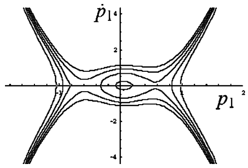

Fig. 1. Phase portrait of the extremal curve p_1.

are constant, which in turn implies that helices are optimal paths for weights of the cost function $c_2 = c_3 = 1$. In addition we observe that the solution (29) at $\dot{p}_1 = 0$ is also a solution of the dynamic equations of motion (12) when $c_2 = c_3$. Therefore, these particular optimal motions are also solutions of the dynamic equations of motion of a laterally constrained axisymmetric AUV. This particular set of optimal motions make available useful and simplistic reference paths for a real AUV to track. Moreover, the particular optimal helical motions could provide tracking trajectories for ascending and descending conventional, slender AUVs.

VI. Conclusion

This note has investigated the motion control of an AUV. The AUV is treated as a nonholonomic system as any lateral motion of a conventional, slender AUV is quickly damped out. The problem is formulated as a fixed end point optimal control problem on the Euclidean Group of Motions $SE(3)$, where the cost function to be minimized is equal to the integral of a quadratic function of the velocity components. An application of the Maximum Principle to this optimal control problem yields the appropriate Hamiltonian and, along with the Poisson bracket, we derive the Hamiltonian vector fields, which define the necessary conditions for optimality. For the special case of the cost function (two of the weights are equal) the necessary condition for optimality can be characterized more easily and we proceed to investigate these solutions. Finally, it is shown that a particular set of optimal motions trace helical paths. Throughout the note, we highlight a particular case where the quadratic cost function is weighted in such a way that it equates to the Lagrangian (Kinetic Energy) of the AUV. Further constraining the extremal curves to equate to the components of momentum defines an almost-Poisson bracket and the resulting vector fields are equivalent to the d'Alembert-Lagrange equations in Hamiltonian form.

References

[1] C. Woolsey and N. Leonard, "Stabilizing underwater vehicle motion using internal rotors," *Automatica*, vol. 38, no. 12, pp. 2053–2062.
[2] N. E. Leonard and J. G. Graver, "Model-based feedback control of autonomous underwater vehicles," *IEEE J. Ocean. Eng.*, vol. 26, no. 4, pp. 633–645, Oct. 2001.
[3] C. Schultz and C. Woolsey, "An experimental platform for validating internal actuator control strategies," in *Proc. IFAC workshop on Guidance and Control of Underwater Vehicles*, 2003, pp. 209–214.
[4] N. Leonard and P. S. Krishnaprasad, "Motion control of drift free, left-invariant systems on Lie groups," *IEEE Trans. Automat. Control*, vol. 40, no. 9, pp. 1539–1554, Sep. 1995.
[5] P. Holmes, J. Jenkins, and N. E. Leonard, "Dynamics of the kirchhoff equations: Coincident centers of gravity and buoyancy," *Phys. D*, vol. 118, pp. 311–342, 1998.
[6] V. Jurdjevic, "Geometric control theory," in *Advanced Studies in Mathematics*. Cambridge, U.K.: Cambridge University Press, 1997, vol. 52.
[7] H. J. Sussmann, "An introduction to the coordinate-free maximum principle," in *In Geometry of Feedback and Optimal Control*, B. Jakubczyk and W. Respondek, Eds. New York: Marcel Dekker, 1997, pp. 463–557.
[8] P. Crouch and S. F. Leite, "The dynamic interpolation problem: On riemannian manifolds, lie groups and symmetric spaces," *J. Dynam. Control Syst.*, vol. 1, no. 2, pp. 177–202, 1995.
[9] A. M. Bloch, *Nonholonomic Mechanics and Control*. New York: Springer-Verlag, 2003.
[10] A. Bloch, P. Crouch, and T. Ratui, "Sub-Riemannian optimal control problems," *Fields Inst. Commun.*, vol. 3, pp. 35–48, 1994.
[11] I. I. Hussein and A. M. Bloch, "Constrained optimal trajectory tracking on the group of rigid body motions," in *Proc. IEEE CDC-ECC*, Seville, Spain, 2005, pp. 2152–2157.
[12] A. J. van der Schaft and B. M. Maschke, "On the Hamiltonian formulation of nonholonomic mechanical systems," *Reports on Mathemat. Phys.*, vol. 34, 1994.
[13] J. D. Biggs, W. Holderbaum, and V. Jurdjevic, "Singularities of optimal control problems on some 6-D lie groups," *IEEE Trans. Autom. Control*, vol. 53, no. 6, pp. 1027–1038, Jun. 2007.
[14] D. Lawden, *Elliptic Functions and Applications*. New York: Springer-Verlag, 1989.

Probabilistic Constrained MPC for Multiplicative and Additive Stochastic Uncertainty

Mark Cannon, Basil Kouvaritakis, and Xingjian Wu

***Abstract*—The technical note develops a receding horizon control strategy to guarantee closed-loop convergence and feasibility in respect of soft constraints. Earlier results [1] addressed the case of multiplicative uncertainty only. The present technical note extends these to the more general case of additive and multiplicative uncertainty and proposes a method of handling probabilistic constraints. The results are illustrated by a simple design study considering the control of a wind turbine.**

***Index Terms*—Constrained control, fatigue life, optimization, stochastic control.**

I. Introduction

Model predictive control (MPC) optimizes predicted behavior in the presence of constraints, thus providing a practicable solution to a closed-loop infinite dimensional optimization problem [2]. However, most applications, besides being constrained, are also subject to uncertainty. This could be handled by robust MPC techniques [3], [4], which assume an uncertainty description in terms of bounds on unknown parameters. Less conservative approaches using available information on the distribution of uncertainty have been proposed (e.g., for LQR and robust control [5], [6], and for constrained MPC [7]). The difficulty in the constrained case is that a non-conservative, efficient means for propagating the effects of uncertainty over a prediction horizon has yet to emerge.

A computationally convenient approach which also reduces the degree of conservativeness was proposed in [1], based on an autonomous description of the prediction dynamics [8]. Use was made of the concept of probabilistic invariance in order to handle soft constraints in the case of multiplicative uncertainty. The current technical note extends the approach to the more challenging problem of systems subject

Manuscript received September 26, 2008; revised February 10, 2009. First published June 10, 2009; current version published July 09, 2009. Recommended by Associate Editor M. Fujita

The authors are with the Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, U.K. (e-mail: mark.cannon@eng.ox.ac.uk).

Digital Object Identifier 10.1109/TAC.2009.2017970

to both additive and multiplicative uncertainty. The presence of additive uncertainty implies that it is not possible to design control laws on the basis of mean square stability (MSS), under which the convergence of state predictions would be ensured. This technical note determines the nonzero asymptotic limit of predicted behavior and redefines the control objective based on deviations away from this limit. It then uses probabilistic invariance to construct an algorithm with desirable closed-loop properties: convergence of the state variance and feasibility of particular types of soft constraints.

Motivated by problems involving fatigue constraints, the soft constraints considered here take the form of limits on the expected number of samples at which a system output exceeds specified bounds over a given horizon. A simplified model is used to demonstrate the effectiveness of probabilistic invariance in converting these constraints into probabilistic state constraints. The results of the technical note are illustrated by a design example based on a simulated wind turbine control problem of maximizing power capture while respecting high cycle fatigue damage constraints.

II. Problem Formulation

Consider the uncertain linear system described by

$$x_{k+1} = A_k x_k + B_k u_k + d_k, \quad x \in \mathbb{R}^{n_x}, \ u \in \mathbb{R}^{n_u}. \tag{1}$$

Let $\theta_k \in \mathbb{R}^m$ be the vector of uncertain elements in A_k, B_k, d_k such that $\{\theta_k, \ k = 0, 1, \ldots\}$ is an independent and identically distributed (i.i.d.) sequence with mean θ and covariance Σ_θ. Then $\theta_k = \bar{\theta} + U\Lambda^{1/2} q_k$, $q_k \sim \mathcal{N}(0, I)$, where U, Λ are the eigen-matrices of Σ_θ, so that

$$[\,A_k \quad B_k \quad d_k\,] = [\,\bar{A} \quad \bar{B} \quad 0\,] + \sum_{j=1}^{m} [\,\tilde{A}_j \quad \tilde{B}_j \quad \tilde{g}_j\,]\, q_{k,j} \tag{2}$$

with $q_k = [\,q_{k,1} \quad \ldots \quad q_{k,m}\,]^T$. Define also an output ψ_k

$$\psi_k = C_k x_k + D_k u_k + \eta_k, \quad \psi_k \in \mathbb{R}^{n_\psi}$$
$$[\,C_k \quad D_k \quad \eta_k\,] = [\,\bar{C} \quad \bar{D} \quad 0\,] + \sum_{j=1}^{m} [\,\tilde{C}_j \quad \tilde{D}_j \quad \tilde{\eta}_j\,]\, q_{k,j} \tag{3}$$

which is subject to soft constraints, namely ψ_k may lie outside a desired interval $I_\psi = [\psi^L, \psi^U]$ but the expected number of samples over a horizon N_c at which this occurs cannot exceed a bound $N_{\max}$

$$\frac{1}{N_c} \sum_{i=0}^{N_c - 1} \Pr\{\psi_{k+i} \notin I_\psi\} \le \frac{N_{\max}}{N_c}. \tag{4}$$

This can be translated into probabilistic constraints on the model state (as discussed in Section IV), and hence into constraints invoked in the online MPC optimization (discussed in Section VI).

Remark 1: The particular distribution of q in (2) and (3) is unimportant for the results of the technical note. However, to derive a guarantee of satisfaction of the probabilistic constraint (4), it is assumed that q is finitely supported (either innately or as a result of truncating an infinitely supported distribution at a sufficiently high probability level). It is furthermore assumed that q is i.i.d. and that it is possible to construct polytopic sets $\mathcal{Q} = \mathrm{Co}\{q^{(i)}, \ i = 1, \ldots, \nu\}$ such that $\Pr\{q \in \mathcal{Q}\} \ge p$. A method of computing such a set $\mathcal{Q}$ is described the example of Section VII. Knowledge of $\mathcal{Q}$ will be assumed in the sequel.

Remark 2: Though relevant for some applications (e.g., the example of Section VII), the assumption that q is i.i.d. is restrictive since it limits the applicability of the results of this technical note to problems in which disturbances are rapidly varying. For the case of additive disturbances alone, temporal correlation could be accounted for by incorporating filtered white noise in the model, however the lifting of this restriction for the general case of multiplicative uncertainty is an area for further work. Note also that the case of additive uncertainty with non-zero unknown mean can be handled through the introduction of integral action.

We define predicted control sequences using a dual mode prediction paradigm [4], according to which the control inputs over the first N steps are free, and a prescribed state feedback law is assumed over the subsequent infinite prediction horizon. The control sequence predicted at time k, $\{u_{k+i}, \ i = 0, 1, \ldots\}$, is formulated (e.g., [8]) as

$$u_{k+i} = \begin{cases} K x_{k+i} + c_{i|k} & i = 0, \ldots, N-1 \\ K x_{k+i} & i = N, N+1, \ldots. \end{cases} \tag{5}$$

The feedback and feedforward structure of (5) suggests an autonomous description of the prediction dynamics [8] with predictions at time k generated by

$$z_{i+1|k} = \Psi_{k+i} z_{i|k} + \delta_{k+i}, \quad i = 0, 1, \ldots \tag{6a}$$
$$z_{0|k}^T = [\,x_k^T \quad f_k^T\,], \ f_k^T = [\,c_{0|k}^T \quad \cdots \quad c_{N-1|k}^T\,] \tag{6b}$$
$$[\,\Psi_k \quad \delta_k\,] = [\,\bar{\Psi} \quad 0\,] + \sum_{j=1}^{m} [\,\tilde{\Psi}_j \quad \tilde{\gamma}_j\,]\, q_{k,j} \tag{6c}$$
$$\bar{\Psi} = \begin{bmatrix} \bar{\Phi} & \bar{B}E \\ 0 & M \end{bmatrix}, \quad \tilde{\Psi}_j = \begin{bmatrix} \tilde{\Phi}_j & \tilde{B}_j E \\ 0 & 0 \end{bmatrix}$$
$$\tilde{\gamma}_j = \begin{bmatrix} \tilde{g}_j \\ 0 \end{bmatrix} \tag{6d}$$

where the plant state x_k is assumed to be known at time k, $\bar{\Phi} = \bar{A} + \bar{B}K$, $\tilde{\Phi}_j = \tilde{A}_j + \tilde{B}_j K$, and

$$M = \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad E = [\,I \quad 0 \quad \cdots \quad 0\,]$$

with I denoting the identity matrix in $\mathbb{R}^{n_u \times n_u}$. In this formulation, the plant state x_{k+i} predicted at time k is then given by $x_{k+i} = \Gamma^T z_{i|k}$, where $\Gamma = [\,I \quad 0\,]^T$.

The advantage of (6) is that constraints on predictions at time k can be invoked conveniently through constraints on the initial state $z_{0|k}$. Since (6) is autonomous, the implied constraints can be computed using one-step-ahead invariance considerations. This affords significant computational advantages in the deterministic case, but in the stochastic case it is of vital importance: it removes the need to propagate the effects of uncertainty over the prediction horizon. Such propagation would require the numerical computation of the distributions of future states (see [7] and [9]), implying a prohibitive computational load.

III. Receding Horizon Performance Objective

Let $\mathbb{E}_k(X)$ denote the expected value of a random variable X, conditional on information available at time k and, since in this section predictions will be made at a single instant k, let $z_i = z_{i|k}$. Under an MSS assumption on (6), it can be shown [10] that $\mathbb{E}_k\left(z_i z_i^T\right)$ converges to a finite limit as $i \to \infty$ along the predicted trajectories of (6), provided that Ψ_{k+i} and δ_{k+i} are independent. An extension to the case of dependent Ψ_{k+i} and δ_{k+i} is given below.

Lemma 1: The sequence $\{z_i, \ i = 0, 1, \ldots\}$ generated by (6) satisfies $\lim_{i\to\infty} \mathbb{E}_k(z_i) = 0$ and $\lim_{i\to\infty} \mathbb{E}_k\left(z_i z_i^T\right) = \Theta$, where Θ is the solution of the Lyapunov equation

$$\Theta - \bar{\Psi}\Theta\bar{\Psi}^T - \sum_{j=1}^{m} \tilde{\Psi}_j \Theta \tilde{\Psi}_j^T = \sum_{j=1}^{m} \tilde{\gamma}_j \tilde{\gamma}_j^T \tag{7}$$

if and only if there exists $P \succ 0$ satisfying

$$P - \bar{\Psi}^T P \bar{\Psi} - \sum_{j=1}^{m} \tilde{\Psi}_j^T P \tilde{\Psi}_j \succ 0. \tag{8}$$

Proof: By the linearity of (6), the sequence $\{z_i\}$ is the sum of $\{\zeta_i\}$ and $\{\xi_i\}$ generated by

$$\zeta_{i+1} = \Psi_{k+i}\zeta_i, \quad \zeta_0 = z_0 \tag{9a}$$
$$\xi_{i+1} = \Psi_{k+i}\xi_i + \delta_{k+i}, \quad \xi_0 = 0. \tag{9b}$$

The necessity and sufficiency of (8) for MSS of (9a) (see e.g., [11]) ensures that $\mathbb{E}_k\left(\zeta_i\zeta_i^T\right) \to 0$, and hence $\zeta_i \to 0$ almost surely, as $i \to \infty$. From (9b), we have $\mathbb{E}_k(\xi_i) = 0$ for all i, so $\mathbb{E}_k(z_i) \to 0$ as $i \to \infty$. Since ξ_i and Ψ_{k+i} are independent, (9b) also gives

$$\begin{aligned}\mathbb{E}_k(\xi_{i+1}\xi_{i+1}^T) &= \mathbb{E}_k\left\{(\Psi_{k+i}\xi_i + \delta_{k+i})(\Psi_{k+i}\xi_i + \delta_{k+i})^T\right\} \\ &= \mathbb{E}_k\left(\Psi_{k+i}\xi_i\xi_i^T\Psi_{k+i}^T\right) + \mathbb{E}_k\left(\delta_{k+i}\delta_{k+i}^T\right) \\ &= \bar{\Psi}\mathbb{E}_k(\xi_i\xi_i^T)\bar{\Psi}^T + \sum_{j=1}^{m}\tilde{\Psi}_j\mathbb{E}_k(\xi_i\xi_i^T)\tilde{\Psi}_j^T + \sum_{j=1}^{m}\tilde{\gamma}_j\tilde{\gamma}_j^T.\end{aligned} \tag{10}$$

Let $\hat{\Theta}_i = \mathbb{E}_k\left(\xi_i\xi_i^T\right) - \Theta$, then from (7) and (10) we have $\hat{\Theta}_{i+1} = \bar{\Psi}\hat{\Theta}_i\bar{\Psi}^T + \sum_{j=1}^{m}\tilde{\Psi}_j\hat{\Theta}_i\tilde{\Psi}_j^T$. From the MSS condition (8), it follows that $\hat{\Theta}_i \to 0$, so that $\mathbb{E}_k\left(\xi_i\xi_i^T\right) \to \Theta$ as $i \to \infty$. Finally note that $\mathbb{E}_k\left(z_iz_i^T\right) \to \mathbb{E}_k\left(\xi_i\xi_i^T\right)$ since $\zeta_i \to 0$ almost surely as $i \to \infty$. ■

For the case of no additive uncertainty, i.e., if $\tilde{\gamma}_j = 0$ for $j = 1,\ldots,m$ in (6), the existence of P satisfying (8) implies $\lim_{i\to\infty}\mathbb{E}\left(z_iz_i^T\right) = 0$. In this case therefore, the predicted cost

$$\begin{aligned}J_k &= \sum_{i=0}^{\infty}\mathbb{E}_k(x_{k+i}^TQx_{k+i} + u_{k+i}^TRu_{k+i}) \\ &= \sum_{i=0}^{\infty}\mathbb{E}_k(z_i^T\tilde{Q}z_i), \\ \tilde{Q} &= \begin{bmatrix} Q + K^TRK & K^TRE \\ E^TRK & E^TRE\end{bmatrix}\end{aligned} \tag{11}$$

is well-defined, and hence is suitable for online minimization by a receding horizon control law. However, for the case of persistent nonzero additive uncertainty, it follows from Lemma 1 that the stage cost of (11) converges to a nonzero limit along trajectories of (6)

$$\lim_{i\to\infty}\mathbb{E}_k\left(z_i^T\tilde{Q}z_i\right) = \mathrm{tr}(\Theta\tilde{Q})$$

so that the cost (11) is infinite. To obtain a finite cost, (11) must be modified as

$$J_k = \sum_{i=0}^{\infty}\mathbb{E}_k(L_i), \quad L_i = z_i^T\tilde{Q}z_i - \mathrm{tr}(\Theta\tilde{Q}). \tag{12}$$

Theorem 2: The cost (12), evaluated along trajectories of (6), is given by

$$J_k = \begin{bmatrix} z_0 \\ 1\end{bmatrix}^T \tilde{P}\begin{bmatrix} z_0 \\ 1\end{bmatrix}, \quad \tilde{P} = \begin{bmatrix} P_z & P_{z1} \\ P_{1z} & P_1\end{bmatrix}$$

where P_z, $P_{1z} = P_{z1}^T$, and P_1 are defined uniquely by

$$P_z - \bar{\Psi}^TP_z\bar{\Psi} - \sum_{j=1}^{m}\tilde{\Psi}_j^TP_z\tilde{\Psi}_j = \tilde{Q} \tag{13a}$$

$$P_{1z} = \sum_{j=1}^{m}\tilde{\gamma}_j^TP_z\tilde{\Psi}_j(I - \bar{\Psi})^{-1} \tag{13b}$$
$$P_1 = -\mathrm{tr}(\Theta P_z). \tag{13c}$$

Proof: Define $V_i = z_i^TP_zz_i + z_i^TP_{z1} + P_{1z}z_i + P_1$, then from $z_{i+1} = \Psi_{k+i}z_i + \delta_{k+i}$, we have

$$\begin{aligned}&\mathbb{E}_k(V_i) - \mathbb{E}_k(V_{i+1}) \\ &\quad = \mathbb{E}_k\left(z_i^T\left[P_z - \mathbb{E}(\Psi_{k+i}^TP_z\Psi_{k+i})\right]z_i\right) \\ &\qquad + 2\left[P_{1z}(I - \bar{\Psi}) - \mathbb{E}(\delta_{k+i}^TP_z\Psi_{k+i})\right]\mathbb{E}_k(z_i) \\ &\qquad - \mathbb{E}(\delta_{k+i}^TP_z\delta_{k+i}).\end{aligned} \tag{14}$$

But (13a) gives the first term of the RHS as

$$\mathbb{E}_k\left(z_i^T\left[P_z - \mathbb{E}(\Psi_{k+i}^TP_z\Psi_{k+i})\right]z_i\right) = \mathbb{E}_k(z_i^T\tilde{Q}z_i). \tag{15}$$

Post-multiplying (7) by P_z and extracting the trace gives

$$\begin{aligned}&\mathrm{tr}\left(\Theta P_z - \bar{\Psi}\Theta\bar{\Psi}^TP_z - \sum_{j=1}^{m}\tilde{\Psi}_j\Theta\tilde{\Psi}_j^TP_z\right) \\ &\quad = \mathrm{tr}\left(\Theta[P_z - \bar{\Psi}^TP_z\bar{\Psi} - \sum_{j=1}^{m}\tilde{\Psi}_j^TP_z\tilde{\Psi}_j]\right) \\ &\quad = \sum_{j=1}^{m}\mathrm{tr}\left(\tilde{\gamma}_j\tilde{\gamma}_j^TP_z\right)\end{aligned}$$

and (13a) therefore implies

$$\mathrm{tr}(\Theta\tilde{Q}) = \sum_{j=1}^{m}\tilde{\gamma}_j^TP_z\tilde{\gamma}_j = \mathbb{E}\left(\delta_{k+i}^TP_z\delta_{k+i}\right). \tag{16}$$

From (14)–(16), and (13b) we obtain $\mathbb{E}_k(V_i) - \mathbb{E}_k(V_{i+1}) = \mathbb{E}_k(z_i^T\tilde{Q}z_i) - \mathrm{tr}(\Theta\tilde{Q})$. Summing this recursion over all $i \geq 0$ gives $V_0 - \lim_{i\to\infty}\mathbb{E}_k(V_i) = \sum_{i=0}^{\infty}\mathbb{E}_k(L_i) = J_k$. Finally we show that $\mathbb{E}_k(V_i) \to 0$ as $i \to \infty$; this follows from the definition of V_i and (13c), which imply

$$\mathbb{E}_k(V_i) = \mathbb{E}_k(z_i^TP_zz_i) + 2P_{1z}\mathbb{E}_k(z_i) - \mathrm{tr}(\Theta P_z)$$

and therefore $\lim_{i\to\infty}\mathbb{E}_k(V_i) = \lim_{i\to\infty}\mathbb{E}_k(z_i^TP_zz_i) - \mathrm{tr}(\Theta P_z) = \lim_{i\to\infty}\mathrm{tr}\left[\mathbb{E}_k(z_iz_i^T)P_z\right] - \mathrm{tr}(\Theta P_z) = 0$, where $\lim_{i\to\infty}\mathbb{E}_k(z_i) = 0$ and $\lim_{i\to\infty}\mathbb{E}_k\left(z_iz_i^T\right) = \Theta$ (from Lemma 1) have been used. ■

IV. A Framework for Handling Soft Constraints

This section describes an analysis technique that enables the soft constraints (4) to be converted into probabilistic state constraints. Use is made of a discrete Markov chain model, according to which x_k can lie in either $S_1 = \mathcal{E}_1$ or $S_2 = \mathcal{E}_2 - \mathcal{E}_1$ where $\mathcal{E}_1 \subset \mathcal{E}_2 \subset \mathbb{R}^{n_x}$. This scenario relies on the assumption (see Remark 1) that the uncertainty in (2) has finite support, though in practice we simply require that $\mathcal{E}_2$ is defined so that the probability of $x_k \notin \mathcal{E}_2$ is negligible. Furthermore, the analysis could be made less conservative (and more realistic) by considering a nested sequence $\mathcal{E}_1 \subset \cdots \subset \mathcal{E}_r$, but $r = 2$ is used here to simplify presentation.

Define the conditional probabilities

$$\Pr(\psi_k \notin I_\psi \mid x_k \in S_j) = p_j, \quad j = 1, 2. \tag{17}$$

For p_1 small, so that S_1 is the *safe* region of state space, it is convenient to assume that S_2 is *unsafe* and set p_2 at some high value (close to, or equal to 1). For a fixed state feedback law and for i.i.d. uncertainty distribution, it is possible to define the time-invariant matrix of transition probabilities Π with elements p_{ij}, where p_{ij} is the probability that the online algorithm steers the state in one step from S_j to S_i. Then over i steps, we have

$$\begin{bmatrix}\Pr(x_{k+i}\in S_1)\\ \Pr(x_{k+i}\in S_2)\end{bmatrix}=\Pi^i\begin{bmatrix}\Pr(x_k\in S_1)\\ \Pr(x_k\in S_2)\end{bmatrix} \tag{18}$$

so that the probability of a constraint violation at time $k+i$ is given by

$$\Pr(\psi_{k+i}\notin I_\psi)=[\,p_1\quad p_2\,]\,\Pi^i\begin{bmatrix}\Pr(x_k\in S_1)\\ \Pr(x_k\in S_2)\end{bmatrix}.$$

Because of the special structure of Π (see e.g., [10]) its eigenvector decomposition is given by

$$\Pi=[\,w_1\quad w_2\,]\begin{bmatrix}1&0\\0&\lambda_2\end{bmatrix}\begin{bmatrix}v_1^T\\ v_2^T\end{bmatrix},\quad 0\le\lambda_2<1.$$

Therefore, as $i\to\infty$, the rate at which constraint violations accumulate given $x_k\in S_j$ tends to

$$R_j=[\,p_1\quad p_2\,]\,w_1v_1^Te_j,\quad j=1,2 \tag{19}$$

where e_j denotes the jth column of the 2×2 identity matrix. If R_1 and R_2 are less than the limit $N_{\max}/N_c$ of (4), then there will exist finite i^* such that, for all $i\ge i^*$, the total expected number of constraint violations will be less than $iN_{\max}/N_c$; i^* is defined as the smallest i for which the cumulative number of constraint violations is less than or equal to $iN_{\max}/N_c$. Provided i^* does not exceed the horizon N_c, it then follows that the probabilistic formulation (17), (18) ensures that the soft constraints (4) on ψ are satisfied.

V. Probabilistic Invariance and Probabilistic Constraints

This section proposes a procedure for the design of the regions $\mathcal{E}_i$ of Section IV by making use of the concept of probabilistic invariance.

Definition 1 ([1]): A set $S\subset\mathbb{R}^{n_x}$ is invariant with probability p (i.w.p. p) under a given control law if $x_{k+1}\in S$ with probability p whenever $x_k\in S$.

The approach is based on ellipsoidal sets, $\mathcal{E}\subset\mathbb{R}^{n_x+Nn_u}$ and $\mathcal{E}_x\subset\mathbb{R}^{n_x}$, defined as

$$\mathcal{E}=\{z:z^T\hat{P}z\le 1\},\ \mathcal{E}_x=\{x:x^T\hat{P}_xx\le 1\},\ \hat{P}_x=(\Gamma^T\hat{P}^{-1}\Gamma)^{-1}$$

so that $\mathcal{E}_x$ is the projection of $\mathcal{E}$ onto the x-subspace. Let $\mathcal{Q}$ denote a confidence region that contains the vector of uncertain parameters in (2) with a specified probability p:

$$\Pr\{q_k\in\mathcal{Q}\}\ge p. \tag{20}$$

Lemma 3: $\mathcal{E}_x$ is i.w.p. p under (5) for any f_k such that $z_{0|k}\in\mathcal{E}$ if scalar $\lambda\in[0,1]$ exists with

$$\begin{bmatrix}\hat{P}_x^{-1} & \Gamma^T\Psi(q^{(i)})\hat{P}^{-1} & \Gamma^T\gamma(q^{(i)})\\ \hat{P}^{-1}\Psi(q^{(i)})^T\Gamma & \lambda\hat{P}^{-1} & 0\\ \gamma(q^{(i)})^T\Gamma & 0 & 1-\lambda\end{bmatrix}\succeq 0 \tag{21}$$

for $i=1,\dots,\nu$, where $\Psi(q_k)=\bar{\Psi}+\sum_{j=1}^m\tilde{\Psi}_jq_{k,j}$ and $\gamma(q_k)=\sum_{j=1}^m\tilde{\gamma}_jq_{k,j}$.

Proof: Equation (20) implies that $\Pr\left(x_{k+1}^T\hat{P}_xx_{k+1}\le 1\right)\ge p$ if

$$x_{k+1}^T\hat{P}_xx_{k+1}\le 1\quad\forall\, z_{0|k}\in\mathcal{E},\ \forall\, q_k\in\mathcal{Q} \tag{22}$$

where x_{k+1} is given by $x_{k+1}=\Gamma^T\Psi(q_k)z_{0|k}+\Gamma^T\gamma(q_k)$. By the $\mathcal{S}$-procedure, (22) is equivalent to the existence of $\lambda\ge 0$ satisfying

$$1-(\Psi(q)z+\gamma(q))^T\Gamma\hat{P}_x\Gamma^T(\Psi(q)z+\gamma(q))\ge\lambda(1-z^T\hat{P}z)$$

for all z and all $q\in\mathcal{Q}$, or equivalently

$$\begin{bmatrix}\lambda\hat{P}&0\\0&1-\lambda\end{bmatrix}-\begin{bmatrix}\Psi(q)^T\\ \gamma(q)^T\end{bmatrix}\Gamma\hat{P}_x\Gamma^T\,[\,\Psi(q)\quad\gamma(q)\,]\succeq 0,\quad \forall\, q\in\mathcal{Q}.$$

This can be expressed using Schur complements as a LMI in q which, when invoked for all $q\in\mathcal{Q}$, is equivalent to (21) for some $\lambda\in[0,1]$. ■

Additional constraints on $\hat{P}$ are needed so as to constrain the conditional probability that ψ_k lies outside the desired interval I_ψ given that $z_{0|k}$ lies in $\mathcal{E}$. Rewriting (3) in the form

$$\psi_k=\hat{C}(q_k)z_{0|k}+\eta(q_k)$$
$$\hat{C}(q_k)=\bar{C}\Gamma^T+\bar{D}\hat{K}+\sum_{j=1}^m(\tilde{C}_j\Gamma^T+\tilde{D}_k\hat{K})q_{k,j}$$

with $\hat{K}=[\,K\quad E\,]$, the following result is based on the confidence region $\mathcal{Q}$.

Lemma 4: $\Pr(\psi_k\notin I_\psi\mid z_{0|k}\in\mathcal{E})\le 1-p$ if

$$\psi^L\le\eta(q^{(i)})\le\psi^U \tag{23}$$

for $i=1,\dots,\nu$ and

$$\left[\hat{C}(q^{(i)})\hat{P}^{-1}\hat{C}(q^{(i)})^T\right]_{jj}\le\left[\psi^U-\eta(q^{(i)})\right]_j^2 \tag{24a}$$
$$\left[\hat{C}(q^{(i)})\hat{P}^{-1}\hat{C}(q^{(i)})^T\right]_{jj}\le\left[\eta(q^{(i)})-\psi^L\right]_j^2 \tag{24b}$$

for $i=1,\dots\nu$, $j=1,\dots n_\psi$, where $[\,]_{ij}$ denotes element ij.

Proof: For any given q it is easy to show that $\max_{z\in\mathcal{E}}[\hat{C}(q)z]_j=[\hat{C}(q)\hat{P}^{-1}\hat{C}(q)^T]_{jj}^{1/2}$. It follows from (20) that $\Pr(\psi_k\in I_\psi)\ge p$ whenever $z_{0|k}\in\mathcal{E}$ if

$$\left[\hat{C}(q)\hat{P}^{-1}\hat{C}(q)^T\right]_{jj}^{1/2}\le\left[\psi^U-\eta(q)\right]_j \tag{25a}$$
$$\left[\hat{C}(q)\hat{P}^{-1}\hat{C}(q)^T\right]_{jj}^{1/2}\le\left[\eta(q)-\psi^L\right]_j \tag{25b}$$

for all $q\in\mathcal{Q}$ and $j=1,\dots,n_\psi$. Since the conditions (25a) and (25b) are convex in q, the equivalent constraints (23)–(24) are obtained by invoking (25) at each vertex of the polytope $\mathcal{Q}$. ■

The two lemmata above are given for the general case, the first for invariance with probability p and the second for constraint violation probability $1-p$. However, for the particular case of the ellipsoids $\mathcal{E}_1$ and $\mathcal{E}_2$, the values for these probabilities (and the corresponding probabilities for the confidence polytopic region $\mathcal{Q}$) will be different. Thus for $\mathcal{E}_1$ the relevant p is p_{11} in Lemma 3 and $1-p_1$ in Lemma 4, whereas for $\mathcal{E}_2$ the corresponding values are 1 and $1-p_2$. For the appropriate choice of these probabilities then it is possible to maximize (as seems desirable) the volume of the projection $\mathcal{E}_x$ of $\mathcal{E}_1$ or $\mathcal{E}_2$ by performing once and offline the optimization

$$\underset{\hat{P}^{-1},\ \lambda\in[0,1]}{\text{maximize}}\ \det(\hat{P}_x^{-1}) \text{ subject to } (21),(23),(24). \tag{26}$$

Remark 3: If λ is a constant, then the constraints in (26) are LMIs in $\hat{P}^{-1}$. Therefore $\mathcal{E}_x$ can be optimized by successively maximizing det $(\hat{P}_x^{-1})$ over the variable $\hat{P}^{-1}$ subject to (21), (23), (24), with the scalar λ fixed at a sequence of values in the interval $[0,1]$.

Remark 4: For input constraints: $u^L \le u_k \le u^U$, (23)–(24) are equivalent to $u^L \le 0 \le u^U$ and

$$\left[\hat{K}^T\hat{P}^{-1}\hat{K}\right]_{jj} \le \left[u^U\right]_j^2, \quad \left[\hat{K}^T\hat{P}^{-1}\hat{K}\right]_{jj} \le \left[u^L\right]_j^2 \tag{27}$$

for $j=1,\ldots,n_u$. In the example of Section IV with $\mathcal{E}_1=\mathcal{E}_x$, this implies $p_1=0$.

VI. Receding Horizon Control

If the plant state is retained within $\mathcal{E}_x$ (taken here to be the projection onto x-space of $\mathcal{E}_1$) with probability p_{11} and returned to $\mathcal{E}_x$ with probability greater than or equal to p_{12}, then by the arguments of Section IV, both the predicted and the closed-loop responses are guaranteed to satisfy the constraints (4). The aim of the online MPC is thus two-fold: (a) minimize the cost (12) subject to $z_{0|k}\in\mathcal{E}$ whenever $x_k\in\mathcal{E}_x$; or (b) return the plant state to $\mathcal{E}_x$ as quickly as possible whenever $x_k\notin\mathcal{E}_x$. The latter is achieved by driving $\mathbb{E}_k(x_{k+1})$ as close to (or as far inside) $\mathcal{E}_x$ as possible. This strategy is an indirect (and possibly suboptimal) but computationally convenient means of increasing p_{12}. The algorithm can be stated as follows.

Algorithm 1: At time k implement $u_k=Kx_k+Ef_k^*$, where

1. if $x_k\in\mathcal{E}_x$, then

$$f_k^* = \arg\min_{f_k}\begin{bmatrix} z_{0|k}\\ 1\end{bmatrix}^T\tilde{P}\begin{bmatrix} z_{0|k}\\ 1\end{bmatrix}$$
$$\text{subject to } z_{0|k}^T\hat{P}z_{0|k}\le 1 \tag{28}$$

2. otherwise (i.e., if $x_k\notin\mathcal{E}_x$),

$$f_k^* = \arg\min_{f_k} z_{0|k}^T\bar{\Psi}\Gamma\hat{P}_x\Gamma^T\bar{\Psi}z_{0|k}$$
$$\text{subject to } \begin{bmatrix} z_{0|k}\\ 1\end{bmatrix}^T\tilde{P}\begin{bmatrix} z_{0|k}\\ 1\end{bmatrix}\le\begin{bmatrix} x_k\\ Mf_{k-1}^*\\ 1\end{bmatrix}^T\tilde{P}\begin{bmatrix} x_k\\ Mf_{k-1}^*\\ 1\end{bmatrix}. \tag{29}$$

Both (28) and (29) require the minimization of a convex quadratic cost subject to a convex quadratic constraint, which can be solved efficiently using the technique of [12]. The constraint in (29) is introduced in order to ensure that the time-average of the expected value of $z_{0|k}^T\tilde{Q}z_{0|k}$ converges to a finite limit, as shown by the following theorem.

Theorem 5: The closed-loop response of (1) under Algorithm 1 satisfies

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}\mathbb{E}_0\left(\begin{bmatrix} x_k\\ f_k^*\end{bmatrix}^T\tilde{Q}\begin{bmatrix} x_k\\ f_k^*\end{bmatrix}\right)\le\operatorname{tr}(\Theta\tilde{Q}). \tag{30}$$

Proof: Let J_k^* denote the value of J_k corresponding to the optimal f_k^* computed by Algorithm 1 at time k. Then

$$\mathbb{E}_k(J_{k+1}^*) = \mathbb{E}_k(J_{k+1}^*|x_{k+1}\in\mathcal{E}_x)\Pr(x_{k+1}\in\mathcal{E}_x) + \mathbb{E}_k(J_{k+1}^*|x_{k+1}\notin\mathcal{E}_x)\Pr(x_{k+1}\notin\mathcal{E}_x)$$

where J_{k+1}^* necessarily satisfies

$$J_{k+1}^*\le\begin{bmatrix}\Psi_k z_{0|k}+\delta_k\\ 1\end{bmatrix}^T\tilde{P}\begin{bmatrix}\Psi_k z_{0|k}+\delta_k\\ 1\end{bmatrix}$$

(because of the objective in (28) if $x_{k+1}\in\mathcal{E}_x$ or the constraint in (29) if $x_{k+1}\notin\mathcal{E}_x$), and hence

$$\mathbb{E}_k(J_{k+1}^*)\le\begin{bmatrix} z_{0|k}\\ 1\end{bmatrix}^T\mathbb{E}\left(\begin{bmatrix}\Psi_k & \delta_k\\ 0 & 1\end{bmatrix}^T\tilde{P}\begin{bmatrix}\Psi_k & \delta_k\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix} z_{0|k}\\ 1\end{bmatrix}. \tag{31}$$

However, from (13a)–(13c) it follows that

$$\mathbb{E}\left(\begin{bmatrix}\Psi_k & \delta_k\\ 0 & 1\end{bmatrix}^T\tilde{P}\begin{bmatrix}\Psi_k & \delta_k\\ 0 & 1\end{bmatrix}\right)=\tilde{P}-\begin{bmatrix}\tilde{Q} & 0\\ 0 & -\operatorname{tr}(\Theta\tilde{Q})\end{bmatrix}$$

and (31) therefore implies that

$$J_k^*-\mathbb{E}_k(J_{k+1}^*)\ge z_{0|k}^T\tilde{Q}z_{0|k}-\operatorname{tr}(\Theta\tilde{Q}).$$

Recursion of this equation for $k=0,1,\ldots$ gives

$$\lim_{n\to\infty}\frac{1}{n}\Big(J_0^*-\mathbb{E}_0\left(J_n^*\right)\Big)\ge\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}\mathbb{E}_0\left(z_{0|k}^T\tilde{Q}z_{0|k}\right)-\operatorname{tr}(\Theta\tilde{Q})$$

and, since J_n^* is lower bounded [because $P_z>0$ in (13a)], it follows that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}\mathbb{E}_0\left(z_{0|k}^T\tilde{Q}z_{0|k}\right)\le\operatorname{tr}(\Theta\tilde{Q})$$

which implies the time-average bound (30). To complete the proof, note that the constraint in (28) is feasible for $x_k\in\mathcal{E}_x$ due to the definition of $\mathcal{E}_x$, and similarly the feasibility of $f_k=Mf_{k-1}^*$ implies that the constraint in (29) is necessarily feasible. ■

Corollary 6: If the probabilities p_{11}, p_{12} are such that, for the conditional constraint violation probability p_1, the expected rates R_1, R_2 of accumulation of constraint violations are within allowable limits, then the bound (4) will be satisfied in closed-loop operation under Algorithm 1.

Proof: This follows from the assumptions on R_1, R_2 and the arguments of Section IV. ■

The algorithm must be initialized by computing $\hat{P}$. A possible procedure for this is as follows: specify initial values p_{11}^0, p_{12}^0 for p_{11}, p_{12}. Then, from the maximum allowed rate $N_{\max}/N_c$ of constraint violations, the analysis of Section IV can be used to compute the minimum permissible value for p_1. Given p_{11}, p_{12} and p_1, the uncertainty set $\mathcal{Q}$ can be constructed and the constraints (21), (23), and (24) formulated, allowing $\hat{P}$ to be optimized by solving (26). The actual value of p_{12} can be computed given $\hat{P}$ (e.g., by using a sufficiently large number of randomized samples of uncertainty and searching over the boundary of $\mathcal{E}_2$). To ensure satisfaction of (4) we require $p_{12}\ge p_{12}^0$, otherwise $\hat{P}$ must be recomputed using reduced values for p_{11}^0, p_{12}^0.

Note that this trial and error procedure may involve several iterations and hence may be computationally demanding, particularly for high-dimensional problems. However, the computation is performed once, offline, and can moreover be greatly reduced by reducing the number of degrees of freedom. Specifically, the optimization of (26) can be used to to design $\mathcal{E}_2$, with the probability of invariance set to 1 and constraint violation probability equal to 1. Then, by defining $\mathcal{E}_1$ as a scaled version of $\mathcal{E}_2$, i.e., $\mathcal{E}_1=\alpha\mathcal{E}_2$ for scaling factor α, a one-dimensional search over values of $\alpha\in(0,1)$ can be used to determine the smallest achievable value for i^*. For each value of α the values of p_1, p_2, p_{12}, p_{22} can be determined by randomized simulations and used in the determination of i^*. This is the procedure employed in the example of Section VII.

Remark 5: If several desired intervals are specified for ψ, each with a bound on the expected number of violations, then the appropriate value for p_1 can be computed based on a weighted average rate of constraint violation. This situation is common when constraints on fatigue damage due stress cycles of varying amplitudes are considered (e.g., using Miner's rule).

VII. Simulation Example

Consider the problem of maximizing the power capture of a variable pitch wind turbine subject to constraints on turbine blade fatigue due to wind fluctuations. It is reasonable to assume that the wind speed statistics are constant over a period of the order of ten minutes [13]. Below rated average wind speed the control objective is to maximize efficiency, achieved by regulating blade pitch angle about a given setpoint. In order to achieve a specified fatigue life, this is performed subject to constraints on the stress cycles experienced by the blades.

A simplified model of blade pitch rotation is given by

$$J\frac{d^2\beta}{dt^2} + c\frac{d\beta}{dt} = T_m - T_p \tag{32}$$

where β is the blade pitch angle, T_m is a torque applied by an actuator used to adjust β, and T_p is the pitching torque due to fluctuations in wind speed, which is a known function of wind speed and the blade's angle of attack, α. Note that α is related in a known manner to wind speed and β. Therefore the model (32) is subject to additive stochastic uncertainty (due to the dependence of T_p on wind speed) and multiplicative uncertainty (due to the dependence of T_p on β), and furthermore these two sources of uncertainty are statistically dependent. Blade fatigue damage depends on the resultant torque, so soft constraints are invoked on an output: $\psi = T_m - T_p$.

By considering variations about a given setpoint for β, a linear discrete model approximation was identified in the form of an ARMA model (with sampling interval 1 s)

$$\psi_{k+1} = a_{k,1}\psi_k + a_{k,0}\psi_{k-1} + b_{k,1}u_k + b_{k,0}u_{k-1} + w_k \tag{33}$$

using data applied to a continuous-time model of the NACA 632–215(V) blade [13]. Least squares estimates of $\theta = [\,a_1 \quad a_0 \quad b_1 \quad b_0 \quad w\,]^T$ were obtained from a set of 1000 open loop simulations, each with a fixed wind speed sampled from the Weibull distribution. The mean $\bar{\theta}$ and covariance Σ_θ of a Gaussian distribution for θ were identified from these simulations.

The ARMA model (33) can be written in the form (1), with

$$A_k = \begin{bmatrix} 0 & a_{k,2} \\ 1 & a_{k,1} \end{bmatrix}, \quad B_k = \begin{bmatrix} b_{k,2} \\ b_{k,1} \end{bmatrix}, \quad d_k = \begin{bmatrix} 0 \\ w_k \end{bmatrix}.$$

The identified $(\bar{\theta}, \Sigma_\theta)$ indicate that B has negligible uncertainty and give an uncertainty class

$$\begin{aligned}
\begin{bmatrix} A_k & d_k \end{bmatrix} &= \begin{bmatrix} \bar{A} & 0 \end{bmatrix} + \sum_{j=1}^{3}\begin{bmatrix} \tilde{A}_j & \tilde{g}_j \end{bmatrix} q_{k,j} \\
\bar{A} &= \begin{bmatrix} 0 & -0.97 \\ 1 & 1.56 \end{bmatrix} \\
\begin{bmatrix} \tilde{A}_1 & \tilde{d}_1 \end{bmatrix} &= \begin{bmatrix} 0 & -0.09 & 0 \\ 0 & 0.13 & 0.02 \end{bmatrix} \\
\begin{bmatrix} \tilde{A}_2 & \tilde{d}_2 \end{bmatrix} &= \begin{bmatrix} 0 & 0.21 & 0 \\ 0 & -0.009 & -0.06 \end{bmatrix} \\
\begin{bmatrix} \tilde{A}_3 & \tilde{d}_3 \end{bmatrix} &= \begin{bmatrix} 0 & -0.06 & 0 \\ 0 & 0.02 & 0.05 \end{bmatrix} \\
\bar{B} &= \begin{bmatrix} -0.20 \\ -0.21 \end{bmatrix}.
\end{aligned}$$

The Gaussian distribution for q_k was validated by the Jarque-Bera test at the 5% level. A linear approximation of the sampled output ψ_k was estimated using a similar approach. The uncertainty in D was found to be negligible, and the uncertainty class for $[\,C \quad \eta\,]$ was formulated as

$$\begin{gathered}
\begin{bmatrix} C_k & \eta_k \end{bmatrix} = \begin{bmatrix} \bar{C} & 0 \end{bmatrix} + \sum_{j=1}^{2}\begin{bmatrix} \tilde{C}_j & \tilde{\eta}_j \end{bmatrix} q_{k,j} \\
\bar{C} = \begin{bmatrix} 0 & 729 \end{bmatrix}, \quad \bar{D} = 959 \\
\begin{bmatrix} \tilde{C}_1 & \eta_1 \end{bmatrix} = \begin{bmatrix} 0 & 300 & 50 \end{bmatrix}, \quad \begin{bmatrix} \tilde{C}_2 & \eta_2 \end{bmatrix} = \begin{bmatrix} 0 & 50 & 100 \end{bmatrix}.
\end{gathered}$$

On the basis of the Gaussian distribution for q, a chi-square distribution for $\|q\|_2^2$ with three degrees of freedom was used to compute polytopic sets $\mathcal{Q}$ by over-bounding the sphere $\{q : \|q\|_2^2 \le r^2\}$, for r satisfying $\Pr(\chi^2(3) \le r^2) = p$.

A prediction horizon of $N = 4$ was employed, and $N_c = 40$ was used as the horizon over which to invoke the upper bound $N_{\max}$ on the permissible number of constraint violations. Miner's rule was used to determine $N_{\max}/N_c$, assuming (for simplicity) a single threshold on the torque $T_m - T_p$. Accordingly, for $p_{11}^0 = 0.9$, $p_{12}^0 = 0.8$, $N_{\max}/N_c = 0.3$, the permissible value for p_1 was found to be 0.2. For these values, the optimization (26) gave

$$\hat{P}_x = \begin{bmatrix} 0.03 & 0.04 \\ 0.04 & 0.069 \end{bmatrix}.$$

Closed-loop simulations of Algorithm 1, performed for an initial condition $x_0 = [\,-7.88 \quad 7.31\,]^T$ (which is close to the boundary of $\mathcal{E}_x$), gave an average number of constraint violations of 3 over a horizon of 40 steps, while the maximum number of constraint violations on any one simulation run was 4. From these simulations, the actual value of p_{12} was found to be 0.85, which exceeds p_{12}^0, indicating that Algorithm 1 satisfies the fatigue constraints.

To establish the efficacy of Algorithm 1, closed-loop simulations were performed for 1000 sequences of uncertainty realizations, and compared in terms of cost and constraint satisfaction with the mean square stabilizing linear feedback law $u_k = Kx_k$. Algorithm 1 gave an average closed-loop cost of 257, whereas the average cost for $u_k = Kx_k$ was 325. Algorithm 1 achieves this improvement in performance by driving (during transients) the predictions hard against the limits of the soft constraints. Both Algorithm 1 and $u_k = Kx_k$ on average resulted a total number of constraint violations within the specified limit over a 40-step horizon. This is to be expected since both control laws achieve acceptable rates of constraint violation in steady state. However, detailed examination of the average numbers of constraint violations over n steps, for $0 < n \le 16$, shows that $u_k = Kx_k$ exceeded the allowable limits during transients, whereas Algorithm 1 gave average constraint violation rates less than $N_{\max}/N_c = 0.3$ for all $n \ge i^* = 13$.

VIII. Conclusions

Robust MPC can be conservative in cases where the distribution of uncertainty is known. Earlier work addressed this issue for the case of multiplicative uncertainty and the current technical note extends the approach to the more challenging case of multiplicative and additive uncertainty. The proposed MPC strategy is shown to have desirable closed-loop stability and feasibility properties. A key ingredient is the concept of probabilistic invariance, which in earlier work was handled conservatively through the use of confidence ellipsoids in the state space. The work discussed here proposes a significant improvement through the use of the vertices of polytopic sets that are known to contain the vector of uncertain parameters with a given probability. The effectiveness of the method is illustrated through a simulation example concerned with maximizing power capture by a wind turbine while respecting soft constraints dictated by limits on fatigue damage. The extension of the work to the case of uncertainty distributions which are not independently identically distributed is a subject of future research.

REFERENCES

[1] M. Cannon, B. Kouvaritakis, and X. Wu, "Model predictive control for systems with stochastic multiplicative uncertainty and probabilistic constraints," *Automatica*, vol. 45, no. 1, pp. 167–172, Jan. 2009.

[2] P. Scokaert and J. Rawlings, "Constrained linear quadratic regulation," *IEEE Trans. Autom. Control*, vol. 43, no. 8, pp. 1163–1169, Aug. 1998.

[3] M. Kothare, V. Balakrishnan, and M. Morari, "Robust constrained model predictive control using linear matrix inequalities," *Automatica*, vol. 32, no. 10, pp. 1361–1379, 1996.

[4] D. Mayne, J. Rawlings, C. Rao, and P. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[5] B. Polyak and R. Tempo, "Probabilistic robust design with linear quadratic regulators," *Syst. Control Lett.*, vol. 43, no. 5, pp. 343–353, 2001.

[6] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized Algorithms for Analysis and Control of Uncertain Systems*. New York/Berlin: Springer, 2005.

[7] I. Batina, A. Stoorvogel, and S. Weiland, "Optimal control of linear, stochastic systems with state and input constraints," in *Proc. 41st IEEE Conf. Decision Control*, 2002, pp. 1564–1569.

[8] B. Kouvaritakis, J. Rossiter, and J. Schuurmans, "Efficient robust predictive control," *IEEE Trans. Autom. Control*, vol. 45, no. 8, pp. 1545–1549, Aug. 2000.

[9] P. Li, M. Wendt, and G. Wozny, "A probabilistically constrained model predictive controller," *Automatica*, vol. 38, no. 7, pp. 1171–1176, 2002.

[10] H. Kushner, *Introduction to Stochastic Control*. New York: Holt, Rinehart and Winston, 1971.

[11] S. Boyd, L. G. El, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: SIAM, 1994.

[12] B. Kouvaritakis, M. Cannon, and J. Rossiter, "Who needs QP for linear MPC anyway?," *Automatica*, vol. 38, no. 5, pp. 879–884, 2002.

[13] T. Burton, D. Sharpe, N. Jenkins, and E. Bossanyi, *Wind Energy Handbook*. New York: Wiley, 2001.

$\mathcal{L}_2$ Gain of Periodic Linear Switched Systems: Fast Switching Behavior

D. Gray Roberson and Daniel J. Stilwell

***Abstract*—We investigate the $\mathcal{L}_2$ gain of periodic linear switched systems under fast switching. For systems that possess a suitable notion of a time-average system, we characterize the relationship between the $\mathcal{L}_2$ gain of the switched system and the $\mathcal{L}_2$ gain of its induced time-average system when the switching rate is sufficiently fast. We show that the switched system $\mathcal{L}_2$ gain is in general different from the average system $\mathcal{L}_2$ gain if the input or output coefficient matrix switches. If only the state coefficient matrix switches, the input-output energy gain for a fixed $\mathcal{L}_2$ input signal is bounded by the $\mathcal{L}_2$ gain of the average system as the switching rate grows large. Additionally, for a fixed $\mathcal{L}_2$ input, the maximum pointwise in time difference between the switched and average system outputs approaches zero as the switching rate grows.**

***Index Terms*—Fast switching, $\mathcal{L}_2$ gain analysis, switched systems.**

I. INTRODUCTION

Switched systems comprise an important class of dynamical systems, and a rich body of literature has grown out of research efforts in this field. Excellent surveys of switched system theory can be found for example in [1]–[4]. In examining these works and the broader literature, one observes that stability analysis and design stabilization have been dominant themes in switched system research. Other system properties, such as controllability and robustness to different switching regimes, have been areas of considerable focus as well. In contrast, input-output behavior, and particularly $\mathcal{L}_2$ gain properties, of switched systems have received relatively little attention, even though such characteristics can provide valuable design criteria and useful performance measures for system properties such as disturbance attenuation.

Recent investigations have begun to address the need for characterization of switched system input-output performance. For example, a bound for the root-mean-square gain (defined similarly to the $\mathcal{L}_2$ gain) is reported in [5] for a system that switches among asymptotically stable subsystems. The bound is derived by considering the stabilizing and destabilizing solutions of algebraic Riccati equations associated with the subsystems. In [6], multiple Lyapunov function and dwell time analysis is used to specify a weighted $\mathcal{L}_2$ gain for a system that switches among possibly unstable subsystems, so long as the stable subsystems are active for a large time relative to the unstable subsystems. Multiple Lyapunov function analysis is also used in [7], where the $\mathcal{L}_2$ gain of systems with state-dependent switching is studied. Under state-dependent switching, subsystems are allowed to be unstable.

Our goal in this technical note is to address $\mathcal{L}_2$ gain of fast switching systems, a class of switched systems for which $\mathcal{L}_2$ gain has yet to be addressed. These systems are composed of subsystems that are not necessarily stable, but the subsystems induce a time-average system that is stable. We consider switched systems with the form

$$\begin{aligned} \dot{x} &= A_{\rho(t)}x + B_{\rho(t)}u \quad x(0) = x_0, \;\; t \geq 0 \\ y &= C_{\rho(t)}x \end{aligned} \tag{1}$$

Manuscript received February 21, 2008; revised September 17, 2008. First published June 10, 2009; current version published July 09, 2009. This work was supported by the National Science Foundation under Grant IIS-0238092 and by the Office of Naval Research under Grant N00014-03-1-0444. Recommended by Associate Editor D. Angeli.

The authors are with the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061 USA (e-mail: grayr@vt.edu; stilwell@vt.edu).

Digital Object Identifier 10.1109/TAC.2009.2017974