

Flexible polarization demultiplexing method based on an adaptive process noise covariance Kalman filter

Jun Ge (葛军), Lianshan Yan (闫连山)*, Anlin Yi (易安林), Yan Pan (盘艳),
Lin Jiang (蒋林), Liangliang Dai (代亮亮), Wei Pan (潘炜), and Bin Luo (罗斌)

Center for Information Photonics & Communications, School of Information Science and Technology,
Southwest Jiaotong University, Chengdu 610031, China

*Corresponding author: lsyang@home.swjtu.edu.cn

Received January 8, 2018; accepted April 16, 2018; posted online May 25, 2018

A flexible polarization demultiplexing method based on an adaptive Kalman filter (AKF) is proposed in which the process noise covariance has been estimated adaptively. The proposed method may significantly improve the adaptive capability of an extended Kalman filter (EKF) by adaptively estimating the unknown process noise covariance. Compared to the conventional EKF, the proposed method can avoid the tedious and time consuming parameter-by-parameter tuning operations. The effectiveness of this method is confirmed experimentally in 128 Gb/s 16QAM polarization-division-multiplexing (PDM) coherent optical transmission systems. The results illustrate that our proposed AKF has a better tracking accuracy and a faster convergence (about 4 times quicker) compared to a conventional algorithm with optimal process noise covariance.

OCIS codes: 060.0660, 060.1660, 060.2330.

doi: 10.3788/COL201816.060601.

Increasing the spectral efficiency through the use of higher-order modulation formats in combination with polarization-division-multiplexing (PDM) has historically been one of the most effective methods to realize higher channel capacities at a comparable or lower cost. However, how to realize demultiplexing for high-order modulation formats (e.g., 16QAM) is the key point in current research. For such a problem, a variety of polarization demultiplexing (PolDemux) techniques have been presented, such as constant modulus algorithm (CMA), multimodulus algorithm (MMA), and their variants^[1-5] that have the disadvantages of complexity and the singularity problem. Therefore, several alternative PolDemux algorithms based on Stokes space (SS)^[6-8] and Kalman filter (KF)^[9-12] are proposed to solve the above problems. SS-based algorithms are transparent to modulation formats. However, the SS-based algorithm with plane fitting will degrade the system performance under a time-varying state of polarization, which can be improved by an adaptive SS algorithm^[13]. In Ref. [14], Muga *et al.* investigated and compared the convergence ratio, tracking, computational complexity, and system performance of an extended Kalman filter to an SS-based polarization demultiplexing method with a geometrical approach previously proposed for adaptive computation of the best fit plane.

In all the KF-based schemes mentioned above, the measurement and the process noise covariance play very important roles in the estimation accuracy. The measurement noise covariance can be directly inferred from the available measurement devices while the process noise covariance needs to be assessed, as a function of time, to optimize the KF performances^[15,16]. In general, it is difficult to determine the effect of multiple changes to the tuning parameters of process noise covariance, so each

parameter will be tuned individually while the others hold constant. This parameter-by-parameter tuning method requires a very large number of manual trials. When the dynamics is nonlinear and the number of parameters is more than two or three, the trial-and-error technique can become quite tedious and time consuming^[16]. Therefore, it would be highly desirable to either pre-know the process noise covariance or adopt a blind strategy to ensure the effectiveness of the KF algorithm.

In this Letter, an effective PolDemux method based on an adaptive process noise covariance Kalman filter^[15] is proposed. We call this approach AKF. By using AKF, the tedious parameter-by-parameter tuning can be avoided. The performance of proposed scheme is experimentally investigated in 128 Gb/s PDM-16QAM systems. The results show that the proposed method may significantly improve the adaptive capability of EKF by adaptively estimating the process noise covariance. Meanwhile, such an approach exhibits a faster and more accurate tracking performance compared to conventional algorithms with optimal process noise covariance.

The block diagram of our proposed AKF is shown in Fig. 1. The independently modulated dual-polarization signals $\mathbf{Z}_{in} = [e_x, e_y]^T$ can be represented in Stokes space using the relations^[6]

$$\begin{cases} s_0 = (e_x e_x^* + e_y e_y^*)/2 \\ s_1 = (e_x e_x^* - e_y e_y^*)/2 \\ s_2 = (e_x^* e_y + e_x e_y^*)/2 \\ s_3 = (-j e_x^* e_y + j e_x e_y^*)/2 \end{cases}, \quad (1)$$

where e_x and e_y are the complex optical-field amplitudes of the two orthogonal polarization states. The first Stokes component s_0 represents the total power; the remaining three components (s_1, s_2, s_3) represent 0° linearly, 45° linearly,

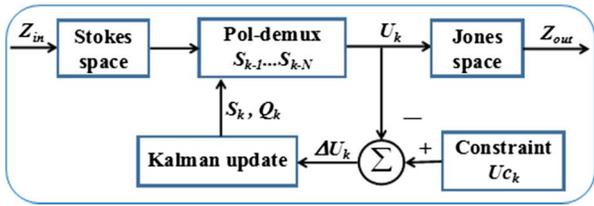


Fig. 1. Block diagram of the proposed PolDemux method based on an AKF.

and circularly polarized light, respectively, and * denotes the complex conjugation operator. At the receiver, the inverse polarization rotation matrix is represented by

$$\mathbf{F} = \begin{bmatrix} \cos(p) \exp(iq/2) & \sin(p) \exp(-iq/2) \\ -\sin(p) \exp(iq/2) & \cos(p) \exp(-iq/2) \end{bmatrix}, \quad (2)$$

where $p = 1/2a \tan(a, \sqrt{b^2 + c^2})$, $q = a \tan(b, c)$, and $\mathbf{n} = (a, b, c)^T$. Applying Eq. (2) to the signal samples, the normal vector \mathbf{n} will be forced to $(1, 0, 0)^T$. An arbitrary plane in Stokes space can be defined as

$$as_1 + bs_2 + cs_3 + d = 0, \quad (3)$$

where a, b , and c are the components of normal vector \mathbf{n} , and d represents the distance from the plane to the origin. However, the parameters are not unique as Eq. (3) can be multiplied by an arbitrary constant and still represent the same plane. So, Eq. (4) is used to enforce the representation of direction cosines by a, b , and c ^[4],

$$a^2 + b^2 + c^2 = 1. \quad (4)$$

The state vector is defined as $\mathbf{S}_k = (a_k, b_k, c_k, d_k)^T$,

$$\mathbf{U}_k = \begin{bmatrix} a_k S_{1,k} + b_k S_{2,k} + c_k S_{3,k} + d_k \\ a_k^2 + b_k^2 + c_k^2 \end{bmatrix}, \quad (5)$$

$$\mathbf{H}_k = \frac{\partial \mathbf{U}_k}{\partial \mathbf{S}_k} = \begin{bmatrix} S_{1,k} & S_{2,k} & S_{3,k} & 1 \\ 2a_k & 2b_k & 2c_k & 0 \end{bmatrix}, \quad (6)$$

$$\mathbf{S}_k = \mathbf{S}_{k-1} + \boldsymbol{\omega}_{k-1}, \quad (7)$$

$$\mathbf{S}_k^- = \mathbf{S}_{k-1}, \quad (8)$$

$$\mathbf{P}_k^- = \mathbf{P}_{k-1} + \mathbf{Q}_k, \quad (9)$$

where \mathbf{U}_k is the measurement vector, and \mathbf{H}_k is the Jacobian matrix formed by the partial derivatives of \mathbf{U}_k with respect to the state variables. The process equation is presented in Eq. (7). \mathbf{S}_k^- is the prediction state vector, and $\boldsymbol{\omega}_k$ is the process noise that is assumed as a white Gaussian sequence. The Kalman prediction is presented in Eqs. (8) and (9). \mathbf{P}_k^- and \mathbf{P}_k are called priori estimate error covariance and posteriori estimate error covariance^[10]. \mathbf{Q}_k is the process noise covariance matrix, which has a great influence on the convergence speed and estimate accuracy

of Kalman filter. In general, \mathbf{Q}_k might be affected by many factors, such as the transmission link and polarization rotation frequency^[12]. Therefore, \mathbf{Q}_k might be different in different transmission systems. In the conventional KF method, \mathbf{Q}_k is simplified to be a diagonal matrix, and the diagonal elements (represented by Q) are selected as an identical value. And, improper values of Q will degrade the filter performance^[16,17].

In this Letter, we propose an AKF method to replace the tedious trial-and-error operations. The forecast of \mathbf{Q}_k is performed by using the last N estimated states at time step k . In Eq. (7), each element of $\boldsymbol{\omega}_k$ represents the process noise of each element in state vector \mathbf{S}_k , respectively. In general, the elements of $\boldsymbol{\omega}_k$ can be different from each other. Accordingly, the diagonal elements in \mathbf{Q}_k are not all equal, assuming that all the elements in \mathbf{S}_k are independent of each other. The procedure of calculating \mathbf{Q}_k is shown below.

1) At time step k , the last N state vectors are calculated to form a matrix

$$(\mathbf{S}_{k-N-1}, \mathbf{S}_{k-N-2}, \dots, \mathbf{S}_k) = \begin{pmatrix} a_{k-N-1} & a_{k-N-2} & \dots & a_k \\ b_{k-N-1} & b_{k-N-2} & \dots & b_k \\ c_{k-N-1} & c_{k-N-2} & \dots & c_k \\ d_{k-N-1} & d_{k-N-2} & \dots & d_k \end{pmatrix}. \quad (10)$$

2) Calculating the variance $\sigma_a^2, \sigma_b^2, \sigma_c^2, \sigma_d^2$ of each row in Eq. (10), then \mathbf{Q}_k is written as

$$\mathbf{Q}_k = \text{diag}(\sigma_a^2, \sigma_b^2, \sigma_c^2, \sigma_d^2), \quad (11)$$

where diag is a diagonal matrix, and the Kalman updating process is presented in Eqs. (12)–(15)^[17],

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T (\mathbf{H} \mathbf{P}_k^- \mathbf{H}^T + \mathbf{R})^{-1}, \quad (12)$$

$$\Delta \mathbf{U}_k = \mathbf{U}_k c_k - \mathbf{U}_k, \quad (13)$$

$$\mathbf{S}_k = \mathbf{S}_k^- + \mathbf{K}_k \Delta \mathbf{U}_k, \quad (14)$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_k^-, \quad (15)$$

where \mathbf{K}_k , $\Delta \mathbf{U}_k$, and $\mathbf{U}_k c_k$ are the Kalman gain, residual, and actual measurement. $\mathbf{U}_k c_k$ is equal to $(0, 1)^T$, and \mathbf{R} is the so-called measurement noise covariance, which can be regarded as a constant diagonal matrix where the diagonal values are set to 0.1. Therefore, \mathbf{Q}_k is used to optimize the filter performance.

The performance of the proposed AKF is experimentally investigated in 128 Gb/s PDM-16QAM coherent optical communication systems as shown in Fig. 2. At the transmitter, an external cavity laser (ECL) oscillating at 1550 nm with ~ 100 kHz linewidth is used as a transmitter laser source. The arbitrary waveform generator (AWG) provides 16×10^9 symbol/s 4-level electrical signals to drive an optical I/Q modulator for obtaining the 16QAM signal. The encoded signals are polarization

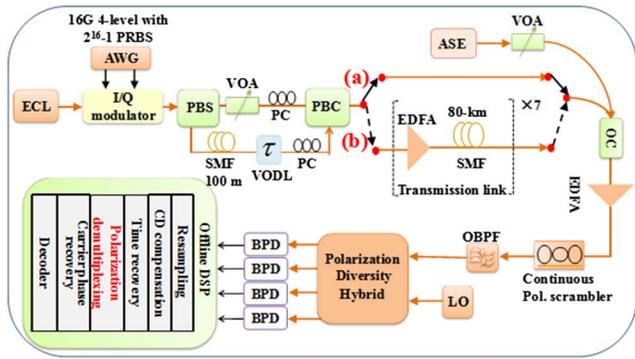


Fig. 2. Experimental setup and DSP flow of PDM-16QAM systems. (a) The B-to-B case; (b) the 560 km transmission link. ECL: external cavity laser; AWG: arbitrary waveform generator; PC: polarization controller; PBS: polarizing beam splitter; VODL: variable optical delay line; PBC: polarizing beam combiner; EDFA: erbium-doped fiber amplifier; ASE: amplified spontaneous emission; VOA: variable optical attenuator; SMF: single mode fiber; OC: 3 dB optical coupler; OBPF: optical bandpass filter; LO: local oscillator; BPD: balanced photodetector.

multiplexed to generate 128 Gb/s PDM-16QAM signals employing an interleave scheme that is composed of a coupler, two polarization controllers (PCs), a variable optical attenuator (VOA), 100 m single mode fiber (SMF), a variable optical delay line (VODL), and a polarization beam combiner (PBC). Here, two PCs, 100 m SMF, the VODL, and the VOA are used to generate two data streams with orthogonal states of polarization (SOPs), decorrelate two data streams, synchronize the bit sequences of two data streams, and balance the optical power between two arms, respectively. The signals are transmitted in the B-to-B case and 560 km SMF transmission case. The VOA and erbium-doped fiber amplifier (EDFA) are deployed to adjust the optical signal to noise ratio (OSNR) of the received signals. At the receiver, the polarization state of the input PDM-16QAM signal is continuously scrambled by Agilent N7788B. Then, the signal is filtered using an optical bandpass filter (OBPF) with 0.8 nm bandwidth centered at the set transmission wavelength. Subsequently, the signal is coherent detected, and then sampled by a real-time oscilloscope for offline processing. After time recovery, the samples are down-sampled to 1 sample per symbol for the PolDemux technique. Either the conventional EKF or our proposed AKF algorithm is used in the PolDemux stage. The bit error rate (BER) is calculated from 2^{16} symbols.

First, the BER performance and the required convergence symbol as a function of the process noise Q is investigated. The polarization rotation frequency is 2474.004 rad/s, achieved by a continuous polarization scrambler. As shown in Fig. 3(a), for a conventional EKF in Ref. [14], the required convergence symbol decreases as the Q increases. The convergence would be reached within 300 symbols in the case of $Q = 10^{-3}$, while more than 30000 symbols are required for convergence in

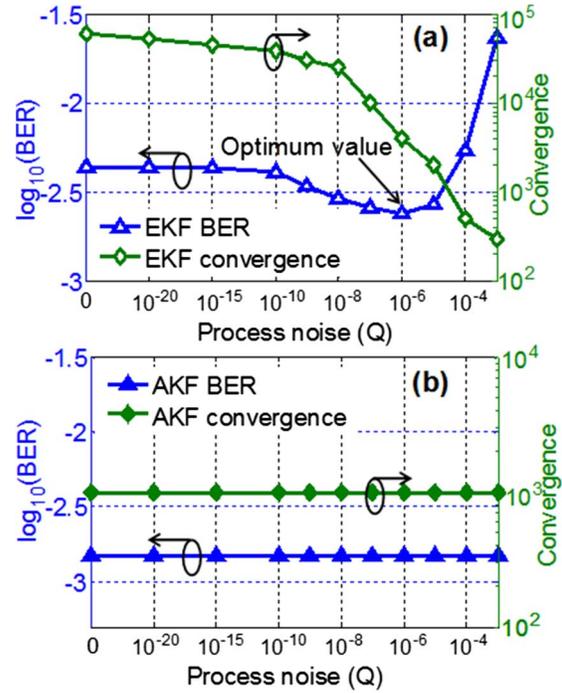


Fig. 3. BER performance and required symbol for convergence vs. process noise (Q) in (a) a conventional EKF and (b) the proposed AKF for a PDM-16QAM signal in the B-to-B configuration, for OSNR = 20 dB.

the case of $Q = 10^{-10}$. We can also observe that the optimal BER performance is obtained when $Q = 10^{-6}$, and 4000 symbols are required for convergence. Therefore, there exists a trade-off between the convergence speed and tracking accuracy for a conventional EKF. Subsequently, the BER and convergence performance for the proposed AKF is investigated in Fig. 3(b). The block length N is set to 8. We can see that by using the AKF, similar or better BER performance is achieved regardless of the initial process noise. Meanwhile, the required symbol for convergence is less than 1000 when Q is in the range of $[0, 10^{-3}]$. Compared to the optimum value ($Q = 10^{-6}$) for the EKF, our method is approximately four times quicker to reach convergence. In the following experiments, $Q = 10^{-6}$ is chosen as the optimum value for EKF.

Then, robustness of the proposed AKF against the polarization rotation frequency is discussed, as shown in Fig. 4. The endless polarization rotation is digitally achieved by a polarization rotation matrix $\mathbf{J} = [\cos(kwT_s) \sin(kwT_s), -\sin(kwT_s) \cos(kwT_s)]$ in the digital signal processing (DSP) section before the PolDemux^[11,13] technique. k is the sample number, w represents the polarization rotation frequency, and T_s is the symbol duration. Here, we measure the BER performance of PDM-16QAM with different polarization rotation speeds under 20 dB OSNR. As seen from Fig. 4, no performance penalty is observed when the rotation is less than 7×10^4 rad/s for the proposed AKF and conventional EKF. In addition, the AKF method shows a better BER performance than the EKF method, which can be attributed to the adaptive estimation of

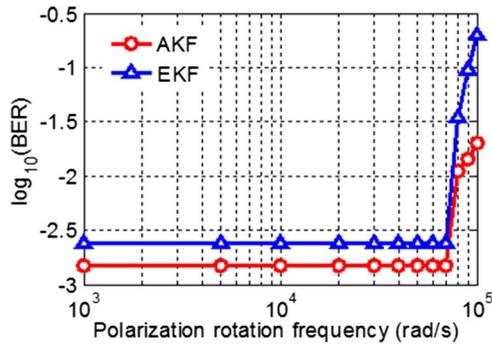


Fig. 4. BER performance vs. the polarization rotation frequency for the PDM-16QAM signal in a B-to-B configuration, for OSNR = 20 dB.

the process noise covariance. The performance of both approaches decreases at the same polarization rotation frequency (7×10^4 rad/s) just because the KF is unable to track such a fast polarization rotation in Stokes space. Nevertheless, the tolerance of the polarization rotation is quite adequate for current commercial optical communication systems. Notably, the maximum polarization rotation frequency in commercial systems is 10^3 rad/s^[18].

Subsequently, we also investigate the convergence performance by calculating the mean square error (MSE) of the modulus of the demultiplexed signal. The MSE is defined in Eq. (16),

$$\text{MSE} = \frac{\sum_{i=1}^n (R_{\text{obs},i} - R_{\text{ref},i})^2}{n}, \quad (16)$$

where $R_{\text{obs},i}$ represents the module of the i th actual received symbol, and $R_{\text{ref},i}$ represents the ideal module of the i th received symbol. The module of the ideal 16QAM signal forms three circles, therefore, $R_{\text{ref},i} \in \{\sqrt{P_{\text{ave}}/5}, \sqrt{P_{\text{ave}}}, \sqrt{9P_{\text{ave}}/5}\}$, where P_{ave} is the average power of the signal. Figure 5(a) shows the MSE performance of the conventional EKF and our proposed AKF for the PDM-16QAM signal in the B-to-B case. For conventional EKF, a smaller process noise (e.g., $Q = 10^{-9}$ and 10^{-6}) leads to a slower convergence. However, a larger process noise (e.g., $Q = 10^{-3}$) leads to a larger MSE. Our proposed AKF algorithm can clearly reach a faster convergence speed and a smaller steady-state error under arbitrary initial process noise covariance. This improvement is attributed to the process noise estimation according to the estimated states. With the faster convergence and better estimation accuracy, the AKF is more suitable for a burst-mode coherent receiver. Figure 5(b) depicts the BER curves as a function of OSNR for the EKF and AKF algorithms in the B-to-B configuration. Because of its faster convergence speed and fine estimation accuracy, the AKF obtains ~ 0.5 dB and ~ 2.2 dB OSNR gains at the 7% FEC (i.e., BER = 0.0044) threshold, compared to $Q = 10^{-6}$ (optimal value) and $Q = 10^{-4}$ for the EKF, respectively.

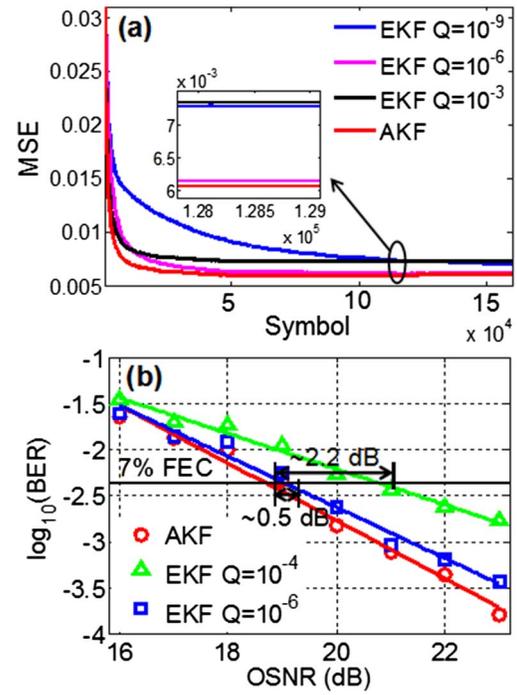


Fig. 5 Comparison of (a) the MSE performance of the modulus of the demultiplexed signal, for OSNR = 20 dB; (b) the B-to-B BER performance.

Figure 6 illustrates the BER performance after transmission over the 560 km SMF. As shown in Fig. 6, the OSNR penalty is only about 2.7 dB, compared to the B-to-B at the 7% FEC threshold. Then, we evaluate the performance of the proposed AKF method with polarization scrambling rates of 1237.002 and 2474.004 rad/s, as shown in Figs. 7(a) and 7(b). The OSNRs are set to be 20 dB for the B-to-B and 23 dB for the 560 km transmission, respectively. It shows that the BER fluctuations are less than 0.3 dB and 0.2 dB under the B-to-B and 560 km transmission, respectively. The results indicate that our proposed method can present good polarization tracking stability.

The computational complexity of two algorithms is shown in Table 1. Compared to the conventional EKF

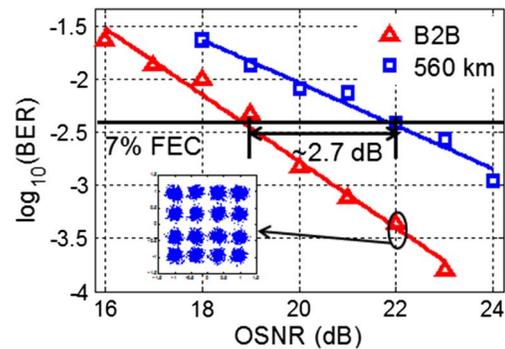


Fig. 6. Comparison of the B-to-B and 560 km transmission performance for the AKF.

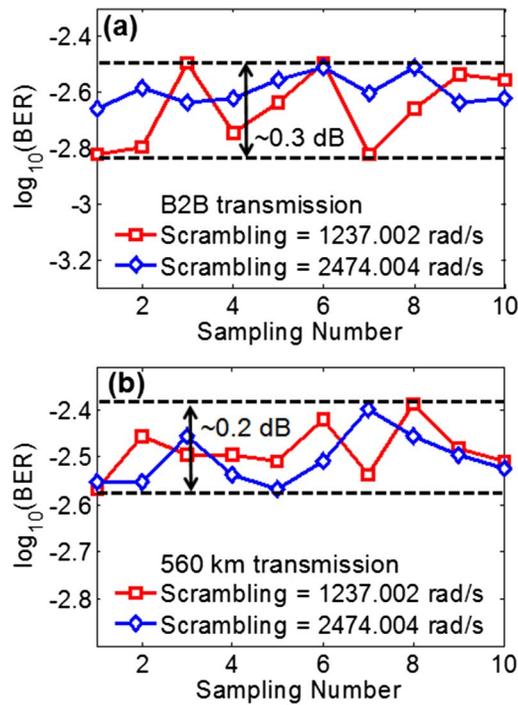


Fig. 7. Polarization tracking stability performance for the proposed method. (a) B-to-B BER performance (OSNR = 20 dB); (b) 560 km transmission BER performance (OSNR = 23 dB).

Table 1. Computational Complexity

Operations	Real Multiplication	Real Addition
Conventional EKF [14]	241	199
Proposed AKF	$245 + 4N$	$195 + 8N$

in Ref. [14], some additional operations (real multiplication and real addition) are taken to calculate the process noise covariance \mathbf{Q} in AKF. The required number of states $N = 8$ is used to calculate \mathbf{Q} in the experiment. But, for the conventional EKF method, \mathbf{Q} needs to be pre-known. A parameter-by-parameter tuning method can be used to obtain a relatively optimal value. However, this method requires a very large number of manual trials. For the proposed AKF, the tolerance of the initial error in \mathbf{Q} is greatly improved. Meanwhile, a faster convergence speed and smaller steady-state error are obtained. Therefore, the complexity of the proposed AKF is acceptable.

In conclusion, an AKF is proposed and experimentally demonstrated for the PolDemux technique. The results show that the proposed method may significantly improve the adaptive capability of an EKF by adaptively estimating the process noise covariance. By using an AKF, tedious and time-consuming parameter-by-parameter tuning methods can be avoided. Not only can the proposed AKF achieve better tracking accuracy, but the convergence is found to be about 4 times quicker than that of a conventional EKF. It is believed that the novel PolDemux method based on an AKF is a very attractive candidate for using PolDemux in flexible and elastic optical networks.

This work was supported by the National Natural Science Foundation of China (NSFC) (Nos. 61335005, 61325023, and 61401378).

References

- X. Zhou, *Chin. Opt. Lett.* **8**, 863 (2010).
- K. Kikuchi, *Opt. Express* **19**, 9868 (2011).
- J. Xiao, C. Tang, X. Li, J. Yu, X. Huang, C. Yang, and N. Chi, *Chin. Opt. Lett.* **12**, 050603 (2014).
- V. B. Ribeiro, J. C. Oliveira, J. C. Diniz, E. S. Rosa, R. Silva, E. P. Silva, L. H. Carvalho, and A. C. Bordonalli, in *OFC* (2012), paper OW3H.4.
- I. Fatadin, D. Ives, and S. J. Savory, *J. Lightwave Technol.* **27**, 3042 (2009).
- B. Szafraniec, B. Nebendahl, and T. Marshall, *Opt. Express* **18**, 17928 (2010).
- N. J. Muga and A. N. Pinto, *J. Lightwave Technol.* **31**, 2122 (2013).
- B. Szafraniec, T. S. Marshall, and B. Nebendahl, *J. Lightwave Technol.* **31**, 648 (2013).
- T. Marshall, B. Szafraniec, and B. Nebendahl, *Opt. Lett.* **35**, 2203 (2010).
- Y. Yang, G. Cao, K. Zhong, X. Zhou, Y. Yao, A. P. T. Lau, and C. Lu, *Opt. Express* **23**, 19673 (2015).
- J. Jignesh, B. Corcoran, C. Zhu, and A. Lowery, *Opt. Express* **24**, 22282 (2016).
- Y. Yang, Q. Zhang, Y. Yao, G. Cao, K. Zhong, X. Zhou, A. P. T. Lau, and C. Lu, *Chin. Opt. Lett.* **14**, 110601 (2016).
- N. J. Muga and A. N. Pinto, *J. Lightwave Technol.* **32**, 3290 (2014).
- N. J. Muga and A. N. Pinto, *J. Lightwave Technol.* **33**, 4826 (2015).
- L. Zanni, S. Sarri, M. Pignati, R. Cherkaoui, and M. Paolone, in *Proceedings of International Conference on Probabilistic Methods Application Power System (PMAAPS)* (2014).
- T. D. Powell, *J. Guidance Control Dyn.* **25**, 901 (2002).
- R. E. Kalman, *J. Basic Eng.* **82**, 35 (1960).
- Z. Xiao, S. Fu, S. Yao, M. Tang, P. Shum, and D. Liu, *J. Lightwave Technol.* **34**, 5526 (2016).