Robles, R.; Sala, A.; Bernal, M. (2019). Performance-oriented quasi-LPV modeling of nonlinear systems. International Journal of Robust and Nonlinear Control. 29(5):1230-1248. https://doi.org/10.1002/rnc.4444

Performance-Oriented Quasi-LPV Modelling of Nonlinear Systems

Ruben Robles[1], Antonio Sala[11] and Miguel Bernal[2]

[1] *Instituto U. de Automática e Informática Industrial, Universitat Politècnica de Valǹcia, Camino de Vera S/N, E-46022 Valencia, Spain*

[2]*Instituto Tecnológico de Sonora, 5 de Febrero 818 Sur, 85000, Ciudad Obregón, Mexico*

SUMMARY

Polytopic quasi-LPV models of nonlinear processes allow using linear matrix inequalities (LMI) to guarantee some performance goal on them (in most cases, locally, over a so-called modelling region). In order to get a finite number of LMIs, nonlinearities are embedded on the convex hull of a finite set of linear models. However, for a given system, the quasi-LPV representations are not unique, yielding different performance bounds depending on the model choice. To avoid such drawback, earlier literature on the topic used annihilator-based approaches, which require gridding on the modelling region, and non-convex BMI conditions for controller synthesis: optimal performance bounds are obtained, but with a huge computational burden. This paper proposes building a model by minimising the projection of the nonlinearities onto directions which are deleterious for performance. For a small modelling region, these directions are obtained from LMIs with the linearised model. Additionally, these directions will guide the selection of the polytopic embedding's vertices. The procedure allows gridding-free LMI controller synthesis, as in standard LPV setups, with a very reduced performance loss with respect to the above BMI+gridding approaches, at a fraction of the computational cost.

Received . . .

KEY WORDS: Quasi-LPV systems, Takagi-Sugeno systems, Linear Matrix Inequalities, Robust Control, linear-parameter-varying systems, gain scheduling

1. INTRODUCTION

Many smooth nonlinear systems can be embedded into a linear parameter-varying (LPV) dynamics, resulting in the so-called *quasi*-LPV [33, 19] or Takagi-Sugeno (TS) [45, 17, 23] systems; both denominations will be considered equivalent in the context in this paper [32], although the TS denomination is usually associated to polytopic bounding and convex interpolation expressions.

[1]Correspondence to: Instituto U. de Automática e Informática Industrial (AI[2]), Universitat Politècnica de València, Camino de Vera S/N, 46022, Valencia, Spain. E-mail: asala@isa.upv.es,

0

Contract/grant sponsor: Spanish ministry of Economy and European Union (AEI/FEDER, UE), grant DPI2016-81002-R; Generalitat Valenciana, grant GRISOLIA/2014/006; ITSON PFCE 2017-18; PROFAPI CA-18 ITSON 2018-1062

Efficient linear matrix inequality (LMI) techniques for LPV systems have been developed in order to analyse their properties or synthesise gain-scheduled controllers for them, see the above-cited works, [2, 40, 16, 26] and references therein, albeit with some sources of conservatism [36].

Alternatively, nonlinearities can also be embedded into the convex hull of polynomials [34, 9], amenable to convex sum-of-squares optimisation [28], but these approaches are intentionally out of the scope of this work. Identification-based LPV models [46] will also be left out of the present discussion.

Rewriting a nonlinear differential equation as a quasi-LPV model consists in decomposing a $\mathcal{C}^1$ nonlinearity, present in a first-principle model, say $\rho(z)$, as $\rho(z) = Q(z)z$, see [33, 45, 30] and references therein, being z a linear function of exogenous or state variables. The knowledge about the relationship between the factors $Q(z)$ and z is usually disregarded, so $Q(z(t))$ is considered as an uncertain linear time-varying element $\Delta(t)$ (assumed measurable in gain-scheduled designs [3, 33]).

Subsequently, some LMIs involving $Q(z)$ must hold for all z in a so-called modelling region Ω_z. If feasible LMI decision variables are found, system properties will be proven inside some bounded invariant set (usually a Lyapunov level set in the undisturbed case, or an inescapable set if disturbances are present and amplitude or integral bounds for them are known [38]). As the above problem would require solving an infinite set of LMIs, in order to get computable results, either evaluating $Q(z)$ in a dense enough grid [49, 19, 35] or obtaining a polytopic bound with a finite number of vertices [45, 20] are routinely used in applications.

Note that, in some LPV developments, stability conditions are relaxed via *a priori* bounds on the time-derivatives of the arguments of Q, see [13, 19, 48]. In the quasi-LPV setup, this might be justified in stability analysis if bounding $\dot{z}$ is possible; however, in controller synthesis, $\dot{z}$ may depend on the to-be-designed control u, hence, such derivative bound assumption may imply *a posteriori* checks of their fulfillment (for instance, by simulation [18, Chap. 4]). In some cases, the time-derivative components may be cancelled [15]. Actually, conditions on the gradient on the polytopic interpolation coefficients [21, 4] seem more natural in the quasi-LPV context than plain time-derivative bounds. Nevertheless, derivative-related issues will not be considered in the scope of the present work.

The main motivation of this work is the fact that the above decomposition $\rho(z) = Q(z)z$ is, in general, not unique. For instance, in $\rho(z) = \sin(z_1 z_2)$ both $(\sin(z_1 z_2)/z_1 \ 0)$ and $(0 \ \sin(z_1 z_2)/z_2)$ can be equally used as two alternative options for $Q(z)$. However, different performance results for the ensuing analysis or control design might be proved for each choice. This is one of the several causes of the inherent conservatism of the quasi-LPV/TS approach with respect to an ideal nonlinear controller [36]. Actually, there are infinitely many of options for $Q(z)$: the use of annihilators $N(z)z = 0$, such that $\rho(z) = (Q(z) + N(z))z$, has been proposed in [19, 8, 11, 47], where decision variables in $N(z)$ are used to diminish conservatism; however, most results require gridding and are BMI in control synthesis problems, so these approaches will be not pursued here. An alternative option is a selection of the optimal factorisation in order to diminish the "spread" of $Q(z)$ over some directions; these idea is exploited in [8], and in the works by the authors' team [29] and [30], which present preliminary approaches to the developments here. Annihilator and minimum-spread options will be briefly reviewed in next section.

DOI: 10.1002/rnc

The objective of this work is reversing the usual approach on the quasi-LPV framework: instead of writing LMIs for a given model, it assumes that LMIs for a given performance optimisation problem are already available and sets out for a quest to determine the optimal model that keeps the proven performance as close as possible to the linearised one (indeed, as the linearisation is included in the LPV embedding, the proven performance cannot be better than the linearised one). A general subspace-based method is presented, determining a worst-case matrix perturbation to the linearised LMIs so the projection of $Q(z)$ onto it should be minimised.

The structure of this paper is as follows: section 2 introduces preliminaries and motivates the problem; section 3 presents the main result of this paper and the assumptions involved. Examples and conclusions appear in sections 4 and 5, respectively. For readability, some ideas and notation are presented in an appendix.

2. PRELIMINARIES AND PROBLEM STATEMENT

Let us consider a class of nonlinear dynamic systems expressed as:

$$\dot{x} = Ax + M_x v + Bu + Ew \tag{1}$$

$$z = C_z x + D_z u + F_z w \tag{2}$$

$$y = C_y x + M_y v + D_y u + F_y w \tag{3}$$

$$v = \rho(z) \tag{4}$$

where $x \in \mathbb{R}^n$ is the vector of state variables, $u \in \mathbb{R}^q$ has the manipulated input variables, $w \in \mathbb{R}^p$ is a disturbance input vector, and $y \in \mathbb{R}^r$ gathers the output variables; accordingly, matrices A, M_x, B, E, C_z, D_z, F_z, C_y, M_y, D_y, and F_y are of adequate size. Additionally, $v \in \mathbb{R}^m$ is a dummy variable representing the output of the nonlinearities $\rho(z)$, being their input $z \in \mathbb{R}^s$ assumed to belong to a known *modelling region* $\Omega_z \subset \mathbb{R}^s$, with $0 \in \Omega_z$, and $\rho(z)$ fulfills the following assumption:

Assumption 1. *The function $\rho : \Omega_z \mapsto \mathbb{R}^m$ is a vector of continuously differentiable nonlinear functions, i.e., $\rho \in \mathcal{C}^1$, with $\rho(0) = 0$. Also, without loss of generality, we will assume that ρ has null Jacobian.*

Under the above assumption, $x = 0$ is an equilibrium point of the above system for $u = w = 0$ and the classical Jacobian linearisation of (1)–(4) is obtained replacing (4) by $v = 0$.

Note that, if ρ had non-zero Jacobian, denoting it as $H := \frac{\partial \rho}{\partial z}$, we can write $\rho = Hz + \tilde{\rho}(z)$, with $\tilde{\rho}(z) := \rho(z) - Hz$ having, evidently, null Jacobian. Replacing M_x and M_y by $\tilde{M}_x = M_x + H$ and $\tilde{M}_y = M_y + H$, respectively, and using $\tilde{\rho}$ instead of ρ in (4), a model (1)–(4) fulfilling Assumption 1 will be readily obtained.

From Taylor's theorem (Peano form), the following result is well known [19, Lemma 3.1]:

Proposition 1. *Under Assumption 1, there exists a factorisation*

$$\rho(z) = Q(z) \cdot z \tag{5}$$

with $Q(z)$ being a continuous matrix-valued function $Q : \Omega_z \mapsto \mathbb{R}^{m \times s}$, with $Q(0) = 0$.

DOI: 10.1002/rnc

Finding $Q(z)$ is trivial in the case $s = 1$, as the only option for $Q(z)$ is $Q(z) = \rho(z)/z$; abusing the notation, for $z = 0$ the quotient expression should be understood as the limit when $z \to 0$, guaranteed to exist by the assumptions. For $\rho : \mathbb{R}^2 \mapsto \mathbb{R}$, the quotient idea is generalised by expressing $\rho(z_1, z_2) = \zeta_1(z) + \rho(0, z_2)$, with $\zeta_1(z) = \rho(z_1, z_2) - \rho(0, z_2)$. Proposition 1 ensures the continuity at $z_1 = 0$ of $q_1(z) := \zeta_1(z)/z_1$, and that of $q_2(z) = \rho(0, z_2)/z_2$ at $z_2 = 0$; subsequently, $\rho(z) = q_1(z)z_1 + q_2(z)z_2$. The idea is generalised to more than two input variables in the Appendix of this work.

Nevertheless, the above decomposition is not unique if the dimension of z is equal or larger than two, as infinitely many $\aleph(z)$ such that $\aleph(z)z = 0$ can be found [8, 19] so that given $\rho(z) = Q_0(z)z$, the decomposition below is still valid:

$$\rho(z) = (Q_0(z) + \aleph(z)) \cdot z \tag{6}$$

As an example, with $s = 2$, $\aleph(z) := \zeta(z)(-z_2 \;\; z_1)$ would be a valid choice for any continuous $\zeta(z)$.

Once a decomposition (5) is chosen, the representation $\{(1), (2), (3), (5)\}$ is named in the literature as quasi-LPV system [19, 6], because the original nonlinear dynamics can be embedded into the LPV system conformed by equations (1)–(3) plus

$$v = \Delta(t) \cdot z \tag{7}$$

defining the time-varying quantity $\Delta(t)$ to be $\Delta(t) := Q(z(t))$.

Note that, in most works in literature, the LMIs arising from posing control problems using the LPV model arising from (1)–(3) and (7) disregard the actual relationship between $\Delta \equiv Q$ and z, i.e., the LMIs do not exploit the explicit "shape" of $Q(z(t))$, and use only *bounds* (vertex models) or *samples* of Δ. This, of course, introduces the so-called *shape-independence* conservatism [36] with respect to an ideal nonlinear control using the *explicit expressions* in $Q(z)$. Limited approaches to shape-dependent results appear in, for instance, [37, 5]; the use of gradient information, discussed in the introduction, may also be considered a shape-dependent approach.

Shape-independence conservatism is the reason on why different choices of $Q(z)$ influence the final LPV obtained performance, and the main motivation of the work here presented. Of course, (7) and (5) are the same, but we use notation Δ to emphasize its use in conservative shape-independent LPV analysis, whereas using $Q(z)$ introduced in (5) emphasizes the initial non-conservative rewriting of (4); thus, Δ and Q can be exchanged at will.

As discussed in the introduction, in order to get a finite number of LMI conditions, either a dense enough grid [22] of values $Q(z_k)$, $z_k \in \Omega_z$, or a polytopic embedding [1] are pursued in many cases. The most straightforward polytopic bounding would be obtaining the $2^{m \times s}$ vertices from the infimum and supremum of each element of matrix $Q(z)$ when z ranges in Ω_z [45]; for instance, these vertices are guaranteed to exist if Ω_z is compact.

If the linear (time-varying) expression (7) is substituted on (1)–(3), a representation:

$$\dot{x} = \tilde{A}(\Delta)x + \tilde{B}(\Delta)u + \tilde{E}(\Delta)w \tag{8}$$

$$y = \tilde{C}(\Delta)x + \tilde{D}(\Delta)u + \tilde{F}(\Delta)w \tag{9}$$

DOI: 10.1002/rnc

can be obtained where system matrices $\tilde{A}(\Delta) := A + M_x \Delta C_z$, $\tilde{B}(\Delta) = B + M_x \Delta D_z, \ldots$ are *affine* in Δ (i.e., affine in $Q(z)$ in the quasi-LPV setup). Quasi-LPV systems (8)–(9) with polytopic bounds for Δ are also denoted as Takagi-Sugeno ones in literature [45, 17, 36, 32]. If (2) were replaced by $z = C_z x + D_z u + M_z v + F_z w$ the resulting model (8)–(9) would have *rational* dependence on Δ, yielding linear-fractional representations [27]. For instance, $\tilde{A}_{LFT}(\Delta) := A + M_x \Delta (1 - M_z \Delta)^{-1} C_z$, etc. Affine dependence on Δ will be later required in some of the results; nevertheless, the 1st-order approximation for small Δ is $\tilde{A}_{LFT}(\Delta) \approx A + M_x \Delta C_z$ so the results will apply to LFT models for small enough Ω_z, as these will yield a small Q (i.e., small Δ) from the continuity assumptions.

Annihilator approaches in prior literature. Options for selecting $\aleph(z)$ in (6) have been proposed in [19, 8, 11, 47], where $\aleph(z)$ is denoted as *annihilator*. Note that, if $\aleph(z) z = 0$, we have $R(z)\aleph(z)z = 0$ for any matrix function $R(z)$, so elements in $R(z)$ can be decision variables. For instance, the annihilator approach, in a state-feedback control synthesis problem pursuing decay rate ς, translates into finding decision variables in $R(z)$, as well as matrices X and F, such that:

$$X \succeq I, \qquad -\left(A + M_x Q_0(z) + M_x R(z) \aleph(z)\right) X - BF - \varsigma X - (*) \succeq 0 \qquad \forall\, z \in \Omega_z \tag{10}$$

where $(*)$ denotes the transpose needed to conform a symmetric matrix, and $\succeq$ denotes positive semidefiniteness. $C_z = I$, $D_z = F_z = 0$ has also been assumed to write (10).

However, the above expression hints the main drawbacks of the annihilator proposal:

- Non-convex controller synthesis problems, due to the products between decision variables in $R(z)$ and those in X.
- Need of setting a grid on Ω_z to approximate (10) by a finite set of constraints. Then, there is the risk of not hitting close enough to the "worst-case" point[2].

Due to the above computational concerns, the annihilator approach will not be pursued in this work; only some comparisons with it will be made in the example section.

Minimum-spread approaches. In [8] stability analysis problems were discussed and, denoting as $\mathcal{A}(z) = A + M_x R(z) \aleph(z)$, the authors proposed minimizing $\lambda := \min \max_{\psi, \phi \in \Omega_z} \bar{\sigma}(\mathcal{A}(\psi) - \mathcal{A}(\phi))$ where $\bar{\sigma}$ denotes the maximum singular value; as a variation, they also proposed minimizing $\bar{\sigma}\left(\frac{\partial \mathcal{A}}{\partial z}\right)$. However, in a generic case, gridding in the parameter space is needed to carry out such minimizations. In [30], a Hessian based minimum-spread result is suggested to solve a problem similar to the above minimization of λ, avoiding gridding, which can be proved exact for quadratic nonlinearities with a single output. The caveat of these approaches is that they are *problem-independent*: minimising the worst-case uncertainty induced in $\mathcal{A}$ by the choice of Q may be conservative if a particular problem's performance is not sensitive in that worst-case direction. This motivates the present contribution: a performance optimisation problem will be defined next, and the problem to be addressed will be obtaining the best performance-oriented quasi-LPV model for it, tailoring the choice of $Q(z)$ to such specific performance optimization setup.

[2] Gridding can be avoided, with polytopic Ω_z, on some classes of polynomial systems, setting constant R and affine $\aleph$ [11, 47]. That makes the approach related to the sum-of-squares techniques in [28], as pinpointed in [47]. Anyway, the relationship with polynomial approaches is intentionally left out of the scope of this paper, restricted to LPV/TS settings.

DOI: 10.1002/rnc

State-subspace approach. In our work [30, Prop.1,$VI.B], a quasi-LPV modelling idea was presented, based on analyzing the null space of the optimal solution, for a single Lyapunov inequality, for a linear system, in the form $x^T\Upsilon(A,D,\gamma)x \geq 0$, being D some decision variables.

This paper generalises the above problem to arbitrary inequalities $\zeta^T\mathcal{MI}(\cdot)\zeta \geq 0$ (arising from Schur, Finsler or other widely used manipulations) so that the assumption that $\zeta \equiv x$ is not needed, and multiple inequalities can be considered.

Let us, thus, propose a generic performance optimisation problem for (1)–(4) and state the quasi-LPV modelling problem to be addressed in this manuscript.

2.1. Performance optimisation problem

Under the conditions discussed above, considering that $\Delta(t) := Q(z(t))$ in the usual LPV analysis of nonlinear systems, some matrix inequalities (actually, in many cases, LMIs) can be asserted on the model matrices arising from (8)–(9); for convenience, we will in the sequel denote the set of model matrices as $L(Q) := \{\tilde{A}(Q), \tilde{B}(Q), \tilde{C}(Q), \tilde{D}(Q), \tilde{E}(Q), \tilde{F}(Q)\}$. The structure of problem to be solved over the original nonlinear system is in the form given by the assumption below:

Assumption 2. *The pursued objective over the model* (1)–(4)*, whose LPV embedding is* (8)–(9)*, is the optimisation of a performance measure γ subject to some constraints in the form of a matrix inequality, i.e., obtaining γ^{LPV} below:*

$$\begin{aligned} \gamma^{LPV} := \inf_{D_1, D_2(\cdot)} \quad & \gamma, \\ \textit{subject to} \quad & \mathcal{MI}(L(Q(z)), \{D_1, D_2(Q(z))\}, \gamma) \succeq 0 \; \forall z \in \Omega_z \end{aligned} \tag{11}$$

where $\{D_1, D_2(Q(z))\}$ are the decision variables (usually some matrix elements), with D_1 denoting a set of decision variables which cannot depend on z (usually matrix variables associated to a Lyapunov function), and $D_2(Q(z))$ are decision variables which can depend on $Q(z)$ (for instance, those related to controller gains in gain-scheduled setups). $\mathcal{MI}(\cdot,\cdot,\cdot) \succeq 0$ will be understood as $\mathcal{MI}$ being positive semi-definite, with $\mathcal{MI}(\cdot,\cdot,\cdot)$ assumed to be a continuous matrix-valued expression, monotonic in γ, i.e., for fixed L, D_1 and D_2, condition $\gamma_1 \geq \gamma_2$ implies $\mathcal{MI}(L, \{D_1, D_2\}, \gamma_1) \succeq \mathcal{MI}(L, \{D_1, D_2\}, \gamma_2)$. The modelling region Ω_z will be assumed to be the s-dimensional unit ball around the origin (if Ω_z were an hyperellipsoid, linear changes of variable would trivially transform it to the unit ball), i.e., $\Omega_z := \{z \in \mathbb{R}^s : z^T z \leq 1\}$.

Of course, problems stated as multiple matrix inequalities will be assumed to be equivalently cast as a single block-diagonal one. Note, for instance, that (10) is a particular example of the generic problem stated in Assumption 2 if minimization of $\gamma := -\varsigma$ is pursued (decay-rate maximization). In a robust control case, $D_1 := \{X, F\}$, in a gain-scheduled case F would be replaced by $F(Q)$ and $D_1 := \{X\}$, $D_2(Q) := \{F(Q)\}$. Shape-independence conservatism arises because, for instance, matrix inequalities $\mathcal{MI}(Q, X, F) \succeq 0$ in (10) are obtained from scalar ones $x^T\mathcal{MI}(Q(x), X, F)x \geq 0$, being the former inequalities a sufficient condition for the latter scalar ones (but not necessary).

DOI: 10.1002/rnc

Modelling aspects of robust versus gain-scheduled control problems. In stability analysis or robust linear control (i.e., assuming u independent of $Q(z(t))$) there is no need of assuming the values of $Q(z(t))$ being explicitly known at any time instant [41].

On the other hand, let us now consider a gain-scheduled controller, say $u(Q(z), x)$. Given the fact that, from (2), z is a function of three arguments, $z(x, u, w)$, the factorisation (5) must forcedly be set up so that unmeasurable components of disturbance vector w are not present in $Q(z)$ to enable actual implementation with available sensors. Also, if z depended on u, then an algebraic loop may appear, needing such loop to be solved at each instant of time in order to compute u (and well-posedness conditions would need to be added), which is quite inconvenient in implementation.

Hence, to avoid these two implementation issues, in gain-scheduled designs, (5) must be in the form $\rho(z) = Q(x, \xi)z$ being ξ the measurable elements of w, if any. For instance, a nonlinearity $v = (1 + x^2)u$ will be forcedly factored with $Q = (1 + x^2)$. As there is no freedom in choosing such "pinned" components of Q, the discussion is this work will be centered on giving good options for the unrestricted elements of Q, so pinned components of Q (i.e., factors multiplying the input) will not be considered any further. Note, too, that there are intermediate options in literature regarding imperfect knowledge of scheduling parameters [39], uncertain vertex models [31], or sampled-data setups with knowledge of scheduling parameters only at sampling instants [25]; nevertheless, these possible extensions will not be considered in the scope of this paper for brevity.

2.2. Problem statement

Due to the non-uniqueness of the decomposition (5), different performance results γ^{LPV} can be obtained for different choices of Q. The main goal of this paper is providing guidelines on how to select a suitable $Q(z)$ so that "performance loss" with respect to the linear case (i.e., the obtained performance when posing (11) with just the linearised model) is reduced, at least when the modelling region Ω_z is small. Also, as a by-product, some worst-performance directions are obtained; these directions can, too, contribute to minimising the performance loss when using polytopic embeddings (projection of $Q(z)$ over them must be minimised).

3. MAIN RESULT

Consider the linearisation of (1)–(4) at the origin, i.e.:

$$\dot{x} = Ax + Bu + Ew \tag{12}$$

$$y = C_y x + D_y u + F_y w \tag{13}$$

The restriction to $Q(0)=0$ of the problem stated in Assumption 2 gives rise to the following definition:

Definition 1 (linearised performance)**.** The linearised performance γ^{lin} is defined to be the result of the optimisation of a performance measure γ subject to constraints in the form of matrix inequalities

DOI: 10.1002/rnc

in decision variables D_1 and D_2, obtaining γ^{lin} below:

$$\begin{aligned} \gamma^{\text{lin}} := \quad & \inf_{D_1, D_2} \quad \gamma, \\ & \text{subject to} \quad \mathcal{MI}(L(0), \{D_1, D_2\}, \gamma) \succeq 0 \end{aligned} \tag{14}$$

being the above constraints $\mathcal{MI}(\cdot, \cdot, \cdot)$ the same as in Assumption 2, and $L(0)$ the set of constant model matrices in (12)–(13).

Note that, actually, the above definition is the particularization of Assumption 2 for $\Omega_z := \{0\}$; with a slight abuse of notation, D_2 in (14) should be understood as $D_2(0)$ in (11) .

Proposition 2. *For any choice of $Q(z)$ fulfilling* (5)*, the proven performance γ^{LPV} in the setting in Assumption 2 will be worse than or equal to the linearised performance γ^{lin}, i.e., $\gamma^{LPV} \geq \gamma^{lin}$.*

Proof
As the modelling region includes the origin, and $Q(0) = 0$ from Proposition 1, the constraint inequalities in (11) must hold for $L(0)$ and $D_2(0)$ and, obviously, also for the rest of values of $z \in \Omega_z$. However, (14) is the particularisation of (11) for the single point $z = 0$; thus, the minimisation over a single point will achieve an equal or lower optimal γ than the said minimisation over the whole ball Ω_z. □

The next assumption formally states the intuition that, in order to get meaningful LPV solutions of (11), the linearised solution must be well-posed.

Assumption 3. *We will assume that the above linearised performance problem is feasible and that the objective function is bounded from below ($\gamma^{lin} > -\infty$). The decision variable values achieving γ^{lin}, i.e., $\arg\inf_{D_1, D_2} \gamma$, will be denoted by $\{D_1^{lin}, D_2^{lin}\}$ (or any arbitrary selection of them if non-unique). By assumption, too, conditions $\mathcal{MI}$ will be such that the linearised problem has a finite solution $\{D_1^{lin}, D_2^{lin}\}$ achieving optimality.*

Let us denote the null space of the optimal solution as:

$$\mathcal{C} := \text{null}(\mathcal{MI}(L(0), \{D_1^{\text{lin}}, D_2^{\text{lin}}\}, \gamma^{\text{lin}})) \tag{15}$$

Proposition 3. *$\mathcal{C}$ is not empty, i.e., there exist ζ such that $\zeta^T \mathcal{MI}(L(0), \{D_1^{lin}, D_2^{lin}\}, \gamma^{lin})\zeta = 0$.*

Proof
Let us assume that the optimal γ^{lin} fulfills $\mathcal{MI}(L(0), \{D_1^{lin}, D_2^{lin}\}, \gamma^{lin}) \succ 0$. Then, by continuity of $\mathcal{MI}$ and, too, continuity of the sorted eigenvalues of a matrix with respect to its coefficients [7], there would exist an interval $\mathcal{I} := (\gamma_1, \gamma_2)$, with $\gamma_1 < \gamma^{lin} < \gamma_2$ fulfilling $\mathcal{MI}(L(0), \{D_1^{lin}, D_2^{lin}\}, \gamma) \succ 0$ for all $\gamma \in \mathcal{I}$. Monotonicity of $\mathcal{MI}$ would entail that $\mathcal{MI}(L(0), \{D_1^{lin}, D_2^{lin}\}, \gamma) \leq \mathcal{MI}(L(0), \{D_1^{lin}, D_2^{lin}\}, \gamma^{lin})$, for $\gamma < \gamma^{lin}$ and given that the constraints for $\gamma \in \mathcal{I}$ are feasible, this contradicts the assumption that γ^{lin} was the optimal solution, because such optimal solution must forcedly be lower than or equal to γ_1. Thus, the assumption that $\mathcal{MI}(L(0), \{D_1^{lin}, D_2^{lin}\}, \gamma^{lin}) \succ 0$ cannot be true. □

DOI: 10.1002/rnc

Now, let us consider the nonlinear model (1)–(4), and consider expressing $\rho(z) = Q(z)z$ with

$$Q(z) = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ q_{m1} & q_{m2} & \cdots & q_{ms} \end{pmatrix} \in \mathbb{R}^{m \times s} \tag{16}$$

For notational brevity, dependence on z of each element q_{ij} has been omitted. Now, the problem in in Assumption 2 will be equivalently recast as:

$$\begin{aligned} \gamma^{\text{LPV}} = \inf_{D_1, D_2(\cdot)} \quad & \gamma, \\ \text{subject to} \quad & \mathcal{MI}(L(Q), \{D_1, D_2(Q)\}, \gamma) \succeq 0 \quad \forall Q \in Q(\Omega_z) \end{aligned} \tag{17}$$

where, with a slight abuse of notation, we define $Q(\Omega_z)$ as

$$Q(\Omega_z) := \{\xi \in \mathbb{R}^{m \times s} | \exists z \in \Omega_z \; s.t. \; \xi = Q(z)\}. \tag{18}$$

As earlier discussed, evidently $\gamma^{\text{LPV}} \geq \gamma^{\text{lin}}$. Now, let us assert additional assumptions for the conditions (17), i.e., (11), which will allow us to bound $\gamma^{\text{LPV}} - \gamma^{\text{lin}}$ for small Ω_z. On the sequel, notation $\| \cdot \|_F$ will denote the Frobenius norm of a matrix.

Now, keeping the linearised solution for D_1, i.e., D_1^{lin} and introducing an incremental notation, $\gamma := \gamma^{lin} + \Delta\gamma$, $D_2 = D_2^{lin} + \Delta D_2$, a conservative version of problem (17) may be written as:

$$\Delta\gamma^* := \sup_{Q \in Q(\Omega_z)} \Delta\bar{\gamma}(Q) \tag{19}$$

where

$$\Delta\bar{\gamma}(Q) := \inf_{\Delta D_2} \quad \Delta\gamma \tag{20}$$

$$\text{subject to } \mathcal{MI}(L(Q), \{D_1^{\text{lin}}, D_2^{\text{lin}} + \Delta D_2\}, \gamma^{\text{lin}} + \Delta\gamma) \succeq 0 \tag{21}$$

where $\Delta\bar{\gamma}(Q)$ is the performance loss (or gain, if negative) if system matrices were modified to a particular value of Q (keeping D_1^{lin}); indeed, note that the for-all clause in (17) is absent in (21), as it has moved to (19).

Obviously, a different ΔD_2 for each Q can be used in (21), analogously to (17). Conservatism in the result of (19) arises from fixing D_1 to the linearised solution, instead of searching for a common D_1 in (17), which, obviously, was not necessarily equal to D_1^{lin}.

As a result of the above discussion, we can assert that $\gamma^{\text{LPV}} \leq \gamma^{\text{lin}} + \Delta\gamma^*$. In exchange for such conservatism, matrix inequality (21) can be, conceptually, *independently* optimized for each $Q \in Q(\Omega_z)$, motivating the definition of $\Delta\bar{\gamma}(Q)$ in (20); this was not possible for the set of inequalities in (17), forcedly solved as a whole.

Assumption 4. *We will consider $\mathcal{MI}$ to be:*

1. *(Jointly) affine* [3] *in arguments (Q, D_2, γ),*

[3] As earlier discussed, rational LFT expressions can be approximated to affine ones for small Ω_z, yielding small Q by continuity. In an analogous way, in order to approximately fulfill Assumption 4.1 in gain-scheduled cases where $B(Q)u(Q, x)$ appears, for small Q we can linearise $B(Q)u(Q, x) \approx B(0)(u(Q, x) - u(0, x)) + B(Q)u(0, x)$.

DOI: 10.1002/rnc

2. *The linearised solution is numerically robust, in the sense that for any $\varepsilon > 0$, there exists $\delta > 0$ such that for any Q such that $\|Q\|_F \leq \delta$, condition* (21) *is feasible and the optimal $\Delta\bar{\gamma}(Q)$ and its associated ΔD_2 in* (20) *fulfil $\|\Delta D_2\|_F \leq \varepsilon$, $|\Delta\bar{\gamma}(Q)| \leq \varepsilon$.*

For instance, the robust decay-rate problem (10) does fulfill Assumption 4.1, as well as its gain-scheduled version using $D_2(Q) \equiv F(Q)$. If the linearised system is fully controllable, some extra constraints are needed to fulfill assumptions 3 and 4.2 (otherwise, decay rate can be made infinitely fast), such as well-knwon constraints on control input amplitude, or pole-region ones (details left to the reader).

Let H be a matrix whose vectors form a basis of the subspace $\mathcal{C}$ defined in (15), guaranteed to exist by Proposition 3. Let $H_\perp$ be a basis of the orthogonal subspace to $\mathcal{C}$. Hence, considering the optimal conditions for the linearised model, introducing notation in incremental variables and a congruence matrix, we can write an equivalent condition to (21), in the form:

$$\begin{pmatrix} H_\perp^T \\ H^T \end{pmatrix} \mathcal{MI}(L(Q), \{D_1^{\text{lin}}, D_2^{\text{lin}} + \Delta D_2\}, \gamma^{\text{lin}} + \Delta\gamma) \begin{pmatrix} H_\perp & H \end{pmatrix} \\ = \begin{pmatrix} \Xi_{11}(Q, \Delta D_2, \Delta\gamma) & \Xi_{12}(Q, \Delta D_2, \Delta\gamma) \\ \Xi_{12}(Q, \Delta D_2, \Delta\gamma)^T & \mathcal{W}(Q, \Delta D_2, \Delta\gamma) \end{pmatrix} \succeq 0 \quad (22)$$

where dependence of submatrices Ξ_{11}, Ξ_{12} and $\mathcal{W}$ on D_1^{lin}, D_2^{lin} and γ^{lin} has been omitted from the notation because they will be considered fixed in the sequel. Now, given the definitions of H and $H_\perp$, we have: $\Xi_{11}(0,0,0) \succ 0$, $\Xi_{12}(0,0,0) = 0$, $\mathcal{W}(0,0,0) = 0$.

The sorted eigenvalues of a matrix are Lipschitz continuous with respect to its elements [7]; therefore, there exists ε such that for all $\|Q\|_F \leq \varepsilon$, $\|\Delta D_2\|_F \leq \varepsilon$, $|\Delta\gamma| \leq \varepsilon$ the inequality $\Xi_{11}(Q, \Delta D_2, \Delta\gamma) \succ 0$ holds. Now, from Assumption 4.(2), for all $\|Q\|_F \leq \min(\varepsilon, \delta)$, the LMIs (22), equivalent to (21), are feasible, and the optimal solution $\Delta\bar{\gamma}(Q)$ in (20) is attained with $\|\Delta D_2\|_F \leq \varepsilon$, $\Delta\bar{\gamma}(Q) \leq \varepsilon$. Then, as $\Xi_{11}(\cdot) \succ 0$, by Schur complement, the optimisation problem of minimising $\Delta\gamma$ subject to:

$$\Xi_{11}(Q, \Delta D_2, \Delta\gamma) \succ 0 \quad (23)$$

$$\mathcal{W}(Q, \Delta D_2, \Delta\gamma) - \Xi_{12}(\cdot)\Xi_{11}(\cdot)^{-1}\Xi_{12}(\cdot)^T \succeq 0 \quad (24)$$

will be feasible, too, for all $\|Q\|_F \leq \min(\varepsilon, \delta)$, and the optimal solution $\Delta\bar{\gamma}(Q)$ will be limited only by active constraints in (24). The above reasoning, then, proves the following proposition:

Proposition 4. *Under assumptions 1 to 4, there exists $\varepsilon > 0$ such that, for all Q verifying $\|Q\|_F \leq \varepsilon$, the solution of* (20) *subject to* (21) *is the same as* (20) *subject to* (24)*.*

Note, however, that (24) is a nonlinear matrix inequality. Let us now prove that, for small enough ε, only $\mathcal{W}$ is relevant.

First, note that the eigenvalues of Ξ_{12} are $O(\varepsilon)$, by the aforementioned Lipschitz continuity and the fact that $\Xi_{12}(0,0,0) = 0$. Now, expressing the affine $\Xi_{11}(Q, \Delta D_2, \Delta\gamma)$ as $\Xi_{11}(Q, \Delta D_2, \Delta\gamma) = M + \mathcal{N}(Q, \Delta D_2, \Delta\gamma)$, being $M := \Xi_{11}(0,0,0)$, and $\mathcal{N}$ a linear function of its arguments, the

DOI: 10.1002/rnc

nonlinear term in (24) can be written as:

$$\begin{aligned}\widehat{\Xi} :&= -\Xi_{12}(\cdot)\Xi_{11}(\cdot)^{-1}\Xi_{12}(\cdot)^T \\ &= -\Xi_{12}(\cdot)M^{-1}\Xi_{12}(\cdot)^T + \Xi_{12}(\cdot)M^{-1}\mathcal{N}(\cdot)M^{-1}\Xi_{12}(\cdot)^T + \dots\end{aligned}$$

which shows that it contains quadratic, cubic, etc. terms in Q and the decision variables, which will be negligible for small enough ε: as the eigenvalues of Ξ_{12} are $O(\varepsilon)$, the eigenvalues of $\widehat{\Xi}$ are $O(\varepsilon^2)$. Thus, the solution $\Delta\gamma^*$ to (19)–(21), for small enough Ω_z (subsequently yielding small enough $Q(z)$), will be approximately equal to $\Delta\gamma_L^*$ obtained from the next optimisation problem:

$$\Delta\gamma_L^* := \sup_{Q\in Q(\Omega_z)} \quad \Delta\bar{\gamma}_L(Q) \tag{25}$$

$$\Delta\bar{\gamma}_L(Q) := \inf_{\Delta D_2} \quad \Delta\gamma \tag{26}$$

$$\text{subject to } \mathcal{W}(Q, \Delta D_2, \Delta\gamma) \succeq 0 \tag{27}$$

because the difference between (27) and (24) shrinks as $O(\varepsilon^2)$.

Note, however, that the above $\mathcal{W}$ is linear in all its arguments by Assumption 4.1, so if we have a solution of (26) subject to (27) for some Q, the solution for a scaled $Q' := \lambda Q$ would be $\Delta\bar{\gamma}_L(Q') = \lambda\Delta\bar{\gamma}_L(Q)$, $\lambda \in \mathbb{R}$. Feasibility is guaranteed for any Q because Assumption 4.2 guaranteed such feasibility for the more conservative (24) for a small enough $\|Q\|_F$, and (27) fulfills the just-mentioned scaling property.

The above scaling and convexity arguments will allow us to give a bound on $\Delta\gamma_L^*$ considering the solution of the above problem over a canonical basis of Q. Let us define as $E_{jk} \in \mathbb{R}^{m\times s}$ the matrix, with the same size as Q, whose element at position (j,k) is one, being the rest of elements equal to zero (j denotes row, k denotes column number).

Lemma 1. *Let us denote by $\Delta\gamma^{[jk]}$ the solution to the LMI problem*

$$\Delta\gamma^{[jk]} := \inf_{\Delta D_2} \Delta\gamma \tag{28}$$

$$\textit{subject to } \mathcal{W}(E_{jk}, \Delta D_2, \Delta\gamma) \succeq 0 \tag{29}$$

Let us define $\Delta\Gamma \in \mathbb{R}^{m\times s}$ as the matrix whose element (j,k) is $\Delta\gamma^{[jk]}$, and the set of matrices

$$\mathcal{J} := \{Q \in \mathbb{R}^{m\times s} : tr(\Delta\Gamma \cdot Q^T) \le \delta\}.$$

It follows that, if $Q(\Omega_z) \subset \mathcal{J}$, then $\Delta\gamma_L^* \le \delta$.

Proof

Conditions (27) are LMIs, so it is easy to prove that $\Delta\bar{\gamma}_L(Q)$ is a convex function of Q. Any arbitrary Q with the structure (16) can be trivially expressed as $Q = \sum_{j=1}^{m}\sum_{k=1}^{s} q_{jk}E_{jk}$. Let us denote the sum of absolute values of q_{jk} by $\varrho := \sum_{j=1}^{m}\sum_{k=1}^{s}|q_{jk}|$. We can state that the optimal $\Delta\bar{\gamma}_L(Q)$ in (26) will fulfill, by scaling and convexity argumentations:

$$\Delta\bar{\gamma}_L(Q) = \varrho \cdot \Delta\bar{\gamma}_L\left(\sum_{j=1}^{m}\sum_{k=1}^{s}\frac{|q_{jk}|}{\varrho} sign(q_{jk})E_{jk}\right)$$

DOI: 10.1002/rnc

$$\leq \varrho \cdot \sum_{j=1}^{m}\sum_{k=1}^{s} \frac{|q_{jk}|}{\varrho} sign(q_{jk}) \Delta\bar{\gamma}_L\left(E_{jk}\right) = \sum_{j=1}^{m}\sum_{k=1}^{s} \left(q_{jk} \cdot \Delta\gamma^{[jk]}\right) = \text{tr}\left(\Delta\Gamma \cdot Q^T\right). \quad (30)$$

Thus, for any $Q \in \mathcal{J}$ we have $\Delta\bar{\gamma}_L(Q) \leq \delta$, which implies the assertion in the lemma. □

3.1. Quasi-LPV modelling methodology

The above argumentations sum up asserting that for a small enough Ω_z, the solution $\Delta\gamma_L^*$ from (25) will be approximately equal to $\Delta\gamma^*$ from (19), so $\gamma^{\text{LPV}} \leq \gamma^{\text{lin}} + \Delta\gamma^* \approx \gamma^{\text{lin}} + \Delta\gamma_L^*$. Thus, the choice of the factorisation $\rho(z)$=$Q(z)z$ should try to minimise $\max_{Q\in Q(\Omega_z)} tr(\Delta\Gamma \cdot Q^T)$.

From the fact that $\text{tr}(\Delta\Gamma \cdot Q^T)$ is the scalar product on the vector space of matrices, and $\|Q\|_F$=$\text{tr}(QQ^T)^{1/2}$, we can assert that, for a given Q such that $\|Q\|_F \leq \varepsilon$, $\text{tr}(\Delta\Gamma Q^T) \leq \|\Delta\Gamma\|_F \|Q\|_F$ (Schwartz inequality), and equality holds if Q=$\alpha\Delta\Gamma$, being α a positive scalar. So the largest performance loss would be incurred if the components of Q "collinear" with $\Delta\Gamma$ are large.

Thus, the problem addressed next is finding a factorisation $\rho(z)$=$Q(z)z$, such that $Q(\Omega_z)$ has the smallest possible orthogonal projection over the 1-dimensional subspace (in the vector space of $\mathbb{R}^{m\times s}$ matrices) spanned by $\Delta\Gamma$. As ρ is a vector of nonlinearities, each row of matrix $\Delta\Gamma$ will induce a so-called *modelling direction* for the corresponding row of Q, i.e., for each individual nonlinearity composing ρ, so its projection over it should be minimised. Thus, the case $m = 1$ will be considered next, and the procedures to be detailed should be repeated for each of the nonlinearities comprising ρ in (4).

Consider a real-valued nonlinearity $v = \rho(z)$, $\rho : \mathbb{R}^s \mapsto \mathbb{R}$, fulfilling assumptions in Section 2 (i.e., being $\mathcal{C}^1$ and $\rho(0) = 0$) and a modelling direction given by column vector $\psi \in \mathbb{R}^s$, normalised so that $\psi^T\psi = 1$ without loss of generality[4].

Let us denote the 1-dimensional vector subspace generated by ψ as $\Xi := \{\eta\psi, \eta \in \mathbb{R}\}$. Denote the projection matrix onto Ξ as $\Pi_\Xi := \psi\psi^T$. The orthogonal projection of a set S of row vectors over Ξ will be understood as:

$$\text{proj}(S, \Xi) := \{\zeta : \exists y \in S \text{ s.t. } \zeta = y\Pi_\Xi\} \quad (31)$$

Recall now that Ω_z was the s-dimensional unit ball (Assumption 2). Defining the line segment $\mathbb{L}$:=$\Omega_z \cap \Xi$, evidently $\text{proj}(Q(\mathbb{L}), \Xi) \subseteq \text{proj}(Q(\Omega_z), \Xi)$ because $\mathbb{L} \subseteq \Omega_z$.

Definition 2. A factorisation $\rho(z) = Q(z)z$, $Q(z) = [q_1(z) \ \dots \ q_s(z)]$ is *tight* on subspace Ξ if

$$\text{proj}(Q(\Omega_z), \Xi) = \text{proj}(Q(\mathbb{L}), \Xi).$$

Proposition 5. *Under the assumptions in Section 2, given any arbitrary vector ψ, generating the 1-dimensional subspace Ξ, there exists a factorisation $\rho(z) = Q(z)z$ which is tight on Ξ.*

Proof

Considering Π_Ξ, and $\Pi_\perp := I - \Pi_\Xi$, z can be decomposed on two orthogonal components $z = z_\Xi + z_\perp$, being $z_\Xi := \Pi_\Xi z$ and $z_\perp := \Pi_\perp z$. Also, if $z \in \Omega_z$, $z_\Xi \in \mathbb{L}$, because Ω_z is the unit ball

[4]From the above-discussed geometric considerations, the actual model direction to be evaluated comes from transposing a $s \times 1$ row of $\Delta\Gamma$ defined in Lemma 1.

DOI: 10.1002/rnc

(Assumption 2), and $\mathbb{L}$ is a diameter of such hypersphere. Consider now the decomposition $\rho(z) = \rho(z_\Xi)+\delta(z)$, being $\delta(z) := \rho(z)-\rho(\Pi_\Xi z)$. Obviously, for $z \in \mathbb{L}$, we have $z = z_\Xi$ so $\delta(z) = 0$.

If, considering z_Ξ as fixed, we define $\bar{\delta}(z_\perp) := \delta(z_\perp + z_\Xi)$, then $\bar{\delta}$ fulfils Assumption 1 so, from Proposition 1, there exists a factorisation

$$\bar{\delta}(z_\perp) = \overline{Q}(z_\perp, z_\Xi)z_\perp = \delta(z). \tag{32}$$

Thus, we can decompose:

$$\rho(z) = \rho(\Pi_\Xi z)\frac{\psi^T z}{\psi^T z} + \overline{Q}(z_\perp, z_\Xi)\Pi_\perp z$$

Denoting:

$$Q(z) := \left(\frac{\rho(\Pi_\Xi z)}{\psi^T z}\psi^T + \overline{Q}(z_\perp, z_\Xi)\Pi_\perp\right) \tag{33}$$

we can, indeed, express $\rho(z) = Q(z)z$.

Now, due to the presence of $\Pi_\perp$, as $\Pi_\perp \Pi_\Xi = 0$,

$$\text{proj}(Q(z_\Xi + z_\perp), \Xi) = \frac{\rho(z_\Xi)}{\psi^T z_\Xi}\psi^T = \text{proj}(Q(z_\Xi), \Xi).$$

So, the projection of $Q(z)$ only depends on the component $z_\Xi \in \mathbb{L}$. Thus, $\text{proj}(Q(\Omega_z), \Xi) \subseteq \text{proj}(Q(\mathbb{L}), \Xi)$ and, as inclusion in the other sense comes from $\Omega_z \supseteq \mathbb{L}$, the proof is complete. □

The above proof constructs the component of $Q(z)$ collinear with ψ^T, but does not consider how to build $\bar{Q}$ arising in the decomposition (32). Note that, actually, choice of $\bar{Q}$ is, again, an instance of the same modelling problem on an $(s-1)$-dimensional space spanned by $z_\perp$. Based on this, a recursive application of the ideas yielding (33) is outlined in the Appendix to completely specify Q, after making a change of variable $z = T\eta$ where the last row of T is ψ^T and the remaining rows are built to conform an orthogonal matrix. From this change of variable, an hyper-rectangular bounding box readily ensues; such bounds will be discussed in Section 3.2.

Once the methodology to obtain $Q(z)$ has been discussed, Algorithm 1 is proposed to minimise the performance bound in Lemma 1. In this way, the spread of $Q(z)$ in the directions more deleterious for performance (at least for small enough modelling regions) will be minimised. Examples in Section 4 will show the performance improvement achieved with the resulting model from the referred algorithm.

Algorithm 1 Factorisation $\rho(z) = Q(z)z$

1: Solve the linearised performance problem, obtaining $\{D_1^{\text{lin}}, D_2^{\text{lin}}\}$ and γ^{lin}.
2: Using D_1^{lin}, obtain matrix $\Delta\Gamma$ defined on Lemma 1.
3: Denoting as $\Delta\Gamma^{[i]}$ the i-th row of $\Delta\Gamma$ and, likewise, $Q^{[i]}$ the i-th row of Q, model each nonlinearity as $\rho_i(z) = Q^{[i]}(z)z$, in such a way so that $Q^{[i]}$ is tight in the direction $\Delta\Gamma^{[i]}$, in the sense of Definition 2.

Note that only the geometry of the null space $\mathcal{C}$ is used in the factorisation proposal (Algorithm 1), and not any shape information on $\rho(z)$. Thus, final performance loss will depend on (a) the actual range of $Q(z)$, (b) the improvement a new D_1 in problem (11) might be able to achieve compared

DOI: 10.1002/rnc

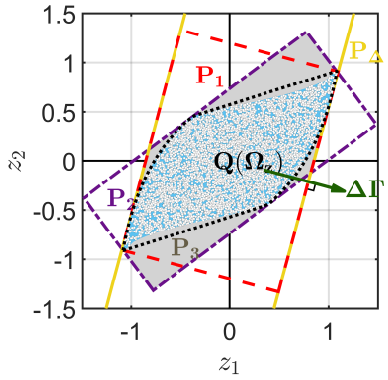

Figure 1. Illustration of options for polytopic bounding of $Q(\Omega_z)$ [blue] for $Q \in \mathbb{R}^{2\times 1}$. Best bounding is the convex hull [dotted-black line]. Given direction $\Delta\Gamma$ [green arrow], the rectangle with minimal projection on $\Delta\Gamma$, equation (52), is P_1. The principal-component box from [20] is P_2 [dash-dotted violet rectangle]. The intersection between P_2 and P_Δ from (35) is the shaded gray region P_3.

to D_1^{lin} and, (c) the influence of other directions not in the null space $\mathcal{C}$ as modelling region size increases (so $Q(z)$ significantly departs from zero). Hence, the ideas inspiring Algorithm 1 only apply for small enough modelling regions.

3.2. Bounding $Q(z)$ by polytopic embeddings

Definition 3. Once $Q(z)$ has been chosen to fulfill (5), a so-called polytopic embedding (a.k.a. TS model) of $Q(z)$ is a set of n_v vertex points $\hat{Q}_i$ such that $Q(z) \in \text{Co}\{\hat{Q}_i, i{=}1,\ldots,n_v\}$ for all $z{\in}\Omega_z$.

The usefulness of the polytopic embeddings is the fact that most references in the literature use well-known convexity argumentations and LMI relaxations [42] to pose a (conservative) approximation of (17) based on matrix inequalities involving only the vertices $\hat{Q}_i$.

Of course, the best embedding would be $\text{Co}(\{Q(z)|z \in \Omega_z\})$ but that set might not have a finite number of vertices. Figure 1 illustrates the idea for a case of a single nonlinearity $\rho(z_1, z_2){=}Q(z)z{=}(q_{11}(z)\ q_{12}(z))z$, where $Q(\Omega_z)$ is depicted as a blue region, and its convex hull is shown with a dotted-black boundary. In practice, the convex hull of a set of (dense enough) grid points might be the best reasonable approximation but it might, nevertheless, have a large number of vertices (for illustration, such grid appears as a collection of hundreds of white dots on Figure 1 inside the blue representation of $Q(\Omega_z)$).

Alternatively, there exists some simple options to avoid gridding, by bounding $Q(\Omega_z)$ in an hyperrectangle, obtaining infimum and supremum over some orthogonal directions, as follows:

(a) *Worst-performance directions:* expression (52) in the Appendix provides a polytopic hyper-rectangle bounding, given some change of variable. For the case $\Delta\Gamma = (0.966\ -0.259)$, which appears as a green arrow in the figure, the resulting bound would be the rectangle delimited by dashed-red lines, labelled as P_1 in the figure.

(b) *Principal-component (PCA) directions:* the work [20] proposes vectorising $Q(z)$ and obtaining the principal directions (eigenvectors of the covariance matrix), to reduce possible overbounding arising from choosing boxes not "aligned" with the principal components of $Q(\Omega_z)$; indeed, in many cases this does improve the obtained performance, compared to other box orientations, see

DOI: 10.1002/rnc

the cited work for details and examples. Informally, principal components align the rectangle with the spread of the data. The result of Kwatowski's PCA algorithm is the rectangle whose edges are outlined in dash-dotted violed line on Figure 1, labelled P_2.

Combined PCA plus worst-performance direction embedding. The above approaches have drawbacks arising from the fact that principal directions of the dataset $Q(z)$ might not be aligned with the worst-performance directions arising from the rows $\Delta\Gamma^{[i]}$. Thus, choosing just one of the two options may lead to either overbounding, in case (a) due to the misaligment with the data, or non-minimal projection over $\Delta\Gamma^{[i]}$, in case (b) losing the "tightness" pursued in the previous subsection. Both situations may yield unwanted performance loss.

Thus, to avoid any overbounding in the most influential direction, our proposal is to obtain both the Kwiatowski principal directions and the pair of supporting hyperplanes given by

$$l_1 := \inf_{z\in\Omega_z} tr(\Delta\Gamma Q(z)^T), \quad l_2 := \sup_{z\in\Omega_z} tr(\Delta\Gamma Q(z)^T) \tag{34}$$

so that the principal-component polytope P_2, is intersected with the set

$$P_\Delta := \{l_1 \leq tr(\Delta\Gamma Q^T) \leq l_2\} \tag{35}$$

being P_Δ represented as a band delimited by yelow lines on the figure. So we can state that $Q(\Omega_z) \subseteq P_3$, $P_3 := P_2 \cap P_\Delta$. The set P_3 is the shaded gray region on the figure[5] and, as a conclusion of the above ideas, such set is the proposed polytopic embedding for subsequent vertex-based LMI conditions.

Let us summarise, in the form of an algorithm, the proposed polytopic bounding, see Algorithm 2 on next page. Note: in the algorithm, we denote as $vec : \mathbb{R}^{m\times s} \mapsto \mathbb{R}^{1\times(ms)}$ the vectorisation operation so that a matrix is transformed to a *row* vector by concatenating its rows. The inverse operation (building the matrix again) is thus denoted as vec^{-1}.

4. EXAMPLES

Consider a nonlinear system whose state equation (1) is

$$\dot{x} = \begin{pmatrix} -2.6 & 0.7 \\ -3.4 & -3.5 \end{pmatrix} x + \begin{pmatrix} 1.2 & -1.7 \\ -0.1 & 0.5 \end{pmatrix} v + \begin{pmatrix} -0.5 \\ 0.9 \end{pmatrix} u + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} w \tag{36}$$

setting $z = x$ in (2) and splitting the output equation (3) onto two components:

$$y := \begin{pmatrix} \nu_\infty \\ \nu_2 \end{pmatrix}, \quad \nu_\infty = C_\infty x + Du, \quad \nu_2 = C_2 x + Du,$$

[5]Obviously, the polyhedral bound could be further trimmed by actually intersecting P_1 and P_2 or obtaining supporting hyperplanes in more directions; nevertheless, that would increase the number of vertices but their effect on the performance (for small Ω_z) would likely not be significant, according to the ideas in Lemma 1. Thus, the recommended option is the set P_3 as a sensible compromise between performance loss and model complexity.

DOI: 10.1002/rnc

Algorithm 2 Performance-oriented polytopic embedding

1: Obtain $\Delta\Gamma$ from the linearised problem using Algorithm 1, as well as the explicit expression of $Q(z)$, using formulae in the Appendix.
2: Denote the set of values of a vectorized Q by;

$$\Theta := \{vec(Q(z)) : z \in \Omega_z\} \subset \mathbb{R}^{1\times ms}$$

Obtain the covariance matrix of Θ, defined as $\Sigma = \int_\Theta \phi^T\phi\, d\phi$, see [20]. The dimensions of Σ are $ms \times ms$.
3: Obtain the eigenvalue decomposition $\Sigma = V\Lambda V^T$.
4: Denoting the i-th column of V as $V^{[i]}$, compute the bounds:

$$\lambda_{i1} = \inf_{\phi\in\Theta} \phi V^{[i]} \qquad \lambda_{i2} = \sup_{\phi\in\Theta} \phi V^{[i]}$$

5: Using the above bounds and (34), form the polytopic subset of $\mathbb{R}^{1\times ms}$ given by:

$$P_3^{vec} := \{l_1 \le vec(\Delta\Gamma)\phi \le l_2, \quad l_{i1} \le (V^{[i]})^T\phi \le l_{i2}, \quad i = 1,\dots,ms\}$$

6: Obtain the vertices of P_3^{vec}, so $P_3^{vec} = Co(\{v^{[1]},\dots v^{[j]}\})$ for some finite j.
7: Obtain the polytopic bounding of $Q(\Omega_z)$ inverting the vectorisation, so we can state that:

$$Q(z) \in P_3 := Co(\{vec^{-1}(v^{[1]}),\dots,vec^{-1}(v^{[j]})\}) \quad \forall z\in\Omega_z$$

with

$$C_\infty := \begin{pmatrix} 0.9 & -0.5 \\ 0 & 0 \end{pmatrix}, \quad C_2 := \begin{pmatrix} 0.9 & 0.5 \\ 0 & 0 \end{pmatrix}, \quad D := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and, last, a nonlinearity (4) given by:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \rho_1(x) \\ \rho_2(x) \end{pmatrix} = \begin{pmatrix} x_1x_2 \\ \sin(0.4x_1x_2) \end{pmatrix}$$

Assuming a quasi-LPV representation $v = Q(x)x$ is available, with $\Omega_z = \{x : x^Tx \le 1\}$, in this example, the performance objective will be a combined $\mathcal{H}_2/\mathcal{H}_\infty$ performance optimisation problem formally stated as minimising γ subject to conditions $\mathcal{MI}(\cdot)$ given by:

$$X \succeq \epsilon I \tag{37}$$

$$\gamma - 3\gamma_\infty - \gamma_2 \ge 0 \tag{38}$$

$$\begin{pmatrix} A(Q)X + BF(Q) + (A(Q)X + BF(Q))^T & (C_2X + DF(Q))^T \\ C_2X + DF(Q) & -\gamma_2 I \end{pmatrix} \preceq 0 \tag{39}$$

$$\begin{pmatrix} A(Q)X + BF(Q) + (A(Q)X + BF(Q))^T & (C_\infty X + DF(Q))^T & E \\ C_\infty X + DF(Q) & -\gamma_\infty I & 0 \\ E^T & 0 & -\gamma_\infty I \end{pmatrix} \preceq 0 \tag{40}$$

being $\epsilon = 10^{-4}$ and

$$A(Q) = \begin{pmatrix} -2.6 & 0.7 \\ -3.4 & -3.4 \end{pmatrix} + \begin{pmatrix} 1.2 & -1.7 \\ -0.1 & 0.5 \end{pmatrix} Q \tag{41}$$

DOI: 10.1002/rnc

If the above conditions are feasible, they guarantee that a gain-scheduled state feedback controller $u(Q,x) := F(Q)X^{-1}x$ ensures $\|\nu_\infty\|_2 \leq \gamma_\infty \|w\|_2$ under zero initial conditions and, too, that with $w = 0$ but nonzero initial state $x(0)$, $\|\nu_2\|_2 \leq \gamma_2 x^T(0)x(0)$, where 2-norm of signals are understood in the integral sense $\|\nu\|_2^2 = \int_0^\infty \nu(t)^T\nu(t)\,dt$. In the proposed problem, the importance weight of the γ_∞ performance bound has been arbitrarily set to be three times that of γ_2, in order to write (38).

This problem verifies the conditions in Assumption 2, with $D_1 \equiv X$, $D_2 \equiv F(Q)$.

Solving the above problem for the linearized system ($Q = 0$), the obtained optimal linearised performance is $\gamma^{\text{lin}} = 1.4121$, jointly with Lyapunov function decision variable $D_1^{\text{lin}} \equiv X^{\text{lin}}$ and linear controller decision variable $D_2^{\text{lin}} \equiv F^{\text{lin}}$, omitted for brevity; so, Assumption 3 is fulfilled. LMIs were set up using YALMIP-20160930 [24], and the used solver was SeDuMi 1.3 [44], in Matlab R2014a. YALMIP and SeDuMi options were set to default values.

Now, the proposed methodology for the choice of $Q(x)$ will be illustrated. With the optimal variables from the linearised solution, the suitable basis H of the relevant null space $\mathcal{C}$ of the optimal LMIs was obtained, so an expression for $\mathcal{W}$ was suitably programmed in the LMI solver..

Next, checking four vertex values for Q:

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

on the incremental version of the LMIs (28) and (29), sensitivity matrix $\Delta\Gamma$ in Lemma 1 results in

$$\Delta\Gamma = \begin{bmatrix} 0.6435 & -0.7654 \\ -0.6333 & 0.7739 \end{bmatrix}$$

Hence, the proposal in this work suggests that the nonlinearities ρ_1 and ρ_2 are modelled for minimal projection onto the subspace spanned for the first and second row of $\Delta\Gamma$, respectively, following Algorithm 1. As a result, the finally proposed factorisation is (42), understanding the fractions at $\xi_3 = 0$ or $\xi_4 = 0$ in a limit sense, see the Appendix.

$$\begin{aligned} \rho_1(x) &= \begin{pmatrix} 0.493\xi_1 - 0.172\xi_2 & -0.493\xi_2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \\ \rho_2(x) &= \begin{pmatrix} \frac{\sin(0.196\xi_4^2)+\sin(\boldsymbol{\kappa})}{\xi_3} & -\frac{\sin(0.196\xi_4^2)}{\xi_4} \end{pmatrix} \begin{pmatrix} \xi_3 \\ \xi_4 \end{pmatrix}, \end{aligned} \quad \begin{aligned} \xi_1 &:= 0.7654x_1 + 0.6435x_2, \\ \xi_2 &:= \Delta\Gamma^{[1]}x = 0.6435x_1 - 0.7654x_2, \\ \xi_3 &:= 0.7739x_1 + 0.6333x_2, \\ \xi_4 &:= \Delta\Gamma^{[2]}x = -0.6333x_1 + 0.7739x_2, \\ \boldsymbol{\kappa} &:= 0.196\xi_3^2 + 0.0792\xi_3\xi_4 - 0.196\xi_4^2 \end{aligned} \tag{42}$$

4.1. Comparison with alternative options for $Q(z)$.

In order to compare the above proposal with alternative approaches, a grid of 481 points $\{x^{[k]}, k = 1, \ldots, 481\}$ over a modelling region Ω_x being a circle of radius r in the state space $\mathbb{R}^2$ has been chosen: the origin plus a grid of 480 points in 10 radius values $\{0.1, 0.2, \ldots, 1\} \times r$ and 48 equally-spaced angle steps. In gridding-based approaches, the proposed $Q(z)$ would be evaluated at these 481 points, and compared with other possibilities for $Q(z)$; of course, later on, gridding will be avoided by embedding $Q(z)$ in a polytope with a small number of vertices. Obviously, a naive scaling transforms Ω_z to the unit ball required in the assumptions.

Gain-scheduled controller parametrisation. In gridding-based approaches, one controller gain will be sought for each grid point. In polytopic-based embeddings, one controller gain F_i will be obtained for each vertex: given the structure of LMIs (39) and (40), it is well-known [12] that these vertex gains can be used to synthesise a gain-scheduled controller with the same interpolation coefficients as the model matrices.

Annihilator approach. Let us first compare with BMI+gridding annihilator approaches in earlier literature [10]. A factorisation $v = (Q_0(x) + \aleph(x))x$ has been used, for any annihilator $\aleph(x)$ such that $\aleph(x)x = 0$. We chose $Q_0(x)$ to be the one suggested by our proposed subspace approach. The choice for $\aleph(x)$, at each grid point, was the linear annihilator:

$$\aleph(x^{[k]}) := \begin{pmatrix} -l_1^{[k]} x_2^{[k]} & l_1^{[k]} x_1^{[k]} \\ -l_2^{[k]} x_2^{[k]} & l_2^{[k]} x_1^{[k]} \end{pmatrix} \tag{43}$$

being $l_1^{[k]}$ and $l_2^{[k]}$ decision variables for each grid point. As discussed in Section 2, product of $\aleph$ and X forces the problem to be solved as a bilinear matrix inequality (BMI). The total number of BMI decision variables is 1928: 481×2 for $l_i^{[k]}$, 481×2 for $F^{[k]} \in \mathbb{R}^{1\times 2}$, three for the symmetric Lyapunov matrix X, plus γ, γ_∞ and γ_2. The inequalities resulting from (37), (38) and repeating 481 times, for each $Q(x^{[k]})$ the LMIs (39) and (40) can be expressed as a big block-diagonal BMI of size 4813×4813. Thus, the annihilator approach entails solving a large BMI problem with a lot of decision variables. For the sake of comparison, in our approach to search for the best $Q(z)$, each element of $\Delta\Gamma$ was obtained with a 13×13 LMI, with 8 decision variables.

Note that the resulting BMI search, if properly converged, should always yield a better performance bound than our proposal: this is intentional, as the issue is determining whether the achieved improvement is significant or worthwhile, given the increase of computational cost arising from the iterative BMI steps. Actual numerical results will be later discussed, after other alternatives options are also presented.

Inspection-based factorisation. For the sake of comparison, four other easy "common sense" factorisation options for $Q(x)$ will be evaluated, either extracting x_1 or x_2 as a factor, as follows:

$$v = \begin{pmatrix} \frac{\rho_1(x)}{x_1} & 0 \\ \frac{\rho_2(x)}{x_1} & 0 \end{pmatrix} x = \begin{pmatrix} x_2 & 0 \\ \frac{\sin(0.4x_1x_2)}{x_1} & 0 \end{pmatrix} x \tag{44}$$

$$v = \begin{pmatrix} 0 & \frac{\rho_1(x)}{x_2} \\ \frac{\rho_2(x)}{x_1} & 0 \end{pmatrix} x = \begin{pmatrix} 0 & x_1 \\ \frac{\sin(0.4x_1x_2)}{x_1} & 0 \end{pmatrix} x \tag{45}$$

$$v = \begin{pmatrix} \frac{\rho_1(x)}{x_1} & 0 \\ 0 & \frac{\rho_2(x)}{x_2} \end{pmatrix} x = \begin{pmatrix} x_2 & 0 \\ 0 & \frac{\sin(0.4x_1x_2)}{x_2} \end{pmatrix} x \tag{46}$$

$$v = \begin{pmatrix} 0 & \frac{\rho_1(x)}{x_2} \\ 0 & \frac{\rho_2(x)}{x_2} \end{pmatrix} x = \begin{pmatrix} 0 & x_1 \\ 0 & \frac{\sin(0.4x_1x_2)}{x_2} \end{pmatrix} x \tag{47}$$

Trivially, these options can give a simple 4-vertices polytopic embedding representation, by combining minimum and maximum bounds of the two non-zero matrix elements.

DOI: 10.1002/rnc

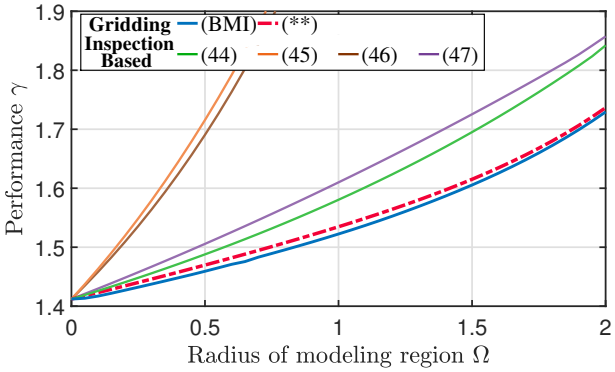

Figure 2. Peformance comparison of choices for $Q(z)$, gridding approach. Our proposal from Algoritm 1, i.e., $Q(z)$ in (42), labelled as (**) achieves a very good performance with significantly less computational cost than the (BMI) one.

Comparison of results (gridding approach). Comparing all six candidate options for $Q(z)$ in the same grid of 481 points results in Figure 2. It can be clearly seen that our proposal for $Q(z)$, labelled as (**), provides the best results except the marginal improvements of the costly annihilator-BMIs, as expected (at least for small modelling region radius). The common-sense factorisations above incur in a clear performance penalty. Note that, as intuitively expected, performance is close to the linearised one (1.4121) for small modelling region radius, but it worsens as such radius increases (indeed, the set $Q(\Omega_z)$ grows larger as the radius of Ω_z expands).

Computation time (gridding approach). On a Core I5-4690 processor, computation of $\Delta\Gamma$ with our LMI proposals takes around 0.088 seconds. Once $Q(z)$ is crafted from the resulting $\Delta\Gamma$, solving the LMIs with SeDuMi 1.3 over the 481-point grid to obtain the new gain-scheduled controller takes 8.76 seconds. On the other hand, using PenBMI2.1 32 bit with default configuration, a 481-point grid takes around 420 seconds in average for each tested radius.

Note that the computational cost would exacerbate for higher-order systems because of the larger dimension of the matrix inequalities and the likely need of an exponentially larger number of grid points; thus, polytopic bounding to avoid gridding is clearly important, and it will be discussed below in the context of this example.

As a conclusion, our subspace approach has two worthwhile advantages (at least in this example): first, it clearly outperforms *ad-hoc* inspection-based factorisations (44)–(47); second, BMI+gridding options in other literature achieve only marginal improvements and the computational cost is much higher.

4.2. Effect of polytopic embedding

Once $Q(z)$ is chosen from our proposal, in order to avoid gridding, performance of polytopic embeddings will be tested. On Figure 3, the three different approaches discussed on Section 3.2 are compared with the gridding-based "best-possible" performance (in this example, denser grids were tested with no appreciable performance improvement). First, the simpler options:

DOI: 10.1002/rnc

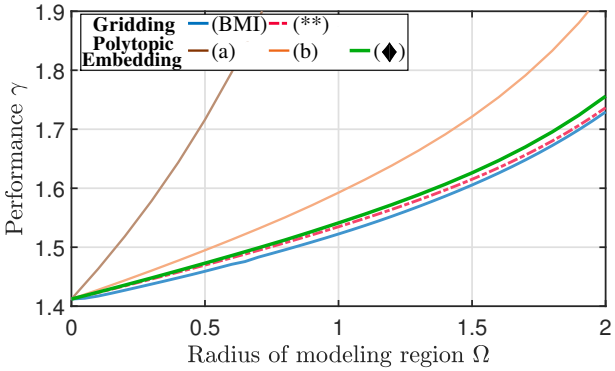

Figure 3. Performance degradation due to different options for polytopic embedding, for the fixed $Q(z)$ arising from Algorithm 1 in (42). See enumeration in the text for interpretation of labels (a), (b), (♦). Our proposal (♦) achieves comparable performance with a 10-fold improvement in computational time over the gridding results (**) and 700-fold over (BMI), both copied from Figure 2.

Table I. Performance versus computation time (radius 2)

Method	Num. of vertices	Computation time (s)	Performance (penalty%)	Figure
Earlier literature:				
BMI-Gridding [10]	Grid 481	420	1.729 (*best*)	Fig.1&2, blue
LMI-Gridding (44)	Grid 481	8.73	1.842 (6.5%)	Fig.1, green
PCA bound (44)	4	0.14	1.929 (11.5%)	–
inf/sup bound of (44)	4	0.11	11.60 (570%)	–
LMI-Gridding (45)	Grid 481	7.14	3.832 (121%)	Fig.1,light brown
LMI-Gridding (46)	Grid 481	7.19	4.819 (178%)	Fig.1,dark brown
LMI-Gridding (47)	Grid 481	6.68	1.857 (7.4%)	Fig.1, violet
Proposals in this work:				
Alg. 1 $Q(z)$+LMI	Grid 481	8.76	1.742 (0.7%)	Fig.1&2, red
Alg. 1+Alg. 2. LMI	24	0.56	1.756 (1.5%)	Fig.2, green

(a) the 16-vertices box arising from an orthogonal transformation induced by worst-performance directions in $\Delta\Gamma$, using (52); we recall that the methodology is analogous to the one yielding P_1 in the simplified illustrative example on Figure 1.

(b) the PCA approach from [20] (with, too, 16 vertices), symbolically depicted as P_2 on Figure 1.

and, finally, our proposed combined approach, labelled as (♦) on Figure 3, intersecting the principal-component box with the supporting hyperplanes (34), which results in a polytope with 24 vertices (i.e., the analogous option to the one depicted as P_3 on Figure 1).

In order to better assess the performance loss due to the polytopic embedding, the original BMI+grid approach and the LMI+grid on our proposed $Q(z)$ have been copied over from Figure 2 to Figure 3 with the same line styles, and additional lines added. So, in total, Figure 3 depicts three polytopic embeddings plus two griding-based calculations (LMI/BMI) over the same choice of $Q(z)$.

Computation time comparison. As intuitively expected, polytopic embedding achieves more conservative performance bounds than gridding options, but at a much lower computational cost.

Table I depicts an overall comparison between the different approaches showing computation time and performance (and the number of considered vertices/grid points), for the rightmost point of the figure (states in the region $x^T x \leq 2$). Indication of figure number and colour of some performance entries are also presented in the table. Let us discuss the different table entries below:

The first rows describe proposals in prior literature:

- Row 1 lists data for the already discussed BMI-gridding, whose performance is, obviously, the best one.
- Rows 2 to 4 use the inspection-based factorisation (44): directly on a grid (row 2), bounding it with the PCA-based approach in [20] (row 3), or using the most straightforward bounding by computing the minimum and maximum over Ω_z of the matrix elements, routinely proposed in many applications [43, 20] in row 4 (in the particular case in this example, in the ball of radius 2 we can bound $-2 \leq x_2 \leq 2$, and $-0.8 \leq \frac{\sin(0.4x_1x_2)}{x_1} \leq 0.8$ yielding a 4-vertices polytopic LPV model, details left for the reader).
- Rows 5 to 7 are the results of alternative inspection-based options.

After presenting prior literature results, the last two rows present the data from our proposals: first, a gridding approach (row 8) to point out the effect of the choice of $Q(z)$ with respect to the other gridding options (rows 1, 2, 5, 6, 7). Apart from the reference row 1 (BMI), the performance from our proposal beats by a large margin the rest of the referred rows. Finally, row 9 presents the combination of our choice of $Q(z)$ plus the polytopic embedding from Algorithm 2, with a minor performance loss with respect to rows 1 and 8, and a significantly faster execution.

For sake of comparison, the inspection-based polytopic embeddings were unable to approach the optimal performance points. For instance, (44) was the best factorisation when evaluated on the 481-point grid, yielding a performance of 1.84 (6.4% penalty); however, the 4-vertices PCA bound of (44) incurred in a 11.5% penalty, and a much larger penalty figure appeared for the straightforward inf/sup bound of each matrix component over the circle Ω_z. Similar results were achieved with the other options (45)–(47), omitted for brevity.

In conclusion, our proposals beat inspection-based models; they yield 0.7% performance penalty with a 46 times lower computation time (Algorithm 1 with LMI gridding), and 1.5% performance penalty with more than 700 times lower computation time (Algorithm 2, polytopic embedding) with respect to the ideal BMI-gridding results. These computational advantages would, likely, accentuate for higher dimensions.

5. CONCLUSIONS

This paper has addressed the problem of choosing a good quasi-LPV model for a nonlinear system. Our proposals are able to mitigate the performance loss from a reference linearised design, compared to other inspection-based modelling choices, while avoiding BMIs needed in prior literature for controller synthesis problems. The main idea is based on obtaining the more harmful perturbations of the model matrices in the linearised design, proposing then a coordinate transformation so that the projection of the uncertainty over these model perturbations is minimised. This direction plays, too, a fundamental role in subsequent polytopic embedding options which are

DOI: 10.1002/rnc

built to avoid the need of gridding. As the root of the procedure is the linearised design, the results are only valid, formally, for small modelling regions around the origin. An illustrative example shows that, indeed, the proposal seems to strike an interesting balance between computational cost and achieved performance: BMIs are avoided, but our LMI results are very close to the ideal BMI ones from prior works, significantly improving over *ad-hoc* factorisations used in many applications.

REFERENCES

[1] David Angeli, Alessandro Casavola, and Edoardo Mosca. Constrained predictive control of nonlinear plants via polytopic linear system embedding. *International Journal of Robust and Nonlinear Control*, 10(13):1091–1103, 2000.

[2] Pierre Apkarian and Pascal Gahinet. A convex characterization of gain-scheduled h_∞ controllers. *IEEE Transactions on Automatic Control*, 40(5):853–864, 1995.

[3] Pierre Apkarian, Pascal Gahinet, and Greg Becker. Self-scheduled h_∞ control of linear parameter-varying systems: a design example. *Automatica*, 31(9):1251–1261, 1995.

[4] M. Bernal and T. M. Guerra. Generalized non-quadratic stability of continuous-time Takagi-Sugeno models. *IEEE Transactions on Fuzzy Systems*, 18(4):815–822, 2010.

[5] Miguel Bernal, Thierry Marie Guerra, and Alexandre Kruszewski. A membership-function-dependent approach for stability analysis and controller synthesis of takagi–sugeno models. *Fuzzy Sets and Systems*, 160(19):2776 – 2795, 2009. Theme: Control.

[6] F.D. Bianchi, R.J. Mantz, and C.F. Christiansen. Gain scheduling control of variable-speed wind energy conversion systems using quasi-lpv models. *Control Engineering Practice*, 13(2):247–255, 2005.

[7] MD Bronshtein. Smoothness of roots of polynomials depending on parameters. *Siberian Mathematical Journal*, 20(3):347–352, 1979.

[8] Fredrik Bruzelius, Stefan Pettersson, and Claes Breitholtz. Linear parameter-varying descriptions of nonlinear systems. In *American Control Conference, 2004. Proceedings of the 2004*, volume 2, pages 1374–1379. IEEE, 2004.

[9] Graziano Chesi. Estimating the domain of attraction for non-polynomial systems via lmi optimizations. *Automatica*, 45(6):1536–1541, 2009.

[10] D. F. Coutinho, M. Fu, A. Trofino, and P. Danes. $\mathcal{L}2$-gain analysis and control of uncertain nonlinear systems with bounded disturbance inputs. *International Journal of Robust and Nonlinear Control*, 18(1):88–110, 2008.

[11] Daniel Ferreira Coutinho, Minyue Fu, and Alexandre Trofino. Robust analysis and control for a class of uncertain nonlinear discrete-time systems. *Systems & control letters*, 53(5):377–393, 2004.

[12] Jan De Caigny, Juan F Camino, Ricardo CLF Oliveira, Pedro Luis Dias Peres, and Jan Swevers. Gain-scheduled $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control of discrete-time polytopic time-varying systems. *IET control theory & applications*, 4(3):362–380, 2010.

[13] Pascal Gahinet, Pierre Apkarian, and Mahmoud Chilali. Affine parameter-dependent lyapunov functions for real parametric uncertainty. *Automatic Control, IEEE Transactions on*, 41(3):436–442, 1996.

[14] Q. Gao, X.J. Zeng, G. Feng, Y. Wang, and J. Qiu. T–S-Fuzzy-Model-Based approximation and controller design for general nonlinear systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 42(4):1143–1154, 2012.

[15] Temoatzin González, Miguel Bernal, Antonio Sala, and Braulio Aguiar. Cancellation-based nonquadratic controller design for nonlinear systems via takagi–sugeno models. *IEEE transactions on cybernetics*, 2016.

[16] Thierry Marie Guerra, Victor Estrada-Manzo, and Zsófia Lendek. Observer design for takagi–sugeno descriptor models: An lmi approach. *Automatica*, 52:154–159, 2015.

[17] Thierry Marie Guerra and Laurent Vermeiren. Lmi-based relaxed nonquadratic stabilization conditions for nonlinear systems in the takagi–sugeno's form. *Automatica*, 40(5):823–829, 2004.

[18] Yun Huang. *Nonlinear optimal control: an enhanced quasi-LPV approach*. PhD thesis, California Institute of Technology, 1999.

[19] Yun Huang and Ali Jadbabaie. Nonlinear $\langle_\infty$ control: An enhanced quasi-lpv approach. In *Proceedings of the IFAC World Congress*, pages 85–90. Citeseer, 1999.

[20] Andreas Kwiatkowski and Herbert Werner. Pca-based parameter set mappings for lpv models with fewer parameters and less overbounding. *IEEE Transactions on Control Systems Technology*, 16(4):781–788, 2008.

[21] D. H. Lee and D. W. Kim. Relaxed LMI conditions for local stability and local stabilization of continuous-time Takagi-Sugeno fuzzy systems. *IEEE Transactions on Cybernetics*, 44(3):394–405, 2014.

[22] Valter JS Leite and Pedro LD Peres. An improved lmi condition for robust d-stability of uncertain polytopic systems. *IEEE Transactions on Automatic Control*, 48(3):500–504, 2003.

[23] Z. Lendek, T.M Guerra, R. Babuška, and B. De-Schutter. *Stability Analysis and Nonlinear Observer Design Using Takagi-Sugeno Fuzzy Models*. Springer-Verlag, Netherlands, 2010.

[24] J. Löfberg. YALMIP : A toolbox for modeling and optimization in Matlab. In *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[25] Adrian McKernan, Antonio Sala, Carlos Ariño, and George W Irwin. Sampled-data gain scheduling of continuous ltv plants. *Automatica*, 45(10):2451–2453, 2009.

[26] Javad Mohammadpour and Carsten Scherer, editors. *Control of Linear Parameter Varying Systems with Applications*. Springer, New York, USA, 2012.

[27] A. Packard. Gain scheduling via linear fractional transformations. *Systems & Control Letters*, 22(2):79–92, 1994.

[28] S. Prajna, A. Papachristodoulou, and F. Wu. Nonlinear control synthesis by sum of squares optimization: A Lyapunov-based approach. In *Control Conference, 2004. 5th Asian*, volume 1, pages 157–165, 2004.

[29] Ruben Robles, Antonio Sala, Miguel Bernal, and Temoatzin Gonzalez. Choosing a Takagi-Sugeno model for improved performance. In *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*, pages 1–6, Istanbul, Turkey, 2015. IEEE.

[30] Ruben Robles, Antonio Sala, Miguel Bernal, and Temoatzin González. Subspace-based Takagi-Sugeno modeling for improved LMI performance. *IEEE Transactions on Fuzzy Systems*, In Press, 2016.

[31] Damiano Rotondo, Fatiha Nejjari, and Vicenç Puig. Robust state-feedback control of uncertain lpv systems: An lmi-based approach. *Journal of the Franklin Institute*, 351(5):2781 – 2803, 2014.

[32] Damiano Rotondo, Vicenç Puig, Fatiha Nejjari, and Marcin Witczak. Automated generation and comparison of takagi–sugeno and polytopic quasi-lpv models. *Fuzzy Sets and Systems*, 277:44–64, 2015.

[33] Wilson J Rugh and Jeff S Shamma. Research on gain scheduling. *Automatica*, 36(10):1401–1425, 2000.

[34] A. Sala and C. Ariño. Polynomial fuzzy models for nonlinear control: a taylor series approach. *IEEE Transactions on Fuzzy Systems*, 17(6):1284–1295, 2009.

[35] Antonio Sala. Computer control under time-varying sampling period: An lmi gridding approach. *Automatica*, 41(12):2077–2082, 2005.

[36] Antonio Sala. On the conservativeness of fuzzy and fuzzy-polynomial control of nonlinear systems. *Annual Reviews in Control*, 33(1):48–58, 2009.

[37] Antonio Sala and Carlos Arino. Relaxed stability and performance lmi conditions for takagi–sugeno fuzzy systems with polynomial constraints on membership function shapes. *IEEE Transactions on Fuzzy Systems*, 16(5):1328–1336, 2008.

[38] Antonio Sala and José Luis Pitarch. Optimisation of transient and ultimate inescapable sets with polynomial boundaries for nonlinear systems. *Automatica*, 73:82–87, 2016.

[39] Masayuki Sato and Dimitri Peaucelle. Gain-scheduled output-feedback controllers using inexact scheduling parameters for continuous-time lpv systems. *Automatica*, 49(4):1019–1025, 2013.

[40] Carsten Scherer. Lpv control and full block multipliers. *Automatica*, 37(3):361–375, 2001.

[41] Carsten Scherer. *Linear Matrix Inequalities in Control Theory*. Delf University, Delf, The Netherlands, 2004.

[42] Carsten W Scherer. Lmi relaxations in robust control. *European Journal of Control*, 12(1):3–29, 2006.

[43] Jeff S. Shamma and James R. Cloutier. A linear parameter varying approach to gain scheduled missile autopilot design. In *American Control Conference, 1992*, pages 1317–1321. IEEE, 1992.

[44] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11-12:625–653, 1999.

[45] Kazuo Tanaka and Hua O Wang. *Fuzzy control systems design and analysis: a linear matrix inequality approach*. John Wiley & Sons, New York, 2001.

[46] Roland Tóth, Peter SC Heuberger, and Paul MJ Van den Hof. Prediction-error identification of lpv systems: Present and beyond. In *In [26]*, pages 27–58. Springer, 2012.

[47] A Trofino and TJM Dezuo. Lmi stability conditions for uncertain rational nonlinear systems. *International Journal of Robust and Nonlinear Control*, 24(18):3124–3169, 2014.

[48] Fen Wu and Ke Dong. Gain-scheduling control of lft systems using parameter-dependent lyapunov functions. *Automatica*, 42(1):39–50, 2006.

[49] Fen Wu, Xin Hua Yang, Andy Packard, and Greg Becker. Induced $\mathcal{L}_2$-norm control for LPV systems with bounded parameter variation rates. *International Journal of Robust and Nonlinear Control*, 6(9-10):983–998, 1996.

A. SYSTEMATIC FACTORISATION OF A NONLINEARITY

Let us consider a $\mathcal{C}^1$ nonlinearity $\tilde{\rho} : \mathbb{R}^s \to \mathbb{R}$:

$$v = \tilde{\rho}(\eta) \tag{48}$$

and present a factorisation option $v = Q(\eta) \cdot \eta$, based in [29, 14] and the proof of Proposition 5.

Let us define $\xi_1 := \tilde{\rho}(\eta_1, \dots, \eta_s)$ and $\xi_i(\eta_i, \dots, \eta_s) := \tilde{\rho}(0, \dots, 0, \eta_i, \dots, \eta_s)$ for $i = \{2, \dots s\}$, $\xi_{s+1} := 0$ and let us denote, for $i = \{1, \dots, s\}$:

$$\zeta_i(\eta_i, \dots, \eta_s) := \xi_i(\eta_i, \dots, \eta_s) - \xi_{i+1}(\eta_{i+1}, \dots, \eta_s). \tag{49}$$

With the above definitions, we have

$$\zeta_i(0, \eta_{i+1}, \dots, \eta_s) = \xi_i(\dots) - \xi_{i+1}(\dots) = \tilde{\rho}(0, \dots, 0, \eta_{i+1}, \dots, \eta_s) - \tilde{\rho}(0, \dots, 0, \eta_{i+1}, \dots, \eta_s) = 0,$$

and $\zeta_i(\cdot) \in \mathcal{C}^1$. Hence, by Proposition 1, there exists a factorisation

$$\zeta_i(\eta_i, \ldots, \eta_s) = \frac{\zeta_i(\eta_i, \ldots, \eta_s)}{\eta_i} \cdot \eta_i$$

taking suitable limits when $\eta_i \to 0$, being $\zeta_i(\ldots)/\eta_i$ a continuous function. Thus, as trivially:

$$\xi_i(\eta_i, \ldots, \eta_s) = \zeta_i(\eta_i, \ldots, \eta_s) + \xi_{i+1}(\eta_{i+1}, \ldots, \eta_s)$$

with these definitions, we have $\xi_1(\ldots) = \zeta_1(\ldots) + \xi_2(\ldots) = \zeta_1(\ldots) + \zeta_2(\ldots) + \xi_3(\ldots) = \ldots$, i.e.:

$$\tilde{\rho}(\eta) = \sum_{i=1}^{s} \zeta_i(\eta_i, \ldots, \eta_s) = \sum_{i=1}^{s} \frac{\zeta_i(\cdot)}{\eta_i} \cdot \eta_i \tag{50}$$

so (50) entails that we can write $v = Q(\eta)\eta$, being $Q(\eta)$ the expression below:

$$Q(\eta) = \frac{\zeta_s}{\eta_s} \begin{pmatrix} 0 & \ldots & 0 & 1 \end{pmatrix} + \begin{pmatrix} \frac{\zeta_1}{\eta_1} & \ldots & \frac{\zeta_{s-1}}{\eta_{s-1}} & 0 \end{pmatrix}. \tag{51}$$

It can be easily proved that the resulting $Q(\eta)$ is tight on the subspace generated by $\psi := (0 \ \ldots \ 0 \ 1)^T$, because the rightmost term in the summation is orthogonal to ψ, so the above expression is in the form (33), details left to the reader.

Polytopic embedding via inf/sup bounds. A straightforward embedding with $n_v{=}2^s$ vertices can be built by obtaining the infimum and supremum, for η in some modelling region $\Omega_\eta \subseteq \mathbb{R}^s$, of each element $\zeta_i(\eta_i, \ldots, \eta_s)/\eta_i$ conforming $Q(\eta)$ in (51). Let us denote them $\underline{\zeta_i^*}$ and $\overline{\zeta_i^*}$, respectively (assuming they exist; by the continuity assumption they will for sure do if Ω_η is compact). Indeed, straightforward interpolation allows us to write the well-known expression:

$$\tilde{\rho}(\eta) = \sum_{i=1}^{s} \left(\mu_{i1}(\eta)\underline{\zeta_i^*} + \mu_{i2}(\eta)\overline{\zeta_i^*} \right) \eta_i \tag{52}$$

with $\mu_{i1}(\eta) + \mu_{i2}(\eta) = 1$, $\mu_{i1}(\eta) \geq 0$, $\mu_{i2}(\eta) \geq 0$ for all $\eta \in \Omega_\eta$.

Note that, apart from this "box" embedding of (48), other possibilities exist, as discussed in Section 3.2 on the main text.