exail

RECURSIVE KALMANNET: DEEP LEARNING-AUGMENTED KALMAN FILTER

H. Mortada, C. Falcon, Y. Kahil, M. Clavaud, J.-P. Michel

exail

33rd European Signal Processing Conference, Palermo, September 9th, 2025

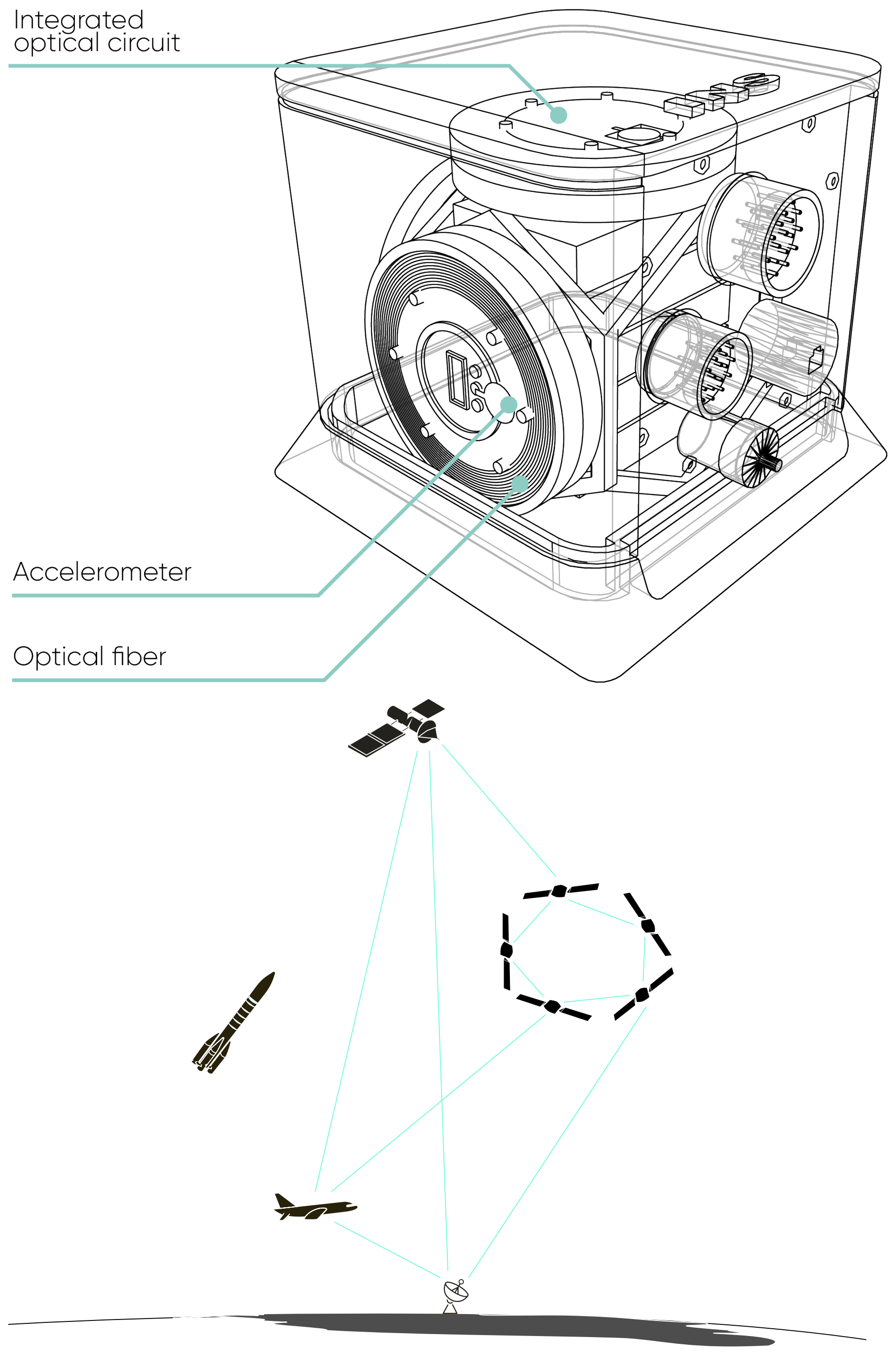
Table of contents

About Exail

- **High-tech French industrial group, with more than 2000 employees in 80 countries. Specialized in:**
  - Robotics
  - **Navigation**
  - Photonics

About Exail

- **High-tech French industrial group, with more than 2000 employees in 80 countries. Specialized in:**
  - Robotics
  - **Navigation**
  - Photonics
- **The navigation division produces extremely precise inertial navigation systems:**
  - Based on Fiber-Optic Gyrometers (FOG) and MEMS accelerometers
  - Output position, velocity and attitude
  - Internal and external sensors (*e.g.* GNSS) are fused using **Kalman Filter (KF)** variants



exail

Problem statement: State and Uncertainty estimation

Linear state-space model:

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$$
$$\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t)$$

- $\mathbf{F}_t$ is the state transition matrix
- $\mathbf{H}_t$ is the observation matrix
- The noise terms $\mathbf{v}_t$ and $\mathbf{w}_t$ are white and independent

exail

Problem statement: State and Uncertainty estimation

- **Linear state-space model:**

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$$
$$\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t)$$

  - $\mathbf{F}_t$ is the state transition matrix
  - $\mathbf{H}_t$ is the observation matrix
  - The noise terms $\mathbf{v}_t$ and $\mathbf{w}_t$ are white and independent

- **The KF is a linear recursif estimator fusing the predictions and the observations via a closed-from gain in a prediction-correction scheme**

exail

Problem statement: State and Uncertainty estimation

- **Linear state-space model:**

$$\mathbf{x}_t = \mathbf{F}_t\mathbf{x}_{t-1} + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$$
$$\mathbf{z}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t)$$

  - $\mathbf{F}_t$ is the state transition matrix
  - $\mathbf{H}_t$ is the observation matrix
  - The noise terms $\mathbf{v}_t$ and $\mathbf{w}_t$ are white and independent

- **The KF is a linear recursif estimator fusing the predictions and the observations via a closed-from gain in a prediction-correction scheme**

- **Noise and KF performance:**

	White	Independant	Gaussian
Best Linear Unbiased Estimator	✗	✗	
Minimal Mean Square Error	✗	✗	✗

- **These noise requirements are rarely met in real-world applications**

Problem statement: State and Uncertainty estimation

› **Linear state-space model:**

$$\mathbf{x}_t = \mathbf{F}_t\mathbf{x}_{t-1} + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$$
$$\mathbf{z}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t)$$

- $\mathbf{F}_t$ is the state transition matrix
- $\mathbf{H}_t$ is the observation matrix
- The noise terms $\mathbf{v}_t$ and $\mathbf{w}_t$ are white and independent

› **The KF is a linear recursif estimator fusing the predictions and the observations via a closed-from gain in a prediction-correction scheme**

› **Noise and KF performance:**

	White	Independant	Gaussian
Best Linear Unbiased Estimator	✗	✗	
Minimal Mean Square Error	✗	✗	✗

› **These noise requirements are rarely met in real-world applications**

› **Goal: overcome the noise requirements and without the prior knwoledge of $\mathbf{Q}_t$ and $\mathbf{R}_t$**

Deep Learning-Augmented Kalman Filtering methods

- **Goal: address the limitations of the traditional KF**
- **Main idea: learn the Kalman gain by leveraging the estimated states in a supervised fashion**
- **Architecture: a Recurrent Neural Network (RNN) with a set of features such as the innovation**

Deep Learning-Augmented Kalman Filtering methods

- **Goal: address the limitations of the traditional KF**
- **Main idea: learn the Kalman gain by leveraging the estimated states in a supervised fashion**
- **Architecture: a Recurrent Neural Network (RNN) with a set of features such as the innovation**
- **State-of-the-art:**
  - KalmanNet [1]: pionneer paper tested on model mismatch and system nonlinearities. It lacks the estimation of the error covariance matrix and the tracking of time-varying gains

[1] G. Revach, N. Shlezinger, X. Ni, A. L. Escoriza, R. J. Van Sloun, and Y. C. Eldar, "KalmanNet: Neural Network Aided Kalman Filtering for Partially Known Dynamics," IEEE Trans. Signal Process., vol. 70, pp. 1532–1547, 2022.

exail

Deep Learning-Augmented Kalman Filtering methods

- **Goal: address the limitations of the traditional KF**
- **Main idea: learn the Kalman gain by leveraging the estimated states in a supervised fashion**
- **Architecture: a Recurrent Neural Network (RNN) with a set of features such as the innovation**
- **State-of-the-art:**
  - KalmanNet [1]: pionneer paper tested on model mismatch and system nonlinearities. It lacks the estimation of the error covariance matrix and the tracking of time-varying gains
  - Split KalmanNet [2]: estimates the gain using two RNN's estimating its components: the predicted error covariance matrix and the inverse of the innovation covariance

[1] G. Revach, N. Shlezinger, X. Ni, A. L. Escoriza, R. J. Van Sloun, and Y. C. Eldar, "KalmanNet: Neural Network Aided Kalman Filtering for Partially Known Dynamics," IEEE Trans. Signal Process., vol. 70, pp. 1532–1547, 2022.
[2] G. Choi, J. Park, N. Shlezinger, Y. C. Eldar, and N. Lee, "Split- KalmanNet: A Robust Model-Based Deep Learning Approach for State Estimation," IEEE Trans. Veh. Technol., vol. 72, pp. 12 326–12 331, 2023.

Deep Learning-Augmented Kalman Filtering methods

- **Goal: address the limitations of the traditional KF**
- **Main idea: learn the Kalman gain by leveraging the estimated states in a supervised fashion**
- **Architecture: a Recurrent Neural Network (RNN) with a set of features such as the innovation**
- **State-of-the-art:**
  - KalmanNet [1]: pionneer paper tested on model mismatch and system nonlinearities. It lacks the estimation of the error covariance matrix and the tracking of time-varying gains
  - Split KalmanNet [2]: estimates the gain using two RNN's estimating its components: the predicted error covariance matrix and the inverse of the innovation covariance
  - Cholesky KalmanNet [3]: add-on split KalmanNet to output the error covariance matrix by calculating its Cholesky factor to ensure its Cholesky factor. Trade-off between state estimation and covariance error matrix accuracy due to the cost function choice

[1] G. Revach, N. Shlezinger, X. Ni, A. L. Escoriza, R. J. Van Sloun, and Y. C. Eldar, "KalmanNet: Neural Network Aided Kalman Filtering for Partially Known Dynamics," IEEE Trans. Signal Process., vol. 70, pp. 1532–1547, 2022.
[2] G. Choi, J. Park, N. Shlezinger, Y. C. Eldar, and N. Lee, "Split- KalmanNet: A Robust Model-Based Deep Learning Approach for State Estimation," IEEE Trans. Veh. Technol., vol. 72, pp. 12 326–12 331, 2023.
[3] M. Ko and A. Shafieezadeh, "Cholesky-KalmanNet: Model-Based Deep Learning With Positive Definite Error Covariance Structure," IEEE Signal Process. Lett., vol. 32, pp. 326–330, 2025.

Proposed method: Recursive KalmanNet (RKN)

- **Deep learning model inspired by Kalman filtering, preserving the prediction–correction scheme**
- **Operates without prior knowledge of noise covariance matrices $\mathbf{Q}_t$ and $\mathbf{R}_t$, and without assuming any noise model**
- **Composed of two RNNs:**
  - One dedicated to direct gain estimation
  - Another contributes to error covariance estimation using the generic Joseph's formulation
- **First method to yield accurate state estimation and consistent error covariance thanks to the architecture and tailored cost function**

Kalman Filter

- **Prediction:**

$$\widehat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \widehat{\mathbf{x}}_{t-1|t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^{\top} + \mathbf{Q}_t$$

Kalman Filter

- **Prediction:**

$$\widehat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \widehat{\mathbf{x}}_{t-1|t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}_t$$

- **Gain:**

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^\top \left(\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^\top + \mathbf{R}_t \right)^{-1}$$

exail

Kalman Filter

- **Prediction:**

$$\widehat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \widehat{\mathbf{x}}_{t-1|t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^{\top} + \mathbf{Q}_t$$

- **Gain:**

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^{\top} \left(\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^{\top} + \mathbf{R}_t \right)^{-1}$$

- **Correction:**

$$\begin{aligned} \widehat{\mathbf{x}}_{t|t} &= \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t (\mathbf{z}_t - \mathbf{H}_t \widehat{\mathbf{x}}_{t|t-1}) \\ &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \mathbf{z}_t \end{aligned}$$

$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1} (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^{\top} + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^{\top}$$

Kalman Filter

- **Prediction:**

$$\widehat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \widehat{\mathbf{x}}_{t-1|t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}_t$$

- **Gain:**

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^\top \left(\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^\top + \mathbf{R}_t \right)^{-1}$$

- **Correction:**

$$\begin{aligned} \widehat{\mathbf{x}}_{t|t} &= \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t (\mathbf{z}_t - \mathbf{H}_t \widehat{\mathbf{x}}_{t|t-1}) \\ &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \mathbf{z}_t \end{aligned}$$

$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1} (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^\top \quad \text{[Joseph's formula: valid for any gain]}$$

$$= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1} \quad \text{[Simplified formulation after plugging-in the gain]}$$

Kalman Filter

- **Prediction:**

$$\widehat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \widehat{\mathbf{x}}_{t-1|t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}_t$$

- **Gain:**

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^\top \left(\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^\top + \mathbf{R}_t \right)^{-1}$$

- **Correction:**

$$\begin{aligned} \widehat{\mathbf{x}}_{t|t} &= \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t (\mathbf{z}_t - \mathbf{H}_t \widehat{\mathbf{x}}_{t|t-1}) \\ &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \mathbf{z}_t \end{aligned}$$

$$\begin{aligned} \mathbf{P}_{t|t} &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1} (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^\top \\ &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) (\mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}_t) (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^\top \end{aligned}$$

exail

Kalman Filter

- **Prediction:**

$$\widehat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \widehat{\mathbf{x}}_{t-1|t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}_t$$

- **Gain:**

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^\top \left(\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^\top + \mathbf{R}_t \right)^{-1}$$

- **Correction:**

$$\begin{aligned} \widehat{\mathbf{x}}_{t|t} &= \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}_t \widehat{\mathbf{x}}_{t|t-1}) \\ &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \mathbf{z}_t \end{aligned}$$

$$\begin{aligned} \mathbf{P}_{t|t} &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1} (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^\top \\ &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)(\mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}_t)(\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^\top \\ &= \underbrace{(\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top}_{\mathbf{A}_t} + \underbrace{(\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{Q}_t (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^\top}_{\mathbf{B}_t} \end{aligned}$$

Kalman Filter

- **Prediction:**

$$\widehat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \widehat{\mathbf{x}}_{t-1|t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}_t$$

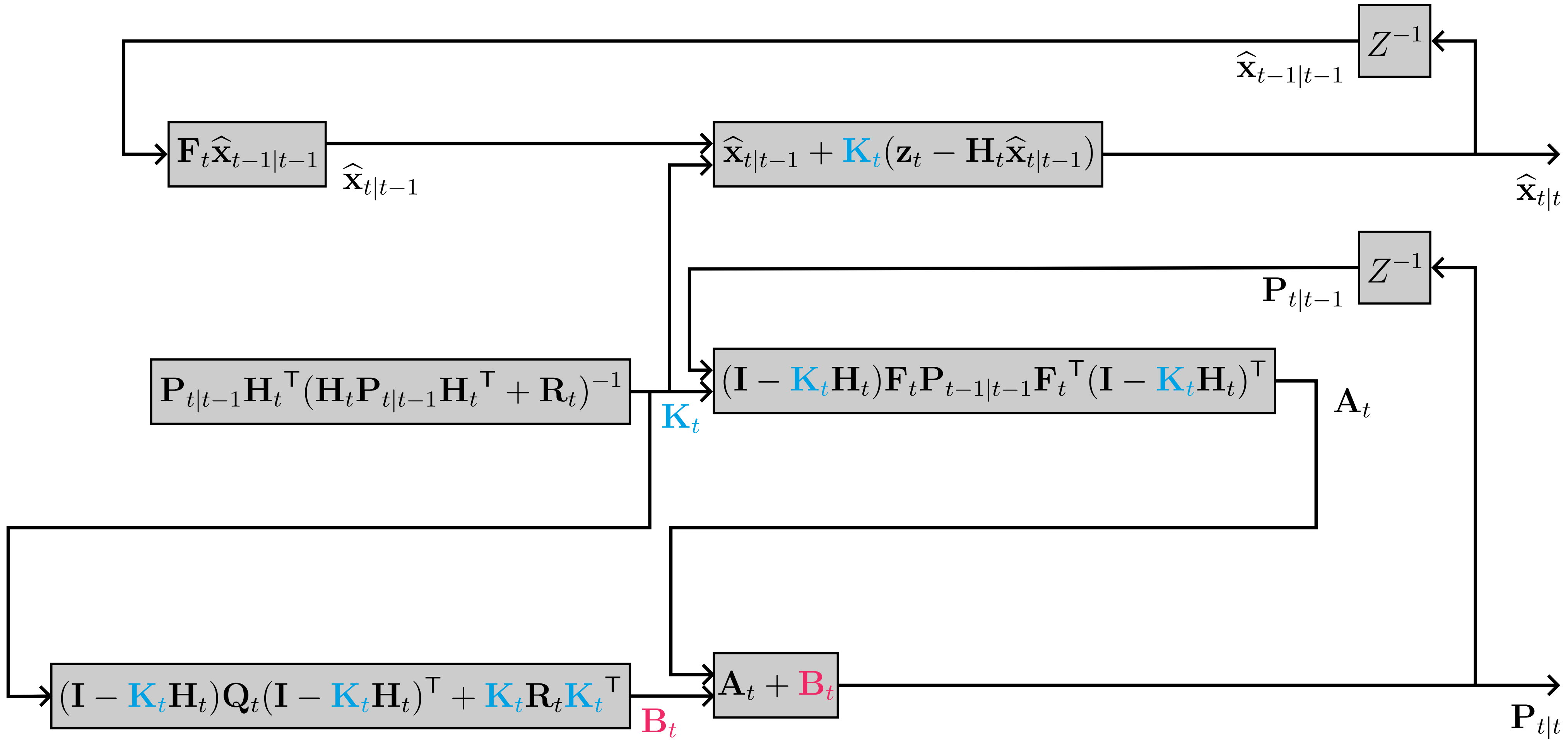
- **Gain:**

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^\top \left(\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^\top + \mathbf{R}_t \right)^{-1}$$

- **Correction:**

$$\begin{aligned} \widehat{\mathbf{x}}_{t|t} &= \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t (\mathbf{z}_t - \mathbf{H}_t \widehat{\mathbf{x}}_{t|t-1}) \\ &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \widehat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \mathbf{z}_t \end{aligned}$$

$$\begin{aligned} \mathbf{P}_{t|t} &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1} (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^\top \\ &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)(\mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}_t)(\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^\top \\ &= \underbrace{(\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top}_{\mathbf{A}_t} + \underbrace{(\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{Q}_t (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)^\top + \mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^\top}_{\mathbf{B}_t} \end{aligned}$$



Recursive KalmanNet (RKN) Architecture

- **Prediction:**

  $\widehat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \widehat{\mathbf{x}}_{t-1|t-1}$

  ~~$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^{\mathsf{T}} + \mathbf{Q}_t$~~

- **Gain:**

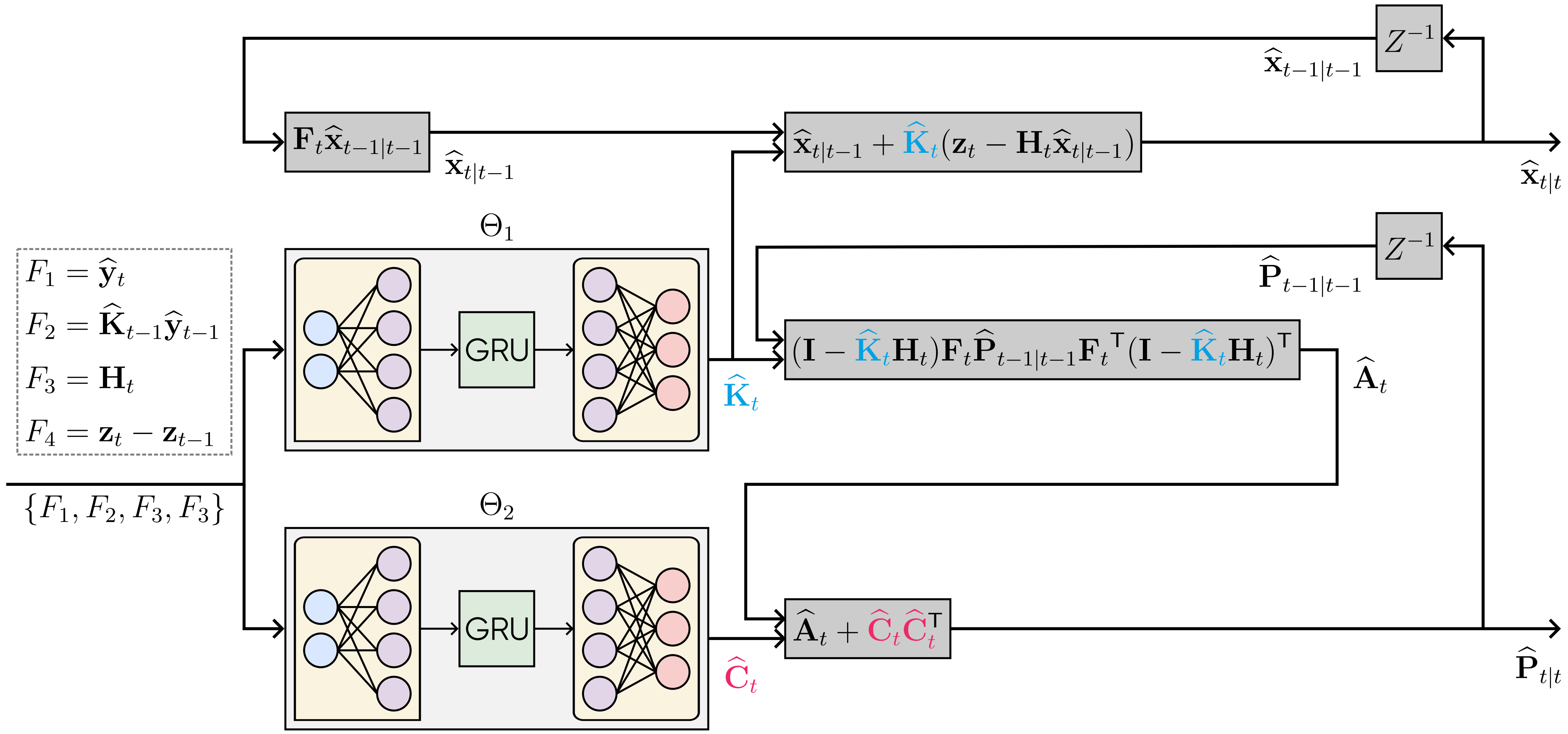
  $\widehat{\mathbf{K}}_t = \mathrm{RNN}_{\Theta_1}(F_1,\ F_2,\ F_3,\ F_4)$

- **Correction:**

  $\widehat{\mathbf{x}}_{t|t} = \widehat{\mathbf{x}}_{t|t-1} + \widehat{\mathbf{K}}_t(\mathbf{z}_t - \mathbf{H}_t \widehat{\mathbf{x}}_{t|t-1})$

  $\widehat{\mathbf{C}}_t = \mathrm{RNN}_{\Theta_2}(F_1,\ F_2,\ F_3,\ F_4)$

  $$\widehat{\mathbf{P}}_{t|t} = \underbrace{(\mathbf{I} - \widehat{\mathbf{K}}_t \mathbf{H}_t)\mathbf{F}_t \widehat{\mathbf{P}}_{t-1|t-1} \mathbf{F}_t^{\mathsf{T}} (\mathbf{I} - \widehat{\mathbf{K}}_t \mathbf{H}_t)^{\mathsf{T}}}_{\widehat{\mathbf{A}}_t} + \underbrace{\widehat{\mathbf{C}}_t \widehat{\mathbf{C}}_t^{\mathsf{T}}}_{\widehat{\mathbf{B}}_t}$$



- **Training is performed by gradient descent on the negative Gaussian log-likelihood of the error estimation:**

$$\mathbf{e}_t^{(i)}(\Theta_1)^{\mathsf{T}} \widehat{\mathbf{P}}_{t|t}^{(i)}(\Theta_1, \Theta_2)^{-1} \mathbf{e}_t^{(i)}(\Theta_1) + \log\det \widehat{\mathbf{P}}_{t|t}^{(i)}(\Theta_1, \Theta_2),$$

averaged over time and batch, where $\mathbf{e}_t^{(i)} = \widehat{\mathbf{x}}_t^{(i)}(\Theta_1) - \mathbf{x}_t$.

exail

Case study: Heavy-tailed bimodal-Gaussian measurement noise

- 1D constant-speed linear model, with a position measurement

$$
\begin{aligned}
\mathbf{x}_t &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}_{t-1} + \mathbf{v}_t, \quad && \mathbf{x}_t \in \mathbb{R}^2 \\
z_t &= \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}_t + w_t, && z_t \in \mathbb{R}.
\end{aligned}
$$

Case study: Heavy-tailed bimodal-Gaussian measurement noise

- **1D constant-speed linear model, with a position measurement**

$$\mathbf{x}_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}_{t-1} + \mathbf{v}_t, \quad \mathbf{x}_t \in \mathbb{R}^2$$
$$z_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}_t + w_t, \qquad z_t \in \mathbb{R}.$$

- **The process noise $\mathbf{v}_t$ is zero-mean Gaussian white noise with covariance $\mathbf{Q}_t = \begin{bmatrix} 0 & 0 \\ 0 & {\sigma_v}^2 \end{bmatrix}$**

exail

Case study: Heavy-tailed bimodal-Gaussian measurement noise

- **1D constant-speed linear model, with a position measurement**

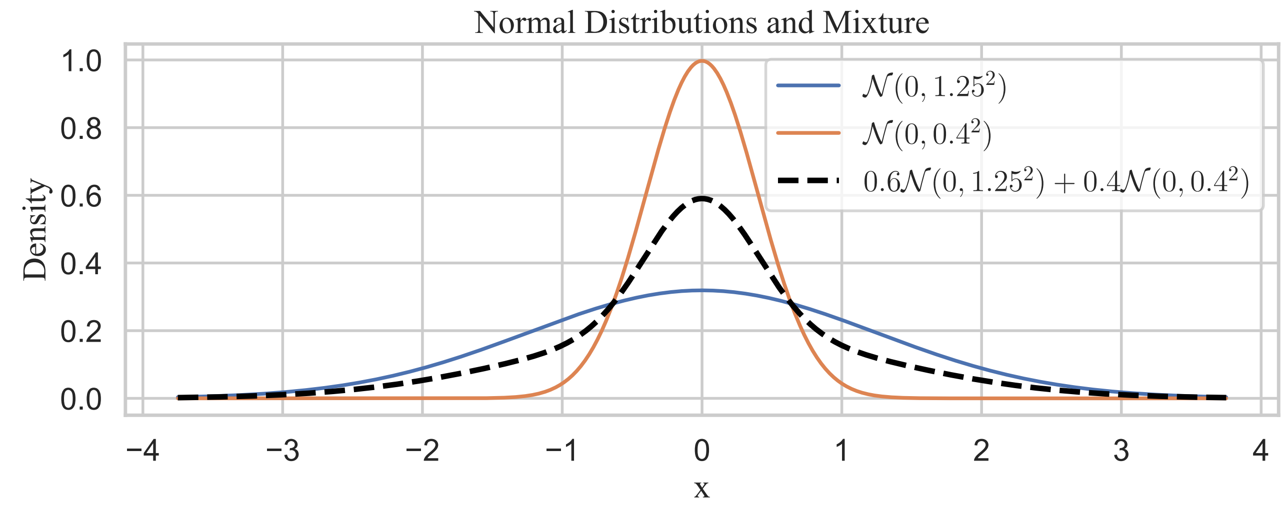
$$\mathbf{x}_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}_{t-1} + \mathbf{v}_t, \quad \mathbf{x}_t \in \mathbb{R}^2$$
$$z_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}_t + w_t, \quad z_t \in \mathbb{R}.$$

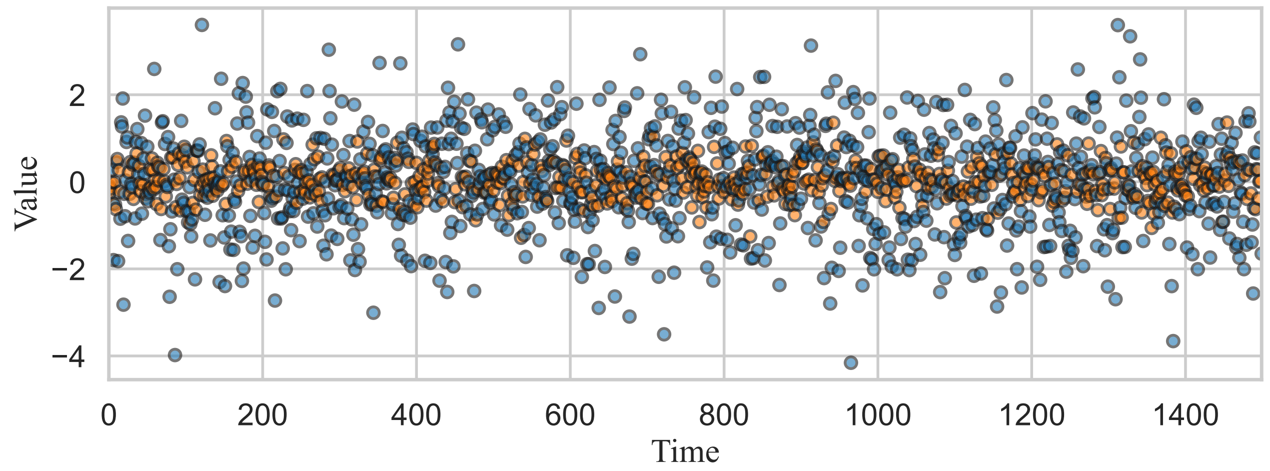
- **The process noise $\mathbf{v}_t$ is zero-mean Gaussian white noise with covariance $\mathbf{Q}_t = \begin{bmatrix} 0 & 0 \\ 0 & {\sigma_v}^2 \end{bmatrix}$**

- **The measurement noise w_t follows a heavy-tailed bimodal-Gaussian distribution:**

$$w_t = Z_t X_t + (1 - Z_t) Y_t$$

  - X_t, Y_t, and Z_t are independent white noise processes
  - X_t and Y_t are Gaussian with variances ${\sigma_1}^2$ and ${\sigma_2}^2$ respectively
  - Z_t is Bernoulli with parameter p
  - w_t is distributed as $p\mathcal{N}(0, {\sigma_1}^2) + (1-p)\mathcal{N}(0, {\sigma_2}^2)$. It has variance $\mathbf{R}_t = Z_t {\sigma_1}^2 + (1 - Z_t){\sigma_2}^2$, and an expected variance of ${\sigma_w}^2 = p{\sigma_1}^2 + (1-p){\sigma_2}^2$





exail

Results: Performance at varying noise heterogeneity levels

- **Numerical application:** ${\sigma_v}^2 = 10^{-4}, \quad {\sigma_1}^2 = 1.56{\sigma_w}^2, \quad p = 0.6, \quad N = 1000$ (test set size)

Results: Performance at varying noise heterogeneity levels

- **Numerical application:** ${\sigma_v}^2 = 10^{-4}, \quad {\sigma_1}^2 = 1.56{\sigma_w}^2, \quad p = 0.6, \quad N = 1000$ (test set size)
- **Compared methods:**
  - O-KF: Optimal KF with time-varying $\mathbf{R}_t = Z_t{\sigma_1}^2 + (1 - Z_t){\sigma_2}^2$ (reference method)
  - SO-KF: Sub-Optimal KF with constant $\mathbf{R}_t = {\sigma_w}^2 = p{\sigma_1}^2 + (1 - p){\sigma_2}^2$
  - Cholesky KalmanNet (CKN)
  - Recursive KalmanNet (RKN)

exail

Results: Performance at varying noise heterogeneity levels

- **Numerical application:** ${\sigma_v}^2 = 10^{-4}, \quad {\sigma_1}^2 = 1.56{\sigma_w}^2, \quad p = 0.6, \quad N = 1000$ (test set size)
- **Compared methods:**
  - O-KF: Optimal KF with time-varying $\mathbf{R}_t = Z_t{\sigma_1}^2 + (1 - Z_t){\sigma_2}^2$ (reference method)
  - SO-KF: Sub-Optimal KF with constant $\mathbf{R}_t = {\sigma_w}^2 = p{\sigma_1}^2 + (1 - p){\sigma_2}^2$
  - Cholesky KalmanNet (CKN)
  - Recursive KalmanNet (RKN)
- **Metrics:** $\textbf{MSE} = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\sum_{i=1}^{N}\mathbf{e}_t^{(i)\top}\mathbf{e}_t^{(i)} \quad \textbf{MSMD} = \frac{1}{T}\sum_{t=1}^{T}\underbrace{\frac{1}{N}\sum_{i=1}^{N}\mathbf{e}_t^{(i)\top}{\mathbf{P}_{t|t}^{(i)}}^{-1}\mathbf{e}_t^{(i)}}_{\textbf{CLT}\Rightarrow\sim\mathcal{N}(2,\ 4/N)}$

exail

Results: Performance at varying noise heterogeneity levels

- **Numerical application:** $\sigma_v{}^2 = 10^{-4}, \quad \sigma_1{}^2 = 1.56\sigma_w{}^2, \quad p = 0.6, \quad N = 1000$ (test set size)
- **Compared methods:**
  - O-KF: Optimal KF with time-varying $\mathbf{R}_t = Z_t\sigma_1{}^2 + (1 - Z_t)\sigma_2{}^2$ (reference method)
  - SO-KF: Sub-Optimal KF with constant $\mathbf{R}_t = \sigma_w{}^2 = p\sigma_1{}^2 + (1 - p)\sigma_2{}^2$
  - Cholesky KalmanNet (CKN)
  - Recursive KalmanNet (RKN)
- **Metrics:** $\text{MSE} = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N}\sum_{i=1}^{N}\mathbf{e}_t^{(i)\mathsf{T}}\mathbf{e}_t^{(i)}$ $\quad \text{MSMD} = \frac{1}{T}\sum_{t=1}^{T}\underbrace{\frac{1}{N}\sum_{i=1}^{N}\mathbf{e}_t^{(i)\mathsf{T}}\mathbf{P}_{t|t}^{(i)-1}\mathbf{e}_t^{(i)}}_{\text{CLT}\Rightarrow\sim\mathcal{N}(2,\,4/N)}$
- **MSE and MSMD for varying noise heteroginity** $\nu = 10\log\frac{\sigma_w{}^2}{\sigma_v{}^2}$

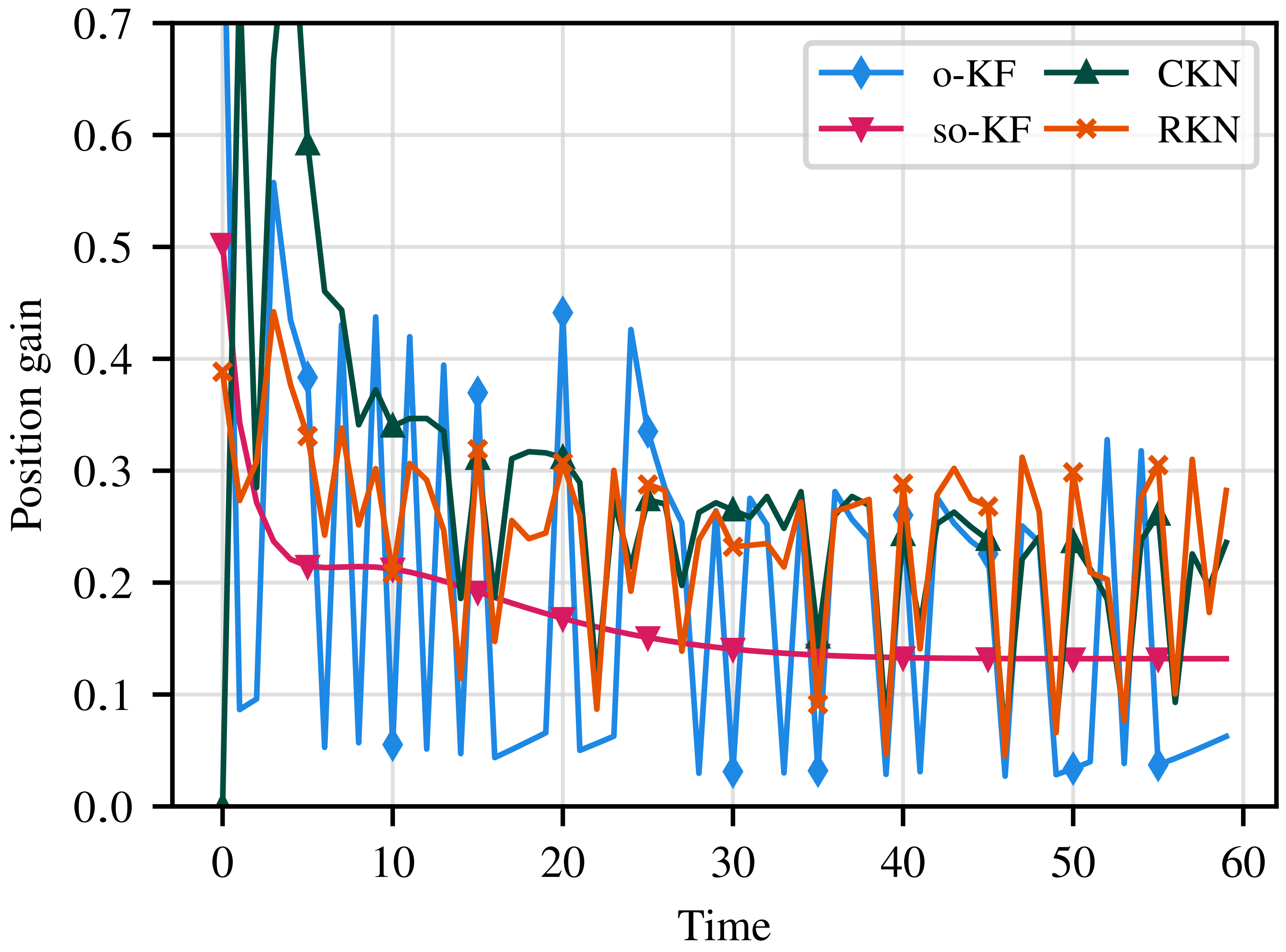
ν	20		30		40		50		60	
	MSE	**MSMD**	**MSE**	**MSMD**	**MSE**	**MSMD**	**MSE**	**MSMD**	**MSE**	**MSMD**
o-KF	-28	2.0	-21	2.0	-14	2.0	-6.9	2.0	0.1	2.0
so-KF	**−26**	2.0	-18	2.0	-11	2.0	-3.9	2.0	3.4	2.0
CKN	-20	4.0	-17	27	-11	3.2	-4.8	6.2	2.4	21
RKN	**−26**	2.0	**−19**	2.0	**−12**	2.0	**−5.1**	1.9	**2.3**	2.2

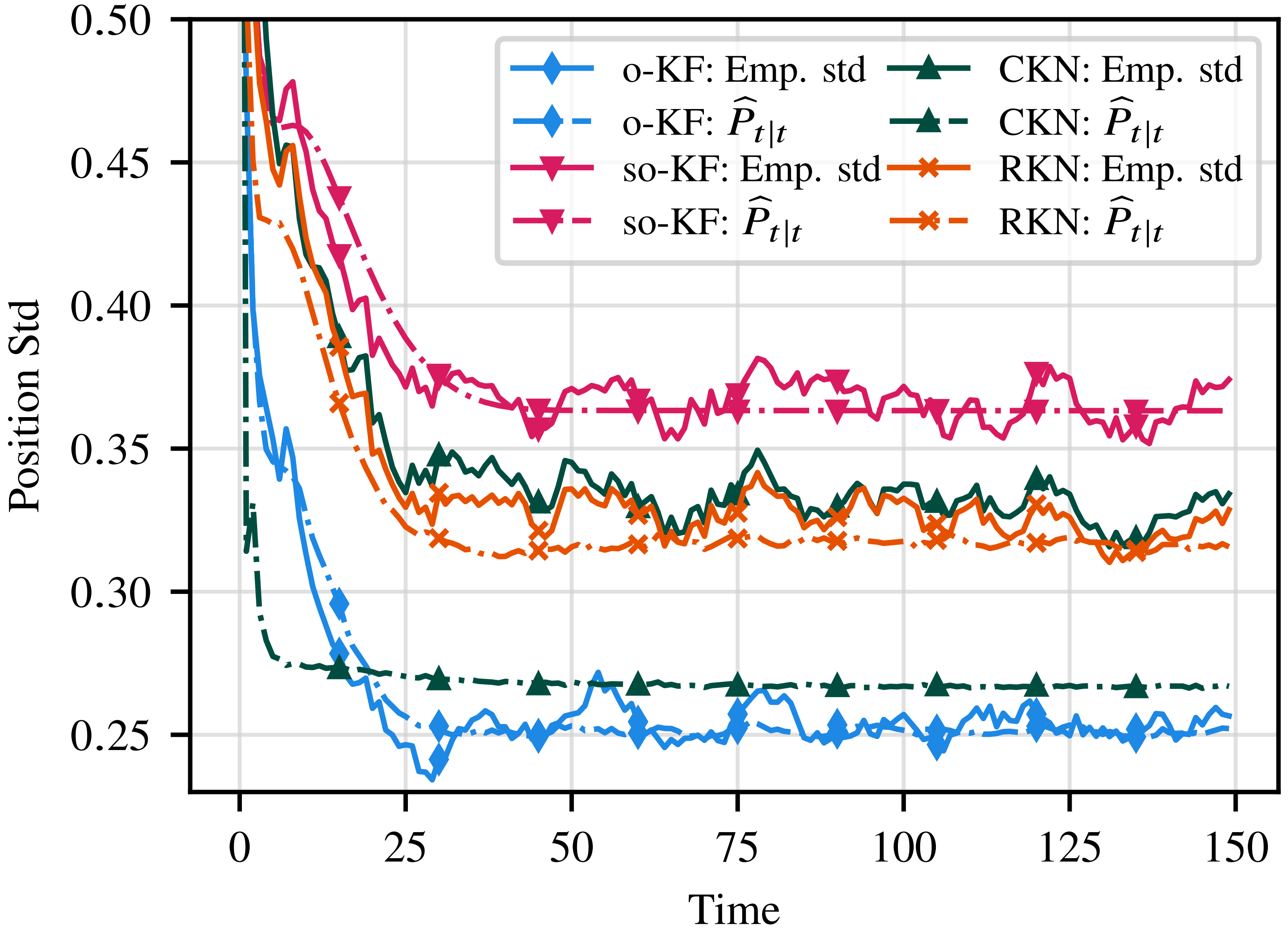
- **RKN gives closest MSE to O-KF with consistent error covariance reflected by MSMD close to** 2

exail

Results: Position gain and error standard deviation estimations

- **fixed** $\nu = 40$ **dB**



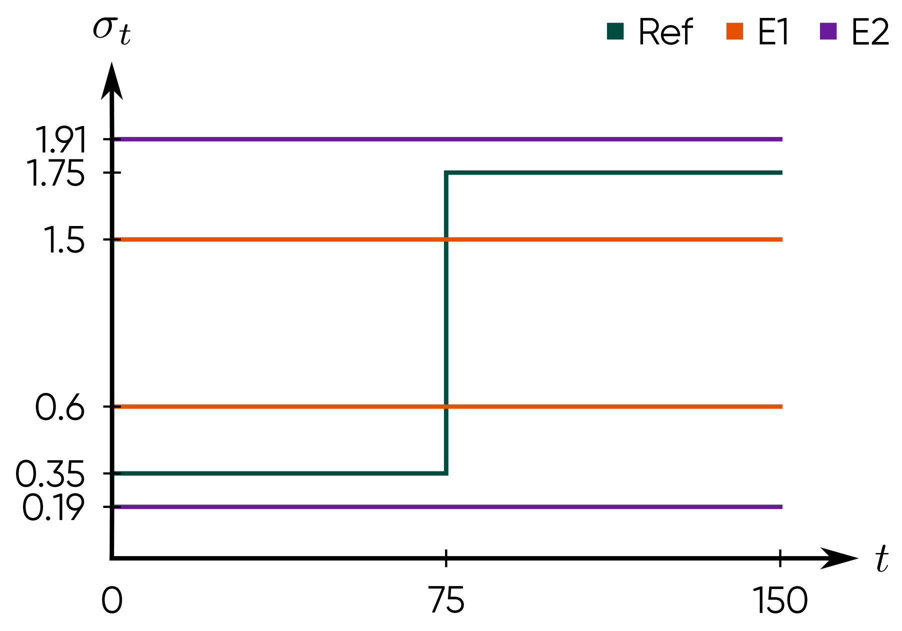


- **RKN is able to track challenging time-varying gain behavior**
- **RKN provides consistent position error covariance of the state error and is better with the conventional so-KF**
- **CKN error covariance is not consistent with the empirical error**

exail

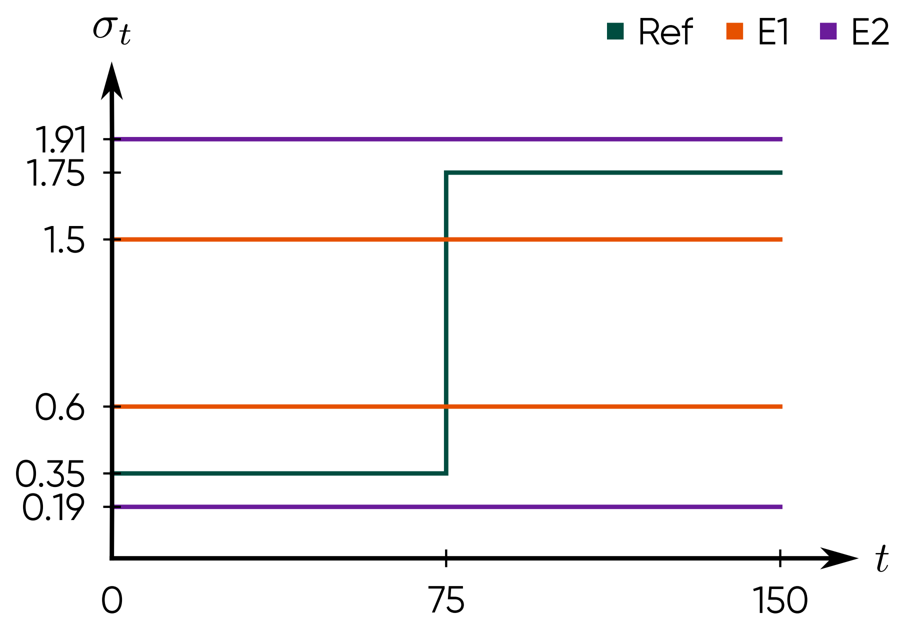
Results: Model generalization capabilities

- $\mathbf{R}_t = {\sigma_t}^2$

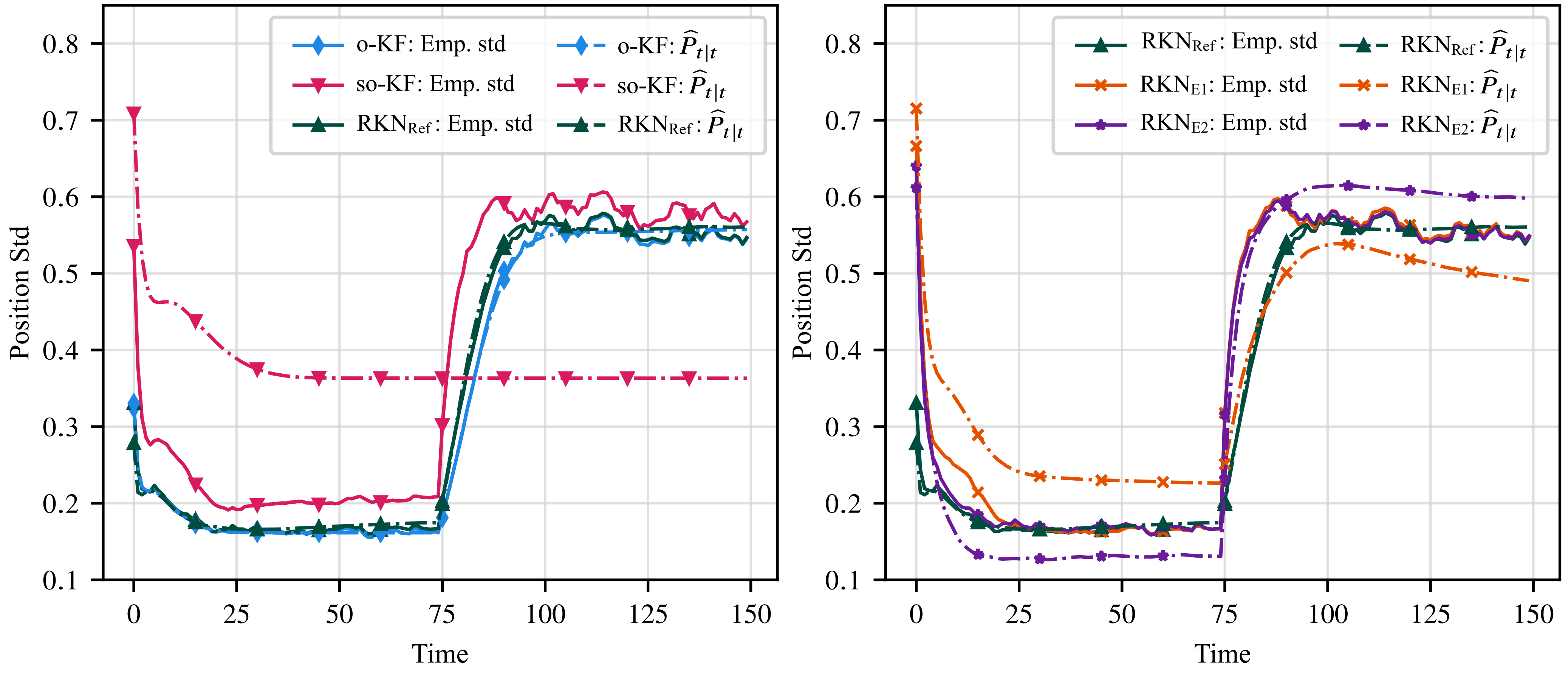


Results: Model generalization capabilities

- $\mathbf{R}_t = {\sigma_t}^2$



- **Training with REF, E1 or E2. Test with REF**



- **No degradation in estimation precision.**
- **Loss of uncertainty consistency in error covariance estimation.**

[4] C. Falcon, H. Mortada, M. Clavaud, and J.-P. Michel, "Recursive KalmanNet : Analyse des capacités de généralisation d'un réseau de neurones récurrent guidé par un filtre de Kalman," in 30e Colloque sur le traitement du signal et des images. GRETSI, 2025.

Conclusion

- RKN, a deep-learning augmented Kalman filter that leverages two recurrent neural networks to estimate gain and error covariance using a formulation derived from Joseph's equation
- In scenarios with bimodal Gaussian noise, RKN outperformes conventional Kalman filters and a state-of-the-art deep learning approach, delivering accurate state and covariance estimates that closely reflect actual errors
- RKN shows strong potential in cases where classical Kalman filtering falls short, such as in non-linear systems or with incomplete state-space models

Thank you for your attention!

exail

code + articles