Risk-Aware Sparse Predictive Control ⋆

Zhicheng Zhang * and **Yasumasa Fujisaki** *

* *Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan (email: {zhicheng-zhang, fujisaki}@ist.osaka-u.ac.jp)*

Abstract: This note presents a risk-aware sparse predictive control for linear discrete-time systems subject to model parametric uncertainty and additive disturbances. We consider the case of probabilistic constraints on system state and hard constraints on control actions, which is equivalent to solving a chance-constrained sparse optimization problem. We then adopt a data-driven sampling approach to seek a "randomized" solution by solving a random convex program that approximates the risk-aware sparse solution with a high probability. We also give an explicit sample complexity to ensure the probabilistic robustness. Finally, the numerical example illustrates the effectiveness of the proposed control strategy.

Keywords: sparse predictive control, probabilistic robustness, data-driven sampling

1. INTRODUCTION

Control effort minimization is a critical issue in modern control and applications, where a sparsity promoting control index is performed to reduce resource consumption. By maximizing the time interval over which the control signal is exactly zero, the resulting control paradigm is referred to as "sparse control", "hands-off control", or "ℓ^1 optimal control" (Nagahara et al., 2016). It also has been extended to model predictive control called sparse predictive control (Gallieri and Maciejowski, 2012; Nagahara et al., 2016; Mishra et al., 2018; Darup et al., 2021).

In this note, we propose a risk-aware sparse predictive control for an uncertain linear discrete-time system, which is closely related to stochastic model predictive control (Calafiore and Fagiano, 2012; Mesbah, 2016). We here examine a more general framework of uncertain systems, where the uncertainties include not only "additive" disturbances but also "model parametric" instances. We became aware of the recent results (Nagahara et al., 2016; Darup et al., 2021), where a worst-case robust scenario of sparse predictive control were considered. However, such constraint setting requires to satisfy for *all* uncertainties and the obtained solution may be overly conservative and pessimistic. To this end, we slightly relax the hard constraint to take a stochastic model into account, where the state constraints are assumed to be probabilistic sense. By this, it ensures *probabilistic robustness* if the constraint holds for "most" of the instances of the uncertainty.

In particular, we follow a data-driven sampling approach called "scenario approach" (Campi and Garatti, 2018) or "randomized algorithms" (Tempo et al., 2012) to address the risk-aware sparse predictive control problem underlying a chance-constrained sense. We then seek a *data-driven* sparse solution to approximate the *risk-aware* sparse solution with a high probability, which is a random convex program, and hence is computationally tractable. The numerical example demonstrates the effectiveness of the proposed control strategy.

⋆ This work was supported by JSPS KAKENHI Grant Number JP20K04547.

2. PROBLEM FORMULATION

2.1 System Dynamics, Constraints, and Assumptions

Consider an uncertain linear discrete-time system

$$x_{t+1} = A(\delta)x_t + B(\delta)u_t + Ew_t, \qquad x_0 \neq 0, \tag{1}$$

where $x_t \in \mathbb{R}^n$ is the system state, $u_t \in \mathbb{R}^m$ is the control input, $w_t \in \mathcal{W} \subseteq \mathbb{R}^{n_w}$ is the additive disturbance with $E \in \mathbb{R}^{n\times n_w}$, and $\delta \in \Delta \subseteq \mathbb{R}^{n_\delta}$ stands for time invariant modeling uncertain parameters of $A(\delta) \in \mathbb{R}^{n\times n}$ and $B(\delta) \in \mathbb{R}^{n\times m}$.

Assumption 1. (Stabilizability). The pair $(A(\delta), B(\delta))$ is stabilizable for any uncertain realizations $\delta \in \Delta$.

Assumption 2. (Stochastic Uncertainty). Assume that the sets Δ and $\mathcal{W}$ are bounded. The uncertain realizations δ and w_t are independent and identically distributed (i.i.d.), randomly extracting from the probability distributions $\mathbb{P}_\delta$ on Δ and $\mathbb{P}_w$ on $\mathcal{W}$, respectively.

Next, we present a prediction system. Given the state x_t observed at sampling time t, let $N \in \mathbb{N}$ be *prediction horizon*, then the predicted, future states are modelled as

$$x_{j+1|t} = A(\delta)x_{j|t} + B(\delta)u_{j|t} + Ew_{j|t}, \qquad j = 0, 1, \ldots, N-1. \tag{2}$$

Here, we mark the subscript $\bullet_{j|t}$ as predictive instants, e.g., $x_{j|t}$ denotes the jth step forward prediction of the state at time t that is initialized at $x_{0|t} = x_t$, and similarly, $u_{j|t}$ and $w_{j|t}$ are the related predictive control input and disturbance, respectively.

In this note, the uncertain prediction system (2) is subject to the *soft state* constraint (i.e., probabilistic or chance constrained) and the *hard control input* constraint

$$\mathbb{P}\{Cx_{j+1|t} \leq c,\ j = 0, 1, \ldots, N-1\} \geq 1 - \epsilon, \tag{3a}$$

$$Du_{j|t} \leq d, \tag{3b}$$

where C and D are state and input constraint matrices governed by the regarding vectors c, d with appropriate sizes, and $\epsilon \in (0, 1)$ is the desired level of accuracy or risk. For simplicity, we write the random variables of (2) as $\theta \doteq (\delta, \bar{w})$ according to probability $\mathbb{P} \doteq \mathbb{P}_\delta \times \mathbb{P}_w^N$, where

$$\bar{w} = \left[w_{0|t}^T \ w_{1|t}^T \ \cdots \ w_{N-1|t}^T\right]^T \in \mathbb{R}^{n_w N}.$$

Here θ takes value in $\Theta \doteq \Delta \times \mathcal{W}^N$, i.e., $\theta \in \Theta \sim \mathbb{P}$.

Due to the presence of the uncertainties θ, regulation of the system state to the origin can not be achieved in general. Conversely, we intend to seek a "risk-aware" predictive control sequences

$$\bar{u} = \left[u_{0|t}^T \ u_{1|t}^T \ \cdots \ u_{N-1|t}^T\right]^T \in \mathbb{R}^{mN} \tag{4}$$

of length $N \in \mathbb{N}$ such that hedge against the stochastic perturbations θ and drive the uncertain prediction state (2) from initial state $x_{0|t}$ towards a prescribed terminal set. In other words, we replace (3a) with

$$\mathbb{P}\{C_f x_{N|t} \leq c_f,\, Cx_{j+1|t} \leq c,\, j = 0, 1, \ldots, N-1\} \geq 1 - \epsilon, \tag{5}$$

and need to inspect whether the condition (5) holds or not, where C_f is a state constraint matrix governed by the regarding a vector c_f with an appropriate size. The essence of the term (5) claims that the (terminal) state constraint allows a small risk level to violate itself, and it is still feasible for *most* of the uncertain instances.

2.2 Sparse Predictive Control

Similar to Darup et al. (2021), the control objective of this note is to implement a sparse predictive control, we then introduce a sparsity-promoting predictive control cost as

$$J(\bar{u}) = \sum_{j=0}^{N-1} \|u_{j|t}\|_1 = \|\bar{u}\|_1. \tag{6}$$

Note that this control cost is a convex ℓ_1 optimal (predictive) control that maximizes the time interval over which the predictive signals is exactly zero under a certain condition (Nagahara et al., 2016). Besides, here the predictive control cost (6) is a direct real-time control iterations rather than the stochastic tubed-based or parametric affine control schemes (Mesbah, 2016), since determining a feedback gain of pure sparse optimal (predictive) control problem is a nontrivial task.

To proceed the sparse predictive control, a *receding horizon strategy* (Maciejowski, 2002; Rawlings et al., 2017) is used to execute model predictive control. That is, no matter how long lengths the controller can predict, only the *first action of predictive inputs* (4)

$$u_t = u_{0|t} = F\bar{u} \tag{7}$$

is applied to the system at each sampling instant t, where

$$F \doteq [I_m \ 0] \in \mathbb{R}^{m \times Nm}$$

indicates a receding horizon scheme matrix and $I_m \in \mathbb{R}^{m \times m}$ is the identify matrix.

3. MAIN RESULTS

This section studies the main results of identifying the risk-aware sparse optimal solution $\bar{u}^*$ that minimizes an ℓ_1 optimal control cost over a finite prediction horizon. This problem is equivalent to dealing with a chance-constrained sparse optimization problem. Meanwhile, we follow a data-driven sampling approach to seek a "randomized" solution $\bar{u}_K^*$ that approximates the original chance-constrained or risk-aware sparse solution with a high probability.

3.1 Risk-Aware Sparse Predictive Control

In this subsection, we present a model predictive control with *risk-aware* sense and *sparse* manner, called risk-aware sparse predictive control, which amounts to solving a *chance-constrained sparse optimization* problem

$$\min_{\bar{x}, \bar{u}} \quad \|\bar{u}\|_1 \tag{8a}$$

$$\text{s.t.} \quad x_{j+1|t} = A(\delta)x_{j|t} + B(\delta)u_{j|t} + Ew_{j|t},$$
$$x_{0|t} = x_t, \quad j = 0, 1, \ldots, N-1, \tag{8b}$$
$$\mathbb{P}\{C_f x_{N|t} \leq c_f,\ Cx_{j+1|t} \leq c,$$
$$j = 0, 1, \ldots, N-1\} \geq 1 - \epsilon, \tag{8c}$$
$$Du_{j|t} \leq d, \quad j = 0, 1, \ldots, N-1, \tag{8d}$$

where

$$\bar{x} = \left[x_{1|t}^T \ x_{2|t}^T \ \cdots \ x_{N|t}^T\right]^T \in \mathbb{R}^{nN}.$$

The risk-aware sparse predictive control problem (8) is well-defined, however, such chance-constrained solution u^* is extremely difficult to calculate exactly. In particular, the feasible set described by chance constraints (8c) is general difficult to evaluate either analytically or numerically, since it involves the multivariate integrals computing and the relevant set is *non-convex* even under the convexity assumption for hard constraint set.

3.2 Data-Driven Sampling Approach

In what follows, we introduce a "data-driven" sampling approach called *scenario optimization* (Campi and Garatti, 2018) to deal with the risk-aware sparse predictive control problem, i.e., chance-constrained sparse program (8), which is essentially a *randomized algorithms* (Tempo et al., 2012) that generates a finite number of random instances of an uncertain variable θ according to the probability $\mathbb{P}$.

First, we randomly collect a finite number of "samples" or "scenarios" $\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(K)}\}$ from $\theta \doteq (\delta, \bar{w})$ in an i.i.d. fashion according to the probability distribution $\mathbb{P}$. We now substitute the scenario counterpart into the chance constrained sparse optimization (8). In this way, we explore the risk-aware sparse predictive control via data-driven scenario approach as

$$\min_{\bar{x}, \bar{u}} \quad \|\bar{u}\|_1 \tag{9a}$$

$$\text{s.t.} \quad x_{j+1|t}^{(i)} = A(\delta^{(i)})x_{j|t}^{(i)} + B(\delta^{(i)})u_{j|t} + Ew_{j|t}^{(i)},$$
$$x_{0|t}^{(i)} = x_t, \quad j = 0, 1, \ldots, N-1,$$
$$i = 1, 2, \ldots, K, \tag{9b}$$
$$C_f x_{N|t}^{(i)} \leq c_f, \quad Cx_{j|t}^{(i)} \leq c, \quad j = 0, 1, \ldots, N-1,$$
$$i = 1, 2, \ldots, K, \tag{9c}$$
$$Du_{j|t} \leq d, \quad j = 0, 1, \ldots, N-1, \tag{9d}$$

where the superscript i denotes the random sampling with K scenarios. It is clear that the data-driven scenario approach gives rise to a "random convex program" that can be solved by means of the off-the-shelf packages, like `CVX` or `YALMIP` in MATLAB. This data-driven sampling method generates a sparse scenario solution $\bar{u}_K^*$ to problem (9), and $\bar{u}_K^*$ is a "randomized" solution since it depends on the collected random samples $\theta^{(1)}$, $\theta^{(2)}$, ..., $\theta^{(K)}$. Without loss of generality, we here assume that the random convex program (9) admits a unique solution. Besides, the feasible set of (9) is an "inner approximation" of the feasible set of the original chance-constrained sparse optimization (8).

We next focus on the uncertain quantification.

Lemma 3. The uncertain part of the feasible set defined by (9b) and (9c) is represented as

$$h(\bar{u},\theta) \leq 0 \tag{10}$$

with a function $h(\bar{u},\theta) : \mathbb{R}^{mN} \times \Theta \to \mathbb{R}$ which is convex in $\bar{u}$ for any fixed $\theta \in \Theta$.

See Appendix A for details.

3.3 A Finite-Sample Guarantee

In general, the obtained data-driven solution $\bar{u}_K^*$ does not always successfully approximate the risk-aware solution $\bar{u}^*$ thanks to the randomness of the data sampling. In order to hold the probabilistic robustness guarantee with a risk ϵ and a confidence level β, sufficient data-driven samples must be collected. In what follows, we state the sample complexity to ensure the data-driven sparse solution $\bar{u}_K^*$ is a good approximation for risk-aware sparse solution $\bar{u}^*$.

Theorem 4. Suppose that an uncertain constraint function $h(\bar{u},\theta)$ defined in (10) is given. Assume that the uncertain variable $\theta \doteq (\delta,\bar{w})$ has the probability distribution $\mathbb{P}$ supported on Θ. Let $\Theta^K \doteq \{\theta^{(1)},\theta^{(2)},\ldots,\theta^{(K)}\}$ be a multi-sample of θ, where K is selected so that it satisfies

$$\sum_{i=0}^{mN-1} \binom{K}{i} \epsilon^i (1-\epsilon)^{K-i} \leq \beta \tag{11}$$

for given specified risk $\epsilon \in (0,1)$ and confidence $\beta \in (0,1)$. Suppose that there exists a unique optimal solution for the random convex program (9), i.e.,

$$\begin{aligned} \min_{\bar{x},\bar{u}} \quad & \|\bar{u}\|_1 \\ \text{s.t.} \quad & h(\bar{u},\theta^{(i)}) \leq 0, \quad i = 1,2,\ldots,K \\ & Du_{j|t} \leq d, \quad j = 0,1,\ldots,N-1. \end{aligned} \tag{12}$$

Then, with confidence $1-\beta$, a probabilistic guarantee

$$\mathbb{P}^K\{h(\bar{u}_K^*,\theta) \leq 0\} \geq 1-\epsilon$$

holds true for the optimal input $\bar{u}_K^*$ obtained by (12).

Remark 5. (Sample Complexity). The satisfaction of term (11) implies that the following sample complexity holds

$$K \geq \frac{mN-1+\ln(1/\beta)+\sqrt{2(mN-1)\ln(1/\beta)}}{\epsilon}, \tag{13}$$

where the right-hand-side term in (13) can be defined as $\mathcal{K}(mN,\epsilon,\beta)$ (Alamo et al., 2010).

Then a direct result is summarized as follows.

Corollary 6. Let $K \geq \mathcal{K}(mN,\epsilon,\beta)$ and $\bar{u}_K^*$ be a feasible optimal control for Problem (9). If $u_{j|t}^*$ is applied to the uncertain system (2) with a finite prediction horizon N, then, with probability $1-\beta$, the risk-aware sparse predictive control (8) is achieved for $j = 0,1,\ldots,N-1$.

4. SIMULATIONS

In this section, we illustrate the risk-aware sparse predictive control on a benchmark two-mass-spring system (Kothare et al., 1996) shown in Fig. 1. The discrete dynamics by Euler's approximation with sampling time $t_s = 0.1$ s is of the form as follows

$$A(\delta) = \begin{bmatrix} 1 & 0 & t_s & 0 \\ 0 & 1 & 0 & t_s \\ -\dfrac{k_s t_s}{m_1} & \dfrac{k_s t_s}{m_1} & 1 & 0 \\ \dfrac{k_s t_s}{m_2} & -\dfrac{k_s t_s}{m_2} & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ \dfrac{t_s}{m_1} \\ 0 \end{bmatrix},$$

$$E = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \dfrac{t_s}{m_1} & 0 \\ 0 & \dfrac{t_s}{m_2} \end{bmatrix}, \qquad x_t = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \qquad w_t = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

The model state x_t has four dimensions, in which the first two components $\{x_1,x_2\}$ stand for positions of two masses, the rest components $\{x_3,x_4\}$ represent velocities. Assume that the additive disturbance w_t has a Gaussian distribution $w \sim \mathrm{N}(0,\Sigma_w)$ with zero mean and covariance matrix $\Sigma_w = \mathrm{diag}(0.02^2, 0.02^2)$. The masses of system are set as $m_1 = m_2 = 1$ kg, and the "elastic constant" k_s is related to parametric uncertainty that follows a uniform probability distribution $k_s \sim \mathrm{Unif}([0.5, 2.0])$.

We take the initial state as

$$x_0 = [0.15 \;\; 0.15 \;\; -0.15 \;\; -0.1]^T,$$

and the prediction horizon is $N = 6$, and the sampling time $t = 50$. For the data-driven sampling setup, we choose the risk $\epsilon = 0.05$, confidence level $\beta = 10^{-6}$ and by means of sample complexity (13), we take scenarios $K = 695$. Next, we use these parameters to prescribe the constraints.

In two-mass-spring system, the standard setting for the mass positions (i.e., the first two states x_1 and x_2) are often free. Hence, we only give the sampled-based velocities constraints $|x_3| \leq 0.15$, $|x_4| \leq 0.15$, and hard control inputs $|u_t| \leq 1$. Meanwhile, we conduct a simulation for quadratic MPC, the ℓ_2 optimal control (i.e., $\min_{\bar{u}} \|\bar{u}\|_2$) so as to make a comparison with sparse predictive control. Fig. 2 displays the control inputs with ℓ_2 optimal and ℓ_1 optimal controllers, respectively. It is clearly that sparse predictive control promotes more zero inputs than quadratic MPC, therefore it enjoys sparsity.

Fig. 3 illustrates the four different state trajectories under risk-aware sparse predictive control. Due to the nonzero uncertainties, the state trajectories will tend to the neighborhood of the origin. In particular, we can see that the velocity x_3 tends to the assigned upper bound 0.15 as well as the velocity x_4 gets close to the pre-specified

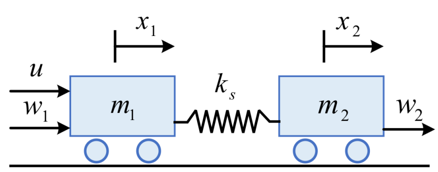

Fig. 1. Two-mass-spring system

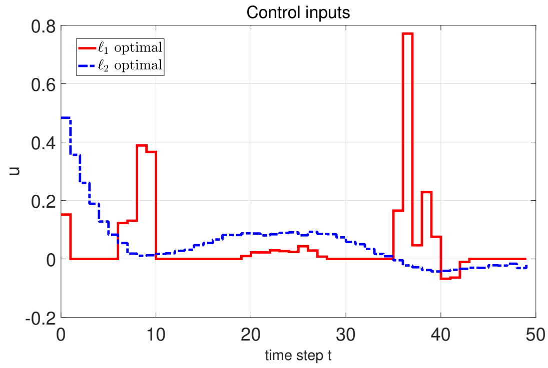

Fig. 2. Profiles of sparse (ℓ_1) predictive control (solid) and quadratic (ℓ_2) model predictive control (dashed).

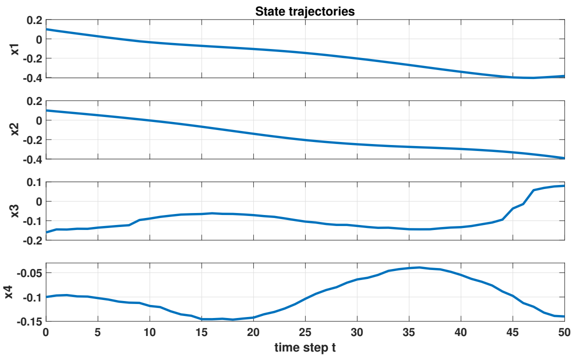

Fig. 3. State trajectories with prediction horizon $N = 6$ and time step $t = 50$ s.

lower bound -0.15. This implies that the sparse predictive control offers a good robustness.

5. CONCLUSION

In this note, we presented a risk-aware sparse predictive control, where the state constraints are considered as probabilistic forms. To solve such chance-constrained sparse optimization problem, we took a data-driven scenario approach to seek a randomized solution by solving a random convex program such that it approximates the original risk-aware sparse solution with a high probability. Meanwhile, we provided a finite sample guarantee for the obtained data-driven scenario sparse solution. Although technical difficulties were not detected for introducing sparse control into the scenario approach to MPC so far, the numerical example showed that the usefulness of this approach.

REFERENCES

Alamo, T., Tempo, R., and Luque, A. (2010). On the sample complexity of randomized approaches to the analysis and design under uncertainty. In *Proceedings of the 2010 American Control Conference*, 4671–4676. IEEE.

Calafiore, G.C. and Fagiano, L. (2012). Robust model predictive control via scenario optimization. *IEEE Transactions on Automatic Control*, 58(1), 219–224.

Campi, M.C. and Garatti, S. (2018). *Introduction to the scenario approach*. SIAM.

Darup, M.S., Book, G., Quevedo, D.E., and Nagahara, M. (2021). Fast hands-off control using ADMM real-time iterations. *IEEE Transactions on Automatic Control*, 67(10), 5416–5423.

Gallieri, M. and Maciejowski, J.M. (2012). ℓasso MPC: Smart regulation of over-actuated systems. In *2012 American Control Conference (ACC)*, 1217–1222.

Kothare, M.V., Balakrishnan, V., and Morari, M. (1996). Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32(10), 1361–1379.

Maciejowski, J.M. (2002). *Predictive control: with constraints*. Pearson education.

Mesbah, A. (2016). Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems Magazine*, 36(6), 30–44.

Mishra, P.K., Chatterjee, D., and Quevedo, D.E. (2018). Sparse and constrained stochastic predictive control for networked systems. *Automatica*, 87, 40–51.

Nagahara, M., Østergaard, J., and Quevedo, D.E. (2016). Discrete-time hands-off control by sparse optimization. *EURASIP Journal on Advances in Signal Processing*, 2016(1), 1–8.

Rawlings, J.B., Mayne, D.Q., and Diehl, M. (2017). *Model predictive control: theory, computation, and design*, volume 2. Nob Hill Publishing Madison, WI.

Tempo, R., Calafiore, G., and Dabbene, F. (2012). *Randomized algorithms for analysis and control of uncertain systems: with applications*. Springer.

Appendix A. PROOF OF LEMMA 3

For a given prediction horizon N, substituting the prediction dynamics (2) recursively, we have the evolution of system in a compact form

$$\bar{x} = \bar{A}_\delta x_{0|t} + \bar{B}_\delta \bar{u} + \bar{E}_\delta \bar{w},$$

where parametric matrices $\bar{A}_\delta$, $\bar{B}_\delta$ and $\bar{E}_\delta$ are given by

$$\bar{A}_\delta = \begin{bmatrix} A(\delta) \\ A^2(\delta) \\ \vdots \\ A^N(\delta) \end{bmatrix}, \ \bar{E}_\delta = \begin{bmatrix} E & 0 & \cdots & 0 \\ A(\delta)E & E & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}(\delta)E & A^{N-2}(\delta)E & \cdots & E \end{bmatrix},$$

$$\bar{B}_\delta = \begin{bmatrix} B(\delta) & 0 & \cdots & 0 \\ A(\delta)B(\delta) & B(\delta) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}(\delta)B(\delta) & A^{N-2}(\delta)B(\delta) & \cdots & B(\delta) \end{bmatrix}.$$

Thus, with

$$\bar{C} = \begin{bmatrix} C & 0 & \cdots & 0 \\ 0 & C & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C \\ 0 & 0 & \cdots & C_f \end{bmatrix}, \qquad \bar{c} = \begin{bmatrix} c \\ c \\ \vdots \\ c \\ c_f \end{bmatrix},$$

we have

$$h(\bar{u}, \theta) \doteq \max(\bar{C}(\bar{A}_\delta x_{0|t} + \bar{B}_\delta \bar{u} + \bar{E}_\delta \bar{w}) - \bar{c}),$$

which completes the proof of Lemma 3. Here $\max(v)$ for a vector v means that the greatest element of v.