An Uncertainty Model Validation Approach to Adaptive Robust Linear Control

Robert L. Kosut *
SC Solutions
3211 Scott Blvd., Santa Clara, CA 95054
kosut@scsolutions.com

Abstract The uncertainty model validation paradigm – a morepreciseterm is *unfalsification* – is reviewed. It is argued that unfalsification of an uncertainty model consisting of disturbance and dynamic uncertainty is the natural replacement for system identification when the intended use of the model is robust control. The result is that any finite data record will generate a family of uncertainty models, all of which are unfalsified, *i.e.*, each one could have reproduced the data. It is shown an "uncertainty" tradeoff curve can be computed quantifying the relation between the dynamic uncertainty and disturbance uncertainty of each unfalsified model. Hence, the family of unfalsified models can be used in an iterative approach to system identification and robust control design.

1 Introduction

Many attempts have been made to transfer the rules by which we learn to an automated procedure for improving the performance of our machines and systems, *e.g.*, the dual control concept [4], [1, Ch. 7], the windsurfer approach to adaptation and learning, [11, 12]. A survey of schemes based on iterative identification and control, sometimes referred to as *iterative adaptive control*, is given by Gevers in [6].

The work described here is a review of [8, 9] and a preview of [10] which address the problem of how uncertainty model unfalsification, as first described by Poolla *et al.*[17], could be used to replace the system identification step in iterative adaptive control. A variety of other methods for using unfalsification in iterative adaptation have been suggested, *e.g.*, [3], [13], [18].

Notation Let $\mathbf{S}^\ell$ denote the set of real sequences of length ℓ, *i.e.*, $x = \{x_1, \ldots, x_\ell\} \in \mathbf{S}^\ell$. The norm of $x \in \mathbf{S}^\ell$ is defined as $\|x\| = \left(\sum_{t=1}^{\ell} x_t^2\right)^{1/2}$. A subsequence is denoted by $x_{1:t} = \{x_1, \ldots, x_t\}$, and the norm of a subsequence is denoted by $\|x\|_{1:t} = \|x_{1:t}\|$. Let $\mathbf{S}$ denote the set of infinite sequences with finite norm.

*Research supported by DARPA, Applied Computation & Mathematics Program under AFOSR Contract No. F49620-93-C-0019.

2 Uncertainty Model Validation Problem

The generic uncertainty model validation problem is as follows:

Given scalar "data" sequences

$$e = \{e_1, \ldots, e_\ell\} \quad \textit{and} \quad u = \{u_1, \ldots, u_\ell\}, \tag{1}$$

establish necessary and sufficient conditions for the existence of a sequence

$$w = \{w_1, \ldots, w_\ell\} \in \mathbf{W}(\sigma) \tag{2}$$

and a causal system

$$\Delta \in \mathbf{\Delta}(\delta) \tag{3}$$

which are consistent with the data, i.e.,

$$e_t = w_t + (\Delta u)_t, \quad \forall t = 1 : \ell \tag{4}$$

The sets $\mathbf{W}(\sigma)$ and $\mathbf{\Delta}(\delta)$ denote, respectively, a set of sequences with norm bounded by σ and a set of causal systems with gain bounded by δ. The data sequence e is often obtained as the "error" between the output of a model of the system with input sequence u and the sensed output of the actual system, with the same input sequence u. The sensed output could be replaced by the simulation output of a high order or more complicated model, in which case, uncertainty model validation is a step in model reduction.

Observe that the error is modeled as consisting of the *sum* of a "noisy" sequence w plus a term depending on the input, Δu, which arises from modeling errors. Certainly other forms could be considered, this being the most basic.

2.1 Disturbance Uncertainty

There are many ways to characterize the disturbance set $\mathbf{W}(\sigma)$. For example:

- **Rms-bounded noise**

$$\mathbf{W}_{\text{rms}} = \left\{ w \in \mathbf{S}^{\ell} \;\middle|\; \frac{1}{\ell} \|w\|^2 \leq \sigma^2 \right\} \tag{5}$$

- **Time-domain white noise** [16]

$$\mathbf{W}_{\text{wht\_time}} = \left\{ w \in \mathbf{S}^{\ell} \mid |r_w(\tau)| \leq \gamma r_w(0) \right\} \tag{6}$$

where $r_w(\tau)$ is the auto-correlation of w,

$$r_w(\tau) = \frac{1}{\ell} \sum_{t=1}^{\ell-\tau} w_t w_{t+\tau}, \quad \forall \tau = 0 : m-1 \leq \ell \tag{7}$$

Observe that $r_w(0) = \|w\|^2 / \ell$.

- **Frequency-domain white noise** [15]

$$\mathbf{W}_{\text{wht\_freq}} = \left\{ w \in \mathbf{S}^{\ell} \mid |\lambda \{R_m(w)\} / \sigma^2 - 1| \leq \epsilon \right\} \tag{8}$$

where $\lambda\{\cdot\}$ denotes eigenvalues and

$$R_m(w) = \begin{bmatrix} r_w(0) & \cdots & r_w(m-1) \\ \vdots & \ddots & \vdots \\ r_w(m-1) & \cdots & r_w(0) \end{bmatrix} \in \mathbf{R}^{m \times m} \tag{9}$$

The disturbance set $\mathbf{W}_{\text{rms}}(\sigma)$ is the simplest of choices for deterministically characterizing "noise." The main advantage is that it is a convex set and therefor easy to handle in optimization. However, there are no restrictions preventing correlation with inputs and so the "worst-case" can occur. As shown above, characterizations of deterministic sets which resemble white noise have been examined in [15] in the frequency domain with application to system identification and in [16] for both time and frequency domains with application to robust control. The set $\mathbf{W}_{\text{wht\_time}}(\gamma, m)$ is essentially one of the standard white noise test where γ is chosen from χ^2 distribution tables; m is the *lag window* used to smooth the correlation function. The set $\mathbf{W}_{\text{wht\_freq}}(\sigma, m, \epsilon)$ is shown in [15] to also be useful for white noise testing; m again is the lag window, σ^2 is the rms-level of w and hence, the average level of the spectrum of w, and $\epsilon \in (0,1)$ determines the "flatness" of the spectrum. Clearly these latter sets do preserve the character of white noise, but they are not convex. However, they are no worse than quadratic and so may be quite amenable to cojugate-gradient methods of optimization.

2.2 Dynamic Uncertainty

Uncertain dynamics can also be characterized in a number of ways. Consider the following causal gain-bounded dynamic uncertainty sets:

- **Gain bounded, linear time invariant (LTI)**

$$\mathbf{\Delta}_{\text{LTI}}(\delta) = \{ \Delta \in \mathbf{LTI} \mid \|Gu\| \leq \delta \|u\|, \forall u \in \mathbf{S} \} \tag{10}$$

Since $G \in \mathbf{LTI}$, the gain bound condition is equivalent to

$$\|\Delta\|_\infty \leq \delta \tag{11}$$

- **Incrementally gain bounded, nonlinear, time-invariant (INTI)**

$$\mathbf{\Delta}_{\text{INTI}} = \{ \Delta \in \mathbf{TI} \mid \|\Delta u - \Delta v\| \leq \delta \|u - v\|, \forall u, v \in \mathbf{S} \} \tag{12}$$

- **Gain bounded, nonlinear, time-invariant (NTI)**

$$\mathbf{\Delta}_{\text{NTI}} = \{ \Delta \in \mathbf{TI} \mid \|\Delta u\| \leq \delta \|u\|, \forall u \in \mathbf{S} \} \tag{13}$$

There are clearly many variations one could include, as well as considering combinations and uncertainty structures descibed by the more inclusive linear fractional representation familiar in robust control design.

3 Uncertainty Model Unfalsification

In this section we use state the necessary and sufficient conditions for solving the complete uncertainty model validation problem (1)-(4) with disturbance uncertainty set $\mathbf{W}_{\text{rms}}$ and dynamic uncertainty sets $\mathbf{\Delta}_{\text{LTI}}$, $\mathbf{\Delta}_{\text{INTI}}$, and $\mathbf{\Delta}_{\text{NTI}}$. We also present the optimal uncertainty tradeoff curves for each of the uncertainty sets. The interested reader should refer to [8, 9, 10] and [17] for the proofs of the results.

3.1 Unfalsification Test

Given sequences

$$u = \{u_1, \ldots, u_\ell\}, \quad e = \{e_1, \ldots, e_\ell\} \tag{14}$$

there exists a sequence

$$w = \{w_1, \ldots, w_\ell\} \tag{15}$$

and a causal system Δ which are consistent with the data, that is,

$$e = w + \Delta u \tag{16}$$

with $w \in \mathbf{W}_{\mathrm{rms}}(\sigma)$ if and only if

$$\frac{1}{\ell}\|w\|^2 \leq \sigma^2 \tag{17}$$

and such that:

- $\Delta \in \mathbf{\Delta}_{\mathbf{LTI}}(\delta)$ *if and only if,*

  $$(\mathcal{E} - \mathcal{W})^T(\mathcal{E} - \mathcal{W}) - \delta^2\mathcal{U}^T\mathcal{U} \leq 0 \tag{18}$$

  with $(\mathcal{E},\mathcal{U},\mathcal{W})$ the $\ell \times \ell$ Toeplitz matrices formed from the sequences (e,u,w), respectively, e.g.,

  $$\mathcal{E} = \begin{bmatrix} e_1 & 0 & \cdots & 0 \\ e_2 & e_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ e_\ell & e_{\ell-1} & \cdots & e_1 \end{bmatrix}$$

- $\Delta \in \mathbf{\Delta}_{\mathbf{INTI}}(\delta)$ *if and only if, $\forall m{-}n \neq 0 : \ell$ and $\forall t = 1 : \ell$,*

  $$\|(z^n - z^m)(e - w)\|_{1:t} \leq \delta\,\|(z^n - z^m)u\|_{1:t} \tag{19}$$

  where z^k is the k-forward shift operator, i.e., if $x = \{x_1, x_2, \ldots\}$ then $z^k x = \{0, \ldots, 0, x_1, x_2, \ldots\}$ with k-zeros.

- $\Delta \in \mathbf{\Delta}_{\mathbf{NTI}}(\delta)$ *if and only if $\forall t = 1 : \ell$,*

  $$\|y\|_{1:t} \leq \delta\,\|u\|_{1:t}\,, \tag{20}$$

3.2 Uncertainty Tradeoff

As shown in [8, 9, 10], a tradeoff between disturbance and system uncertainty is obtained by solving the following optimization problem.

Given data (e,u), fix δ and perform the optimization:

$$\hat{\sigma}(\delta) := \min_{w,\sigma} \sigma \quad \textit{subject to} \quad \begin{cases} w \in \mathbf{W}(\sigma) \\ \Delta \in \mathbf{\Delta}(\delta) \\ e = w + \Delta u \end{cases} \tag{21}$$

The graph of $\hat{\sigma}(\delta)$ versus δ, referred to as the *uncertainty tradeoff curve*, establishes a tradeoff between the model uncertainty bound, δ, and the minimum possible corresponding disturbance uncertainty bound $\hat{\sigma}(\delta)$. Every point on the curve depends on a different choice of the uncertainty pair (w, Δ). The shape of the curve depends on the choice of the uncertainty sets $\mathbf{W}(\sigma), \mathbf{\Delta}(\delta)$. We can examine $\mathbf{W}_{\mathrm{rms}}(\sigma)$ together with any one of $\mathbf{\Delta}_{\mathbf{LTI}}(\delta)$,$\mathbf{\Delta}_{\mathbf{INTI}}(\delta)$, or $\mathbf{\Delta}_{\mathbf{NTI}}(\delta)$, thus leading to three tradeoff curves: $\hat{\sigma}_{\mathbf{LTI}}(\delta)$, $\hat{\sigma}_{\mathbf{INTI}}(\delta)$, and $\hat{\sigma}_{\mathbf{NTI}}(\delta)$. Since the uncertainty sets are convex, it follows that (21) is a convex optimization, and hence, all the tradeoff curves are convex functions. In addition, as shown in [10] they are nested, i.e.,

$$\hat{\sigma}_{\mathbf{NTI}}(\delta) < \hat{\sigma}_{\mathbf{INTI}}(\delta) < \hat{\sigma}_{\mathbf{LTI}}(\delta), \quad \forall\delta > 0 \tag{22}$$

The nesting occurs because the extremes of the convex functions are similarly ordered. At $\delta = 0$, the optimal w is equal to e, and hence, all three minimum rms levels are identically zero, *i.e.*, at $\delta = 0$,

$$\hat{\sigma}_{\mathbf{LTI}}(0) = \hat{\sigma}_{\mathbf{INTI}}(0) = \hat{\sigma}_{\mathbf{NTI}}(0) = \|e\|_{\mathrm{rms}} \tag{23}$$

At the other extreme when the rms level is zero ($w = 0$), the corresponding uncertainty bounds, denoted by $\delta_{\mathbf{NTI}}, \delta_{\mathbf{INTI}}, \delta_{\mathbf{LTI}}$, are the minimum possible to satisfy the constraints. But as shown in [10] these are ordered as follows:

$$\delta_{\mathbf{NTI}} < \delta_{\mathbf{INTI}} < \delta_{\mathbf{LTI}} \tag{24}$$

This follows from the fact that the condition for which $\Delta \in \mathbf{\Delta}_{\mathbf{LTI}}$ is more restrictive than for $\Delta \in \mathbf{\Delta}_{\mathbf{INTI}}$, which is more restrictive than for $\Delta \in \mathbf{\Delta}_{\mathbf{NTI}}$.

4 Unfalsification with Parametric Models

To apply the unfalsification approach to models that include unknown parameters consider the single-actuator, single-sensor *prediction error (PE) uncertainty model* [8, 9]:

$$y = G(\theta)u + H(\theta)(w + \Delta u) \quad \begin{cases} \theta \in \mathbf{\Theta} \\ w \in \mathbf{W}(\sigma) \\ \Delta \in \mathbf{\Delta}(\delta) \end{cases} \tag{25}$$

where y and u are, respectively, the observed output and input sequences, $G(\theta)$ and $H(\theta)$ are causal, linear-time-invariant systems, initially at rest, each dependent on a parameter vector $\theta \in \mathbf{\Theta}$. The prediction error associated with the above uncertainty model is,

$$e(\theta) := H(\theta)^{-1}(y - G(\theta)u) \tag{26}$$

and clearly decomposes into:

$$e(th) = w + \Delta u \tag{27}$$

The set $\mathbf{\Theta}$ is a subset of

$$\mathbf{\Theta}_{\mathrm{stab}} = \left\{\theta \in \mathbf{R}^p \;\middle|\; H(\theta)^{-1} \text{ and } H(\theta)^{-1}G(\theta) \text{ are stable}\right\} \tag{28}$$

Parameters in $\mathbf{\Theta}_{\text{stab}}$ insure that the *predictor* associated with (25) is stable [14].

Observe that the PE uncertainty model (25) is characterized by *three* types of parameters, (θ, σ, δ), *i.e.*, uncertainty arises from transfer function parameters, disturbance, and dynamics. In contrast, the standard PE model is characterized by two types of parameters, (θ, σ), *i.e.*, uncertainty is due to transfer function parameters and disturbance only.

The form of the model implies that the dominant plant dynamics are well approximated by the LTI system $G(\theta), H(\theta)$. If $\mathbf{\Delta} = \mathbf{\Delta}_{\text{LTI}}$, then the model implies that the system is LTI but uncertain. This may be a reasonable assumption in some cases, *e.g.*, flexible systems undergoing small dispalcements. But in many circumstances the model error is due to inherent nonlinearities. For example, consider the case where the *center* of the uncertainty model, $(G(\theta), H(\theta))$, is LTI, but dynamic uncertainty is possibly nonlinear.

For robust control design, (θ, σ, δ) are known. In some cases θ can also be uncertain,*i.e.*, $\theta \in \mathbf{\Theta}$. This is important particularly if θ represents uncertain physical parameters. In most cases of system identification, θ is used to encode the LTI system $(G(\theta), H(\theta))$, the center of the uncertainty set.

Classical system identification poses an *optimization problem* in (θ, σ), and does not deal with the dynamic uncertainty set $\mathbf{\Delta}(\delta)$ which is of critical importance for robust control. In contrast, a model is said to be *validated* if and only if it could have produced the data. Validation is perhaps a misnomer, as one can never *prove* that a model will be able to accurately predict the future. More precisely, the data can *falsify* a model, *i.e.*, the model may prove to be incapable of fully explaining the data. Hence, instead of validation, as already discussed, we use the more precise, but awkward term: *unfalsification.* Clearly unfalsification is a *feasibility problem* – find a the model set whose members are consistent with the data. As shown in [8, 17], and reviewed in the previous sections, this philosophical shift allows the dynamic uncertainty bound δ to be estimated (unfalsified) along with θ and σ.

Using the specific uncertainty sets $\mathbf{W}_{\text{rms}}(\sigma)$ and $\mathbf{\Delta}_{\text{NTI}}(\delta)$, the uncertainty model (25) is unfalsified by the ℓ-point data sequences (y, u) – or equivalently by the data sequences $(e(\theta), u)$ – if and only if there exists $\theta \in \mathbf{\Theta}$ and an ℓ-point sequence w, such that:

$$\begin{aligned} \|w\| &\leq \sigma\sqrt{\ell} \\ \|e(\theta) - w\|_t &\leq \delta \|u\|_t, \quad \forall t \in [1:\ell] \end{aligned} \tag{29}$$

Uncertainty models in the unfalsified (feasibility) set satisfying the above inequalities may not be unique. In addition, there is no easy parametrization, and attempts to discretize the space may lead to huge dimensions even for a few parameters. For this reason we restrict attention to a more easily paramtrized set of unfalsified uncertainty models as described next.

As shown in [8, 9], an uncertainty tradeoff between dynamic and disturbance uncertainty is obtained by solving the following optimization problem:

Fix δ and perform the optimization:

$$\hat{\sigma}(\delta) := \min_{\theta, \sigma, w} \ \sigma, \quad \textit{subject to (29)} \tag{30}$$

The minimizing values of θ and w are given by:

$$\hat{\sigma}(\delta) = \max\left\{ \max_{\tau \in [1:\ell]} \left(\left\| e(\hat{\theta}(\tau)) \right\|_\tau - \delta \|u\|_\tau \right), (\theta) \right\} \tag{31}$$

$$\theta(\delta) = \hat{\theta}(\tau) \Big|_{\tau = \tau(\delta)} \tag{32}$$

where

$$\hat{\theta}(\tau) = \arg\min_\theta \|e(\theta)\|_\tau, \quad \tau \in [1:\ell] \tag{33}$$

$$\tau(\delta) = \arg \max_{\tau \in [1:\ell]} \left(\left\| e(\hat{\theta}(\tau)) \right\|_\tau - \delta \|u\|_\tau \right) \tag{34}$$

Remarks

1. The graph of $\hat{\sigma}(\delta)$ versus δ, referred to as the *uncertainty tradeoff curve*, establishes the tradeoff between model uncertainty, δ, and the minimum possible corresponding disturbance uncertainty $\hat{\sigma}(\delta)$.

2. To every point on the tradeoff curve there is a different set of nominal transfer functions $(G(\theta), H(\theta))$ because $\theta \equiv \theta(\delta)$.

3. The tradeoff curve separates the unfalsified and falsified uncertainty models based on the current data.

4. The model uncertainty bound, δ, can range between the two extremes: when the prediction error is due only to the disturbance ($\delta = 0$) and when it is due only to the dynamic uncertainty ($\sigma = 0$).

5. For the special case when $\delta = 0$, the corresponding uncertainty model on the tradeoff curve, $\hat{\theta}(0), \hat{\sigma}(0)$, is precisely the usual least-squares prediction error transfer function estimate.

6. In general, the tradeoff curve is not convex because $e(\theta)$ is not affine in θ except in the case where $(G(\theta), H(\theta))$ are paramtrized via an ARX model.

7. Notice that no mention has been made about the "true" system which generated the data. All that is claimed is that there exists a set of *possibly nonlinear uncertainty models*, centered at an LTI model, each of which *could* have produced the data.

References

[1] K. J. Astrom and B. Wittenmark, *Adaptive Control,* Addison-Wesley, 1995.

[2] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory,* SIAM Studies in Applied Mathematics, 1994.

[3] M.A. Dahleh and J.C. Doyle, "From data to control", *Proc. Workshop on Modeling of Uncertainty in Control Systems*, Springer-Verlag, 1992.

[4] A. A. Feldbaum, *Optimal Control Theory,* New York, Academic Press.

[5] C. Foias and A. E. Frazho, *The Commutant Lifting Approach to Interpolation Problems*, Birkhauser-Verlag, 1990.

[6] M. Gevers, "Towards a joint design of identification and control," *Essays on Control: Perspectives in the Theory and its Applications*, editors: H.L. Trentelman and J.C. Willems, pp. 111-151, Birkhauser, Boston, MA, 1993.

[7] G.C. Goodwin, M. Gevers, and B. Ninness, "Quantifying the error in estimated transfer functions with applications to model order selection," *IEEE Trans. on Automatic Control*, vol. 37, no.7, pp. 913-928, July 1992.

[8] R.L. Kosut, "Uncertainty model unfalsification: a system identification paradigm compatible with robust control design," *Proc. 1995 CDC*, New Orleans, LA, Dec. 1995.

[9] R. L. Kosut, "Iterative adaptive robust control via uncertainty model unfalsification," *Proc. 1996 IFAC World Congress*, June, 1996, San Francisco, CA

[10] R. L. Kosut and B. D. O. Anderson, "Uncertainty model unfalsification," submitted, 1997 CDC.

[11] W.S. Lee, B.D.O. Anderson, R.L. Kosut, and I.M.Y. Mareels, "A new approach to adaptive robust control," *Int. J. of Adaptive Control and Signal Processing*, vol. 7, pp. 183-211, 1993.

[12] W.S. Lee, B.D.O. Anderson, I.M.Y. Mareels, and R.L. Kosut, "On some key issues in the windsurfer approach to robust adaptive control", *Automatica*, Vol. 31, No. 11, pp. 1619-1636, 1995.

[13] M.M. Livestone, M.A. Dahleh, and J.A. Farrell, "A framework for robust control based model invalidation", *Proc. 1994 ACC*, pp. 3017-3020, Baltimore, MD, June 1994.

[14] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, 1987.

[15] M. Massoumnia and R. L. Kosut, "A family of norms for system identification problems," *Proc. 1993 ACC*, San Francisco, CA, June 1993, and *IEEE Trans. Aut. Control*, vol. 39, no. 5, pp. 1027-1031, May 1994.

[16] F. Paganini, "A set-based approach for white noise modeling," *IEEE Trans. Aut. Contr.*, vol. 41, no. 10, pp. 1453-1465, Oct. 1996.

[17] K. Poolla, P. Khargonnekar, A. Tikku, J. Krause, and K. Nagpal, "A time-domain approach to model validation", *Proc. 1992 ACC*, Chicago, IL, June 1992, and *IEEE Trans. Aut. Contr.*, vol. 39, no. 5, pp. 951-959, May 1994.

[18] M. G. Safonov and T. C. Tsao, "The unfalsified control concept: a direct path from experiment to controller", presented at *Conference on Feedback Control, Nonlinear Systems, and Complexity*, McGill University, Montreal,Canada, May 6-7, 1994.