Joint Carrier Frequency Offset and Channel Estimation for Uplink MIMO-OFDMA Systems Using Parallel Schmidt Rao-Blackwellized Particle Filters

Kyeong Jin Kim, Man-On Pun, Ronald Iltis

TR2010-095 November 2010

Abstract

Joint carrier frequency offset (CFO) and channel estimation for uplink MIMO-OFDMA systems over time-varying channels is investigated. To cope with the prohibitive computational complexity involved in estimating multiple CFOs and channels, pilot-assisted and semi-blind schemes comprised of parallel Schmidt Extended Kalman filters (SEKFs) and Schmidt-Kalman Approximate Particle Filters (SK-APF) are proposed. In the SK-APF, a Rao-Blackwellized particle filter (RBPF) is developed to first estimate the nonlinear state variable, i.e. the desired user's CFO, through the sampling-importance-resampling (SIRS) technique. The individual user channel responses are then updated via a bank of Kalman filters conditioned on the CFO sample trajectories. Simulation results indicate that the proposed schemes can achieve highly accurate CFO/channel estimates, and that the particle filtering approach in the SK-APF outperforms the more conventional Schmidt Extended Kalman Filter.

IEEE Transactions on Communications

Joint Carrier Frequency Offset and Channel Estimation for Uplink MIMO-OFDMA Systems Using Parallel Schmidt Rao-Blackwellized Particle Filters

Kyeong Jin Kim, *Member, IEEE*, Man-On Pun, *Member, IEEE*, and Ronald A. Iltis, *Senior Member, IEEE*

***Abstract*—Joint carrier frequency offset (CFO) and channel estimation for uplink MIMO-OFDMA systems over time-varying channels is investigated. To cope with the prohibitive computational complexity involved in estimating multiple CFOs and channels, pilot-assisted and semi-blind schemes comprised of parallel Schmidt Extended Kalman filters (SEKFs) and Schmidt-Kalman Approximate Particle Filters (SK-APF) are proposed. In the SK-APF, a Rao-Blackwellized particle filter (RBPF) is developed to first estimate the nonlinear state variable, i.e. the desired user's CFO, through the sampling-importance-resampling (SIRS) technique. The individual user channel responses are then updated via a bank of Kalman filters conditioned on the CFO sample trajectories. Simulation results indicate that the proposed schemes can achieve highly accurate CFO/channel estimates, and that the particle filtering approach in the SK-APF outperforms the more conventional Schmidt Extended Kalman Filter.**

***Index Terms*—Orthogonal frequency division multiple access (OFDMA), multiple-input multiple-output (MIMO), channel estimation, frequency synchronization, Schmidt extended Kalman filter (SEKF), Rao-Blackwellized particle filter (RBPF).**

I. Introduction

ORTHOGONAL Frequency Division Multiple Access (OFDMA) is a leading technology for broadband wireless networks[1]. In addition to its robustness to multipath fading and high spectral efficiency, OFDMA offers flexibility in allocating subcarriers to different users based on quality of service (QoS) requirements and channel conditions [2]. Advances in multiple-input multiple-output (MIMO) techniques have led to considerable interest in MIMO-OFDMA in which substreams of a broadband source are transmitted over multiple antennas and subcarriers.

In this paper we consider an uplink MIMO-OFDMA system over *time-varying* Rayleigh fading channels. In the uplink, the MIMO-OFDMA system requires that active users must be synchronized in frequency in order to maintain orthogonality.[1] Furthermore, accurate channel estimation is required when coherent data detection is employed. Several schemes have been proposed to perform joint synchronization and channel estimation for uplink OFDMA systems, e.g. [4]–[8]. Despite their good performance, these existing schemes assume that channels experience block-fading and CFOs are *time-invariant*. However, such approaches may not be suitable for outdoor channels with high user mobility where large Doppler shifts are present. In [9], an extended Kalman filter (EKF)-based scheme has been proposed to track the time variations of CFOs and channels for single-user MIMO-OFDM systems impaired by multiple CFOs. By exploiting a set of pilots, [9] can provide accurate estimation of CFOs and channels. In [10], parallel EKFs are proposed to perform CFO estimation for uplink OFDMA systems as well as multiple access interference (MAI) cancellation. However, in addition to its high complexity required in computing the Kalman gain, the EKF approach in [9] suffers from potential divergence caused by the highly nonlinear dependence of the received signal on the CFOs [11]. To circumvent this problem, [11] has proposed a particle filtering (PF)-based scheme [12] to track a single time-varying CFO for OFDM systems, assuming perfect channel information is available. Recently, joint frequency offset and channel estimation has been proposed in [13] using Gauss-Hermite integration. Unfortunately, this approach is only valid for single-user systems.

In this work, both pilot-aided and semi-blind schemes are proposed to jointly estimate time-varying frequency offsets and channels for multiuser uplink MIMO-OFDMA systems. The challenge in the uplink MIMO-OFDMA case can be appreciated by comparing with the simpler SISO-OFDMA case. For instance, consider user separation techniques. For SISO-OFDMA systems it has been shown that iterative direct compensation of one user's CFO while treating other users' signals as interference can provide effective user separation [14]. However, for MIMO-OFDMA with each user distorted by multiple CFOs, direct compensation of only one CFO may not lead to signal improvement due to the existence of other CFOs of the same user. To circumvent this problem, two

Paper approved by N. Al-Dhahir, the Editor for Space-Time, OFDM and Equalization of the IEEE Communications Society. Manuscript received January 15, 2009; revised September 15, 2009 and January 15, 2010.

The work of M. O. Pun was done when he was a Croucher post-doctoral fellow at Princeton University, Princeton, NJ.

The work of R. A. Iltis was supported in part by a gift from Nokia.

K. J. Kim was with Nokia Inc., 6000 Connection Drive, Irving, TX 75039 USA (e-mail: kyeong.j.kim@hotmail.com).

M. O. Pun is with Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139 USA (e-mail: mpun@merl.com).

R. A. Iltis is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560 USA (e-mail: iltis@ece.ucsb.edu).

Digital Object Identifier 10.1109/TCOMM.2010.080310.090053

[1]Interested readers should refer to [3] for a comprehensive tutorial on OFDMA synchronization.

techniques are proposed here. First, to achieve user separation at low computational complexity, pilot-aided methods employ the Schmidt-Kalman filters to break down the multiuser CFO and channel estimation problem into separable sub-problems by exploiting a few training blocks [15]. Each estimation sub-problem is then solved by a Rao-Blackwellized particle filter where the desired user's CFO is estimated through sampling-importance-resampling (SIRS) and the channel response is updated via a EKF conditioned on the CFO samples generated by SIRS. By using the SIRS technique, the divergence problem due to non-linearity introduced by the CFOs is reduced. Finally, since training blocks are usually only available at the beginning of each data frame, a semi-blind approach exploiting the tentatively detected data symbols is developed to improve estimation accuracy. Specifically a QR decomposition (QRD)-based detector is incorporated into the receivers to provide reliable tentative data decisions. Simulation results confirm the effectiveness of the proposed joint estimators for MIMO-OFDMA.

The rest of the paper is organized as follows. Section II presents the signal model for the MIMO-OFDMA system, followed by the suboptimal parallel Schmidt extended Kalman filter (SEKF) in Section III. Based on the parallel SEKF, a parallel Schmidt-Kalman approximate-Rao-Blackwellized particle filter (SK-APF) is proposed in Section IV. The QRD-M detector semi-blind approach is given in Section V. Simulation results are presented in Section VI and conclusions are given in Section VII.

Notation: Vectors and matrices are denoted by boldface letters. $\|\cdot\|$ represents the Euclidean norm of the enclosed vector and $|\cdot|$ denotes the amplitude of the enclosed complex-valued quantity. Finally, we use $\mathrm{E}\{\cdot\}$, $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ for expectation, complex conjugation, transposition and Hermitian transposition.

II. Signal and Channel Models for Uplink MIMO-OFDMA Systems

We consider an uplink OFDMA system with N subcarriers and K active users. The base station (BS) and each active user are equipped with M_r and M_t antennas, respectively. Each user is assigned N_k exclusive subcarriers, where $\sum_{k=1}^{K} N_k \leq N$. We denote the index set of carriers assigned to the k-th user as $\mathcal{I}_k \triangleq \{i_1, i_2, \ldots, i_{N_k}\}$ where $1 \leq i_l \leq N$ for $l = 1, 2, \cdots, N_k$. The proposed algorithms are applicable to any carrier assignment scheme (CAS). Note that even though we are using orthogonal subcarriers, self-interference and MAI are inevitable due to CFOs [16]. Denote by $\tilde{\boldsymbol{d}}_k^p(n) \in \mathbb{C}^N$ the data symbols transmitted by the k-th user from the p-th transmit antenna over the n-th OFDMA block. For convenience, we assume that the data symbols are taken from the same complex-valued finite alphabet and independently identically distributed (i.i.d). The i-th entry of $\tilde{\boldsymbol{d}}_k^p(n)$, $\tilde{d}_{k,i}^p(n)$, is non-zero if and only if $i \in \mathcal{I}_k$. Next, $\tilde{\boldsymbol{d}}_k^p(n)$ is converted to the corresponding time-domain vector by an N-point inverse discrete Fourier transform (IDFT):

$$\boldsymbol{d}_k^p(n) = \boldsymbol{W}^H \tilde{\boldsymbol{d}}_k^p(n), \tag{1}$$

where $\boldsymbol{W}^H$ is the IDFT matrix. To prevent inter-symbol interference (ISI), a cyclic prefix (CP) of N_g symbols is appended in front of each IDFT output block. The resulting vector of length $N_d^g = N + N_g$ is digital-to-analog converted by a pulse-shaping filter with a finite support on $[0, T_d)$ where $T_d = N_d^g T_s$ with $1/(NT_s)$ being the subcarrier spacing. Finally, the analog signal from the pulse-shaping filter is transmitted from the p-th antenna over the channel.

The channel between the p-th transmit antenna of the k-th user and the q-th receive antenna of the BS during the n-th block, $\{h_{k,l}^{p,q}(n), 0 \leq l \leq L_k^{p,q} - 1\}$, is modeled as a tapped delay line with $L_k^{p,q}$ being the channel order. Since $L_k^{p,q}$ is generally unknown, in practice we replace $L_k^{p,q}$ with L_f for all users and antenna pairs. For the widely accepted ITU channel model, an upper bound for L_f can be specified. We assume that the length of the CP is sufficient to comprise the maximum path delay, i.e., $L_f \leq N_g$. Furthermore, we assume that $\{h_{k,l}^{p,q}(n)\}$ is constant over a OFDMA block n of length $N_d^g T_s$ s but varies between blocks.

A. Signal Model in the Presence of CFOs

Due to the Doppler effect and oscillator mismatch between transmitter and receiver pairs, the received signal is usually distorted by carrier frequency offsets [17]. Since the antenna separation on each user's mobile terminal is generally much smaller than that on the BS, it can be reasonably assumed that the CFO between transmit antennas of the same user and a specific receive antenna on the BS is identical. That is, $\{\delta f_k^q = \delta f_k^{p,q}, \forall p\}$. Let $\varepsilon_k^q \triangleq \delta f_k^q N T_s$ be the normalized carrier frequency offset with respect to (w.r.t.) the carrier spacing $1/(NT_s)$ between transmit antennas of the k-th user and the q-th receive antenna of the BS. In the *presence* of CFOs, the received vector signal after removing the guard interval becomes [6], [14], [17], [18]

$$\begin{aligned}\boldsymbol{r}^q(n) &= \sum_{k=1}^{K} \boldsymbol{\Delta}(\varepsilon_k^q(n)) \sum_{p=1}^{M_t} \boldsymbol{D}_k^p(n) \boldsymbol{h}_k^{p,q}(n) + \boldsymbol{v}^q(n), \\ &= \sum_{k=1}^{K} \tilde{\boldsymbol{D}}_{\varepsilon,k}^q(n) \boldsymbol{h}_k^q(n) + \boldsymbol{v}^q(n),\end{aligned} \tag{2}$$

where

$$\begin{aligned}
&\boldsymbol{h}_k^{p,q}(n) \triangleq \left[h_{k,0}^{p,q}(n), h_{k,1}^{p,q}(n), \ldots, h_{k,L_f-1}^{p,q}(n)\right]^T, \\
&\boldsymbol{D}_k^p(n) \triangleq \begin{bmatrix} d_{k,0}^p(n) & d_{k,N-1}^p(n) & \ldots & d_{k,N-L_f+1}^p(n) \\ d_{k,1}^p(n) & d_{k,0}^p(n) & \ldots & d_{k,N-L_f+2}^p(n) \\ \vdots & \vdots & \ldots & \ldots \\ d_{k,N-1}^p(n) & d_{k,N-2}^p(n) & \ldots & d_{k,N-L_f}^p(n) \end{bmatrix}, \\
&\boldsymbol{\Delta}(\varepsilon_k^q(n)) \triangleq e^{j\theta_k^q(n)} \mathrm{diag}(1, e^{j\frac{2\pi\varepsilon_k^q(n)}{N}}, .., e^{j\frac{2\pi(N-1)\varepsilon_k^q(n)}{N}}), \\
&\theta_k^q(n) \triangleq 2\pi \sum_{l=0}^{n-1} \varepsilon_k^q(l), \\
&\boldsymbol{h}_k^q(n) \triangleq \left[\boldsymbol{h}_k^{1,q}(n)^T, \boldsymbol{h}_k^{2,q}(n)^T, .., \boldsymbol{h}_k^{M_t,q}(n)^T\right]^T, \\
&\tilde{\boldsymbol{D}}_{\varepsilon,k}^q(n) \triangleq \left[\boldsymbol{\Delta}(\varepsilon_k^q(n)) \boldsymbol{D}_k^1(n), .., \boldsymbol{\Delta}(\varepsilon_k^q(n)) \boldsymbol{D}_k^{M_t}(n)\right].
\end{aligned} \tag{3}$$

Recall that the CP length equals the channel length plus timing offset. Under such an assumption, the timing errors do not explicitly appear in the received signal model [14]. Thus, we have suppressed the timing errors in (2).

Several approaches have been proposed to model the time-varying channels and frequency offsets in mobile environments. Since we assume a normalized Doppler spread $f_D T_d \ll 1$,we adopt the following first-order autoregressive (AR) parametric model widely used in [10], [13], [18]–[20] to characterize the time-varying frequency offset and channel responses.

$$\begin{aligned}\varepsilon_k^q(n) &= \alpha_{k,\varepsilon}^q \varepsilon_k^q(n-1) + w_{k,\varepsilon}^q(n), \\ \boldsymbol{h}_k^q(n) &= \alpha_{k,h}^q \boldsymbol{h}_k^q(n-1) + \boldsymbol{w}_{k,h}^q(n), \end{aligned} \tag{4}$$

where $w_{k,\varepsilon}^q(n) \sim \mathcal{N}(w_{k,\varepsilon}^q(n); 0, \eta_{k,\varepsilon}^q)$ and $\boldsymbol{w}_{k,h}^q(n) \sim \mathcal{CN}(\boldsymbol{w}_{k,h}^q(n); \mathbf{0}, \eta_{k,h}^q \boldsymbol{I}_{M_t L_f})$. Furthermore, we assume that $\left|\alpha_{k,\varepsilon}^q\right| < 1$ and $\left|\alpha_{k,h}^q\right| < 1$. Note that we assume *independent* fading across transmit antennas and multipaths. The time-variation in CFO in (4) arises from (a) local oscillator instability due to temperature/voltage variations and (b) changes in relative platform Doppler velocity. In the next section, we propose a pilot-aided parallel Schmidt extended Kalman filter approach to estimate $\varepsilon_k^q(n)$ and $\boldsymbol{h}_k^q(n)$ based on the received signal $\boldsymbol{r}^q(n)$, assuming exact knowledge of $\{\boldsymbol{D}_k^p(n)\}$, $\left\{\alpha_{k,\varepsilon}^q\right\}$, and $\left\{\alpha_{k,h}^q\right\}$ in the BS.

III. Parallel Schmidt Kalman Filter for Joint Channel and CFO Estimation

A. Kalman Filter Formulation

To facilitate the derivation of the Schmidt Kalman filter, we need to convert the complex signal model developed in the previous section into a *real-valued* form. We begin by defining the following *real-valued* quantities.

$$\begin{aligned}
&\mathbb{H}(\varepsilon(n)) \triangleq \mathbb{H}_k^q(\varepsilon_k^q(n)) \\
&= \begin{bmatrix} \mathrm{Re}\{\tilde{\boldsymbol{D}}_{\varepsilon,k}^q(n)\} & -\mathrm{Im}\{\tilde{\boldsymbol{D}}_{\varepsilon,k}^q(n)\} \\ \mathrm{Im}\{\tilde{\boldsymbol{D}}_{\varepsilon,k}^q(n)\} & \mathrm{Re}\{\tilde{\boldsymbol{D}}_{\varepsilon,k}^q(n)\} \end{bmatrix}, \\
&\boldsymbol{f}(n) \triangleq \boldsymbol{f}_k^q(n) = [\mathrm{Re}\{\boldsymbol{h}_k^q(n)\}^T, \mathrm{Im}\{\boldsymbol{h}_k^q(n)\}^T]^T, \\
&\mathbb{H}_{\backslash k}(\varepsilon_{\backslash k}(n)) = \\
&\left[\mathbb{H}_1^q(\varepsilon_1(n)), .., \mathbb{H}_{k-1}^q(\varepsilon_{k-1}(n)), \mathbb{H}_{k+1}^q(\varepsilon_{k+1}(n)), ..\mathbb{H}_K^q(\varepsilon_K(n))\right], \\
&\boldsymbol{f}_{\backslash k}(n) = [\boldsymbol{f}_1^q(n)^T, .., \boldsymbol{f}_{k-1}^q(n)^T, \boldsymbol{f}_{k+1}^q(n)^T, .., \boldsymbol{f}_K^q(n)^T]^T, \\
&\boldsymbol{z}(n) \triangleq \boldsymbol{z}_k^q(n) \sim \mathcal{N}(\boldsymbol{z}_k^q(n); \mathbf{0}, N_0/T_s \boldsymbol{I}_{2N}),
\end{aligned} \tag{5}$$

where the subscript $\backslash k$ stands for the exclusion of the parameters associated with the k-th user. For example, $\varepsilon_{\backslash k}(n)$ denotes all $\varepsilon_i(n)$s except $\varepsilon_k(n)$.

Utilizing the real-valued quantities defined above, (2) can be rewritten as the following real-valued system and observation equations:

$$\begin{aligned}\boldsymbol{y}(n) &= \mathbb{H}(\varepsilon(n))\boldsymbol{f}(n) + \mathbb{H}_{\backslash k}(\varepsilon_{\backslash k}(n))\boldsymbol{f}_{\backslash k}(n) + \boldsymbol{z}(n), \\ \boldsymbol{f}(n) &= \boldsymbol{A}_{k,h}^q \boldsymbol{f}(n-1) + \boldsymbol{w}_{k,h}^r(n), \\ \varepsilon(n) &= \alpha_{k,\varepsilon}^q \varepsilon(n-1) + w_{k,\varepsilon}^q(n), \end{aligned} \tag{6}$$

where $\boldsymbol{A}_{k,h}^q = [\boldsymbol{I}_2 \otimes \alpha_{k,h}^q \boldsymbol{I}_{M_t L_f}]$, and $\boldsymbol{w}_{k,h}^r(n)$ is similarly the real/imaginary partitioned version of the process noise in (4). Note that $\boldsymbol{y}(n)$ in (6) corresponds to the received vector only from the q-th antenna and $\boldsymbol{f}(n)$ and $\mathbb{H}(\varepsilon(n))$ are quantities related to the k-th user as received at antenna q. It should be emphasized that although the k-th user uses orthogonal sets of subcarriers in $\mathcal{I}_k$, self-interference and MAI occur due to CFOs [16]. Thus, vectors $\mathbb{H}(\varepsilon(n))\boldsymbol{f}(n)$ and $\mathbb{H}_{\backslash k}(\varepsilon_{\backslash k}(n))\boldsymbol{f}_{\backslash k}(n)$ cannot remain orthogonal. The EKF can be employed to jointly estimate $\boldsymbol{f}_k^q(n)$ and $\epsilon_k^q(n)$ by concatenating all those vectors for $k = 1, 2, \ldots K$. However, such EKF approaches are susceptible to divergence problems due to the nonlinear nature of CFO [9], [10]. Furthermore, direct computation of the Kalman gain for (6) using the concatenated state vector of all users is highly inefficient. In the following, we propose to use the suboptimal parallel Schmidt extended Kalman filters to develop a feasible estimator.

B. Parallel Schmidt Extended Kalman Filter (SEKF)

Following the notation in [15], we propose a parallel bank of K Schmidt extended Kalman filters (SEKFs) with each SEKF related to one desired user. Without loss of generality, we assume that the k-th user is the desired user for the k-th SEKF. In the k-th SEKF, we divide the state vector into the essential state vector (ESV) related to the q-th antenna of the k-th user,

$$\boldsymbol{x}(n) \triangleq [\varepsilon(n), \boldsymbol{f}(n)^T]^T, \tag{7}$$

and the *nuisance* state vector (NSV) $\boldsymbol{x}_{\backslash k}(n)$ for the interfering users. Note that $\boldsymbol{x}(n)$ in (7) is understood to represent $\boldsymbol{x}_k^q(n)$.

Substituting (7) and $\boldsymbol{x}_{\backslash k}(n)$ into (6) and applying a first-order approximation, we can obtain the following new pair of linearized system and observation equations as

$$\begin{aligned}
&\boldsymbol{y}(n) \approx \mathbb{H}(\hat{\varepsilon}(n|n-1))\hat{\boldsymbol{f}}(n|n-1) \\
&\quad + \mathbb{H}_{\backslash k}(\hat{\varepsilon}_{\backslash k}(n|n-1))\hat{\boldsymbol{f}}_{\backslash k}(n|n-1) \\
&\quad + \begin{bmatrix}\boldsymbol{J}(n) & \boldsymbol{J}_{\backslash k}(n)\end{bmatrix} \begin{bmatrix} \boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n|n-1) \\ \boldsymbol{x}_{\backslash k}(n) - \hat{\boldsymbol{x}}_{\backslash k}(n|n-1) \end{bmatrix} + \boldsymbol{z}(n), \\
&\boldsymbol{x}(n) = \boldsymbol{A}\boldsymbol{x}(n-1) + \boldsymbol{w}(n), \\
&\boldsymbol{x}_{\backslash k}(n) = \boldsymbol{A}_{\backslash k}\boldsymbol{x}_{\backslash k}(n-1) + \boldsymbol{w}_{\backslash k}(n),
\end{aligned} \tag{8}$$

where $\boldsymbol{A} = \mathrm{blkdiag}\left(\alpha_{k,\varepsilon}^q, \boldsymbol{A}_{k,h}^q\right)$ and $\boldsymbol{x}_{\backslash k}(n)$ denotes all state vectors excluding $\boldsymbol{x}_k(n)$ of the desired user. The noise term in (8) is given as $\boldsymbol{w}(n) \sim \mathcal{N}(\boldsymbol{w}(n); \mathbf{0}, \boldsymbol{Q})$ with $\boldsymbol{Q} = \mathrm{blkdiag}\left(\eta_{k,\varepsilon}^q, \boldsymbol{I}_2 \otimes \eta_{k,h}^q \boldsymbol{I}_{M_t L_f}/2\right)$. Similarly we can compute $\boldsymbol{A}_{\backslash k}$, $\boldsymbol{w}_{\backslash k}(n) \sim \mathcal{N}(\boldsymbol{w}_{\backslash k}(n); \mathbf{0}, \boldsymbol{Q}_{\backslash k})$ and $\boldsymbol{Q}_{\backslash k}$. Furthermore, the Jacobian matrix $\boldsymbol{J}(n)$ in (8) is the gradient w.r.t. the state vector of the nonlinear measurement function and equals

$$\boldsymbol{J}(n) \triangleq [\boldsymbol{J}_\varepsilon(n) \ \boldsymbol{J}_h(n)] \ \in \mathbb{R}^{2N \times (2M_t L_f + 1)}, \tag{9}$$

See (10) at the top of the next page.

The Jacobian $\boldsymbol{J}_{\backslash k}(n)$ can be similarly obtained and reads

$$\begin{aligned}&\boldsymbol{J}_{\backslash k}(n) = [\boldsymbol{J}_1(n), .., \boldsymbol{J}_{k-1}(n), \boldsymbol{J}_{k+1}(n), .., \boldsymbol{J}_K(n)] \\ &\in \mathbb{R}^{2N \times (K-1)(2M_t L_f + 1)}.\end{aligned} \tag{11}$$

For the conventional EKF, the covariance matrix $\boldsymbol{P}(n|n-1)$

$$\boldsymbol{J}_\varepsilon(n)=\left[\begin{array}{c}-\mathrm{Im}\{\boldsymbol{\Lambda\Delta}(\hat\varepsilon(n|n-1))\boldsymbol{D}^q_{\varepsilon,k}(n)\},-\mathrm{Re}\{\boldsymbol{\Lambda\Delta}(\hat\varepsilon(n|n-1))\boldsymbol{D}^q_{\varepsilon,k}(n)\}\\ \mathrm{Re}\{\boldsymbol{\Lambda\Delta}(\hat\varepsilon(n|n-1))\boldsymbol{D}^q_{\varepsilon,k}(n)\},-\mathrm{Im}\{\boldsymbol{\Lambda\Delta}(\hat\varepsilon(n|n-1))\boldsymbol{D}^q_{\varepsilon,k}(n)\}\end{array}\right]\times\hat{\boldsymbol{f}}(n|n-1),$$
$$\boldsymbol{J}_h(n)=\mathbb{H}(\hat\varepsilon(n|n-1)),\boldsymbol{\Lambda}\triangleq\frac{2\pi}{N}\,\mathrm{diag}\,(0,1/N,..(N-1)/N)\,. \tag{10}$$

and the Kalman gain matrix $\boldsymbol{K}(n)$ take the following forms.

$$\boldsymbol{P}(n|n-1)=\left[\begin{array}{cc}\boldsymbol{P}_{k,k}(n|n-1) & \boldsymbol{P}_{k,\backslash k}(n|n-1)\\ \boldsymbol{P}_{\backslash k,k}(n|n-1) & \boldsymbol{P}_{\backslash k,\backslash k}(n|n-1)\end{array}\right],$$
$$\boldsymbol{K}(n)=\left[\begin{array}{c}\boldsymbol{K}_k(n)\\ \boldsymbol{K}_{\backslash k}(n)\end{array}\right]=\left[\begin{array}{c}\boldsymbol{P}_{k,k}(n|n-1)\boldsymbol{J}(n)^T+\boldsymbol{P}_{k,\backslash k}(n|n-1)\boldsymbol{J}_{\backslash k}(n)^T,\\ \boldsymbol{P}_{\backslash k,k}(n|n-1)\boldsymbol{J}(n)^T+\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1)\boldsymbol{J}_{\backslash k}(n)^T\end{array}\right]\mathcal{A}, \tag{12}$$

where

$$(\mathcal{A})^{-1}=\boldsymbol{J}(n)\boldsymbol{P}_{k,k}(n|n-1)\boldsymbol{J}(n)^T+\boldsymbol{J}(n)\boldsymbol{P}_{k,\backslash k}(n|n-1)\boldsymbol{J}_{\backslash k}(n)^T+\boldsymbol{J}_{\backslash k}(n)\boldsymbol{P}_{\backslash k,k}(n|n-1)\boldsymbol{J}(n)^T+\boldsymbol{J}_{\backslash k}(n)\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1)\boldsymbol{J}_{\backslash k}(n)^T+N_0/T_s\boldsymbol{I}. \tag{13}$$

In (12), $\boldsymbol{K}(n)$ is the total EKF Kalman gain for joint tracking of all uplink users.

In lieu of computing the composite $\boldsymbol{K}(n)$ in (12), the SEKF sets the NSV part of the Kalman gain $\boldsymbol{K}_{\backslash k}(n)$ to zero, which results in

$$\hat{\boldsymbol{K}}(n)=[\boldsymbol{K}_{k,\mathrm{SKF}}(n)^T,\boldsymbol{0}^T]^T. \tag{14}$$

Approximation (14) is exact when estimation errors between the k-th user and its interfering users are uncorrelated and the error covariance of the interfering users is zero, i.e. $\boldsymbol{P}_{\backslash k,k}(n|n-1)=\boldsymbol{0}$ and $\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1)=\boldsymbol{0}$.

Now the estimate for the ESV becomes

$$\boldsymbol{K}_{k,\mathrm{SKF}}(n)=\left(\boldsymbol{P}_{k,k}(n|n-1)\boldsymbol{J}(n)^T+\boldsymbol{P}_{k,\backslash k}(n|n-1)\boldsymbol{J}_{\backslash k}(n)^T\right)\mathcal{A},$$
$$\hat{\boldsymbol{x}}(n|n)=\hat{\boldsymbol{x}}(n|n-1)+\boldsymbol{K}_{k,\mathrm{SKF}}(n)\left(\boldsymbol{y}(n)-\mathbb{H}(\hat\varepsilon(n|n-1))\hat{\boldsymbol{f}}(n|n-1)-\mathbb{H}_{\backslash k}(\hat\varepsilon_{\backslash k}(n|n-1))\hat{\boldsymbol{f}}_{\backslash k}(n|n-1)\right), \tag{15}$$

where the *predicted* estimate $\hat{\boldsymbol{x}}(n+1|n)$ and the corresponding error covariance matrix $\boldsymbol{P}_{k,k}(n+1|n)$ for the desired user are computed as follows.

$$\begin{aligned}\hat{\boldsymbol{x}}(n+1|n)&=\boldsymbol{A}\hat{\boldsymbol{x}}(n|n),\\ \boldsymbol{P}_{k,k}(n+1|n)&=\boldsymbol{A}\boldsymbol{P}_{k,k}(n|n)(\boldsymbol{A})^T+\boldsymbol{Q}.\end{aligned} \tag{16}$$

To evaluate (16), we need to compute the joint NSV and ESV error covariance matrix based on (15), $\boldsymbol{P}(n|n)$ which following [15] is

$$\boldsymbol{P}(n|n)=\left[\begin{array}{cc}\boldsymbol{P}_{k,k}(n|n) & \boldsymbol{P}_{k,\backslash k}(n|n)\\ \boldsymbol{P}_{\backslash k,k}(n|n) & \boldsymbol{P}_{\backslash k,\backslash k}(n|n)\end{array}\right], \tag{17}$$

with each sub-matrix given by

$$\begin{aligned}&\boldsymbol{P}_{k,k}(n|n)=\\ &\mathcal{B}\boldsymbol{P}_{k,k}(n|n-1)(\mathcal{B})^T-\mathcal{B}\boldsymbol{P}_{k,\backslash k}(n|n)(\mathcal{C})^T-\mathcal{C}\boldsymbol{P}_{\backslash k,k}(n|n)(\mathcal{B})^T+\\ &\mathcal{C}\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1)(\mathcal{C})^T+\frac{N_0}{T_s}\boldsymbol{K}_{k,\mathrm{SKF}}(n)\boldsymbol{K}_{k,\mathrm{SKF}}(n)^T,\\ &\boldsymbol{P}_{k,\backslash k}(n|n)=\boldsymbol{P}_{\backslash k,k}(n|n)^T\\ &=\mathcal{B}\boldsymbol{P}_{k,\backslash k}(n|n-1)-\mathcal{C}\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1),\\ &\boldsymbol{P}_{\backslash k,\backslash k}(n|n)=\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1),\end{aligned} \tag{18}$$

where $\mathcal{B}\triangleq(\boldsymbol{I}-\boldsymbol{K}_{k,\mathrm{SKF}}(n)\boldsymbol{J}(n))$ and $\mathcal{C}\triangleq\boldsymbol{K}_{k,\mathrm{SKF}}(n)\boldsymbol{J}_{\backslash k}(n)$. Note that the error covariance for the interfering users $\boldsymbol{P}_{\backslash k,\backslash k}(n|n)$ is not updated by the k-th SEKF.

To summarize the SEKF, the k-th filter produces $\{\hat{\boldsymbol{x}}_k(n+1|n),\boldsymbol{P}_{k,k}(n+1|n)\}$ by exploiting the predictions $\{\hat{\boldsymbol{x}}_k(n|n-1),\hat{\boldsymbol{x}}_{\backslash k}(n|n-1),\boldsymbol{P}_{k,k}(n|n-1),\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1)\}$. If the cross error covariance is zero, i.e. $\boldsymbol{P}_{k,\backslash k}(n|n-1)=\boldsymbol{0}$, then we have the so-called inflated covariance system [21]. It is worth noting that $\boldsymbol{P}_{k,\backslash k}(n|n)=\boldsymbol{0}$ only when $\boldsymbol{P}_{k,\backslash k}(n|n-1)=\boldsymbol{0}$ and $\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1)=\boldsymbol{0}$. In general, we have $\boldsymbol{P}_{k,\backslash k}(n|n)\neq\boldsymbol{0}$ due to the error covariance of the interfering users.

IV. Parallel Approximate-Rao-Blackwellized Particle Filter

A Rao-Blackwellized particle filter using the first-order approximation to the observation is proposed for each SEKF with structure suggested in [12], [22]–[24]. The state vector is factorized into two terms related to the channel coefficients $\boldsymbol{f}(n)$ and the CFO state variable $\varepsilon(n)$, respectively.

Let $\boldsymbol{f}(n)$ be the channel vector and $\epsilon_i^n=\{\epsilon_i(n),\epsilon_i(n-1)\ldots\epsilon_i(0)\}$ be the i-th trajectory of frequency offset particles for the k-th user and the q-th receive antenna with k and q again suppressed for clarity. We define the cumulative observation sequences as $\boldsymbol{y}^n\triangleq\{\boldsymbol{y}(n),\boldsymbol{y}(n-1),\cdots,\boldsymbol{y}(0)\}$. The importance sampling estimate of $p(\boldsymbol{f}(n)|\boldsymbol{y}^n)$ is given by [24]

$$p(\boldsymbol{f}(n)|\boldsymbol{y}^n)=\sum_{i=1}^{N_p}\beta_i(n)p(\boldsymbol{f}(n)|\boldsymbol{y}^n,\epsilon_i^n). \tag{19}$$

We assume N_p particles in the Rao-Blackwellization. In [24], the importance sampling weight for unbiased estimation is given by

$$\beta_i(n)=\frac{p(\epsilon_i^n|\boldsymbol{y}^n)}{\pi(\epsilon_i(n)|\boldsymbol{y}^n,\epsilon_i^{n-1})\pi(\epsilon_i^{n-1}|\boldsymbol{y}^{n-1})}, \tag{20}$$

where $\pi(\cdot)$ denotes the sampling density. Note that the prior sampling density $\pi(\epsilon_i^{n-1}|\boldsymbol{y}^{n-1})$ is a causal function of measurements up to time $n-1$. The minimum variance choice [25] for the sample density of current sample $\epsilon_i(n)$ is $\pi(\epsilon_i(n)|\boldsymbol{y}^n,\epsilon_i^{n-1})=p(\epsilon_i(n)|\boldsymbol{y}^n,\epsilon_i^{n-1})$, i.e. the true density function of the CFO given the past trajectory ϵ_i^{n-1} and cumulative measurements $\boldsymbol{y}^n$. The conditional distribution under this Rao-Blackwellization of $\boldsymbol{f}(n)$ is thus given by :

$$p(\boldsymbol{f}(n)|\boldsymbol{y}^n, \epsilon_i^n) = \mathcal{N}(\boldsymbol{f}(n); \hat{\boldsymbol{f}}_i(n|n), \boldsymbol{P}_{k,k,i}(n|n)), \quad (21)$$

where $\hat{\boldsymbol{f}}_i(n|n), \boldsymbol{P}_{k,k,i}(n|n)$ are the Kalman filter estimate and covariance using the particular CFO trajectory ϵ_i^n, respectively. These quantities will be approximated using the Schmidt-Kalman filter for the multi-user case.

To utilize Rao-Blackwellization, we rewrite the linearized system in (8) by removing the terms related to $\boldsymbol{J}(n)$, i.e. only the first-order linearization w.r.t. NSV is employed.

$$\begin{aligned}\boldsymbol{y}(n) &\approx \mathbb{H}(\varepsilon(n))\boldsymbol{f}(n)+\\ &\mathbb{H}_{\backslash k}(\hat{\varepsilon}_{\backslash k}(n|n-1))\hat{\boldsymbol{f}}_{\backslash k}(n|n-1)+\\ &\boldsymbol{J}_{\backslash k}(n)\left[\boldsymbol{x}_{\backslash k}(n)-\hat{\boldsymbol{x}}_{\backslash k}(n|n-1)\right]+\boldsymbol{z}'(n).\end{aligned} \quad (22)$$

Using the Kalman filter quantities yields

$$\begin{aligned}&\pi(\epsilon_i(n)|\boldsymbol{y}^n, \epsilon_i^{n-1}) \propto p(\boldsymbol{y}(n)|\boldsymbol{y}^{n-1}, \epsilon_i^n)p(\epsilon_i(n)|\epsilon_i(n-1)) =\\ &\mathcal{N}(\boldsymbol{y}(n); \hat{\boldsymbol{y}}_i(n|n-1), \Sigma_i(n|n-1))\mathcal{N}(\epsilon_i(n); \alpha_{k,\epsilon}^q\epsilon_i(n-1), \eta_{k,\epsilon}^q),\end{aligned} \quad (23)$$

where

$$\begin{aligned}&\hat{\boldsymbol{y}}_i(n|n-1) =\\ &\mathbb{H}(\epsilon_i(n))\hat{\boldsymbol{f}}_i(n|n-1)+\mathbb{H}_{\backslash k}(\hat{\varepsilon}_{\backslash k}(n|n-1))\hat{\boldsymbol{f}}_{\backslash k}(n|n-1),\\ &\boldsymbol{\Sigma}_i(n|n-1) =\\ &N_0/T_s\boldsymbol{I}+\mathbb{H}(\epsilon_i(n))\boldsymbol{P}_{k,k,i}(n|n-1)\mathbb{H}(\epsilon_i(n))^T+\\ &\boldsymbol{J}_{\backslash k}(n)\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1)\boldsymbol{J}_{\backslash k}(n)^T.\end{aligned} \quad (24)$$

The importance weight corresponding to the minimum-variance sampling distribution is [25]

$$\beta_i(n) = \frac{1}{c_1}p(\boldsymbol{y}(n)|\epsilon_i^{n-1}, \boldsymbol{y}^{n-1})\beta_i(n-1), \quad (25)$$

where c_1 is a constant. The density function in (25) can be approximated using the samples $\epsilon_j(n)$ as

$$\begin{aligned}&p(\boldsymbol{y}(n)|\epsilon_i^{n-1}, \boldsymbol{y}^{n-1}) =\\ &\int p(\boldsymbol{y}(n)|\varepsilon(n), \epsilon_i^{n-1}, \boldsymbol{y}^{n-1})p(\varepsilon(n)|\epsilon_i^{n-1}, \boldsymbol{y}^{n-1})d\varepsilon(n) \approx\\ &\frac{1}{c_2}\sum_{j=1}^{N_p}\mathcal{N}\Big(\boldsymbol{y}(n); \mathbb{H}(\epsilon_j(n))\hat{\boldsymbol{f}}_i(n|n-1)+\\ &\mathbb{H}_{\backslash k}(\hat{\varepsilon}_{\backslash k}(n|n-1)\ \hat{\boldsymbol{f}}_{\backslash k}(n|n-1), \boldsymbol{\Sigma}_i(n|n-1)\Big)\times\\ &\mathcal{N}\left(\epsilon_j(n); \alpha_{k,\epsilon}^q\epsilon_i(n-1), \eta_{k,\epsilon}^q\right),\end{aligned} \quad (26)$$

where the constant $1/c_2$ normalizes (26) to a valid probability density function. The weight $\beta_i(n)$ is *independent* of the current sample $\epsilon_i(n)$, but dependent on the past trajectory ϵ_i^{n-1}.

A. Sampling With M-Statistics

To generate the frequency offset sample $\{\epsilon_i(n)\}$, we assume in the following that we have ϵ_i^{n-1}, $\{\hat{\boldsymbol{f}}_i(n-1|n-1)\}$, and $\{\boldsymbol{P}_{k,k,i}(n-1|n-1)\}$. Using the dynamic model, we generate a tentative sample that will be updated later after the sampling process :

$$\epsilon_i(n) = \alpha_{k,\epsilon}^q\epsilon_i(n-1)+w_{k,\epsilon}^q. \quad (27)$$

Conditioned on $\epsilon_i(n)$, we compute $\hat{\boldsymbol{f}}_i(n|n-1)$ and $\boldsymbol{P}_{k,k,i}(n|n-1)$ via (16) with $\hat{\boldsymbol{f}}_i(n-1|n-1)$ and $\boldsymbol{P}_{k,k,i}(n-1|n-1)$ computed from the trajectory ϵ_i^{n-1}.

The CFO uncertainty region is approximated by a $2M+1$ point grid on the interval $[-M\Delta_\epsilon, M\Delta_\epsilon)$ with $\delta_m = m\Delta_\epsilon$, where $m = -M, -M+1, \cdots, M$. Next, we compute the sampling distributions at the grid points using a new measurement $\boldsymbol{y}(n)$ as follows :

$$\begin{aligned}&\pi_i(m) \triangleq \mathcal{N}(\boldsymbol{y}(n); \hat{\boldsymbol{y}}_{(i,m)}(n|n-1), \Sigma_{(i,m)}(n|n-1))\times\\ &\mathcal{N}(\delta_m; \alpha_{k,\epsilon}^q\epsilon_i(n-1), \eta_{k,\epsilon}^q),\end{aligned} \quad (28)$$

where

$$\begin{aligned}\hat{\boldsymbol{y}}_{(i,m)}(n|n-1) &= \mathbb{H}(\delta_m)\hat{\boldsymbol{f}}_i(n|n-1)\\ &+ \mathbb{H}_{\backslash k}(\hat{\varepsilon}_{\backslash k}(n|n-1))\hat{\boldsymbol{f}}_{\backslash k}(n|n-1)\end{aligned}$$

and

$$\begin{aligned}\Sigma_{(i,m)}(n|n-1) &= N_0/T_s\boldsymbol{I}\\ &+ \mathbb{H}(\delta_m)\boldsymbol{P}_{k,k,(i,m)}(n|n-1)\mathbb{H}(\delta_m)^T\\ &+ \boldsymbol{J}_{\backslash k}(n)\boldsymbol{P}_{\backslash k,\backslash k}(n|n-1)\boldsymbol{J}_{\backslash k}(n)^T.\end{aligned}$$

In this equation, $\boldsymbol{P}_{k,k,(i,m)}(n|n-1)$ is the corresponding covariance for particle i and grid point δ_m. Upon computing all $\pi_i(m)$, we choose a random value $\tilde{m}$ satisfying the following inequality:

$$\sum_{l=-M}^{\tilde{m}}\frac{\pi_i(l)}{\sum_{l'=-M}^{M}\pi_i(l')} > u > \sum_{l=-M}^{\tilde{m}-1}\frac{\pi_i(l)}{\sum_{l'=-M}^{M}\pi_i(l')}, \quad (29)$$

where u is a uniform r.v. on $[0,1]$. With $\tilde{m}$ obtained above, we select $\epsilon_i(n) = \delta_{\tilde{m}}$ and update $\epsilon_i^n = \{\epsilon_i(n), \epsilon_i^{n-1}\}$.

Finally, the resulting conditional distribution given in (21) can be computed as

$$p(\boldsymbol{f}(n)|\boldsymbol{y}^n, \epsilon_i^n) = \mathcal{N}(\boldsymbol{f}(n); \hat{\boldsymbol{f}}_i(n|n), \boldsymbol{P}_{k,k,i}(n|n)), \quad (30)$$

where

$$\begin{aligned}&\hat{\boldsymbol{f}}_i(n|n) = \hat{\boldsymbol{f}}_i(n|n-1)+\boldsymbol{K}_i(n)(\boldsymbol{y}(n)-\hat{\boldsymbol{y}}_i(n|n-1)),\\ &\boldsymbol{K}_i(n) =\\ &\left(\boldsymbol{P}_{k,k,i}(n|n-1)\boldsymbol{J}_i(n)^T+\boldsymbol{P}_{(k,i),\backslash k}(n|n-1)\boldsymbol{J}_{\backslash k}(n)^T\right)\mathcal{A}_i\end{aligned} \quad (31)$$

The matrices $\boldsymbol{P}_{(k,i),\backslash k}(n|n-1)$, $\boldsymbol{J}_i(n)$, and $\mathcal{A}_i$ correspond to $\boldsymbol{P}_{k,\backslash k}(n|n-1)$, $\boldsymbol{J}(n)$, and $\mathcal{A}$ for the i-th particle. These terms are defined in equations (18),(9), and (13). However, the Jacobian $\boldsymbol{J}_i(n)$ reduces to $\mathbb{H}(\epsilon_i(n))$ for the particle filter case. These particle dependent terms are updated independently by the k users. Thus, the unconditional distribution is given by

$$p(\boldsymbol{f}(n)|\boldsymbol{y}^n) = \sum_{i=1}^{N_p}\beta_i(n)\mathcal{N}(\boldsymbol{f}(n); \hat{\boldsymbol{f}}_i(n|n), \boldsymbol{P}_{k,k,i}(n|n)). \quad (32)$$

Since the proposed scheme consists of suboptimal parallel filters, it is referred to as the Schmidt-Kalman approximate-Rao-Blackwellized particle filter (SK-APF) in the sequel. SK-APF is summarized in Algorithm 1.

Algorithm 1 Parallel Schmidt-Kalman Approximate-Rao-Blackwellized Particle Filter

1. Given ϵ_i^{n-1}, $\hat{\boldsymbol{f}}_i(n-1|n-1)$, $\boldsymbol{P}_{k,k,i}(n-1|n-1)$ for the k-th user
(a) Generate a sample $\epsilon_i(n) = \alpha^q_{k,\varepsilon}\epsilon_i(n-1) + w^q_{k,\varepsilon}$.
(b) Compute interfering channel prediction $\hat{\boldsymbol{f}}_{\setminus k}(n|n-1)$ and covariance $\boldsymbol{P}_{\setminus k,\setminus k}(n|n-1), \forall k$.

$$\begin{aligned}\boldsymbol{P}_{\setminus k,\setminus k}(n|n-1) &= \boldsymbol{A}_{\setminus k}\boldsymbol{P}_{\setminus k,\setminus k}(n-1|n-1)\boldsymbol{A}^T_{\setminus k} + \boldsymbol{Q}_{\setminus k}\\ \boldsymbol{P}_{\setminus k,\setminus k}(n|n) &= \boldsymbol{P}_{\setminus k,\setminus k}(n|n-1)\end{aligned} \tag{33}$$

3. Receive next measurement $\boldsymbol{y}(n)$.
4. Compute Jacobian matrix $\boldsymbol{J}_{\setminus k}(n)$ for interfering users.
for $i = 1$ to N_p **do**
(c) Compute the conditional channel estimate $\hat{\boldsymbol{f}}_i(n|n-1)$ with covariance $\boldsymbol{P}_{k,k,i}(n|n-1)$ conditioned on ϵ_i^{n-1}.
for $m = 1$ to $2M+1$ **do**
(d) Define a grid: $\delta_m = \Delta_\epsilon(-M+m-1)$.
(e) Compute $\pi_i(m) \propto \mathcal{N}(\boldsymbol{y}(n); \hat{\boldsymbol{y}}_{(i,m)}(n|n-1), \Sigma_{(i,m)}(n|n-1))\mathcal{N}(\delta_m; \alpha^q_{k,\epsilon}\epsilon_i(n-1), \eta^q_{k,\epsilon})$ according to (28).
end for
(f) Find $\tilde{m}$ with a random variable $u = rand$ satisfying

$$\sum_{l=-M}^{\tilde{m}} \frac{\pi_i(l)}{\sum_{l'=-M}^{M}\pi_i(l')} > u > \sum_{l=-M}^{\tilde{m}-1} \frac{\pi_i(l)}{\sum_{l'=-M}^{M}\pi_i(l')}. \tag{34}$$

(g) Update a trajectory $\epsilon_i^n = \{\delta_{\tilde{m}}, \epsilon_i^{n-1}\}$.
(h) Update $\boldsymbol{P}_{\setminus k,k}(n|n)$ using eq. (18).
(i) Compute channel estimate and its error covariance matrix

$$\begin{aligned}\boldsymbol{P}_{k,k,i}(n|n) &= \mathcal{B}_i\boldsymbol{P}_{k,k,i}(n|n-1)(\mathcal{B}_i)^T - \mathcal{B}_i\boldsymbol{P}_{k,\setminus k}(n|n)(\mathcal{C}_i)^T\\ &\quad -\mathcal{C}_i\boldsymbol{P}_{\setminus k,k}(n|n)(\mathcal{B}_i)^T + \mathcal{C}_i\boldsymbol{P}_{\setminus k,\setminus k}(n|n-1)(\mathcal{C}_i)^T\\ &\quad +\frac{N_0}{T_s}\boldsymbol{K}_i(n)\boldsymbol{K}_i(n)^T,\\ \hat{\boldsymbol{f}}_i(n|n) &= \hat{\boldsymbol{f}}_i(n|n-1) + \boldsymbol{K}_i(n)(\boldsymbol{y}(n) - \hat{\boldsymbol{y}}_i(n|n-1)),\\ \mathcal{B}_i &= \boldsymbol{I} - \boldsymbol{K}_i(n)\mathbb{H}(\epsilon_i(n)),\\ \mathcal{C}_i &= \boldsymbol{K}_i(n)\boldsymbol{J}_{\setminus k}(n),\end{aligned} \tag{35}$$

where the Kalman gain is defined in (31).
(j) Compute importance weight $\beta_i(n)$ according to (25) and (26).
end for
5. Compute the unconditional channel and frequency offset estimates.

$$\hat{\boldsymbol{f}}(n|n) = \sum_{i=1}^{N_p}\beta_i(n)\hat{\boldsymbol{f}}_i(n|n) \text{ and } \hat{\varepsilon}(n) \approx \sum_{i=1}^{N_p}\beta_i(n)\epsilon_i(n). \tag{36}$$

6. Compute the unconditional error covariance for user k

$$\begin{aligned}\boldsymbol{P}_{k,k}(n|n) = \sum_{i=1}^{N_p}\beta_i(n)[(\hat{\boldsymbol{f}}_i(n|n)\hat{\boldsymbol{f}}_i(n|n)^T\\ +\boldsymbol{P}_{k,k,i}(n|n)] - \hat{\boldsymbol{f}}(n|n)\hat{\boldsymbol{f}}(n|n)^T.\end{aligned} \tag{37}$$

V. Joint Data Detection and Parameter Estimation

To this point, we have assumed that an entire training block is employed for joint channel and frequency offset estimation. In practical OFDMA systems, such training blocks are only available in the beginning of each frame, and only $N_{sp} < N$ pilot symbols are embedded into each data block to track the time-variations of CFOs and channels. However, for high-mobility applications, the pilots in each data block may not provide satisfactory tracking performance. Recently, an EM-based iterative joint data detection and channel/CFO estimation approach has been proposed in [14], where tentative data decisions are exploited to improve the channel/CFO estimates in an iterative manner. We next similarly combine joint data detection based on QRD-M [26], [27] with the SK-APF channel/CFO estimator.

Without loss of generality, we focus on data detection for the k-th user in the sequel. Furthermore, $\mathcal{I}_{k,d}$ and $\mathcal{I}_{k,p}$ denote the data and pilot carrier index sets for the k-th user, respectively. The DFT of the received signal at antenna q is

$$\begin{aligned}\tilde{\boldsymbol{r}}^q(n) = \boldsymbol{W}\boldsymbol{\Delta}(\varepsilon^q_k(n))\sum_{p=1}^{M_t}\boldsymbol{D}^p_k(n)\boldsymbol{h}^{p,q}_k(n) +\\ \boldsymbol{W}\sum_{k'=1,k'\neq k}^{K}\boldsymbol{\Delta}(\varepsilon^q_{k'}(n))\sum_{p=1}^{M_t}\boldsymbol{D}^p_{k'}(n)\boldsymbol{h}^{p,q}_{k'}(n) + \tilde{\boldsymbol{v}}^q(n).\end{aligned} \tag{38}$$

Consider detection of the m-th data carrier assigned to the k-th user. In the presence of CFO, the corresponding received signal distorted by ICI can be expressed as

$$\begin{aligned}\tilde{r}^q_{k,m|\mathcal{I}_{k,d}}(n) &= \tilde{r}^q_{k,s,m|\mathcal{I}_{k,d}}(n) + \tilde{r}^q_{k,ici,m}(n)\\ &+ \tilde{v}^q_{k,m|\mathcal{I}_{k,d}}(n),\end{aligned} \tag{39}$$

where $\tilde{r}^q_{k,s,m|\mathcal{I}_{k,d}}(n)$ and $\tilde{r}^q_{k,ici,m}(n)$ are the desired signal and the ICI signal from other subcarriers, respectively. These terms can be explicitly written as

$$\begin{aligned}&\tilde{r}^q_{k,s,m|\mathcal{I}_{k,d}}(n) =\\ &e^{j\theta^q_k(n)}\mu(\varepsilon^q_k(n))\sum_{p=1}^{M_t}H^{p,q}_{k,m|\mathcal{I}_{k,d}}(n)\tilde{d}^p_{k,m|\mathcal{I}_{k,d}}(n),\\ &\tilde{r}^q_{k,ici,m}(n) =\\ &\sum_{k'=1}^{K}e^{-j\theta^q_{k'}(n)}\sum_{\substack{m'=0\\ m'\neq m}}^{N-1}\mu(\varepsilon^q_{k'}(n)+m'-m)\times\\ &\sum_{p=1}^{M_t}H^{p,q}_{k',m'}(n)\tilde{d}^p_{k',m'}(n),\end{aligned} \tag{40}$$

where $\mu(\theta)\stackrel{\triangle}{=}\frac{1}{N}e^{j\pi\theta(N-1)/N}\frac{sin(\pi\theta)}{sin(\pi\theta/N)}$ and $H^{p,q}_{k,m}(n)\stackrel{\triangle}{=}(\boldsymbol{W})_{(m,1:L_f)}\boldsymbol{h}^{p,q}_k(n)\stackrel{\triangle}{=}(\boldsymbol{W}_{L_f})_m\boldsymbol{h}^{p,q}_k(n)$. Note that in (40), $\tilde{r}^q_{k,ici,m}(n)$ includes CFO-induced self-interference and MAI [16].

Since the channel and CFO are unknown to the receiver, they are approximated by their predictions. Furthermore, some subcarriers are employed to modulate pilot symbols for all users. The estimate of the received signal using the channel/offset predictions to be employed for detection is

$$\begin{aligned}\hat{r}^q_{k,m|\mathcal{I}_{k,d}}(n|n-1) = \frac{e^{-j\hat{\theta}^q_k(n)}}{\mu(\hat{\epsilon}^q_k(n|n-1))}\tilde{r}^q_{k,s,m|I_{k,d}}(n) \approx\\ \hat{r}^q_{k,s,m|\mathcal{I}_{k,d}}(n|n-1) + n^q_{k,m|\mathcal{I}_k}(n|n-1),\end{aligned} \tag{41}$$

where the desired signal, interferers and noise are modeled as

[28]

$$\begin{aligned}
&\hat{r}^q_{k,s,m|\mathcal{I}_{k,d}}(n|n-1) = \sum_{p=1}^{M_t} \hat{H}^{p,q}_{k,m|\mathcal{I}_{k,d}}(n|n-1)\tilde{d}^p_{k,m|\mathcal{I}_{k,d}}(n),\\
&n^q_{k,m|\mathcal{I}_k}(n|n-1) =\\
&\frac{e^{-j\hat{\theta}^q_k(n|n-1)}}{\mu(\hat{\varepsilon}^q_k(n|n-1))}\left(\hat{r}^q_{k,ici,m}(n|n-1) + \tilde{v}^q_{k,m|\mathcal{I}_{k,d}}(n)\right),\\
&\hat{r}^q_{k,ici,m}(n|n-1) =\\
&\sum_{k'=1}^{K} e^{-j\hat{\theta}^q_{k'}(n|n-1)} \sum_{\substack{m'=0\\m'\neq m}}^{N-1} \mu(\hat{\varepsilon}^q_{k'}(n|n-1)+m'-m)\times\\
&\sum_{p=1}^{M_t} \hat{H}^{p,q}_{k',m'}(n|n-1)\tilde{d}^p_{k',m'}(n).
\end{aligned} \tag{42}$$

In (42), $\hat{\theta}^q_k(n|n-1) = 2\pi\sum_{l=0}^{n-1}\hat{\varepsilon}^q_k(l|l-1)$ with $\hat{\epsilon}^q_k(0|-1) = \hat{\epsilon}^q_k(0)$. Using time updated samples $\epsilon_i(n|n-1)$, we approximate the prediction by $\hat{\varepsilon}^q_k(n|n-1) \approx \sum_{i=1}^{N_p}\beta_i(n)\epsilon_i(n|n-1)$, with indices q, k suppressed in $\epsilon_i(n|n-1)$.

Proposition 1: The covariance of $n^q_{k,m|\mathcal{I}_k}(n|n-1)$ is given by

$$\begin{aligned}
&\sigma^2_{n_{q,k}}(n|n-1) \triangleq E\{|n^q_{k,m|\mathcal{I}_k}(n|n-1)|^2\} =\\
&\frac{2N_0}{T_s|\mu(\hat{\varepsilon}^q_k(n|n-1))|^2}+\\
&\frac{1}{M_t|\mu(\hat{\varepsilon}^q_k(n|n-1))|^2}\sum_{k'=1}^{K}\sum_{\substack{m'=0,\\m'\neq m}}^{N-1}|\mu(\hat{\varepsilon}^q_{k'}(n|n-1)+m'-m)|^2\times\\
&\mathrm{Tr}\{(\boldsymbol{I}_{M_t}\otimes(\boldsymbol{X}_W)_{m'})\,\hat{\boldsymbol{h}}^q_{k'}(n|n-1)\hat{\boldsymbol{h}}^q_{k'}(n|n-1)^H\},
\end{aligned} \tag{43}$$

where $(\boldsymbol{X}_W)_m \triangleq (\boldsymbol{W}_{L_f})^H_m(\boldsymbol{W}_{L_f})_m$.

The derivation of Proposition 1 is given in the Appendix.

Let $m \in \mathcal{I}_{k,d}$. According to (41), we have

$$\begin{aligned}
&\hat{\boldsymbol{r}}_{k|m}(n|n-1) \triangleq\\
&[\hat{r}^1_{k|m}(n|n-1), \hat{r}^2_{k|m}(n|n-1), \ldots, \hat{r}^{M_r}_{k|m}(n|n-1)]^T, \approx\\
&\begin{bmatrix} \hat{H}^{1,1}_{k|m}(n|n-1) & \ldots & \hat{H}^{M_t,1}_{k|m}(n|n-1)\\ \vdots & \ldots & \vdots\\ \hat{H}^{1,M_r}_{k|m}(n|n-1) & \ldots & \hat{H}^{M_t,M_r}_{k|m}(n|n-1)\end{bmatrix}\begin{bmatrix}\tilde{d}^1_{k|m}(n)\\ \vdots\\ \tilde{d}^{M_t}_{k|m}(n)\end{bmatrix}+\\
&\begin{bmatrix} n^1_{k,m|\mathcal{I}_k}(n|n-1)\\ \vdots\\ n^{M_r}_{k,m|\mathcal{I}_k}(n|n-1)\end{bmatrix},\\
&\triangleq \hat{\boldsymbol{H}}_{k|m}(n|n-1)\tilde{\boldsymbol{d}}_{k|m}(n) + \boldsymbol{n}_{k,m|\mathcal{I}_k}(n|n-1).
\end{aligned} \tag{44}$$

In (44), $\boldsymbol{n}_{k,m|\mathcal{I}_k}(n) \sim \mathcal{N}(\boldsymbol{n}_{k|m}(n); \boldsymbol{0}, \hat{\boldsymbol{\Sigma}}_{k|m}(n|n-1))$ with $\hat{\boldsymbol{\Sigma}}_{k,m|\mathcal{I}_k}(n|n-1) \triangleq \mathrm{diag}\{\sigma^2_{n_{1,k}}(n|n-1), \ldots, \sigma^2_{n_{M_r,k}}(n|n-1)\}$. This form of the covariance approximates the ICI at different receive antennas as independent, and is similar to that of [26] in the detection problem. The main difference is that the noise terms in (44) no longer have equal variances. To take advantage of QRD-M [26], we employ a a diagonal scaling of the received/corrected vector $\hat{\boldsymbol{r}}_{k|m}(n|n-1)$ for better detection performance [29]. Let $\hat{\boldsymbol{L}}_{k|m}(n|n-1)$ be the diagonal square root of $\hat{\boldsymbol{\Sigma}}_{k|m}(n|n-1)$. Premultiplying $\hat{\boldsymbol{r}}_{k|m}(n|n-1)$ by $\hat{\boldsymbol{L}}^{-1}_{k|m}(n|n-1)$ yields

$$\begin{aligned}
&\hat{\boldsymbol{r}}_{k|m}(n|n-1)^\dagger \approx\\
&\hat{\boldsymbol{L}}^{-1}_{k|m}(n|n-1)\hat{\boldsymbol{H}}_{k|m}(n|n-1)\tilde{\boldsymbol{d}}_{k|m}(n) + \boldsymbol{e}_{k|m}(n|n-1),
\end{aligned} \tag{45}$$

where $E\{\boldsymbol{e}_{k|m}(n|n-1)\boldsymbol{e}_{k|m}(n|n-1)^H\} \approx \boldsymbol{I}$.

Data detection of $\{\tilde{\boldsymbol{d}}_{k|m}(n)\}$ can be performed based on (45) by employing QRD-M. The QRD-M detector approximates the maximum-likelihood decision.

$$\begin{aligned}
&\hat{\boldsymbol{d}}_{k|m}(n) = \arg\min_{\tilde{\boldsymbol{d}}_{k|m}(n)\in|S|^{M_t}}\\
&\left\|\hat{\boldsymbol{r}}_{k|m}(n|n-1)^\dagger - \hat{\boldsymbol{L}}_{k|m}(n|n-1)^{-1}\hat{\boldsymbol{H}}_{k|m}(n|n-1)\tilde{\boldsymbol{d}}_{k|m}(n)\right\|^2,
\end{aligned} \tag{46}$$

where $\mathcal{S}$ denotes the signal constellation. For QRD-M, the condition $M_r \geq M_t$ is required, which can be easily satisfied in practical uplink MIMO-OFDMA systems.

Now let

$$\hat{\boldsymbol{L}}_{k|m}(n|n-1)^{-1}\hat{\boldsymbol{H}}_{k|m}(n|n-1) = \hat{\boldsymbol{Q}}_{\mathrm{QR}}(n|n-1)\hat{\boldsymbol{R}}_{\mathrm{QR}}(n|n-1)$$

be the QR decomposition, where $\hat{\boldsymbol{Q}}_{\mathrm{QR}}(n|n-1)$ is a unitary matrix and $\hat{\boldsymbol{R}}_{\mathrm{QR}}(n|n-1)$ is an upper triangular matrix. Substituting $\hat{\boldsymbol{L}}_{k|m}(n|n-1)^{-1}\hat{\boldsymbol{H}}_{k|m}(n|n-1) = \hat{\boldsymbol{Q}}_{\mathrm{QR}}(n|n-1)\hat{\boldsymbol{R}}_{\mathrm{QR}}(n|n-1)$ into (46), we have

$$\begin{aligned}
&\hat{\boldsymbol{d}}_{k|m}(n)_{\text{QRD-M}} \approx\\
&\arg\min_{\tilde{\boldsymbol{d}}_{k|m}(n)\in|S|^{M_t}}\left\|\hat{\boldsymbol{r}}_{k|m}(n|n-1)^\dagger_{\mathrm{QR}} - \hat{\boldsymbol{R}}_{\mathrm{QR}}(n|n-1)\tilde{\boldsymbol{d}}_{k|m}(n)\right\|^2,
\end{aligned} \tag{47}$$

where $\hat{\boldsymbol{r}}_{k|m}(n|n-1)^\dagger_{\mathrm{QR}} = \hat{\boldsymbol{Q}}_{\mathrm{QR}}(n|n-1)^H\hat{\boldsymbol{r}}_{k|m}(n|n-1)^\dagger$. Using the upper-triangular property of $\hat{\boldsymbol{R}}_{\mathrm{QR}}(n|n-1)$ and combining the M-algorithm, we can detect MIMO user data efficiently. The details of the QRD-M algorithm can be found in [26] and the references therein.

VI. Simulation Results

In this section, computer simulations are performed to confirm the performance of the proposed schemes. The simulated system has $N = 128$ subcarriers. Unless otherwise specified, information bits are mapped onto uncoded QPSK symbols through a Gray map and all users experience the same normalized Doppler shift of $F_dT_d = 0.001$. Furthermore, we set $\alpha^q_{k,\varepsilon} = 0.9999$ and $\eta^q_{k,\varepsilon} = 10^{-5}\ \forall k, q$ in (4). The SEKF is initialized with $\hat{\boldsymbol{f}}^q_k(1|0) = \boldsymbol{0}$ and $\boldsymbol{P}^q_k(1|0) = \boldsymbol{I},\ \forall k, q$ whereas the SK-APF is initialized with $\hat{\boldsymbol{f}}^q_{k,i}(1|0) = \boldsymbol{0}$ and $\boldsymbol{P}^q_{k,i}(1|0) = \boldsymbol{I},\ \forall k, i, q$. In the following, four examples will be presented. In the first two examples, we focus on the pilot-aided schemes to exploit training blocks only whereas the last examples investigate the semi-blind schemes.

Example 1: Pilot-Aided Estimation With Equal User Power

In this example, we consider a system with $M_t = M_r = 4$ and $K = 5$ users. Each user has equal user power and occupies $N_k = 12$ subcarriers. Figs. 1 and 2 show the channel estimation mean-squared error (MSE) and the average absolute (ABS) CFO estimation error over all users at E_s/N_0 of 20 dB, respectively. For the proposed pilot-aided SK-APF,

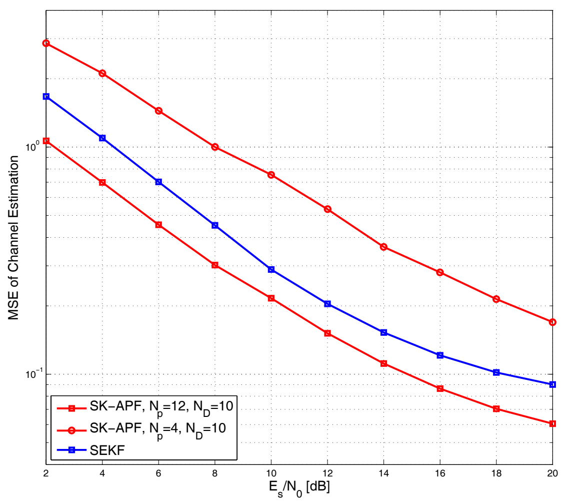

Fig. 1. Channel estimation performance of the proposed pilot-aided schemes as a function of E_s/N_0 in a system with equal-power users.

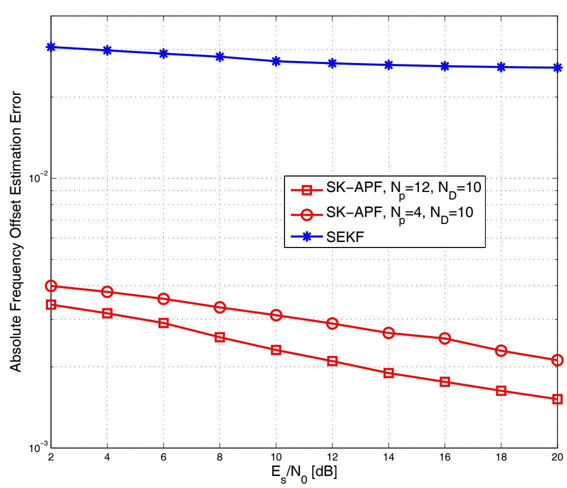

Fig. 2. CFO estimation performance of the proposed pilot-aided schemes as a function of E_s/N_0 in a system with equal-power users.

$N_p = 4, 12$ particles are employed. Furthermore, we use $2M + 1 = 21$ points to approximate the CFO in the first symbol interval. In the remaining symbol intervals, we use $2M+1 = 11$ points. Finally, the multipath intensity profile is $E\{|(\boldsymbol{h}_k^{p,q}|)_l|^2\} = \{0.6, 0.4\}\ \forall p, q, k$ with $l = 1, 2 = L_f$.

Inspection of Figs. 1 and 2 indicates that the SK-APF substantially outperforms the pilot-aided SEKF in terms of CFO estimation errors but has similar channel estimation performance as the SEKF. Furthermore, it is evident from Fig. 2 that the SK-APF provides more accurate CFO estimates as the number of particles employed, N_p, increases from 4 to 12.

Example 2: Pilot-Aided Estimation With Unequal User Power

Next, we consider the same system in Example 1, except that users have unequal signal powers due to the near-far effect with the first user $10 \log(K)$ dB stronger than the others. Figs. 3 and 4 show the channel estimation MSE and the

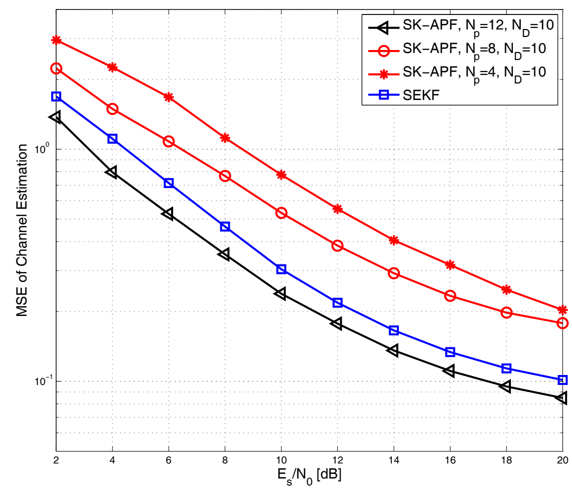

Fig. 3. Channel estimation performance of the proposed pilot-aided schemes for weaker users as a function of E_s/N_0 in a system with unequal-power users.

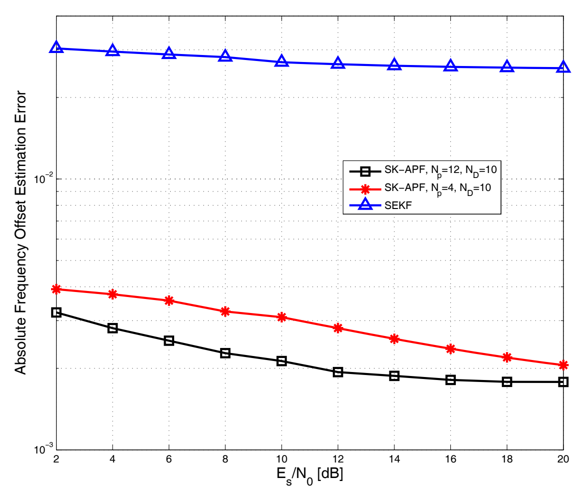

Fig. 4. CFO estimation performance of the proposed pilot-aided schemes for weaker users as a function of E_s/N_0 in a system with unequal-power users.

average ABS CFO estimation error over all weaker users (excluding the first user) at E_s/N_0 of 20 dB, respectively. Comparison of Figs. 3-4 and Figs.1-2 suggests that the near-far effect has only marginal impact on the estimation performance of the proposed schemes.

Example 3: Semi-Blind Estimation With Equal User Power

The semi-blind SK-APF is now employed with one training block followed by data blocks consisting of both pilots and data symbols. To approximate the CFO uncertainty region, we generate $2M + 1 = 101$ and 11 grid points for the first symbol interval and remaining symbol intervals, respectively. Furthermore, $M_t = M_r = 2$, $K = 2$ users and $N_k = 32$ subcarriers are employed in the simulation. The multipath intensity profile is $E\{|(\boldsymbol{h}_k^{p,q})_l|^2\} = \{0.5610,\ 0.2520,\ 0.1132,\ 0.0509,\ 0.0229\}$ for $\forall p, q, k$ with $l = 1, \ldots, 5 = L_f$.

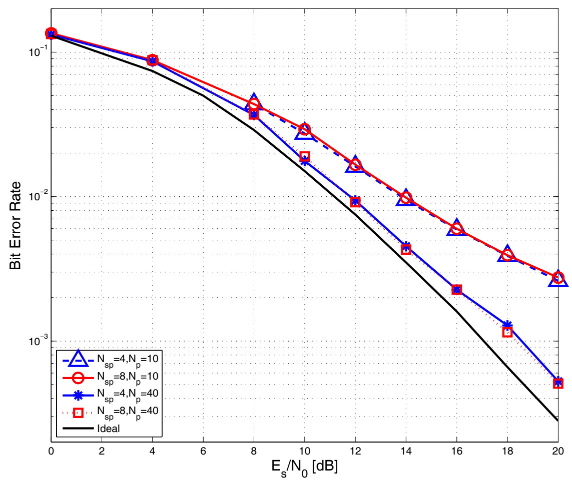

Fig. 5. Uncoded BER performance of the proposed semi-blind SK-APF as a function of E_s/N_0 with N_{sp} scattered pilots and N_p particles.

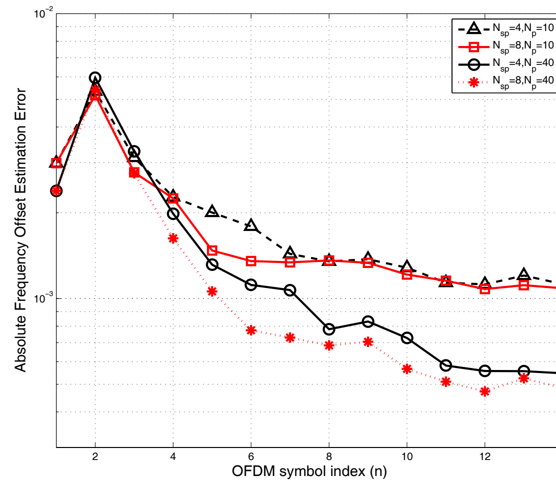

Fig. 6. CFO estimation performance of the proposed semi-blind SK-APF as a function of symbol index n at $E_s/N_0 = 20$ dB.

Fig. 5 shows the uncoded bit error rate (BER) performance of the proposed methods with different numbers of scattered pilots, $N_{sp} = 4, 8$ and particles, $N_p = 10, 40$. The curve labeled "**Ideal**" is obtained with perfect knowledge of channel and frequency offset. It is evident from Fig. 5 that more particles result in better BER performance for semi-blind SK-APF whereas more pilots lead to marginal performance improvement. Furthermore, inspection of Fig. 5 reveals that the semi-blind SK-APF with $N_{sp} = 8$ and $N_p = 40$ has about 1.5 dB loss w.r.t. the ideal curve at BER of 10^{-3}.

Figs. 6 and 7 show the channel and CFO estimation performance as a function of OFDM symbol index n at E_s/N_0 of 20dB, respectively. Figs. 6 and 7 suggest that increasing the number of particles from $N_p = 10$ to $N_p = 40$ can provide significantly more accurate CFO and channel estimation at the cost of higher computational complexity. Furthermore, inspection of Fig. 6 suggests that the maximum CFO estimation error occurs at $n = 2$. This is due to the fact that the first symbol is solely composed of pilots, and hence the initial CFO estimate is more accurate. On symbol $n = 2$, the estimator begins to diverge due to data errors, but with a sufficient number of measurements the CFO error begins decreasing on symbol $n = 3$.

Example 4 : Semi-Blind Estimation for Systems With More Users

In this last example, we investigate a system with $M_t = M_r = 4$ using the SK-APF. The same parameters are employed to generate frequency offset samples as shown in the previous example. Figs. 8 and 9 show the BER performance for $K = 2$ and $K = 4$, respectively. Clearly, more MAI is induced by the presence of more users in the system. As a result, the system performance degrades as K increases from 2 to 4, which is evidenced by comparing the curves in Figs. 8 and 9 obtained with $N_{sp} = 4$ scattered pilots and $N_p = 40$ particles. More specifically, Fig. 9 indicates that semi-blind SK-APF with $N_{sp} = 4$ and $N_p = 40$ entails 3.5 dB loss w.r.t. the ideal curve at BER of 10^{-4}. Furthermore, Fig. 9 shows

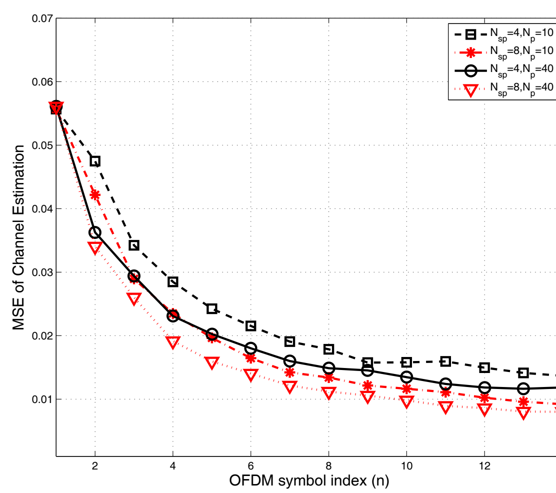

Fig. 7. Channel estimation performance of the proposed semi-blind SK-APF as a function of symbol index n. at $E_s/N_0 = 20$ dB.

that only marginal performance improvement can be obtained by increasing the number of particles beyond $N_p = 40$.

VII. Conclusion

Pilot-aided and semi-blind joint data detection and frequency offset/channel estimation schemes have been proposed for the uplink MIMO-OFDMA systems. The proposed schemes employ the parallel Schmidt Kalman filter to decompose the multiuser estimation problem into more tractable subproblems, each of which deals with only one desired user. Following the decomposition, the Schmidt Rao-Blackwellized particle filter is employed to track the time varying channel and CFO of the desired user in each subproblem. Simulation results have shown that the resulting scheme can provide accurate CFO and channel estimates at affordable computational complexity.

Throughout this work, pilot design has not been taken into account. However, as shown in [6], [30], [31], optimally designed pilots can substantially improve the system perfor-

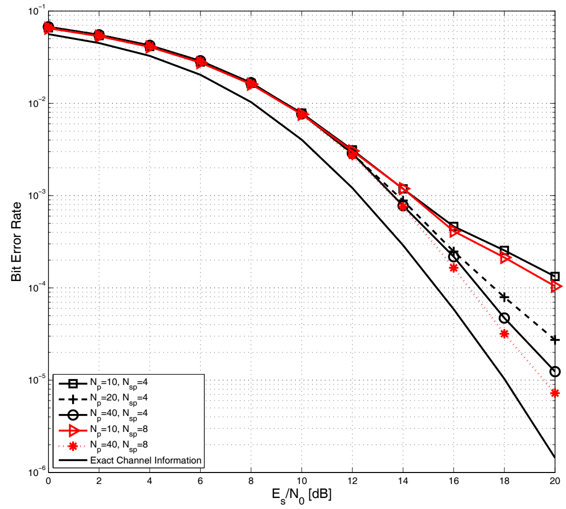

Fig. 8. Uncoded BER performance of the SK-APF as a function of E_s/N_0 with $M_t = M_r = 4$, $N_{sp} = \{4, 8\}$ and $K = 2$.

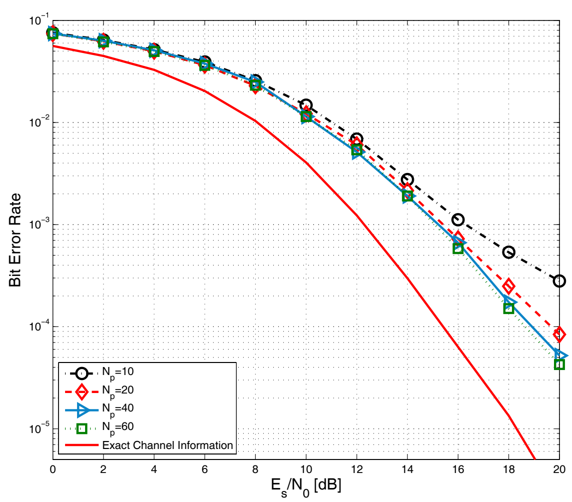

Fig. 9. Uncoded BER performance of the SK-APF as a function of E_s/N_0 with $M_t = M_r = 4$, $N_{sp} = 4$ and $K = 4$.

mance. It is anticipated that the performance of both SEKF and SK-APF estimators will improve with such optimized pilots and this should be investigated in actual system applications.

Appendix: Proof of Proposition 1

This Appendix summarizes the proof of Proposition 1. For simplicity, define $\lambda_1 \stackrel{\triangle}{=} \hat{\varepsilon}_k^q(n|n-1) + m' - m$. Recalling that users' symbols are assumed i.i.d., we have

$$\sigma^2_{n_{q,k,m}}(n) \stackrel{\triangle}{=} E\{|n^q_{k,m|\mathcal{I}_k}(n)|^2\} = \frac{1}{|\mu(\hat{\varepsilon}_k^q(n|n-1))|^2} \times E\left\{ \sum_{k'=1}^{K} \sum_{\substack{m'=0,\\ m'\neq m}}^{N-1} |\mu(\lambda_1)|^2 [\tilde{d}^1_{k',m'}(n), \ldots, \tilde{d}^{M_t}_{k',m'}(n)] \boldsymbol{a}_{k'} \times \boldsymbol{a}^H_{k'} [\tilde{d}^1_{k',m'}(n), \ldots, \tilde{d}^{M_t}_{k',m'}(n)]^H \right\} + 2N_0/T_s, \tag{A.1}$$

where

$$\boldsymbol{a}_{k'} \stackrel{\triangle}{=} \begin{bmatrix} (\boldsymbol{W}_{L_f})_{m'} \hat{\boldsymbol{h}}^{1,q}_{k'}(n|n-1) \\ \vdots \\ (\boldsymbol{W}_{L_f})_{m'} \hat{\boldsymbol{h}}^{M_t,q}_{k'}(n|n-1) \end{bmatrix} = (\boldsymbol{I}_{M_t} \otimes (\boldsymbol{W}_{L_f})_{m'}) \hat{\boldsymbol{h}}^q_{k'}(n|n-1). \tag{A.2}$$

Using (A.2), we have (A.3), see next page, where we have exploited the following equalities

$$E\{||\tilde{\boldsymbol{d}}^p_{k,m|\mathcal{I}_k}(n)||^2\} = 1/M_t, \tag{A.4}$$
$$(\boldsymbol{A} \otimes \boldsymbol{B})^H (\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{A}^H \boldsymbol{C}) \otimes (\boldsymbol{B}^H \boldsymbol{D}). \tag{A.5}$$

References

[1] P. W. Chan, E. S. Lo, V. K. N. Lau, R. S. Chen, K. B. Letaief, R. D. Murch, and W. H. Mow, "Performance comparison of downlink multiuser MIMO-OFDMA and MIMO-MC-CDMA with transmit side information—multi-cell analysis," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2193–2203, June 2007.

[2] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, pp. 1747–1758, Oct. 1999.

[3] M. Morelli, C.-C. J. Kuo, and M. Pun, "Synchronization techniques for OFDMA: a tutorial review," *Proc. IEEE*, vol. 95, pp. 1394–1427, July 2007.

[4] M. O. Pun, M. Morelli, and C.-C. J. Kuo, "Maximum likelihood synchronization and channel estimation for OFDMA uplink transmissions," *IEEE Trans. Commun.*, vol. 54, pp. 726–736, Apr. 2006.

[5] Z. Cao, U. Tureli, and Y. Yao, "Deterministic multiuser carrier-frequency offset estimation for interleaved OFDMA uplink," *IEEE Trans. Commun.*, vol. 52, pp. 1585–1594, Sep. 2004.

[6] S. Sezginer, P. Bianchi, and W. Hachem, "Asymptotic Cramer-Rao bounds and training design for uplink MIMO-OFDMA systems with frequency offsets," *IEEE Trans. Signal Process.*, vol. 55, pp. 3606–3622, July 2007.

[7] A. Saemi, J. Cances, and V. Meghdadi, "Synchronization algorithms for MIMO OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 4441–4451, Dec. 2007.

[8] D. Huang and K. B. Letaief, "An interference-cancellation scheme for carrier frequency offsets correction in OFDMA systems," *IEEE Trans. Commun.*, vol. 53, pp. 1155–1165, July 2005.

[9] T. Roman, M. Enescu, and V. Koivunen, "Joint time-domain tracking of channel and frequency offsets for MIMO OFDM systems," *Wireless Personal Commun.*, vol. 31, pp. 181–200, Dec. 2004.

[10] P. Zhao, L. Kuang, and J. Lu, "Carrier frequency offset estimation using extended Kalman Filter in uplink OFDMA systems," in *Proc. IEEE International Conf. Commun.*, May 2006, pp. 2870–2874.

[11] T. Nyblom, T. Roman, M. Enescu, and V. Koivunen, "Time-varying carrier offset tracking in OFDM systems using particle filtering," in *Proc. 4th IEEE International Symposium Signal Process. Information Technol.*, Dec. 2004, pp. 217–220.

[12] R. A. Iltis, "A sequential Monte Carlo filter for joint linear/nonlinear state estimation with application to DS-CDMA," *IEEE Trans. Signal Process.*, vol. 51, pp. 417–426, Feb. 2003.

[13] K. J. Kim and R. A. Iltis, "Frequency offset synchronization and channel estimation for the MIMO-OFDM system using Rao-Blackwellized Gauss-Hermite Filter," in *Proc. Asilomar Conf. Signals Systems Computers*, Oct. 2005, pp. 829–833.

[14] M. O. Pun, M. Morelli, and C.-C. J. Kuo, "Iterative detection and frequency synchronization for generalized OFDMA uplink transmissions," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 629–639, Feb. 2007.

[15] M. S. Grewal and A. P. Andrews, *Kalman Filtering Theory and Practice using Matlab*. New York: John Wiley and Sons, 2001.

[16] K. Raghunath and A. Chockalingam, "SIR analysis and interference cancellation in uplink OFDMA with large carrier frequency and timing offsets," in *Proc. IEEE Wireless Commun. Networking Conf.*, Mar. 2007, pp. 997–1002.

[17] Z. Zhang and C. Tellambura, "The effect of imperfect carrier frequency offset estimation on OFDMA uplink transmission," in *Proc. IEEE International Conf. Commun.*, May 2007, pp. 6281–6286.

[18] T. Kang and R. A. Iltis, "Iterative decoding, offset and channel estimation for OFDM using the unscented Kalman Filter," in *Proc. Asilomar Conf. Signals Systems Computers*, Nov. 2007, pp. 1728–1732.

TABLE I
SUMMARY OF SYMBOLS FOR USER k AND RECEIVING ANTENNA q

Symbol	Definition
p,q	indices for transmit and receive antennas, respectively
$\mathcal{I}_k, \mathcal{I}_{k,d}, \mathcal{I}_{k,p}$	indices for subcarrier, data symbols and pilot symbols, respectively
$\boldsymbol{W}$	DFT matrix
$\boldsymbol{h}_k^{p,q}(n)$	complex-valued channel vectors between transmit antenna p and receive antenna q
$\varepsilon(n) \triangleq \varepsilon_k^q(n)$	normalized frequency offset
$\boldsymbol{\Delta}(\varepsilon_k^q(n))$	frequency offset matrix for one OFDM symbol
$\boldsymbol{d}_k^p(n)$, $\tilde{\boldsymbol{d}}_k^p(n)$	time- and frequency-domain data symbols from antenna p, respectively
$\boldsymbol{D}_k^p(n)$, $\tilde{\boldsymbol{D}}_{\varepsilon,k}^p(n)$	circulant data matrix without and with frequency offset, respectively
$\alpha_{k,\varepsilon}^q, \alpha_{k,h}^q$	first-order dynamic model coefficients
$\mathbb{H}(\varepsilon(n)) \triangleq \mathbb{H}_k^q(\varepsilon_k^q(n)), \mathbb{H}_{\setminus k}(\varepsilon_{\setminus k}(n))$	real-valued equivalent channel matrix, interfering matrix
$\boldsymbol{f}(n), \boldsymbol{f}_{\setminus k}(n)$	real-valued channel vector, interfering channel vector, respectively
$\boldsymbol{x}(n), \boldsymbol{x}_{\setminus k}(n)$	essential state vector, interfering nuisance state vector
$\boldsymbol{A}, \boldsymbol{A}_{\setminus k}$	Kalman dynamic matrix, interfering Kalman dynamic matrix
$\boldsymbol{Q}, \boldsymbol{Q}_{\setminus k}$	Covariance matrix for dynamic equation, interfering covariance matrix
$\boldsymbol{J}(n), \boldsymbol{J}_{\setminus k}(n)$	Jacobian matrix, interfering Jacobian matrix
$\boldsymbol{P}_{k,k}(n\|n-1), \boldsymbol{P}_{\setminus k,k}(n\|n-1), \boldsymbol{P}_{\setminus k,\setminus k}(n\|n-1)$	Kalman covariance matrices
$\boldsymbol{K}(n), \boldsymbol{K}_{k,\text{SKF}}(n)$	Kalman gain matrix, Schmidt-Kalman gain matrix
$\hat{\boldsymbol{x}}(n\|n-1), \hat{\boldsymbol{x}}(n\|n)$	Kalman state prediction, Kalman state estimation
ϵ_i^n	i-th trajectory of the frequency offset particles
$\boldsymbol{y}^n$	cumulative observation sequences
$\pi(.), \beta_i(n)$	sampling density, i-th importance sampling weight
δ_m	m-th grid point
$\hat{\boldsymbol{f}}_i(n-1\|n-1)$	Kalman channel estimation conditioned on ϵ_i^{n-1}
$\boldsymbol{P}_{k,k,i}(n-1\|n-1), \boldsymbol{P}_{k,k,(i,m)}(n-1\|n-1)$	Kalman covariance matrices conditioned on ϵ_i^{n-1}
$\boldsymbol{K}_i(n)$	Schmidt-Kalman gain
$\pi_i(\cdot)$	sampling density

$$
\begin{aligned}
&M_t|\mu(\hat{\varepsilon}_k^q(n|n-1))|^2\sigma^2_{n_{q,k,m}}(n) - \frac{2M_tN_0}{T_s} = \sum_{k'=1}^{K}\sum_{\substack{m'=0,\\ m'\neq m}}^{N-1}|\mu(\lambda_1)|^2\times \\
&\text{Tr}\left(\left(\boldsymbol{I}_{M_t}\otimes(\boldsymbol{W}_{L_f})_{m'}\right)\hat{\boldsymbol{h}}_{k'}^q(n|n-1)\hat{\boldsymbol{h}}_{k'}^q(n|n-1)^H\left(\boldsymbol{I}_{M_t}\otimes(\boldsymbol{W}_{L_f})_{m'}\right)^H\right), \\
&= \sum_{k'=1}^{K}\sum_{\substack{m'=0,\\ m'\neq m}}^{N-1}|\mu(\lambda_1)|^2\times\text{Tr}\Big([\boldsymbol{I}_{M_t}\otimes(\boldsymbol{X}_W)_{m'}]\{\hat{\boldsymbol{h}}_{k'}^q(n|n-1)\hat{\boldsymbol{h}}_{k'}^q(n|n-1)^H\}\Big),
\end{aligned}
\tag{A.3}
$$

[19] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, pp. 1065–1076, May 2002.

[20] M. A. Vázquez, M. F. Bugallo, and J. Míguez, "Sequential Monte Carlo methods for complexity-constrained MAP equalization of dispersive MIMO channels," *Signal Process.*, vol. 88, pp. 1017–1034, 2008.

[21] G. A. Watson, D. H. McCabe, and T. R. Rice, "Multisensor-multisite composite tracking in the presence of sensor residual bias," in *Proc. SPIE Conf. Signal Data Processing Small Targets*, July 1999, pp. 370–381.

[22] N. de Freitas, "Rao-Blackwellised particle filtering for fault diagnosis," in *Proc. IEEE Aerospace Conf.*, Mar. 2002.

[23] A. Doucet, S. J. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, pp. 197–208, 2000.

[24] C. Andrieu, M. Davy, and A. Doucet, "Efficient particle filtering for jump Markov systems: application to time-varying autoregressions," *IEEE Trans. Signal Process.*, vol. 41, pp. 1762–1770, July 2003.

[25] A. Doucet, N. Gordon, and V. Krishnamurthy, "Particle filter for state estimation of jump Markov linear system," *IEEE Trans. Signal Process.*, vol. 49, pp. 613–624, Mar. 2001.

[26] K. J. Kim, Y. Yue, R. A. Iltis, and J. D. Gibson, "A QRD-M/Kalman Filter-based detection and channel estimation algorithm for MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 710–721, Mar. 2005.

[27] H. Kawai, K. Higuchi, N. Maeda, and M. Sawahashi, "Adaptive control of surviving symbol replica candidates in QRM-MLD for OFDM MIMO multiplexing," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 1130– 1140, June 2006.

[28] F. Yang, K. H. Li, and K. C. Teh, "Performance analysis of the blind minimum output variance estimator for carrier frequency offset in OFDM systems," *EURASIP J. Applied Signal Process.*, vol. 2006, article ID 49 257, pp. 1–8, 2006.

[29] C. Tsai, R. A. Iltis, and R. E. Cagley, "Joint interference suppression and QRD-M detection for spatial multiplexing MIMO systems in a Rayleigh fading channel," in *Proc. International Telemetering Conf.*, Oct. 2006.

[30] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, pp. 1615–1624, June 2003.

[31] H. Minn, N. Al-Dhahir, and Y. Li, "Optimal training signals for MIMO OFDM channel estimation in the presence of frequency offset and phase noise," *IEEE Trans. Commun.*, vol. 54, pp. 1754–1759, Oct. 2006.

Kyeong Jin Kim received the M.S. degree from the Korea Advanced Institute of Science and Technology (KAIST) in 1991 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California, Santa Barbara in 2000. During 1991-1995, he was a research engineer at the video research center of Daewoo Electronics, Ltd., in Korea. In 1997, he joined the Data Transmission and Networking Laboratory at the University of California, Santa Barbara. After receiving his degrees, he joined the Nokia Research Center in Dallas as a senior research engineer. From 2005 to 2009, he worked with the Nokia Corporation in Dallas as a L1 specialist. His research has focused on transceiver design, resource management, and scheduling in the wireless communication systems.

Man-On Pun (M'06) received the BEng. (Hon.) degree in electronic engineering from the Chinese University of Hong Kong in 1996, the MEng. degree in Computer Science from the University of Tsukuba, Japan in 1999, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 2006, respectively. He is a research scientist at Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA. He held research positions at Princeton University, Princeton, NJ from 2006 to 2008 and Sony Corporation in Tokyo, Japan, from 1999 to 2001. Dr. Pun's research interests are in the areas of statistical signal processing and wireless communications. Among his publications in these areas is the recent book *Multi-Carrier Techniques for Broadband Wireless Communications: A Signal Processing Perspective* (Imperial College Press, 2008).

Dr. Pun received the best paper award - runner-up from the IEEE Conference on Computer Communications (Infocom), Rio de Janeiro, Brazil in 2009; the best paper award from the IEEE International Conference on Communications, Beijing, China in 2008; and the IEEE Vehicular Technology Fall Conference, Montreal, Canada in 2006. He is a recipient of several scholarships including the Japanese Government (Monbusho) Scholarship, the Sir Edward Youde Memorial fellowship for Overseas Studies and the Croucher postdoctoral fellowship.

Ronald A. Iltis received the B.A. (biophysics) from The Johns Hopkins University in 1978, the M.Sc in engineering from Brown University in 1980, and the Ph.D in electrical engineering from the University of California, San Diego in 1984. Since 1984, he has been with the University of California, Santa Barbara, where he is currently a Professor in the Department of Electrical and Computer Engineering. His current research interests are in OFDM, underwater acoustic communications, MIMO radar, and nonlinear estimation. He has also served as a consultant to government and private industry in the areas of adaptive arrays, neural networks, and spread-spectrum communications.