Adaptive Stochastic Systems
Estimation, Filtering, and Noise Attenuation

Araz Hashemi

Wayne State University
Detroit, MI

March 19, 2014
OUTLINE

Sign-Regressor Algorithms
   Motivation and Adaptive Filtering
   Formulation
   Limit Convergence
   Asymptotic Distribution

Sign-Error Algorithms
   Formulation
   Limit Convergence
   Limit Distributions
   Simulated Performance Results

Noise Attenuation
   Formulation
   Robust Noise Attenuation with Unmodeled Dynamics
   Impact of Signal Estimation Errors
   Examples
LINEAR MODELS

\[ y_n = \varphi_n^T \alpha + e_n \]

- \( \varphi_n \in \mathbb{R}^r = \) input vector (at ‘time’ \( n \)).
- \( \alpha \in \mathbb{R}^r = \) system parameter
- \( e_n \in \mathbb{R} = \) stochastic system noise (random error at \( n \))
- \( y_n \in \mathbb{R} = \) output observation (at \( n \))

Estimation problem: Use (known) inputs \( \varphi_n \) with observed outputs \( y_n \) to estimate \( \alpha \).
Adaptive Filtering

- Suppose \( \{\varphi_n\}, \{e_n\} \) are stationary sequences with
  - \( \mathbb{E}[\varphi_n\varphi_n'] = R \) ("covariance matrix")
  - \( \mathbb{E}[\varphi_ny_n] = \mathbb{E}[\varphi_n(\varphi_n'\alpha + e_n)] = q \)

- Let \( \theta \) represent our current estimate for \( \alpha \)

- Write \( J(\theta) = \mathbb{E} |y_n - \varphi_n'\theta|^2 \).

- \( \alpha \) minimizes \( J(\cdot) \), with \( J(\alpha) = 0 \) (\( \alpha = R^{-1}q \))

- Hence we want to construct a SA algorithm for \( \theta_n \) such that \( \theta_n \to \alpha \)
Least Mean-Squares Algorithm

Algorithm 1 (LMS)

A Least Mean Squares (LMS) Algorithm is of the form:

\[
\theta_{n+1} = \theta_n + \mu_n \varphi_n [y_n - \varphi_n' \theta_n] \\
= \theta_n + \mu_n \varphi_n [\varphi_n' \alpha + e_n - \varphi_n' \theta_n] \\
= \theta_n - \mu_n \varphi_n \varphi_n' [\theta_n - \alpha] + \varphi_n e_n
\]

where \( \mu_n \) is a sequence of step-sizes (gain-sizes).

It is adaptive in the sense that it adjusts estimates based on the residual error \( y_n - \varphi_n' \theta_n \).
TIME-VARYING PARAMETER

- What if $\alpha$ is time-varying? E.g. $\alpha = \alpha_n$ so that
  
  $$y_n = \varphi'_n \alpha_n + e_n$$

- Take a constant gain size $\mu_n \equiv \mu$

- Since the adaptive filtering algorithm adjusts by
  
  $$\theta_{n+1} = \theta_n - \mu \varphi_n \varphi'_n [\theta_n - \alpha_n] + \varphi_n e_n$$

  it should be able to track the parameter $\alpha_n$ as it changes
Markovian Parameter

- If $\alpha_n$ evolves by small increments (e.g. $\alpha_{n+1} = \alpha_n + \tilde{e}_n$ where $\mathbb{E}\tilde{e}_n = 0$) then asymptotic convergence is known.

- What if $\alpha_n$ can “jump” by arbitrarily large increments?

- Suppose $\alpha_n$ is a Markov chain with state space $\mathcal{M} = \{a_1, a_2, \ldots, a_{m_0}\}$ and transition matrix

\[
P^\varepsilon = I + \varepsilon Q
\]
**Multiple Access Networks**

**TDMA Callers Use Time Slices of a Frequency**

**CDMA Users are Separated by Codes**
SIGN-ALGORITHMS

- Often want to speed computations when filtering, especially in communication networks
  - High dimensional data
  - Frequent data shuffling
  - Limited resources
- One way to speed computations is to reduce data complexity during estimation
- By minimizing $J(\theta) = \mathbb{E}|y_n - \varphi'_n \theta|^2$ we obtained:

  \[
  \text{LMS: } \quad \theta_{n+1} = \theta_n + \mu \varphi_n (y_n - \varphi'_n \theta_n)
  \]

- By minimizing $J(\theta) = \mathbb{E}|y_n - \varphi'_n \theta|$ one obtains the Sign-Error algorithm

  \[
  \text{SE: } \quad \theta_{n+1} = \theta_n + \mu \varphi_n \text{sgn}(y_n - \varphi'_n \theta_n)
  \]
**Sign-Error to Sign-Regressor**

- By just using the sign of the residuals $y_n - \varphi_n' \theta_n$, computations are reduced to simple bit shifts with substantial speed improvement.

- However, the non-linear (non-continuous) $\text{sgn}(\cdot)$ operator makes analysis difficult.

- Also, by ‘throwing away’ much of the information in the residuals, the SE algorithm tends to converge more slowly than the LMS.

- Instead, take $\text{sgn}(\cdot)$ on the regression vector $\varphi_n$. 
Algorithm 2 (SR)

A Sign-Regressor (SR) algorithm is of the form:

$$\theta_{n+1} = \theta_n + \mu \text{Sgn}(\varphi_n)(y_n - \varphi'_n \theta_n)$$

- ‘Compromise’ of SE and LMS
- Uses full error information from residuals $y_n - \varphi'_n \theta_n$
- ‘Clips’ direction information from regressor $\varphi_n$
- Still shows improved computation speed from LMS
- Linear form helps facilitate analysis
\[ y_n = \varphi' n \alpha_n + e_n \]  \hspace{1cm} (1)

**A 1.1 (Markov Chain Assumptions)**

- \( \alpha_n \) is a discrete-time homogeneous Markov chain, state space \( \mathcal{M} = \{a_1, a_2, \ldots, a_{m_0}\}, a_i \in \mathbb{R}^r \).
- There exists a small \( \varepsilon > 0 \) such that the transition probability matrix of \( \alpha_n \) is given by
  \[ P^{\varepsilon} = I + \varepsilon Q \]
  where \( Q \) is an irreducible generator of a continuous-time Markov chain.
- The initial distribution
  \[ \pi_0 = [\mathbb{P}\{\alpha_0 = a_1\}, \mathbb{P}\{\alpha_0 = a_2\}, \ldots, \mathbb{P}\{\alpha_0 = a_{m_0}\}] \] is independent of \( \varepsilon \).
A 1.2 (Mixing-type signals)

- The sequences \( \{\varphi_n\}, \{e_n\} \) are independent of the parameter process \( \{\alpha_n\} \).
- The sequence of signals \( \{ (\varphi_n, e_n) \} \) is bounded.
- There exists a stable matrix \( H \in \mathbb{R}^{r \times r} \) and a constant \( K > 0 \) such that for all \( n \)

\[
\left| \sum_{j=n}^{\infty} E_n \left[ \text{Sgn}(\varphi_j)\varphi_j' - H \right] \right| \leq K \\
\left| \sum_{j=n}^{\infty} E_n \left[ \text{Sgn}(\varphi_j)e_j \right] \right| \leq K
\]

where \( E_n \) is conditional expectation w.r.t. \( \{\varphi_j, e_j, \alpha_j : j < n; \alpha_n\} \)
A 1.3 (Ergodic properties)

For the matrix $H$ as in A1.1 and for each $m \in \mathbb{N}$, as $n \rightarrow \infty$

\[
\frac{1}{n} \sum_{k=m}^{m+n} \text{Sgn} \varphi_k \varphi'_k \overset{p}{\rightarrow} H
\]

\[
\frac{1}{n} \sum_{k=m}^{m+n} \text{Sgn}(\varphi_k)e_k \overset{p}{\rightarrow} 0
\]

- Unbounded signals $(\varphi_n, e_n)$ can also be treated (use martingale difference signals or expanding truncation device). Assumed only for notational simplicity.

- Stationary, mixing signals ensure A1.2, A1.3
Mean-Square Error Bounds

Theorem 1.1

Let $\theta_n$ be given by the SR algorithm. Under assumptions A1.1 and A1.2, there exists $N_{\mu,\varepsilon} > 0$ such that for all $n > N_{\mu,\varepsilon}$ we have

$$E|\tilde{\theta}_n|^2 = E|\theta_n - \alpha_n|^2 = O(\mu + \varepsilon + \varepsilon^2/\mu).$$

- $\varepsilon$ = “transition rate” of parameter $\alpha_n$
- $\mu$ = “adaptation rate” of estimate $\theta_n$
**Infinitesimal Limit Behavior**

- Optimally, take $\mu = \frac{\varepsilon}{\sqrt{1+\varepsilon}}$ to minimize expected error bound.

- In practice, one chooses $\mu$ without knowledge of $\varepsilon$.

- Since smaller $\varepsilon$, $\mu$ give smaller expected error, one wonders what happens at the infinitesimal level as $\varepsilon, \mu \to 0$.

- Suppose $\varepsilon$ is a function of $\mu$, and then examine limit behavior as $\mu \to 0$.

- Depending on the relationship of $\varepsilon$ to $\mu$, we see very different behavior in the limit system.
Limit Analysis breaks down into 3 cases:

1. \( \varepsilon = O(\mu) \) “On-Line”
   - e.g. \( \varepsilon = \mu \)
   - \( \alpha_n \) can jump about as quickly as \( \theta_n \) can track it
   - Limit dynamics have occasional jumps in \( \alpha_n \), but \( \theta_n \) is still able to track it closely

2. \( \varepsilon \ll \mu \) “Slower Markov Chain”
   - e.g. \( \varepsilon = \mu^2 \)
   - \( \alpha_n \) jumps so infrequently it may as well be constant
   - Limit behavior determined by initial distribution
     \[
     \pi_0 = [\mathbb{P}\{\alpha_0 = a_1\}, \mathbb{P}\{\alpha_0 = a_2\}, \ldots, \mathbb{P}\{\alpha_0 = a_{m_0}\}\]
     \]

3. \( \varepsilon \gg \mu \) “Fast Markov Chain”
   - e.g. \( \varepsilon = \sqrt{\mu} \)
   - \( \alpha_n \) jumps too quickly for \( \theta_n \) to track it
   - Frequent jumping of \( \alpha_n \) means that it quickly comes to the stationary distribution \( \nu = [\nu_1, \ldots, \nu_{m_0}] \)
To examine the limit behavior, interpolate the discrete \( \theta_n, \alpha_n \) to continuous-time processes:

\[
\theta^\mu(t) \overset{\Delta}{=} \theta_n, \quad \alpha^\mu(t) \overset{\Delta}{=} \alpha_n, \quad \text{for } t \in [n\mu, n\mu + \mu)
\]

Then consider the limit of the random processes

\[
\theta^\mu(t) \xrightarrow{w} \theta(t)
\]
\[
\alpha^\mu(t) \xrightarrow{w} \alpha(t) \quad \text{as } \mu \to 0
\]

\[\xrightarrow{w}\] denotes “weak convergence”

i.e. \( X_n \xrightarrow{w} X \) if for any bounded, continuous function \( f \),

\[
E f(X_n) \to E f(X)
\]

\( \alpha_n \) is interpolated with step-size \( \mu \) when it actually changes at rate \( \varepsilon \). So we see different limits for \( \alpha(t) \) depending on the relationship of \( \varepsilon \) to \( \mu \).
ON-LINE LIMIT

Theorem 1.2

Let \( \varepsilon = O(\mu) \), and take \( \theta_n \) by the SR algorithm with assumptions A1.1, A1.2, and A1.3. Then as \( \mu \to 0 \)

\[
(\theta^\mu(t), \alpha^\mu(t)) \xrightarrow{w} (\theta(t), \alpha(t))
\]

such that \( \alpha(t) \) is a continuous-time Markov chain generated by \( Q \)
and \( \theta(t) \) satisfies the Markov-switching ODE

\[
\frac{d}{dt} \theta(t) = H(\alpha(t) - \theta(t)), \quad \theta(0) = \theta_0
\]
**SLOWER M.C. LIMIT**

When $\varepsilon \ll \mu$, the limit goes to the average against the initial distribution $\alpha_* = \sum_{i=1}^{m_0} a_i \pi_{0,i}$.

**Theorem 1.3**

Let $\varepsilon = \mu^{1+\eta}$ for some $\eta > 0$. Then as $\mu \to 0$, $\theta^\mu(t) \xrightarrow{w} \theta(t)$ such that $\theta(t)$ is the solution to the ODE

$$\frac{d}{dt} \theta(t) = H(\alpha_* - \theta(t)), \quad \theta(0) = \theta_0$$

**Corollary 1.4**

For any increasing sequence of time shifts $t_\mu \to \infty$ as $\mu \to 0$, $\theta^\mu(\cdot + t_\mu) \xrightarrow{w} \alpha_*$ as $\mu \to 0$. 

Fast M.C. Limit

When $\varepsilon \gg \mu$, the limit goes to the average against the stationary distribution $\bar{\alpha} = \sum_{i=1}^{m_0} a_i \nu_i$.

**Theorem 1.5**

Let $\varepsilon = \mu^\gamma$ for some $1/2 < \gamma < 1$, Then as $\mu \to 0$, $\theta^\mu(t) \xrightarrow{w} \theta(t)$ such that $\theta(t)$ is the solution to the ODE

$$
\frac{d}{dt} \theta(t) = H(\bar{\alpha} - \theta(t)), \quad \theta(0) = \theta_0
$$

**Corollary 1.6**

For any increasing sequence of time shifts $t_\mu \to \infty$ as $\mu \to 0$, $\theta^\mu(\cdot + t_\mu) \xrightarrow{w} \bar{\alpha}$ as $\mu \to 0$. 
Since \((\theta^\mu, \alpha^\mu) \xrightarrow{w} (\theta, \alpha)\), we wish to establish the rate of convergence.

- Given by appropriate scaling factor \(\gamma\) such that \(\frac{\theta^\mu - \alpha^\mu}{\mu^\gamma}\) converges to a non-zero limit.
- From \(\mathbb{E}|\theta_n - \alpha_n|^2 = O(\mu)\) (when \(\varepsilon = \mu\)), one expects \(\gamma = 1/2\) is the rate.

### A 1.4 (Mixing Signals CLT)

The scaled signals

\[
\sqrt{\mu} \sum_{j=0}^{t/\mu-1} \text{Sgn}(\varphi_j)e_j \xrightarrow{w} \tilde{w},
\]

where \(\tilde{w}(t)\) is a Brownian motion with covariance \(\tilde{\Sigma}t\) with \(\tilde{\Sigma} \in \mathbb{R}^{r \times r}\) positive definite.
ON-LINE LIMIT DISTRIBUTION

- With $\varepsilon = O(\mu)$, take $N_{\mu,\varepsilon} = N_{\mu}$ such that
  $\mathbb{E}|\theta_n - \alpha_n|^2 = O(\mu)$ for $n > N_{\mu}$. Then define the scaled error
  
  $u_n \triangleq \theta_n - \alpha_n \sqrt{\mu},$

  
  
  $u^\mu(t) \triangleq u_n$ for $t \in [(n - N_{\mu})\mu, (n - N_{\mu})\mu + \mu)$

Theorem 1.7

Let $\varepsilon = O(\mu)$ and assume A1.1 – A1.4. Then $u^\mu(\cdot) \xrightarrow{w} u(\cdot)$ such that $u(\cdot)$ is a solution to the stochastic differential equation

$$du = Hudt + \tilde{\Sigma}^{1/2}dw$$

where $w(\cdot) \in \mathbb{R}^r$ is a standard Brownian motion.
SLOOWER M.C. LIMIT DISTRIBUTION

- With $\varepsilon \ll \mu$, take error from $\alpha_*$. Define

$$v_n \triangleq \frac{\theta_n - \alpha_*}{\sqrt{\mu}},$$

$$v^\mu(t) \triangleq v_n \quad \text{for } t \in [(n - N_\mu)\mu, (n - N_\mu)\mu + \mu)$$

**Theorem 1.8**

Let $\varepsilon = \mu^{1+\eta}$ for $0 < \eta \leq 1$ and assume A1.1 – A1.4. Then $v^\mu(\cdot) \xrightarrow{w} v(\cdot)$ such that $v(\cdot)$ is a solution to the stochastic differential equation

$$dv = Hvdt + \tilde{\Sigma}^{1/2}dw$$

where $w(\cdot) \in \mathbb{R}^r$ is a standard Brownian motion.
**Fast M.C. Limit Distribution**

- With \( \varepsilon \gg \mu \), take error from \( \bar{\alpha} \). Define

\[
    z_n \triangleq \frac{\theta_n - \bar{\alpha}}{\sqrt{\mu}},
\]

\[
    z^\mu(t) \triangleq z_n \quad \text{for } t \in [(n - N_\mu)\mu, (n - N_\mu)\mu + \mu)
\]

**Theorem 1.9**

Let \( \varepsilon = \mu^\gamma \) for \( 1/2 \leq \gamma < 1 \) and assume A1.1 – A1.4. Then

\[
    z^\mu(\cdot) \xrightarrow{w} z(\cdot) \quad \text{such that } z(\cdot) \text{ is a solution to the stochastic differential equation}
\]

\[
    dz = Hzdt + \tilde{\Sigma}^{1/2}dw
\]

where \( w(\cdot) \in \mathbb{R}^r \) is a standard Brownian motion.
INTERPRETATION OF LIMIT DISTRIBUTION

- Theorems 1.7, 1.8, and 1.9 characterize errors $\theta_n - \alpha_n$, $\theta_n - \alpha_*$, and $\theta_n - \bar{\alpha}$ respectively.

- For each case the theorems imply the asymptotic error is mean 0 with variance $\mu S$,

- $S$ is the solution to the Lyapunov equation $HS + SH' = -\tilde{\Sigma}$.

- More explicitly $S = \int_0^\infty \exp(Hs)\tilde{\Sigma} \exp(H's)ds$. 
**SIGN-ERROR ALGORITHM**

We now return to the Sign-Error algorithm for analysis.

**Algorithm 3 (SE)**

A Sign-Error (SE) algorithm is of the form:

\[ \theta_{n+1} = \theta_n + \mu \varphi_n \text{sgn}(y_n - \varphi'_n \theta_n). \]

- \( \text{sgn}(\cdot) \) operator taken on the residuals
- With appropriate choice of ‘training’ sequence for \( \varphi_n \), computations can be reduced to simple bit shifts
- Significant speed improvement from LMS and SR
- Hard operator on the residuals makes analysis more difficult
ASSUMPTIONS

A 2.1 (Markov Chain)

*Same as A 1.1; \( \alpha_n = \text{Markov chain with transition matrix} \)

\[ P^\varepsilon = I + \varepsilon Q, \text{ etc...} \]

A 2.2 (Stationary Signals)

- \{ (\varphi_n, e_n) \} is stationary and independent of \{ \alpha_n \}.
- \( \varphi_n \) is bounded and \{ e_k \} is zero-mean.

Let \( \mathcal{F}_n \) be the \( \sigma \)-algebra generated by \{ (\varphi_j, e_j), \alpha_j : j < n; \alpha_n \}.
Denote the conditional expectation with respect to \( \mathcal{F}_n \) by \( E_n \).
A 2.3 (Locally Linear)

For each \( i = 1, \ldots, m_0 \), define

\[
\begin{align*}
    g_n & \triangleq \varphi_n \text{sgn}(\varphi'_n[\alpha_n - \theta_n] + e_n) \\
    g_n(\theta, i) & \triangleq \varphi_n \text{sgn}(\varphi'_n[a_i - \theta] + e_n) \mathbb{I}_{\{\alpha_n = a_i\}} \\
    \tilde{g}_n(\theta, i) & \triangleq E_n g_n(\theta, i)
\end{align*}
\]

For each \( n \) and \( i \), there is an \( A_n^{(i)} \in \mathbb{R}^{r \times r} \) such that given \( \alpha_n = a_i \),

\[
\begin{align*}
    \tilde{g}_n(\theta, i) &= A_n^{(i)}(a_i - \theta) \mathbb{I}_{\{\alpha_n = a_i\}} + o(|a_i - \theta| \mathbb{I}_{\{\alpha_n = a_i\}}) \\
    \mathbb{E}A_n^{(i)} &= A^{(i)}
\end{align*}
\]
A 2.4 (Linearization Mixing)

There is a sequence of non-negative real numbers \(\{\phi(k)\}\) with \(\sum_k \phi^{1/2}(k) < \infty\) such that for each \(n\) and each \(j > n\), and for some \(K > 0\),

\[
|E_n A_{j}^{(i)} - A^{(i)}| \leq K \phi^{1/2}(j - n) \tag{2}
\]

uniformly in \(i = 1, \ldots, m_0\).

- Boundedness only assumed on \(\varphi_n\), which is often deterministic anyway. Can also use expanding truncation to remove boundedness assumption.

- While \(g_n(\theta, i)\) is not smooth in \(\theta\), \(\tilde{g}_n(\theta, I) = E_n g_n(\theta, i)\) can be

- \(\tilde{g}_n(\theta, i)\) is locally (near \(a_i\)) linearizable if conditional joint density of \((\varphi_n, e_n)\) differentiable with bounded derivatives
With the stronger assumptions, we can obtain the same error bounds as before:

**Theorem 2.1 (Mean-Square Error Bounds)**

Assume A2.1 – A2.4. Then there is an $N_{\mu,\varepsilon} > 0$ such that for all $n \geq N_{\mu,\varepsilon}$,

$$
\mathbb{E}|\tilde{\theta}_n|^2 = \mathbb{E}|\alpha_n - \theta_n|^2 = O\left(\mu + \varepsilon + \varepsilon^2/\mu\right).
$$

For limit analysis, interpolate as before:

$$
\theta^\mu(t) \triangleq \theta_n, \quad \alpha^\mu(t) \triangleq \alpha_n \quad \text{for } t \in [n\mu, n\mu + \mu)
$$
**On-Line Case:** \( \varepsilon = O(\mu) \)

**Theorem 2.2 (On-Line Limit)**

Take \( \theta_n \) by the SE algorithm. Let \( \varepsilon = O(\mu) \) and assume A2.1 – A2.4. Then

\[
(\theta^\mu(\cdot), \alpha^\mu(\cdot)) \xrightarrow{w} (\theta(\cdot), \alpha(\cdot))
\]

such that \( \alpha(\cdot) \) is a continuous-time Markov chain with generator \( Q \) and \( \theta(\cdot) \) satisfies the Markov-switched ODE

\[
\frac{d}{dt} \theta(t) = A^{(\alpha(t))} (\alpha(t) - \theta(t)), \quad \theta(0) = \theta_0
\]
Slower M.C. Case: $\varepsilon \ll \mu$

In the case $\varepsilon \ll \mu$, the limit is again characterized by the initial distribution $\pi_0$ of $\alpha_0$.

**Theorem 2.3 (Slower M.C. Limit)**

Let $\varepsilon = \mu^{1+\eta}$ for $0 < \eta \leq 1$ and assume A2.1 – A2.4. Then $\theta^\mu(\cdot) \xrightarrow{w} \theta(\cdot)$ such that $\theta(\cdot)$ is the solution to the ODE

$$
\frac{d}{dt} \theta(t) = \sum_{i=1}^{m_0} A^{(i)} \left( a_i - \theta(t) \right) \pi_{0,i}, \quad \theta(0) = \theta_0
$$
**Fast M.C. Case**: $\varepsilon \gg \mu$

In the case $\varepsilon \gg \mu$, the limit is characterized by the stationary distribution $\nu$ associated with $Q$.

**Theorem 2.4**

*Fast M.C. Limit* Let $\varepsilon = \mu^\gamma$ for $1/2 \leq \gamma < 1$ and assume $A2.1 – A2.4$. Then $\theta^\mu(\cdot) \xrightarrow{w} \theta(\cdot)$ such that $\theta(\cdot)$ is the solution to the ODE

$$
\frac{d}{dt} \theta(t) = \sum_{i=1}^{m_0} A^{(i)} (a_i - \theta(t)) \nu_i, \quad \theta(0) = \theta_0
$$
**ON-LINE:** $\varepsilon = O(\mu)$

Again define scaled error

$$u_n \triangleq \tilde{\theta}_n/\sqrt{\mu} = (\alpha_n - \theta_n)/\sqrt{\mu}$$

Then interpolate to

$$u^\mu(t) \triangleq u_n \text{ for } t \in [(n - N_\mu)\mu, (n - N_\mu)\mu + \mu]$$
**ON-LINE:** $\varepsilon = O(\mu) \ II$

**Lemma 2.5 (CLT for Mixing Processes)**

Define $\varpi_k \triangleq \varphi_k \text{sgn}(e_k)$. Then

$$
\sqrt{\mu} \sum_{j=0}^{(t/\mu)-1} \varpi_j \xrightarrow{w} \tilde{w}(t)
$$

where $\tilde{w}(t)$ is a Brownian motion with covariance $\tilde{\Sigma}t$ given by

$$
\tilde{\Sigma} \triangleq \mathbb{E}\varpi_0\varpi_0' + \sum_{j=1}^{\infty} \mathbb{E}\varpi_j\varpi_0' + \sum_{j=1}^{\infty} \mathbb{E}\varpi_0\varpi_j'.
$$
Theorem 2.6 (On-Line Limit Distribution)

If $\varepsilon = O(\mu)$ and under A2.1 – A2.4 $u^\mu(\cdot) \xrightarrow{w} u(\cdot)$ such that

$$du = -A^{(\alpha)} u dt - \tilde{\Sigma}^{1/2} dw,$$

where $w(\cdot)$ is a standard Brownian motion and $\alpha = \alpha(\cdot)$ is the continuous-time Markov chain associated with $Q$. 
SLOWER M.C. : $\varepsilon \ll \mu I$

For the case $\varepsilon \ll \mu$, define

$$\alpha* \triangleq \sum_{i=1}^{m_0} a_i \pi_{0,i},$$

$$\nu_n \triangleq \frac{\alpha* - \theta_n}{\sqrt{\mu}}$$

$$\nu^\mu(t) \triangleq \nu_n \text{ for } t \in [(n - N_\mu)\mu, (n - N_\mu)\mu + \mu)$$

$$A^{(*)} \triangleq \sum_{i=1}^{m_0} A^{(i)} \pi_{0,i}.$$ 

Then we have the following.
**Slower M.C. : \( \varepsilon \ll \mu \)**

**Theorem 2.7**

**Slower M.C. Limit Distribution**  
If \( \varepsilon = \mu^{1+\eta} \) for some \( 0 < \eta \leq 1 \) and under A2.1 – A2.4 \( v^\mu(\cdot) \xrightarrow{w} v(\cdot) \) such that

\[
dv = -A^{(*)}vdt - \tilde{\Sigma}^{1/2}dw
\]

where \( w(\cdot) \) is a standard Brownian motion.
Fast M.C. : $\varepsilon \gg \mu$

For the case $\varepsilon \gg \mu$, define

$$\bar{\alpha} \triangleq \sum_{i=1}^{m_0} a_i \nu_i$$

$$z_n \triangleq \frac{\bar{\alpha} - \theta_n}{\sqrt{\mu}}$$

$$z'^\mu(t) \triangleq z_n \quad \text{for} \quad t \in [(n - N_\mu)\mu, (n - N_\mu)\mu + \mu)$$

$$\bar{A} \triangleq \sum_{i=1}^{m_0} A^{(i)} \nu_i.$$
**Theorem 2.8**

**Fast M.C. Limit Distribution** If \( \varepsilon = \mu^\gamma \) for some \( \frac{1}{2} \leq \gamma < 1 \) and under A2.1 – A2.4, \( z^\mu(\cdot) \overset{w}{\to} z(\cdot) \) such that

\[
dz = -\bar{A}zdt - \bar{\Sigma}^{1/2}dw
\]

where \( w(\cdot) \) is a standard Brownian motion.
SIMULATED PERFORMANCE RESULTS

- Here we ran simulations to demonstrate the performance (and convergence properties) of the algorithms in each of the cases.
- We fix the step size $\mu = .05$ and consider three cases:
  - $\epsilon = \frac{3}{5}\mu$ ($\epsilon = O(\mu)$)
  - $\epsilon = \mu^2$ (Slow Markov chain)
  - $\epsilon = \sqrt{\mu}$ (Fast Markov chain).
- Take state space $\mathcal{M} = \{-1, 0, 1\}$ with transition matrix $P^\epsilon = I + \epsilon Q$, where

$$Q = \begin{bmatrix}
-0.6 & 0.4 & 0.2 \\
0.2 & -0.5 & 0.3 \\
0.4 & 0.1 & -0.5
\end{bmatrix}.$$ 

and thus has $\nu = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. 
Hence \( \bar{\alpha} = \sum_{i=1}^{3} a_i \nu_i = 0 \)

Initial distribution \( \pi_0 = (3/4, 1/8, 1/8) \). So \( \alpha^* = \sum_{i=1}^{3} a_i \pi_{0,i} = -0.625 \).

Take \( \{\varphi_n\} \) and \( \{e_n\} \) are i.i.d. \( \mathcal{N}(0, 1) \) and \( \mathcal{N}(0, .25) \), respectively.

We proceed to observe 1000 iterations of the algorithm for the cases \( \varepsilon = O(\mu) \) and \( \varepsilon \gg \mu \), and 10,000 iterations for the case \( \varepsilon \ll \mu \) (in order to illustrate some variations of the parameter).
Parameter and Estimates: $\varepsilon = O(\mu)$
Parameter and Estimates: $\varepsilon \ll \mu$
Parameter and Estimates: $\varepsilon \gg \mu$
Average of Parameter and Estimates over time: $\varepsilon \gg \mu$: Convergence to stationary mean
Scaled error $z_n$ with $\varepsilon \gg \mu$: Diffusion Behavior
**Impact of Unmodeled Dynamics I**

- Often, models can be mismatched; e.g. the true system dynamics are higher dimensional than the modeled system

\[ y_n = \varphi'_n \alpha_n + e_n \]
\[ = \tilde{\varphi}'_n \tilde{\alpha}_n + \tilde{\varphi}'_n \tilde{\alpha}_n + e_n \]

- \( \tilde{\varphi}_n, \tilde{\alpha}_n \) modeled parts ; \( \tilde{\varphi}_n, \tilde{\alpha}_n \) unmodeled parts

- Take \( \varphi_n \) i.i.d. 7-dimensional \( \mathcal{N}(3, 1) \) with modeled part \( \tilde{\varphi}_n \)
  4-dimensional, errors \( e_n \sim \mathcal{N}(0, 0.25) \) as before

- Markov chain has state space \( \mathcal{M} = \{-\rho, 0, \rho\} \) where
  \( \rho = [1, 2^{-1}, \ldots, 2^{-6}] \in \mathbb{R}^7 \).
Impact of Unmodeled Dynamics II

- Transition matrix $P^{ε} = I + εQ$, initial $\pi_0$ and stationary $ν$ distributions as before

- Examine the SE algorithm for computing estimates of the modeled part of the parameter $\hat{α}$

$$\theta_{n+1} = \theta_n + \bar{φ}_n \text{sgn}(y_n - \bar{φ}_n' \theta_n) \in \mathbb{R}^4$$
Norm difference of modeled parameter and estimate $||\tilde{\alpha}_n - \theta_n||$ for $\varepsilon = O(\mu)$. 
Average Norm Difference $||\bar{\alpha}_n - \bar{\theta}_n||$ over time for $\varepsilon = O(\mu)$. 
Average Norm Difference $||\tilde{\alpha} - \tilde{\alpha}_n||$ and $||\hat{\alpha} - \hat{\theta}_n||$ over time for $\varepsilon \gg \mu$. 
**NOISE ATTENUATION WITH UNMODELED DYNAMICS**

- We can see mismatched models results in larger deviation from the limit
- Oftentimes, output $y_t$ is a combination of *all* previous input signals $x_t$, e.g.
  \[ y_t = \sum_{j=0}^{\infty} x_{t-j} \alpha_j \]
- However, in practice usually assume a finite model order $n$; that is
  \[ y_t = \sum_{j=0}^{n} x_{t-j} \alpha_j \]
- Difference between actual system order and model order introduces bias
- For tractability, one assumes some bound $\rho_n$ on the unmodeled bias ($\rho_n \rightarrow 0$ as $n$ increases)
LINEAR REGULATOR PROBLEM

The original regulation problem

- $P$: Linear Time-Invariant Plant
  $F$: Feedback Controller
  $d$: output disturbance (stochastic)

- Want to control so output signal $x$ follows constant reference value $x_r$
Sign-Regressor Algorithms

Sign-Error Algorithms

Noise Attenuation

Formulation

- Since system is LTI, we can write

\[ X(z) = \frac{F(z)P(z)}{1 + F(z)P(z)} X_r(z) + \frac{1}{1 + F(z)P(z)} D(z) \]

\[ = U(z) + \frac{1}{1 + F(z)P(z)} D(z) \]

- Writing \( y_k = x_k - x_r \), we have

\[ Y(z) = (U(z) - X_r(z)) + \frac{1}{1 + F(z)P(z)} D(z) \]

- If controller is \( F \) stabilizing, first term goes to zero exponentially fast. For purposes of persistent noise attenuation, ignore first term and write

\[ Y(z) = \frac{1}{1 + F(z)P(z)} D(z) \]
A basic feedback configuration for noise attenuation

- Assume $P(z)$ is exponentially stable and write

$$P(z) = p_0 + p_1 z^{-1} + \cdots + p_n z^{-n} + \delta(z)$$

where $\delta(z) = \sum_{j=n+1}^{\infty} p_j z^{-j}$ and $\sum_{j=n+1}^{\infty} |p_j| \leq \rho_n$

- By exp. stability, $|\rho_n| \leq \kappa \lambda^n$ for some $\kappa > 0$ and $0 < \lambda < 1$
Write:

\[ p = [p_0, \ldots, p_n]' = \text{modeled plant} \]

\[ p^* = [p_{n+1}, p_{n+2}, \ldots]' = \text{unmodeled plant} \]

\[ \psi'_k = [d_k, d_{k-1}, \ldots, d_{k-n}] \]

\[ \tilde{\psi}'_k = [d_{k-(n+1)}, \ldots] \]

\[ w_k = \sum_{j=0}^{\infty} p_j d_{k-j} = \psi'_k p + \tilde{\psi}'_k p^* \]

Taking \( Q = \frac{F}{1+FP} \) and assuming \( Q \) is a stable FIR filter of order \( m \), we can write

\[ y_k = d_k - Q * w_k \]

\[ = d_k - [w_k, w_{k-1}, \ldots, w_{k-m}][q_0, q_1, \ldots, q_m]' \]

\[ = d_k - \phi'_k \theta, \]

where \( \phi'_k = [w_k, w_{k-1}, \ldots, w_{k-m}] \).
A 3.1 (Disturbance and measurement errors)

(1) $d_k$ is estimated by $\hat{d}_k = d_k + e_k$.
$e_k$ is stationary, $\mathbb{E}e_k = 0$, $\mathbb{E}e_k^2 \leq \sigma^2 < \infty$.
(2) The modeled part $p$ is known. The unmodeled dynamics $p^*$ has a uniform norm bound $\rho_n$.

Write

$$\xi'_k = [e_k, e_{k-1}, \ldots, e_{k-n}] \quad \hat{\psi}'_k = \psi'_k + \xi'_k$$
$$\hat{w}_k = \hat{\psi}'_k p \quad \tilde{\varepsilon}_k = -\xi'_k p + \hat{\psi}'_k p^*$$
$$\hat{\phi}'_k = [\hat{w}_k, \hat{w}_{k-1}, \ldots, \hat{w}_{k-m}] \quad \zeta'_k = [\tilde{\varepsilon}_k, \tilde{\varepsilon}_{k-1}, \ldots, \tilde{\varepsilon}_{k-m}]$$

Then

$$y_k = d_k - [w_k, w_{k-1}, \ldots, w_{k-m}] [q_0, q_1, \ldots, q_m]'$$
$$= \hat{d}_k - e_k - [\hat{w}_k + \tilde{\varepsilon}_k, \hat{w}_{k-1} + \tilde{\varepsilon}_{k-1}, \ldots, \hat{w}_{k-m} + \tilde{\varepsilon}_{k-m}] [q_0, q_1, \ldots, q_m]'$$
$$= \hat{d}_k - e_k - \hat{\phi}'_k \theta - \zeta'_k \theta,$$
**Matrix Expansion**

For estimation, after $N$ observations available regression data are

$$\hat{D}_N = \begin{bmatrix} \hat{d}_1 \\ \vdots \\ \hat{d}_N \end{bmatrix}; \quad \hat{\Phi}_N = \begin{bmatrix} \hat{\phi}'_1 \\ \vdots \\ \hat{\phi}'_N \end{bmatrix}. $$

Writing

$$\Phi_N = \begin{bmatrix} \phi'_1 \\ \vdots \\ \phi'_N \end{bmatrix}; \quad \Xi_N = \begin{bmatrix} \zeta'_1 \\ \vdots \\ \zeta'_N \end{bmatrix}; \quad E_N = \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix},$$

so that $\Phi_N = \hat{\Phi}_N + \Xi_N$, we have

$$\gamma_N = \hat{D}_N - E_N - \hat{\Phi}_N \theta - \Xi_N \theta.$$
**Signal Estimation Phase**

- Suppose $d_k$ stationary with annihilating filter $H(z)$ such that $H(z)D(z) \approx 0$

- Then the plant output $v_k$ will be $V(z) = \frac{FP}{1+FPH}HD \approx 0$.

- So subsequent control design should use signal

$$y_k = \hat{d}_k = d_k + e_k$$
After a controller $F$ is (successfully) designed and implemented, the output $y_k = x_k - x_r$ will be small due to the rejection of disturbance by the feedback system.

In this case $y_k$ will have (nearly) no information which can be utilized for the control design.

Use of open-loop control for signal estimation when the plant is stable. $K = P(1)$
**TWO-PHASE APPROACH**

Signal Estimation and Control Design:
- Noises pass through to the output to be estimated.
- Control is designed by an LS-type algorithm.

Two Phase Approach: (1) Signal Estimation and (2) Noise Rejection
To focus on unmodeled dynamics, take $e_k \equiv 0$

Then observation equation is simplified to

$$Y_N = D_N - (\hat{\Phi}_N + \Xi_N)\theta,$$

with

$$\hat{\Phi}_N = \begin{bmatrix} \hat{\phi}_1' \\ \vdots \\ \hat{\phi}_N' \end{bmatrix}, \quad \Xi_N = \begin{bmatrix} \zeta_1' \\ \vdots \\ \zeta_N' \end{bmatrix}$$

and $\hat{\phi}_k' = [\psi_k'p, \psi_{k-1}'p, \ldots, \psi_{k-m}'p]$ and $\zeta_k' = [\tilde{\psi}_k'p^*, \tilde{\psi}_{k-1}'p^*, \ldots, \tilde{\psi}_{k-m}'p^*]$.

$\Gamma \subset \mathbb{R}^{N \times m}$ denotes the uncertainty set for $\Xi_N$ which accommodates all unmodeled dynamics $p^* = \{p_j\}_{j=n+1}^\infty$ with $\sum_{j=n+1}^\infty |p_j| \leq \rho_n$. 
To minimize the mean-square error
\[ \min_{\theta} (D_N - \hat{\Phi}_N \theta)'(D_N - \hat{\Phi}_N \theta), \]
design control parameter
\[ \theta_N \triangleq \left( \hat{\Phi}_N' \hat{\Phi}_N \right)^{-1} \hat{\Phi}_N' D_N. \]

Analyze performance by residual
\[ \mu_N(\Xi_N, D_N) \triangleq \frac{1}{N} (D_N - (\hat{\Phi}_N + \Xi_N)\theta_N)'(D_N - (\hat{\Phi}_N + \Xi_N)\theta_N) \]
\[ \mu_N(D_n) \triangleq \max_{\Xi_N \in \Gamma} \mu(\Xi_N, D_N). \]

A 3.2 (Bounded Disturbance Variance)

The N-sample path of the disturbances \( D_N \) satisfies
\[ D_N \in M_D \triangleq \{ \| D_N / \sqrt{N} \|_2 \leq \sigma^2 \}. \]
Let $\|D_N/\sqrt{N}\|_2 = \lambda$ and define the $v_N \triangleq \frac{D_N/\sqrt{N}}{\lambda}$, so $\hat{\Phi}_N(D_N) = \sqrt{N}\lambda \hat{\Phi}_N(v_N)$.

Denote $\sigma_{\text{min}}$ as the smallest singular value of a matrix and $b_{\text{min}} \triangleq \min_{\|v_N\|_2=1} \sigma_{\text{min}}(\hat{\Phi}_N(v_N))$.

Write $f(\rho_N) \triangleq \max_{\Xi_N \in \Gamma} \frac{\|\Xi_N\|}{\sqrt{N}}$.

**Theorem 3.1 (Worst-Case Performance)**

The worst-case disturbance attenuation performance is given by

$$
\mu \triangleq \max_{D_N \in M_D} \mu_N(D_N) \leq \frac{f(\rho_N)}{b_{\text{min}}}
$$
One might also consider the “min-max” performance

\[
\eta_N(D_N, \theta_N) \triangleq \frac{1}{N} \max_{\Xi_N \in \Gamma} (D_N - (\hat{\Phi}_N + \Xi_N)\theta_N)'(D_N - (\hat{\Phi}_N + \Xi_N)\theta_N)
\]

\[
\eta_N(D_N) \triangleq \min_{\theta_N} \eta_N(D_N, \theta_N)
\]

Note that \( \eta_N(D_N) \leq \mu_N(D_N) \)

“min max” designs often lead to non-linear (even non-quadratic) optimization problems, where only numerical solutions feasible

Use gradient-descent approach:

\[
G(D_N, \theta_N) \triangleq \frac{\partial \eta_N(D_N, \theta_N)}{\partial \theta_N} = \frac{2}{N} \max_{\Xi_N \in \Gamma} (\hat{\Phi}_N + \Xi_N)'(D_N - (\hat{\Phi}_N + \Xi_N)\theta_N).
\]
Algorithm 4 (Two-Phase Algorithm)

The following algorithm searches for $\theta^*_N = \arg \min_{\theta_N} \eta(D_n, \theta_N)$ in two phases:

- **Initial Value.**
  The initial value $\theta^0$ is given by the nominal design
  \[
  \theta^0 = \left( \hat{\Phi}_N \hat{\Phi}_N \right)^{-1} \hat{\Phi}_N D_N.
  \]

- **Iteration Steps.**
  For $k = 0, 1, 2, \ldots$,
  \[
  \theta^{k+1} = \theta^k - \beta_k \hat{G}(D_N, \theta^k)
  \]
  where $\beta_k$ is the step size at the $k$th iteration, $\hat{G}(D_N, \theta^k)$ is an approximate gradient (obtained by using Monte Carlo methods or grid calculation in place of the uncertainty set $\Gamma$).
Impact of Signal Estimation Errors

A 3.3 (Measurement Errors)

1. \(\{d_k\}\) is a sequence of i.i.d. random variables satisfying \(\mathbb{E}d_k = 0\) and \(\mathbb{E}d_k^2 = \sigma_d^2 < \infty\). The fourth moment of \(d_k\) is finite: \(\mathbb{E}d_k^4 < \infty\).

2. \(\{d_k\}\) is estimated by \(\hat{d}_k = d_k + e_k\) such that \(\{e_k\}\) is a sequence of independent and identically distributed (i.i.d.) random variables with \(\mathbb{E}e_k = 0\) and \(\mathbb{E}e_k^2 = \sigma_e^2 < \infty\). \(\{e_k\}\) is independent of \(\{d_k\}\).

3. The modeled part \(p\) is known. The unmodeled dynamics \(p^*\) has a uniform norm bound \(\rho_n\).
**LIMIT WITH MEASUREMENT ERRORS I**

Define nominal design with measurement errors and unmodeled dynamics

\[
\theta^e_n \equiv (\hat{\Phi}'_N \hat{\Phi}_N)^{-1} \hat{\Phi}'_N \hat{D}_N = ((\Phi_N - \Xi_N)'(\Phi_N - \Xi_N))^{-1} (\Phi'_N - \Xi'_N)(D_N + E_N)
\]

Write

\[
P^0_n = \begin{bmatrix}
\sum_{j=0}^{n-|l_2-l_1|} p_j p_j + |l_2-l_1| \\
\end{bmatrix}
\]

Then we can formulate the limit of the estimate \(\theta^e_N\) in terms of \(P^0_n\) as follows.
Limit with Measurement Errors II

Proposition 3.2

Under A3.3, assuming $P_0 \mathbf{n}$ is full rank, we have

$$\theta^e_N = \left[ \hat{\Phi}'_N \hat{\Phi}_N \right]^{-1} \hat{\Phi}'_N \hat{D}_N \xrightarrow{a.s.} \left[ P_0^0 \right]^{-1} \begin{bmatrix} p_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ as } N \to \infty.$$
**LIMIT WITHOUT MEASUREMENT ERRORS I**

Without measurement errors, the estimates are simplified to

\[ \theta_N^0 = (\Phi_N' \Phi_N)^{-1} \Phi_N' D_N \]

Denote

\[ P_n = \left[ \sum_{j=0}^{\infty} p_j p_j + |l_2 - l_1| \right]_{l_1, l_2 = 0, 1, \ldots, n} \]

As before, we can formulate the limit of \( \theta_N^e \) in terms of \( P_n \) as follows.
**Proposition 3.3**

Under A3.3 and assuming $P_n$ is full rank, we have

$$
\theta^0_N = \left[ \Phi'_N \Phi_N \right]^{-1} \Phi_N D_N \xrightarrow{a.s.} [P_n]^{-1} \begin{bmatrix} p_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{as } N \to \infty.
$$
**Theorem 3.4**

Under the assumptions of Propositions 3.2 and 3.3 and assuming that $P_0^n - P_n$ is invertible, we have

$$
\theta^e_N - \theta^0_N \xrightarrow{a.s.} [P_0^n - P_n]^{-1} \begin{bmatrix}
p_0 \\
0 \\
\vdots \\
0
\end{bmatrix}
$$

where

$$
-[P_0^n - P_n]_{l_1,l_2} = \sum_{j=n-|l_2-l_1|+1}^{\infty} p_j p_{j+|l_2-l_1|}.
$$
Defining

\[ \rho_n^{(l)} \triangleq \sum_{j=n+1}^{\infty} p_{j-l} p_j \leq \sum_{j=n+1}^{\infty} |p_j| \leq \rho_n \]

for sufficiently large \( n \), we see that

\[
[P^0_n - P_n]^{-1} = -\begin{bmatrix}
\rho_n^{(0)} & \rho_n^{(1)} & \rho_n^{(2)} & \cdots & \rho_n^{(n)} \\
\rho_n^{(1)} & \rho_n^{(0)} & \rho_n^{(1)} & \cdots & \rho_n^{(n-1)} \\
\rho_n^{(2)} & \rho_n^{(1)} & \rho_n^{(0)} & \cdots & \rho_n^{(n-2)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\rho_n^{(n)} & \rho_n^{(n-1)} & \cdots & \rho_n^{(1)} & \rho_n^{(0)}
\end{bmatrix}^{-1}
\]

for \( l = |l_2 - l_1| \in \{0, 1, \ldots, n\} \).
Example: Nominal Design with Unmodeled Dynamics

- 7th order system \( P(z) = p_0 + p_1 z^{-1} + \cdots + p_7 z^{-7} \)
- 3rd order model: \( P_0(z) = p_0 + p_1 z^{-1} + p_1 z^{-2} + p_3 z^{-3} \)
- Take \( \rho = 0.6 \)
- \( d_k \) i.i.d. uniformly distributed in \([-1, 1]\) (observable);
- \( N = 1000 \) observations
Uncertainty set $\Gamma$ generated by Monte Carlo method. Randomly generate 200 values of $p^*$, normalized so that $|p_4| + |p_5| + |p_6| + |p_7| = 0.6$. Corresponding $\Xi_N$ matrices give $\Gamma$.

Controller has order $m = 20$ ($\theta \in \mathbb{R}^{21}$),

$$\theta_N = \left(\hat{\Phi}_N^\prime \hat{\Phi}_N\right)^{-1} \hat{\Phi}_N^\prime D_N$$

Measure performance by noise-attenuation factor

$$\gamma = \frac{\|Y_N\|_2}{\|D_N\|_2/N}$$

$\gamma < 1$ indicates noise attenuation, the smaller the better

$\rho = 0 \sim \gamma = 0.0148$ (98.5% noise attenuation)

$\rho = 0.6 \sim \gamma = 0.2943$ (70.1% noise attenuation)
Impact of Unmodeled Dynamics

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>0.1</th>
<th>0.3</th>
<th>0.5</th>
<th>0.7</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>0.0570</td>
<td>0.1464</td>
<td>0.2512</td>
<td>0.3459</td>
<td>0.4493</td>
</tr>
<tr>
<td>Reduction</td>
<td>94.3%</td>
<td>85.4%</td>
<td>74.9%</td>
<td>65.4%</td>
<td>55.1%</td>
</tr>
</tbody>
</table>
EXAMPLE: IMPACT OF MEASUREMENT ERRORS

- Plant IIR, $p_k = (0.5)^k$ for $k = 0, 1, \ldots, \infty$

- Model order $n = 10$, Controller order $m = 10$

- Thus $\rho_n = 2 - \sum_{k=0}^{10} p_k = (0.5)^{10} \approx 9.8 \times 10^{-4}$

- Observe estimates for $\theta^e_N$, $\theta^0_N$ for $N = 10, 20, \ldots, 1010$ (100 updates).
Impact of estimation errors given by $||\theta^e_N - \theta^0_N||$, $N = Kn = 10, \ldots, 1010$
Thanks!