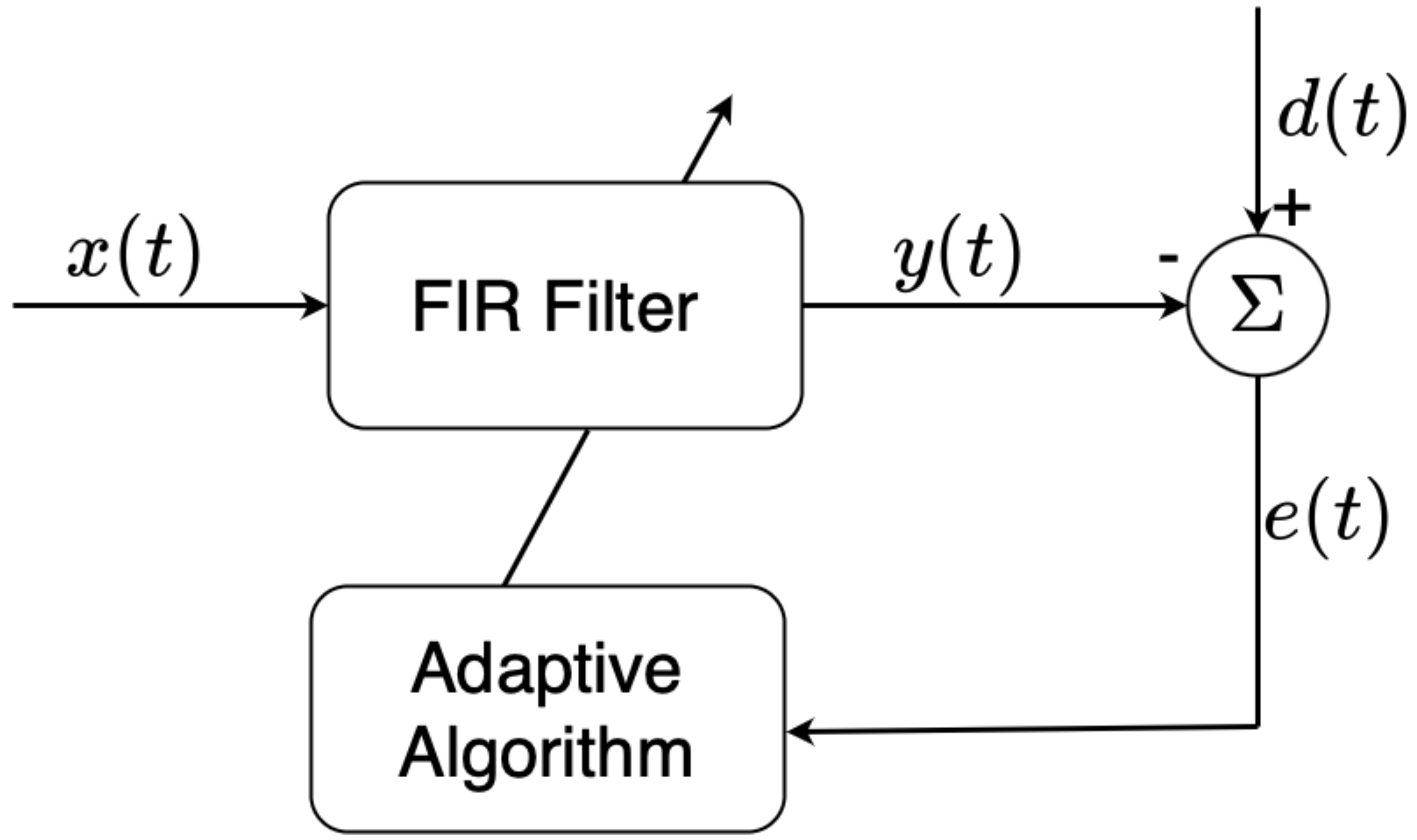
STATE SPACE APPROACH TO ADAPTIVE FILTERS

KAUSHIK IYER (1223696175)
MS RAS (EE)
EEE 606

MOTIVATION



No knowledge of system used

No provision to account for noise in the adaptive algorithm

Dynamic gain (or adaptation) is generally a direct scaling of error or input signal

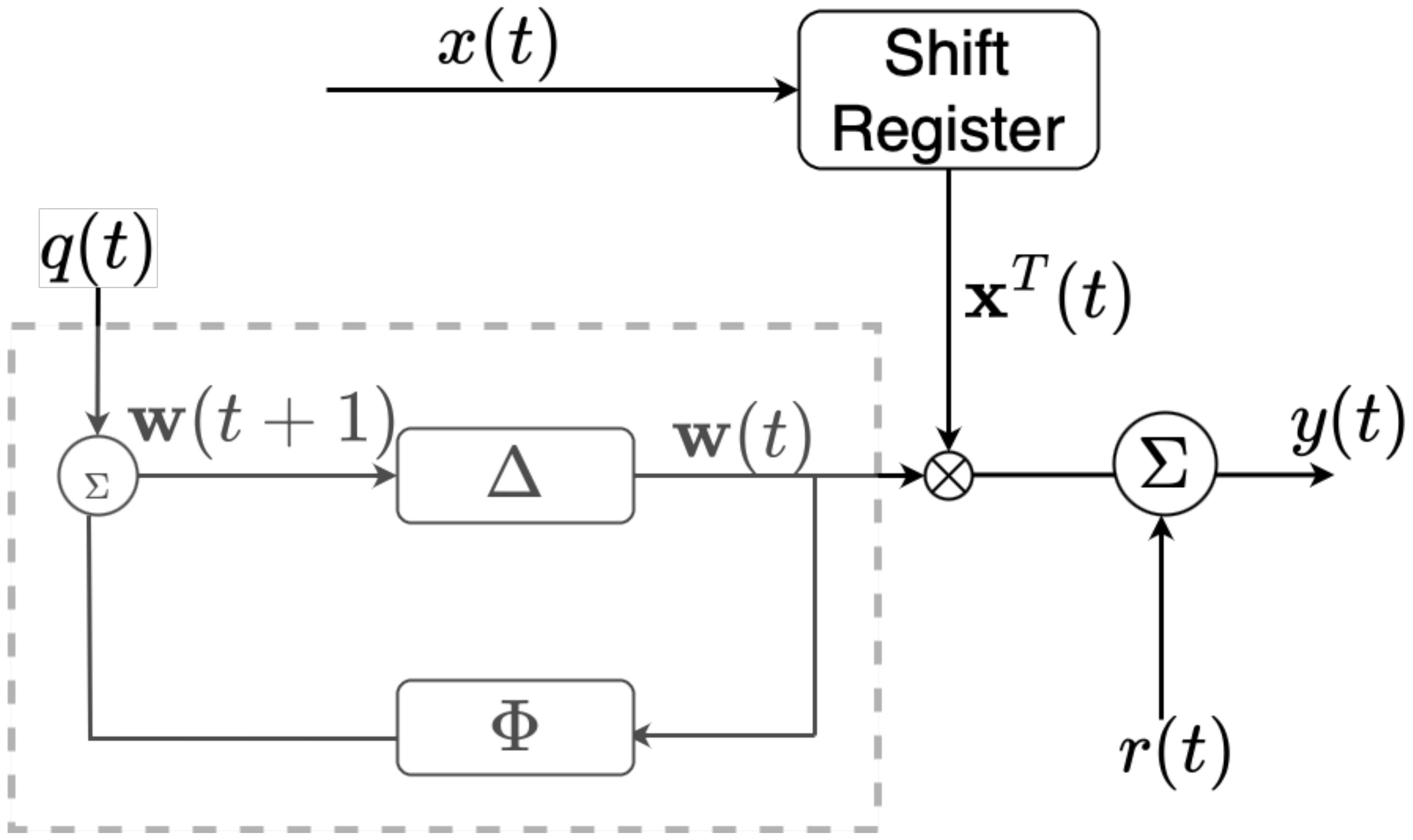


Figure idea taken from [1]

$\mathbf{w}(t)$ – State vector (here, coefficients of L^{th} order FIR filter)

Φ – State Transition matrix (captures system dynamics)

$q(t) \sim \mathcal{N}\left(0,\mathbf{Q}\right)$ is the noise in predicting system dynamics (eg. Gradient noise, linearisation error etc.)

$r(t) \sim \mathcal{N}\left(0,\mathbf{R}\right)$ is the noise in the output or the noise in the observed noisy signal (can be estimated from signal characteristics)

STATE SPACE LEAST MEAN SQUARES (SS-LMS)

Approach presented in [2]. It is a state space generalisation of conventional LMS. Idea is to estimate the unknown states of a system given a set of observations available. The theory is presented for discrete time systems, but valid for continuous time systems too.

$y_{obs}(k)$ — set of observations available
$\mathbf{w}(k)$ — unknown state vector (to be estimated)

Notation:
$\bar{a}$ **— predicted value (theoretical, using system dynamics)**
$\hat{a}$ **— estimated value (after using available observations)**

Equations of SS-LMS

$$\bar{\mathbf{w}}(k) = \mathbf{\Phi}_{k|k-1}\hat{\mathbf{w}}(k-1)$$

$\epsilon(k) = y_{obs}(k) - \bar{y}(k)$ is the prediction error, where, $\bar{y}(k) = \mathbf{C}(k)\bar{\mathbf{w}}(k)$ is the predicted output

$e(k) = y_{obs}(k) - \hat{y}(k)$ is the estimation error, where, $\hat{y}(k) = \mathbf{C}(k)\hat{\mathbf{w}}(k)$ is the estimated output (after using estimator)

$e(k) = \epsilon(k) - \mathbf{C}(k)\left(\hat{\mathbf{w}}(k) - \bar{\mathbf{w}}(k)\right)$ (Relation between estimation and prediction error)

State Update form

$\hat{\mathbf{w}}(k) = \bar{\mathbf{w}}(k) + \mathbf{K}(k)\epsilon(k)$ where, $\mathbf{K}(k)$ is the observer gain which is to be found so that $e(k) = 0$

$\mathbf{K}(k) = \mathbf{C}^T(k)\left(\mathbf{C}(k)\mathbf{C}^T(k)\right)^{-1}$ — Minimum norm solution

To ensure complete controllability from output to states of the system [2], i.e., the states can achieve any output given an initial input, an additional gain matrix $\mathbf{G}$ is added.

$\mathbf{G}$ ensures that $\left[\mathbf{\Phi}_{k|k-1} - \mathbf{K}(k)\mathbf{C}(k)\mathbf{\Phi}_{k|k-1}, \mathbf{K}(k)\right]$ has full rank (required condition for valid estimator controllability).

$$\mathbf{K}(k) = \mu\mathbf{G}\mathbf{C}^T(k)\left(\mathbf{C}(k)\mathbf{C}^T(k)\right)^{-1}$$

μ — step size (for faster convergence)
$\mathbf{G}$ — controllability gain (for valid estimator)

Alternate forms of Observer gain

$\mathbf{K}(k) = \mu\mathbf{G}\mathbf{C}^T(k)\left(\lambda\mathbf{I} + \mathbf{C}(k)\mathbf{C}^T(k)\right)^{-1}$ — for better numerical stability (λ - damping factor)

$\mathbf{K}(k) = \mu\mathbf{G}\mathbf{C}^T(k)$ — for simplification (without normalisation term)

ANALOGY WITH LMS

$\mathbf{w}(k) \implies \mathbf{b}(k)$ (State vector represents filter coefficients of L^{th} order FIR system)

$\mathbf{\Phi}_{k|k-1} \implies \mathbf{I}$ (State transition matrix is just identity)

$\mathbf{C}(k) \implies \mathbf{x}^T(k) = \left[x(k) \quad x(k-1) \quad \dots \quad x(k-L)\right]$ (observation matrix is the vector of past inputs)

$y_{obs}(k) \implies d(k)$ (Observed signal is the desired signal)

The gradient in LMS is incorporated in the observer gain

SIMULATION TO TRACK SINUSOID

Track the phase and amplitude of a sinusoid corrupted by noise, given by the equation: $y(k) = \sigma_s \cos(\omega_0 k T_s + \phi) + \nu(kT_s)$.

The state vector consists of 2 components: amplitude and phase. The state transition matrix is given by [3]:

$$\mathbf{\Phi} = \begin{bmatrix} \cos(\omega_0 T_s) & \sin(\omega_0 T_s) \\ -\sin(\omega_0 T_s) & \cos(\omega_0 T_s) \end{bmatrix}$$

The observation matrix is given by: $\mathbf{C} = [1 \quad 0]$

$\sigma_s = \dfrac{a^2}{2}$ — signal power

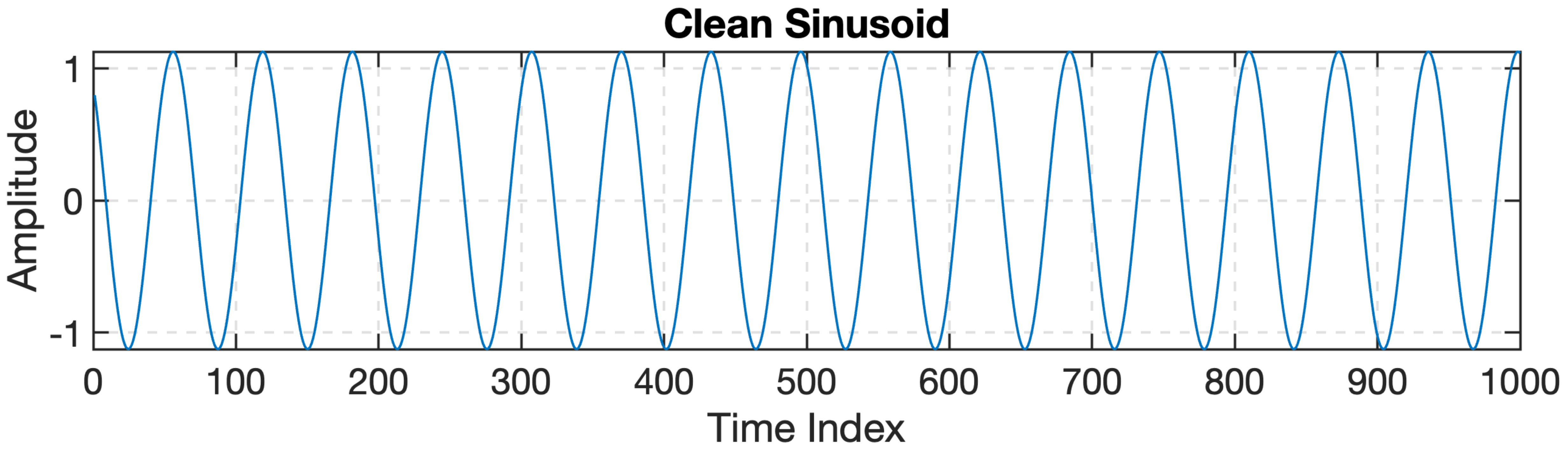
ω_0 — known signal frequency

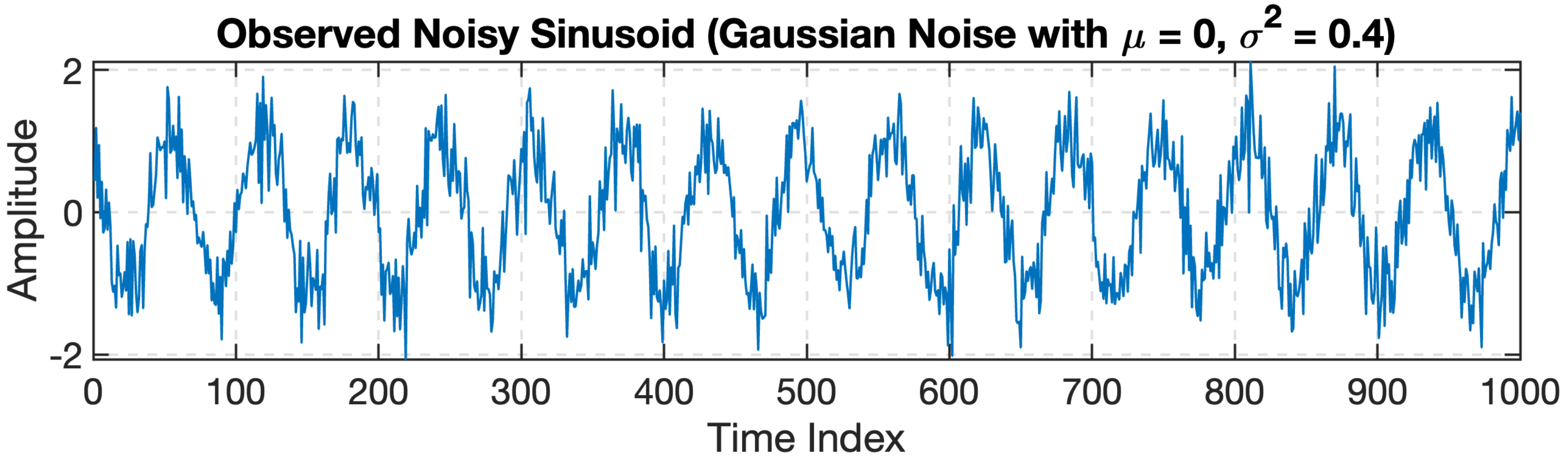
T_s — sampling time

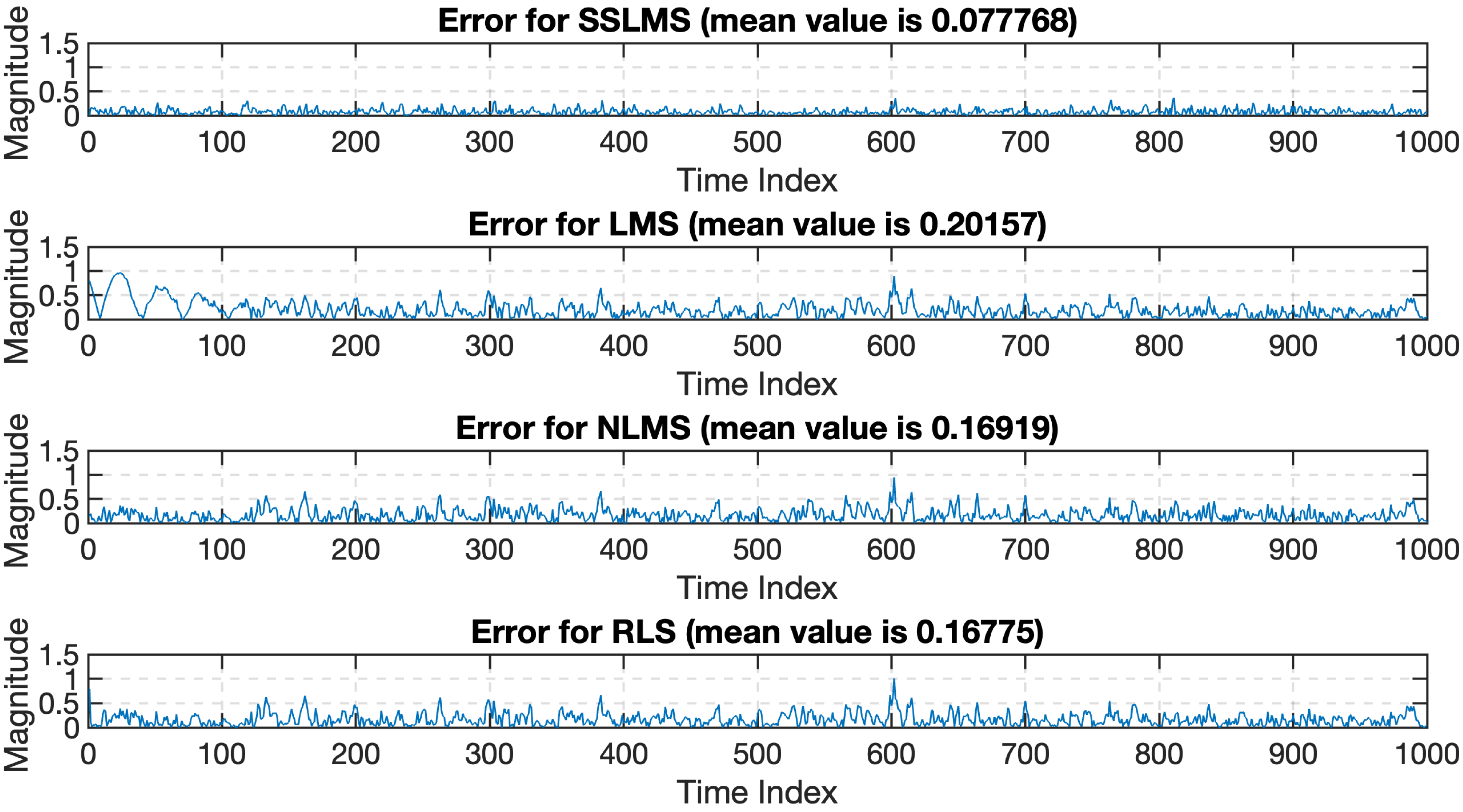
$\nu \sim \mathcal{N}(0,\sigma^2)$ — Gaussian noise

The sinusoid parameters have been tracked using:

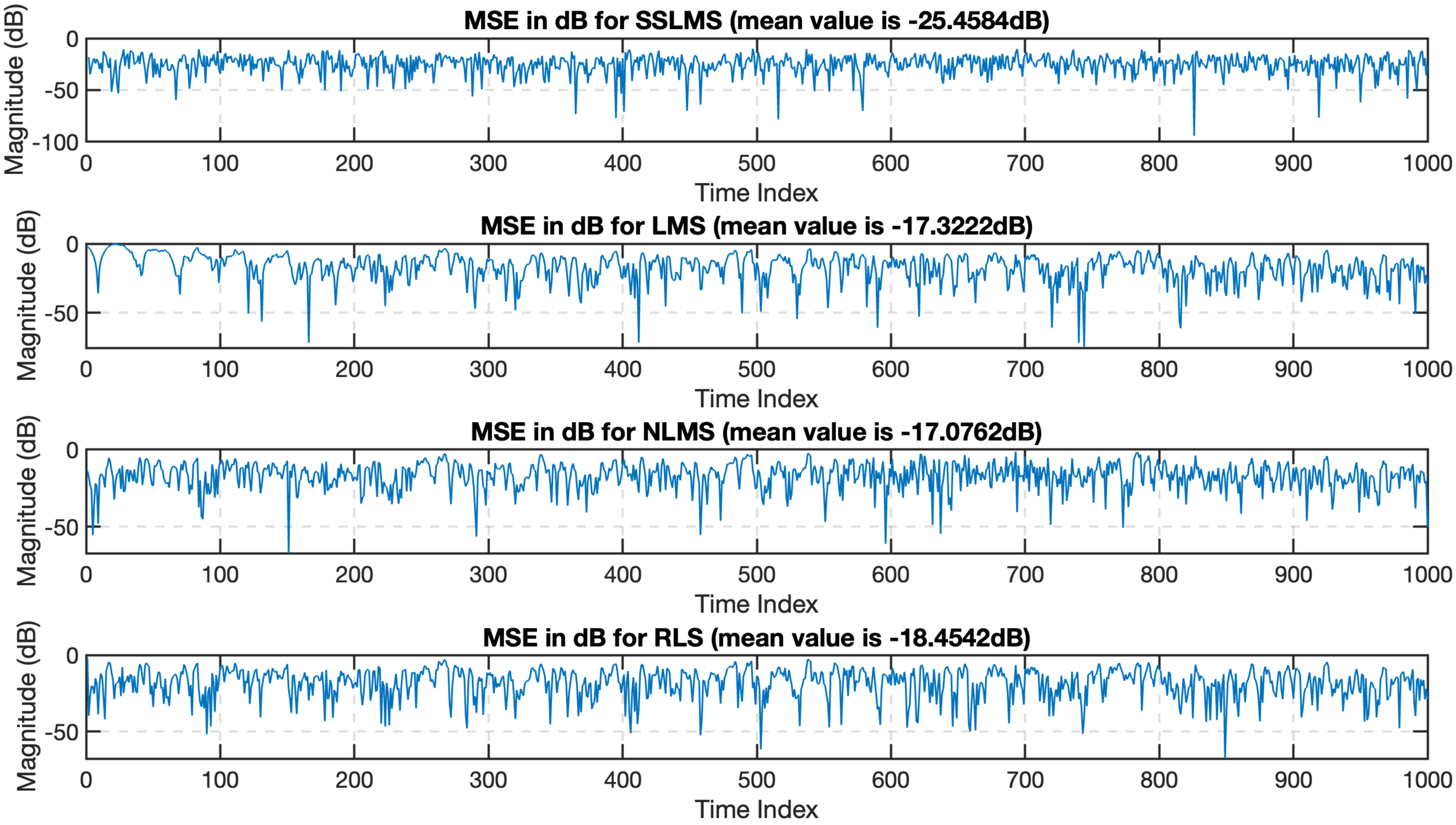
1. SS-LMS $(\mathbf{G} = \mathbf{I}, \mu = 0.2)$
2. LMS $(L = 5, \mu = 0.005)$
3. NLMS $(L = 5, \mu = 0.1)$
4. RLS $(L = 5, \beta = 0.98, \delta = 0.01)$







Error for SSLMS (mean value is 0.077768)
Error for LMS (mean value is 0.20157)
Error for NLMS (mean value is 0.16919)
Error for RLS (mean value is 0.16775)
Magnitude
0
0.5
1
1.5
0
100
200
300
400
500
600
700
800
900
1000
Time Index



MSE in dB for SSLMS (mean value is -25.4584dB)
Magnitude (dB)
0
-50
-100
0
100
200
300
400
500
600
700
800
900
1000
Time Index
MSE in dB for LMS (mean value is -17.3222dB)
Magnitude (dB)
0
-50
Time Index
MSE in dB for NLMS (mean value is -17.0762dB)
Magnitude (dB)
0
-50
Time Index
MSE in dB for RLS (mean value is -18.4542dB)
Magnitude (dB)
0
-50
Time Index

KALMAN FILTER BASED STATE SPACE LEAST MEAN SQUARES

The SSLMS and SSLMSWAM do not explicitly include noise models in the feedback or in the gain matrices. The Kalman filter is a special case of a state space approach. It incorporates the noise covariances and attempts to minimise the state error covariance. In the context of FIR filters, the state error covariance is defined as:

$$\mathbf{P}(k) = E\left[\left(\mathbf{w}(k) - \mathbf{w}_{opt}(k)\right)^T \left(\mathbf{w}(k) - \mathbf{w}_{opt}(k)\right)\right]$$

The Kalman filter algorithm is divided into two stages:

1. Prediction Step
    1. Define the output error: $e_{meas}(k) = y_{obs}(k) - \mathbf{C}(k)\bar{\mathbf{w}}(k)$
    2. $\bar{\mathbf{w}}(k) = \mathbf{I}\hat{\mathbf{w}}(k-1)$ ($\mathbf{C}(k), \mathbf{x}(k)$ as defined in slide 4)
    3. $\bar{\mathbf{P}}(k) = \hat{\mathbf{P}}(k-1) + \mathbf{Q}$

2. Update Step
    1. $\mathbf{K}(k) = \bar{\mathbf{P}}(k)\mathbf{C}^T(k)\left(\mathbf{C}(k)\bar{\mathbf{P}}(k)\mathbf{C}^T(k) + \mathbf{R}\right)^{-1}$ (Kalman Gain)
    2. $\hat{\mathbf{w}}(k) = \bar{\mathbf{w}}(k) + \mathbf{K}(k)e_{meas}(k);$ (Update the predicted state using the observation)
    3. $\hat{\mathbf{P}}(k) = \left(\mathbf{I} - \mathbf{K}(k)\mathbf{C}(k)\right)\bar{\mathbf{P}}(k-1)$ (Update the state error covariance)

Initialisations: $\mathbf{P}(0), \mathbf{w}(0), \mathbf{Q}, \mathbf{R}$

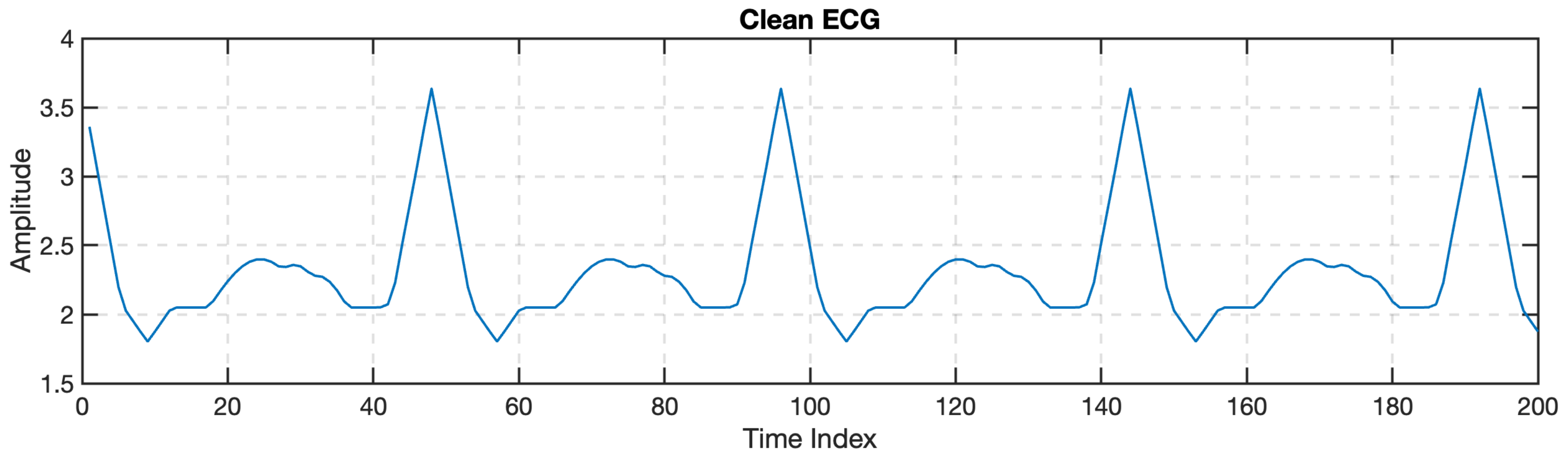
SIMULATION TO TRACK ECG SIGNAL

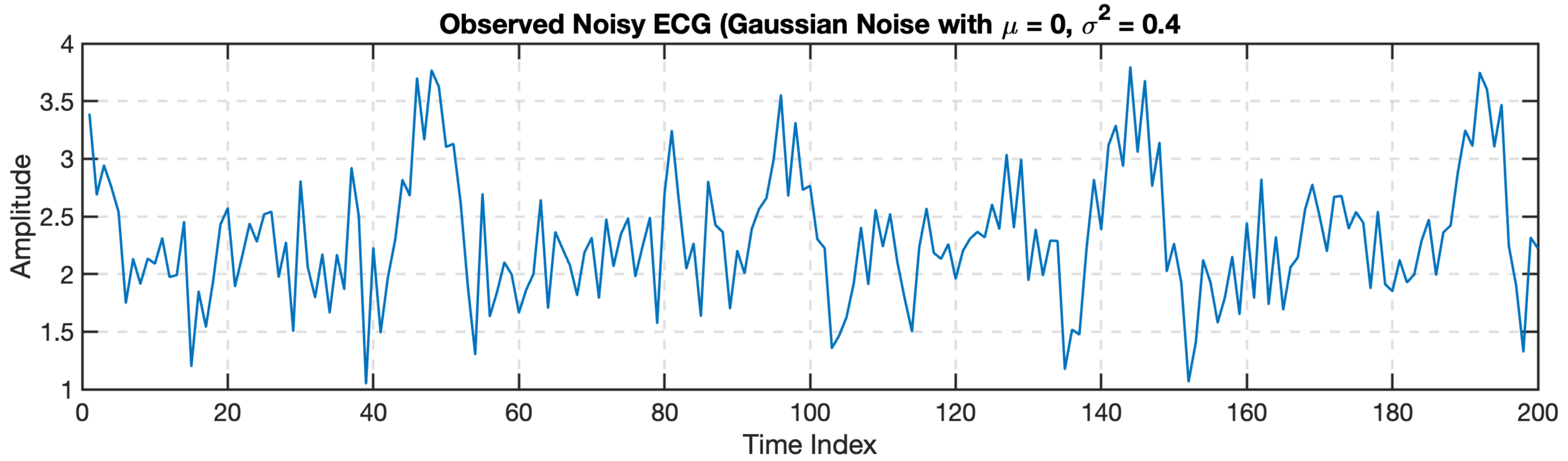
Track an ECG signal which is corrupted by noise. The data has been taken from the MIT-BIH Arrhythmia database [4].

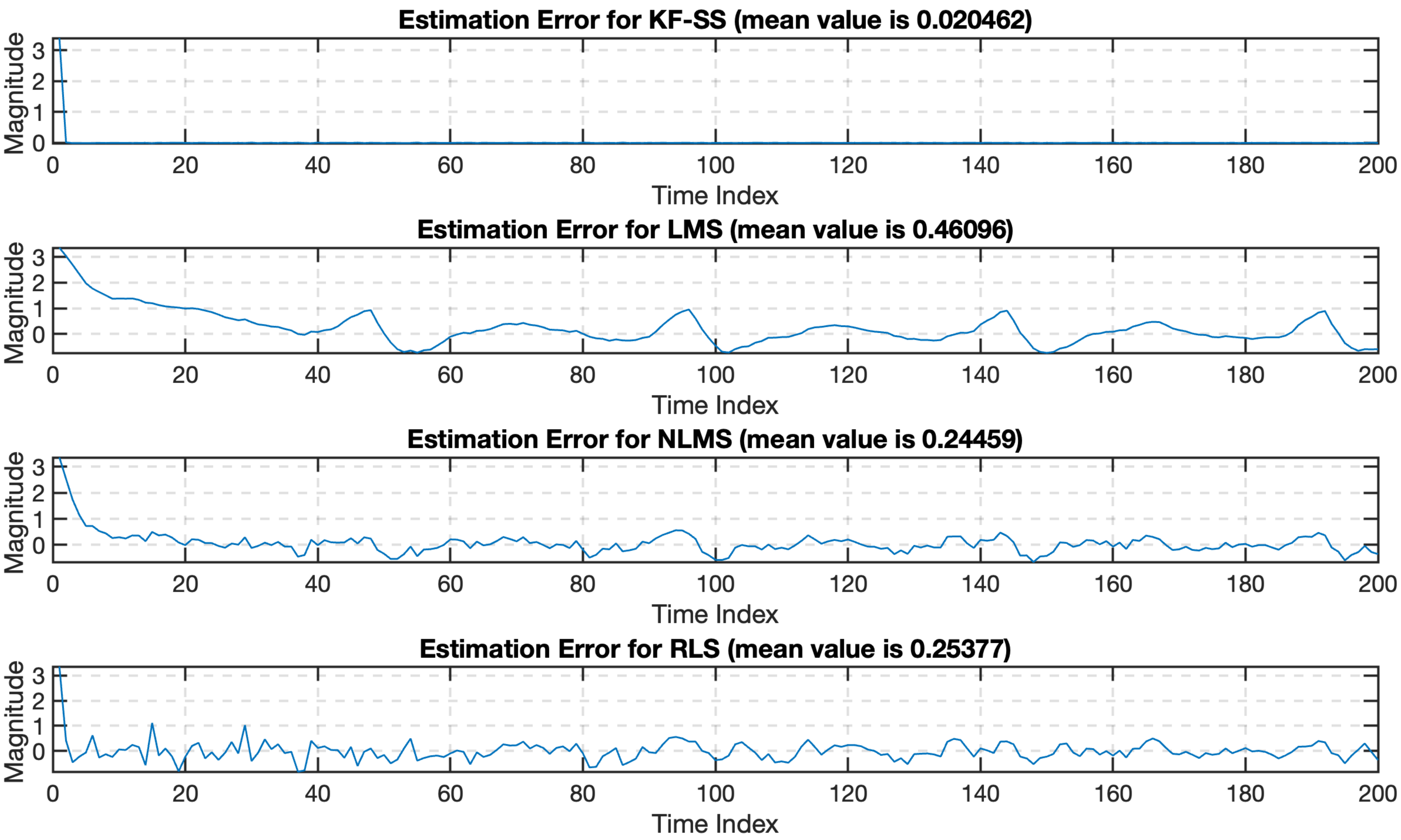
Different parameter based Kalman filter approaches have also been presented in [5]

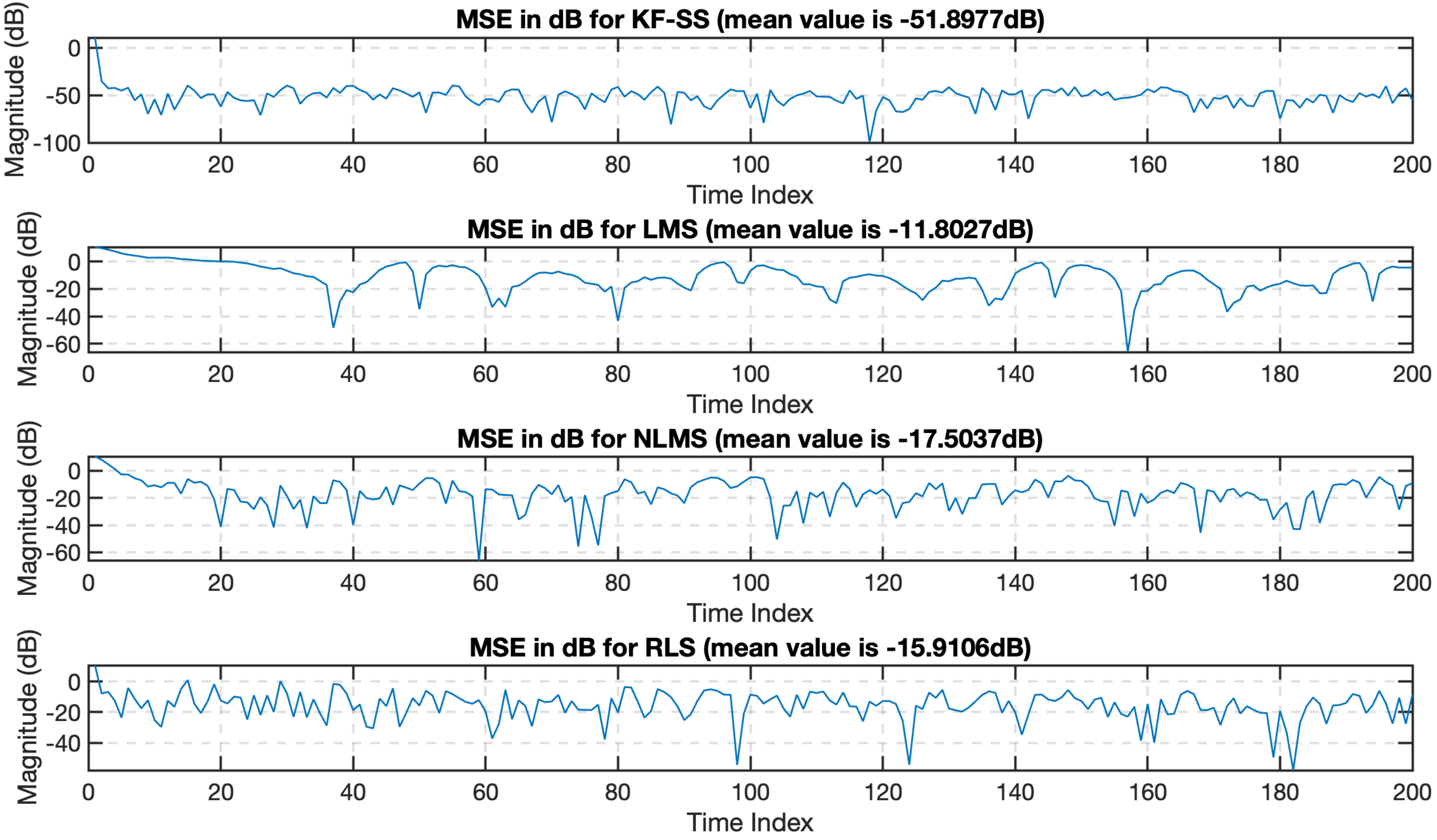
Initialisations:

1. KF ($L = 15, P = 10\mathbf{I}_{L+1}, \mathbf{Q} = 2\mathbf{I}_{L+1}, \mathbf{R} = \left(\frac{1}{\sigma^2}\right)$, $\mathbf{w}(0) = \mathbf{0}$)
2. LMS ($L = 15, \mu = 0.001$)
3. NLMS ($L = 15, \mu = 0.2$)
4. RLS ($L = 15, \beta = 0.99, \delta = 0.01$)






Estimation Error for KF-SS (mean value is 0.020462)
Magnitude
Time Index
Estimation Error for LMS (mean value is 0.46096)
Magnitude
Time Index
Estimation Error for NLMS (mean value is 0.24459)
Magnitude
Time Index
Estimation Error for RLS (mean value is 0.25377)
Magnitude
Time Index


MSE in dB for KF-SS (mean value is -51.8977dB)
Magnitude (dB)
0
-50
-100
0
20
40
60
80
100
120
140
160
180
200
Time Index
MSE in dB for LMS (mean value is -11.8027dB)
Magnitude (dB)
0
-20
-40
-60
Time Index
MSE in dB for NLMS (mean value is -17.5037dB)
Magnitude (dB)
0
-20
-40
-60
Time Index
MSE in dB for RLS (mean value is -15.9106dB)
Magnitude (dB)
0
-20
-40
Time Index

DISCUSSION AND KEY TAKEAWAYS

MERITS

- SS-based algorithms are more robust because they incorporate both system dynamics and noise models.
- They show a faster convergence compared to conventional adaptive filtering algorithms. Also have a relatively low computation per sample.
- There is more control to the designer to configure the adaptive feedback. They also offer more insight into physical aspects of a system, if needed.

DEMERITS

- SS-based algorithms are heavily dependent on the quality of the observations available to them.
- They are also sensitive to initialisations, which require knowledge of the system concerned.

LEARNING FOR THE CLASS

Introduce concepts of feedback control from control theory applied to applications like signal tracking, system ID, noise cancellation etc. Since they have been heavily studied in target tracking and navigation algorithms, such algorithms are extremely stable and implementable on real time.
It is worthy considering them in applications involving adaptive signal processing

REFERENCES

[1] Lippuner, D., & Moschytz, G. S. (2004). The Kalman filter in the context of adaptive filter theory. *International journal of circuit theory and applications*, *32*(4), 223-253.

[2] Malik, M. B., & Salman, M. (2008). State-space least mean square. *Digital Signal Processing*, *18*(3), 334-345.

[3] SALMAN, M. (2009). *Adaptive estimation using state-space methods* (Doctoral dissertation, National University of Sciences and Technology, Pakistan).

[4] Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. IEEE Eng in Med and Biol 20(3):45-50 (May-June 2001). (PMID: 11446209)

[5] Vullings, R., De Vries, B., & Bergmans, J. W. (2010). An adaptive Kalman filter for ECG signal enhancement. *IEEE transactions on biomedical engineering*, *58*(4), 1094-1103.