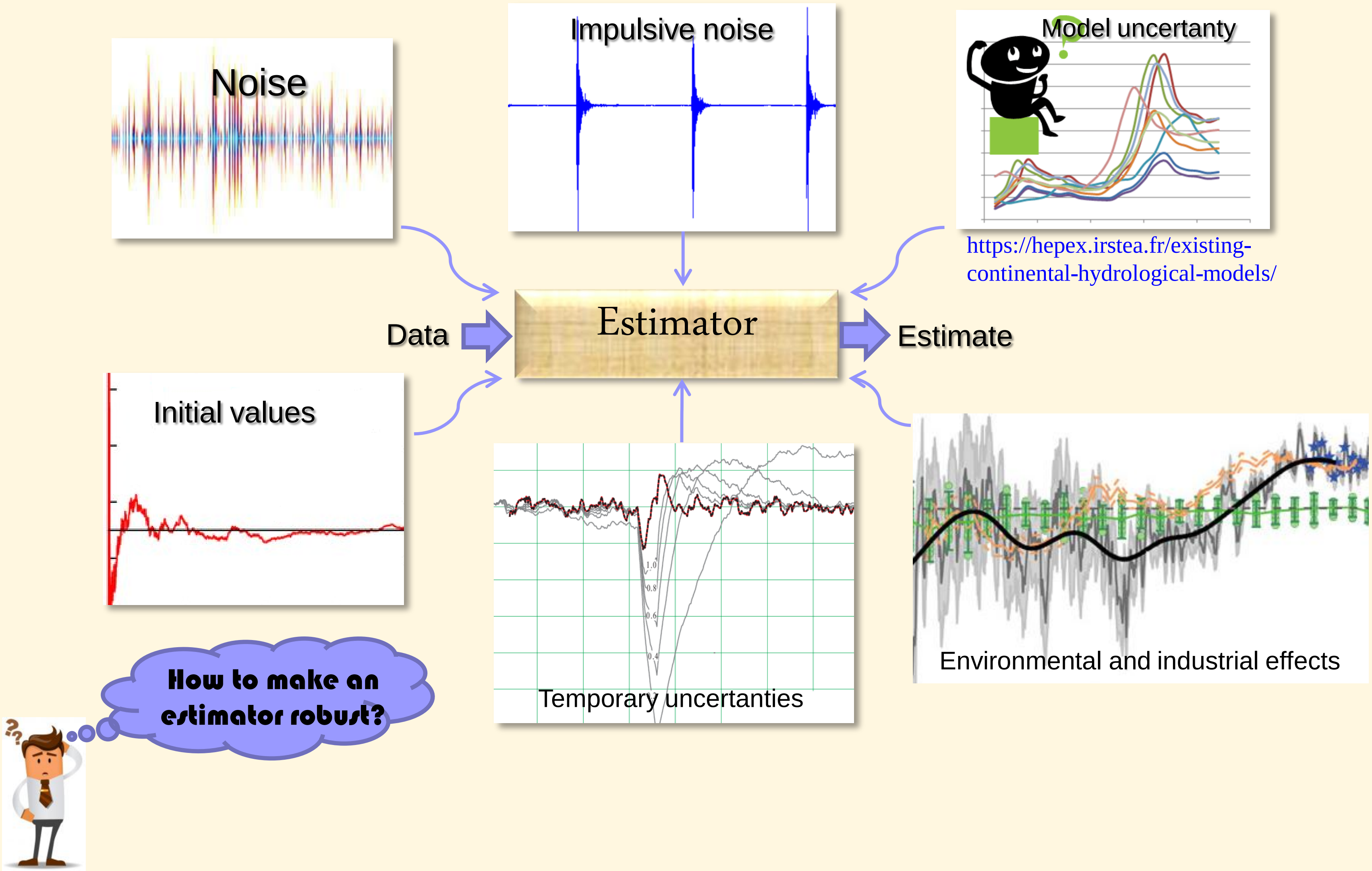
State Estimation with Colored Measurement Noise using Kalman and UFIR Filtering

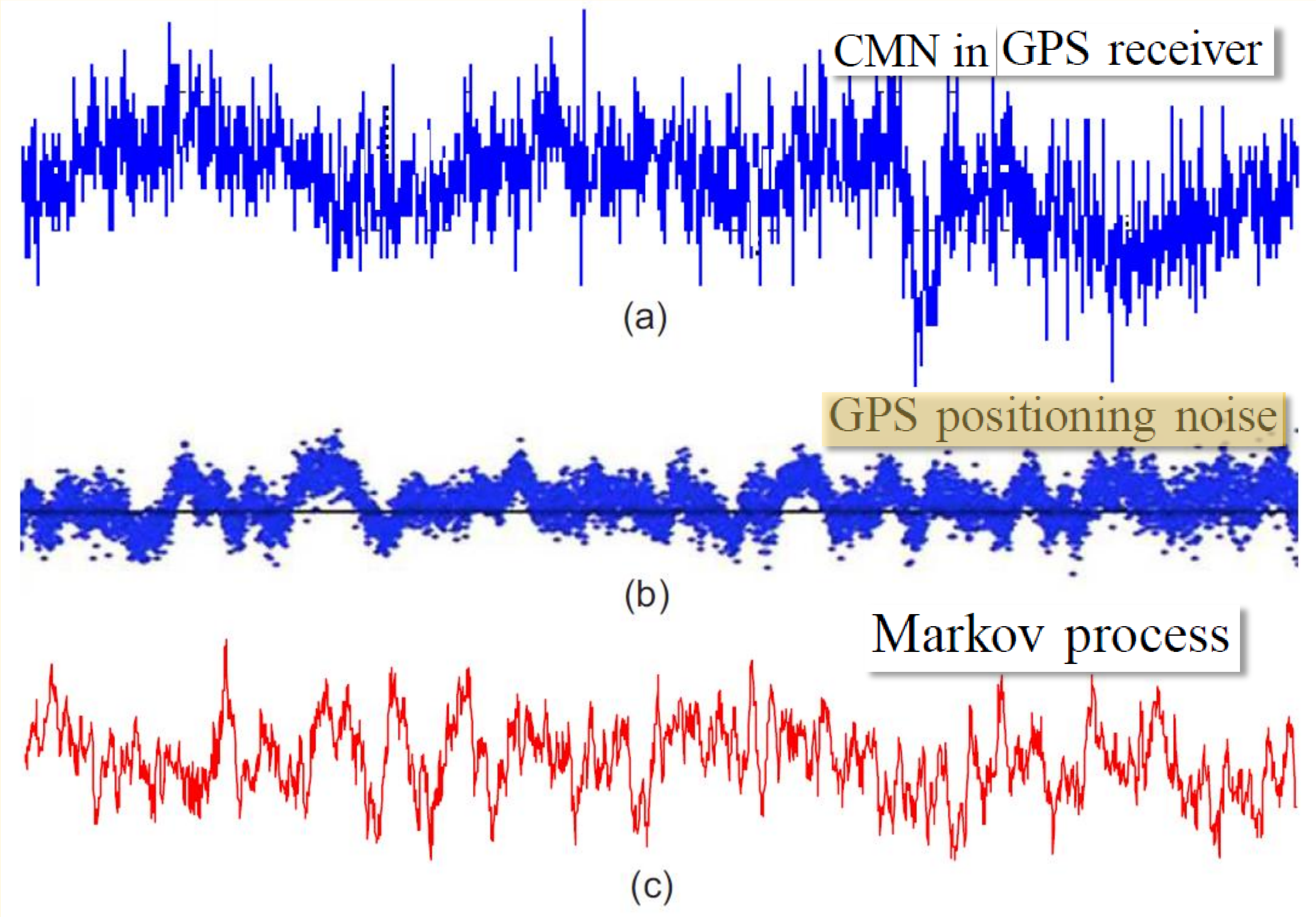
Prof. Yuriy S. Shmaliy
PhD, DSc, IEEE Fellow
Universidad de Guanajuato
shmaliy@ugto.mx

TELECOM SudParis, France – 29, August, 2018

Factor Affecting Estimation Accuracy



Colored Noise in Electronic Systems



State-Space Model with Colored Measurement Noise

We will view the problem with the CMN as the *triplet Markov model* (TMM):

$$\begin{bmatrix} x_n \\ v_n \\ y_{n-1} \end{bmatrix} = \begin{bmatrix} F_n & 0 & 0 \\ 0 & \Psi_n & 0 \\ H_{n-1} & I & 0 \end{bmatrix} \begin{bmatrix} x_{n-1} \\ v_{n-1} \\ y_{n-1} \end{bmatrix} + \begin{bmatrix} B_n w_n \\ \xi_n \\ 0 \end{bmatrix}$$

where $x_n \in \mathbb{R}^K$ is the *state vector*,
$y_n \in \mathbb{R}^M$ is the *observation vector*
$v_n \in \mathbb{R}^M$ is the *Markov noise*

Error vectors, $w_n \sim \mathcal{N}(0, Q_n) \in \mathbb{R}^P$ and $\xi_n \sim \mathcal{N}(0, R_n) \in \mathbb{R}^M$, are *zero mean white Gaussian* and *uncorrelated* with known *covariances.*
All matrices are supposed to be known and time-varying.
Matrix $\Psi_n \in \mathbb{R}^{M \times M}$ is selected such that noise v_n becomes stationary.

Two approaches are known to derive an estimator for CMN

- Augmented state
- Measurement differencing

Important Observations:

1) Known results employ the prediction state model $x_n = F_n x_{n-1} + w_{n-1}$, $w_n \sim \mathcal{N}(0, Q_n)$, used in feedback control. Tracking may fit better the real-time model [29] $x_n = F_n x_{n-1} + w_n$, which real-time noise w_n suits FIR filtering [30] as a causal equivalent (or innovations sequence) to x_n [31]. Estimates from both models are equal when $Q_n = Q_{n-1}$, but differ otherwise, especially for Markov jumps.
2) State-space modeling with colored measurement noise can be provided using the triplet Markov chains (TMC) introduced in [32] [33] and described for KF as (14, in [34]).
3) The problem of colored noise can be viewed as robust against errors in noise vectors. The unbiased finite impulse response (UFIR) filter [35] completely ignores noise and may be a better choice than the KF [36] [37].

Augmented State Approach

Separate TMM into the augmented state and regular observation equations

$$\begin{bmatrix} x_n \\ v_n \end{bmatrix} = \begin{bmatrix} F_n & 0 \\ 0 & \Psi_n \end{bmatrix} \begin{bmatrix} x_{n-1} \\ v_{n-1} \end{bmatrix} + \begin{bmatrix} B_n & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} w_n \\ \xi_n \end{bmatrix}$$

$$y_n = \begin{bmatrix} H_n & I \end{bmatrix} \begin{bmatrix} x_n \\ v_n \end{bmatrix} + 0\,,$$

and rewrite as

$$\boxed{\begin{aligned} \tilde{x}_n &= \tilde{F}_n \tilde{x}_{n-1} + \tilde{B}_n \tilde{w}_n\,, \\ y_n &= \tilde{H}_n \tilde{x}_n + \tilde{v}_n\,, \end{aligned}}$$

where $\tilde{w}_n$ and $\tilde{v}_n$ have the covariances

$$E\{\tilde{w}_n \tilde{w}_n^T\} = \begin{bmatrix} Q_n & 0 \\ 0 & R_n \end{bmatrix} \qquad E\{\tilde{v}_n \tilde{v}_n^T\} = 0$$

Remark: The KF can be applied straightforwardly, but becomes ill-conditioned because of $E\{\tilde{v}_n \tilde{v}_n^T\} = 0$.

Measurement Differencing

To avoid colored noise, consider a new observation

$$\begin{aligned} z_n &= y_n - \Psi_n y_{n-1} \, , \\ &= H_n x_n + v_n - \Psi_n H_{n-1} x_{n-1} - \Psi_n v_{n-1} \end{aligned}$$

take x_{n-1} and v_{n-1} from the TMM, and obtain

$$\boxed{z_n = D_n x_n + \bar{v}_n}$$

New observation

where $D_n = H_n - \Gamma_n$, $\bar{v}_n = \Gamma_n B_n w_n + \xi_n$, $\Gamma_n = \Psi_n H_{n-1} F_n^{-1}$

$$E\{\bar{v}_n \bar{v}_n^T\} = \Gamma_n \Phi_n + R_n \qquad E\{\bar{v}_n w_n^T\} = \Gamma_n B_n Q_n$$

$$\Phi_n = B_n Q_n B_n^T \Gamma_n^T$$

Remark: Because $\bar{v}_n$ correlates w_n , two different KF algorithms can be derived, but only one unique UFIR algorithm.

Optimal KF Algorithm

1. Define the prior state estimate as $\hat{x}_n^- = F_n \hat{x}_{n-1}$ and find the prior error covariance:

$$P_n^- = F_n P_{n-1} F_n^T + B_n Q_n B_n^T$$

2. Define the measurement residual as

$$s_n = z_n - D_n \hat{x}_n^- = D_n F_n \epsilon_{n-1} + D_n B_n w_n + \bar{v}_n$$

and find the innovation covariance $S_n = E\{s_n s_n^T\}$ as

$$\begin{aligned} S_n &= D_n F_n P_{n-1} F_n^T D_n^T + D_n B_n Q_n B_n^T D_n^T \\ &\quad + \Gamma_n \Phi_n + R_n + D_n \Phi_n + \Phi^T D_n^T \\ &= D_n P_n^- D_n^T + R_n + H_n \Phi_n + \Phi_n^T D_n^T . \end{aligned}$$

3. Write the KF recursive estimate as

$$\begin{aligned}\hat{x}_n &= \hat{x}_n^- + K_n s_n \\ &= F_n \hat{x}_{n-1} + K_n (z_n - D_n F_n \hat{x}_{n-1})\end{aligned}$$

Remark: The optimal bias correction gain K_n can now be found in two ways:

1. Derive a new Kalman gain for correlated noise sources.
2. De-correlate the noise sources and apply the standard Kalman gain.

KF for Correlated Noise Sources

1. Define the estimation error as $\epsilon_n = x_n - \hat{x}_n$ and write

$$\epsilon_n = (I - K_n D_n) F_n \epsilon_{n-1} + (I - K_n D_n) B_n w_n - K_n \bar{v}_n$$

2. Observe that $\epsilon_{n-1},\ w_n,\ \text{and}\ \bar{v}_n$ are mutually uncorrelated and find the error covariance $P_n = E\{\epsilon_n \epsilon_n^T\}$ as

$$\begin{aligned}
P_n &= (I - K_n D_n) P_n^- (I - K_n D_n)^T \\
&\quad + K_n (\Gamma_n \Phi_n + R_n) K_n^T - (I - K_n D_n) \Phi_n K_n^T \\
&\quad - K_n \Phi_n^T (I - K_n D_n)^T \qquad (19a) \\
&= (I - K_n D_n) P_n^- (I - K_n D_n)^T \\
&\quad + K_n [(H_n - D_n) \Phi_n + R_n] K_n^T - (I - K_n D_n) \Phi_n K_n^T \\
&\quad - K_n \Phi_n^T (I - K_n D_n)^T \\
&= (I - K_n D_n) P_n^- (I - K_n D_n)^T + K_n (H_n \Phi_n + R_n) K_n^T \\
&\quad - \Phi_n K_n^T - K_n \Phi_n^T (I - K_n D_n)^T \\
&= (I - K_n D_n) P_n^- (I - K_n D_n)^T + K_n R_n K_n^T \\
&\quad - K_n \Phi_n^T (I - K_n D_n)^T - (I - K_n H_n) \Phi_n K_n^T \\
&= P_n^- - (P_n^- D_n^T + \Phi_n) K_n^T - K_n (P_n^- D_n^T + \Phi_n)^T \\
&\quad + K_n S_n K_n^T\,, \qquad (19b)
\end{aligned}$$

3. Find the optimal bias correction gain by minimizing the trace of P_n, which is equal to the MSE, as

$$\frac{\partial \,\mathrm{tr} P_n}{\partial K_n} = -2(P_n^- D_n^T + \Phi_n) + 2K_n S_n = 0$$

$$\Downarrow$$

$$\boxed{K_n = (P_n^- D_n^T + \Phi_n) S_n^{-1}}$$

4. Using the optimal gain derived, transform the error covariance to

$$\boxed{P_n = P_n^- - K_n (D_n P_n^- + \Phi_n^T)}$$

5. Modify the KF algorithm as shown below with a pseudo code.

Pseudo Code of the KF for Correlated Noise Sources

Algorithm 1: KF for Correlated w_n and $\bar{v}_n$

Data: y_n, $\hat{x}_0$, P_0, Q_n, R_n

Result: $\hat{\mathbf{x}}_n$, P_n

```
begin
    for n = 1, 2, ··· do
        z_n = y_n − Ψ_n y_{n−1} ;
        P_n^− = F_n P_{n−1} F_n^T + B_n Q_n B_n^T ;
        S_n = D_n P_n^− D_n^T + R_n + H_n Φ_n + Φ_n^T D_n^T ;
        K_n = (P_n^− D_n^T + Φ_n) S_n^{−1} ;
        x̂_n^− = F_n x̂_{n−1} ;
        x̂_n = x̂_n^− + K_n (z_n − D_n x̂_n^−) ;
        P_n = (I − K_n D_n) P_n^− − K_n Φ_n^T ;
    end for
end
```

$$z_n = y_n - \Psi_n y_{n-1}$$
$$P_n^- = F_n P_{n-1} F_n^T + B_n Q_n B_n^T$$
$$S_n = D_n P_n^- D_n^T + R_n + H_n \Phi_n + \Phi_n^T D_n^T$$
$$K_n = (P_n^- D_n^T + \Phi_n) S_n^{-1}$$
$$\hat{x}_n^- = F_n \hat{x}_{n-1}$$
$$\hat{x}_n = \hat{x}_n^- + K_n (z_n - D_n \hat{x}_n^-)$$
$$P_n = (I - K_n D_n) P_n^- - K_n \Phi_n^T$$

Remark: By $\psi = 0$, the algorithm becomes the standard KF

KF for De-correlated Noise Sources

1. Use the Bar-Shalom's trick and represent the state equation as

$$\begin{aligned} x_n &= F_n x_{n-1} + B_n w_n + \boxed{\Lambda_n (z_n - D_n x_n - \bar{v}_n)} \\ &= A_n x_{n-1} + u_n + \zeta_n \,, \end{aligned}$$

where

$$\begin{aligned} A_n &= (I - \Lambda_n D_n) F_n \,, \\ u_n &= \Lambda_n z_n \,, \\ \zeta_n &= (I - \Lambda_n D_n) B_n w_n - \Lambda_n \bar{v}_n \end{aligned}$$

and noise $\zeta_n \sim \mathcal{N}(0, \mathcal{Q}_n) \in \mathbb{R}^K$ has the covariance

$$\begin{aligned} \mathcal{Q}_n &= E\{[(I - \Lambda_n D_n) B_n w_n - \Lambda_n \bar{v}_n][\ldots]^T\} \\ &= E\{\{[(I - \Lambda_n D_n) B_n - \Lambda_n B_n^T \Gamma_n^T] w_n - \Lambda_n \xi_n\}\{\ldots\}^T\} \\ &= [I - \Lambda_n (D_n + \Psi_n H_{n-1} F_n^{-1})] B_n Q_n B_n^T \\ &\quad \times [\ldots]^T + \Lambda_n R_n \Lambda_n^T \\ &= (I - \Lambda_n H_n) B_n Q_n B_n^T (I - \Lambda_n H_n)^T + \Lambda_n R_n \Lambda_n^T \,. \quad (27) \end{aligned}$$

2. Find matrix Λ_n to make ζ_n and $\bar{v}_n$ uncorrelated by considering $E\{\zeta_n \bar{v}_n^T\} = 0$ as

$$
\begin{aligned}
E\{\zeta_n \bar{v}_n^T\} &= E\{[(I - \Lambda_n D_n)B_n w_n - \Lambda_n \bar{v}_n] \\
&\quad \times (w_n^T B_n^T \Gamma_n^T + \xi_n^T)\}, \\
&= E\{[(I - \Lambda_n D_n)B_n w_n - \Lambda_n \Gamma_n B_n w_n - \Lambda_n \xi_n] \\
&\quad \times (w_n^T B_n^T \Gamma_n^T + \xi_n^T)\}, \\
&= [(I - \Lambda_n D_n) - \Lambda_n \Gamma_n] B_n Q_n B_n^T \Gamma_n^T \\
&\quad - \Lambda_n R_n = 0 \\
&= (I - \Lambda_n H_n)\Phi_n - \Lambda_n R_n = 0 \qquad (28)
\end{aligned}
$$

that yields

$$\boxed{\Lambda_n = \Phi_n (H_n \Phi_n + R_n)^{-1}}$$

3. By the transformations, obtain

$$\boxed{\mathcal{Q}_n = (I - \Lambda_n H_n) B_n Q_n B_n^T (I - \Lambda_n D_n)^T}$$

4. Apply the standard KF algorithm

Pseudo Code of the KF for De-Correlated Noise Sources

Algorithm 2: KF for De-correlated w_n and $\bar{v}_n$

Data: $y_n,\ \hat{x}_0,\ P_0,\ Q_n,\ R_n$

Result: $\hat{\mathbf{x}}_n,\ P_n$

begin

  for $n = 1, 2, \cdots$ **do**

    $z_n = y_n - \Psi_n y_{n-1}$;

    $P_n^- = A_n P_{n-1} A_n^T + Q_n$;

    $S_n = D_n P_n^- D_n^T + R_n$;

    $K_n = P_n^- D_n^T S_n^{-1}$;

    $\hat{x}_n^- = A_n \hat{x}_{n-1} + \Lambda_n z_n$;

    $\hat{x}_n = \hat{x}_n^- + K_n (z_n - D_n \hat{x}_n^-)$;

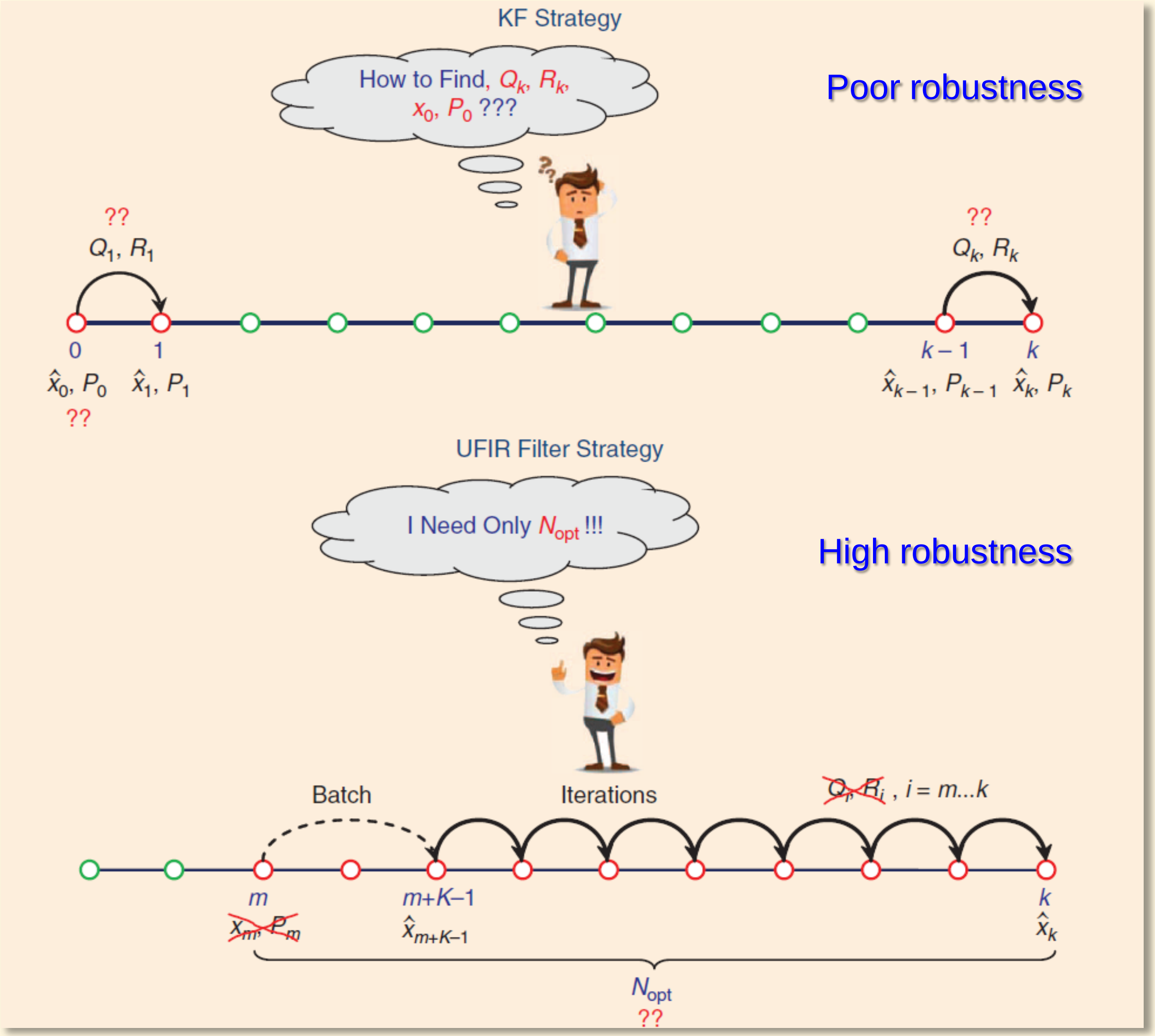
    $P_n = (I - K_n D_n) P_n^-$;

  end for

end

Remark: By $\psi = 0$, the algorithm becomes the standard KF

Remark: KF requires all information about mode, noise, and initial values and has thus poor robustness. Better robustness provides the UFIR filter.



Y. S. Shmaliy, S. Zhao, C. K. Ahn, Unbiased FIR filtering: an iterative alternative to Kalman filtering ignoring noise and initial conditions, *IEEE Control Systems Mag.*, **37**, 5, 70-89, 2017.

Unique UFIR Filtering Algorithm

Unlike the KF, the UFIR filter can be applied universally on [*m*,*n*] for correlated and de-corelated noise sources

Algorithm 3: UFIR Filter for CMN

Data: N, y_n
Result: $\hat{x}_n$

```
begin
    for k = N − 1, N, ··· do
        m = k − N + 1 , s = k − N + K ;
        G_s = (C^T_{m,s} C_{m,s})^{-1} ;
        x̄_s = G_s C^T_{m,s} Y_{m,s} ;
        for l = s + 1 : n do
            z_l = y_l − Ψ_l y_{l−1} ;
            G_l = [D_l^T D_l + (F_l G_{l−1} F_l^T)^{-1}]^{-1} ;
            K_l = G_l D_l^T ;
            x̄_l^- = F_l x̄_{l−1} ;
            x̄_l = x̄_l^- + K_l (z_l − D_l x̄_l^-) ;
        end for
        x̂_n = x̄_n ;
    end for
end
```

$$Y_{m,s} = [\, y_m \ \ \ldots \ \ y_s \,]^T$$

$$C_{m,s} = \begin{bmatrix} D_m (F_s \ldots F_{m+1})^{-1} \\ \vdots \\ D_{s-1} F_s^{-1} \\ D_s \end{bmatrix}$$

Remark: No information about noise and initial values is required.

Y. S. Shmaliy, Linear optimal FIR estimation of discrete time-invariant state-space models, *IEEE Trans. Signal Process.*, **58**, 6, 3086-3096, 2010.

Numerical Example of Tracking

Consider the two-state and three state tracking problems with

Two-state:

$$F = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} \frac{\tau^2}{2} \\ \tau \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$\sigma_w = 50\,\mathrm{m/s^2}$ $\sigma_\xi = 4\,\mathrm{m}$ $x_0 = [1\ 1]^T$ $P_0 = 0$

$$N_{\mathrm{opt}} = \sqrt{\frac{12\sigma_\xi}{\tau^2\sigma_w}} \cong 20$$

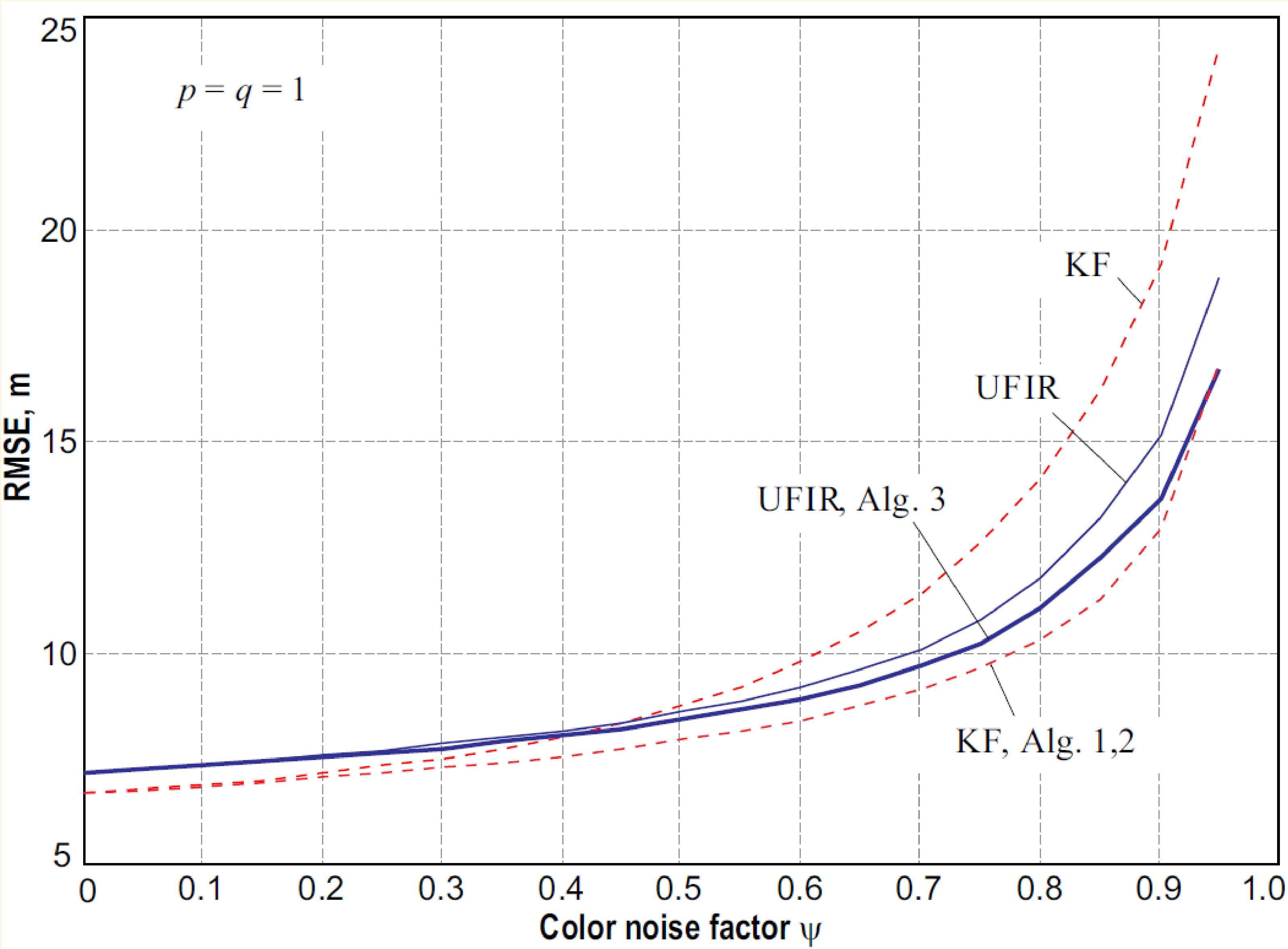
Three-state:

$$F = \begin{bmatrix} 1 & \tau & \frac{\tau^2}{2} \\ 0 & 1 & \tau \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} \frac{\tau^2}{2} \\ \tau \\ 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

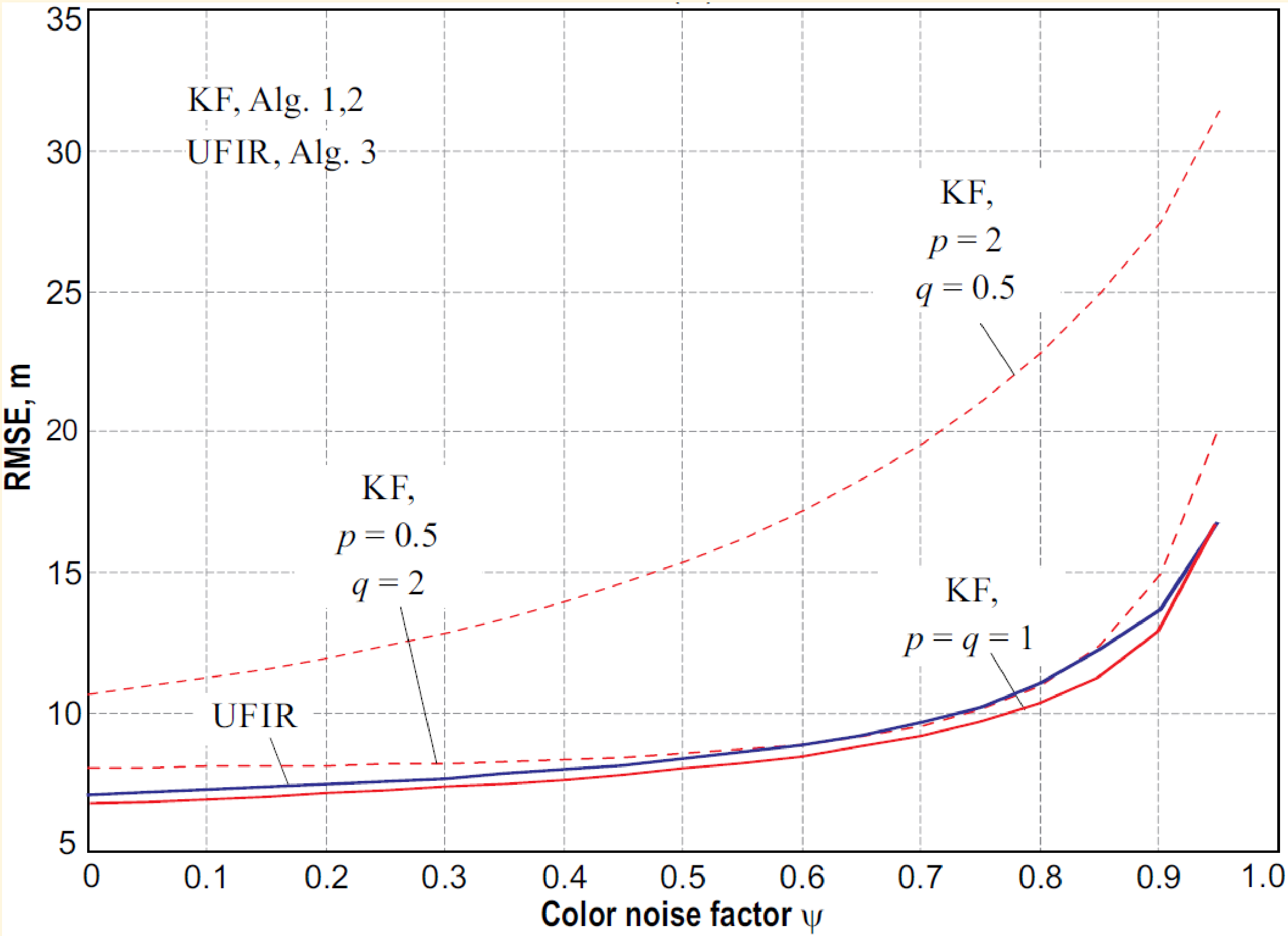
$Q = \sigma_w^2 = 1\,\mathrm{m/s^2}$ $R = \sigma_\xi^2 = 20^2\,\mathrm{m}$ $x_0 = [1\ 1\ 1]^T$

$P_0 = 0$

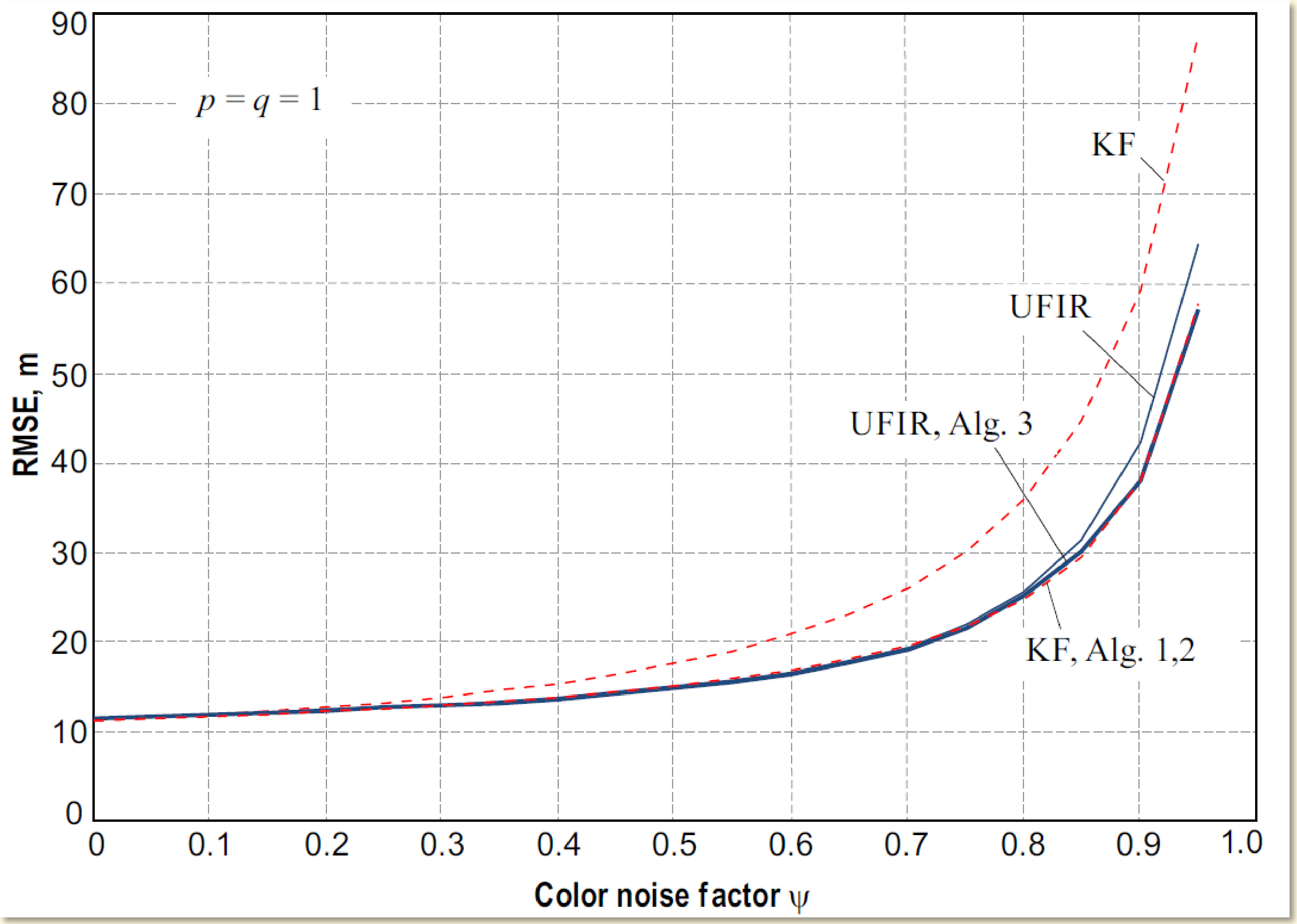
2-States: Performance under the ideal conditions of $p = q = 1$



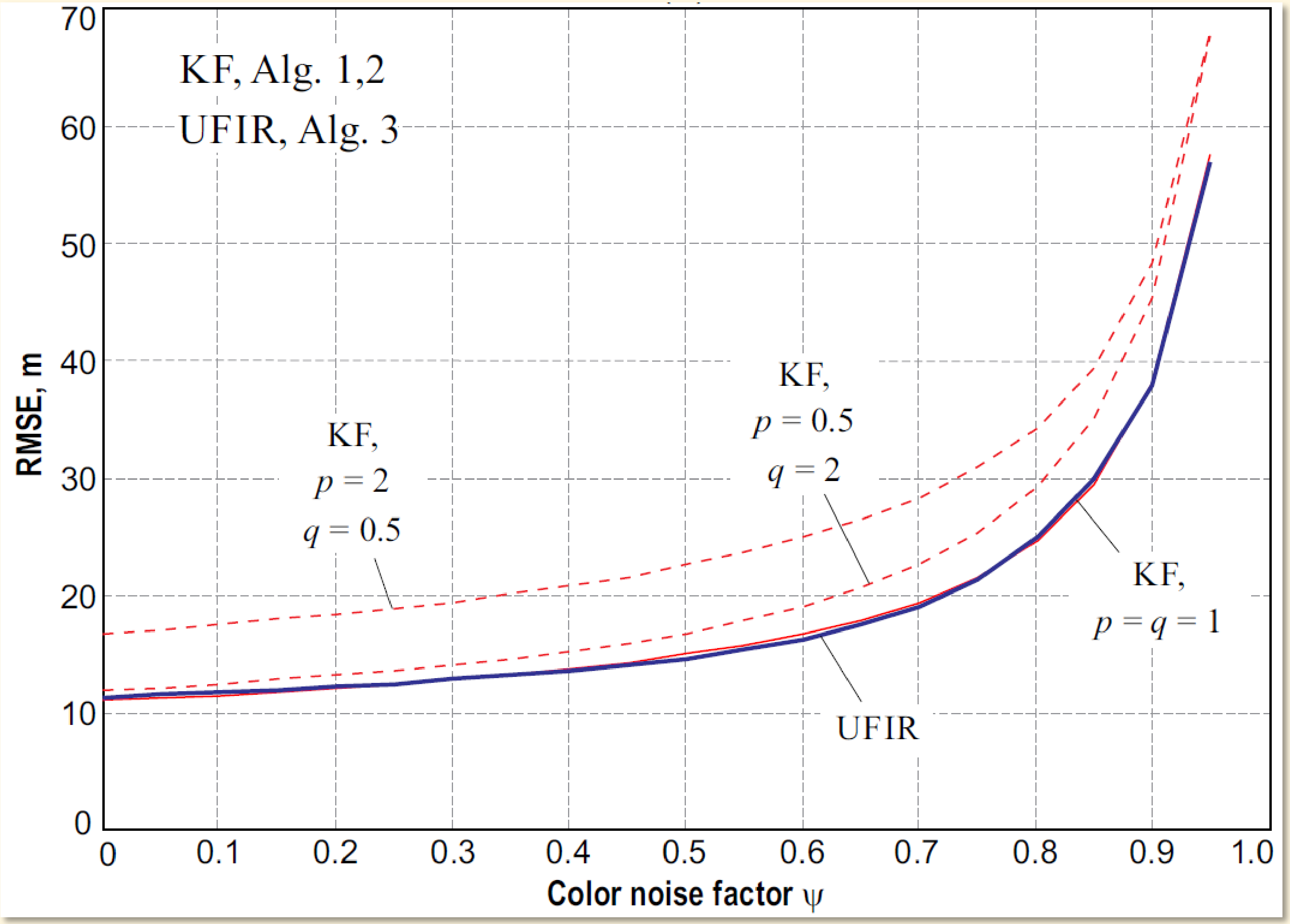
2-States: Performance under the real-world conditions of $p \neq q \neq 1$



3-States: Performance under the ideal conditions of p = q =1



3-States: Performance under the real-world conditions of $p \neq q \neq 1$

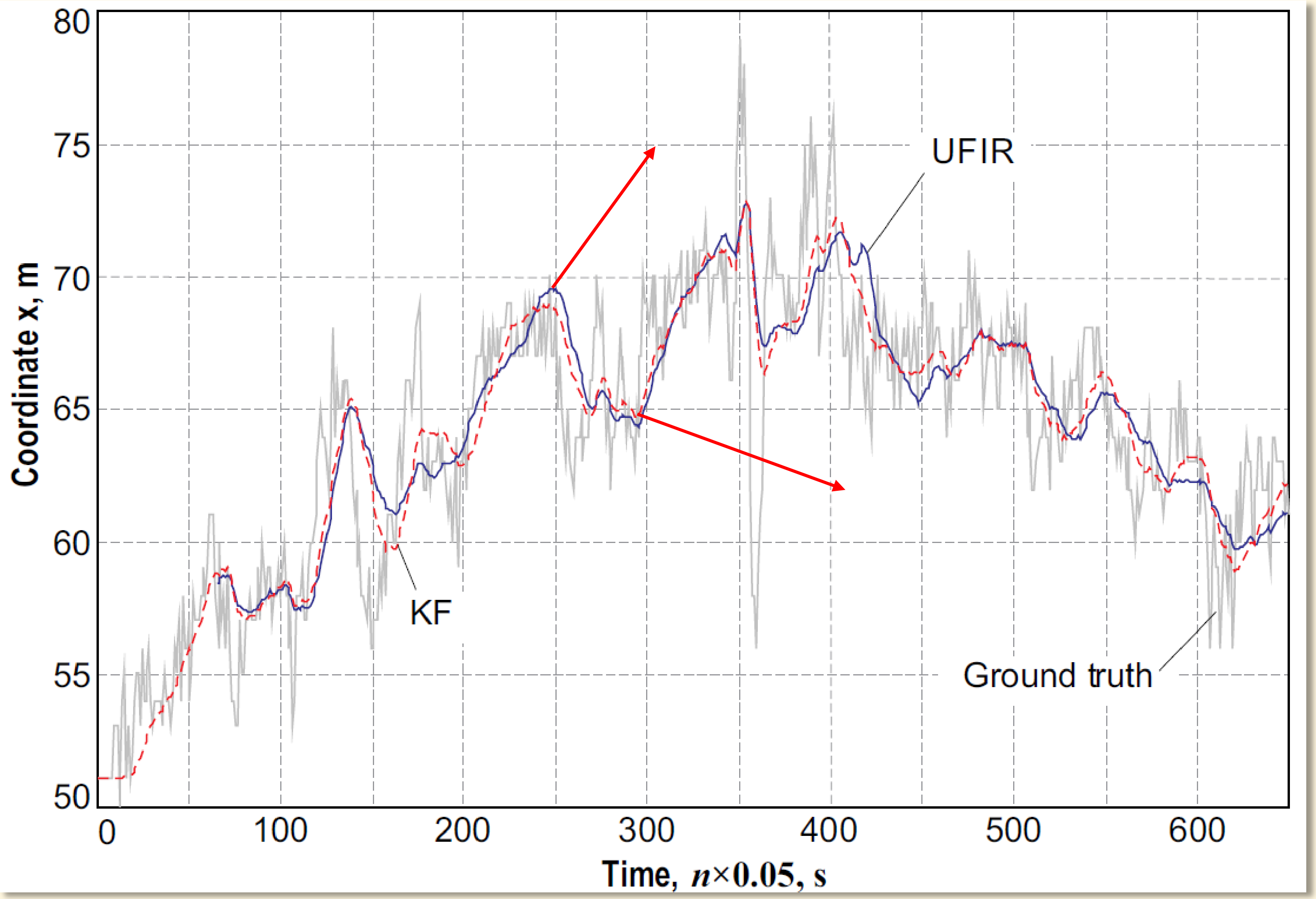


Visual Target Tracking

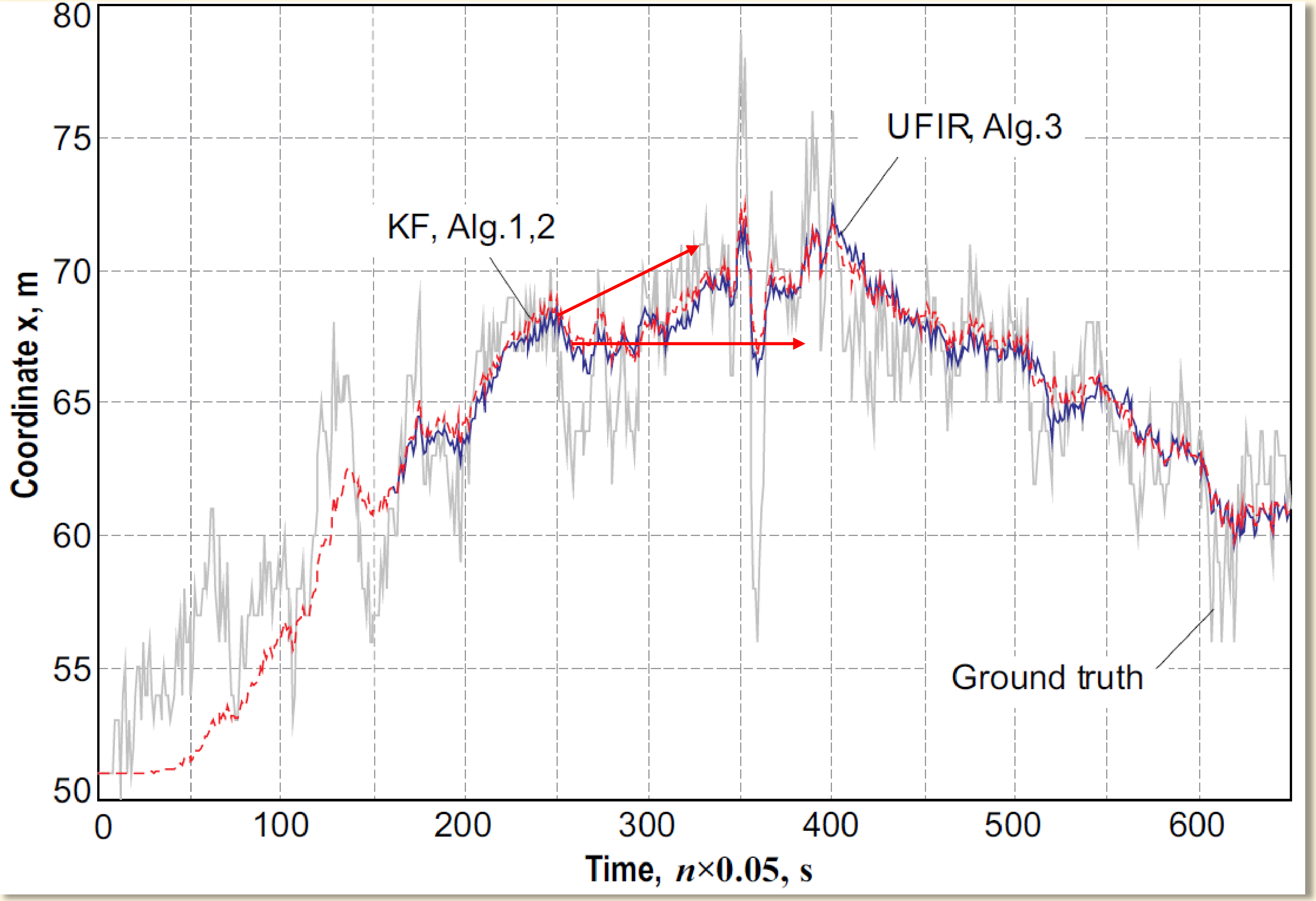
Consider the Car4 benchmark of the visual target tracking

Measurement noise in tracking is colored. But not much information is available. Then set $\tau = 0.05\,\mathrm{s}$ that corresponds to 20 frames/sec, $\sigma_w = 3\,\mathrm{m/s^2}$, $\sigma_\xi = 2\,\mathrm{m}$, and find $N \cong 65$.

Filtering with the standard KF and UFIR filter



Filtering with the KF and UFIR filter modified for CMN



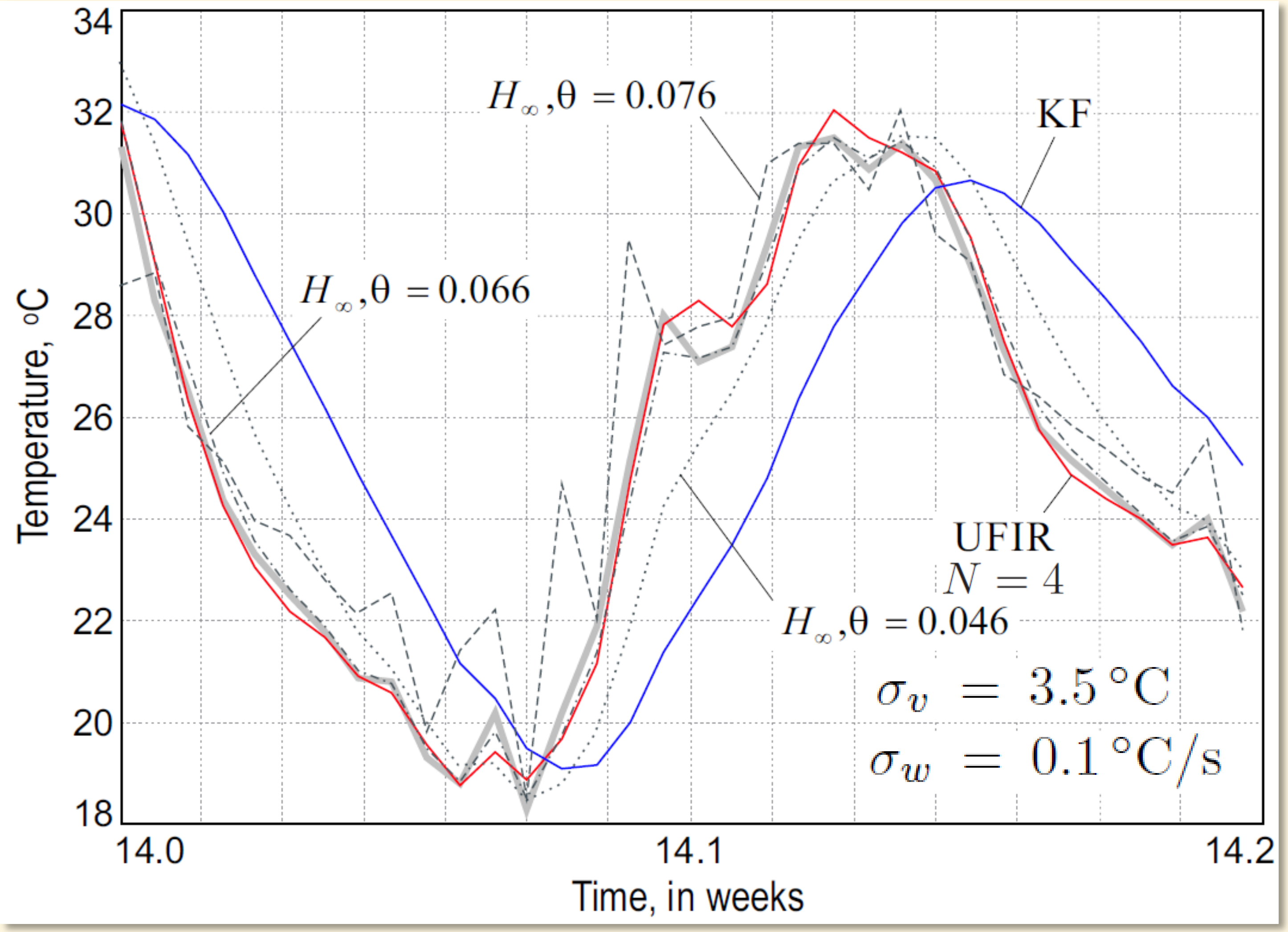
Resume

1. Kalman filter:
    1. For CMN, the modified KF algorithm is general.
    2. All three KF algorithms modified for augmented vectors, corelated noise, and de-corelated noise produced identical estimates and are thus equivalent.
    3. No “illness” was observed for the augmented vectors.
    4. The proof of identities of the KF algorithms for correlated and de-correlated noise sources is still challenging.

1. UFIR filter:
    1. For CMN, the modified UFIR algorithm has appeared to be ill-conditioned by the augmented matrices and unique for correlated and de-correlated noise sources.
    2. The UFIR filter has demonstrated better performance (higher robustness) than the KF algorithms both in simulations and experimentally.

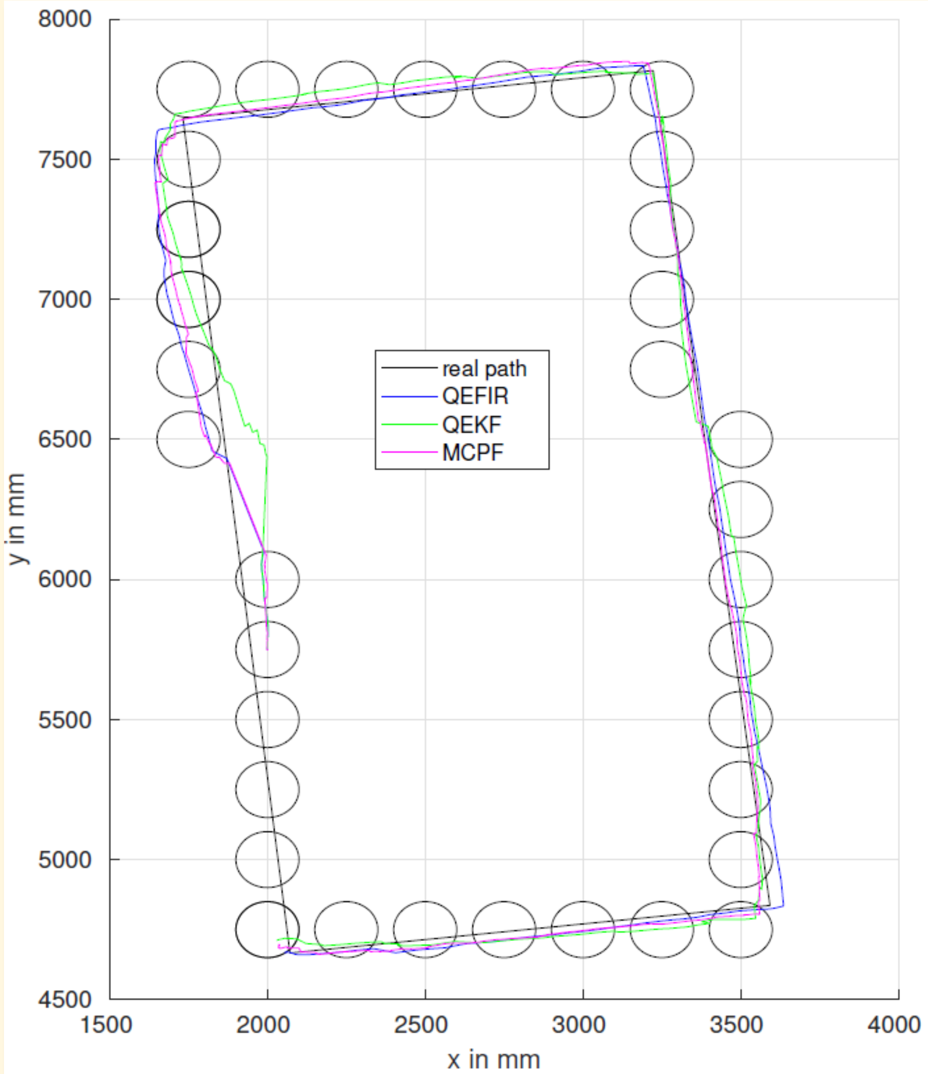
Some Other Applications

Temperature State Tracking



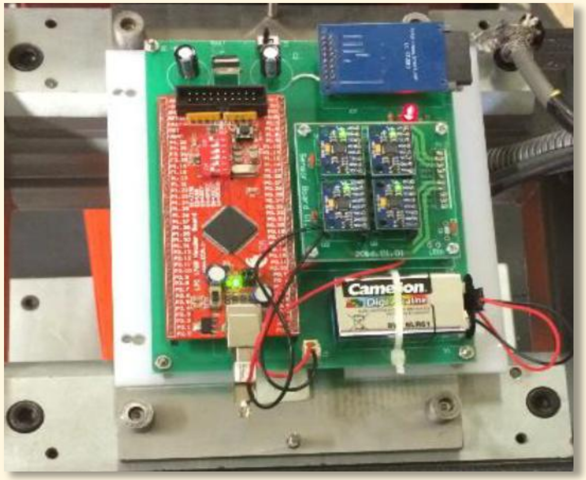
Y. S. Shmaliy, F. Lehmann, S. Zhao, C. K. Ahn, Comparing robustness of the Kalman, Hinf, and UFIR filters, *IEEE Trans. Signal Process.*, **66**, 13, 3447-3458, 2018.

Pose Estimation of Mobile Robots

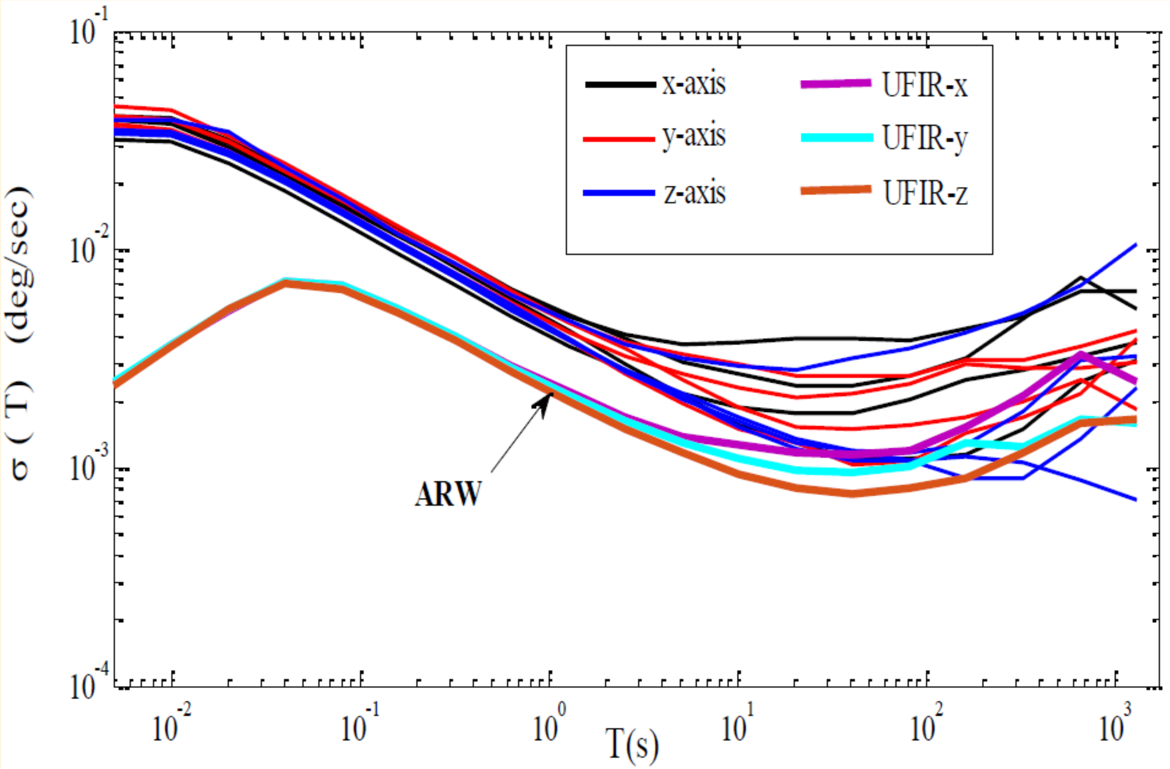


D. Heb, C. Rohrig, Pose estimation of mobile robots with quantized measurements using EFIR filtering: Experimental comparison with the EKF, *Proc. 50th Int. Symp. on Robotics (ISR 2018),* Munich, Germany, June 2018.

Accuracy Improvement of a Multi-MEMS Inertial Measurement Unit by Using an Iterative UFIR Filter



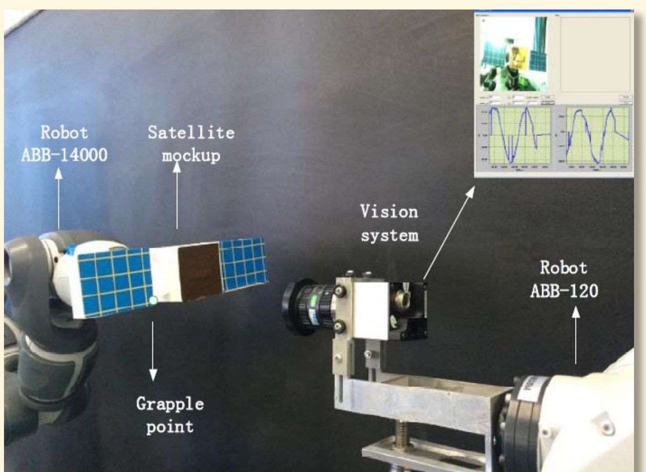
Sensor	ARW ($^\circ/\sqrt{s}$)	Improvement Single sensor/UFIR
Gx1	0.0044	1.83
Gx2	0.0049	2.04
Gx3	0.0057	2.37
Gx4	0.0055	2.29
UFIR	0.0024	

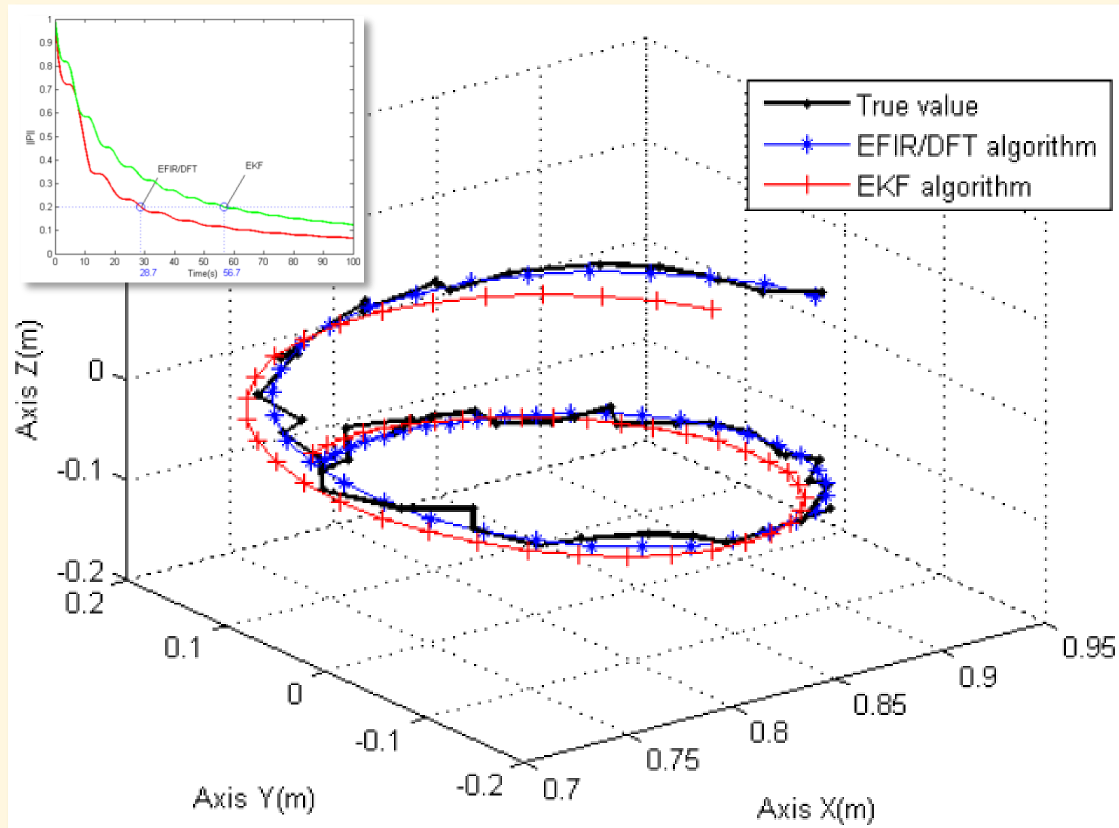


Note that optimal filtering was not applied here in view of *unknown noise.*

R. Rasoulzadeh, A. M. Shahri, Accuracy Improvement of a Multi-MEMS Inertial Measurement Unit by Using an Iterative UFIR Filter with the EKF, *Proc. 2017 European Navigation Conf. (ENC),* Lausanne, Switzerland, May 9-12, 2017, pp. 279-286.

Trajectory Prediction of Space Robot for Capturing Non-cooperative Target





It is stated experimentally that the EFIR filter is more robust than the EKF and can predict trajectory fast and accurately.

D. Han, P. Huang, Trajectory prediction of space robot for capturing non-cooperative target, *Proc. 18th Int. Conf. on Advanced Robotics (ACAR)*, Hong Kong, China, July 2017, pp. 328-333.

Conclusions:

- Under the CMN, the modified *KF* and *UFIR* algorithms produce *better estimates than the standard algorithms*.
- For CMN, all three *modified KF algorithms are equivalent*. The *UFIR* algorithm *is ill-conditioned by the augmented matrices* and has a *unique* form *for correlated and de-correlated noise* sources.
- Modified for CMN, the *UFIR algorithm demonstrates better performance* (higher robustness) than the KF.