Rational Basis Functions for Robust Identification

from Frequency and Time Domain Measurements

Hüseyin Akçay

Department of Electrical and Electronics Engineering
Anadolu University, Eskişehir, Turkey

October 19, 2008

Outline

The estimation of dynamic models on the basis of observed input–output measurements:

- Stochastic modeling of disturbances and representation of errors as averages over ensembles of possible noise realisations (Ljung:1999).
- Characterization of disturbances and also estimation errors according to a deterministic model under which the worst–case amplitude is quantified (Makila *etal.*:1995) –robust or worst-case identification–
  - Significant advantages. For example, errors due to non–linearities are easily accommodated, and the resultant models and error bounds are of a form suitable for subsequent robust control design.

System classes

Let the discrete-time system with impulse-response sequence $g(k)$ be represented via an associated power series

$$G(z) = \sum_{k=0}^{\infty} g(k) z^k. \tag{1}$$

- The system is stable if G is analytic on $D_1 = D$ where $D_R = \{z \in \mathbf{C} : |z| < R\}$. In addition, if $G(z)$ is continuous on the boundary of D denoted by T, then it is an element of the disc algebra $\mathcal{A}(D)$. If $|G(z)|^p$ is integrable on T then $G(z)$ is an element of the Hardy space $\mathcal{H}_p(D)$.
- The system can also be characterised according to the space ℓ_p that $g(k)$ lives in. Common choices are ℓ_1 in which $\sum_k |g(k)| < \infty$ and ℓ_∞ in which $\sup_k |g(k)| < \infty$.

Linear–in–parameters model structures

$$\widehat{G}_N(z,\theta) = \sum_{k=0}^{n-1} \theta_k \mathcal{B}_k(z)$$

where $\mathcal{B}_k(z)$ are rational basis functions with prescribed poles.

- Laguerre basis functions

$$\mathcal{B}_k(z) \triangleq \frac{\sqrt{1-a^2}}{1-az}\left(\frac{z-a}{1-az}\right)^k, \quad k = 0, 1, \cdots$$

  where $-1 < a < 1$. By choosing a according to prior knowledge of the relative stability of $G(z)$, the undermodelling error can be reduced in comparison to the use of an FIR model structure (Makila:1990), which is a special case of the Laguerre structure when $a = 0$.

- Two–parameter Kautz basis is defined as follows. Let

$$\zeta(z) = \frac{cz^2 + bz + 1}{z^2 + bz + c}$$

where $b, c \in \mathbf{R}$ satisfy $b^2 - 4c < 0$ and ζ has no poles in the closed unit disk. Let $\psi_0(z) = 1$, $\psi_1(z) = 1/(z^2 + bz + c)$ and $\psi_2(z) = z/(z^2 + bz + c)$. Let $\psi_k(z) = \zeta(z)\psi_{k-2}(z)$ for $k > 2$. Then Kautz models are obtained by orthonormalizing ψ_0, ψ_1 and ψ_2. It is more appropriate to employ them when prior knowledge of a resonant mode exists (Wahlberg:1994).

- Laguerre and Kautz models are special cases of general orthonormal bases (Heuberger *etal.*:1995) where the poles are again restricted to a finite set.
- The rational wavelet basis

$$\mathcal{B}_w(z) \triangleq \frac{1}{1-\overline{w}z}, \quad w \in W$$

where W is a set of discrete points in Dhas been suggested in Dudley Ward and Partington: 1996.

  - Denoting the linear span of the set $\{\mathcal{B}_0, \mathcal{B}_1, \cdots, \mathcal{B}_{n-1}\}$ by X_n, the wavelet basis enjoys the advantage of generalising the FIR, Laguerre, two–parameter Kautz, and general orthonormal bases. Intuitively, this should lead to smaller undermodelling induced error when employed for the purposes of system identification.

Fundamental set A set $\{\mathcal{B}_w : w \in W\}$ is fundamental in a given space X if its closure under the norm on X is equal to X.

- A sufficient condition for the wavelet basis functions to be fundamental in $\mathcal{A}(D)$) is that W be a dyadically spaced lattice of the form (Dudley Ward and Partington:1996):

$$W = \left\{ \xi_{p,k} : \xi_{p,k} = (1 - 2^{-p}) e^{j2\pi k/2^p},\ 0 \le k < 2^p;\ p \ge 0 \right\}.$$

  The dyadically spaced lattice satisfies the so-called ‘Hayman-Lyons condition’ considered in Hayman and Lyons:1990.

Re–parameterise the linear space $\text{sp}\{\mathcal{B}_w; w \in W\}$ as

$$\begin{aligned} X_n &= \text{sp}\left\{\frac{1}{1-\overline{\xi_k}z}\ ; k = 0, 1, 2\cdots, n-1\right\} \\ &= \text{sp}\left\{\mathcal{B}_k(z)\ ; k = 0, 1, 2, \cdots, n-1\right\} \end{aligned}$$

where the functions $\{\mathcal{B}_k(z)\}$, which have been considered in detail in Ninness and Gustafsson:1997 are defined by

$$\begin{aligned} \mathcal{B}_0(z) &= \sqrt{1-|\xi_0|^2}/(1-\overline{\xi_0}z), \\ \mathcal{B}_k(z) &= \frac{\sqrt{1-|\xi_k|^2}}{1-\overline{\xi_k}z}\prod_{m=0}^{k-1}\frac{z-\xi_m}{1-\overline{\xi_m}z}, \quad k = 1, 2, \cdots . \end{aligned} \tag{2}$$

They are orthonormal in $\mathcal{H}_2(D)$ with respect to the inner–product

$$\langle \mathcal{B}_n, \mathcal{B}_m \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{B}_n(e^{j\omega}) \overline{\mathcal{B}_m(e^{j\omega})} \, \mathrm{d}\omega = \begin{cases} 1 & ; m = n \\ 0 & ; m \neq n. \end{cases}$$

- The Laguerre, two–parameter Kautz, general orthonormal, rational wavelet bases are special cases of (2).

Objectives

- To show that a necessary and sufficient condition for (2) to be fundamental in $\mathcal{A}(D)$ and in $\mathcal{H}_p(D)$ for all $1 \leq p < \infty$ is that $\sum_{k=0}^{\infty}(1 - |\xi_k|) = \infty$, a much milder condition than the Hayman-Lyons condition!
- To study robust estimation using the minimax scheme proposed by Mäkilä: 1991 and derive conditions for robust convergence, and to explicitly quantify estimation error for each of the spaces $\mathcal{A}(D), \mathcal{H}_p(D)$ and ℓ_1 and for both frequency–domain and time–domain measurements.

Consider the problem of identifying an LTI, SISO, discrete-time system with impulse response $g(k)$.

- The system is ℓ_2 bounded–input/bounded–output stable and real. Then, $G(z) = \sum_{k=0}^{\infty} g(k)z^k \in \mathcal{H}_\infty$.
- If the system is ℓ_∞ bounded–input/bounded–output stable, then $g \in \ell_1$ and $G(z)$ is continuous on T so that in fact $G(z) \in \mathcal{A}(D)$.

The identification of $G(z)$ is performed on the basis of the observed and possibly noise corrupted input–output behavior of the system.

Frequency-domain identification

Measurement set–up

$$E_k = G(e^{j\omega_k}) + \eta_k; \quad k = 0, \cdots, N$$

where η_k is a corruption to the true frequency response $G(e^{j\omega_k})$ assumed to be bounded as $\|\eta\|_\infty \leq \epsilon$.

The robust identification objective is to produce, on the basis of the observed response E_k, an approximate model $\widehat{G}_N \in \mathcal{A}(D)$ for G in such a way that the following condition is satisfied:

$$\lim_{\substack{N \to \infty \\ \epsilon \to 0}} \sup_{\|\eta\|_\infty \leq \epsilon} \|\widehat{G}_N - G\|_\infty = 0, \quad \text{for all } G \in \mathcal{A}(D). \tag{3}$$

Time-domain identification

Measurement set–up

$$y(t) = (g \circledast u + \eta)(t) = \sum_{k=0}^{\infty} g(k)u(t-k) + \eta(t),$$

where the input signal $u(t)$ is bounded as $\|u\|_\infty \leq 1$ (with $u(t) = 0$ for $t < 0$) and $y(t)$ is the measured output corrupted by a bounded disturbance $\|\eta\|_\infty \leq \epsilon$.

The robust estimation objective is to again satisfy (3) or the following condition under the constraint that $\widehat{g}_N \in \ell_1$

$$\lim_{\substack{N\to\infty \\ \epsilon\to 0}} \sup_{\|\eta\|_\infty \leq \epsilon} \|\widehat{g}_N - g\|_1 = 0 \text{ for all } g \in \ell_1. \tag{4}$$

- An identification algorithm that satisfies either one of the above properties is called *convergent* and *robustly convergent* if it does not rely on *a–priori* information about the unknown system and noise.
- We will call η noise although it may be present due to nonlinearities, time variations etc.

The identification algorithms to be studied are a result of the works of Partington (1994) and Mäkilä (1991) who have derived a general framework to solve robust estimation problems of the form just posed.

Given the linear model structure

$$\widehat{G}(z,\theta) = \sum_{k=0}^{n-1} \theta_k \mathcal{B}_k(z) \tag{5}$$

and linear subspaces $X_k = \text{sp}\{\mathcal{B}_0, \cdots, \mathcal{B}_{k-1}\}$, for frequency domain measurements the robust estimate $\widehat{G}_N(z)$ is found as the solution of the minimax problem:

$$\widehat{G}_N(z) \triangleq \arg\min_{H \in X_n} \max_{0 \le k \le N} |H(e^{j\omega_k}) - E_k|. \tag{6}$$

A sufficient condition on the model structure X_n and the denseness of the frequency evaluation points ω_k such that (6) results in a robust estimator satisfying (3) is that there exists a fixed $0 < \delta < 1$ such that for each n

$$\max_{0 \le k \le N} |G(e^{j\omega_k})| \ge \delta \, \|G\|_\infty \quad \text{for all } G \in X_n. \tag{7}$$

In the case of time–domain data, the algorithm (6) takes the following (similar) form

$$\widehat{g}_N \triangleq \arg\min_{g\in X_n} \max_{0\leq t\leq N-1} |(g \circledast u)(t) - y(t)| \quad (8)$$

where $\widehat{g}_N$ denotes the impulse response of the identified model and a sufficient condition on the model structure X_n and the input u such that (8) results in a robust estimator satisfying (4) is that there exists a fixed $0 < \delta < 1$ such that for each n

$$\max_{0\leq t\leq N-1} |(g \circledast u)(t)| \geq \delta \, \|G\|_\infty \text{ (or } \|g\|_1) \quad (9)$$

and for all G (or g) $\in X_n$.

Estimation error bounds

Provided that the conditions (7) and (9) hold respectively, it is possible to specify explicit bounds on the estimation error as (Partington:1994,Dudley Ward and Partington:1996)

$$\|G - \widehat{G}_N\|_\infty \leq \left(\frac{2}{\delta} + 1\right) d(G, X_n; \mathcal{A}(D)) + \frac{2}{\delta}\epsilon \qquad (10)$$

$$\|g - \widehat{g}_N\|_1 \leq \left(\frac{2}{\delta} + 1\right) d(g, X_n; \ell_1) + \frac{2}{\delta}\epsilon \qquad (11)$$

where $d(f, X_n; X)$ defined as

$$d(f, X_n; X) \triangleq \inf_{h \in X_n} \|h - f\|_X \tag{12}$$

represents the error in approximating f by some function from the model set X_n.

- The conditions (7) and (9) are also necessary for robust recovery of systems in the spaces $\mathcal{A}(D)$ and ℓ_1 (Partington:1996).
- Many other estimation approaches are possible other than (6) or (8). If the evaluation points ω_k are uniformly spaced, then a class of two–stage non–linear methods are available (Gu and Khargonekar:1992) for which worst-case error bounds are comparable to (10).

The sub-optimality property of the two–stage schemes crucially depends on the uniform frequency spacing. If this uniformity requirement is dropped, the error due to undermodelling will decrease polynomially in model order n (Akçay *etal.*:1994) even if G is extremely smooth.

- whereas using (6), the undermodelling error will decrease according to $d(G, X_n, \mathcal{A}(D))$ which decreases exponentially in n for exponentially stable discrete-time systems.
- The formulation (6) or (8) reduces the worst-case identification problem to a choice of complete model sets.
- It is desirable to choose (via prior knowledge of $G(z)$) basis functions $\mathcal{B}_k(z)$ such that the distance $d(G, X_n; X)$ from $G(z)$ to $X_n = \text{sp}\{\mathcal{B}_0, \cdots, \mathcal{B}_{n-1}\}$ is as small as possible.

- Once a complete model set $\{X_n\}$ for X has been chosen, it is necessary to check that it is compatible with the measurement set–up in that the sufficient conditions (7) or (9) for robust convergence are satisfied.
- In the frequency domain, the link between n and N is provided by Bernstein's inequality (Zygmund:1959) for FIR models and for rational models, it can be derived by means of a sharp inequality due to Borwein and Erdelyi (1996).
- In the time domain, the dependence between n and N is referred to as the *sample complexity* (Poolla and Tikku:1994).

Outline

We will establish that the orthonormal set (2) is a fundamental set for $\mathcal{A}(D)$ iff $\sum_{n=0}^{\infty}(1-|\xi_n|)=\infty$.

- It is known (Dewilde and Dym:1981) that $\{\mathcal{B}_k\}_{k\geq 0}$ is fundamental in $\mathcal{H}_2$ under the same necessary and sufficient condition, and we show that this may be extended to all $\mathcal{H}_p$ spaces for $1\leq p<\infty$.

Finite Blaschke products

$$\varphi_n(z) \triangleq \prod_{m=0}^{n-1} \frac{z - \xi_m}{1 - \overline{\xi_m}\, z}. \tag{13}$$

Lemma (Christoffel-Darboux Identity) For all $z, \zeta \in D$,

$$\sum_{k=0}^{n-1} \overline{\mathcal{B}_k(\zeta)}\, \mathcal{B}_k(z) = \frac{1 - \overline{\varphi_n(\zeta)}\, \varphi_n(z)}{1 - \bar{\zeta} z}. \tag{14}$$

- A key consequence of this result is that it facilitates a simple integral formulation of $d(G, X_n; \mathcal{A}(D))$ as follows.

Lemma Let $G \in \mathcal{A}(D)$. Let $\widehat{G}_n$ be the projection

$$\widehat{G}_n(\zeta) \triangleq \sum_{k=0}^{n-1} \langle G, \mathcal{B}_k \rangle \, \mathcal{B}_k(\zeta), \quad \zeta \in \mathbf{D}. \tag{15}$$

Then

$$G(\zeta) - \widehat{G}_n(\zeta) = \frac{\varphi_n(\zeta)}{2\pi j} \oint_T \frac{G(z)}{z - \zeta} \, \overline{\varphi_n(z)} \, \mathrm{d}z. \tag{16}$$

- In order to use this result to provide $\mathcal{L}_\infty$ error bounds, it is necessary to derive an upper bound on $|\varphi_n(z)|$.

Lemma Let φ_n be as in (13). Then for each $z \in \mathbf{D}$

$$|\varphi_n(z)| \leq \exp\left(-\frac{1}{2} (1 - |z|) \sum_{k=0}^{n-1} (1 - |\xi_k|) \right).$$

Let $\mathcal{A}(D_R, K)$ denote the set of all $G(z)$ which are magnitude bounded by $K < \infty$ and analytic on $D_R = \{z \in \mathbf{C} : |z| < R\}$ for some $R > 1$.

- For a given $G \in \mathcal{A}(D_R, K)$, the $\mathcal{L}_\infty$ distance $d(G, X_n; \mathcal{A}(D))$ from $X_n = \text{sp}\{\mathcal{B}_0, \cdots, \mathcal{B}_{n-1}\}$ to an arbitrary $G(z)$ can be bounded.

Lemma Let $G \in \mathcal{A}(D_R, K)$. Then,

$$d(G, X_n; \mathcal{A}(D)) \leq \frac{KR}{R-1} \exp\left(-\frac{R-1}{2R}\sum_{k=0}^{n-1}(1-|\xi_k|)\right).$$

A set $A \subset X$ is fundamental in X if and only if any bounded linear functional vanishing on A also vanishes on X.

- This fact will be used firstly to examine the case of $X = \mathcal{A}(D)$ and the Laguerre basis.

Lemma The Laguerre basis obtained by letting $\xi_k = a$, $\forall k$ in (2) is fundamental in $\mathcal{A}(D)$ for all $-1 < a < 1$.

Theorem The orthonormal set (2) is fundamental in $\mathcal{A}(D)$ if and only if $\sum_{k=0}^{\infty}(1 - |\xi_k|) = \infty$.

Corollary The set (2) is fundamental in $\mathcal{H}_p$ for all $1 \leq p < \infty$ if and only if $\sum_{k=0}^{\infty}(1 - |\xi_k|) = \infty$.

- The set (2) is a *minimal spanning set* in $\mathcal{H}_2$ since its elements are orthonormal and removal of any element from the set diminishes the span.
- The above results can also be obtained directly from the results in Achieser:56. Nevertheless, our results are self contained and further results will be based on the explicit error bounds derived.

The rational wavelet basis is defined by

$$\mathcal{B}_w \triangleq \frac{1}{1 - \bar{w}z}, \qquad w \in W.$$

where W is an arbitrary subset of D.

Corollary Consider the rational wavelet basis. Then, $\{\mathcal{B}_w\}_{w \in W}$ is a fundamental set in $\mathcal{A}(D)$ if and only if $\sum_{w \in W}(1 - |w|) = \infty$.

- The lattice W may be modified so as to contain an element w a finite or infinite number of times if each repeated w can be associated uniquely with a basis function in the form $(1 - \overline{w}z)^{-k}$ for some integer $k > 1$.
- The base constructed in this manner does not contain polynomials in its linear span. This deficiency can be remedied by adjoining polynomials into the base.
- In Ward and Partington: 1995, model sets containing both rationals and polynomials are generated by the Hardy-Sobolev norm on *smooth* subset of $\mathcal{A}(D)$.

Outline

We consider now fundamental model sets applicable for robust estimation from time-domain data.

Theorem The set (2) is fundamental in ℓ_1 if

$$\lim_{m\to\infty} \exp\left(-\frac{1}{2}\sum_{k=0}^{m-1}(1-|\xi_k|)\right)\sum_{k=0}^{m-1}\frac{1}{1-|\xi_k|}=0.$$

Corollary The set (2) is fundamental in ℓ_1 if $k^{-\alpha}=O(1-|\xi_k|)$ for some $0<\alpha<1$.

- A base $\{\mathcal{B}_k\}$ can be fundamental in ℓ_1 without requiring the set of points $\{\xi_k\}$ have an accumulation point in D.
- When the set $\{\xi_k\}$ has an accumulation point in D, the conclusion of the theorem easily follows from Theorem 2 in Dudley Ward and Partington:1995, since in this case the set of basis functions $\{\mathcal{B}_k\}$ will be a fundamental set in the (Hardy-Sobolev) $H^{2,1}$ norm, which dominates the ℓ_1 norm.
- The corollary provides a rather tight criterion. For example, if $1 - |\xi_k| = O(1/k(\log k)^{\beta})$ for some $\beta > 1$, then the base $\{\mathcal{B}_k\}$ will not even be fundamental in $\mathcal{A}D$.

The following result will be required later.

Corollary Let $\{X_n\}_{n\geq 0}$ be the model set spanned by the orthonormal set in (2), where the chosen poles lie in the complement of D_r for some fixed $r > 1$. Let g denote the impulse response of a transfer function $G \in \mathcal{A}(\mathcal{D}_{\mathcal{R}}, \mathcal{K})$. Then

$$d(g, X_n; \ell_1) \leq \frac{KR}{R-1}\left(\frac{r+1}{r-1}n + \frac{R}{R-1}\right)e^{-(R-1)(r-1)n/2Rr}.$$

Outline

A solution to the problem of identification from frequency domain data is presented.

- The model sets are assumed to be complete in $\mathcal{A}(D)$ but arbitrary.
- The frequency response data are not required to be uniformly spaced.
- Given fundamental model sets $\{X_n\}$, it is necessary to derive the relationship $n(N)$ (known as a sampling theorem for $\mathcal{A}(D)$) such that the sufficient condition (7) holds for robust convergence of the scheme (6) to exist.

Maximum angular gap:

$$\delta_N \triangleq \max_{0 \leq k \leq N} \min_{\substack{\ell \neq k \\ 0 \leq \ell \leq N}} |\omega_k - \omega_\ell| \tag{17}$$

where $\omega_0 = 0$ and $\omega_1 = \pi$.

Lemma Let $\{\mathcal{B}_k\}_{k=0}^{n-1}$ be a set of rational functions analytic in D_r for some $r > 1$ and let p denote the number of poles of $\{\mathcal{B}_k\}_{k=0}^{n-1}$ (including poles at ∞). Let δ_N be as in (17). Then (7) holds for some $0 < \delta < 1$ provided that

$$\delta_N \leq 2\,\frac{(1-\delta)(r-1)}{p(r+1)}. \tag{18}$$

Corollary Consider the orthonormal set in (2). Let $r_n = \max_{k<n} |\xi_k|$. Then (7) holds for some $0 < \delta < 1$ if

$$\delta_N \leq 2\,\frac{1-\delta}{n}\left(\frac{1-r_n}{1+r_n}\right). \tag{19}$$

- If ω_k are uniformly spaced and the chosen poles outside $|z| < r$ for some fixed $r > 1$, (7) is satisfied provided

$$N \geq \frac{(r+1)\pi n}{2(1-\delta)(r-1)}.$$

- This condition is weaker than the requirement for the rational wavelets in Dudley Ward and Partington:1996.

Theorem Consider the orthonormal set in (2). Suppose $\{e^{j\omega_k}\}_{k\geq 0}$ is dense in T. Let δ_N be as in (17).Then the algorithm given in (6) is robustly convergent if $\sum_{n=0}^{\infty}(1-|\xi_n|)=\infty$ and δ_N satisfies (19) with $r_n=\max_{k<n}|\xi_k|$. In particular for each fixed $r>1$, an orthonormal set of rational functions with poles in the complement of D_r can be chosen such that the algorithm given in (6) is robustly convergent if

$$\delta_N \leq 2\,\frac{1-\delta}{n}\left(\frac{r-1}{r+1}\right)$$

or when the frequencies are uniformly spaced

$$N \geq \frac{\pi n}{2(1-\delta)}\left(\frac{r+1}{r-1}\right).$$

Theorem Consider the orthonormal set in (2). Let $r_n = \max_{k<n} |\xi_k|$. Suppose $\{e^{j\omega_k}\}_{k\geq 0}$ is dense in T. Let δ_N be as in (17). For each N, choose an n such that (19) is satisfied. Let $\widehat{G}_n$ be the estimate of $G \in \mathcal{A}(D_R, K)$, $R > 1$ by the algorithm given in (6). Then

$$\|G - \widehat{G}_n\|_\infty \leq \left(\frac{2}{\delta} + 1\right) \frac{KR}{R-1} \exp\left(-\frac{R-1}{2R} \sum_{k=0}^{n-1} (1 - |\xi_k|)\right) + \frac{2}{\delta}\,\epsilon.$$

In particular for each fixed $r > 1$, an orthonormal set of rational functions with poles in the complement of D_r can be chosen such that if δ_N satisfies (5) or N satisfies (5) when the frequencies are uniformly spaced, then

$$\|G - \widehat{G}_n\|_\infty \leq \left(\frac{2}{\delta} + 1\right) \frac{KR}{R-1}\, e^{-(R-1)(r-1)n/2Rr} + \frac{2}{\delta}\,\epsilon.$$

- The last theorem extends the Laguerre and Kautz results in Dudley Ward and Partington:1996 to arbitrary orthonormal bases.
- Notice that The upper bound on the estimation error is minimized for $r = \infty$. This conforms with the *n-width* result Pinkus:1985 that polynomial models are optimal for the class $\mathcal{A}(D_R, K)$.

Outline

Solutions to the problems of identification from time-domain data are presented.

- The condition (9) places severe restrictions on inputs.

An input signal u satisfying (9) is called a δ-cover of X_n (Harrison etal.:1997). The length of the shortest δ-cover of X_n is the *sampling size* for u.

- The sampling sizes and δ-covers are known for polynomials and certain compact subsets of $\mathcal{A}(D)$ and ℓ_1:
  - The sampling size for the set of nth order polynomials denoted by $\mathcal{P}_n$ in the ℓ_1-norm is $O(\beta^n)$ for some $\beta \in (1, 2]$ (Dahleh *etal.*:1993).
  - The sampling size of $\mathcal{P}_n$ in the $\mathcal{H}_\infty$-norm is $O(n^2)$ (Harrison *etal.*:1996).

- The δ-covers for polynomial models can be used in the construction of δ-covers for compact rational model sets (with the same norm).
  - **Example:** The set of *n*-th order, strictly proper transfer functions which are analytic on D_r for some fixed $r > 1$, denoted by $\mathcal{V}(n, r^{-1})$. Every $0.2 + 0.8\delta$-cover of $\mathcal{P}_m$ is also a δ-cover of $\mathcal{V}(n, r^{-1})$ (Harrison *etal.*:1997) where *m* can be chosen to be any integer satisfying

$$m \geq \frac{4nr}{r-1} \ln \left(\frac{20r}{(1-\delta)(r-1)} \right).$$

Lemma Let $\{X_n\}_{n\geq 0}$ be the model set spanned by the orthonormal set in (2), where the chosen poles lie in the complement of D_r for some fixed $r > 1$. For each n choose an integer m satisfying (3). Let u be the $0.2 + 0.8\delta$-cover of $\mathcal{P}_m$ in X, where X denotes either $\mathcal{A}(D)$ or ℓ_1, and let N be the length of u. Then

$$\max_{0\leq t\leq N-1} |(g \circledast u)(t)| \geq \delta \, \|g\|_X \quad \text{for all } g \in X_n.$$

Theorem Consider the orthonormal set in (2), where the chosen poles lie in the complement of D_r for some fixed $r > 1$. Let X denote either ℓ_1 or $\mathcal{A}(D)$ and let the inputs be chosen as in the lemma. Then the algorithm given in (8) robustly converges in X.

Theorem Consider the orthonormal set in (2), where the chosen poles lie in the complement of D_r for some fixed $r > 1$. Let the inputs be chosen as in the lemma. Let $\widehat{G}_n$ be the estimate of $G \in \mathcal{A}(D_R, K)$, by the algorithm given in (6). Then

$$\|G - \widehat{G}_n\|_\infty \leq \left(\frac{2}{\delta} + 1\right) \frac{KR}{R-1}\, e^{-(R-1)(r-1)n/2Rr} + \frac{2}{\delta}\, \epsilon.$$

Theorem Consider the orthonormal set in (2), where the chosen poles lie in the complement of D_r for some fixed $r > 1$. Let the inputs be as in the lemma. Let $\widehat{g}_n$ be the estimate of g, the impulse response of $G \in \mathcal{A}(D_R, K)$, by the algorithm given in (8). Then

$$\|g - \widehat{g}_n\|_1 \leq \left(\frac{2}{\delta} + 1\right) \frac{KR}{R-1} \left(\frac{r+1}{r-1}\, n + \frac{R}{R-1}\right) \cdot e^{-(R-1)(r-1)n/2Rr} + \frac{2}{\delta}\,\epsilon.$$

When $X = \mathcal{H}_2(D)$ (or ℓ_2), a necessary and sufficient condition for the existence of robustly convergent algorithms in X is that there exists a fixed $0 < \delta < 1$ such that for each n

$$\max_{0 \le t \le N-1} |(g \circledast u)(t)| \ge \delta \, \|g\|_2 = \delta \, \|G\|_2 \quad \text{for all } g \text{ (or } G) \ \in X_n.$$

Then , for the minimax algorithms an explicit bound on the ℓ_2 norm of the estimation error is obtained from Partington:1994 as

$$\|g - \widehat{g}_N\|_2 \le \left(\frac{2}{\delta} + 1 \right) d(g, X_n, \ell_2) + \frac{2}{\delta} \, \epsilon.$$

- The persistence of excitation condition places mild restrictions on the choice of input signal despite the fact that it appears to be stronger than the usual persistence of excitation condition in Ljung:1999.
  - The sampling size of δ-covers for polynomial models is $O(n)$ (Partington:1994).
  - As well, the least-squares algorithm has robust convergence property in ℓ_2 (or $\mathcal{H}_2(D)$) identification. Moreover, u can be chosen such that to identify a system $G \in \rfloor alP_n$ with an error of $O(\epsilon)$ one requires only $O(n)$ measurements (Partington:1994).

- An analysis of the use of rational model structures in a robust estimation context was provided. A key result of this analysis was the provision of necessary and sufficient conditions on the poles of the rational model structures for them to form a fundamental set in $\mathcal{A}(D)$ and $\mathcal{H}_p(D)$ $(1 \leq p < \infty)$.
- It was shown how robust estimation algorithms using both time and frequency domain data could be constructed together with explicit error bounds on the estimation accuracy. These results have implications for mixed parametric/non–parametric estimation, model reduction and may be extended to the multi-variable setting.