the proposed approximation method with more conservative approaches such as the one presented in [3].

References

[1] M. A. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. National Bureau of Standards, US Government Printing Office, Washington DC, 1964.

[2] B. C. Arnold and R. A. Groeneveld. Bounds on expectations of linear systematic statistics based on dependent samples. *Mathematics of Operations Research*, 4(4):441–447, 1979.

[3] A. Bemporad, F. Borrelli, and M. Morari. Min-max control of constrained uncertain discrete-time linear systems. *IEEE Transactions on Automatic Control*, 48(9):1600–1606, September 2003.

[4] A. Bemporad, G. Ferrari-Trecate, and M. Morari. Observability and controllability of piecewise affine and hybrid systems. *IEEE Transactions on Automatic Control*, 45(10):1864–1876, 2000.

[5] D. Bertsimas, K. Natarajan, and C. Teo. Tight bounds on expected order statistics. *Probability in Engineering and Information Systems*, 20(4):667–686, 2006.

[6] A. A. Borovkov. *Asymptotic Methods in Queueing Theory*. John Wiley & Sons, New York, NY, 1984.

[7] L. O. Chua and A. C. Deng. Canonical piecewise-linear representation. *IEEE Transactions on Circuits and Systems*, 35(1):101–111, January 1988.

[8] B. De Schutter and T. J. J. van den Boom. Model predictive control for max-min-plus-scaling systems — Efficient implementation. In M. Silva, A. Giua, and J.M. Colom, editors, *Proceedings of the 6th International Workshop on Discrete Event Systems (WODES'02)*, pages 343–348, Zaragoza, Spain, October 2002.

[9] B. De Schutter and T. J. J. van den Boom. MPC for continuous piecewise-affine systems. *Systems & Control Letters*, 52(3–4):179–192, July 2004.

[10] F. M. Dekking, C. Kraaikamp, H. P. Lopuhaä, and L. E. Meester. *A Modern Introduction to Probability and Statistics*. Springer, London, UK, 2005.

[11] S. S. Farahani, T. van den Boom, and B. De Schutter. Model predictive control for stochastic max-min-plus-scaling systems – An approximation approach. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 391–396, Orlando, USA, December 2011.

[12] S. S. Farahani, T. van den Boom, and B. De Schutter. Exact and approximate approaches to the identification of stochastic max-plus-linear systems. *Discrete Event Dynamic Systems: Theory and Applications*, 24(4):447–471, December 2014.

[13] S. S. Farahani, T. van den Boom, H. van der Weide, and B. De Schutter. An approximation method for computing the expected value of max-affine expressions. *European Journal of Control*, 27:17–27, January 2016.

[14] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.

[15] V. V. Gorokhovik and O. I. Zorko. Piecewise affine functions and polyhedral sets. *Optimization*, 31(3):209–221, 1994.

[16] J. Harris, J. L. Hirst, and M. Mossinghoff. *Combinatorics and Graph Theory*. Undergraduate Texts in Mathematics. Springer, New York, 2nd edition, 2008.

[17] W. P. M. H. Heemels, B. De Schutter, and A. Bemporad. Equivalence of hybrid dynamical models. *Automatica*, 37(7):1085–1091, July 2001.

[18] M. Johansson. *Piecewise Linear Control Systems*, volume 284 of *Lecture Notes in Control and Information Sciences*. Springer, Berlin, Germany, 2003.

[19] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Vol. 2*. Wiley, New York, 2nd edition, 1995.

[20] D. M. W. Leenaerts and W. M. G. van Bokhoven. *Piecewise Linear Modeling and Analysis*. Kluwer Academic Publishers, Boston, Massachusetts, 1998.

[21] W. M. McEneaney. Robust/H_∞ filtering for nonlinear systems. *Systems & Control Letters*, 33:315–325, 1998.

[22] I. Necoara, B. De Schutter, T. van den Boom, and H. Hellendoorn. Model predictive control for uncertain max-min-plus-scaling systems. *International Journal of Control*, 81(5):701–713, May 2008.

[23] G. J. Olsder, J. A. C. Resing, R. E. de Vries, M. S. Keane, and G. Hooghiemstra. Discrete event systems with stochastic processing times. *IEEE Transactions on Automatic Control*, 35(3):299–302, March 1990.

[24] S. Ovchinnikov. Max-min representation of piecewise linear functions. *Beiträge zur Algebra und Geometrie/Contributions to Algebra and Geometry*, 43(1):297–302, 2002.

[25] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, Singapore, 1991.

[26] S. M. Ross. *Probability models*. Academic Press, San Diego, CA, 8th edition, 2003.

[27] S. Simić. On an upper bound for Jensen's inequality. *Journal of Inequalities in Pure and Applied Mathematics*, 10(2), 2009. Article 60.

[28] J. E. Tierno, R. M. Murray, J. C. Doyle, and I. M. Gregory. Numerically efficient robustness analysis of trajectory tracking for nonlinear systems. *Journal of Guidance, Control, and Dynamics*, 20(4):640–647, 1997.

[29] Z. Wang, F. Yang, D.W.C. Ho, and X. Liu. Robust finite-horizon filtering for stochastic systems with missing measurements. *IEEE Signal Processing Letters*, 12(6):437–440, 2005.

[30] W. Whitt. Queues with service times and interarrival times depending linearly and randomly upon waiting times. *Journal of Queueing Systems*, 6(1):335–351, December 1990.