Spectral Effects on the Rate of Convergence of the LMS Adaptive Algorithm

Aarón E. Flores
Information Systems Laboratory
Outline

- LMS algorithm and its eigenvalue spread problem
Outline

- LMS algorithm and its eigenvalue spread problem
- Need to predict its performance in practice
Outline

- LMS algorithm and its eigenvalue spread problem
- Need to predict its performance in practice
- LMS Transient Efficiency in terms of spectra
Outline

- LMS algorithm and its eigenvalue spread problem
- Need to predict its performance in practice
- LMS Transient Efficiency in terms of spectra
- LMS Nonstationary Efficiency in terms of spectra
Outline

- LMS algorithm and its eigenvalue spread problem
- Need to predict its performance in practice
- LMS Transient Efficiency in terms of spectra
- LMS Nonstationary Efficiency in terms of spectra
- Application examples and simulations
Outline

- LMS algorithm and its eigenvalue spread problem
- Need to predict its performance in practice
- LMS Transient Efficiency in terms of spectra
- LMS Nonstationary Efficiency in terms of spectra
- Application examples and simulations
- Conclusion
Background and Motivation

- Adaptive Linear Combiner
- Mean Square Error (MSE)
- Minimum MSE Solution
- LMS Algorithm
- Eigenvalue Spread Problem of LMS
- Benchmark: LMS/Newton Algorithm

Learning Curves

LMS Transient Efficiency

LMS Transient Efficiency in Terms of Spectra

Simulation Examples

LMS Nonstationary Efficiency in Terms of Spectra

Nonstationary Simulations

Conclusion
Adaptive Linear Combiner

- Weight vector: \( \mathbf{w}_k = [w_{1k} \ w_{2k} \ \cdots \ w_{Lk}]^T \)
- Input vector: \( \mathbf{x}_k = [x_{1k} \ x_{2k} \ \cdots \ x_{Lk}]^T \)
- Output: \( y_k = \mathbf{x}_k^T \mathbf{w}_k \)
- Error: \( \epsilon_k = d_k - y_k \)
Mean Square Error (MSE)

- Mean Square Error $\xi_k \triangleq E[\epsilon_k^2]$
- Assuming stationarity: $R \triangleq E[x_k x_k^T], \quad p \triangleq E[x_k d_k]$

$$\xi_k = E[d_k^2] - 2p^T w_k + w_k^T R w_k$$
Minimum MSE Solution

- Find fixed $w$ that minimizes $\xi_k$

  MMSE solution: $w^* = R^{-1}p$

- In practice $R$ and $p$ are usually unknown

- Use data samples to iteratively adjust $w_k$
**LMS Algorithm**

- Based on instantaneous gradient descent adaptation

\[ w_{k+1} = w_k - \mu \frac{\partial \varepsilon_k^2}{\partial w_k} \]

**LMS algorithm:**

\[ w_{k+1} = w_k + 2\mu \varepsilon_k x_k \]
LMS Algorithm

- Based on instantaneous gradient descent adaptation

\[ w_{k+1} = w_k - \mu \frac{\partial \epsilon_k^2}{\partial w_k} \]

LMS algorithm:

\[ w_{k+1} = w_k + 2\mu \epsilon_k x_k \]

- Simple and robust \( \rightarrow \) widely used
Eigenvalue Spread Problem of LMS

- Speed of convergence depends on:
  - distribution of eigenvalues of $R$
  - chosen initial condition $w_0$
Eigenvalue Spread Problem of LMS

- Speed of convergence depends on:
  - distribution of eigenvalues of $R$
  - chosen initial condition $w_0$

- In practice: good (or bad) initial conditions not known
Eigenvalue Spread Problem of LMS

- Speed of convergence depends on:
  - distribution of eigenvalues of $R$
  - chosen initial condition $w_0$

- In practice: good (or bad) initial conditions not known

- Need a way to predict LMS speed of convergence in practice
Benchmark: LMS/Newton Algorithm

- LMS/Newton algorithm: LMS-like adaptation using Newton’s method
  \[ w_{k+1} = w_k + 2\mu\lambda_{\text{avg}}R^{-1}\epsilon_k x_k \]

- Uses \( R \) to whiten input \( x_k \), eliminating eigenvalue spread problem

- Ideal, cannot be implemented in practice due to lack of knowledge of \( R \)

- Usually used as theoretical benchmark for adaptive algorithms
Learning Curves
Preliminaries

- $\lambda_i$ are eigenvalues of $R$
- $\lambda_{\text{avg}} \triangleq \frac{\sum_{i=1}^{L} \lambda_i}{L}$
- $\lambda \triangleq [\lambda_1 \lambda_2 \ldots \lambda_L]^T$, $\Lambda \triangleq \text{Diag}(\lambda)$
- Eigendecomposition: $R = Q\Lambda Q^T$, where $Q$ is a unitary matrix with the eigenvectors of $R$ as columns.
- $v_k \triangleq w_k - w^*$, $v'_k \triangleq Q^Tv_k$
- $F_k \triangleq \text{diag}(E[v'_k v'_k^T])$
- $1 \in \mathbb{R}^L$, $1 \triangleq [1 1 \ldots 1]^T$
**MSD and MSE**

- **Mean Square Deviation (MSD):**

\[
\eta_k \triangleq E[\|v_k\|^2] = E[\|v'_k\|^2] = 1^T F_k
\]
Mean Square Deviation (MSD):

\[
\eta_k \triangleq E[\|v_k\|^2] = E[\|v_k'\|^2] = 1^TF_k
\]

Assuming \( \mu \) is small enough, the Mean Square Error (MSE) is given by

\[
\xi_k \triangleq E[\epsilon_k^2] \approx \xi^* + \lambda^TF_k,
\]

where \( \xi^* \) is the MMSE, achieved using \( w^* \)
Dynamics of $F_k$

Assuming $\mu$ is small enough and some other technical conditions, it has been shown that

LMS: $F_{k+1} \approx (I - 4\mu \Lambda)F_k + 4\mu^2 \xi^* \lambda$

$$F_k \approx (I - 4\mu \Lambda)^k(F_0 - \mu \xi^* 1) + \mu \xi^* 1$$
Dynamics of $F_k$

Assuming $\mu$ is small enough and some other technical conditions, it has been shown that

- **LMS**: $F_{k+1} \approx (I - 4\mu\Lambda)F_k + 4\mu^2\xi^*\lambda$

  $$F_k \approx (I - 4\mu\Lambda)^k(F_0 - \mu\xi^*1) + \mu\xi^*1$$

- **LMS/Newton**: $F_{k+1} \approx (1 - 4\mu\lambda_{avg})^kF_k + 4\mu^2\lambda_{avg}^2\xi^*\lambda^{-1}$

  $$F_k \approx (1 - 4\mu\lambda_{avg})^k(F_0 - \mu\lambda_{avg}\xi^*\lambda^{-1}) + \mu\lambda_{avg}\xi^*\lambda^{-1}$$
MSE Learning Curves

Substituting \( F_k \) in \( \xi_k \approx \xi^* + \lambda^T F_k \) we obtain

- **LMS:**
  \[
  \xi_k \approx \xi_\infty + \lambda^T (I - 4\mu \Lambda)^k (F_0 - \mu \xi^* 1)
  \]
MSE Learning Curves

Substituting $F_k$ in $\xi_k \approx \xi^* + \lambda^T F_k$ we obtain

- **LMS:**
  $$\xi_k \approx \xi_\infty + \lambda^T (I - 4\mu \Lambda)^k (F_0 - \mu \xi^* 1)$$

- **LMS/Newton:**
  $$\xi_k \approx \xi_\infty + \lambda^T (1 - 4\mu \lambda_{\text{avg}})^k (F_0 - \mu \lambda_{\text{avg}} \xi^* \lambda^{-1})$$
MSE Learning Curves

Substituting $F_k$ in $\xi_k \approx \xi^* + \lambda^T F_k$ we obtain

- **LMS:**
  \[
  \xi_k \approx \xi_\infty + \lambda^T (I - 4\mu\Lambda)^k (F_0 - \mu\xi^*1)
  \]

- **LMS/Newton:**
  \[
  \xi_k \approx \xi_\infty + \lambda^T (1 - 4\mu\lambda_{\text{avg}})^k (F_0 - \mu\lambda_{\text{avg}}\xi^*\lambda^{-1})
  \]

- For both, LMS and LMS/Newton:
  \[
  \xi_\infty \triangleq \lim_{k \to \infty} \xi_k \approx \xi^*(1 + \mu \text{Tr} (R))
  \]
**MSD Learning Curves**

Substituting $F_k$ in $\eta_k = 1^T F_k$ we obtain

- **LMS:**

  \[
  \eta_k \approx \eta_\infty + 1^T(I - 4\mu \Lambda)^k(F_0 - \mu \xi^* 1)
  \]

  \[
  \eta_\infty \triangleq \lim_{k \to \infty} \eta_k \approx \mu L \xi^*
  \]
Substituting $F_k$ in $\eta_k = 1^T F_k$ we obtain

**LMS:**

$$\eta_k \approx \eta_\infty + 1^T (I - 4\mu \Lambda)^k (F_0 - \mu \xi^* 1)$$

$$\eta_\infty \triangleq \lim_{k \to \infty} \eta_k \approx \mu L \xi^*$$

**LMS/Newton:**

$$\eta_k \approx \eta_\infty + 1^T (1 - 4\mu \lambda_{avg})^k (F_0 - \mu \lambda_{avg} \xi^* \lambda^{-1})$$

$$\eta_\infty \triangleq \lim_{k \to \infty} \eta_k \approx \mu L \xi^* \lambda_{avg} \left(\frac{1}{\lambda}\right)_{avg}$$
MSD Learning Curves

Substituting $F_k$ in $\eta_k = 1^T F_k$ we obtain

**LMS:**

$$
\eta_k \approx \eta_\infty + 1^T (I - 4\mu \Lambda)^k (F_0 - \mu \xi^* 1)
$$

$$
\eta_\infty \triangleq \lim_{k \to \infty} \eta_k \approx \mu L \xi^*
$$

**LMS/Newton:**

$$
\eta_k \approx \eta_\infty + 1^T (1 - 4\mu \lambda_{avg})^k (F_0 - \mu \lambda_{avg} \xi^* \lambda^{-1})
$$

$$
\eta_\infty \triangleq \lim_{k \to \infty} \eta_k \approx \mu L \xi^* \lambda_{avg} \left(\frac{1}{\lambda}\right)_{avg}
$$

**LMS and LMS/Newton have same asymptotic MSD if**

$$
\mu_{\text{LMS}} = \mu_{\text{LMS/Newton}} \lambda_{avg} \left(\frac{1}{\lambda}\right)_{avg} \geq \mu_{\text{LMS/Newton}}
$$
Eigenvalue Spread Problem

![Graph showing learning curves for different algorithms]

- Eigenvalue Spread Problem
- MSE Curve with Random Initial Conditions
- Dynamics of F
- MSE Learning Curves
- MSD Learning Curves
- Preliminaries
- MSD and MSE
- LMS Transient Efficiency
- LMS Transient Efficiency in Terms of Spectra
- Nonstationary Simulations
- Simulation Examples
- LMS Nonstationary Efficiency in Terms of Spectra
- Background and Motivation
- Learning Curves

![Graph with iterations and MSE values for LMS and LMS/Newton]

- LMS (slow mode)
- LMS (fast mode)
- LMS (average)

\[ \xi_\infty \]

\[ \text{MSE} \]

\[ \text{Iterations} \]
MSE Curve with Random Initial Conditions

Typical MSE learning curve averaged over random initial conditions:

\[ E[\xi_k] < \frac{1}{2\mu\lambda_{\text{max}}} \quad \text{and} \quad \frac{1}{2\mu\lambda_{\text{min}}} < k < \frac{1}{2\mu\lambda_{\text{max}}} \]

**Graph:**
- MSE vs. Iterations
- LMS and LMS/Newton trajectories
- Log-log scale for MSE

**Equations:**
- \( E[\xi_k] \)
- \( \frac{1}{2\mu\lambda_{\text{max}}} \)
- \( \frac{1}{2\mu\lambda_{\text{min}}} \)
LMS Transient Efficiency
Area under the Learning Curve (MSE)

- Speed of convergence measure:

\[ J \triangleq \sum_{k=0}^{\infty} \xi_k - \xi_{\infty} \]
Area under the Learning Curve (MSE)

- Speed of convergence measure:

\[ J \triangleq \sum_{k=0}^{\infty} \xi_k - \xi_{\infty} \]

- LMS:

\[ J \approx \frac{v_0^T v_0}{4\mu} - \frac{L\xi^*}{4} \]
Area under the Learning Curve (MSE)

- Speed of convergence measure:
  \[ J \triangleq \sum_{k=0}^{\infty} \xi_k - \xi_{\infty} \]

- LMS:
  \[ J \approx \frac{v_0^T v_0}{4\mu} - \frac{L\xi^*}{4} \]

- LMS/Newton:
  \[ J \approx \frac{v_0^T Rv_0}{4\mu\lambda_{avg}} - \frac{L\xi^*}{4} \]
Area under the Learning Curve (MSD)

- Speed of convergence measure:

\[ D \triangleq \sum_{k=0}^{\infty} \eta_k - \eta_{\infty} \]
Area under the Learning Curve (MSD)

- Speed of convergence measure:
  \[
  D \triangleq \sum_{k=0}^{\infty} \eta_k - \eta_{\infty}
  \]

- LMS:
  \[
  D \approx \frac{v_0^T R^{-1} v_0}{4\mu} - \frac{\text{Tr } R^{-1} \xi^*}{4}
  \]
Area under the Learning Curve (MSD)

- **Speed of convergence measure:**

  \[ D \triangleq \sum_{k=0}^{\infty} \eta_k - \eta_\infty \]

- **LMS:**

  \[ D \approx \frac{v_0^T R^{-1} v_0}{4\mu} - \frac{\text{Tr } R^{-1} \xi^*}{4} \]

- **LMS/Newton:**

  \[ D \approx \frac{v_0^T v_0}{4\mu \lambda_{avg}} - \frac{\text{Tr } R^{-1} \xi^*}{4} \]
LMS Transient Efficiency (MSE)

MSE performance metric:

\[ \text{LMS MSE Transient Efficiency} \triangleq \frac{J_{\text{LMS}/\text{Newton}}}{J_{\text{LMS}}} \]
LMS Transient Efficiency (MSE)

- MSE performance metric:

\[ \text{LMS MSE Transient Efficiency} \triangleq \frac{J_{LMS/Newton}}{J_{LMS}} \]

- Assuming \( v_0^T v_0 \gg \mu L \xi^* \) and \( \frac{v_0^T R v_0}{\lambda_{\text{avg}}} \gg \mu L \xi^* \)

\[ \text{LMS MSE Transient Efficiency} \approx \frac{1}{\lambda_{\text{avg}}} \frac{v_0^T R v_0}{v_0^T v_0} \]
LMS Transient Efficiency (MSE)

- MSE performance metric:

\[
\text{LMS MSE Transient Efficiency} \triangleq \frac{J_{\text{LMS/Newton}}}{J_{\text{LMS}}}
\]

- Assuming \(v_0^T R v_0 \gg \mu L \xi^*\) and \(\frac{v_0^T R v_0}{\lambda_{\text{avg}}} \gg \mu L \xi^*\)

\[
\text{LMS MSE Transient Efficiency} \approx \frac{1}{\lambda_{\text{avg}}} \frac{v_0^T R v_0}{v_0^T v_0}
\]

- Lemma 1. If \(v_0\) and \(R\) are scaled such that \(\lambda_{\text{avg}} = ||v_0|| = 1\), then

\[
\text{LMS MSE Transient Efficiency} \approx v_0^T R v_0
\]
LMS Transient Efficiency (MSD)

- MSD performance metric:

\[
\text{LMS MSD Transient Efficiency} = \frac{D_{LMS}}{D_{LMS/Newton}}
\]
LMS Transient Efficiency (MSD)

- MSD performance metric:

\[
\text{LMS MSD Transient Efficiency} \triangleq \frac{DL_{\text{LMS}}}{DL_{\text{LMS/Newton}}}
\]

- If \( \mu^{\text{LMS}} = \mu^{\text{LMS/Newton}} \lambda_{\text{avg}} \left( \frac{1}{\lambda} \right)_{\text{avg}} \) and assuming

\[
\nu_0^T \nu_0 \gg \mu^{\text{LMS/Newton}} \lambda_{\text{avg}} \text{Tr} \ R^{-1} \xi^* \quad \text{and} \quad \nu_0^T R^{-1} \nu_0 \gg \mu^{\text{LMS}} \text{Tr} \ R^{-1} \xi^*
\]

\[
\text{LMS MSD Transient Efficiency} \approx \frac{1}{\left( \frac{1}{\lambda} \right)_{\text{avg}}} \frac{\nu_0^T R^{-1} \nu_0}{\nu_0^T \nu_0}
\]
LMS Transient Efficiency (MSD)

- MSD performance metric:
  \[ \text{LMS MSD Transient Efficiency} = \frac{D_{LMS}}{D_{LMS/Newton}} \]

- If \( \mu^{LMS} = \mu^{LMS/Newton} \lambda_{avg} \left( \frac{1}{\lambda} \right)_{avg} \) and assuming
  \[ v_0^T v_0 \gg \mu^{LMS/Newton} \lambda_{avg} \text{Tr} R^{-1} \xi^* \]
  \[ v_0^T R^{-1} v_0 \gg \mu^{LMS} \text{Tr} R^{-1} \xi^* \]
  then
  \[ \text{LMS MSD Transient Efficiency} \approx \frac{1}{\left( \frac{1}{\lambda} \right)_{avg}} \frac{v_0^T R^{-1} v_0}{v_0^T v_0} \]

- Lemma 2. If \( v_0 \) and \( R \) are scaled such that \( \left( \frac{1}{\lambda} \right)_{avg} = \| v_0 \| = 1 \), then
  \[ \text{LMS MSD Transient Efficiency} \approx v_0^T R^{-1} v_0 \]
LMS Transient Efficiency in Terms of Spectra
Tapped Delay Line

\[
\mathbf{x}_k = \begin{bmatrix} x_k & x_{k-1} & x_{k-2} & \cdots & x_{k-L+1} \end{bmatrix}^T
\]

\[
\phi_{xx}[n] \triangleq E[x_k x_{k+n}]
\]

\[ R \text{ has Toeplitz structure: } R_{k,l} = \phi_{xx}[k - l] \]
Fourier Spectra

Continuous spectrum:

\[ \Phi_{xx}(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \phi_{xx}[n] e^{-j\omega n} \]

\[ V_0(e^{j\omega}) = \sum_{n=0}^{L-1} v_0[n] e^{-j\omega n} \]
Fourier Spectra

- Continuous spectrum:
  \[ \Phi_{xx}(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \phi_{xx}[n] e^{-j\omega n} \]
  \[ V_0(e^{j\omega}) = \sum_{n=0}^{L-1} v_0[n] e^{-j\omega n} \]

- Discrete spectrum:
  \[ \Phi_{xx}[m] = \sum_{n=-N+1}^{N-1} \phi_{xx}[n] e^{-2\pi j \frac{mn}{M}} \]
  \[ V_0[m] = \frac{1}{\sqrt{M}} \sum_{n=0}^{L-1} v_0[n] e^{-2\pi j \frac{mn}{M}} , \]

where \( m = 0, 1, \ldots, M - 1 \), \( N \geq L \), \( M \geq N + L - 1 \)
**Inner Product Preservation**

- It can be shown that

\[

v_o^T R v_0 = \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}(e^{j\omega}) |V_0(e^{j\omega})|^2 d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m] |V_o[m]|^2

\]

where \( \Phi_{xx}(e^{j\omega}) \) is the Fourier transform of the cross-correlation function, and \( V_0(e^{j\omega}) \) is the Fourier transform of the input signal.
Inner Product Preservation

- It can be shown that

\[ \mathbf{v}_o^T \mathbf{R} \mathbf{v}_0 = \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}(e^{j\omega}) |\mathbf{V}_0(e^{j\omega})|^2 d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m] |\mathbf{V}_0[m]|^2 \]

- and \( \lambda_{\text{avg}} = \| \mathbf{v}_0 \| = 1 \) implies

\[ \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}(e^{j\omega}) d\omega = \frac{1}{2\pi} \int_0^{2\pi} |\mathbf{V}_0(e^{j\omega})|^2 d\omega = 1 \]

\[ \frac{1}{M} \sum_{m=0}^{M-1} \Phi_{xx}[m] = \sum_{m=0}^{M-1} |\mathbf{V}_0[m]|^2 = 1 \]
**LMS Transient Efficiency (MSE)**

**Theorem 1.**

\[
\frac{J_{LMS/Newton}}{J_{LMS}} \approx \frac{1}{2\pi} \int_{0}^{2\pi} \Phi_{xx}(e^{j\omega}) |V_{0}(e^{j\omega})|^2 d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m] |V_{0}[m]|^2
\]
**LMS Transient Efficiency (MSE)**

- **Theorem 1.**

\[
\frac{J_{LMS/Newton}}{J_{LMS}} \approx \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}(e^{j\omega}) |V_0(e^{j\omega})|^2 d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m] |V_0[m]|^2
\]

- **Corollary:** \(w_0 = 0 \implies v_o = -w^*\)

\[
\frac{J_{LMS/Newton}}{J_{LMS}} \approx \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}(e^{j\omega}) |W^*(e^{j\omega})|^2 d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m] |W^*[m]|^2
\]
**LMS Transient Efficiency (MSE)**

- **Theorem 1.**

\[
\frac{J_{LMS/Newton}}{J_{LMS}} \approx \frac{1}{2\pi} \int_{0}^{2\pi} \Phi_{xx}(e^{j\omega})|V_{0}(e^{j\omega})|^2 d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m]|V_{0}[m]|^2
\]

- **Corollary:** \(w_{0} = 0 \Rightarrow v_{0} = -w^{*}\)

\[
\frac{J_{LMS/Newton}}{J_{LMS}} \approx \frac{1}{2\pi} \int_{0}^{2\pi} \Phi_{xx}(e^{j\omega})|W^{*}(e^{j\omega})|^2 d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m]|W^{*}[m]|^2
\]

- **LMS MSE Transient Efficiency** is equal to the weighted average of the input psd using the Wiener solution spectrum as the weighting function.
LMS Transient Efficiency (MSD)

Theorem 2. For sufficiently large $L$

$$\frac{D^{LMS}}{D^{LMS/Newton}} \approx \frac{1}{2\pi} \int_{0}^{2\pi} \Phi_{xx}^{-1}(e^{j\omega})|V_0(e^{j\omega})|^2 d\omega \approx \sum_{m=0}^{M-1} \Phi_{xx}^{-1}[m]|V_0[m]|^2$$
LMS Transient Efficiency (MSD)

- **Theorem 2.** For sufficiently large $L$

\[
\frac{D_{LMS}}{D_{LMS/Newton}} \approx \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}^{-1}(e^{j\omega}) |V_0(e^{j\omega})|^2 d\omega \approx \sum_{m=0}^{M-1} \Phi_{xx}^{-1}[m] |V_0[m]|^2
\]

- **Corollary:** $w_0 = 0 \implies v_o = -w^*$

\[
\frac{D_{LMS}}{D_{LMS/Newton}} \approx \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}^{-1}(e^{j\omega}) |W^*(e^{j\omega})|^2 d\omega \approx \sum_{m=0}^{M-1} \Phi_{xx}^{-1}[m] |W^*[m]|^2
\]
LMS Transient Efficiency (MSD)

- **Theorem 2.** For sufficiently large $L$

  \[
  \frac{D_{\text{LMS}}}{D_{\text{LMS/Newton}}} \approx \frac{1}{2\pi} \int_{0}^{2\pi} \Phi^{-1}_{xx}(e^{j\omega}) |V_0(e^{j\omega})|^2 d\omega \approx \sum_{m=0}^{M-1} \Phi^{-1}_{xx}[m] |V_0[m]|^2
  \]

- **Corollary:** $w_0 = 0 \implies v_0 = -w^*$

  \[
  \frac{D_{\text{LMS}}}{D_{\text{LMS/Newton}}} \approx \frac{1}{2\pi} \int_{0}^{2\pi} \Phi^{-1}_{xx}(e^{j\omega}) |W^*(e^{j\omega})|^2 d\omega \approx \sum_{m=0}^{M-1} \Phi^{-1}_{xx}[m] |W^*[m]|^2
  \]

- LMS MSD Transient Efficiency is equal to the weighted average of the reciprocal of the input psd using the Wiener solution spectrum as the weighting function.
Simulation Examples

- System Identification Example
- System ID Simulation 1
- Equalization Example
- Equalization Simulation

LMS Nonstationary Efficiency in Terms of Spectra

Nonstationary Simulations

Conclusion
System Identification Example

- FIR filter and adaptive filter have 16 taps
- Plant noise is independent of input signal
System ID Simulation 1

- Wiener solution spectrum = plant frequency response
- LMS MSE Transient Efficiency from theory = 2.069
- LMS MSE Transient Efficiency from simulation = 2.064

Spectra

- Input psd
- Plant Frequency Response
- Eigenspread = 5709

MSE Learning Curves

- LMS/Newton
- LMS

Background and Motivation
Learning Curves
LMS Transient Efficiency
LMS Transient Efficiency in Terms of Spectra
Simulation Examples
- System Identification Example
- System ID Simulation 1
- Equalization Example
- Equalization Simulation
LMS Nonstationary Efficiency in Terms of Spectra
Nonstationary Simulations
Conclusion
System ID Simulation 2

- Wiener solution spectrum = plant frequency response
- LMS MSE Transient Efficiency from theory = 1.728
- LMS MSE Transient Efficiency from simulation = 1.716

Spectra

Eigenspread = 52.3

MSE Learning Curves

Input psd

Plant Frequency Response

\[ \Phi_{xx}(e^{j\omega}) \]

\[ |W^*(e^{j\omega})|^2 \]

\[ 10^{-1} \quad 10^0 \quad 10^1 \quad 10^2 \quad 10^3 \quad 10^4 \]

\[ \xi_k^2 \]

Iterations

LMS

LMS/Newton

Department of Electrical Engineering
System ID Simulation 2 cont... 

- LMS and LMS/Newton have same asymptotic MSD
- LMS MSD Transient Efficiency from theory = 0.231
- LMS MSD Transient Efficiency from simulation = 0.251

MSE Learning Curves

![MSE Learning Curves Graph]

MSD Learning Curves

![MSD Learning Curves Graph]
System ID Simulation 3

- Wiener solution spectrum = plant frequency response
- LMS MSE Transient Efficiency from theory = 0.227
- LMS MSE Transient Efficiency from simulation = 0.222
System ID Simulation 3 cont. . .

- LMS and LMS/Newton have same asymptotic MSD
- LMS MSD Transient Efficiency from theory = 2.745
- LMS MSD Transient Efficiency from simulation = 2.676

MSE Learning Curves

<table>
<thead>
<tr>
<th>Iterations</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>10^3</td>
</tr>
<tr>
<td>400</td>
<td>10^2</td>
</tr>
<tr>
<td>600</td>
<td>10^1</td>
</tr>
<tr>
<td>800</td>
<td>10^0</td>
</tr>
<tr>
<td>1000</td>
<td>10^-1</td>
</tr>
<tr>
<td>1200</td>
<td>10^-2</td>
</tr>
<tr>
<td>1400</td>
<td>10^-3</td>
</tr>
<tr>
<td>1600</td>
<td>10^-4</td>
</tr>
<tr>
<td>1800</td>
<td>10^-5</td>
</tr>
<tr>
<td>2000</td>
<td>10^-6</td>
</tr>
</tbody>
</table>

MSD Learning Curves

<table>
<thead>
<tr>
<th>Iterations</th>
<th>MSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>10^2</td>
</tr>
<tr>
<td>1000</td>
<td>10^1</td>
</tr>
<tr>
<td>1500</td>
<td>10^0</td>
</tr>
<tr>
<td>2000</td>
<td>10^-1</td>
</tr>
<tr>
<td>2500</td>
<td>10^-2</td>
</tr>
<tr>
<td>3000</td>
<td>10^-3</td>
</tr>
</tbody>
</table>

Equalization Example

- Channel is a 4 taps FIR filter and adaptive equalizer has 16 taps
- Input signal is white \(\rightarrow\) psd of input to adaptive filter is magnitude squared of channel frequency response
Equalization Simulation

- Wiener solution spectrum \(\approx\) Inverse of plant frequency response
- LMS MSE Transient Efficiency from theory = 0.137
- LMS MSE Transient Efficiency from simulation = 0.132

![Spectra](image)

- Input psd
- Wiener Solution Spectrum

![MSE Learning Curves](image)

- \(\xi_k\)
- LMS
- LMS/Newton

DEPARTMENT OF ELECTRICAL ENGINEERING
LMS Nonstationary Efficiency in Terms of Spectra

LMS Nonstationary Efficiency in Terms of Spectra

- Nonstationarity Model
- LMS Nonstationary Efficiency
- Average Spectrum of Wiener Solution Steps
- LMS Nonst. Efficiency in terms of Spectra

Nonstationary Simulations

Conclusion
Nonstationarity Model

- First-order Markov process on Wiener solution

\[ w_{k+1}^* = w_k^* + \gamma_k, \]

where \( \gamma_k \) is a white zero mean stationary random process.
Nonstationarity Model

- First-order Markov process on Wiener solution

\[ w_{k+1}^* = w_k^* + \gamma_k, \]

where \( \gamma_k \) is a white zero mean stationary random process.

- Assume the power of \( \gamma_k \) to be small so that time variations of \( w_k^* \) are slow.
Nonstationarity Model

- First-order Markov process on Wiener solution

\[ w_{k+1}^* = w_k^* + \gamma_k, \]

where \( \gamma_k \) is a white zero mean stationary random process.

- Assume the power of \( \gamma_k \) to be small so that time variations of \( w_k^* \) are slow.

- For each, LMS and LMS/Newton, there will be an optimum value of \( \mu \) that minimizes the steady state MSE, and another value of \( \mu \) that minimizes the steady state MSD. We denote such optimum MSE and MSD by \( \xi_\infty \) and \( \eta_\infty \) respectively.
LMS Nonstationary Efficiency

- MSE performance metric

\[
\left( \frac{\xi_{LMS/Newton} - \xi^*}{\xi_{\infty} - \xi^*} \right)^2 = E[\gamma_k^T R \gamma_k],
\]

where \( R \) and \( \gamma_k \) are scaled s.t. \( \lambda_{\text{avg}} = E[\|\gamma_k\|^2] = 1 \)
LMS Nonstationary Efficiency

- **MSE performance metric**

\[
\left( \frac{\xi_{\infty}^{\text{LMS/Newton}} - \xi^*}{\xi_{\infty}^{\text{LMS}} - \xi^*} \right)^2 = E[\gamma_k^T R \gamma_k],
\]

where \( R \) and \( \gamma_k \) are scaled s.t. \( \lambda_{\text{avg}} = E[\|\gamma_k\|^2] = 1 \)

- **MSD performance metric**

\[
\left( \frac{\eta_{\infty}^{\text{LMS}}}{\eta_{\infty}^{\text{LMS/Newton}}} \right)^2 = E[\gamma_k^T R^{-1} \gamma_k],
\]

where \( R \) and \( \gamma_k \) are scaled s.t. \( \left( \frac{1}{\lambda} \right)_{\text{avg}} = E[\|\gamma_k\|^2] = 1 \)

Department of Electrical Engineering

Stanford University
Continuous spectrum:

\[ \Gamma(e^{j\omega}) \triangleq E \left[ \left| \sum_{n=0}^{L-1} \gamma_k[n] e^{-j\omega_n} \right|^2 \right] \]
Average Spectrum of Wiener Solution Steps

- **Continuous spectrum:**

\[
\Gamma(e^{j\omega}) \triangleq E \left[ \left| \sum_{n=0}^{L-1} \gamma_k[n] e^{-j\omega n} \right|^2 \right]
\]

- **Discrete spectrum:**

\[
\Gamma[m] \triangleq \frac{1}{M} \Gamma(e^{j\omega}) \bigg|_{\omega=\frac{2\pi m}{M}},
\]

where \(m = 0, 1, \ldots M - 1\), \(N \geq L\), \(M \geq N + L - 1\)
Average Spectrum of Wiener Solution Steps

- **Continuous spectrum:**
  \[
  \Gamma(e^{j\omega}) \triangleq E \left[ \left| \sum_{n=0}^{L-1} \gamma_k[n] e^{-j\omega n} \right|^2 \right]
  \]

- **Discrete spectrum:**
  \[
  \Gamma[m] \triangleq \frac{1}{M} \Gamma(e^{j\omega}) \bigg|_{\omega = \frac{2\pi m}{M}},
  \]
  where \( m = 0, 1, \ldots, M - 1 \), \( N \geq L \), \( M \geq N + L - 1 \)

**Note:** \( E[\|\gamma_k\|^2] = 1 \) implies
\[
\frac{1}{2\pi} \int_0^{2\pi} \Gamma(e^{j\omega}) d\omega = \sum_{m=0}^{M-1} \Gamma[m] = 1
\]
Theorem 3.

\[
\left( \frac{\xi_{\text{LMS/Newton}} - \xi^*}{\xi_{\infty} - \xi^*} \right)^2 \approx \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}(e^{j\omega}) \Gamma(e^{j\omega}) d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m] \Gamma[m]
\]
**LMS Nonst. Efficiency in terms of Spectra**

- **Theorem 3.**

\[
\left( \frac{\xi_{\text{LMS/Newton}} - \xi^*}{\xi_\infty - \xi^*} \right)^2 \approx \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}(e^{j\omega}) \Gamma(e^{j\omega}) d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m] \Gamma[m]
\]

- **Theorem 4. For sufficiently large \(L\)**

\[
\left( \frac{\eta_{\text{LMS}}}{\eta_{\text{LMS/Newton}}} \right)^2 \approx \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}^{-1}(e^{j\omega}) \Gamma(e^{j\omega}) d\omega \approx \sum_{m=0}^{M-1} \Phi_{xx}^{-1}[m] \Gamma[m]
\]
LMS Nonst. Efficiency in terms of Spectra

- **Theorem 3.**

\[
\left( \frac{\xi_{\infty}^{\text{LMS/Newton}} - \xi^*}{\xi_{\infty}^{\text{LMS}} - \xi^*} \right)^2 \approx \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}(e^{j\omega}) \Gamma(e^{j\omega}) d\omega = \sum_{m=0}^{M-1} \Phi_{xx}[m] \Gamma[m]
\]

- **Theorem 4.** For sufficiently large \( L \)

\[
\left( \frac{\eta_{\infty}^{\text{LMS}}}{\eta_{\infty}^{\text{LMS/Newton}}} \right)^2 \approx \frac{1}{2\pi} \int_0^{2\pi} \Phi_{xx}^{-1}(e^{j\omega}) \Gamma(e^{j\omega}) d\omega \approx \sum_{m=0}^{M-1} \Phi_{xx}^{-1}[m] \Gamma[m]
\]

- **LMS Nontationary efficiency is determined by the weighted average of the input psd (or its reciprocal) using the average spectrum of the Wiener solution steps as the weighting function.**
Nonstationary Simulations
Nonstationary Simulation 1

- \( \mu \)'s for LMS and LMS/Newton are optimized in MSE sense
- LMS MSE Nonstationary Efficiency from theory = 1.312
- LMS MSE Nonstationary Efficiency from simulation = 1.332
Nonstationary Simulation 1 cont . . .

- $\mu$'s for LMS and LMS/Newton are optimized in MSD sense
- LMS MSD Nonstationary Efficiency from theory $= 0.5074$
- LMS MSD Nonstationary Efficiency from simulation $= 0.5005$

Spectra

Input psd

Average Spectrum of $\gamma_k$

MSD Learning Curves

Eigenspread = 39.1
Nonstationary Simulation 2

- μ’s for LMS and LMS/Newton are optimized in MSE sense
- LMS MSE Nonstationary Efficiency from theory = 0.576
- LMS MSE Nonstationary Efficiency from simulation = 0.569

Spectra

Eigenspread = 25.7

Input psd

Average Spectrum of γ_k

MSE Learning Curves

<table>
<thead>
<tr>
<th>Iterations</th>
<th>LMS</th>
<th>LMS/Newton</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.105</td>
<td>0.115</td>
</tr>
<tr>
<td>100</td>
<td>0.11</td>
<td>0.115</td>
</tr>
<tr>
<td>200</td>
<td>0.115</td>
<td>0.115</td>
</tr>
<tr>
<td>300</td>
<td>0.115</td>
<td>0.115</td>
</tr>
<tr>
<td>400</td>
<td>0.115</td>
<td>0.115</td>
</tr>
<tr>
<td>500</td>
<td>0.115</td>
<td>0.115</td>
</tr>
<tr>
<td>600</td>
<td>0.115</td>
<td>0.115</td>
</tr>
<tr>
<td>700</td>
<td>0.115</td>
<td>0.115</td>
</tr>
<tr>
<td>800</td>
<td>0.115</td>
<td>0.115</td>
</tr>
<tr>
<td>900</td>
<td>0.115</td>
<td>0.115</td>
</tr>
<tr>
<td>1000</td>
<td>0.115</td>
<td>0.115</td>
</tr>
</tbody>
</table>
Nonstationary Simulation 2 cont . . .

- \( \mu \)'s for LMS and LMS/Newton are optimized in MSD sense
- LMS MSD Nonstationary Efficiency from theory = 1.272
- LMS MSD Nonstationary Efficiency from simulation = 1.295

Spectra

**Input psd**

**Average Spectrum of** \( \gamma_k \)

**MSD Learning Curves**

- LMS
- LMS/Newton

---

**Department of Electrical Engineering**

Stanford University

Oral's Presentation - p. 47/69
Conclusion
Contributions

- LMS speed of convergence is determined by similarity between input psd and spectrum of the Wiener solution, when using zero initial conditions.
Contributions

- LMS speed of convergence is determined by similarity between input psd and spectrum of the Wiener solution, when using zero initial conditions.

- LMS steady state performance in nonstationary environment is determined by similarity between input psd and spectrum of the Wiener solution variations.
Contributions

- LMS speed of convergence is determined by similarity between input psd and spectrum of the Wiener solution, when using zero initial conditions.

- LMS steady state performance in nonstationary environment is determined by similarity between input psd and spectrum of the Wiener solution variations.

- Future work:
  - Apply similar analysis to nonstationary input signal conditions.
Acknowledgments

- Prof. Bernard Widrow, Research Advisor
- Prof. Boris Murmann, Committee Chairman
- Prof. Tom Cover, Associate Research Advisor
- Prof. Julius Smith, Committee Member
- Zoo-mates: Gabriel, Jeff, Jingyan, Juan Carlos, Marcelo, Max, Ngin-Choo, Oscar, Takeshi
- Staff: Denise, Diane, Joice, Marianne
- Friends: Alejandro, Alfredo, Arvindh, Candy, Carlos, Debbie, Eduardo, Edgar, Gabino, Grace, Hector, Jaime, Jorge, Jose, Juan, Kerstin, Miguel, Paul, Pollo, Persa, Susi, Teresa, Zazu
- My Parents: Ramona and Aurelio
- Judith and Silvia
Acknowledgments

- Prof. Bernard Widrow, Research Advisor
- Prof. Boris Murmann, Committee Chairman
- Prof. Tom Cover, Associate Research Advisor
- Prof. Julius Smith, Committee Member
- Zoo-mates: Gabriel, Jeff, Jingyan, Juan Carlos, Marcelo, Max, Ngin-Choo, Oscar, Takeshi
- Staff: Denise, Diane, Joice, Marianne
- Friends: Alejandro, Alfredo, Arvindh, Candy, Carlos, Debbie, Eduardo, Edgar, Gabino, Grace, Hector, Jaime, Jorge, Jose, Juan, Kerstin, Miguel, Paul, Pollo, Persa, Susi, Teresa, Zazu
- My Parents: Ramona and Aurelio
- Judith and Silvia
Acknowledgments

- Prof. Bernard Widrow, Research Advisor
- Prof. Boris Murmann, Committee Chairman
- Prof. Tom Cover, Associate Research Advisor
- Prof. Julius Smith, Committee Member
- Zoo-mates: Gabriel, Jeff, Jingyan, Juan Carlos, Marcelo, Max, Ngin-Choo, Oscar, Takeshi
- Staff: Denise, Diane, Joice, Marianne
- Friends: Alejandro, Alfredo, Arvindh, Candy, Carlos, Debbie, Eduardo, Edgar, Gabino, Grace, Hector, Jaime, Jorge, Jose, Juan, Kerstin, Miguel, Paul, Pollo, Persa, Susi, Teresa, Zazu
- My Parents: Ramona and Aurelio
- Judith and Silvia
Acknowledgments

- Prof. Bernard Widrow, Research Advisor
- Prof. Boris Murmann, Committee Chairman
- Prof. Tom Cover, Associate Research Advisor
- Prof. Julius Smith, Committee Member
- Zoo-mates: Gabriel, Jeff, Jingyan, Juan Carlos, Marcelo, Max, Ngin-Choo, Oscar, Takeshi
- Staff: Denise, Diane, Joice, Marianne
- Friends: Alejandro, Alfredo, Arvindh, Candy, Carlos, Debbie, Eduardo, Edgar, Gabino, Grace, Hector, Jaime, Jorge, Jose, Juan, Kerstin, Miguel, Paul, Pollo, Persa, Susi, Teresa, Zazu
- My Parents: Ramona and Aurelio
- Judith and Silvia
Acknowledgments

- Prof. Bernard Widrow, Research Advisor
- Prof. Boris Murmann, Committee Chairman
- Prof. Tom Cover, Associate Research Advisor
- Prof. Julius Smith, Committee Member
- Zoo-mates: Gabriel, Jeff, Jingyan, Juan Carlos, Marcelo, Max, Ngin-Choo, Oscar, Takeshi
- Staff: Denise, Diane, Joice, Marianne
- Friends: Alejandro, Alfredo, Arvindh, Candy, Carlos, Debbie, Eduardo, Edgar, Gabino, Grace, Hector, Jaime, Jorge, Jose, Juan, Kerstin, Miguel, Paul, Pollo, Persa, Susi, Teresa, Zazu
- My Parents: Ramona and Aurelio
- Judith and Silvia
System ID Simulation 3 cont . . .

- LMS2 and LMS/Newton have same asymptotic MSD
- LMS MSD Transient Efficiency from theory = 0.231
- LMS MSD Transient Efficiency from simulation = 0.251
System ID Simulation 4 cont . . .

- LMS2 and LMS/Newton have same asymptotic MSD
- LMS MSD Transient Efficiency from theory = 2.745
- LMS MSD Transient Efficiency from simulation = 2.676

MSE Learning Curves

MSD Learning Curves
**System ID Simulation 2**

- Wiener solution spectrum = plant frequency response
- LMS MSE Transient Efficiency from theory = 1.413
- LMS MSE Transient Efficiency from simulation = 1.421

**Spectra**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>0</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
<th>3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Φ_{xx}(e^{jω})</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1.5</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

**Input psd**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>0</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
<th>3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

**Plant Frequency Response**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>0</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
<th>3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.015</td>
<td>0.01</td>
<td>0.005</td>
<td>0.01</td>
<td>0.005</td>
<td>0.01</td>
<td>0.015</td>
<td>0.005</td>
</tr>
</tbody>
</table>

**MSE Learning Curves**

<table>
<thead>
<tr>
<th>Iterations</th>
<th>100</th>
<th>200</th>
<th>300</th>
<th>400</th>
<th>500</th>
<th>600</th>
<th>700</th>
<th>800</th>
<th>900</th>
<th>1000</th>
<th>1100</th>
<th>1200</th>
</tr>
</thead>
<tbody>
<tr>
<td>ξ_k</td>
<td>10^{-1}</td>
<td>10^{-2}</td>
<td>10^{-3}</td>
<td>10^{-4}</td>
<td>10^{-5}</td>
<td>10^{-6}</td>
<td>10^{-7}</td>
<td>10^{-8}</td>
<td>10^{-9}</td>
<td>10^{-10}</td>
<td>10^{-11}</td>
<td>10^{-12}</td>
</tr>
</tbody>
</table>

- LMS/Newton
- LMS
Define $\Phi_{xx} \in \mathbb{R}^M$ a vector with entries

$$
\Phi_{xx}[m] = \sum_{n=-N+1}^{N-1} \phi_{xx}[n] e^{-2\pi j \frac{m n}{M}},
$$

$m = 0, 1, \ldots, M - 1, \quad N \geq L, \quad M \geq N + L - 1$
Define $\Phi_{xx} \in \mathbb{R}^M$ a vector with entries

$$\Phi_{xx}[m] = \sum_{n=-N+1}^{N-1} \phi_{xx}[n] e^{-2\pi j \frac{mn}{M}},$$

$m = 0, 1, \ldots, M - 1$, $N \geq L$, $M \geq N + L - 1$

Since $R$ is Toeplitz we can factor $R$ and $R^{-1}$ as

$$R = U \Psi U^* \quad R^{-1} \approx U \Psi^{-1} U^*$$

where $\Psi = \text{Diag}(\Phi_{xx})$ and $U$ is a $L \times M$ matrix with entries

$$U_{l,m} = \frac{1}{\sqrt{M}} e^{2\pi j \frac{ml}{M}}$$
Let $V_0 \triangleq U^* v_0$ be the $M$-point DFT of $v_0$, i.e.

$$V_0[m] = \frac{1}{\sqrt{M}} \sum_{n=0}^{L-1} v_0[n] e^{-2\pi j \frac{mn}{M}} \quad m = 0, 1, \ldots, M - 1$$
Spectrum of Initial Weight Deviation

Let $V_0 \triangleq U^* v_0$ be the $M$-point DFT of $v_0$, i.e.

$$V_0[m] = \frac{1}{\sqrt{M}} \sum_{n=0}^{L-1} v_0[n] e^{-2\pi j \frac{mn}{M}} \quad m = 0, 1, \ldots, M - 1$$

Using the spectral factorization of $R$

$$v_o^T R v_0 = v_o^T U \Psi U^* v_0 = V_0^* \Psi V_0 = \sum_{m=0}^{M-1} \Phi_{xx}[m] |V_0[m]|^2$$
Spectrum of Initial Weight Deviation

- Let $V_0 \triangleq U^*v_0$ be the $M$-point DFT of $v_0$, i.e.

$$V_0[m] = \frac{1}{\sqrt{M}} \sum_{n=0}^{L-1} v_0[n] e^{-2\pi j \frac{mn}{M}} \quad m = 0, 1, \ldots, M - 1$$

- Using the spectral factorization of $R$

$$v_0^T R v_0 = v_0^T U \Psi U^* v_0 = V_0^* \Psi V_0 = \sum_{m=0}^{M-1} \Phi_{xx}[m]|V_0[m]|^2$$

- Note: $\lambda_{avg} = 1 = \|v_0\| = 1$ implies

$$\frac{1}{M} \sum_{m=0}^{M-1} \Phi_{xx}[m] = \sum_{m=0}^{M-1} |V_0[m]|^2 = 1$$
Average Spectrum of $\nu_k$: Simulation 1

**Background and Motivation**

**Learning Curves**

**LMS Transient Efficiency**

**LMS Transient Efficiency in Terms of Spectra**

**Simulation Examples**

**LMS Nonstationary Efficiency in Terms of Spectra**

**Nonstationary Simulations**

**Conclusion**

**Contributions**

---

**Department of Electrical Engineering**

---

**Oral's Presentation - p. 56/69**
Average Spectrum of $v_k$, Simulation 1

$E[|V[m]|^2]$ vs. Frequency $m \frac{\pi}{100}$

Iteration $k = 3$

- LMS
- LMS/Newton
Average Spectrum of $v_k$ Simulation 1

![Graph showing average spectrum of $v_k$ for Simulation 1 with iterations and frequency axes labeled.]
Average Spectrum of $u_k$ Simulation 1
Average Spectrum of $v_k$ Simulation 1

**Iteration $k = 100$**

$E[V[m]^2]$

- Blue line: LMS
- Red line: LMS/Newton

**Frequency** $m \frac{\pi}{100}$

**Contributions**
Average Spectrum of $\nu_k$ Simulation 1

Iteration $k = 500$

![Graph showing the average spectrum of $\nu_k$ for Simulation 1 with iterations up to k = 500. The graph compares LMS and LMS/Newton methods.]
Average Spectrum of $\nu_k$ Simulation 1

![Graph showing the spectrum of $\nu_k$ Simulation 1 with iterations and frequency dimensions.](image)
Average Spectrum of $u_k$ Simulation 3

![Graph showing average spectrum for Simulation 3 with iterations and frequency]

- Background and Motivation
- Learning Curves
- LMS Transient Efficiency
- LMS Transient Efficiency in Terms of Spectra
- Simulation Examples
- LMS Nonstationary Efficiency in Terms of Spectra
- Nonstationary Simulations
- Conclusion
- Contributions
Average Spectrum of $v_k$ Simulation 3

Iteration $k = 50$

- LMS
- LMS/Newton

Frequency $m \frac{\pi}{100}$
Average Spectrum of $v_k$. Simulation 3

Iteration $k = 100$

![Average Spectrum of $v_k$. Simulation 3](chart.png)
Average Spectrum of $\nu_k$ Simulation 3

![Graph showing the average spectrum of $\nu_k$ simulation 3 with iterations and frequency axes.](image)
Average Spectrum of $v_k$ Simulation 3

**Diagram:**
- **Label:** $E[|V[m]|^2]$
- **X-axis:** Frequency $m \cdot \frac{\pi}{100}$
- **Y-axis:** $10^{-3}$ to $10^{-4}$

**Legend:**
- LMS
- LMS/Newton

**Graph Description:**
- The graph compares the average spectrum of $v_k$ for Simulation 3 at iteration $k = 1000$.
- The spectrum is represented for both LMS and LMS/Newton methods.
- The frequencies are plotted on the x-axis, ranging from 10 to 100.
- The y-axis represents the power spectral density, ranging from $10^{-3}$ to $10^{-4}$.

**Contributions:**
- [List of contributions]

---

**Department of Electrical Engineering**

---

Oral's Presentation - p. 67/69
Average Spectrum of $\nu_k$ Simulation 3
Average Spectrum of $\nu_k$ Simulation 3

![Diagram showing the spectrum](image)s