A Unified Perspective on Time-Domain and Frequency-Domain Kalman Filters for Acoustic System Identification

Tobias Kabzinski and Peter Jax
Institute of Communication Systems (IKS), *RWTH Aachen University*, Aachen, Germany
{kabzinski,jax}@iks.rwth-aachen.de

***Abstract*—Time-domain and frequency-domain Kalman filters are applied as adaptive filters to many audio signal processing problems, especially to acoustic system identification. As the formal relationship between these two Kalman filters is not obvious, we connect them using a unified perspective. To this end, we show the equivalence of both filters as a special case and transfer concepts from one domain to the other. Specifically, the concept of considering multiple past observations is transferred to the frequency-domain Kalman filters, and the concept of block-based processing is transferred to the time-domain Kalman filters. Transferring these concepts yields new Kalman filter variants that allow for a different trade-off between convergence speed, steady-state performance and computational complexity.**

***Index Terms*—Adaptive filters, Kalman filtering, system identification**

I. Introduction

Adaptive filters are frequently applied for acoustic system identification. Kalman filters (KFs) represent one popular family of adaptive filters. They are based on a state-space description and estimate a state vector which completely characterizes a linear acoustic system. The applications in audio signal processing are numerous. For example, KFs have been applied in acoustic echo control (AEC) to identify the echo path [1], [2]. Furthermore, KFs have been utilized in active noise control (ANC) [3], [4] and for the online [5] or offline [6] measurement of head-related transfer functions (HRTFs). Due to the specific needs and constraints in these applications, many variants of KFs have been proposed.

On the one hand, there are time-domain KFs, such as in [2]–[4]. Here, the state vector represents impulse response coefficients of the system to be identified. The time-domain KF has been extended to consider multiple past signal samples and has been related to the affine projection algorithm in [7]. Theoretical performance bounds were investigated [8], [9], and an efficient approximation has been proposed in [10].

On the other hand, in frequency-domain KFs the state vector represents complex-valued discrete Fourier transform coefficients. The frequency-domain formulation allows for a different efficient approximation and is related to frequency-domain adaptive filters [1]. There are extensions to multiple-input-single-output systems [11] and non-linear systems [12]. A partitioned-block variant to estimate long filters has been proposed in [13]. More recently, frequency-domain KFs have been extended to incorporate prior knowledge about the systems to be identified [14].

As the time-domain and the frequency-domain approaches have evolved separately in the past, they may appear disconnected. However, it is the purpose of this contribution to connect the two approaches and to provide a unified perspective. After reviewing the time-domain and the frequency-domain KFs in Sec. II, we show that both approaches are equivalent under certain conditions in Sec. III. Then, we establish new KF variants by transferring time-domain concepts to the frequency domain and vice versa in Sec. IV. Finally, we evaluate and compare the different variants (Sec. V), and conclude in Sec. VI.

II. Review of Kalman Filters for System Identification

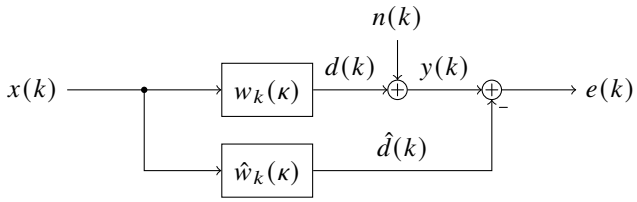

Fig. 1. Signal model for system identification.

We consider the single-channel system identification setup as depicted in Fig. 1. The discrete-time input signal $x(k)$ is filtered with the, possibly time-variant, filter $w_k(\kappa)$ of length L producing the non-distorted, and non-observable, output signal

$$d(k) = \sum_{\kappa=0}^{L-1} w_k(\kappa)\, x(k-\kappa)\,. \tag{1}$$

A noisy output signal $y(k) = d(k) + n(k)$ can be observed. The adaptive filter $\hat{w}_k(\kappa)$ estimates the output signal as $\hat{d}(k)$.

A. Time-Domain Kalman Filter (TKF)

The signal model for the time-domain Kalman filter [2] describes the impulse response coefficient vector

$$\mathbf{w}_{\mathrm{t},k} = \left[w_{\mathrm{t},k}(0)\,, \ldots, w_{\mathrm{t},k}(L-1)\right]^{\mathrm{T}} \in \mathbb{R}^{L} \tag{2}$$

as a multidimensional random process. Here, $(\cdot)^{\mathrm{T}}$ denotes the transpose of a vector or matrix and the subscript "t" indicates a time-domain quantity. Throughout this paper bold capital letters denote matrices, while bold lower case letters denote vectors.

In some cases the vector notation is kept for consistency even if a vector simplifies to a scalar. The coefficients' evolution over time is assumed to be described by the state equation

$$\mathbf{w}_{\mathrm{t},k} = \mathbf{A}_\mathrm{t}\mathbf{w}_{\mathrm{t},k-1} + \mathbf{q}_{\mathrm{t},k} \tag{3}$$

with the state transition matrix $\mathbf{A}_\mathrm{t}$ and zero-mean normally-distributed process noise $\mathbf{q}_{\mathrm{t},k} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_\mathrm{t})$. Typically, a Markov model with scalar fading factor $0 \ll \gamma \leq 1$ is used, i.e., $\mathbf{A}_\mathrm{t} = \gamma\mathbf{I}_L$, where $\mathbf{I}$ is the $L\times L$ identity matrix. The process noise variance per coefficient is denoted by σ_q^2 such that $\mathbf{Q}_\mathrm{t} = \sigma_q^2\mathbf{I}_L$. The observation equation is given by

$$y(k) = \mathbf{y}_{\mathrm{t},k} = \mathbf{C}_{\mathrm{t},k}\mathbf{w}_{\mathrm{t},k} + n(k) = \mathbf{C}_{\mathrm{t},k}\mathbf{w}_{\mathrm{t},k} + \mathbf{n}_{\mathrm{t},k}, \tag{4}$$

where normally-distributed observation noise $n(k) \sim \mathcal{N}(0, \sigma_n^2)$ or, in more general matrix-vector notation, $\mathbf{n}_{\mathrm{t},k} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_\mathrm{t})$ is assumed. These assumptions about $\mathbf{A}_\mathrm{t}$, $\mathbf{Q}_\mathrm{t}$ and $\mathbf{R}_\mathrm{t}$ are common and can also be found in [2], [3], [7], [9]. The observation matrix, which depends on the input signal, is $\mathbf{C}_{\mathrm{t},k} = \overleftarrow{\mathbf{x}}_{[L],k}^\mathrm{T}$. Here, we define $\overrightarrow{\mathbf{x}}_{[L],k} = [x(k-L+1), \ldots, x(k)]^\mathrm{T}$ as the length-L "forward" vector containing the L most recent time samples of $x(k)$ up to time index k. Similarly, $\overleftarrow{\mathbf{x}}_{[L],k} = [x(k), \ldots, x(k-L+1)]^\mathrm{T}$ shall be the "backward" vector.

The Kalman equations constitute the optimal linear minimum-mean-squared-error estimator of the state [2]. We denote the *a priori* state estimate as $\hat{\mathbf{w}}_{\mathrm{t},k}^{(-)}$, the *a posteriori* estimate as $\hat{\mathbf{w}}_{\mathrm{t},k}^{(+)}$, and define the state error covariance matrix, with corresponding *a priori* and *a posteriori* definitions, as $\mathbf{P}_\mathrm{t} = \mathcal{E}\left\{(\mathbf{w}_\mathrm{t} - \hat{\mathbf{w}}_\mathrm{t})(\mathbf{w}_\mathrm{t} - \hat{\mathbf{w}}_\mathrm{t})^\mathrm{H}\right\}$, where $\mathcal{E}\{\cdot\}$ denotes the expectation operator and $(\cdot)^\mathrm{H}$ represents complex conjugate transpose. The prediction equations are given by

$$\hat{\mathbf{w}}_{\mathrm{t},k}^{(-)} = \mathbf{A}_\mathrm{t}\hat{\mathbf{w}}_{\mathrm{t},k-1}^{(+)}, \tag{5}$$

and

$$\mathbf{P}_{\mathrm{t},k}^{(-)} = \mathbf{A}_\mathrm{t}\mathbf{P}_{\mathrm{t},k-1}^{(+)}\mathbf{A}_\mathrm{t}^\mathrm{T} + \mathbf{Q}_\mathrm{t}, \tag{6}$$

while the update steps read

$$\mathbf{K}_{\mathrm{t},k} = \mathbf{P}_{\mathrm{t},k}^{(-)}\mathbf{C}_{\mathrm{t},k}^\mathrm{T}\left(\mathbf{C}_{\mathrm{t},k}\mathbf{P}_{\mathrm{t},k}^{(-)}\mathbf{C}_{\mathrm{t},k}^\mathrm{T} + \mathbf{R}_\mathrm{t}\right)^{-1}, \tag{7}$$

$$\hat{\mathbf{w}}_{\mathrm{t},k}^{(+)} = \hat{\mathbf{w}}_{\mathrm{t},k}^{(-)} + \mathbf{K}_{\mathrm{t},k}\mathbf{e}_{\mathrm{t},k}, \tag{8}$$

with

$$\mathbf{e}_{\mathrm{t},k} = \left(\mathbf{y}_{\mathrm{t},k} - \mathbf{C}_{\mathrm{t},k}\hat{\mathbf{w}}_{\mathrm{t},k}^{(-)}\right), \tag{9}$$

and

$$\mathbf{P}_{\mathrm{t},k}^{(+)} = \left(\mathbf{I}_L - \mathbf{K}_{\mathrm{t},k}\mathbf{C}_{\mathrm{t},k}\right)\mathbf{P}_{\mathrm{t},k}^{(-)} \tag{10}$$

We choose the initial value $\mathbf{P}_{\mathrm{t},-1}^{(+)} = \sigma_p^2\mathbf{I}_L$.

B. General Time-Domain Kalman Filter (GTKF)

In [7], the above model is extended to take multiple past observations into account at each time step, which can be shown to be related to the affine projection algorithm. This variant is referred to as the so-called "general" Kalman filter. Its quantities are indicated by underlines in the following. The modified observation vector, including the past J observations of the output signal, is given by the length-J vector

$$\underline{\mathbf{y}}_{\mathrm{t},k} = \left[\mathbf{y}_{\mathrm{t},k}^\mathrm{T}, \ldots, \mathbf{y}_{\mathrm{t},k-J+1}^\mathrm{T}\right]^\mathrm{T} = [y(k), \ldots, y(k-J+1)]^\mathrm{T} \tag{11}$$

and hence the observation noise $\underline{\mathbf{n}}_{\mathrm{t},k} \sim \mathcal{N}(\mathbf{0}, \underline{\mathbf{R}}_\mathrm{t})$ in (4) becomes J-dimensional. Similarly, a modified observation matrix considering J past observations can be defined by

$$\underline{\mathbf{C}}_{\mathrm{t},k} = \left[\mathbf{C}_{\mathrm{t},k}^\mathrm{T}, \ldots, \mathbf{C}_{\mathrm{t},k-J+1}^\mathrm{T}\right]^\mathrm{T} \in \mathbb{R}^{J\times L}. \tag{12}$$

This model is based on the assumption that the system impulse response does not change a lot within the past J samples. The Kalman equations (5) to (10) can be applied similarly using the underlined quantities.

C. Frequency-Domain Kalman Filter (FKF)

A frequency-domain Kalman filter for system identification has first been proposed for echo cancellation in [1]. Instead of using samplewise processing, the convolution operations in (4) and (9) are conducted for a block of B samples using the overlap-save method. Each block is indexed by block index λ.

The state vector, representing the acoustic system's impulse response, is defined as the discrete Fourier transform (DFT) of the zero-padded impulse response, i.e.,

$$\mathbf{w}_{\mathrm{f},\lambda} = \mathbf{F}\mathbf{H}\mathbf{w}_{\mathrm{t},\lambda B} = \mathbf{F}\begin{bmatrix}\mathbf{w}_{\mathrm{t},\lambda B}\\ \mathbf{0}_{(M-L)\times 1}\end{bmatrix} \in \mathbb{C}^M, \tag{13}$$

where $\mathbf{H} = \begin{bmatrix}\mathbf{I}_L & \mathbf{0}_{L\times(M-L)}\end{bmatrix}^\mathrm{H}$ is the zero-padding matrix, and $\mathbf{F}$ is the unitary $M\times M$ DFT matrix. In contrast to [1], we prefer to use unitary DFT matrices in our formulation to facilitate showing the equivalence in a special case later. Similarly to (3), the state equation can be formulated as

$$\mathbf{w}_{\mathrm{f},\lambda} = \mathbf{A}_\mathrm{f}\mathbf{w}_{\mathrm{f},\lambda-1} + \mathbf{q}_{\mathrm{f},\lambda}. \tag{14}$$

The process noise is assumed to be described by a complex multivariate random process $\mathbf{q}_{\mathrm{f},\lambda} \sim \mathcal{N}_C(\mathbf{0}, \mathbf{Q}_\mathrm{f})$. Here, $\mathbf{A}_\mathrm{f}$ denotes the frequency-domain state transition matrix, which is usually assumed to be $\mathbf{A}_\mathrm{f} = \gamma\mathbf{I}_M$. The frequency-domain observation equation can be written as

$$\mathbf{y}_{\mathrm{f},\lambda} = \mathbf{C}_{\mathrm{f},\lambda}\mathbf{w}_{\mathrm{f},\lambda} + \mathbf{n}_{\mathrm{f},\lambda}. \tag{15}$$

This equation requires the following definitions, namely, the length-M frequency-domain observation vector

$$\mathbf{y}_{\mathrm{f},\lambda} = \mathbf{F}\mathbf{G}\overrightarrow{\mathbf{y}}_{[B],\lambda B}, \tag{16}$$

where $\overrightarrow{\mathbf{y}}_{[B],\lambda B}$ is defined analogously to $\overrightarrow{\mathbf{x}}_{[L],k}$ above. The frequency-domain observation matrix is

$$\mathbf{C}_{\mathrm{f},\lambda} = \sqrt{M}\mathbf{F}\mathbf{G}\mathbf{G}^\mathrm{H}\mathbf{F}^\mathrm{H}\,\mathrm{diag}\left\{\mathbf{F}\overrightarrow{\mathbf{x}}_{[M],\lambda B}\right\} \in \mathbb{C}^{M\times M} \tag{17}$$

with the constraining matrix $\mathbf{G} = \begin{bmatrix}\mathbf{0}_{B\times(M-B)} & \mathbf{I}_B\end{bmatrix}^\mathrm{H}$.

The length-M frequency-domain observation noise vector $\mathbf{n}_{\mathrm{f},\lambda} = \mathbf{F}\mathbf{G}\overrightarrow{\mathbf{n}}_{[B],\lambda B}$ is assumed to be described by a complex random process $\mathbf{n}_{\mathrm{f},\lambda} \sim \mathcal{N}_C(\mathbf{0}, \mathbf{R}_\mathrm{f})$.

Replacing the time-domain quantities, with subscripts "t", by the corresponding frequency-domain quantities, with subscripts "f", and replacing the time index k with the block index λ, the equations (5) to (10) describe the frequency-domain Kalman filter.

D. Diagonalized Frequency-Domain Kalman Filter (DFKF)

A common simplification, introduced in [1] to reduce the computational complexity, is to approximate all the matrices involved as $M \times M$ diagonal matrices. According to [1] and considering the scaling of the DFT matrix used here, the approximated observation matrix is

$$\mathbf{\Gamma}_{\mathrm{f},\lambda} = \frac{B}{\sqrt{M}} \operatorname{diag}\left\{\mathbf{F}\vec{\mathbf{x}}_{[M],\lambda B}\right\}, \tag{18}$$

the Kalman gain is approximated as

$$\mathbf{K}_{\mathrm{f},\lambda} \approx \mathbf{P}_{\mathrm{f},\lambda}^{(-)}\mathbf{\Gamma}_{\mathrm{f},\lambda}^{\mathrm{H}}\sqrt{M}\left(M\mathbf{\Gamma}_{\mathrm{f},\lambda}\mathbf{P}_{\mathrm{f},\lambda}^{(-)}\mathbf{\Gamma}_{\mathrm{f},\lambda}^{\mathrm{H}} + \frac{M}{B}\mathbf{R}_{\mathrm{f}}\right)^{-1}, \tag{19}$$

and the approximated state error covariance update is

$$\mathbf{P}_{\mathrm{f},\lambda}^{(+)} = \left(\mathbf{I} - \mathbf{K}_{\mathrm{f},\lambda}\mathbf{\Gamma}_{\mathrm{f},\lambda}\right)\mathbf{P}_{\mathrm{f},\lambda}^{(-)}. \tag{20}$$

If the state transition matrix, the process noise covariance matrix and the observation noise covariance matrix are also diagonal, an efficient vectorized implementation can be used.

III. Special Case of Equivalence

While the time-domain and the frequency-domain Kalman filters in Secs. II-A and II-C, respectively, may appear disconnected, we will now prove that they are equivalent for $B = 1$ under certain assumptions. From $B = 1$, we obtain $\lambda B = k$, and we want to identify an impulse response of the length $L = M - B + 1 = M$. Therefore, (13) simplifies to $\mathbf{w}_{\mathrm{f},k} = \mathbf{F}\mathbf{w}_{\mathrm{t},k}$. If we choose $\mathbf{A}_{\mathrm{f}} = \mathbf{F}\mathbf{A}_{\mathrm{t}}\mathbf{F}^{\mathrm{H}}$, $\mathbf{Q}_{\mathrm{f}} = \mathbf{F}\mathbf{Q}_{\mathrm{t}}\mathbf{F}^{\mathrm{H}}$, the state equations (3) and (14) are equivalent.

Similarly, we show that the observation equations (4) and (15) are equivalent. Multiplying (15) from the left by $\mathbf{F}^{-1} = \mathbf{F}^{\mathrm{H}}$, we obtain

$$\mathbf{F}^{\mathrm{H}}\mathbf{y}_{\mathrm{f},k} = \begin{bmatrix}\mathbf{0}_{(M-1)\times 1}\\ y(k)\end{bmatrix} = \mathbf{F}^{\mathrm{H}}\mathbf{C}_{\mathrm{f},k}\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{w}_{\mathrm{f},k} + \mathbf{F}^{\mathrm{H}}\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{n}_{\mathrm{f},k} \tag{21}$$

$$= \mathbf{F}^{\mathrm{H}}\mathbf{C}_{\mathrm{f},k}\mathbf{F}\mathbf{w}_{\mathrm{t},k} + \begin{bmatrix}\mathbf{0}_{(M-1)\times 1}\\ n(k)\end{bmatrix}. \tag{22}$$

To prove the equivalence, it is hence sufficient to show that

$$\mathbf{F}^{\mathrm{H}}\mathbf{C}_{\mathrm{f},k}\mathbf{F} = \begin{bmatrix}\mathbf{0}_{(M-1)\times M}\\ \mathbf{C}_{\mathrm{t},k}\end{bmatrix}. \tag{23}$$

Inserting (17) into the left side of (23) yields

$$\mathbf{F}^{\mathrm{H}}\mathbf{C}_{\mathrm{f},k}\mathbf{F} = \mathbf{F}^{\mathrm{H}}\sqrt{M}\mathbf{F}\mathbf{G}\mathbf{G}^{\mathrm{H}}\mathbf{F}^{\mathrm{H}}\operatorname{diag}\left\{\mathbf{F}\vec{\mathbf{x}}_{[M],k}\right\}\mathbf{F} \tag{24}$$

$$= \sqrt{M}\mathbf{G}\mathbf{G}^{\mathrm{H}}\underbrace{\mathbf{F}^{\mathrm{H}}\operatorname{diag}\left\{\mathbf{F}\vec{\mathbf{x}}_{[M],k}\right\}\mathbf{F}}_{=\frac{1}{\sqrt{M}}C\left\{\vec{\mathbf{x}}_{[M],k}\right\}}, \tag{25}$$

where we used the property that any circulant matrix, defined by a vector $\mathbf{c} = [c(0), \ldots, c(M-1)]^{\mathrm{T}}$, i.e.,

$$C\{\mathbf{c}\} = \begin{bmatrix} c(0) & c(M-1) & \ldots & c(1)\\ c(1) & c(0) & \ldots & c(2)\\ \vdots & \ldots & \ddots & \vdots\\ c(M-1) & c(M-2) & \ldots & c(0)\end{bmatrix}, \tag{26}$$

can be factorized [15] as follows:

$$\sqrt{M}\mathbf{F}^{\mathrm{H}}\operatorname{diag}\{\mathbf{F}\mathbf{c}\}\mathbf{F} = C\{\mathbf{c}\}. \tag{27}$$

Then, continuing from (25):

$$\mathbf{F}^{\mathrm{H}}\mathbf{C}_{\mathrm{f},k}\mathbf{F} = \begin{bmatrix}\mathbf{0}_{M-1} & 0\\ \mathbf{0}_{1\times(M-1)} & 1\end{bmatrix} C\left\{\vec{\mathbf{x}}_{[M],k}\right\} = \begin{bmatrix}\mathbf{0}_{(M-1)\times M}\\ \overleftarrow{\mathbf{x}}_{[M],k}^{\mathrm{T}}\end{bmatrix} \tag{28}$$

as multiplying by $\mathbf{G}\mathbf{G}^{\mathrm{H}}$ from the left sets the first $M-1$ rows to zero and extracts the last row of $C\left\{\vec{\mathbf{x}}_{[M],k}\right\}$. This remaining row is the time-reversed version of $\vec{\mathbf{x}}_{[M],k}$ and that is $\overleftarrow{\mathbf{x}}_{[M],k}$ for $M = L$ and $B = 1$.

Since we have shown that both the state equations and observation equations are equivalent, the corresponding Kalman filters must be equivalent, too. To obtain identical initial conditions for both the TKF and the FKF, we set $\mathbf{P}_{\mathrm{f}} = \mathbf{F}\mathbf{P}_{\mathrm{t}}\mathbf{F}^{\mathrm{H}}$ and must choose $\mathbf{R}_{\mathrm{f}}$ such that the variance of the last sample of the time-domain vector $\mathbf{F}^{\mathrm{H}}\mathbf{n}_{\mathrm{f},k}$ in (22) equals σ_n^2. Hence, $\mathbf{R}_{\mathrm{f}} = \sigma_n^2\mathbf{I}_M$.

IV. Proposed Kalman Filter Variants

Having shown the above equivalence, we can transfer concepts from one domain to the other and vice versa. Three analogies will be drawn in the following.

A. General Frequency-Domain Kalman Filter (GFKF)

Inspired from the idea of taking multiple observations into account in the GTKF, we apply this concept to the frequency-domain KFs. The modified frequency-domain observation equation considering J observations is

$$\underline{\mathbf{y}}_{\mathrm{f},\lambda} = \underline{\mathbf{C}}_{\mathrm{f},\lambda}\mathbf{w}_{\mathrm{f},\lambda} + \underline{\mathbf{n}}_{\mathrm{f},\lambda}, \tag{29}$$

where

$$\underline{\mathbf{y}}_{\mathrm{f},\lambda} = \left[\mathbf{y}_{\mathrm{f},\lambda}^{\mathrm{H}}, \ldots, \mathbf{y}_{\mathrm{f},\lambda-J+1}^{\mathrm{H}}\right]^{\mathrm{H}} \in \mathbb{C}^{JM} \tag{30}$$

and

$$\underline{\mathbf{C}}_{\mathrm{f},\lambda} = \left[\mathbf{C}_{\mathrm{f},\lambda}^{\mathrm{H}}, \ldots, \mathbf{C}_{\mathrm{f},\lambda-J+1}^{\mathrm{H}}\right]^{\mathrm{H}} \in \mathbb{C}^{JM\times M}. \tag{31}$$

In this case, the observation noise becomes JM-dimensional, i.e., we assume $\underline{\mathbf{n}}_{\mathrm{f},\lambda} \sim \mathcal{N}_C\left(\mathbf{0}, \underline{\mathbf{R}}_{\mathrm{f}}\right)$ with $\underline{\mathbf{R}}_{\mathrm{f}} = \mathbf{I}_J \otimes \mathbf{R}_{\mathrm{f}}$. Here, $\otimes$ denotes the Kronecker product. Again, the Kalman equations (5) to (10) can be applied with the appropriate replacements.

The proposed generalization is equivalent to the GTKF only if $B = 1$. For $J > 1$, the equivalence can be proven similarly to above. Yet, this generalization can also be employed for any block length B, which will no longer be equivalent to the GTKF. A GFKF-equivalent time-domain KF for $B > 1$ is presented in Sec. IV-C.

B. General Diagonalized Frequency-Domain KF (GDFKF)

The computational cost of the general frequency-domain Kalman filter will be dominated by the cost of computing the $JM \times JM$ inverse in the Kalman gain (analogously to (7)). It approximately corresponds to $O\left((JM)^3\right)$. Therefore, we modify the more efficient diagonalized frequency-domain Kalman filter to consider J past observations by defining

$$\underline{\mathbf{\Gamma}}_{\mathrm{f},\lambda} = \left[\mathbf{\Gamma}_{\mathrm{f},\lambda}^{\mathrm{H}}, \ldots, \mathbf{\Gamma}_{\mathrm{f},\lambda-J+1}^{\mathrm{H}}\right]^{\mathrm{H}} \in \mathbb{C}^{JM\times M}. \tag{32}$$

With this definition, however, not all quantities in the diagonalized frequency-domain Kalman filter remain diagonal matrices. Instead of computing the inverse of the $JM \times JM$ matrix

$$\left(M\underline{\mathbf{\Gamma}}_{\mathrm{f},\lambda}\mathbf{P}_{\mathrm{f},\lambda}^{(-)}\underline{\mathbf{\Gamma}}_{\mathrm{f},\lambda}^{\mathrm{H}} + \frac{M}{B}\underline{\mathbf{R}}_{\mathrm{f}}\right), \tag{33}$$

we suggest to reduce this $JM \times JM$ block matrix of $M \times M$ diagonal matrices above to a block diagonal matrix of M dense matrices of size $J \times J$ by multiplication with appropriate permutation matrices. Computing the inverses of the small matrices can be implemented in a parallel fashion. Note that this approach is structurally similar to the multi-channel diagonalized frequency-domain Kalman filter in [11]. A detailed complexity discussion for all variants is omitted here for brevity.

C. General Block Time-Domain Kalman Filter (GBTKF)

To complete the picture of single-channel Kalman filters in the time domain and in the frequency domain, we can also formulate a block-based KF in the time domain and immediately extend it to the general version. Therefore, we define the state equation, now with block index λ, as

$$\mathbf{w}_{\mathrm{t},\lambda} = \mathbf{A}_{\mathrm{t}}\mathbf{w}_{\mathrm{t},\lambda-1} + \mathbf{q}_{\mathrm{t},\lambda}, \tag{34}$$

in which the state is updated only every B-th sample, similarly to the state equation (14). The observation equation becomes

$$\underline{\underline{\mathbf{y}}}_{\mathrm{t},\lambda} = \underline{\underline{\mathbf{C}}}_{\mathrm{t},\lambda}\mathbf{w}_{\mathrm{t},\lambda} + \underline{\underline{\mathbf{n}}}_{\mathrm{t},\lambda}, \tag{35}$$

where the JB-dimensional observation vector is

$$\underline{\underline{\mathbf{y}}}_{\mathrm{t},\lambda} = \left[\overleftarrow{\mathbf{y}}_{[B],\lambda B}^{\mathrm{T}}, \ldots, \overleftarrow{\mathbf{y}}_{[B],(\lambda-J+1)B}^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{36}$$

The $JB \times L$ observation matrix is given by

$$\underline{\underline{\mathbf{C}}}_{\mathrm{t},\lambda} = \left[\overleftarrow{\mathbf{x}}_{[L],\lambda B}, \ldots, \overleftarrow{\mathbf{x}}_{[L],\lambda B-B+1}, \ldots, \overleftarrow{\mathbf{x}}_{[L],(\lambda-J+1)B-B+1}\right]^{\mathrm{T}} \tag{37}$$

and we have $\underline{\underline{\mathbf{n}}}_{\mathrm{t},\lambda} = \left[\overleftarrow{\mathbf{n}}_{[B],\lambda B}^{\mathrm{T}}, \ldots, \overleftarrow{\mathbf{n}}_{[B],(\lambda-J+1)B}^{\mathrm{T}}\right]^{\mathrm{T}}$. In this case, the observation noise becomes JB-dimensional and we assume that $\underline{\underline{\mathbf{n}}}_{\mathrm{t},\lambda} \sim \mathcal{N}\left(\mathbf{0}, \underline{\underline{\mathbf{R}}}_{\mathrm{t}}\right)$ with $\underline{\underline{\mathbf{R}}}_{\mathrm{t}} = \mathbf{I}_J \otimes \underline{\mathbf{R}}_{\mathrm{t}}$.

For $B = 1$, the general block time-domain Kalman filter reduces to the general time-domain Kalman filter (cf. Sec. II-A), which again reduces to the standard time-domain Kalman filter (cf. Sec. II-A) for $J = 1$. With appropriate choices of $\mathbf{R}_{\mathrm{f}}$, $\mathbf{Q}_{\mathrm{f}}$ and the initial state error covariance matrix $\mathbf{P}_{\mathrm{f}}$, the frequency-domain KF and the block time-domain KF can be shown to be equivalent. For simplicity $J = 1$ is considered, but the result also holds for $J > 1$. Using $\mathbf{n}_{\mathrm{f}} = \mathbf{F}\mathbf{G}\mathbf{n}_{\mathrm{t}}$ and $\mathbf{w}_{\mathrm{f}} = \mathbf{F}\mathbf{H}\mathbf{w}_{\mathrm{t}}$, it can be shown that

$$\mathbf{R}_{\mathrm{f}} = \mathbf{F}\mathbf{G}\mathbf{R}_{\mathrm{t}}\mathbf{G}^{\mathrm{H}}\mathbf{F}^{\mathrm{H}} \approx \sigma_n^2\frac{B}{M}\mathbf{I}_M, \tag{38}$$

$$\mathbf{Q}_{\mathrm{f}} = \mathbf{F}\mathbf{H}\mathbf{Q}_{\mathrm{t}}\mathbf{H}^{\mathrm{H}}\mathbf{F}^{\mathrm{H}} \approx \sigma_q^2\frac{L}{M}\mathbf{I}_M, \tag{39}$$

and

$$\mathbf{P}_{\mathrm{f}} = \mathbf{F}\mathbf{H}\mathbf{P}_{\mathrm{t}}\mathbf{H}^{\mathrm{H}}\mathbf{F}^{\mathrm{H}} \approx \sigma_p^2\frac{L}{M}\mathbf{I}_M. \tag{40}$$

The above diagonal approximations become more precise for $B \to M$ and $L \to M$, respectively. However, they do not yield the original time-domain covariance matrices, we started from, when transforming back to the time-domain due to the impact of multiplying with $\mathbf{G}$ or $\mathbf{H}$. Instead one must choose

$$\mathbf{R}_{\mathrm{f}} = \sigma_n^2\mathbf{I}_M, \tag{41}$$

$$\mathbf{Q}_{\mathrm{f}} = \sigma_q^2\mathbf{I}_M, \tag{42}$$

$$\mathbf{P}_{\mathrm{f}} = \sigma_p^2\mathbf{I}_M. \tag{43}$$

Note that all the variants presented can straightforwardly be extended to the multiple-input-single-output cases.

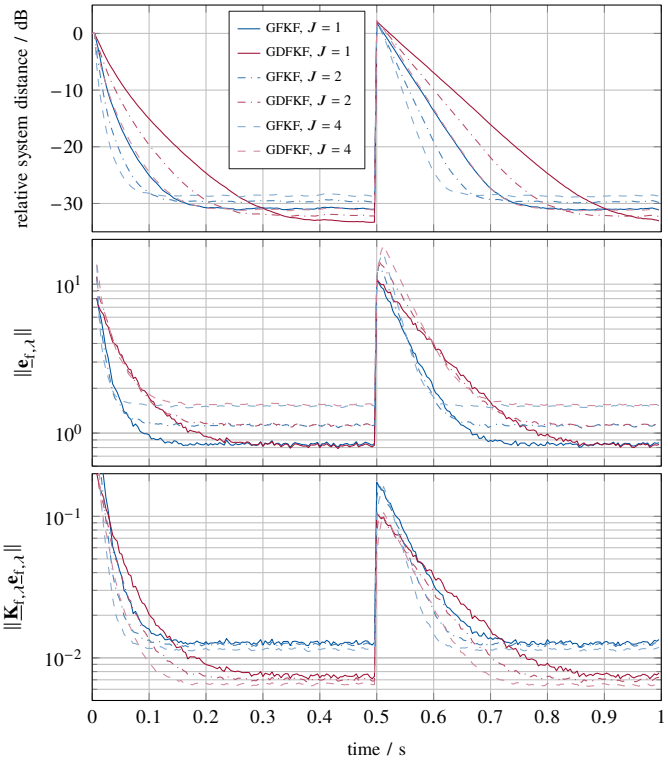

Fig. 2. GFKF vs. GDFKF for $L = 128$, $B = 64$ at 20 dB SNR.

V. Experimental Comparison

A comprehensive evaluation in various applications for time-invariant and time-variant conditions, and using different process noise and observation noise online-estimation methods is out of scope for this paper. Thus, we conducted three simulated experiments to evaluate and compare the different Kalman filter variants using the setup from Fig. 1. The relative system distance metric [2] is used for evaluation. The common parameters for each experiment were: sampling rate 16 kHz, white Gaussian noise signals $x(k)$, $\sigma_p^2 = 10^{-4}$, σ_n^2 matching the true signal-to-noise ratio (SNR) conditions, which was generated adding white Gaussian noise. The impulse response coefficients were drawn from a standard normal distribution, and the impulse response vectors were normalized to unit energy. Each experiment was repeated 20 times and the results were averaged to obtain smoother, more readable curves. For simplicity, we chose: $\mathbf{A}_{\mathrm{f}} = \mathbf{I}_M$ and $\mathbf{A}_{\mathrm{t}} = \mathbf{I}_L$ in all cases.

First, we analyze how considering multiple past observations in the GFKF and GDFKF can be advantageous. Fig. 2 depicts

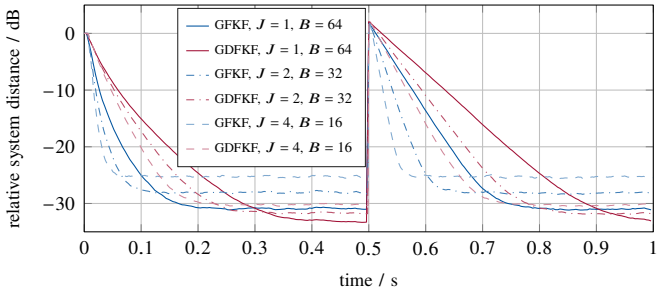

Fig. 3. GFKF vs. GDFKF for $L = 128$ with $JB = 64$ at 20 dB SNR.

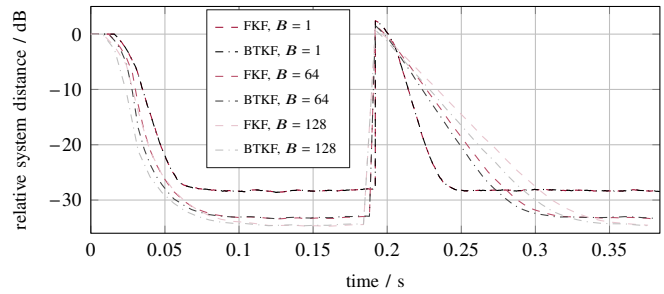

Fig. 4. FKF with covariance matrices (41) to (43) vs. BTKF for different block lengths. The curves of FKF and BTKF coincide for $B = 1$, as expected.

the convergence and steady-state behavior for an example of exchanging the impulse response at 0.5 s. Considering multiple past observations increases the norm of the error vector $\underline{\mathbf{e}}_{\mathrm{f},\lambda}$ and hence increases the effect of the state vector update $\underline{\mathbf{K}}_{\mathrm{f},\lambda}\underline{\mathbf{e}}_{\mathrm{f},\lambda}$, which results in faster convergence. However, this comes at the price of decreased steady-state performance as the larger random error implies to continue updating the state vector into random directions. This result matches the results for the GTKF in [7]. Surprisingly, the steady-state system distance of the GDFKF is lower than the one of the GFKF. This might be a consequence of the block diagonal Kalman gain matrix which does not model dependencies between frequency bins and results in smaller changes overall. In this example, we observe that the GDFKF with $J = 4$ offers comparable performance to the original FKF at a reduced complexity. This demonstrates that the GDFKF can be an attractive choice.

Second, the impact of considering more past observations is investigated when the "time span" of past samples, i.e., JB remains constant, as shown in Fig. 3. Both for the GFKF and the GDFKF, we observe that larger J means faster convergence but worse steady-state performance, similar to Fig. 2.

Fig. 4 compares the FKF and the block time-domain KF (BTKF), i.e. the GBTKF for $J = 1$, for different block lengths. As the FKF becomes equivalent to the BTKF for the exact equations in (38) to (40), it is not depicted because the results are identical to the ones of the BTKF. These approximations become exact for $B = 1$ and yield identical outputs, as shown. Using the diagonal matrices (41) to (43) in the frequency-domain decreases the performance for larger block lengths compared to the BTKF. These frequency-domain diagonal matrices represent nonzero variances in *all* M time-domain samples — mismatching the assumption — as only B or L of the M time-domain samples are nonzero. Note that larger block lengths for both the FKF and the BTKF mean slower convergence but a lower steady-state system distance.

VI. Conclusion

We have provided a unified perspective on time-domain and frequency-domain Kalman filters (KFs). It was shown that the TKF and the FKF are equivalent for the block length $B = 1$. This equivalence leads to the block time-domain KF variant (BTKF), which can also be equivalent to the FKF for $B > 1$. We have transferred the concept of considering multiple past observations from the GTKF to frequency-domain KFs. This yields the GFKF and the GDFKF variants. Further, we extended the BTKF to the GBTKF, which considers multiple past blocks of samples. The proposed variants extend the variety of KFs and allow us to precisely tune the compromises between convergence speed, steady-state performance and computational complexity required in specific applications.

References

[1] G. Enzner and P. Vary, "Frequency-domain adaptive Kalman filter for acoustic echo control in hands-free telephones," *Signal Process.*, vol. 86, no. 6, pp. 1140–1156, 2006.

[2] G. Enzner, "Bayesian inference model for applications of time-varying acoustic system identification," in *18th Eur. Signal Process. Conf. (EUSIPCO)*, 2010, pp. 2126–2130.

[3] P. A. Lopes, J. A. Gerald, and M. S. Piedade, "The random walk model Kalman filter in multichannel active noise control," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2244–2248, 2015.

[4] S. Liebich, J. Fabry, P. Jax, and P. Vary, "Time-domain Kalman filter for active noise cancellation headphones," in *25th Eur. Signal Process. Conf. (EUSIPCO)*, 2017, pp. 593–597.

[5] T. Kabzinski and P. Jax, "An adaptive crosstalk cancellation system using microphones at the ears," in *Audio Eng. Soc. Conv. 147*. Audio Eng. Soc., 2019, pp. 1–10.

[6] ——, "Towards faster continuous multi-channel HRTF measurements based on learning system models," in *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, 2022, pp. 436–440.

[7] C. Paleologu, J. Benesty, and S. Ciochină, "Study of the general Kalman filter for echo cancellation," *IEEE Trans. on Audio, Speech, and Lang. Process.*, vol. 21, no. 8, pp. 1539–1549, 2013.

[8] J. Guo, F. Yang, G. Enzner, and J. Yang, "Tracking analysis and improvement of broadband Kalman filter using the two-echo-path model as a rapid tunnel," in *16th Int. Workshop on Acoustic Signal Enhancement (IWAENC)*, 2018, pp. 341–345.

[9] J. Fabry, S. Kühl, and P. Jax, "On the steady state performance of the Kalman filter applied to acoustical systems," *IEEE Signal Process. Lett.*, vol. 27, pp. 1854–1858, 2020.

[10] J. Fabry, S. Liebich, P. Vary, and P. Jax, "Active noise control with reduced-complexity Kalman filter," in *16th Int. Workshop on Acoustic Signal Enhancement (IWAENC)*, 2018, pp. 166–170.

[11] S. Malik and G. Enzner, "Recursive Bayesian control of multichannel acoustic echo cancellation," *IEEE Signal Process. Lett.*, vol. 18, no. 11, pp. 619–622, 2011.

[12] ——, "Variational Bayesian inference for nonlinear acoustic echo cancellation using adaptive cascade modeling," in *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, 2012, pp. 37–40.

[13] F. Kuech, E. Mabande, and G. Enzner, "State-space architecture of the partitioned-block-based acoustic echo controller," in *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, 2014, pp. 1295–1299.

[14] T. Haubner, A. Brendel, and W. Kellermann, "Online acoustic system identification exploiting Kalman filtering and an adaptive impulse response subspace model," *J. Signal Process. Systems*, vol. 94, no. 2, pp. 147–160, 2022.

[15] G. Golub and C. Van Loan, *Matrix computations*, 3rd ed. Baltimore, MD, USA: The John Hopkins University Press, 1996.