Uncertainty Model Unfalsification

Robert L. Kosut 1
SC Solutions
3211 Scott Blvd.
Santa Clara, CA 95054, USA
kosut@scsolutions.com

Brian D. O. Anderson 2
Australian National University
Canberra, A.C.T., Australia
brian.anderson@anu.edu.au

Abstract

The main contributions presented here are (i) to widen the classes of model sets for which necessary and sufficient conditions for uncertainty model unfalsification can be obtained, and (ii) to display the effect of different assumptions concerning the modeling error on the curves defining the boundary of unfalsified models (the optimal uncertainty tradeoff curve) for the same underlying data set.

1 Introduction

The work described here is a further extension of the uncertainty model unfalsification paradigm as first described by Poolla et al.[10]. Unfalsification has been proposed as a replacement for the system identification step in iterative adaptive control [2], [3], [6]. The paradigm has also been proposed for direct controller unfalsification [11, 12]. Unfalsification helps quantify the connections between adaptation and learning as espoused in the dual control concept [1, Ch. 7], and more recently in the windsurfer approach to adaptation [5].

In this paper we show how different assumptions on the model error type affect the unfalsification test and associated uncertainty tradeoff curves. We specifically examine the effect of assuming that the unknown but bounded dynamic uncertainty (model error) is linear-time-invariant, incrementally nonlinear time-invariant, or nonlinear time-invariant. In all cases the disturbance is assumed to be unknown but RMS bounded.

1 The author wishes to acknowledge the support of the Australian National University and DARPA, Applied Computation & Mathematics Program under AFOSR Contract No. F49620-93-C-0019.

2 Research School of Information Sciences and Engineering and Cooperative Research Centre for Robust and Adaptive Systems, Australian National University, ACT 0200, Australia. The author wishes to acknowledge the funding of the activities of the Cooperative Research Centre for Robust and Adaptive Systems by the Australian Commonwealth Government under the Cooperative Research Centres Program.

Notation

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

The systems considered are scalar, sampled-data systems which are causal, map $S^f$ into itself, and are initially at rest, i.e., they have no memory of input sequences for $t < 1$. For the system $G$, the value of the mapping $Gu$ at time $t$ is denoted by $(Gu)_t$ and a subsequence of the mapping by $(Gu)_{\ell,t}$.

2 Uncertainty Model Unfalsification Problem

The generic uncertainty model unfalsification problem is as follows:

Given scalar data sequences $e, v \in S^f$, establish necessary and sufficient conditions for the existence of a disturbance sequence $w \in S^f$ and a system $\Delta$ such that

$$w \in W(\sigma), \quad \Delta \in \Delta(\delta)$$

and which are consistent with the data error model, i.e.,

$$e_t = w_t + (\Delta v)_t, \quad \forall t = 1 : \ell$$

The sets $W(\sigma)$ and $\Delta(\delta)$ denote, respectively, a set of sequences with norm bounded by $\sigma$ and a set of systems with gain bounded by $\delta$. The data sequence $e$ is often obtained as the prediction error associated with an assumed model of the system, and $v$ is a function of other sensed signals, the choice reflecting the type of dynamic uncertainty, or model error. For example, using the standard form in [7],

$$e = H^{-1}(y - Gu)$$

p. 1
where \((G, H)\) are linear-time-invariant systems, \(u\) is the control input, and \(y\) is the sensed output. If \(v = u\), then \(\Delta\) represents additive model error, i.e., the uncertainty model has the form,

\[
y = Gu + H\Delta u + Hw
\]

If \(v = Gu\), then \(\Delta\) represents multiplicative-model error, i.e., the uncertainty model the form,

\[
y = Gu + H\Delta Gu + Hw
\]

There are clearly many variations one could include, e.g., combinations of additive and multiplicative model errors, co-prime factor uncertainty, and so on, ultimately leading to the uncertainty structures described by the more inclusive linear fractional representation familiar in robust control design. In addition, as shown in section 4.1, the error could be obtained from a parametric prediction error model with parameters associated with transfer function coefficients which characterize the input/output and disturbance dynamics.

### 2.1 Disturbance Uncertainty

There are many ways to characterize the disturbance set \(W(\sigma)\). For example, consider the following sets of finite sequences:

- **Rms-bounded noise**
  \[
  W_{\text{rms}} = \left\{ w \in \mathbb{S}^f \mid \frac{1}{\ell} \|w\|^2 \leq \sigma^2 \right\}
  \]

- **Time-domain white noise** [9]
  \[
  W_{\text{wh.t.time}} = \left\{ w \in \mathbb{S}^f \mid |r_w(\tau)| \leq \gamma r_w(0) \right\}
  \]
  where \(r_w(\tau)\) is the auto-correlation of \(w\),

  \[
  r_w(\tau) = \frac{1}{\ell} \sum_{t=1}^{\ell} w_tw_{t+\tau}, \quad \forall \tau = 0 : m - 1 \leq \ell
  \]
  Observe that \(r_w(0) = \|w\|^2 / \ell\).

- **Frequency-domain white noise** [8]
  \[
  W_{\text{wh.freq}} = \left\{ w \in \mathbb{S}^f \mid \lambda \{R_m(w)\} / \sigma^2 - 1 \leq \epsilon \right\}
  \]
  where \(\lambda\{\cdot\}\) denotes eigenvalues and

  \[
  R_m(w) = \begin{bmatrix}
          r_w(0) & \cdots & r_w(m - 1) \\
          \vdots & \ddots & \vdots \\
          r_w(m - 1) & \cdots & r_w(0)
        \end{bmatrix}
  \]

  The disturbance set \(W_{\text{rms}}(\sigma)\) 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 [8] in the frequency domain with application to system identification and in [9] for both time and frequency domains with application to robust control. The set \(W_{\text{wh.t.time}}(\gamma, m)\) is essentially one of the standard white noise test where \(\gamma\) is chosen from \(\chi^2\) distribution tables; \(m\) is the lag window used to smooth the correlation function. The set \(W_{\text{wh.freq}}(\sigma, m, \epsilon)\) is shown in [8] to also be useful for white noise testing; \(m\) again is the lag window, \(\sigma^2\) 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 conjugate-gradient methods of optimization.

### 2.2 Gain-Bounded Dynamic Uncertainty

Uncertain dynamics can also be characterized in a number of ways. Consider the following gain-bounded, time-invariant (TI) dynamic uncertainty sets:

- **Linear (LTI)**
  \[
  \Delta_{\text{LTI}}(\delta) = \left\{ \Delta \in \text{LTI} \mid \sup_{v \in S} \frac{\|\Delta v\|}{\|v\|} \leq \delta \right\}
  \]

  Since \(\Delta \in \text{LTI}\), the gain bound condition is equivalent to the frequency domain bound:

  \[
  |\Delta\{e^{j\omega}\}| \leq \delta, \quad \forall \omega \in [-\pi, \pi]
  \]

- **Incrementally nonlinear (INTI)**
  \[
  \Delta_{\text{INTI}} = \left\{ \Delta \in \text{TI} \mid \sup_{v_1, v_2 \in S} \frac{\|\Delta v_1 - \Delta v_2\|}{\|v_1 - v_2\|} \leq \delta \right\}
  \]

- **Nonlinear (NTI)**
  \[
  \Delta_{\text{NTI}} = \left\{ \Delta \in \text{TI} \mid \sup_{v \in S} \frac{\|\Delta v\|}{\|v\|} \leq \delta \right\}
  \]

### 3 Uncertainty Model Unfalsification

In this section we state the necessary and sufficient conditions for solving the complete uncertainty model unfalsification problem (section 2) with disturbance uncertainty set \(W_{\text{rms}}\) and dynamic uncertainty sets \(\Delta_{\text{LTI}}, \Delta_{\text{INTI}}, \text{and } \Delta_{\text{NTI}}\). We also show how to obtain the optimal uncertainty tradeoff curves corresponding to each of the uncertainty sets.
3.1 Unfalsification Test

Given data sequences $e, v \in S^t$, there exists a sequence $w \in S^t$ and a system $A$ which are consistent with the data, that is,

$$e_t = w_t + (\Delta v)_t, \quad \forall t = 1 : \ell$$

(13)

with $w \in W_{\text{ms}}(\sigma)$ if and only if

$$\frac{1}{\ell} \|w\|_2^2 \leq \sigma^2$$

(14)

and such that:

- $\Delta \in \Delta_{\text{LTI}}(\delta)$ if and only if,

$$(E - V)(e - V) - \delta^2 V^TY \leq 0$$

(15)

with $(E, V, W)$ the $\ell \times \ell$ Toeplitz matrices formed from the sequences $(e, v, w)$, respectively, e.g.,

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

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

$$\|(e^n - w^n)(e - w)\|_1 : \ell \leq \delta \|(e^n - w^n)v\|_1 : \ell$$

(16)

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

- $\Delta \in \Delta_{\text{INT}}(\delta)$ if and only if $\forall t = 1 : \ell$,

$$\|\varepsilon - w\|_1 : \ell \leq \delta \|v\|_1 : \ell$$

(17)

The result for $\Delta \in \Delta_{\text{LTI}}(\delta)$ and the necessity for $\Delta \in \Delta_{\text{INT}}(\delta)$ (which is the same as the necessary and sufficient conditions for gain-bounded linear-time-varying (LTV) systems) is found in [10]. Proof of the remaining results can be found in [4].

3.2 Inclusion Property

Conditions (15)-(17) satisfy an inclusion, or nesting, property. To see this, define the matrix:

$$\mathcal{R}(w, \delta) = (E - V)^T(E - V) - \delta^2 V^TY$$

(18)

It is shown in [4] that conditions (18)-(17) are equivalent, respectively, to the following:

- $\Delta \in \Delta_{\text{LTI}}(\delta)$ if and only if,

$$\alpha^T \mathcal{R}(w, \delta) \alpha \leq 0, \quad \forall \alpha \in \mathbb{R}^\ell$$

(19)

- $\Delta \in \Delta_{\text{INTI}}(\delta)$ if and only if,

$$\alpha^T \mathcal{R}(w, \delta) \alpha \leq 0, \quad \forall \alpha \in \mathbb{A}_{\text{INTI}}$$

(20)

- $\Delta \in \Delta_{\text{INT}}(\delta)$ if and only if,

$$\alpha^T \mathcal{R}(w, \delta) \alpha \leq 0, \quad \forall \alpha \in \mathbb{A}_{\text{INT}}$$

(21)

where $\mathbb{A}_{\text{INTI}}$ and $\mathbb{A}_{\text{INT}}$ are defined as follows:

$$\mathbb{A}_{\text{INTI}} = \{\alpha_i \in \mathbb{R}^\ell \mid i = 1 : \ell \}$$

(22)

with:

$$\alpha_i^T = [1, 0, 0, \ldots, 0]$$

(23)

and

$$\mathbb{A}_{\text{INT}} = \{\alpha_i \in \mathbb{R}^\ell \mid i = 1 : \ell + 1/2\}$$

(22)

with:

$$\alpha_i^T = [1, 0, 0, \ldots, 0]$$

(23)

The $\ell(\ell + 1)/2$ elements of $\mathbb{A}_{\text{INTI}}$ enumerate all the unique time shifts and subsequence sums possible which satisfy the time-invariance and incremental gain bound condition (11). These of course include the case of no shifts and all subsequence sums which are enumerated as the $\ell$ elements of $\mathbb{A}_{\text{INTI}}$. Hence, for a given $\delta$ and data set, (11) is more restrictive than (12), i.e., $\mathbb{A}_{\text{INTI}} \subseteq \mathbb{A}_{\text{INT}}$. Because $\mathbb{A}_{\text{INTI}} \subseteq \mathbb{R}^\ell$, it follows that $\Delta \in \Delta_{\text{LTI}}(\delta)$ (9) is the most restrictive of the three model error sets. Specifically, (19) insures that all of the eigenvalues of $\mathcal{R}(w, \delta)$ are negative, whereas (20)-(21) allow some of the eigenvalues to be positive.
model uncertainty bound, $\delta$, and the minimum possible corresponding disturbance uncertainty bound $\bar{\delta}(\delta)$. Every point on the curve depends on a different choice of the uncertainty pair $(w, \Delta)$. The tradeoff curve separates the unfalsified and falsified uncertainty models based on the current data. The shape of the curve depends on the choice of the uncertainty sets $W_\text{rs}(\sigma), \Delta(\delta)$.

We can examine $W_\text{rs}(\sigma)$ together with any one of $\Delta_{\text{LTI}}(\delta), \Delta_{\text{INTI}}(\delta), \Delta_{\text{ANTI}}(\delta)$, thus leading to three tradeoff curves: $\delta_{\text{LTI}}(\delta)$, $\delta_{\text{INTI}}(\delta)$, and $\delta_{\text{ANTI}}(\delta)$. Since the uncertainty sets are convex, it follows that (32) is a convex optimization, and hence, all the tradeoff curves are convex functions. In addition, as shown in (4) they are nested, i.e.,

$$\delta_{\text{INTI}}(\delta) < \delta_{\text{ANTI}}(\delta) < \delta_{\text{LTI}}(\delta), \quad \forall \delta > 0 \tag{33}$$

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 identical, i.e., at $\delta = 0$,

$$\delta_{\text{LTI}}(0) = \delta_{\text{INTI}}(0) = \delta_{\text{ANTI}}(0) = \|e\|_{\text{rms}} \tag{34}$$

At the other extreme when the rms level is zero ($w = 0$), the corresponding uncertainty bounds satisfy (28).

4 Unfalsification with Parametric Models

4.1 Parametric Uncertainty Model

To unfalsify models that include unknown parameters, consider the single-actuator, single-sensor prediction error (PE) uncertainty model [2, 3]:

$$y = G(\theta)u + H(\theta)(w + \Delta u) \quad \begin{cases} \theta \in \Theta \\ w \in W(\sigma) \\ \Delta \in \Delta(\delta) \end{cases} \tag{35}$$

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

$$e(\theta) \equiv H(\theta)^{-1}(y - G(\theta)u) \tag{36}$$

which decomposes into,

$$e(\theta) = w + \Delta u \tag{37}$$

The set $\Theta$ is a subset of

$$\Theta_{\text{stab}} = \{ \theta \in \mathbb{R}^p \mid H(\theta)^{-1} \text{ and } H(\theta)^{-1}G(\theta) \text{ are stable} \} \tag{38}$$

Parameters in $\Theta_{\text{stab}}$ insure that the predictor associated with (35) is stable [7]. These assumptions imply that the dominant plant dynamics are well approximated by the LTI system $G(\theta), H(\theta)$ for some value of $\theta$. If $\Delta = \Delta_{\text{LTI}}$, then the model represents the belief that the true system is dominantly LTI, but uncertain. This may be a reasonable assumption in some cases, e.g., flexible systems undergoing small displacements. In other circumstances the model error is due to inherent nonlinearities.

Observe that the PE uncertainty model (35) is characterized by three types of parameters, $(\theta, \sigma, \delta)$, i.e., uncertainty arises from transfer function parameters, disturbance, and dynamics. In contrast, the standard PE model is characterized by two types of parameters, $(\theta, \sigma)$, i.e., uncertainty is due to transfer function parameters and disturbance only. Furthermore, classical system identification poses an optimization problem in $(\theta, \sigma)$, and does not deal with the dynamic uncertainty set $\Delta(\delta)$ which is of critical importance for robust control. In contrast, unfalsification is a feasibility problem — find a model set whose members are consistent with the data. This philosophical shift allows the dynamic uncertainty bound $\delta$ to be estimated (unfalsified) along with $\theta$ and $\sigma$.

4.2 Unfalsification Test

Given data sequences $y, u, v \in S^f$, there exists a sequence $w \in S^f$, a system $\Delta$, and parameter $\theta$ which are consistent with the data, that is, the prediction error satisfies

$$e(\theta)_t = (H(\theta)^{-1})v_t - (H(\theta)^{-1}G(\theta))u_t \tag{39}$$

with $\theta \in \Theta$ and $w \in W(\sigma)$ if and only if

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

and such that:

- $\Delta \in \Delta_{\text{LTI}}(\delta)$ if and only if,
  $$\alpha^T \mathbb{R}(w, \delta, \theta) \alpha \leq \theta, \quad \forall \alpha \in \mathbb{R}^\ell \tag{40}$$

- $\Delta \in \Delta_{\text{INTI}}(\delta)$ if and only if,
  $$\alpha^T \mathbb{R}(w, \delta, \theta) \alpha \leq \theta, \quad \forall \alpha \in \mathbb{A}_{\text{INTI}} \tag{41}$$

- $\Delta \in \Delta_{\text{ANTI}}(\delta)$ if and only if,
  $$\alpha^T \mathbb{R}(w, \delta, \theta) \alpha \leq \theta, \quad \forall \alpha \in \mathbb{A}_{\text{ANTI}} \tag{42}$$

with $\mathbb{A}_{\text{INTI}}$ and $\mathbb{A}_{\text{ANTI}}$ as defined in (22)-(23) and where

$$\mathbb{R}(w, \delta, \theta) = (\mathbb{E}(\theta) - W)^T(\mathbb{E}(\theta) - W) - \delta^2 V^T V$$

and $\mathbb{E}(\theta)$ is the Toeplitz matrix formed from the prediction error sequence $e(\theta) \in S^f$. 

p. 5
4.3 Uncertainty Tradeoff

The optimal uncertainty tradeoff is now obtained by solving the following optimization problem:

\[
\begin{align*}
\text{Fix } \delta \text{ and perform the optimization:} \\
\min_{w, \sigma, \theta} & \quad \sigma \\
\text{subject to } & \quad (39) \text{ and either } (40), (41), \text{ or } (42), \text{ depending on } \theta(\delta) \text{ or } \theta(\delta) \text{ respectively.}
\end{align*}
\]

The graph of \( \delta(\theta) \) versus \( \delta \) establishes the tradeoff between model uncertainty, \( \delta \), and the minimum possible corresponding RMS disturbance uncertainty bound \( \sigma \).

To every point on the tradeoff curve there is a different parameter \( \theta(\delta) \), and hence, a different set of nominal transfer functions, \( G(\theta(\delta)) \), \( H(\theta(\delta)) \).

In general, the tradeoff curve is not convex because \( \sigma(\theta) \) is not affine in \( \theta \) and \( v \) may also depend on \( \theta \), e.g., \( v = G(\theta)u \). The curve is convex, and may be found via convex optimization, when \( v = u \) (additive model error) and \( (G(\theta), H(\theta)) \) are parametrized via an ARX model, i.e.,

\[
\begin{align*}
H(\theta)^{-1} &= 1 + a_1 z^{-1} + \cdots + a_n z^{-n} \\
H(\theta)^{-1}G(\theta) &= b_1 z^{-1} + \cdots + b_m z^{-m}
\end{align*}
\]

where \( \theta^T = [a_1 \ldots a_n, b_1 \ldots b_m] \) and \( z^{-k} \) is the k-delay operator.

5 Iterative Adaptive Control

The goal of uncertainty model unfalsification is to replace the identification step in an iterative adaptive control design. In [2, 3] it is suggested that uncertainty models along the tradeoff curve be used for robust control design in a cautious manner by first selecting the model with the largest value of dynamic uncertainty \( \delta \). Next design a robust controller to achieve the best possible performance with the uncertainty model. Now apply the controller to the actual system. If the resulting closed-loop system performs no worse than predicted by the uncertainty model, then pick another model along the curve with a smaller \( \delta \). Continue decreasing \( \delta \) until performance is worse than predicted. When this occurs, new data is available to re-compute the tradeoff curve. This iterative design approach is similar to that proposed in [8]. The process can be made more efficient by following the procedure presented in a series of papers by Saifunov et al. (see [11, 12] and the references therein), where it is shown how to falsify a candidate controller before it is implemented.

6 Concluding Remarks

We have shown that a number of tradeoff curves can be generated, each consistent with the data, but based on different assumptions about the dynamic uncertainty. Questions remain regarding how to use all these curves in the iterative design procedure outlined above.

References

June 10, 1998

Dear Colleague:

The members of the Operating Committee of the 1997 CDC take this opportunity to thank you for attending the Conference in San Diego last December. Your participation made the Conference a great success continuing the tradition established by past CDCs.

In the midst of an otherwise very pleasant experience is the issue regarding the production of the CD-ROM Proceedings. This greatly disappointing situation has lingered on for too long and has dampened our spirits. Despite valiant efforts the CD-ROM Producer failed to deliver the CD-ROM at the Conference as promised and the last six months have been a long-drawn-out "battle" for its production and delivery. Even though the circumstances surrounding the production of the CD-ROM were beyond our control it was our responsibility to make it available in a timely manner.

We are all very happy and relieved that this lengthy and often contentious process is now over! All of us on the Operating Committee apologize for this delay and the disruption it has caused. Additional information about the CD-ROM can be obtained at: http://www.odu.edu/~ieecss/1997CDCPubs/.

Sincerely,
Operating Committee
1997 CDC