FUSION OF HETEROGENEOUS DATA SOURCES: A QUATERNIONIC APPROACH

Clive Cheong Took and Danilo Mandic

Electrical and Electronic Engineering Department,
Imperial College London, London SW7 2AZ , UK,
Phone: +44-207-594-6271, Emails: {c.cheong-took, d.mandic}@ic.ac.uk.

ABSTRACT

Sequential fusion of three- and four- dimensional heterogeneous data is achieved in the quaternion space $\mathbb{H}$. This way, data from multiple sensors are combined in order to achieve "improved accuracies" and more specific inferences that could not be performed by the use of only a single sensor. To this end, the quaternion LMS (QLMS) is proposed for the *online* fusion of hypercomplex data within the "data fusion via vector spaces" framework. Case studies on real-world signals such as environmental and financial time series are provided to support the proposed approach.

1. INTRODUCTION

Data fusion offers enhanced capability to exploit the available information within homogeneous or heterogenous data sources. This is particularly important when modeling multichannel processes with time varying degree of correlation and coupling between the channels. There are several established data fusion approaches, those include the hierarchical, decentralized and sequential fusion [1, 2].

From the viewpoint of machine learning and signal processing, sequential fusion is most interesting since it facilitates the adaptive online mode of operation, and the use of well understood adaptive filters in this context. One recently proposed sequential data fusion model is so-called "data fusion via vector spaces", where heterogeneous quantities are conveniently modeled as a single higher dimensional quantity. One such example is in wind modeling [3, 4], where wind speed and direction are fused into a complex-valued quantity and the modeling is performed simultaneously for both the speed and direction using modular nonlinear adaptive filtering architectures such as the complex pipelined recurrent neural network (PRNN) [1, 5]. This way, the wind vector $\mathbf{v}(k)$ is expressed as

$$\mathbf{v}(k) = |\mathbf{v}(k)| \exp^{\imath\theta(k)} = v_E(k) + \imath v_N(k) \qquad (1)$$

where θ, $|\mathbf{v}(k)|$, $v_E(k)$, and $v_N(k)$ denote respectively the direction, magnitude, wind speed in the east direction, and wind speed in the north direction. Data fusion via vector spaces is particularly useful for processes which are made complex (or hypercomplex) by *convenience of representation* such as in the case of complex valued wind model, shown in Fig. 1

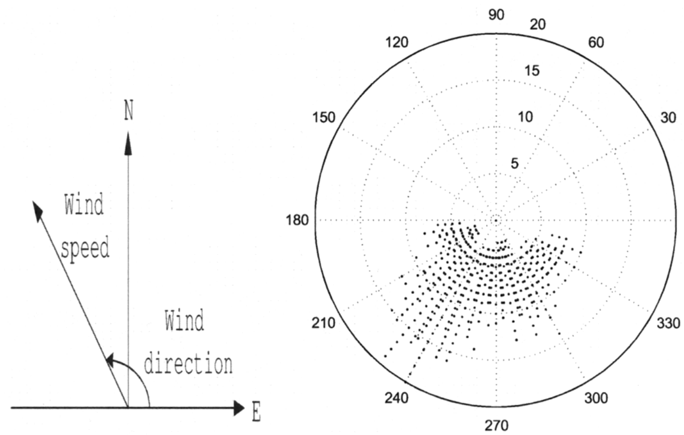

Fig. 1. Wind recordings: *Left*: a complex-valued representation, *Right*: wind lattice.

Here the wind speed and direction exhibit very different dynamics but are correlated (right hand plot of Fig. 1), and the complex model is very natural for their simultaneous modeling. Other multi-dimensional signal processing techniques include the concept of "long vector" [6], and real-valued multichannel LMS [7].

Sequential data fusion via vector spaces becomes even more interesting in the modeling of three- and four-dimensional quantities, for instance, in the simultaneous modeling of a three dimensional wind vector and air temperature. It is therefore natural to ask whether the different dynamics of the wind speed components and air temperature and the different time scales within these components can be simultaneously modeled in a four dimensional quaternion space $\mathbb{H}$.

The aim of this paper is to introduce sequential data fusion via quaternion hypercomplex spaces, in order to enable simultaneous modeling of three- and four-dimensional quantities. To achieve this, we first introduce a quaternion LMS (QLMS) algorithm and then illustrate its usefulness for the modeling of

1-4244-2376-7/08/$20.00 ©2008 IEEE

earthquake time series, environmental time series, and financial indexes.

2. QUATERNION ALGEBRA

Quaternions can be considered as non-commutative extensions of complex numbers, and comprise at most four variables [8]. A quaternion variable $q \in \mathbb{H}$ has a real part (denoted with subscript a), and three imaginary parts (denoted with subscripts b, c, d), and can be expressed as:

$$\begin{aligned} q &= [q_a, \mathbf{q}] \\ &= [q_a, (q_b, q_c, q_d)] \\ &= q_a + q_b\imath + q_c\jmath + q_d\kappa \qquad \{q_a, q_b, q_c, q_d \in \mathbb{R}\} \end{aligned} \tag{2}$$

Properties of the orthogonal unit vectors, $\imath$, $\jmath$, κ describing the three vector dimensions of a quaternion are

$$\begin{aligned} \imath\jmath &= \kappa \\ \jmath\kappa &= \imath \\ \kappa\imath &= \jmath \\ \imath\jmath\kappa &= \imath^2 = \jmath^2 = \kappa^2 = -1 \end{aligned} \tag{3}$$

Due to the non-commutivity of the quaternion, for example, $\jmath\imath \neq \imath\jmath$, instead $\jmath\imath = -\kappa$. Likewise, the product of quaternions w and x is given by

$$\begin{aligned} wx &= [w_a, \mathbf{w}][x_a, \mathbf{x}] \\ &= [w_a x_a - \mathbf{w}\cdot\mathbf{x},\ w_a\mathbf{x} + x_a\mathbf{w} + \mathbf{w}\times\mathbf{x}] \end{aligned} \tag{4}$$

where symbols "$\cdot$" and "$\times$" denote respectively the dot-product and the cross-product. Elementwise, the quaternion product can be evaluated as

$$wx = \begin{pmatrix} w_a x_a - w_b x_b - w_c x_c - w_d x_d \\ w_a x_b + w_b x_a + w_c x_d - w_d x_c \\ w_a x_c + w_c x_a + w_d x_b - w_b x_d \\ w_a x_d + w_d x_a + w_b x_c - w_c x_b \end{pmatrix} \tag{5}$$

It is clear that the quaternion product is non-commutative, however, similarly to the complex numbers, the conjugate of a quaternion $q^* = [q_a, \mathbf{q}]^* = [q_a, -\mathbf{q}]$, and the norm $||q||_2^2 = qq^*$; for more detail, see [8]. For clarity, in this work we consider scalar quaternion quantities, since we can also have a vector of quaternions. We will next exploit the intimate relationships between the components of a quaternion to model the coupling between the channels of a 4D signal.

3. DERIVATION OF QUATERNION LMS (QLMS)

Based on the properties of quaternion algebra and by employing stochastic gradient descent, we shall now introduce the QLMS algorithm for finite impulse response (FIR) adaptive filters. The cost function $\mathcal{J}(n)$ is the instantaneous squared error given by

$$\begin{aligned} \mathcal{J}(n) &= e(n)e^*(n) \\ &= e_a^2(n) + e_b^2(n) + e_c^2(n) + e_d^2(n) \\ &= d(n)d^*(n) + y(n)y^*(n) \\ &\quad -y(n)d^*(n) - d(n)y^*(n) \end{aligned} \tag{6}$$

where the error $e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n)$, with $d(n)$, $\mathbf{w}(n)$, and $\mathbf{x}(n)$ denoting respectively the teaching signal, the adaptive weight vector, and the filter input. Symbols $(\cdot)^T$, and $(\cdot)^*$ denote respectively the vector transpose, and quaternion conjugate operator. The filter output $y(n)$ can be computed as

$$y(n) = \mathbf{w}^T(n)\mathbf{x}(n) = \sum_{m=1}^{L} w_m(n)x(n-m) \tag{7}$$

To update the mth coefficient $w_m(n)$, we need to calculate the following gradient

$$\begin{aligned} \nabla_{w_m}(\mathcal{J}(n)) &= \nabla_{w_m}(y(n)y^*(n)) - \nabla_{w_m}(y(n)d^*(n)) \\ &\quad -\nabla_{w_m}(d(n)y^*(n)) \end{aligned} \tag{8}$$

where

$$\begin{aligned} \nabla_{w_m}(y(n)y^*(n)) &= 4y(n)x^*(n-m) - 2x^*(n-m)y^*(n) \\ \nabla_{w_m}(y(n)d^*(n)) &= -2x^*(n-m)d^*(n) \\ \nabla_{w_m}(d(n)y^*(n)) &= 4d(n)x^*(n-m) \end{aligned} \tag{9}$$

which yields

$$\begin{aligned} \nabla_{w_m}(\mathcal{J}(n)) &= 4y(n)x^*(n-m) - 2x^*(n-m)y^*(n) \\ &\quad +2x^*(n-m)d^*(n) - 4d(n)x^*(n-m) \\ &= 4\big(y(n) - d(n)\big)x^*(n-m) \\ &\quad -2x^*(n-m)\big(y^*(n) - d^*(n)\big) \\ &= -2\big(2e(n)x^*(n-m) - x^*(n-m)e^*(n)\big) \end{aligned} \tag{10}$$

Before calculating the above gradients, notice that the cost function (6) is real valued, and not quaternion-valued as in [9,10]. Furthermore, the non-commutativity of the quaternion product must be taken into account, that is

$$\begin{aligned} \nabla_{w_m}(\mathcal{J}(n)) &= e(n)\nabla_{w_m}(e^*(n)) + \nabla_{w_m}(e(n))e^*(n) \\ &\neq e(n)\nabla_{w_m}(e^*(n)) + e^*(n)\nabla_{w_m}(e(n)) \end{aligned} \tag{11}$$

Finally, the update of the adaptive weight vector of QLMS can be expressed as:

$$\underbrace{w_m(n+1) = w_m(n) + \mu\big(2e(n)x^*(n-m)}_{\text{complex LMS}} - x^*(n-m)e^*(n)\big) \tag{12}$$

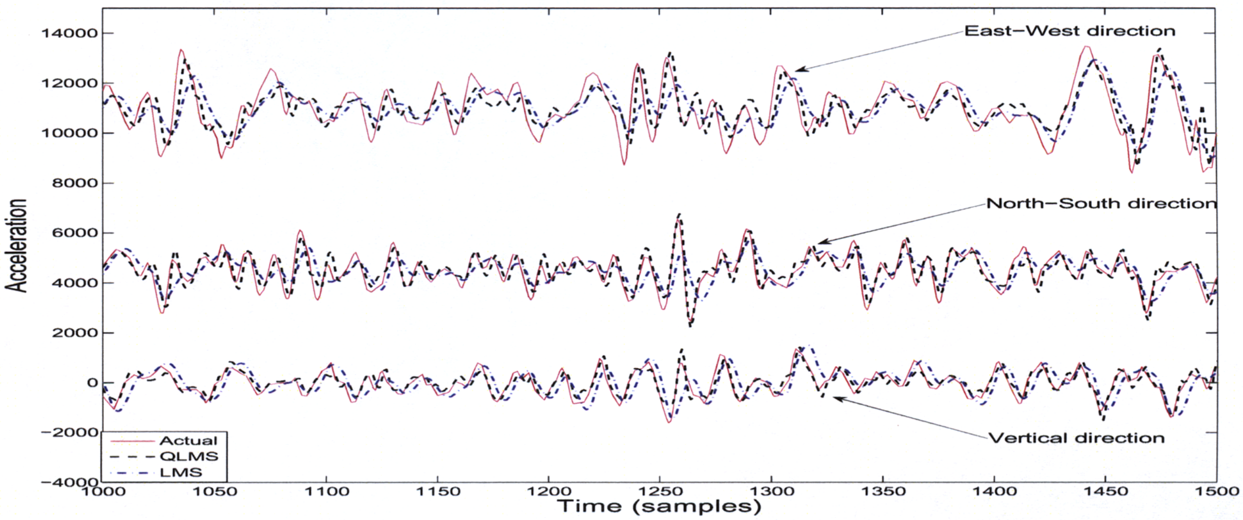

Fig. 2. Accelerometer readings of Loma Prieta Earthquake in the east-west, north-south, and vertical direction. The prediction gains for the triple univariate LMS (a separate real value LMS per channel) and the QLMS were respectively $R_p = 10$ dB and $R_p = 16.66$ dB.

From (12), there is similarity between the updates of complex LMS (CLMS) [11] and QLMS. Therefore, it is natural to ask whether QLMS is a generic extension of CLMS. If the imaginary parts $\jmath$ and κ of quaternions $\mathbf{x}(n-m)$ and $e(n)$ vanish, the quaternion representation has only the real and one imaginary dimension and the QLMS update becomes

$$w_m(n+1) = w_m(n) + \mu\Big([e_a(n)x_a(n-m) + 3e_b(n)x_b(n-m)] + \imath[-e_a(n)x_b(n-m) + 3e_b(n)x_a(n-m)]\Big) \tag{13}$$

whereas the CLMS update is given by

$$w_m(n+1) = w_m(n) + \mu\Big([e_a(n)x_a(n-m) + e_b(n)x_b(n-m)] + \imath[-e_a(n)x_b(n-m) + e_b(n)x_a(n-m)]\Big) \tag{14}$$

A comparison between (13) with (14) shows that QLMS does not degenerate into CLMS, highlighting the fact that QLMS is not a trivial extension of CLMS. The potential advantages of QLMS over multiple LMS for the fusion applications) include

- Since both QLMS and LMS are online algorithms, it is important to utilize fully the available information. From the quaternion product (5), unlike a quadruple univariate LMS, the QLMS exploits the cross-correlation and the coupling between each dimension to process the hypercomplex signal;
- The stepsize plays a key role in the performance of stochastic gradient algorithms. Selection of four stepsizes and their manual "tuning in" for a reasonable performance can be time consuming and problematic when processing each of the four dimensions separately with LMS. QLMS circumvents this problem since it requires only one stepsize.

4. CASE STUDIES

In most applications, the heterogeneous components of a process are modeled separately, as univariate quantities. This is typically due to their radically different natures, e.g. it is difficult to model the air temperature and wind speed within the same model, although they are correlated. Our aim is to illustrate that when heterogenous data sources are modeled as a hypercomplex quantity by "convenience of representation", this results in a greatly improved performance.

We now present case studies, illustrating the potential of information fusion based on the QLMS in the adaptive prediction setting. In multi-step ahead prediction, we need to estimate the signal M-samples ahead based on values in the past. In our simulations, we have set $M = 2$; the extent of multistep ahead prediction is also known as the prediction horizon. For a quantitative evaluation of the performance, we employ the standard prediction gain R_p [5], given by

$$R_p = 10 \log \frac{\sigma_{\mathbf{x}}^2}{\sigma_e^2} \quad \text{(dB)} \tag{15}$$

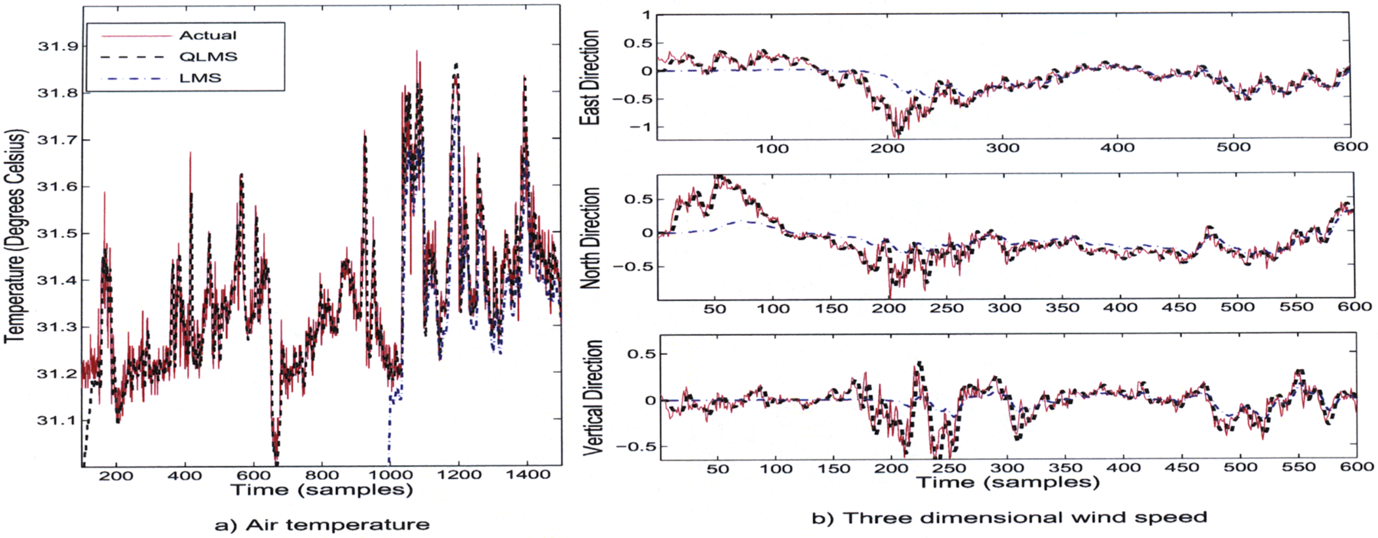

Fig. 3. Comparison of prediction results for the quadruple LMS and QLMS. The prediction gain for the four univariate LMS and the QLMS were respectively $R_p = 9.98$ dB and $R_p = 29.78$ dB.

where $\sigma_{\mathbf{x}}^2$ and σ_e^2 denote respectively the estimated variances of the input and the error.

Prediction plays a pivotal role in the forecasting of earthquakes, wind, and in financial markets and is rarely conducted based on data fusion principles. To this end, we compare the results obtained by four independent LMS algorithms, with those obtained by direct QLMS. For each scenario, the stepsizes were chosen manually to ensure optimal performances.

4.1. The Loma Prieta Earthquake

Seismic signals[1], shown in Fig. 2 occur naturally as three-dimensional, and can therefore be modeled as a *pure* quaternion (a *pure* quaternion has zero real part). In the simulations, the stepsizes for the three independent LMS algorithms were selected as: $\mu_1 = 8 \times 10^{-6}$ (vertical), $\mu_2 = 2 \times 10^{-5}$ (north-south), and $\mu_3 = 3 \times 10^{-6}$ (east-west), while the stepsize for the QLMS was set to $\mu = 5 \times 10^{-9}$.

Fig. 2 shows that estimates for both the LMS (dash-dot line) and QLMS (broken line) were reasonable. In fact, the QLMS and LMS had similar performance for the east-west direction. On the other hand, QLMS outperformed LMS in the north-south and vertical estimates of Loma Prieta Earthquake. The overall prediction gain for the quadruple LMS was $R_p = 10$ dB, whereas that of QLMS was $R_p = 16.66$ dB.

[1]The Loma Prieta Earthquake data was recorded from an accelerometer by Joel Yellin at the Charles F. Richter Seismological Laboratory, University of California.

4.2. Fusion of wind and temperature

The wind data (a segment shown in Fig. 3) was initially sampled at 50 Hz, but re-sampled at 5 Hz for simulation purposes. To illustrate the benefits of the quaternion representation in the fusion of heterogeneous data sources, we combined the air temperature with the three wind directions. The stepsizes for the quadruple univariate LMS were selected as: $\mu_1 = 1.03 \times 10^{-7}$ (temperature), $\mu_2 = 9 \times 10^{-3}$ (vertical), $\mu_3 = 1.5 \times 10^{-3}$ (north-south), and $\mu_4 = 2 \times 10^{-3}$ (east-west), whereas the stepsize for the QLMS was set to $\mu = 10^{-5}$.

Notice that the dynamics of temperature is quite different from that for the three orthogonal wind speeds, that is, the fluctuation of the temperature is relatively small compared to wind speeds. Fig. 3 shows that QLMS converged much faster than the quadruple LMS.

4.3. Stock market prediction

Another field where prediction plays a prominent role is in finance. We will now demonstrate how we can exploit the different features of stocks within the quaternionic data fusion model. The stock market prices analyzed were obtained from "http://biz.yahoo.com/r/", under the tab "Historical Quotes" for the period of 25 April, 2003 to 24 April 2008, with a sampling period of one trading day. We considered two particular cases: 1) Four stock features of Citigroup: the trading volume, and its high, low, and closing prices as illustrated in Fig. 4; 2) Four stocks fused into the quaternionic model,

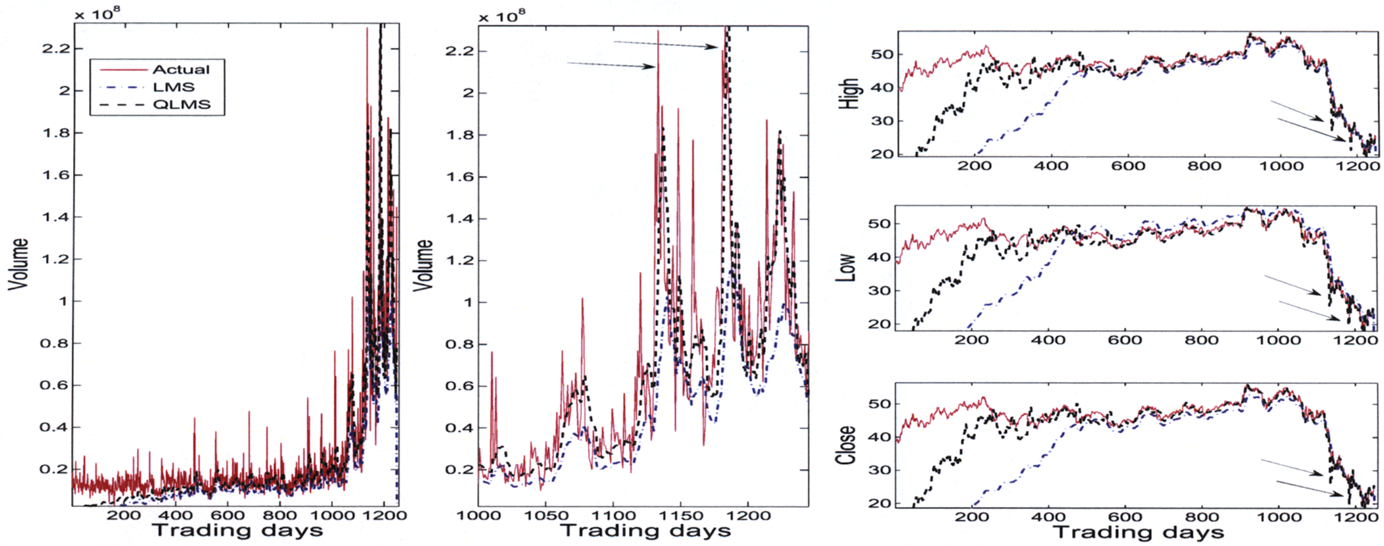

Fig. 4. Prediction results for four features of Citigroup stock: the volume traded, the high, low, and closing prices per trading day. The prediction gains for the quadruple univariate LMS and the QLMS were respectively $R_p = 9.92$ dB and $R_p = 17.97$ dB. The middle plot zooms the volume traded during a high activity period. Notice the *negative* correlation between the prices and the volume traded.

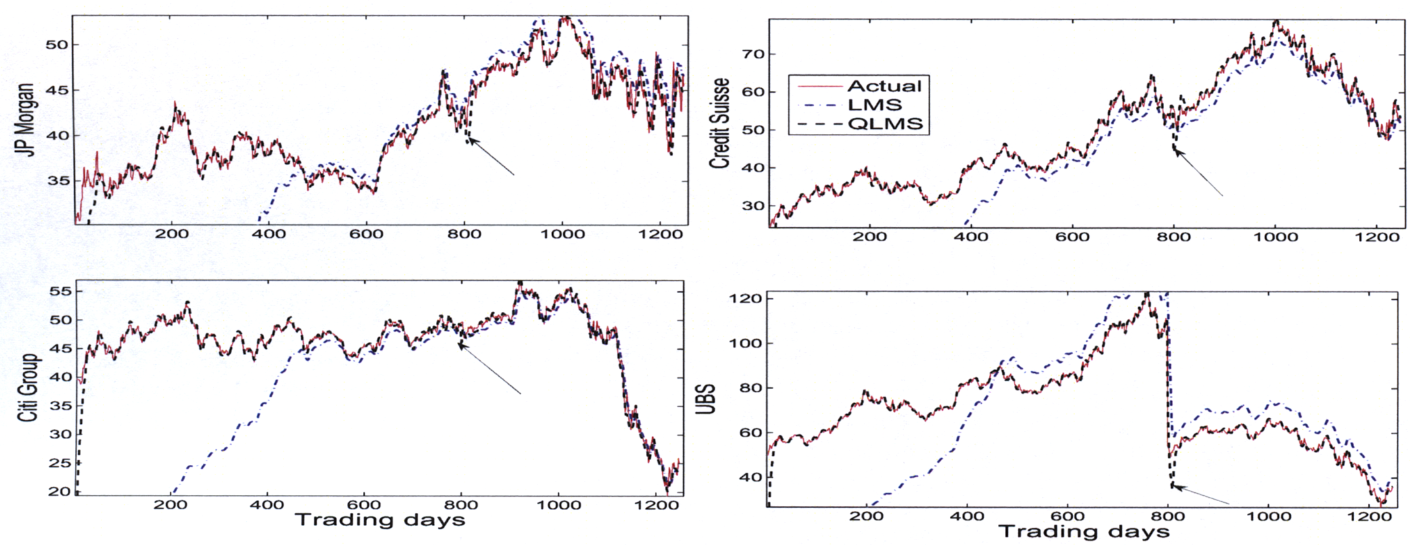

Fig. 5. The stock (high) price of JPMorgan, UBS, Citigroup, and Credit Suisse, with one sample per trading day. The prediction gains for the quadruple univariate LMS and the QLMS were respectively $R_p = 9.99$ dB and $R_p = 35.88$ dB. Observe the price drop in the UBS stock at around the 800th sample; the accurate modeling of this feature enabled the QLMS to predict a dip (shown by arrows) in the other stock prices.

these were the JP Morgan, UBS, Citigroup, and Credit Suisse stocks as illustrated in Fig. 5.

Fig. 4 illustrates the positive correlation between the price features in the right-hand side plots, while the volume traded for Citigroup exhibited a negative correlation with respect to the prices. There was no significant fluctuations in the stock features, until the 1000th trading day. From the 1000th trading day on, the net decrease in price features prompts activity on the trading floor, as shown by the high fluctuations of the volume (middle plot). The quadruple LMS predicting the volume traded failed to follow the dynamics of this feature, while QLMS estimate was much better, especially during the high fluctuations (shown by the arrows in the middle plot). Subsequently, QLMS predicted a net decrease in prices (pointed by arrows in the right-hand plots), owing to the inherent coupling within the quaternionic model. If a trader was to follow the prediction by QLMS, they would gain profit by selling the stock, since the price was still on the decline. For this scenario, the prediction gain for the quadruple univariate LMS was $R_p = 9.92$ dB, while that of QLMS was $R_p = 17.97$ dB. The optimal stepsizes for the four univariate LMS were selected as: $\mu_1 = 8 \times 10^{-19}$ (volume), $\mu_2 = 9 \times 10^{-8}$ (high), $\mu_3 = 1 \times 10^{-7}$ (low), and $\mu_4 = 9 \times 10^{-8}$ (close), highlighting the different dynamics of each feature; while the stepsize of QLMS was set to $\mu = 10^{-5}$.

Fig. 5 illustrates stock (high) prices of JP Morgan, Citigroup, UBS and Credit Suisse. In this scenario, it is not clear whether there exists a degree of correlation between the stock prices. However, if we zoom in from the 1100th trading day, we notice a general decrease in stock prices, except for JP Morgan. This arises in the aftermath of the credit crunch market in the USA. Due to its *top-tier* investment bank status, JP Morgan had more cushion against the credit crunch, and therefore traders did not know if they were to sell or keep JP Morgan stocks; this caused high fluctuations in the last trading days. Despite this uncertainty, the QLMS estimate of the JP Morgan stock was quite reasonable. Also, observe the significant dip in UBS stock price around the 800th trading day (pointed by an arrow): inherent coupling within the dimensions of the quaternionic model enabled QLMS to predict a drop in the other stock prices. The prediction gain for the four univariate LMS was $R_p = 9.99$ dB, whereas that of QLMS was $R_p = 35.88$ dB.

5. CONCLUSION

A vector space based approach for the fusion of heterogeneuos data sources has been proposed. This has been achieved in the quaternion domain $\mathbb{H}$ for three- and four- dimensional data. Next, the quaternion least mean square (QLMS) has been introduced as a convenient collaborative signal processing tool for four-dimensional processes. Due to the inherent coupling between the components within the quaternionic signal model, we have obtained much improved results in the prediction of real world processes such as 3D seismic signals, wind data and air temperature, and financial stocks.

6. ACKNOWLEDGEMENT

We are most grateful to Prof. Kazuyuki Aihara and his team at the University of Tokyo, for providing us with the wind data recorded in an urban environment.

7. REFERENCES

[1] D. Mandic, M. Golz, A. Kuh, D. Obradovic, and T. Tanaka (Eds.), *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, Springer Verlag, 2008.

[2] D. Mandic, D. Obradovic, A. Kuh, T. Adali, U. Trutschel, M. Golz, P. D. Wilde, J. A. Barria, A. Constantinides, and J. A. Chambers, "Data Fusion for Modern Engineering Applications: An Overview," *Proceedings of ICANN '05*, pp. 715–721, Sept. 2005.

[3] D. P. Mandic, S. L. Goh, and K. Aihara, "Sequential Data Fusion via Vector Spaces: Complex Modular Neural Network Approach," *IEEE Workshop on Machine Learning for Signal Processing*, pp. 147 – 151, Sept. 2005.

[4] D. P. Mandic, S. L. Goh, and K. Aihara, "Sequential Data Fusion via Vector Spaces: Fusion of Heterogeneous data in the Complex Domain," *Journal of VLSI Signal Processing*, vol. 48, pp. 99 – 108, 2007.

[5] S. Haykin and L. Li, "Nonlinear Adaptive Prediction of Nonstationary Signals," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 526–535, 1995.

[6] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2481–2491, 1994.

[7] Y. Huang and J. Benesty, *Audio Signal Processing for Next Generation Multimedia Communication Systems*, Kluwer Academic Publishers, 2004.

[8] W. Hamilton, *Elements of Quaternions*, Longmans, Green and Co, 1899.

[9] P. Arena, S. Baglio, S. Fortuna, and M. G. Xibilia, "Chaotic Time-Series Prediction via Quaternionic Multilayer Perceptrons," *IEEE International Conference on System Man and Cybernetics*, pp. 1790–1794, 1995.

[10] P. Arena, L. Fortuna, G. Muscato, and M. G. Xibilia, "Neural Networks in Multidimensional Domains," *Lecture Note in Control and Information Sciences (Springer-Verlag)*, vol. 234, 1998.

[11] B. Widrow, J. McCool, and M. Ball, "The Complex LMS algorithm," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719 – 720, 1975.