Reference and Command Governors: A Tutorial on Their Theory and Automotive Applications

Kolmanovsky, I.; Garone, E.; Di Cairano, S.

TR2014-119 June 2014

Abstract

The paper provides a tutorial overview of reference governors and command governors, which are add-on control schemes for reference supervision and constraint enforcement in closed-loop feedback control systems. Approaches to the development of such schemes for linear and nonlinear systems are described. The treatment of unmeasured disturbances and parametric uncertainties is addressed. Generalizations to extended command governors, feedforward reference governors, reduced order reference governors, parameter governors, networked reference governors and decentralized reference governors are discussed. Examples of applications of these techniques to automotive systems are given. A comprehensive list of references is included. Comments comparing reference and command governor approaches with Model Predictive Control and on future directions in reference and command governor research are included.

American Control Conference (ACC)
Reference and Command Governors: A Tutorial on Their Theory and Automotive Applications

Ilya Kolmanovsky\textsuperscript{1} Emanuele Garone\textsuperscript{2} Stefano Di Cairano\textsuperscript{3}

Abstract—The paper provides a tutorial overview of reference governors and command governors, which are add-on control schemes for reference supervision and constraint enforcement in closed-loop feedback control systems. Approaches to the development of such schemes for linear and nonlinear systems are described. The treatment of unmeasured disturbances and parametric uncertainties is addressed. Generalizations to extended command governors, feedforward reference governors, reduced order reference governors, parameter governors, networked reference governors and decentralized reference governors are discussed. Examples of applications of these techniques to automotive systems are given. A comprehensive list of references is included. Comments comparing reference and command governor approaches with Model Predictive Control and on future directions in reference and command governor research are included.

1. INTRODUCTION

With the advances in control theory, many effective techniques have become available for the design of feedback control systems with the desired stability, performance and disturbance rejection properties. The interest in treating the requirements that have the form of pointwise-in-time state and control constraints has also been growing, in particular, in the automotive domain. Examples of constraints in practical systems include actuator magnitude and rate limits, and bounds imposed on process variables to ensure safe and efficient system operation. Controllers that achieve high performance in systems with constraints are typically nonlinear and are often (but not always) based on prediction and optimization.

A control designer faced with the task of satisfying the constraints has several choices. One route is to re-design the controller within the Model Predictive Control (MPC) framework [96], [92], [22], [89], [52]. Another route is to augment a well-designed nominal controller, that already achieves high performance for small signals, with constraint handling capability for larger signals and transients that have the potential to induce constraint violation. The second route may be attractive to practitioners interested in preserving an existing/legacy controller or concerned with computational effort, tuning complexity, stability and robustness certification, requirements other than constraint handling satisfactorily addressed by the existing controller, etc. Anti-windup compensation and the augmentation of Lyapunov controllers with barrier functions are examples of the second approach and so are the reference and command governors.

As its name suggests, the reference governor is an add-on scheme for enforcing pointwise-in-time state and control constraints by modifying the reference command to a well-designed (for small signals) closed-loop system. See Figure 1. Numerous reference governor like schemes have been proposed. The range of potential options includes scalar and vector reference governors, command governors, extended command governors, incremental reference governors, feedforward reference governors, network reference governors, reduced order reference governors, parameter governors, virtual state governors and others. The common intent of these governors is to preserve, whenever possible, the response of the closed loop system designed by conventional control techniques. Frequently (but not always), they achieve this by ensuring that the modified reference command is as close as possible to the original reference command subject to satisfying the constraints.

The scalar reference governor is attractive as it leads to computationally simple implementations for both linear and nonlinear systems with disturbances and parameter uncertainties. Other reference/command governor like schemes are more complex but provide better performance or address special problems, e.g., networked implementation for systems with communication delays and data drops, or decentralized and reduced order implementation for large scale systems.

The conventional scalar reference governors, vector reference governors, command governors and extended command governors generate \( v(t) \) in Figure 1 based on the current value of the reference, \( r(t) \), the state estimate, \( x(t) \), and, in the case of the extended command governor, based also on an \( n \)-vector state \( \hat{x} \), of a supplementary and fictitious dynamic system.

While the reference governor is attractive as it leads to computationally simple implementations for both linear and nonlinear systems with disturbances and parameter uncertainties,

Reference governors were first proposed as continuous-time algorithms in [65]. Later the discrete-time framework [46], [49] has emerged that has several advantages from the implementation standpoint. The static reference governor [46] used \( v(t) = \kappa(t)r(t) \), where the parameter \( \kappa(t), 0 \leq \kappa(t) \leq 1 \), was maximized subject to conditions that guaran-
ted the constraint enforcement. Specifically, $x(t + 1) \in O_m$, where $O_m$ is the maximum output admissible set [45] of all states that with zero reference command do not lead to subsequent constraint violation. Because of the possibility of oscillations [46], the static reference governor was abandoned and replaced by a dynamic reference governor (see Section II) for which finite-time convergence for constant or nearly constant reference commands was established. Other formulations of reference and command governors have appeared in [9], [10], [47], [11], [23]. See also references therein. The developments included the treatment of linear systems with uncertainties and set-bounded disturbance inputs, and the implementation based on non-positively invariant sets. Extended command governors [50] represented a further generalization with a potential to provide a larger constrained domain of attraction and faster response at the price of increased computational complexity.

For nonlinear systems, approaches to reference governor design have been also developed, see e.g., [7], [48], [95], [51], [15] and references therein. Some of these approaches exploit on-line prediction through simulations or level sets of Lyapunov functions to guard against constraint violation. The robust reference governor that handles parametric uncertainty based on response approximations has been presented in [108]. The parameter governor has been proposed in [80] to adjust constant controller parameters or controller states based on prediction and optimization.

More recently, classical reference and command governor ideas have been extended in several additional directions. These include the treatment of networked systems and large scale systems.

Automotive control systems are designed to satisfy stringent and numerous fuel economy, emissions, safety, performance and drivability requirements [56]. Their traditional implementation involves a family of hierarchically arranged real-time control algorithms, where reference commands from an upper layer controller are passed to a lower level controller, with the highest level commands corresponding to the driver and the road input. The on-board computing power in automotive systems (ROM, RAM and chronometrics) is very limited and successful control solutions must minimize the computational footprint. In addition, these solutions must be easy to understand, modify and calibrate. With the trend towards growing use of model-based control in the automotive industry and increasing importance of handling limits, reference and command governors represent an attractive option for several automotive applications where by supervision and minimum modification of reference commands applied to lower level controllers various constraints can be enforced.

In this paper we survey several basic and more recent reference governor results. We then discuss some of the research on applications of reference governors to automotive control problems. Comments on connections with MPC and on directions for future research are discussed at the end. Parts of Sections II, V, and IV-B are based on [86], however, re-written and modified to improve the discussions and expand references.

II. REFERENCE GOVERNORS FOR LINEAR SYSTEMS

The classical reference governor is designed based on a discrete-time linear system model of the form,

$$x(t + 1) = Ax(t) + Bv(t) + Bu(t),$$
$$y(t) = Cx(t) + Dv(t) + Dw(t).$$  \( (1) \)

In (1), $x$ is an $n$-vector state, $v$ is an $m$-vector input, $w$ is an $l$-vector disturbance, and $y$ is a $p$-vector system output. The disturbance $w(t)$ is unmeasured but has known bounds specified as $w(t) \in W$ for all $t \in Z^+$, where $W$ is a given compact set and $Z^+$ denotes the set of non-negative integers.

Typically, the model (1) represents the closed-loop system thereby reflecting the combined closed-loop dynamics of the plant and of the controller in Figure 1. As normally the closed-loop system is designed to be asymptotically stable, the matrix $A$ is assumed to be a Schur matrix (all eigenvalues are in the interior of the unit disk).

The constraints are imposed on the output variables, $y(t)$, and have the form,

$$y(t) \in Y \text{ for all } t \in Z^+,$$  \( (2) \)

where $Y \subset R^p$ is a prescribed set, with $0 \in intY$. While not required, for computational reasons constraints are often defined so that $Y$ is a polytope (compact set defined by a set of linear inequalities).

Since (1) is a model of the closed-loop system, (2) can represent constraints on either state or control variables. For instance, a control constraint $|u_1(t)| \leq 1$ where the control is generated by a state feedback law, $u = Kx$, can be restated as $y_1(t) \leq 1$, $y_2(t) \leq 1$ with $y_1 = \hat{e}_1 K$, $y_2 = -\hat{e}_1 K$, and $\hat{e}_1 = [1 \ 0 \ldots 0]$.

The Scalar Reference Governor (SRG) uses the following update [46], [49], [47]:

$$v(t) = v(t - 1) + \kappa(t)(r(t) - v(t - 1)).$$  \( (3) \)

In (3), $\kappa(t)$ is a scalar adjustable bandwidth parameter, $0 \leq \kappa(t) \leq 1$. If no danger of constraint violation exists, $\kappa(t) = 1$, and $v(t) = r(t)$ so that the reference governor does not interfere with the operation of the system. If a potential for constraint violation exists, the value of $\kappa(t)$ is decreased by the reference governor. In the extreme case, $\kappa(t) = 0$, $v(t) = v(t - 1)$ so that the reference governor isolates the system from further application of reference command to ensure safety.

![Reference/command governor applied to closed-loop (Plant + Controller) system with constraints.](image-url)
Assuring safety involves a constraint,

\[ (v(t), x(t)) \in P, \]  
where \( P \subset \mathbb{O}_w \subset \mathbb{R}^m \times \mathbb{R}^p \). The set \( \mathbb{O}_w \) is the maximum output admissible set, i.e., the set of all states \( x(t) \) and constant inputs, \( v(t+k) = v(t) = \bar{v} \), such that for all disturbances satisfying \( w(t+k) \in W \), \( k \geq 0 \) the subsequent response satisfies the constraints

\[ \mathbb{O}_w = \{(\bar{v}, x(t)) : y(t+k) \in Y \forall w(t+k) \in W, \ v(t+k) = \bar{v}, \ \forall k \in \mathbb{Z}^+ \}. \tag{5} \]

The choice of \( P = \mathbb{O}_w \), where \( \mathbb{O}_w \) is a slightly tightened version of \( \mathbb{O}_w \) is frequently made in (4), where

\[ \mathbb{O}_w = \mathbb{O}_w \cap \mathbb{O}_w^f \]  
and \( \mathbb{O}_w^f \) is the set of commands such that the associated steady state constrained output \( (D + C(I - A)^{-1}B) \bar{v} \) satisfies constraints with a margin \( \epsilon > 0 \) (typically small),

\[ \mathbb{O}_w^f = \{(\bar{v}, x(t)) : (D + C(I - A)^{-1}B) \bar{v} \in (1 - \epsilon)Y \}. \tag{7} \]

Assuming that \( Y \) is a polytope, the pair \((C, A)\) is observable and the minimum invariant set for (1) with \( v = 0 \) is strictly constraint admissible, i.e.,

\[ CF + DW \subset \text{int}Y, \quad F = \bigoplus_{k=0}^{\infty} A^kBW, \]

it can be shown that \( \mathbb{O}_w \) is non-empty, positively-invariant (with \( v(t) \) maintained constant) and finitely-determined polytope, representable by linear inequalities of the form,

\[ \mathbb{O}_w = \{x(0), v) : H_s x(0) + H_v v \leq s \}. \tag{8} \]

Thus the set \( \mathbb{O}_w \) is a finitely-determined inner approximation of \( \mathbb{O}_w \) which can be made arbitrary close to \( \mathbb{O}_w \) by decreasing \( \epsilon \). Procedures for computing \( \mathbb{O}_w \) are detailed in [45] and [79].

With \( \mathbb{O}_w \) computed off-line as (8) and \( P = \mathbb{O}_w \) in (4), the selection strategy for \( \kappa(t) \) at the time instant \( t \) is based on maximizing \( \kappa(t) \), subject to the constraints that

\[ 0 \leq \kappa(t) \leq 1 \]

and

\[ (v(t), x(t)) \in \mathbb{O}_w. \]

Due to the positive invariance of \( \mathbb{O}_w \), recursive feasibility is maintained at each time step: the value of \( \kappa(t) = 0 \) remains a feasible solution of the above optimization problem provided it is feasible at the initial time. Since only the scalar parameter \( \kappa(t) \) is optimized on-line, the computational complexity of this approach is minimal and, in fact, the optimization problem is explicitly solvable [47]. If the system model or the constraints change, \( H_s, H_v \) and \( s \) can be computed on-line.

To further illustrate the conditions on \( \kappa(t) \), consider the case \( W = \{0\}, \ Y = \{y \in \mathbb{R}^p : 5y \leq s \}. \) Then the condition

\[ (v(t), x(t)) \in \mathbb{O}_w \text{ reduces to} \]

\[ S(CA^k x(t) + C(I - A)^{-1}(I - A^k)Bv(t)) \leq s, \tag{9} \]

\[ S(C(I - A)^{-1}Bv(t)) \leq s(1 - \epsilon), \tag{10} \]

\[ k = 0, \ldots, k^* \tag{11} \]

where \( k^* \) is any upper bound on the finite-determination index [49], [47] and \( 1 > \epsilon > 0 \). Hence the conditions on the scalar \( \kappa(t) \) take the following form,

\[ \kappa(t)H_s(k)(r(t) - \nu(t - 1)) \leq s - H_s(k)x(t) - H_v(k)v(t - 1), \tag{12} \]

\[ \kappa(t)H_v(\infty)(r(t) - \nu(t - 1)) \leq s(1 - \epsilon) - H_v(\infty)v(t - 1), \tag{13} \]

\[ k = 0, \ldots, k^* \]

for appropriately defined matrices \( H_s(k), H_v(\infty) \) and \( H_v(k) \) based on (9). These matrices can be either pre-computed off-line and stored (while eliminating redundant constraints to simplify the representation and the associated computations), or computed on-line in case the model or the constraints undergo on-line changes. Each of the above inequality conditions and 0 \( \leq \kappa(t) \leq 1 \) bound \( \kappa(t) \in [0, \kappa_{\text{max}}(t, k)] \), \( k = 0, 1, \ldots, k^* \) and \( k = \infty \). Therefore, \( \kappa(t) \) is set to the minimum of \( k^* + 1 \) numbers \( \kappa_{\text{max}}(t, k) \). Similar conditions on \( \kappa(t) \) are obtained if \( W \neq \{0\} \), see [47] for details.

Further simplifications occur with \( P \neq \mathbb{O}_w, P \subset \mathbb{O}_w \) in (4). While \( P \) has to satisfy certain assumptions, it is not required to be positively-invariant and can be much simpler than \( \mathbb{O}_w \). With the general \( P \), a situation that no feasible \( \kappa(t) \) yielding (4) exists can occur. In such a case, \( \kappa(t) = 0 \) is chosen. Following this procedure, the constraint enforcement and usual response properties of the reference governor to constant inputs are maintained. In fact, \( P \) can be obtained from \( \mathbb{O}_w \) by the systematic elimination of almost redundant constraint and applying a pull-in procedure, see [47]. This strategy can lead to a ten-fold reduction in the on-line computing effort with some loss in performance, see e.g., [117].

As an illustration, we consider a reference governor designed for a double integrator system. The system model in continuous-time has the following form,

\[ \dot{x}_1 = x_2, \]
\[ \dot{x}_2 = u. \]

The model is converted to discrete-time assuming a sampling period of \( T_s = 0.1 \) sec. A nominal controller is defined as

\[ u = -0.917(x_1 - v) - 1.636x_2, \]

where \( v \) is a set-point for \( x_1 \). The state and control constraints are imposed as

\[ |x_1| \leq 1, \ |x_2| \leq 0.1, \ |u| \leq 0.1. \]

The initial state is \( x_1(0) = x_2(0) = 0 \). For small \( r \), the nominal response with \( v = r \) satisfies the constraints and the reference governor remains inactive. For a larger command, \( v = r = 0.5 \), the nominal controller results in the maximum excursion of \( x_2 \) of about 0.2 and \( u \) of about 0.46, both violating the
imposed constraints by large amounts. To avoid de-tuning the nominal controller (and thus compromising the response for small commands that do not cause constraint violation), we apply SRG to govern \(v\). The \(P = \tilde{O}_\infty\) is used in the reference governor implementation, which is a polytope defined by 36 linear inequalities. The closed-loop responses of the system to the command \(r = 0.5\), after augmentation with SRG, are given in Fig. 2 and compared to the unconstrained response. The effect of the reference governor is to slow down the command and the subsequent system response. Note that the modified reference command signal \(v(t)\) converges to the command \(r = 0.5\) in a finite time.

Fig. 2. The closed-loop response of the double integrator plant with the reference governor and the unconstrained response.

Response properties of the reference governor, including conditions for the finite-time convergence of \(v(t)\) to \(r(t)\), are detailed in references [49], [47]. Essentially, if \(r(t)\) for \(t \geq t_0\) remains constant or varies in a sufficient small neighborhood of a constant value, then \(v(t)\) converges to \(r(t)\) in a finite-time if \(r(t)\) is steady-state constraint admissible. If \(r(t)\) is not steady-state constraint admissible then \(v(t)\) will converge to the closest feasible value in a finite time. Similar finite-time convergence results can be developed for sufficiently slowly varying \(r(t)\). The finite-time convergence is a desirable property indicating that after transients caused by large changes in \(r(t)\), the reference governor becomes inactive and nominal closed-loop system performance is recovered. This or similar properties are retained by other governors discussed next.

A. Vector Reference Governor (VRG)

The Vector Reference Governor (VRG) approach is a modification of (3) for \(m > 1\). The VRG uses a diagonal matrix \(K(t)\) in place of a scalar \(\kappa(t)\) to decouple the governing of different channels [49]:

\[
v(t) = v(t-1) + K(t)(r(t) - v(t-1)),
\]

where \(K(t) = \text{diag}(\kappa_i(t))\). The values of \(\kappa_i(t), i = 1, \cdots, n\), are chosen to minimize \(\langle v(t) - r(t) \rangle^T Q(v(t) - r(t))\), where \(Q = Q^T > 0\) with \(v(t)\) given by (12), subject to the constraints \(0 \leq \kappa_i(t) \leq 1, i = 1, \cdots, n\), and \((v(t),x(t)) \in \tilde{O}_\infty\). This optimization problem can be solved online using quadratic programming techniques. Online use of conventional iterative techniques can be avoided by using explicit multi-parametric quadratic programming [13], [71], [72]. The VRG is superior to the SRG in that it offers more flexibility in the choice of \(v(t)\). In applications, it can provide faster response than the SRG.

B. Command Governor (CG) and Extended Command Governor (ECG)

The command governor and extended command governor approaches were proposed in [10], [23], [50], see also references therein. The simplest variant of the command governor is similar to the VRG. In it, a cost function,

\[
J = ||v(t) - r(t)||_Q^2 = (v(t) - r(t))^T Q (v(t) - r(t)),
\]

where \(Q = Q^T > 0\), is minimized with respect to \(v(t)\) subject to the constraint that

\[
(v(t),x(t)) \in \tilde{O}_\infty.
\]

In case \(\tilde{O}_\infty\) is polyhedral, \(v(t)\) is computed by solving a quadratic programming problem either online or offline as a piecewise affine function of \(x(t), r(t)\) using multi-parametric programming techniques [13].

A variant of the CG, the so called Prioritized Reference Governor (PRG) has been proposed in [62]. The PRG enforces hard constraints and it satisfies soft constraints in the order of priority. The soft constraints are relaxed by slack variables and the penalty on the slack variables is added to the cost with lower weights corresponding to lower priority constraints.

The extended command governor approach of [50] is an extension of the command governor approach that generates \(v(t)\) according to,

\[
v(t) = \tilde{C} \tilde{x}(t) + \rho(t),
\]

where, over the prediction horizon, the fictitious state, \(\tilde{x}(t) \in R^n\), and \(\rho(t)\) evolve as,

\[
\tilde{x}(t+k+1) = \tilde{A} \tilde{x}(t+k), k \geq 0,
\]

\[
\rho(t+k) = \rho(t).
\]

The values of \(\rho(t)\) and \(\tilde{x}(t)\) are optimized using a quadratic cost function,

\[
J = \frac{1}{2} ||\rho(t) - r(t)||_P^2 + \frac{1}{2} ||\tilde{x}(t)||_P^2,
\]

with \(P = P^T > 0\), satisfying \(\tilde{A}^T P \tilde{A} - P < 0\), subject to the constraint that \((\rho(t),\tilde{x}(t),x(t)) \in \tilde{O}_\infty\). The set \(\tilde{O}_\infty\) is a finitely determined inner approximation to the set of all triplets \((\rho(t),\tilde{x}(t),x(t))\) that do not induce subsequent constraint violation when the input sequence \(v(t+k)\) is determined by the fictitious dynamics per (13) and (14). Without the fictitious states, i.e., when \(\bar{n} = 0\), \(\tilde{O}_\infty = \tilde{O}_\infty\), the ECG becomes the simple command governor [50]. The optimization problem can be solved online using conventional quadratic programming.
techniques. Again iterative procedures can be avoided by using explicit multi-parametric quadratic programming [13].

Various choices of $\hat{A}$ and $\hat{C}$ in (13), (14) can be made. The shift sequences used in [50] are generated by the fictitious dynamics with,

$$\hat{A} = \begin{bmatrix} 0 & I_m & 0 & 0 & \cdots \\ 0 & 0 & I_m & 0 & \cdots \\ 0 & 0 & 0 & I_m & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$\hat{C} = \begin{bmatrix} I_m & 0 & 0 & \cdots \end{bmatrix},$$

where $I_m$ is an $m \times m$ identity matrix. In this case, ECG can be re-formulated as a Model Predictive Controller (MPC). Another approach [60], motivated by [102], uses Laguerre sequences. These sequences possess orthogonality properties and are generated by the fictitious dynamics with,

$$\hat{A} = \begin{bmatrix} \alpha I_m & \beta I_m & -\alpha \beta I_m & \alpha^2 \beta I_m & \cdots \\ 0 & \alpha I_m & \beta I_m & -\alpha \beta I_m & \cdots \\ 0 & 0 & \alpha I_m & \beta I_m & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$\hat{C} = \sqrt{\beta} \begin{bmatrix} I_m & -\alpha I_m & \alpha^2 I_m & -\alpha^3 I_m & \cdots \end{bmatrix},$$

where $\beta = 1 - \alpha^2$, and $0 \leq \alpha \leq 1$ is a selectable parameter that corresponds to the time-constant of the fictitious dynamics. Note that with the choice of $\alpha = 0$, (16) coincides with the shift register considered in [50]. The advantage of the ECG is that it can produce larger domain of attraction for $x(0)$ than is possible using the SRG, VRG, or CG.

III. REFERENCE GOVERNORS FOR NONLINEAR SYSTEMS

In the case of nonlinear system models, the design of the reference governor can be based on model linearization at a selected operating point, or by applying the reference governor to the nonlinear model directly.

A. Reference governor design based on model linearization

In the case model linearization is employed in the form (1), the differences between the nonlinear model and the linearized model can be compensated by modifying the output of (1) with a constant offset term, $g(t)$, [117], [60],

$$y(t) = Cx(t) + Dv(t) + Dw(t) + g(t).$$

The offset term is then augmented to the linearized system model as an extra state, with the assumed dynamics,

$$g(t+k) = g(t).$$

During the on-line operation, $g(t)$ is set to the difference between the currently measured or estimated actual output and its prediction based on the linear system model, i.e.,

$$o(t) = y_{\text{nonlinear}}(t) - y_{\text{linear}}(t).$$

This tightens the constraints and serves as a practical measure to protect against constraint violation. In the case constrained outputs are not directly measured, they can be estimated with state or input observers, see e.g., [63] for an example. While errors can be tolerated in prediction, accurate measurements of estimates of current values of constrained outputs are useful for non-conservative treatment of constraints.

B. Reference governor design based on nonlinear model

Approaches to reference governor design for constraint enforcement based on nonlinear closed-loop system models include [4], [7], [48], [51], [95]. Several of the reported approaches exploit predictive on-line simulations or sub-level sets of Lyapunov functions to guard against constraint violation. An incremental step reference governor strategy with one update of fixed magnitude per time step has been investigated in [116]. Landing reference governors have been proposed for systems with terminal constraints, e.g., for reaching a desired position with a small velocity in mechanical systems [76], [77], [67]. Strategies for the design of reference governors based on approximating nonlinear models by piecewise affine models have been reported in [15]. Reference [5] compares direct nonlinear versus linearization-based treatment.

In a typical setup, the nonlinear system model has the form,

$$x(t+1) = f(x(t), v(t), w(t)),$$

with the constraints

$$y(t) = h(x(t)) \leq 0.$$ 

The input $w$ is assumed to be stationary, e.g., it may represent a parametric uncertainty, a set-bounded or a set-bounded and a rate-bounded unmeasured disturbance.

Let a function $S$ be such that with $v(t+k) = \bar{v}$ for all $x(t)$ in the sub-level set of $S$, i.e., satisfying

$$S(x(t), \bar{v}) \leq 0,$$

the subsequent trajectory is guaranteed to be safe and strongly returnable. The trajectory is safe if the constraints $y(t+k) \in Y$ hold for all $k \in Z^+$. The trajectory is strongly returnable [48], if there exists $k^*$ that does not depend on $x(t)$ or $\bar{v}$ and $0 < k < k^*$ such that $S(x(t+k), \bar{v}) < 0$. Note that no requirement is made on the sub-level set (20) to be invariant. Technical conditions require $S$ to be continuous but it can be (and often is) non-smooth.

With $S$ defined, $\kappa(t)$ in the SRG (3) can be chosen based on solving a scalar optimization problem,

$$\kappa(t) = \max \{ \lambda \in [0,1] : S(x(t), v(t-\lambda)) \leq 0 \}.$$ 

If no feasible solution to (21) exists, $\kappa(t) = 0$ is used.

Several methods to construct $S$ exist. If Lyapunov or Input-to-State Stable (ISS) Lyapunov functions, $V$, for the closed-loop system are available then typically

$$S(x, v) = V(x, v) - c,$$

with an appropriately chosen $c > 0$. This approach exploits the positive-invariance properties of sub-level sets of Lyap-
First, instead of developing an explicit functional representation for $S$, $S$ can be also defined implicitly [7] as

$$S(x(t), \bar{v}) \leq \max_{i=1, \ldots, k^*} h_i(x(t+k)(x(t), \bar{v}, w(t))),$$

where $x(t+k)(x(t), \bar{v}, w(t))$ is the predicted trajectory emanating from the state $x(t)$ with $v(t+k) = v(t)$ and $w(t+k) \in W$. The horizon $k^*$ must be sufficiently long, see [7]. The determination of $S$ based on (22) reduces to an on-line simulation if $W = \{0\}$; if $W \neq \{0\}$, in evaluating (22) either multiple scenarios of $w(t)$ are considered or a set of optimal control problems with respect to $w(t)$ (corresponding to $0 < k \leq k^*$ and to different components of the output vector) is solved.

The online optimization of a scalar parameter $\lambda$ in (21) is generally simple. Procedures based on bisections or grid search (while checking $x(t)$ = 1 first) can be employed. The incremental reference governor approach of [116] distributes the solution over time by checking the feasibility of a single value of $v(t)$ that differs from $v(t-1)$ by a fixed step size. In case $S$ is quadratic, an explicit solution for $\lambda$ in (21) is easily derived; in other cases, numerical solution strategies such as the one proposed in [86] based on predictor-corrector form of Newton’s method applied to parameteric root-finding can be used.

The response properties of the nonlinear reference governor are similar to the ones in the linear case: If an initial selection of $v(0)$ exists such that $S(x(0), v(0)) \leq 0$ then constraints (19) will be satisfied for all $t \in \mathbb{Z}^+$ and $v(t)$ will converge to constant steady-state constraint admissible $r(t)$ in a finite-time. See [48] for the precise statement of the relevant conditions.

### C. Robust reference governor

In the robust reference governor approach of [108], the model (18) has no time-dependent disturbance input but depends on uncertain parameters, $\theta \in \mathbb{R}^l$,

$$x(t+1) = f(x(t), v(t), \theta),$$

and constraints

$$y(t) = h(x(t)) \leq 0.$$  

In verifying the feasibility of $x$, the range of $\theta \in \Theta \subset \mathbb{R}^l$ is covered by a grid of values, $\theta^j$, and sets $\Theta^j, j = 1, \ldots, n_\theta$. The admissibility of given $x(t)$, $v$ is tested based on whether the predicted response satisfies the constraints (24). In this prediction, (24) is replaced by a set of conditions

$$y^j(t; \theta) = y(t; \theta^j) + \sum_{i=1}^l z^j_i(t)(\theta_i - \theta_i^j) + M||\theta - \theta^j||^2 \leq 0,$$

where the $j$th condition must hold for all $\theta \in \Theta^j$. In (25), $z^j_i(t) = \frac{\partial f}{\partial \theta_i}(x^j(t), v, \theta^j)$ is computed based on the parametric sensitivity equations,

$$z^j_i(t+1) = \frac{\partial f}{\partial x}(x^j(t), v, \theta^j)z^j_i(t) + \frac{\partial f}{\partial \theta_i}(x^j(t), v, \theta^j),$$

$$z^j_i(0) = 0.$$  

The expression (25) is based on Taylor series expansion of the solution, with $M > 0$ chosen to protect against the inaccuracies of this approximation. Often the implementation with $n_\theta = 1$ suffices.

### D. Reference governor for linear systems with nonlinear constraints

In [59] a special case of reference governor design based on a linear system model with nonlinear constraints has been studied. The nonlinear constraints are given by

$$Y = \{y: \ h_i(y) \leq 0, \ i = 1, \ldots, q\},$$

where $h_i$ are nonlinear functions. Feedback linearization [70] can be applied to many systems to render the closed loop system linear, however, the resulting $Y$ is typically non-polyhedral.

The existing reference governor results in [49], [47] for system (1) with constraints (28) hold for any compact, convex set $Y$ with $0 \in \text{int}Y$. When $Y$ is not polyhedral, it can often be approximated by a polyhedron, however, such approximations may lead to reference governor designs that are either保守 or have high computational complexity especially for systems with multi-dimensional constrained outputs. Instead of employing polyhedral approximations, in [59] the linear model (1) is used to predict $y(t)$ and the predicted violations of $h_i(y(t+k)) \leq 0$ for $k > 0$ are used to constrain $\kappa(t)$ in (3). With this approach, several types of constraints can be treated, including convex constraints, convex quadratic constraints, Mixed-Logical-Dynamic (MLD) constraints of if-then type and concave constraints. For instance, in the case of quadratic constraints, $\kappa(t)$ is given by an explicit formula. In the case of concave constraints, constraint linearization is employed, leading to replacing the condition $y(t+k|t) \in Y$ by

$$y(t+k|t) \in Y_r,$$

where

$$Y_r(t) = \{y: \ h_i(y_i(t)) + h_j(y_j(t)|t) - y_j(t) \leq 0\},$$

and $y_i(t)$ is appropriately chosen, $i = 1, \ldots, r$. Finally, the
landing reference governor handles if-then terminal constraints in systems requiring the soft-landing of the components [76], [77], [67].

**E. Parameter governor**

Parameter governors [80], [81], [82] adjust parameters, \( \theta(t) \in \Theta \), in nominal control laws to optimize predicted system response over a finite, receding horizon subject to constraints. Parameters are assumed to remain constant over the prediction horizon and the cost of the general form,

\[
J(t) = ||\theta||_{\tilde{Q}}^2 + \sum_{k=0}^{\infty} \Omega(x(t+k|t), \theta(t), r(t)),
\]

penalizes the system response as well as parameter deviations. The assumption of constant parameters over the prediction horizon reduces computational and implementation effort, and simplifies the analysis. In fact, \( \Theta \) can be a finite set so the evaluation of (31) and constraints reduces to several simulations.

Specific parameter governor schemes considered in [80] include the feed-forward governor and the gain governor. For these schemes terminal set conditions need not be imposed to assure stability provided the horizon is chosen sufficiently long and in agreement with the appropriate assumptions.

In the feed-forward governor approach of [80], a disturbance free system is considered with an integrator included as a part of the overall system,

\[
\begin{align*}
\dot{x}(t+1) &= f(x(t), u(t)) \\
\dot{x}_c(t+1) &= x_c(t) + z(t) - r \\
z(t) &= h_c(x(t)),
\end{align*}
\]

where \( z \) is an auxiliary output which is, in general, different from the constrained outputs. The control law includes integral action and an adjustable feed-forward offset \( \theta(t) \),

\[
u(t) = u_c(r) - \varepsilon x_c(t) + \tilde{u}_{f0}(x(t), r) + \theta,
\]

where \( x_c(r), u_c(r) \) denote the equilibrium values of state and control variables corresponding to the given \( r \). Due to the use of integral action, if \( \theta \) is constant, then as \( t \to \infty \) it follows that \( z(t) \to r, x_c(t) \to x_c(r), u(t) \to u_c(r) \). The small gain integral control leads to dynamics decomposition into slow and fast modes. The fast dynamics can be made to avoid constraints by changing \( \theta(t) \); consequently, the constraints will be satisfied if slow manifold is well within the constraint admissible region. The feedforward governor of [80] can thus handle large reference changes and recover a large set of initial states. The cost (31) is modified to include the penalty on the integral states.

In the gain governor approach of [80], the adjustable parameters are basically the control gains

\[
\begin{align*}
\dot{x}(t+1) &= f(x(t), u(t)) \\
\dot{u}(t) &= u_c(r) + \tilde{u}_{f0}(x(t), r, \theta(t)),
\end{align*}
\]

where \( \tilde{u}_{f0}(x_c(r), r, \theta) = 0 \) for all \( \theta \in \Theta \).

**F. Other developments**

Other approaches to reference governor design for nonlinear systems are possible. We mention the developments in [39] (and early paper [38] for linear systems) and [55] as specific examples.

**IV. RECENT REFERENCE GOVERNOR DEVELOPMENTS**

In this section, we discuss several more recent reference governor developments; these include feedforward reference governor, reduced order reference governor, reference governor for decentralized systems, and reference governor for networked systems.

**A. Feedforward reference governor**

Most of the command/reference governor approaches presented in the literature make explicit use of state measurements or of suitable state estimations, see e.g. [6], to modify the reference in order to ensure constraints satisfaction.

However, as recently remarked in some papers, sensorless command/reference governor schemes that do not make explicit use of the plant state are possible, at the price of some additional conservativeness. This is not surprising as feedback is not a necessary requirement in many classical solutions to manage the reference, e.g. when filtering the reference signal to avoid high frequency responses.

The main idea behind feedforward command/reference governor approaches is that, if sufficiently slow transitions in the set-point modifications are applied by the reference management unit, one can have a high confidence on the expected value of the state, even in the absence of an explicit measurement of it, because of the asymptotical stability of the pre-compensated system at hand. This feature may be of interest for all those applications where either the measurement or the estimation of the state may be difficult or unsuitable. For instance, it is a feature could help the design of those multi-agent distributed supervisory schemes where to know the entire aggregate state (or part of it) at each time instant can be very costly or even unrealistic.

A feedforward command governor scheme was first introduced in [42] as the first block within a decentralized command governor scheme. In that paper, the main idea is to choose at each time instant an input \( v(t) \) such that the corresponding steady-state equilibrium would satisfy constraints (i.e., \( v(t) \in \hat{O}_c \)) where \( \hat{O}_c \) is defined in (7) and to modify \( v(t) \) ‘slow enough’ so as to ensure that the constrained output is always ‘close’ to the steady-state equilibrium. To do so, two expedients are used:

- The input \( v(t) \) is changed every \( \tau \) steps and kept constant in between, where \( \tau \) is a generalized settling time with parameter \( 0 < \gamma < 1 \) for the system (see [43] for details);
- The variation \( v(t+\tau) - v(t) \) is constrained to belong to a static precomputed set \( \Delta V \) which, in combination with the generalized settling time, ensures that the output trajectory is always inside a ball of radius \( \varepsilon \) from the steady state.
This very early feedforward scheme was proven to have the same theoretical properties as the classical command governor, but also turned out to be quite conservative. In [43], the conservatism of the approach has been substantially reduced by constraining \( v(t + \tau) - v(t) \) to a new set \( \Delta V(v(t - \tau), \rho(t)) \) which depends on the previously applied command \( v(t - \tau) \) and on the scalar \( \rho(t) \), representing an estimate of the maximal possible distance between the actual value of the output and the steady-state value associated with the input \( v(t - \tau) \). A further reduction of conservatism has been achieved in [32], where a scheme that allows the command governor unit to modify the command at each time instant has been presented. This scheme has been proved to be asymptotically equivalent to a command governor with feedback in the case of non-noisy systems. Please note that, as shown in the same paper, feedforward command/reference governors tend to behave quite poorly with plants subject to important disturbances, as the only mean of disturbance rejection is the stability of the matrix \( A \).

B. Reduced order reference governors

The reduced order reference/command governor [61] for systems with states decomposable into “slow” and “fast” states can be based on the reduced order model for “slow” states, provided constraints are tightened to ensure that the contributions of fast states do not cause constraint violation.

Consider system (1), where we assume that \( W = \{0\} \) for the ease of exposition. After appropriate state transformations, the decomposed system has the following state space realization,

\[
\begin{bmatrix}
    x_s(t + 1) \\
    x_f(t + 1)
\end{bmatrix} =
\begin{bmatrix}
    A_s & 0 \\
    0 & A_f
\end{bmatrix}
\begin{bmatrix}
    x_s(t) \\
    x_f(t)
\end{bmatrix} +
\begin{bmatrix}
    B_s \\
    B_f
\end{bmatrix} v(t),
\]

\[
y(t) =
\begin{bmatrix}
    C_s & C_f
\end{bmatrix}
\begin{bmatrix}
    x_s(t) \\
    x_f(t)
\end{bmatrix} + D v(t),
\]

where \( x_s(t) \in \mathbb{R}^{n_s} \) and \( x_f(t) \in \mathbb{R}^{n_f} \) are the slow and fast vector states, respectively, \( n_s + n_f = n \), and the matrices are appropriately sized.

The reference governor design is based on the reduced order system model, representing the dynamics of the slow states and steady-state values of fast states,

\[
x_s(t + 1) = A_s x_s(t) + B_s v(t),
\]

\[
y_r(t) = C_s x_s(t) + (C_f \Gamma_f + D) v(t),
\]

where \( \Gamma_f = (I_{n_f} - A_f)^{-1} B_f \).

In order for the reference governor based on the reduced order model, (37)-(38), to enforce the constraints for the full order model, we tighten the constraints. Specifically, we introduce two “error sets” \( E_x \subseteq \mathbb{R}^{n_f} \) and \( E_y \subseteq \mathbb{R}^{n_f} \). We require that \( E_x \) be \( A_f \)-invariant, i.e.,

\[
A_f E_x \subseteq E_x,
\]

and also satisfy the following inclusion,

\[
C_f E_x \subseteq E_y.
\]

The the reduced order reference governor enforces tightened constraints on the reduced order output,

\[
y_r(t + k) \in Y \sim E_y, \forall k \geq 0.
\]

To ensure that the contributions of the fast states do not lead to constraint violation, the reduced order reference governor enforces an extra constraint of the form,

\[
-A_f \Gamma_f \Delta v(t) \in E_x \sim A_f E_x.
\]

where \( \Delta v(t) = v(t) - v(t - 1) \). Here, \( \sim \) denotes the \( P \)-difference of two sets\(^1\). The constraint (42) can be re-written as a constraint on \( \kappa(t) \),

\[
-A_f \Gamma_f \kappa(t)(v(t) - v(t - 1)) \in E_x \sim A_f E_x.
\]

If \( v(-1) \) is initialized so that \( e_f(0) = x_f(0) - \Gamma_f v(-1) \in E_x \), the reference governor response properties [47] hold, specifically, the recursive feasibility of \( \kappa(t) = 0 \), the guaranteed constraint enforcement and the finite-time convergence for constant set-points. This approach of imposing ancillary constraints on the evolution of \( v(t) \) can be applied similarly to handling observer errors and to the design of command governors and extended command governors [61].

C. Network reference governor handling variable delays

Reference governor-based approaches for networked control systems have been proposed in [8], [25], [26] and more recently in [16], [17]. In these approaches, the controller and the plant are connected via a, usually non-ideal, communication network. In [16], [17], the set-point commands, \( v(t) \), are transmitted through a communication channel that has variable continuous time delay, \( \delta(t) \in [0, \delta] \). When the delay \( \delta(t) \) is a smaller than the reference governor update period, \( T_r \), the effect of the delay on the state is shown in [16] to satisfy the following relation:

\[
x(t + 1) = A x(t) + B v(t) + R(\delta(t)) \Delta v(t),
\]

\[
\Delta v(t) = v(t + 1) - v(t),
\]

where

\[
R(\delta) = - \int_0^\delta e^{A_v(T_r - \tau)} B_v d\tau,
\]

and \( (A_v, B_v) \) are the continuous-time realization of the system controlled through the communication network. Consequently, the effective disturbance introduced by the delay in (43) is modulated by \( \Delta v(t) \), i.e., change in \( \Delta v \).

To overbound \( R(\delta) \Delta v \), suppose a matrix \( P \) and a set of vertices \( \{w_i, i = 1, \cdots, n_w\} \) is given such that \( ||P \Delta v||_\infty \leq 1 \) implies \( R(\tau) \Delta v \in convh\{w_i, i = 1, \cdots, n_w\} \) for all \( 0 \leq \tau \leq T_r \), where \( convh \) denotes the convex hull. Then the reference governor algorithm that ensures constraint enforcement and finite time convergence properties has the following form

\(^1\sim A := \{a \in A : a + b \in A, \forall b \in B\}. \) See [79].
\[
\begin{align*}
\min_{v(t)} & \ |r(t) - v(t)|^2, \\
\text{subject to} & \quad H_x(Ax(t) + Bv(t) + H_v v(t)) \leq s - \max_{i=1,\ldots,n_u} (H_u w_i) \zeta (44)
\end{align*}
\]
where \(H_x\), \(H_v\), and \(s\) are matrices in the representation of \(\mathcal{O}_\infty = \{ (x,v) : H_x x + H_v v \leq s \} \).

While originally developed for network control systems, the network reference governor can prove useful in any application when the time delay is time varying. Several extensions of these results are developed in [16], [17]. They include (see [16]) the treatment of the combined delays in feedback and feed-forward channels, the output measurement case, and a longer (possibly unbounded) delay for which a simple command acceptance/rejection logic is implemented at the plant side. A further extension is developed in [17], where the delay is assumed to be slowly varying with a known bound on the time rate of change and known at the time instant the command is sent. The approach combines the Smith predictor based on the estimated delay value and additive disturbance to cover the effect of the delay uncertainty.

D. Virtual state governor for integrating existing controllers

While based on principles similar to the ones of the reference governor, the virtual state governor (VSG) [21] aims at solving a slightly different problem, namely how to integrate multiple actuators, each equipped with an assigned non-modifiable feedback control law. The obtained control architecture must be capable of effectively exploiting all the actuators at the same time if needed, but it also needs to minimize the use of those actuators that are “expensive” to operate. This problem is of interest in automotive applications, for instance in cornering control [19], engine control [18], and energy management in hybrid powertrains [20], and in aerospace applications such as attitude control [21].

For the case where the actuators are divided into two groups we have
\[
x(t+1) = Ax(t) + B_1 u_1(t) + B_2 u_2(t),
\]
where \(x\) is the local state, \(y_i\) is the local constrained output, and \(v_i\) the local manipulable reference vector which, if no constraints (and no command governor) were present, would coincide with the desired reference \(r_i\). Each subsystem has its own reference signal and is governed by a local reference management unit. The management units are organized in a communication network described by an undirected graph \(\Gamma = (\mathcal{N}, \mathcal{E})\), where the set of edges \(\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}\) describes the existence of communication links among the units. Common assumptions for the schemes developed so far are that each unit may exchange information only with its neighbors (i.e. the agent \(i\) may communicate with the node \(j\) only if \((i,j) \in \mathcal{E}\)) and that each couple of nodes may communicate only once between two time steps.
The goal of decentralized command governor is to determine a distributed reference management strategy able to locally modify the commands \( v_i, i = 1, ..., N \) in such a way that:

- the application of all the \( v_i, i = 1, ..., N \) is such that the aggregated output \( y = [y^T_1, ..., y^T_N]^T \) do not produce violation of global (coupling) constraints \( y(t) \in Y, \forall t \)
- each \( v_i(t) \) approximates as close as possible \( r_i(t) \).

The first solutions proposed to solve this problem made use of the feedforward command governor approach [42]. The use of the feedforward approach in its simpler formulation, allows one to reformulate the decentralized reference management problem as the static problem of locally determining every \( \tau \) steps the commands \( v_i, i = 1, ..., N \) such that the aggregated vector \( \nu = [\nu^T_1, ..., \nu^T_N]^T \) belongs to the static sets \( \mathcal{O}_\nu \) and such that the variation of \( \nu \) within two update times is constrained as \( v(t + \tau) - v(t) \in \mathcal{A} \).

To solve this problem the general idea is that at each time instant:

- on the basis of the information common to all agents, each agent computes locally the same family of local artificial regions \( V_i(t), i = 1, ..., N \). These regions are convex and compact sets containing the origin as an internal point and are such that

\[
\Delta V_i(t) \times \cdots \times \Delta V_N(t) \subseteq \Delta V \cap \{\mathcal{O}_\nu \sim \{v(t - \tau)\}\}.
\]

which means that whenever \( v_i(t) \in \Delta V_i(t), i = 1, ..., N \), constraints are satisfied.

- On the basis of the locally available information, each agent computes the command to be applied minimizing its local cost function \( \|g_i(t) - r_i(t)\|^2 \), with matrix \( \Psi_i = \Psi^T > 0 \), and such that \( v_i(t) \in \Delta V_i(t) \).

Following this philosophy, some sequential and parallel approaches have been introduced. Sequential approaches [33] are schemes where only one agent at the time is allowed to modify its command. Parallel approaches [34] are schemes where all the agents are allowed to move the command at the same time, making worst case assumptions on the choices of the others. This second approach has proven to work quite well when the aggregated command \( \nu \) is far from the borders of \( \mathcal{O}_\nu \), but quite poorly close to them. For this reason, hybrid approaches switching between parallel and sequential modes have been proposed in [44] and [111]. We note that while all the above mentioned schemes ensure constraints satisfaction, they may experience problems of convergence to “good approximations” of the desired reference signals \( r_i(t), i = 1, ..., N \). In fact, as shown in [33], these schemes may experience convergence of the command to Nash equilibria that are not Pareto optimal. In [35] and [110] this phenomenon has been carefully investigated and some algorithms to check the existence of these anomalies and to eliminate them are provided. It is also worth mentioning that, following the same idea of the above mentioned feedforward distributed command governor strategies, decentralized schemes making use of the state have been recently presented in [112] and [36].

\[ F. \] Other developments

We mention also approaches in [109] that combine reference governing and controller switching to improve performance. Related strategy is also used in [83].

\[ V. \] Automotive Applications

In this section we discuss several applications of reference and command governors to practical systems arising in automotive applications.

\[ A. \] Turbocharged automotive engines

As gasoline engines are downsized and turbocharged to improve fuel consumption, protecting the engine from violating constraints without compromising engine response is becoming harder requiring systematic treatment. The constraints include the compressor surge limit, actuator limits on throttle and wastegate, turbine speed and temperature limits, intake pressure overshoot limit, combustion limits, etc.

The surge constraint handling in turbocharged gasoline engines (see the schematic in Fig. 3) using reference governor techniques is addressed in [59], [60].

![Fig. 3. Schematics of a turbocharged gasoline engine.](image)

The application of the reduced order reference governor in [59] is motivated by different time scales of the engine model variables. The model is order five with the following state variables: intake manifold pressure (kPa), boost pressure (kPa), exhaust manifold pressure (kPa), turbocharger speed (rpm), and wastegate flow (g/sec). The eigenvalues of the linearized continuous-time model are,

\[ \{-2.39, -3.16, -24.3, -161, -259\}, \]

suggesting that the dynamics can be decomposed into a second or a third order slow subsystem and, respectively, a third or a second order fast subsystem. The model has 2 outputs (\( y(t) \)): boost pressure (kPa), and compressor flow (g/sec) that to avoid surge are constrained by an affine inequality as \( y(t) \in Y \), where,

\[ Y = \{y : SY \leq s\}. \] (53)

The reference governor is applied to modify the electronic throttle (ETC) command and wastegate command. Specifically, the reduced order reference governor has been de-
signed based on the second order reduced order model without performance loss. The validation results based on the nonlinear model and the observer for unmeasured states are presented in [61]. See Figures 4-5. The reduced order reference governor design was based on the linear model and utilized the procedure for mismatch compensation with the nonlinear model described in in Section III-A.

Fig. 4. The response of a nonlinear system plotted on a compressor map with reduced order (solid) and full order (dashed) reference governors applied.

Fig. 5. The throttle response of a nonlinear system with reduced order (solid) and full order (dashed) reference governors applied.

B. Vehicle dynamics and rollover protection

The applications of the reference and extended command governors to vehicle roll control are considered in [84]. The schemes modify steering angle and operation of the brakes so that vehicle constraints are satisfied. The dynamics of vehicle yaw and roll motion are captured by a discrete-time linear model with four states: lateral velocity of Center of Gravity of the vehicle, yaw rate of the unsprung mass, roll rate of the sprung mass and roll angle of the sprung mass. The constraints on the load transfer ratio, which is the difference between the load on right tires minus the load on left tires divided by total weight of the vehicle, between −1 and 1, are enforced. As shown in [84], both scalar reference governor (SRG) and extended command governor (ECG) prevent rollover, while modifications of the vehicle trajectory are slight. The on-line computing effort associated with the SRG (explicitly solvable scalar optimization problem) is shown to be less than of the extended command governor (simple and explicitly solvable parametric quadratic program), while the domain of recoverable states of the extended command governor can be much larger. Further as Figures 6- 7 illustrate, ECG is robust to a mismatch between the model and actual system dynamics.

C. Electromagnetic actuators

Reference governors were applied to enforce a variety of constraints in electromagnetic actuators. See [67], [59], [95], [76], [77]. In particular, the constraint induced by the limited coil current leads to a constraint expressed by a concave nonlinear function, while the soft landing constraint is of MLD type. These types of constraints can be handled effectively using techniques in [59].

Fig. 6. The time histories of load transfer ratio $y(t)$ with ECG designed based on $v = 22.5$ m/sec model and $\hat{O}_\infty(22.5)$ as terminal set (solid) and without ECG (dashed) for the vehicle maneuver at 30 m/sec. The constraints are shown by the dash-dotted lines.

Fig. 7. The time histories of commanded steering angle $\delta_r(t)$ (dashed) and of $\delta(t)$ (solid) for the vehicle maneuver at 30 m/sec, where $\delta(t)$ is prescribed by the ECG designed based on $v = 22.5$ m/sec model and $\hat{O}_\infty(22.5)$.
D. Fuel cells

The publications [117], [108], [90] consider handling constraints in fuel cell applications using reference governor techniques. In these systems constraints are imposed to maintain the oxygen over hydrogen ratio sufficiently high, thereby preventing oxygen starvation, to avoid compressor surge and choke regions, and to avoid compressor voltage saturation.

The developments in [117] are based on the linearized model of order 10, and exploit the procedure in Section III-A for compensating the mismatch between the linearized model and the actual nonlinear system. The reference governor is applied to control fuel cell load (current) requests. Since in [117] (and also in [108], [90]) the reference governor controls the load, it is referred to as the load governor. The design exploits the use of non-positively invariant $P \subset \hat{O}_\infty$ in (4) that are obtained from $\hat{O}_\infty$ by eliminating almost redundant inequalities and the pull-in procedure. The number of inequalities in the representation of $\hat{O}_\infty \subset \mathbb{R}^{12}$ was 325 and reduced to 62 in the representation for $P$. This led to the reduction of the number of flops from 10100 to 1650. The implementation in the production micro-controller has shown that the reference governor computations require 1.3 msec at 10 msec update rate and 4kb of ROM.

The nonlinear reference governor is applied to fuel cell constraint handling in [108]. Parameter uncertainties in temperature and humidity are handled using the robust reference governor approach of Section III-C. Robust constraint enforcement capability with minimum impact on system response time has been demonstrated.

E. Other automotive applications

Reference [118] has addressed belt restraint systems. References [85], [2] have considered the applications of reference and command governors to engine speed control. Reference governors were applied to handling constraints in electric batteries in [97] and [106]. Applications to constraint handling in HCCI engines were reported in [58] and in free piston engines in [120]. Finally, applications to vehicle dynamic control were considered in [15].

F. Non-automotive applications

This tutorial paper would be incomplete without mentioning several non-automotive applications of reference and command governors. Applications to aerospace systems are treated in [99], [29], [40], [78], [101], [68], [121], [100], [37]. Applications to electric power systems are considered in [30], [27], [54], [28], [113], [104]. Other applications include chemical processes [73], cable robots [98], disk drives [53], rotary cranes [57], open water channel networks [91], gas turbine engines [78], [64]; inverted pendulum [24], cooperative vehicle control [114]; four tank laboratory system [31]; electrostatically actuated membrane mirrors [75]; and tokamak reactors used in thermonuclear fusion [119], [93]. The growing breadth of these applications suggests widening interest in the reference governors for engineering applications.

VI. CONNECTIONS WITH OTHER DESIGN TECHNIQUES

In this Section we comment on connections between reference/command governors and related control techniques. These comments will be further extended in the final draft of the paper.

A. Connections with Model Predictive Control

As predictive control schemes for constrained reference tracking, governors have many common features with Model Predictive Control (MPC) and, in fact, can be designed within MPC framework [103]. At the same time, they are special schemes with unique motivation and several unique properties, results, and simplifications (such as finite-time convergence for constant reference commands or the design based on reduced constraint set) that are not easily available to more general MPC controllers. Furthermore, reference handling in MPC [41], tube MPC [3], [94] and reduced complexity MPC approaches in [87] (arguably) incorporate features similar to reference/command governors. Parameter governors proposed in [80] have similarities with parameterized nonlinear MPC of [1]. Other techniques for reference tracking in constrained systems include [14].

B. Connections with input shaping

The input shaping techniques have been proposed to minimize residual vibrations in flexible structures, see e.g. [105]. Similar to reference/command governors input shapers modify the input to the system, however, they typically are not designed to enforce state and control constraints. The feedforward and reduced order reference governors may be suitable for problems where input shaping has traditionally been used and there are constraints; however, their properties in such applications remain to be further studied.

VII. RESEARCH TOPICS

While the subject of reference governors has been researched for over twenty years and connects naturally with Modern Predictive Control, a variety of new research directions can be identified.

The nonlinear reference governor results (see e.g., [48], [108] and references therein) extend to the case when the system has uncertain set-bounded constant parameters. At the same time, a non-conservative application of reference governors to systems with uncertain parameters being estimated online remains an area to be further explored. Special assumptions appear to be necessary in this case to guarantee recursive feasibility and other reference governor properties.

The treatment of the case when constraints are time-varying or reconfigured dynamically is of interest for many applications. While we have had success in treating specific examples, see e.g., [85], the theory remains largely to be developed.

Traditionally, reference and extended command governors modify the set-points to closed-loop systems. While the theory assumes that the set-points are given, in real systems the set-points may be adjusted by a human operator in
response to external conditions. This dependence of set-points on external conditions can create a feedback loop encompassing the reference governor and nominal closed-loop system. A closely related situation occurs when the governor augments control signals at the actuator command level. This implementation of reference governors is appealing as the design and calibration of the nominal controller can be changed without the need to re-design the governor. The properties of the reference governor in the loop remain to be studied.

Error governors are schemes related to reference governors, however, they act on the tracking error at the controller input and not on the reference command. They are also primarily intended for handling control input constraints and not output constraints. See [45], [66], [115]. As compared to reference governors, error governors have received relatively little attention; obtaining convergence guarantees for them that are similar to reference/command governor has been elusive. It is interesting that the error governor can be applied with relative ease to direct adaptive controllers [69].

ACKNOWLEDGEMENTS

The authors acknowledge their colleagues, coauthors on cited joint publications, for contributions to the material included in this paper.

REFERENCES


