A KERNEL BASED TECHNIQUE FOR MSER EQUALISATION FOR NON-LINEAR CHANNELS

Rangeet Mitra and Vimal Bhatia

Discipline of Electrical Engineering, Indian Institute of Technology Indore, Indore-453441, India

ABSTRACT

Adaptive channel equalisation is a signal processing technique to mitigate inter-symbol interference (ISI) in a time dispersive channel. To this end, the use of least mean squares (LMS) algorithm and its variants is widespread since they minimise the minimum mean squared error (MMSE) criteria by online stochastic gradient algorithms and they asymptotically tend to the optimal Weiner solution for linearly separable channels. The kernel least mean squares (KLMS) algorithm and its variants are based on the MMSE based algorithms for non-linear channels. However, as has been pointed out in the literature, the minimum bit/symbol error rate (MBER/MSER) criterion is a better choice for adapting an equaliser as compared to the traditional approaches based on MMSE criterion. In this paper, we propose a novel equaliser that is inspired from the recently proposed MSER adaptation by Gong et al. using the kernel trick for non-linear channel equalisation.

Index Terms— Minimum symbol-error rate criterion, kernel trick, non-linear equalisation.

1. INTRODUCTION

The aim of any communication link design is to maximise the number of bits that can be sent reliably over that link (which is mathematically bounded by the Shannon capacity). There are three ways to optimise the performance of a communication link: a) Equalisation, b) Diversity and c) Channel-coding [1]. Equalisation, in particular, is the process of inferring the inverse transfer function of the channel so as to counter time dispersion. This inference from the channel can be supervised or unsupervised. Supervised learning assumes the knowledge of labels of training data. Unsupervised learning does not assume knowledge of labels of training data. The scope of this paper is limited to supervised equalisation.

Equalisers are generally adapted by the well known stochastic gradient algorithms using the least mean squares (LMS) algorithm and its variants [2] that are based on the minimum mean squared error (MMSE) criterion. However, as has been pointed out seminally in [3], and reviewed in [4–7] the minimum bit error rate (MBER)/ minimum symbol error rate (MSER) criterion is more suitable for channel equalisation as compared to MMSE approaches as it optimises the symbol error rate directly. Recently, an extension to the MBER paradigm was proposed in [8] based on a normalised filtering paradigm that used soft approximations for signum function and showed superior convergence performance as compared to the original adaptive minimum bit error rate (AMBER) algorithm [3].

All algorithms, reviewed above, work well when the data is affinely separable or "equalisable" as defined in theorems in [3]. However, when there is a non-linear channel, the data ceases to be affinely separable. In such scenarios, the works as in [9], which consider kernel based approaches for equalisation are more appropriate as the data, which is not linearly separable in indigenous space is mapped to kernel space where it can be linearly separable. Later, the work in [10], introduces a complex kernel for complex valued data. However, the works in [9, 10] are kernelised version of the adaptations of stochastic MMSE criterion. Though non-linear MBER based approaches have been explored in [6] using radial basis functions, they need further computations like adapting the centers and spread parameters. Also, the performance of radial basis functions (RBF) is dependent on initialisation of centers. Kernel based approaches do not need these computations and learn the parameters implicitly.

In this work, we propose a stochastic gradient based MSER algorithm, and invoke the kernel trick as in [10] for non-linear complex channel equalisation for the algorithm given in [8]. We find better convergence results in terms of symbol error rate in our preliminary investigations in case of the proposed algorithm as compared to the complex kernel least mean squares (KLMS) algorithm in [10] over non-linear channels.

2. SYSTEM MODEL

In this section, we describe the system model considered in the paper. From the system model depicted in Fig. 1, let s_k denote the input constellation at the k^{th} time instant. It is passed through a finite impulse response (FIR) filter $\{h_i\}_{i=1}^{L}$, where L is the tap length. The received symbol at k^{th} time

Thanks to IIT Indore and MHRD for funding this research.

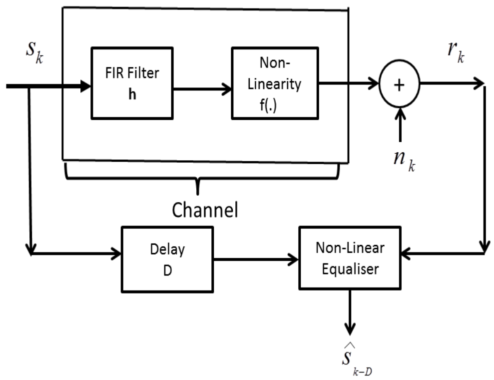

Fig. 1. System Model

instant , r_k, is given as follows:

$$r_k = f(\sum_{i=0}^{L-1} h_i s_{k-i}) + n_k \tag{1}$$

where $f(.)$ is an arbitrary non-linear mapping, n_k is complex additive white Gaussian noise (AWGN) and D denotes the equaliser delay. The channel consists of a linear FIR filter $\mathbf{h} = [h_0, h_1 ... h_{L-1}]$ and a non-linear mapping $f(.)$, which acts on the output of the convolution of input and the FIR filter.

3. BRIEF REVIEW OF NAMBER

The normalised-AMBER (NAMBER), originally proposed in [8], was found to be fast convergent as compared to AMBER owing to the use of a soft approximation of a signum function and an adaptation scheme similar to normalised-LMS (NLMS). In this section, we briefly review NAMBER for channel equalisation. Let us introduce the terminology to be followed in this section. While reviewing the NAMBER (MSER based equaliser) and to denote a linearly separable ISI channel we will use $f(x) = x$ and an FIR equaliser $\mathbf{c}_k$ at k^{th} time instant. Then r_k will purely be the convolution of input symbols with the FIR channel with AWGN added to it. NAMBER solves the following approximate optimisation problem ((26) from [8]):

$$\begin{aligned} &\min_{\mathbf{c}} . \|\mathbf{c}_k - \mathbf{c}_{k-1}\|_2^2 \\ \text{s.t. } &\text{sign}(\beta[\Re(\mathbf{c}_k^T \mathbf{r}_k) - \Re(s_{k-D}) + 1]) \\ &+ \text{sign}(\beta[\Re(\mathbf{c}_k^T \mathbf{r}_k) - \Re(s_{k-D}) - 1]) \quad = 0 \\ &\text{sign}(\beta[\Im(\mathbf{c}_k^T \mathbf{r}_k) - \Im(s_{k-D}) + 1]) \\ &+ \text{sign}(\beta[\Im(\mathbf{c}_k^T \mathbf{r}_k) - \Im(s_{k-D}) - 1]) \quad = 0 \end{aligned} \tag{2}$$

where, $\mathbf{r}_k$ is the vector of past N samples of channel output at k^{th} instant. Also, $\Re(.)$ and $\Im(.)$ indicate the real and imaginary part of the complex quantity respectively. NAMBER plugs in the $\tanh$ function as approximation for the signum function and further approximates its derivative by a first order Taylor approximation to obtain a normalised-AMBER algorithm.

The final adaptation equation is as follows:

$$\mathbf{c}_k = \mathbf{c}_{k-1} - \mu I_k \frac{\mathbf{r}_k^*}{\mathbf{r}_k^H \mathbf{r}_k + \epsilon} \tag{3}$$

where μ is the step-size, H denotes hermitian transpose and $*$ denotes complex conjugation. ϵ is an arbitrarily small value commonly used in normalised adaptive filtering [2].

Using $\tanh$ approximation for signum function, the following value of I_k was derived [8]:

$$\begin{aligned} I_k = &\tanh(\beta(\Omega_R + 1)) + \tanh(\beta(\Omega_R - 1)) \\ &+j(\tanh(\beta(\Omega_I + 1)) + \tanh(\beta(\Omega_I - 1))) \end{aligned} \tag{4}$$

where,

$$\Omega_R = \Re(\mathbf{c}_k^T \mathbf{r}_k) - \Re(s_{k-D})$$

and

$$\Omega_I = \Im(\mathbf{c}_k^T \mathbf{r}_k) - \Im(s_{k-D})$$

.

4. REVIEW OF COMPLEX KLMS

In this section, we review the complex KLMS algorithm. We first review the LMS algorithm and then show how it can be formulated in non-linear scenario using the kernel trick. The well known LMS algorithm adapts as follows:

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \mu e_k \mathbf{r}_k^* \tag{5}$$

where μ is the step size and e_k is the deviation of the filtered output from the desired training symbol. However, this algorithm fails to work in non-linear conditions where the data is not guaranteed to be affinely separable. The complex KLMS algorithm makes the parameter $\mathbf{c}_k$ implicit and writes the above adaptation as:

$$\mathbf{c}_{k+1} = \mu \sum_{i=1}^{k} e_i \mathbf{r}_i^* \tag{6}$$

assuming zero initial conditions. Invoking the kernel trick and writing the output of equaliser as $y_{k+1} =< \mathbf{c}_{k+1}, \mathbf{r}_{k+1} >_K$ (where $< .,. >_K$ denotes inner product in kernel space [9]), the above adaptation is given by:

$$y_{k+1} = \mu \sum_{i=1}^{k} e_i < \mathbf{r}_i, \mathbf{r}_{k+1} >_K \tag{7}$$

where K represents a kernel inner product. The most popular choice of kernels is a Gaussian kernel given in (8) as it maps the dataset to infinite dimensional Hilbert space and renders the data to be affinely seperable in that space. It is given by the following formula [10],

$$< \mathbf{r}_i, \mathbf{r}_{k+1} >_K = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{\sum_{\forall q} (r_i^q - r_{k+1}^{q^*})}{2\sigma^2} \right) \tag{8}$$

σ is the spread parameter and controls the performance of the algorithm. It can be determined by using Silverman's rule as in [11]. Also the superscript q denotes the q^{th} component of vector.

5. PROPOSED ALGORITHM

In this section, we propose a new algorithm for non-linear channel equalisation based on MSER based adaptation. The NAMBER algorithm, which is reviewed in Section 3 ceases to perform well in situations where data is not guaranteed to be linearly separable, for e.g., non-linear channels; because it estimates only an affine parameter. The complex KLMS algorithm, which is reviewed in Section 4 is more suitable for non-linear datasets. However, the complex KLMS algorithm is based on a MMSE criterion in a kernel Hilbert space. Therefore, we try to formulate the problem inspired by an MSER solution in kernel space without the need to estimate centers and spread factors as in [6].

By applying the kernel trick in (3) the adaptation at the k^{th} instant would then become,

$$y_{k+1} = -\mu \sum_{i=1}^{k} I_i < \mathbf{r}_i, \mathbf{r}_{k+1} >_K \tag{9}$$

where I_i is the corresponding error term at i^{th} time instant whose value has been defined in Section 3. The $\|\mathbf{r}_k\|_2^2$ is not included in the denominator for two reasons: a) It is a non-linearity that can be learnt by the Gaussian kernel function, which has been proven to be capable of approximating (also known as universal approximation property) any arbitrary non-linearity, and b) It is computationally less cumbersome as we are saved of kN multiplications and k divisions.

This kernelised MSER algorithm like some of the MSER algorithms reviewed in [6] can be viewed as higher order statistics of data and hence is more suitable for non-linear data when the underlying pdf is non-Gaussian. Hence the MSER criterion can be considered a better cost function as compared to MMSE based approaches capable of handling deep fading scenarios. However, as opposed to some of the RBF based non-linear algorithms in [6], we do not need to estimate the RBF centers as they are optimised implicitly by the kernel inner product. We chose to apply the kernel trick on the algorithm in [8] as it showed significantly better convergence as compared to other MSER based equalisers reviewed in [8] over linear channels.

6. SIMULATIONS

In this section, we present the simulation results to validate the proposed scheme and compare the results with complex KLMS for non-linear equalisation. To evaluate the proposed scheme, we present four simulations. We first observe how the symbol error rate (SER) evolves as a function of adaptation iterations for QPSK and QAM modulation schemes keeping by the signal to noise ratio (SNR) fixed in Fig. 2 and Fig. 3. Consequently, we vary the SNR and see how the converged symbol error rate decays as a function of SNR in Fig. 4 and Fig. 5.

In the first simulation, we considered a generic QPSK constellation and passed it through a three-tap FIR channel $\mathbf{h} = [0.341, 0.876, 0.341]$ ((CH=4) of [12]) for both real and imaginary parts. The $f(x) = x + 0.2x^2 - 0.1x^3$ ((NL=3) of [12]) was used as a non-linearity for both real and imaginary parts of the channels. Independently identically distributed (i.i.d.) complex Gaussian noise at 25dB SNR was added. The convergence plots are shown in Fig. 2. The proposed algorithm converges to a lower SER floor as compared to complex KLMS.

For the second simulation, we considered a 16-QAM constellation and mapped it to a new constellation by the technique proposed in [13]. It was convolved by the FIR filter $\mathbf{h} = [0.26, 0.93, 0.26]$ ((CH=2) of [12]) for both real and imaginary parts. The non-linearity was as same as considered in the first simulation. The convergence plots are shown in Fig. 3. Convergence to a lower error rate metric is seen as compared to complex KLMS algorithm.

In the third simulation, we compare the SER vs SNR performance of the complex KLMS and the proposed algorithms for QPSK in the simulation settings described above in the first simulation. It is observed that in Fig. 4, the proposed algorithm outperforms complex KLMS by half a decade. This curve is obtained by averaging over 200 Monte Carlo simulations. The same channel and non-linearity as in the first simulation are used. SER was monitored (for plotting it vs SNR) after 500 iterations of adaptation of both algorithms.

In the fourth simulation, in Fig. 5 we compare the SER vs SNR performance of the complex KLMS and the proposed algorithm for 16-QAM in the simulation settings discussed above for 16-QAM in the second simulation. It is observed that the proposed algorithm outperforms complex KLMS by

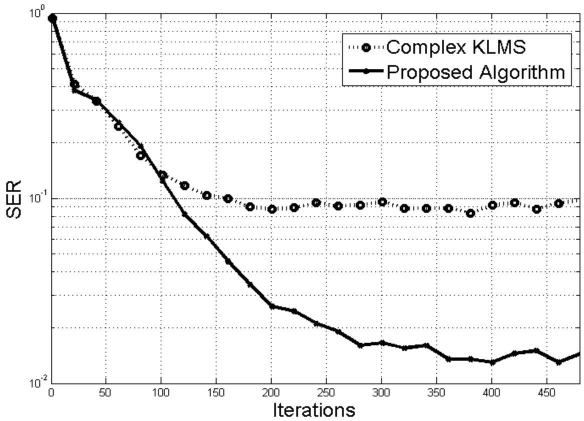

Fig. 2. Error Rate Convergence Comparison for QPSK. $\mathbf{h} = [0.341, 0.876, 0.341]$, $f(x) = x + 0.2x^2 - 0.1x^3$,$\mu = 1$.

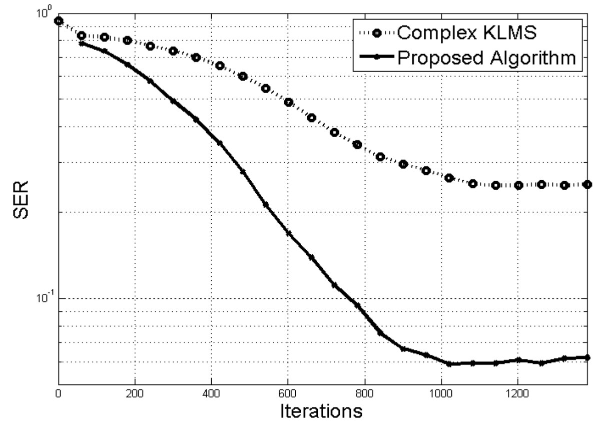

Fig. 3. Error Rate Convergence Comparison for QAM. $\mathbf{h} = [0.26, 0.93, 0.26]$, $f(x) = x + 0.2x^2 - 0.1x^3$,$\mu = 1$.

more than half a decade. SER was monitored (for plotting it vs SNR) after 1000 iterations of adaptation of both algorithms. The same channel and non-linearity as used in Fig. 3 was used in this simulation.

In Fig. 6 the proposed approach is compared with the other algorithms given in [10] like normalised complex KLMS-1 (NCKLMS1) and normalised complex KLMS-2 (NCKLMS2). Faster convergence is observed in case of the proposed algorithm as compared to NCKLMS1 and NCKLMS2 in the same simulation conditions as described as a "hard non-linear channel equalisation" environment in [10] with the same step-size μ and spread factor σ. The channel $\mathbf{h} = [-0.9 + 0.8i, 0.6 - 0.7i]$ and $f(x) = x + (0.2 + 0.25i)x^2 + (0.12 + 0.09i)x^3$ was considered.

Lastly, in Fig. 7, the proposed approach is compared with algorithms in [10] like complex KLMS, NCKLMS1 and NCKLMS2 by varying the SNR and plotting the SER after setting the number of iterations to 500 when all algorithms have converged. It is observed from Fig. 7 that at low SNR, the SER performance of complex KLMS, NCKLMS1, NCKLMS2 and the proposed algorithm are equivalent. A gain of 4dB is observed as compared to NCKLMS1 and NCKLMS2 at an SER of 4×10^-2. However, as the SNR increases, the SER of the proposed algorithm is less than the error rate of complex KLMS, NCKLMS1 and NCKLMS2. Thus we observe that at high SNR the proposed algorithm is a more suitable equalisation algorithm as compared to complex KLMS, NCKLMS1 and NCKLMS2. From Fig. 6 and Fig. 7, we can conclude that the proposed approach has the same computational complexity as compared to complex KLMS and has a better SER performance than more computationally complex algorithms like NCKLMS1 and NCKLMS2.

Please note that the step-sizes for NCKLMS1 and NCKLMS2 have been chosen according to [10], in which the step-size parameter have been "tuned for best performance".

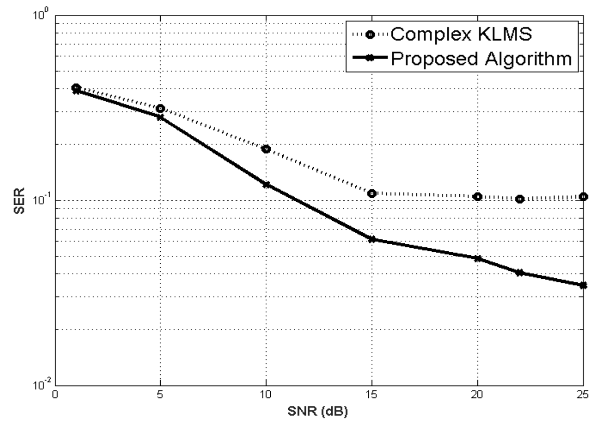

Fig. 4. SER vs SNR Comparison for QPSK. $\mathbf{h} = [0.341, 0.876, 0.341]$, $f(x) = x + 0.2x^2 - 0.1x^3$,$\mu = 1$.

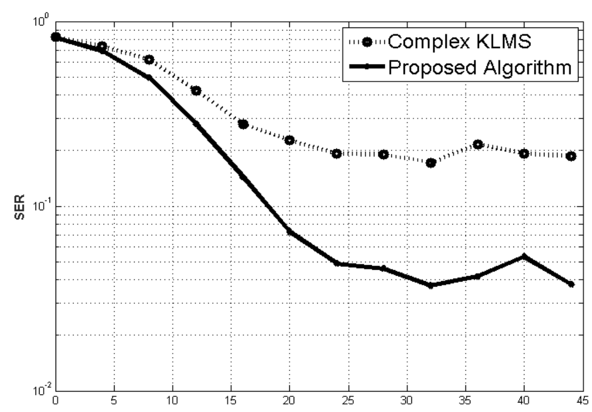

Fig. 5. SER vs SNR Comparison for QAM. $\mathbf{h} = [0.26, 0.93, 0.26]$, $f(x) = x + 0.2x^2 - 0.1x^3$,$\mu = 1$.

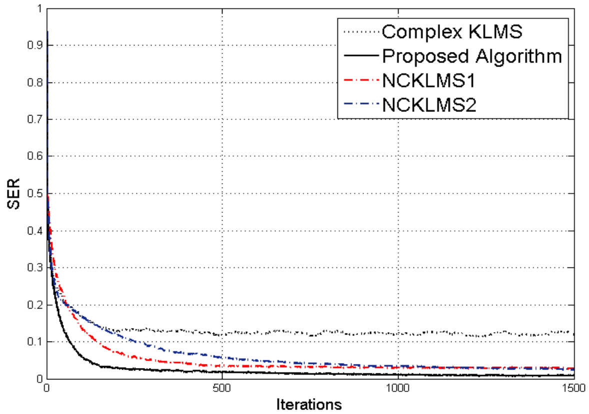

Fig. 6. SER vs iterations for comparison with complex KLMS ($\mu = 0.1$), NCKLMS1 ($\mu = 0.5$) and NCKLMS2 ($\mu = 0.25$) for QPSK for channel in [10], SNR=25dB.

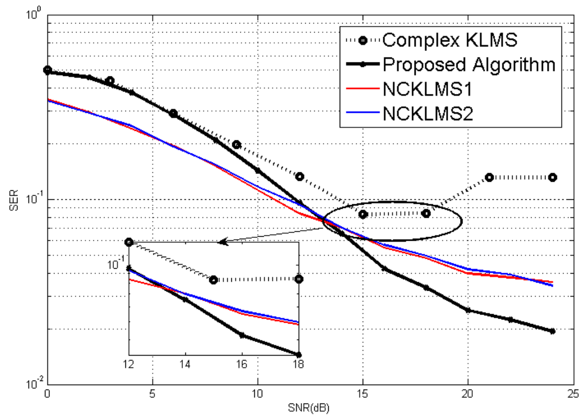

Fig. 7. SER vs SNR comparison with complex KLMS ($\mu = 0.1$), NCKLMS1 ($\mu = 0.5$) and NCKLMS2 ($\mu = 0.25$) for QPSK for channel in [10].

7. CONCLUSION

A new MSER based approach for non-linear channel equalisation has been proposed and compared with complex KLMS, NCKLMS1 and NCKLMS2. Better convergence is observed in case of the proposed algorithm as compared to complex KLMS. Also, we get lower SER as a function of SNR in case of the proposed scheme as compared to complex KLMS based approaches. Hence, it is a better channel equalisation algorithm for QPSK and QAM modulation techniques.

REFERENCES

[1] J. G. Proakis, *Intersymbol Interference in Digital Communication Systems*, Wiley Online Library, 2001.

[2] A. H. Sayed, *Fundamentals of Adaptive Filtering*, John Wiley and Sons, 2003.

[3] C. C. Yeh and J. R. Barry, "Adaptive minimum bit-error rate equalization for binary signaling," *IEEE Transactions on Communications*, vol. 48, no. 7, pp. 1226–1235, 2000.

[4] S. Chen, A. K. Samingan, B. Mulgrew, and L. Hanzo, "Adaptive minimum-ber linear multiuser detection for DS-CDMA signals in multipath channels," *IEEE Transactions on Signal Processing*, vol. 49, no. 6, pp. 1240–1247, 2001.

[5] S. Chen, N. N. Ahmad, and L. Hanzo, "Adaptive minimum bit-error rate beamforming," *IEEE Transactions on Wireless Communications,*, vol. 4, no. 2, pp. 341–348, 2005.

[6] S. Chen, S. Tan, L. Xu, and L. Hanzo, "Adaptive minimum error-rate filtering design: A review," *Signal Processing*, vol. 88, no. 7, pp. 1671–1697, 2008.

[7] V. Bhatia, B. Mulgrew, and A.T. Georgiadis, "Stochastic gradient algorithms for equalisation in α-stable noise," *Signal Processing*, vol. 86, no. 4, pp. 835–845, 2006.

[8] M. Gong, F. Chen, H. Yu, Z. Lu, and L. Hu, "Normalized adaptive channel equalizer based on minimal symbol-error-rate," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1374–1383, 2013.

[9] W. Liu, P. P. Pokharel, and J. C. Principe, "The kernel least-mean-square algorithm," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 543–554, 2008.

[10] P. Bouboulis and S. Theodoridis, "Extension of wirtinger's calculus to reproducing kernel hilbert spaces and the complex kernel LMS," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 964–978, 2011.

[11] B. W. Silverman, *Density estimation for statistics and data analysis*, vol. 26, CRC press, 1986.

[12] J. C. Patra, P. K. Meher, and G. Chakraborty, "Nonlinear channel equalization for wireless communication systems using legendre neural networks," *Signal Processing*, vol. 89, no. 11, pp. 2251–2262, 2009.

[13] W. Rao, "Blind equalization using coordinate mapping for 4-PAM and 16-QAM signals," *Chinese Journal of Electronics*, vol. 23, no. 1, 2014.